

A comprehensive model for disomically inherited metrical traits expressed in triploid tissues

H. S. POONI, ISH KUMAR* & G. S. KHUSH†

School of Biological Sciences, University of Birmingham, PO Box 363, Birmingham, U.K. *Directorate of Rice Research, Rajendra Nagar, Hyderabad 500 030, India and †Irri, PO Box 933, Manila, Philippines

A biometrical genetic model is presented for the analysis of quantitatively varying diploid inherited traits which are expressed in a triploid phase. It shows that gene dispersion and ambidirectionality influences virtually all the components of means and at least four components of variances. Consequently, separate parameters are needed to describe the genetic variation among the second-degree statistics of the selfing and the backcrossing series. It is further shown that the effects of maternal/cytoplasmic inheritance can be separated from those of the nuclear genes both by the scaling tests and by the weighted least squares method. The applicability of the model to experimental data is demonstrated by analysing the amylose content of the generations derived from a cross between two pure breeding lines of rice.

Keywords: amylose content, endosperm traits, gene dosage, grain quality.

Introduction

Biometrical models currently available for the investigation of quantitatively controlled triploid expressed endosperm traits (e.g. Gale, 1976; Huidong, 1987; Bogyo *et al.*, 1988) are constrained in several respects. Firstly, they assume segregation at only one or two loci and therefore cannot be used effectively to analyse crosses which may be segregating at several loci and/or displaying dispersion of the 'increasing' and 'decreasing' alleles in the parental lines. Secondly, they do not take into consideration all possible reciprocals of various generations and various types of the additive \times dominance and the dominance \times dominance interaction effects that may be prevailing in the material. Furthermore, two of the models (e.g. Huidong, 1987 and Bogyo *et al.*, 1988) use the same scale but assign different expectations to several generations. Thus, the analysis of the same set of data by these models is unlikely to yield identical conclusions, especially when non-allelic interactions are detected significant.

In this paper we attempt to remedy these limitations of the current models by developing a model which can accommodate segregation at any number of loci and account for any level of gene association/dispersion in

the parental lines. We extend this model further to incorporate maternal/cytoplasmic inheritance so that their components can be tested/estimated simultaneously with those of the allelic and non-allelic effects of the nuclear genes. Although the new model is still subjected to the ubiquitous assumptions of linkage equilibrium, no genotype environment interaction and no gametic selection, we nevertheless believe that it is comprehensive enough to be applicable to most experimental situations and we demonstrate its field applicability by analysing a set of generations that have been derived from a cross between high and low amylose lines of rice.

Model

Gene effects and scale

Special features of gene expression among the triploid tissues of cereals such as the endosperm have already been explained by various authors (e.g. Gale, 1976; Huidong, 1987; Bogyo *et al.*, 1988) and therefore need not be elaborated again. It is worth mentioning, however, that disomic segregation at any locus (say A/a) in the $3N$ background produces four unique genotypes (namely AAA , AAa , Aaa and aaa) instead of the usual

Table 2 Expectations on a digenic interaction model of generations derived from the association and dispersion crosses

Generations	Parameters															
	<i>m</i>	<i>d_a</i>	<i>d_b</i>	<i>h_{a1}</i>	<i>h_{a2}</i>	<i>h_{b1}</i>	<i>h_{b2}</i>	<i>i_{ab}</i>	<i>j_{ab1}</i>	<i>j_{ab2}</i>	<i>j_{ba1}</i>	<i>j_{ba2}</i>	<i>l_{a1b1}</i>	<i>l_{a1b2}</i>	<i>l_{a2b1}</i>	<i>l_{a2b2}</i>
Association cross (<i>AAABBB</i> × <i>aaabbb</i>)																
<i>P₁</i>	1	1	1					1								
<i>P₂</i>	1	-1	-1					1								
<i>F₁</i> (1 × 2)	1	1/3	1/3	1		1		1/9	1/3		1/3		1			
<i>RF₁</i> (2 × 1)	1	-1/3	-1/3		1		1	1/9		-1/3		-1/3				1
<i>F₂</i>	1			1/4	1/4	1/4	1/4						1/16	1/16	1/16	1/16
<i>RF₂</i>	1			1/4	1/4	1/4	1/4						1/16	1/16	1/16	1/16
<i>F₁</i> × <i>P₁</i>	1	1/3	1/3				1/2	1/9		1/6		1/6			1/4	
<i>P₁</i> × <i>F₁</i>	1	2/3	2/3	1/2			1/2	4/9	1/3		1/3			1/4		
<i>F₁</i> × <i>P₂</i>	1	-1/3	-1/3	1/2			1/2	1/9	-1/6		-1/6			1/4		
<i>P₂</i> × <i>F₁</i>	1	-2/3	-2/3		1/2		1/2	4/9		-1/3		-1/3		1/4		
Dispersion cross (<i>AAAbbb</i> × <i>aaaBBB</i>)																
<i>P₁</i>	1	1	-1													
<i>P₂</i>	1	-1	1													
<i>F₁</i> (1 × 2)	1	1/3	-1/3	1			1	-1/9		1/3	-1/3			1		
<i>RF₁</i> (2 × 1)	1	-1/3	1/3		1	1		-1/9	-1/3			1/3			1	
<i>F₂</i>	1			1/4	1/4	1/4	1/4						1/16	1/16	1/16	1/16
<i>RF₂</i>	1			1/4	1/4	1/4	1/4						1/16	1/16	1/16	1/16
<i>F₁</i> × <i>P₁</i>	1	1/3	-1/3		1/2	1/2		-1/9	1/6			-1/6			1/4	
<i>P₁</i> × <i>F₁</i>	1	2/3	-2/3	1/2			1/2	-4/9		1/3	-1/3			1/4		
<i>F₁</i> × <i>P₂</i>	1	-1/3	1/3	1/2			1/2	-1/9		1/6	-1/6			1/4		
<i>P₂</i> × <i>F₁</i>	1	-2/3	2/3		1/2	1/2		-4/9	-1/3			1/3		1/4		

See text for definitions of parameters.

remaining *K'* loci, we need 10 parameters to explain the nuclear-controlled differences among the means of the early (*P₁*, *P₂*, *F₁*, *F₂*, *B₁* and *B₂*) as well as the advanced generations of a cross between them. These parameters are *m* (mean of the *F_∞* generation), [*d*] (mean additive deviation), [*h*]₁ and [*h*]₂ (dominance components), [*i*] (additive × additive interaction component), [*j*]₁ and [*j*]₂ (additive × dominance interaction components) and [*l*]₁, [*l*]₂ and [*l*] (dominance × dominance interaction components). These parameters take the following definitions.

m as earlier

$$[d] = \sum_{i=1}^{K-K'} d_i - \sum_{i=1}^{K'} d_i$$

$$[h]_1 = \sum_{i=1}^{K-K'} h_{i1} + \sum_{i=1}^{K'} h_{i2}$$

$$[h]_2 = \sum_{i=1}^{K-K'} h_{i2} + \sum_{i=1}^{K'} h_{i1}$$

$$[i] = \sum_{j>i=1}^K i_{ij} - \sum_{i=1}^{K-K'} \sum_{j=1}^{K'} i_{ij}$$

$$[j]_1 = \sum_{j>i=1}^{\frac{1}{2}(K-K')(K-1)} (j_{ij1} + j_{ji1}) - \sum_{j>i=1}^{\frac{1}{2}K'(K-1)} (j_{ij2} + j_{ji2})$$

$$[j]_2 = \sum_{j>i=1}^{\frac{1}{2}(K-K')(K-1)} (j_{ij2} + j_{ji2}) - \sum_{j>i=1}^{\frac{1}{2}K'(K-1)} (j_{ij1} + j_{ji1})$$

$$[l]_1 = \sum_{j>i=1}^{\frac{1}{2}K'(K'-1)} l_{i1j1} + \sum_{j>i=1}^{\frac{1}{2}(K-K')(K-K'-1)} l_{i2j2} + \sum_{j>i=1}^{\frac{1}{2}K'(K-K')} (l_{i1j2} + l_{i2j1})$$

$$[l]_2 = \sum_{j>i=1}^{\frac{1}{2}K'(K'-1)} l_{i2j2} + \sum_{j>i=1}^{\frac{1}{2}(K-K')(K-K'-1)} l_{i1j1} + \sum_{j>i=1}^{\frac{1}{2}K'(K-K')} (l_{i2j1} + l_{i1j2})$$

$$[l] = \sum_{j>i=1}^{\frac{1}{2}K(K-1)} (l_{i1j1} + l_{i1j2} + l_{i2j1} + l_{i2j2})$$

Contributions of the above parameters to the overall means of various generations and their reciprocals are given in Table 3. This model, however, accounts for nuclear controlled variation only. As triploid tissues in general and endosperm in particular are the characteristics of seeds/grains and mother plays a vital role in their nourishment, the differences between generations/families therefore can also be influenced by the delayed effects of the maternal genotypes and/or of cytoplasm. The models that can account for these

Table 3 Contributions of the genetic/cytoplasmic/maternal components to the family means of generations derived from a cross between two pure breeding varieties

Generation	Genetic components										Maternal/cytoplasmic components		
	<i>m</i> *	[<i>d</i>]	[<i>h</i>] ₁	[<i>h</i>] ₂	[<i>i</i>]	[<i>j</i>] ₁	[<i>j</i>] ₂	[<i>l</i>] ₁	[<i>l</i>] ₂	[<i>l</i>]	[<i>d</i> _{<i>m</i>}]	[<i>h</i> _{<i>m</i>}]	[<i>c</i>]
Parental lines													
<i>P</i> ₁	1	1	0	0	1	0	0	0	0	0	1	0	1
<i>P</i> ₂	1	-1	0	0	1	0	0	0	0	0	-1	0	-1
F ₁ hybrid													
<i>F</i> ₁ (1 × 2)	1	1/3	1	0	1/9	1/3	0	1	0	0	1	0	1
<i>RF</i> ₁ (2 × 1)	1	-1/3	0	1	1/9	0	-1/3	0	1	0	-1	0	-1
F ₂ generation													
<i>F</i> ₁ × <i>F</i> ₁ ^a	1	0	1/4	1/4	0	0	0	0	0	1/16	0	1	1
<i>F</i> ₁ × <i>RF</i> ₁ ^b	1	0	1/4	1/4	0	0	0	0	0	1/16	0	1	1
<i>RF</i> ₁ × <i>F</i> ₁ ^c	1	0	1/4	1/4	0	0	0	0	0	1/16	0	1	-1
<i>RF</i> ₁ × <i>RF</i> ₁ ^d	1	0	1/4	1/4	0	0	0	0	0	1/16	0	1	-1
B ₁ generation													
<i>F</i> ₁ × <i>P</i> ₁	1	1/3	0	1/2	1/9	0	1/6	0	1/4	0	0	1	1
<i>RF</i> ₁ × <i>P</i> ₁	1	1/3	0	1/2	1/9	0	1/6	0	1/4	0	0	1	-1
<i>P</i> ₁ × <i>F</i> ₁	1	2/3	1/2	0	4/9	1/3	0	1/4	0	0	1	0	1
<i>P</i> ₁ × <i>RF</i> ₁	1	2/3	1/2	0	4/9	1/3	0	1/4	0	0	1	0	1
B ₂ generation													
<i>F</i> ₁ × <i>P</i> ₂	1	-1/3	1/2	0	1/9	-1/6	0	1/4	0	0	0	1	1
<i>RF</i> ₁ × <i>P</i> ₂	1	-1/3	1/2	0	1/9	-1/6	0	1/4	0	0	0	1	-1
<i>P</i> ₂ × <i>F</i> ₁	1	-2/3	0	1/2	4/9	0	-1/3	0	1/4	0	-1	0	-1
<i>P</i> ₂ × <i>RF</i> ₁	1	-2/3	0	1/2	4/9	0	-1/3	0	1/4	0	-1	0	-1
F ₃ generation													
<i>F</i> ₃ (<i>a</i> & <i>b</i>)	1	0	1/8	1/8	0	0	0	0	0	1/64	0	1/2	1
<i>F</i> ₃ (<i>c</i> & <i>d</i>)	1	0	1/8	1/8	0	0	0	0	0	1/64	0	1/2	-1
F _∞ generation													
<i>F</i> _∞ (<i>a</i> & <i>b</i>)	1	0	0	0	0	0	0	0	0	0	0	0	1
<i>F</i> _∞ (<i>c</i> & <i>d</i>)	1	0	0	0	0	0	0	0	0	0	0	0	-1

*See text for definitions.

effects are already published (Mather & Jinks, 1982 and Pooni *et al.*, 1987) and need no elaboration except that they are represented by components [*c*], [*dm*] and [*hm*] which account for the averaged difference attributable to the parental cytoplasm, the delayed effect of inbred mothers and the delayed effect of the *F*₁ mother respectively. Contributions of these components to the various generations are also given alongside those of the genetic components in Table 3.

Scaling tests

Mather (1949) propounded the principles of scaling tests to establish: (a) the adequacy of the additive/dominance model and (b) the absence of scalar effects/non-allelic interactions. His ideas were further extended by Perkins & Jinks (1970), Jinks (1978) and Pooni *et al.* (1987) to devise tests for complications

such as reciprocal differences and linkage disequilibrium and to identify the most prominent sources/components when several sources are found to contribute to the variation simultaneously. The latter ideas are better applied to the present case where all five sources, namely additive, dominance, epistatic, maternal and cytoplasmic effects are assumed to contribute to the between-generation and the between-reciprocal differences.

When all possible reciprocals of the basic generations are available, the effects of cytoplasm and epistasis can be separated and tested unambiguously using the following comparisons.

Tests of cytoplasmic effects

$$1 \quad \bar{F}_2(F_1 \times F_1) + \bar{F}_2(F_1 \times RF_1) - \bar{F}_2(RF_1 \times F_1) - \bar{F}_2(RF_1 \times RF_1) = 4 [c]$$

$$2 \quad \bar{B}_1(F_1 \times P_1) - \bar{B}_1(RF_1 \times P_1) = 2[c]$$

$$3 \quad \bar{B}_2(F_1 \times P_2) - \bar{B}_2(RF_1 \times P_2) = 2[c]$$

Tests of epistasis

$$4 \quad \begin{aligned} & \bar{B}_1(F_1 \times P_1) + \bar{B}_2(F_1 \times P_2) - \bar{F}_2(F_1 \times F_1) \\ & \quad - \bar{F}_2(F_1 \times RF_1) \\ & = 0.222[i] - 0.167([j]_1 - [j]_2) \\ & \quad + 0.25([l]_1 + [l]_2 - 0.5[l]). \end{aligned}$$

$$5 \quad \begin{aligned} & \bar{B}_1(RF_1 \times P_1) + \bar{B}_2(RF_1 \times P_2) - \bar{F}_2(RF_1 \times F_1) \\ & \quad - \bar{F}_2(RF_1 \times RF_1) \\ & = 0.222[i] - 0.167([j]_1 - [j]_2) \\ & \quad + 0.25([l]_1 + [l]_2 - 0.5[l]). \end{aligned}$$

$$6 \quad \begin{aligned} & \bar{B}_1(P_1 \times F_1) + \bar{B}_1(P_1 \times RF_1) + \bar{B}_2(P_2 \times F_1) \\ & \quad + \bar{B}_2(P_2 \times RF_1) - \bar{F}_1 - \overline{RF}_1 - P_1 - \bar{P}_2 \\ & = 0.333([j]_1 - [j]_2) - 0.5([l]_1 + [l]_2) - 0.444[i]. \end{aligned}$$

Delayed effects of the maternal genotypes, however, cannot be tested unambiguously from the early generations because they are always confounded with some epistatic components. The following comparisons, nevertheless, provide conditional tests of $[hm]$ and $[dm]$ which can be applied after the non-allelic interactions are shown to be absent.

Tests of delayed effects of maternal genes

$$7 \quad \begin{aligned} & 2\bar{B}_1(F_1 \times P_1) + 2\bar{B}_2(RF_1 \times P_2) - \bar{B}_1(P_1 \times F_1) \\ & \quad - \bar{B}_1(P_1 \times RF_1) - \bar{B}_2(P_2 \times F_1) - \bar{B}_2(P_2 \times RF_1) \\ & = 4[h_m]. \end{aligned}$$

$$8 \quad \begin{aligned} & 2\bar{B}_2(RF_1 \times P_2) - 2\bar{B}_1(F_1 \times P_1) + \bar{B}_1(P_1 \times F_1) \\ & \quad + \bar{B}_1(P_1 \times RF_1) - \bar{B}_2(P_2 \times F_1) - \bar{B}_2(P_2 \times RF_1) \\ & = 4[d_m]. \end{aligned}$$

Components of variances. Using the genetic parameters defined in the beginning of this section we derive the expectations of the genetic components of various between- and within-family variances. On an additive/dominance model we require seven components to explain all the genetic variation among the variances of the various generations and their reciprocals. These components have the following definitions.

$$D \text{ (additive variance)} = \sum_{i=1}^K d_i^2$$

$$H_1 \text{ (dominance variance)} = \sum_{i=1}^{K-K'} h_{i1}^2 + \sum_{i=1}^{K'} h_{i2}^2$$

$$H_2 \text{ (dominance variance)} = \sum_{i=1}^{K-K'} h_{i2}^2 + \sum_{i=1}^{K'} h_{i1}^2$$

$$H_{12} \text{ (dominance covariance)} = \sum_{i=1}^K h_{i1} h_{i2}$$

$$F \text{ (add./dom. covariance)} = \sum_{i=1}^K d_i (h_{i1} + h_{i2})$$

$$F' \text{ (add./dom. covariance)} = \sum_{i=1}^{K-K'} d_i h_{i1} - \sum_{i=1}^{K'} d_i h_{i2}$$

$$F'' \text{ (add./dom. covariance)} = \sum_{i=1}^{K-K'} d_i h_{i2} - \sum_{i=1}^{K'} d_i h_{i1}.$$

Contributions of these and the environmental components to the various variances are given in Table 4. As usual, all the genetic components are biased when they are estimated in the presence of epistasis and/or linkage disequilibrium.

Example

Data that we analysed to demonstrate the applicability of model to field experiments were recorded on a set of basic generations derived from a cross between two elite inbred lines of rice IR8 and IR24. The F_1 cross of these lines was produced reciprocally during 1985 and the F_1 and RF_1 plants were selfed and crossed to IR8 and IR24 during the following season. The reciprocal F_1 hybrids were also reproduced during the latter season to obtain seed of the same age. These materials formed a part of a much larger study, the results of which will be published in a separate paper.

Large numbers of seeds were sampled from each of the P_1 , P_2 , F_1 , RF_1 , F_2 (F_1 self), RF_2 (RF_1 self), B_1 ($F_1 \times P_1$), RB_1 ($RF_1 \times P_1$), B_2 ($F_1 \times P_2$) and RB_2 ($RF_1 \times P_2$) families and analysed for amylose content (percentage of endosperm by weight), which is known to be controlled by several loci with unequal effects (Kumar *et al.*, 1987). The amylose levels of individual seeds were measured on a Technicon Analyser following Juliano (1971). The means, variances and degrees of freedom, etc. of various generations are presented in Table 5.

Tests of cytoplasmic effects and non-allelic interactions

Lack of a complete set of reciprocals among these data restricts the application of several tests that are described in the subsection on scaling tests. Furthermore, others need to be modified and still more are devised to determine the significance of some specific components. The results of these tests, together with

Table 4 Genetic and environmental components of within family variances of various generations derived from a cross between two pure breeding varieties

Generation	Genetic components							Environmental components			
	D^*	H_1	H_2	H_{12}	F	F'	F''	E^\dagger	E_1	E_2	E_3
P_1	0	0	0	0	0	0	0	1	1	0	0
P_2	0	0	0	0	0	0	0	1	0	1	0
F_1 and RF_1	0	0	0	0	0	0	0	1	0	0	1
F_2 (all)	5/9	3/16	3/16	-1/8	1/6	0	0	1	1/4	1/4	1/2
B_1 (F_1 mothers)	4/9	0	1/4	0	0	0	-2/3	1	1/2	0	1/2
B_1 (P_1 mother)	1/9	1/4	0	0	0	-1/3	0	1	1/2	0	1/2
B_2 (F_1 mothers)	4/9	1/4	0	0	0	2/3	0	1	0	1/2	1/2
B_2 (P_2 mother)	1/9	0	1/4	0	0	0	1/3	1	0	1/2	1/2
F_3 (all) σ^2b	5/9	1/64	1/64	1/32	0	0	0	0	0	0	0
σ^2w	5/18	3/32	3/32	-1/16	1/12	0	0	1	3/8	3/8	1/4
$F_{..}$ (all) σ^2b	1	0	0	0	0	0	0	0	0	0	0
σ^2w	0	0	0	0	0	0	0	1	1/2	1/2	0

*See text for definitions.

† E_1, E_2 and E_3 apply when $VP_1 \neq VP_2 \neq VF_1$.

Table 5 Means and variances of various generations derived from a cross between two pure breeding lines IR8 and IR24 of rice for amylose content

Generation	\bar{x}	d.f.	$V(x)$	$V(\bar{x})$
P_1 (IR8)	27.26	19	0.81	0.0405
P_2 (IR24)	15.11	19	0.82	0.0410
F_1 ($P_1 \times P_2$)	25.24	19	2.22	0.1110
RF_1 ($P_2 \times P_1$)	27.96	19	0.90	0.0450
F_2 ($F_1 \times F_1$)	23.80	395	27.75	0.0701
RF_2 ($RF_1 \times RF_1$)	24.60	414	30.26	0.0729
B_1 ($F_1 \times P_1$)	27.42	85	5.91	0.0687
RB_1 ($RF_1 \times P_1$)	29.47	99	21.51	0.2151
B_2 ($F_1 \times P_2$)	18.93	99	51.25	0.5125
RB_2 ($RF_1 \times P_2$)	19.58	98	55.74	0.5630

the comparisons on which they are based, are presented in Table 6.

Two of the three tests employed to determine the presence of cytoplasmic effects are highly significant and the third is not. Each comparison, however, takes a negative value and differences between these values are not significant [$\chi^2_{(2)} = 2.47^{ns}$]. This suggests that the results are consistent among tests and differences due to the maternal cytoplasm are highly significant [combined $\chi^2_{(1)} = 8.15^{**}$].

Tests (d) and (e) show that non-allelic interactions are non-significant throughout. However, these tests have the same expectation

$$(e.g. 0.222[i] - 0.167([j]_1 - [j]_2) + 0.25([l]_1 + [l]_2 - 0.5[l]))$$

Table 6 Tests of cytoplasmic effects and digenic interactions

Comparison	Test	c	P
Cytoplasmic effects			
(a) $\bar{F}_2 - \overline{RF}_2$	-0.80 ± 0.38	2.12	*
(b) $\bar{B}_1 - \overline{RB}_1$	-2.07 ± 0.53	3.89	***
(c) $\bar{B}_2 - \overline{RB}_2$	-0.63 ± 1.04	0.61	ns
Non-allelic interactions			
(d) $\bar{B}_1 + \bar{B}_2 - 2\bar{F}_2$	-1.25 ± 0.93	1.35	ns
(e) $\overline{RB}_1 + \overline{RB}_2 - 2\overline{RF}_2$	-0.15 ± 1.03	0.15	ns
(f) $2\bar{F}_2 + 2\overline{RF}_2 - \bar{F}_1 - \overline{RF}_1 - \bar{P}_1 - \bar{P}_2$	1.23 ± 0.90	1.37	ns
(g) $\bar{B}_1 + \overline{RB}_1 - \bar{B}_2 - \overline{RB}_2 + \bar{F}_1 - \overline{RF}_1 - \bar{P}_1 + \bar{P}_2$	3.51 ± 1.26	2.90	**
(h) $2\bar{F}_2 + 2\overline{RF}_2 - 2\bar{B}_1 - 2\overline{RB}_1 - 2\bar{B}_2 - 2\overline{RB}_2 + \bar{F}_1 + \overline{RF}_1 + \bar{P}_1 + \bar{P}_2$	1.57 ± 2.50	0.63	ns

ns $P > 0.05$; * $0.05 \geq P > 0.01$; ** $0.01 \geq P > 0.001$;

*** $P \leq 0.001$.

and it is possible that various components have cancelled out. Comparisons (f), (g) and (h), on the other hand, determine the significances of $-2.222[i]$

$$(assuming 0.25[l] - [l]_1 - [l]_2 - 0.333([j]_1 - [j]_2) is small),$$

$$0.667([j]_1 + [j]_2)$$

$$(assuming 0.5([l]_2 - [l]_1) is small)$$

$$and 1.333[i]$$

$$(assuming [j]_1 - [j]_2 - [l]_1 - [l]_2 + 0.25[l] is small)$$

respectively and show that $[j]_1 + [j]_2$ is significant while $[i]$ is not. Test (f) is also unique in that it detects epistasis only when $[hm]$ is zero. A non-significant value of this test further indicates that either $4[hm]$ is zero or its effect is cancelled by that of $-2.222[i]$.

In addition, comparisons $0.5(\bar{P}_1 - \bar{P}_2)$ and $\bar{F}_1 + R\bar{F}_1 - \bar{P}_1 - \bar{P}_2$ are also significant indicating that $[d] + [dm] + [c]$ and $[h]_1 + [h]_2 - 1.7778[i]$ differ significantly from zero.

Maximum likelihood estimates

The above tests have indicated that parameters $[i]$, $[I]_1$, $[I]_2$, $[I]$ and $[hm]$ are non-significant and therefore need not be included in the model. Consequently, a model including m , $[d]$, $[h]_1$, $[h]_2$, $[j]_1$, $[j]_2$, $[dm]$ and $[c]$ parameters was fitted to the means of the 10 generations using the weighted least squares (Wls) procedures of Cavalli (1952) and Mather & Jinks (1982). While this model gave an adequate fit it included, however, two ($[j]_1$ and $[dm]$) components which were not significant. The model was therefore refitted after excluding the non-significant components. The estimates of the remaining components and the goodness of fit of the model are tabulated in Table 7.

Table 7 Components of generation means and goodness of fit of the model

Component	Wls estimate	<i>c</i>	<i>P</i>
<i>m</i>	21.19 ± 0.14	153	***
$[d]$	6.70 ± 0.20	33	***
$[h]_1$	2.37 ± 0.36	7	***
$[h]_2$	9.08 ± 0.33	28	***
$[j]_2$	1.96 ± 0.84	2.3	*
$[c]$	-0.60 ± 0.15	4.0	***
$\chi^2(4)$	7.59		ns

See Table 6 for probability levels.

Components of variances

According to the expectations in Table 4 we do not expect any heterogeneity between the variances of P_1 and P_2 ; F_1 and RF_1 ; F_2 and RF_2 , B_1 and RB_1 and B_2 and RB_2 families except in the presence of genotype × micro-environmental interactions or differential maternal sensitivity. On the other hand, genetic segregation is expected to make the pooled variances of the F_2 , B_1 and B_2 generations significantly larger than the environmental variance E [= $0.25(VP_1 + VP_2 + VF_1 + VRF_1)$]. Variance ratios to test the differences between the various pairs of statistics are presented below.

Comparison	Variance ratio (d.f.1, d.f.2)	Significance
VP ₁ vs. VP ₂	1.01 (19,19)	<i>P</i> > 0.05
VF ₁ vs. VRF ₁	2.47 (19,19)	<i>P</i> > 0.05
VF ₂ vs. VRF ₂	1.09 (395,414)	<i>P</i> > 0.05
VB ₁ vs. VRB ₁	3.64 (99,85)	<i>P</i> > 0.001
VB ₂ vs. VRB ₂	1.09 (98,99)	<i>P</i> > 0.05
Pooled VF ₂ vs. E	24.45 (809,76)	<i>P</i> > 0.001
Pooled VB ₁ vs. E	12.04 (184,76)	<i>P</i> > 0.001
Pooled VB ₂ vs. E	45.04 (197,76)	<i>P</i> > 0.001

The above tests show that the experimental results agree with the theoretical expectations on every occasion except one. Variances VB₁ and VRB₁ differ significantly indicating the presence of $g \times e$ interaction and/or differential maternal sensitivity. However, this is an isolated result which is not supported by any other test (from this or other crosses) and therefore must be treated as a rare event.

Further analysis of the second degree statistics is not possible without some explicit assumptions. The first assumption that we make is of complete association of alleles between the parental lines. Under this assumption we equate F with $F' + F''$ and $0.1875(H_1 + H_2) + 0.125H_{12}$ with $0.125(H_1 + H_2)$ and modify the F_2 expectation to:

$$VF_2 = 0.5556D + 0.125(H_1 + H_2) + 0.1667(F' + F'') + E.$$

Secondly, we assume that the dominance (H_1 and H_2) components are confounded with the covariance (F' and F'') components in such a way that F' and F'' provide the estimates of $F' + 0.375H_1$ and $F'' - 0.375H_2$, respectively. Now various components can be estimated both by simultaneous equations and by weighted least squares method and these estimates are listed below.

Component	Simultaneous solution	Simultaneous estimate	Wls estimate
<i>E</i>	$0.25(VP_1 + VP_2 + VF_1 + VRF_1)$	1.19	1.19 ± 0.19
<i>D</i>	$1.8(VF_2 - E) - 0.45(VB_1 - VB_2)$	32.49	30.93 ± 2.79
<i>F'</i>	$1.8VB_2 - 1.2VF_2 - 0.3(VB_1 + E)$	56.78	60.86 ± 8.10
<i>F''</i>	$1.2(VF_2 - VB_1) - 0.3(VB_2 - E)$	1.99	—
			$\chi^2(1) = 0.43^{ns}$

Discussion and conclusion

It is apparent from Table 1 that 16 parameters are needed to define completely the intra- and interallelic effects of two loci. These parameters are the same in number as those used by Huidong (1987) but considerably more than those defined by Bogyo *et al.* (1988). Furthermore, the new parameters are not merely the symbolic variants of the previously defined addi-

tive \times dominance and dominance \times dominance interaction components they are also expected to differ in magnitude and direction under various genetic situations. Consequently, the present model is likely to explain the gene action and digenic interaction more effectively than the models of Bogyo *et al.* (1988).

An important consequence of gene expression in the triploid phase is that both nuclear and cytoplasmic/maternal factors can affect the performances of reciprocal crosses simultaneously (see Tables 2 and 3). For example, different types of gene effects, namely additive, dominance and epistasis, etc., now contribute to the differences between the reciprocals of the F_1 and the backcross generations, which are not normally affected when the character is disomically inherited. This means that the F_1 performance is also affected by gene association/dispersion in exactly the same way as that of the parental lines and the backcross families. Furthermore, such differences are expected to persist in the descendent generations of the backcrosses but not of the F_1 . Therefore, separate interaction parameters are needed to explain the genetic differences between the generations of the backcrossing and the selfing series (except F_1).

Gene dispersion is also a major factor which affects eight of the 10 components that summarize the intra- and interallelic effects of the multilocus model (see Table 3). It reduces the apparent magnitude of components like $[d]$ and $[i]$ and confounds the various dominance (h_{i1} and h_{i2}) and interaction components (j and l) with each other. The latter components are also subjected to internal cancellations due to the ambidirectionality of various intra- and interallelic effects. Thus, the true magnitudes of various components will be revealed rarely except when gene association is complete, dominance is unidirectional and all pairs of interacting loci display the same type (complementary or duplicate) of epistasis.

With 13 parameters in the model and many of them correlated with each other, it is absolutely essential to identify the contributing components so that they can be included in the model and estimated by the weighted least squares procedure. Comparisons presented in the sub-section on 'scaling tests' provide much of the requisite information concerning the epistatic components and the cytoplasmic effects. However, while the contributions of these effects can be determined unambiguously prior to WLS analysis, tests of the delayed maternal effects ($[dm]$ and $[hm]$) are not possible from the early generations except when epistasis is non-significant. The presence of the maternal components, therefore, has to be interpreted indirectly from the inadequacy of the model when non-allelic interaction is detected significant.

The impact of gene dispersion on the components of variances is twofold (see Table 4). Firstly, it confounds the two types of dominance effects (h_{i1} and h_{i2}) and their cross products in such a way that both of them appear in the expectations of H_1 , H_2 , F' and F'' components. Thus the significance of any of these components cannot be attributed exclusively to the h_{i1} nor the h_{i2} effects of various loci except when it is known that all the high and low scoring alleles are present in P_1 and P_2 respectively. Secondly, it makes the cross-product components differ between the backcrossing and the selfing series such that F is equal to F' and F'' but only when alleles are completely associated in the parents ($K' = 0$). Components D and H_{12} , on the other hand, are unaffected by gene dispersion and therefore can be interpreted unambiguously irrespective of the level of gene association/dispersion displayed by the parental lines.

Application of the present model to the rice data has revealed for the first time that cytoplasm has a significant effect on the amylose level of various generations. Although its effect is comparatively small, nevertheless replacement of the IR8 cytoplasm with that of IR24 increases the amylose level by up to 0.60 per cent (see Table 7). Furthermore, these results are not exclusive to the IR8 \times IR24 cross because cytoplasmic differences have also been detected in all the remaining crosses of the present study.

A general conclusion that can be drawn from the analysis of first-degree statistics is that amylose content is controlled by a minimum of two, and perhaps more, genes whose alleles display a modest level of digenic interaction. A large value of $[d]$ points to a high degree of gene association in the parental lines while a large $[h]_2$ indicates that a single dose of the dominant allele is often sufficient for the full expression of dominance at the various loci. The marginal significance of $[h]_1$, on the other hand, suggests that the h_{i1} effects are either negligible or ambidirectional. It is also possible that the h_{i1} effects are non-existent throughout and $[h]_1$ mostly represents the h_{i2} effects of those loci (K') that are dispersed in the parents.

At first sight the analysis of second-degree statistics seems to contradict the above conclusions (see section on components of variances). This is particularly apparent from the presence of F' and the absence of F'' among the estimated components, which suggests that perhaps h_{i1} effects of various loci are more important than their h_{i2} counterparts. However, a closer look reveals that F' represents the combined magnitude of $F' + 0.375H_1$, which can take a large value even when both components are small. On the other hand, F'' provides an estimate of $F'' - 0.375H_2$ and it is more likely to be underestimated in most situations except

when dominance is for the lower score and F'' is negative.

Finally, the predictions made from the available parameters following Jinks & Pooni (1976) reveal that the rate of transgression among the recombinant inbred lines is likely to be high. Nearly 27 per cent of these lines are expected to transgress the parental range (13.79 per cent for each parent) and about 11 per cent should also match or supersede the amylose level of the hybrid IR24 × IR8. However, a similar range of amylose levels will also be observed among the second cycle hybrids which can be produced by crossing the recombinant inbreds. While the genetics of this character therefore accrue no clear advantage to hybrids or inbreds as varieties, the same however may not be true for other traits such as yield and maturity.

Acknowledgements

We would like to thank Drs M. J. Kearsey and P. S. Virk for critically reading the manuscript. The statistical analysis was carried out on the University of Birmingham Mainframe Computer.

References

- BOGYO, T. P., LANCE, R. C. M., CHEVALIER, P. AND NILAN, R. A. 1988. Genetic models for quantitatively inherited endosperm characters. *Heredity*, **60**, 61–67.
- CAVALLI, L. L. 1952. An analysis of linkage in quantitative inheritance. In: Reeve, E. C. R. and Waddington, C. H. (eds) *Quantitative Inheritance*, HMSO, London, pp. 135–144.
- GALE, M. D. 1976. High α -amylase breeding and genetical aspects of the problem. *Cereal Res. Comm.*, **4**, 231–243.
- HUIDONG, M. O. 1987. Genetic expression for endosperm traits. *Proceedings of the Second International Conference on Quantitative Genetics*, Sinaur Associates Inc., Massachussetts, pp. 478–487.
- JINKS, J. L. 1978. Unambiguous test for linkage of genes showing nonallelic interactions for a metrical trait. *Heredity*, **40**, 171–173.
- JINKS, J. L. AND POONI, H. S. 1976. Predicting the properties of recombinant inbred lines derived by single seed descent. *Heredity*, **36**, 253–266.
- JULIANO, B. O. 1971. A simplified assay for milled rice amylose. *Cereal Sci. Today*, **16**, 334–338.
- KUMAR, I., KHUSH, G. S. AND JULIANO, B. O. 1987. Genetic analysis of waxy locus in rice (*Oryza sativa* L.). *Theor. Appl. Genet.*, **73**, 481–488.
- MATHER, K. 1949. *Biometrical Genetics* 1st edn, Methuen, London.
- MATHER, K. AND JINKS, J. L. 1982. *Biometrical Genetics* 3rd edn, Chapman and Hall, London.
- PERKINS, J. M. AND JINKS, J. L. 1970. The detection and estimation of genotype environment interaction, linkage and epistatic components for a metrical trait. *Heredity*, **25**, 157–177.
- POONI, H. S., COOMBS, D. T., VIRK, P. S. AND JINKS, J. L. 1987. Detection of epistasis and linkage of interacting genes in the presence of reciprocal differences. *Heredity*, **58**, 257–266.
- SMITH, H. H. AND ROBSON, D. S. 1959. Inheritance of dimensions of flower parts in tobacco. *Proceedings of the International Symposium on Biometrical Genetics*, Ottawa, Pergamon Press.