

# Genetical control of amylose content in selected crosses of *indica* rice

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

School of Biological Sciences, University of Birmingham, P.O. Box 363, Birmingham B15 2TT, U.K. and \*IRRI, P.O. Box 933, Manila, Philippines

Models of Pooni *et al.* (1992) were employed to investigate the genetic control of amylose content in 10 rice crosses produced by the pairwise crossing of five varieties representing almost the whole range of amylose levels from 0 to 28 per cent. Analyses of the first-degree statistics revealed an important role of the additive and the dominance effects in determining the genetic variability in all the crosses. Epistasis and cytoplasmic effects were also observed to contribute significantly to the variability among the generation means of most crosses. Dominance was generally towards the higher score and its effects were enhanced by a complementary dominance  $\times$  dominance interaction in several sets of basic generations.

The predominantly additive nature of the genetic variability was further revealed by the analyses of second-degree statistics. Component *D* was detected significant in all the crosses while components  $H_1$  and  $H_2$  were non-significant throughout. Significance of the covariance components  $F'$  and  $F''$ , however, showed indirectly that dominance contributed significantly to variability at the variance level. Higher levels of transgression and considerable increases in the phenotypic ranges displayed by the segregating generations of various crosses, also pointed to their potency for yielding superior recombinants with diverse levels of amylose.

**Keywords:** cooking quality, endosperm traits, gene action, triploid tissues.

## Introduction

Amylose content determines the cooking quality of rice. It is known to vary from virtually 0 to about 30 per cent in the cultivars around the world. Unlike other cereals, people of various regions prefer to eat certain types of rice (Kumar & Khush, 1987). This obviously means that precise levels of amylose have to be bred into new varieties to satisfy local preferences.

Transmission of amylose content, on the other hand, is very complex indeed. It is stored in the endosperm which is a triploid tissue. Consequently, the standard diploid models of biometrical genetics cannot be applied effectively to study its genetic control. Pooni *et al.* (1992) have recently proposed a comprehensive set of models for triploid tissues which deal specifically with the inheritance of traits such as amylose content. In this paper we use these models to investigate the genetic control of amylose content in a set of crosses which were produced by the pairwise crossing of a stratified sample of varieties representing all the major amylose groups.

## Materials and methods

Materials for the present study consisted of the parental,  $F_1(P_1 \times P_2)$ ,  $RF_1(P_2 \times P_1)$ ,  $F_2(F_1 \text{ selfed})$ ,  $RF_2(RF_1 \text{ selfed})$ ,  $B_1(F_1 \times P_1)$ ,  $RB_1(RF_1 \times P_1)$ ,  $B_2(F_1 \times P_2)$  and  $RB_2(RF_1 \times P_2)$  generations derived from 10 crosses involving the following rice varieties.

Variety	Abbreviation	Category	Amylose class (%)
IR29	IR29	Waxy	0–2
IR37307–8	IR307	Very low amylose	3–9
IR24	IR24	Low amylose	10–19
BPI 121–407	BPI	Intermediate amylose	20–25
IR8	IR8	High amylose	> 25

The varieties were initially tested for genetic uniformity and then self/cross-pollinated in all possible pairwise combinations to produce five selfed and 20  $F_1$  families. Individuals of these  $F_1$ s, and parental selfs, were grown at IRRI during 1985 (dry season) and self/cross-pollinated to obtain the  $F_2$  and the backcross generations. Fresh seeds of the parental lines and the  $F_1$  hybrids were also produced to minimize the environmental and

seed age differences among various seed lots. The seeds obtained were dehulled in a Satake machine and individually milled in a test tube mill. Single-grain analysis for amylose content was performed on 20 randomly sampled seeds of each parental and  $F_1$  family, about 400 seeds of each  $F_2$  generation and between 60 and 100 seeds of various backcross generations, following Juliano (1971).

## Analysis and results

### (i) Generation means and reciprocal differences

Generation means, their standard errors and the  $t$  values determining the significance of differences between the parents and the reciprocal families of  $F_1$ ,  $F_2$ ,  $B_1$  and  $B_2$  generations are presented in Table 1. The results indicate that all the parents differ significantly from each other for amylose content. Significant differences are also detected between the  $F_1$  and  $RF_1$  families of all the crosses except  $BPI \times IR29$ , indicating either the presence of maternal effects or the impact of gene dosage or both. Reciprocals of the  $F_2$  and the  $B_1$  generations, on the other hand, differ significantly on only 14 out of 20 occasions, which suggests that these differences are generally less pronounced than those displayed by the  $F_1$  and the  $RF_1$  families of various crosses.

The  $B_2$  generation, however, contrasts with the rest in that differences between its reciprocals are non-significant in all the crosses except  $BPI \times IR307$  and  $IR8 \times BPI$ .

### (ii) Test of cytoplasmic effects

According to the models of Pooni *et al.* (1992), differences between the reciprocals of  $F_2$ ,  $B_1$  and  $B_2$  generations in the present study can only be attributed to the cytoplasmic (denoted by  $[c]$ ) effects. Furthermore, these differences are expected to remain constant as long as there is no differential interaction between maternal cytoplasm and nuclear genes. We test this assumption by fitting a single  $[c]$  parameter to  $\bar{F}_2 - \overline{RF}_2$ ,  $\bar{B}_1 - \overline{RB}_1$  and  $\bar{B}_2 - \overline{RB}_2$  values by the weighted least squares procedure (see Mather & Jinks, 1982 for details) and determining its goodness of fit as a  $\chi^2$  for two degrees of freedom.

The estimates of  $[c]$ , the  $\chi^2_{2;}$  values, and the visual relationships between the  $[c]$  values of various generations, are presented in Table 2. These results show that  $[c]$  is significant for all the crosses except  $IR24 \times IR29$  and  $BPI \times IR307$ . The significance of  $\chi^2_{2;}$  further indicates that  $[c]$  alone does not explain all the differences between the reciprocals of  $F_2$ ,  $B_1$  and  $B_2$  genera-

tions in seven crosses. Additional parameters will therefore have to be allocated to the final models to account for these differences.

Differences between various pairs of reciprocals, however, remain constant in  $BPI \times IR29$  and  $IR8 \times IR24$  crosses. Consequently, component  $[c]$  explains all the significant variability among the reciprocals of these crosses. No reciprocal effects are detected in cross  $IR24 \times IR29$ .

### (iii) Test of epistasis

The presence of non-allelic interactions is normally detected by scaling tests which assume the absence of complications such as differential interactions between maternal cytoplasm and progeny genotypes. Scaling tests, devised by Pooni *et al.* (1992), therefore cannot be applied effectively in the present case. Instead we determine the presence of epistasis by fitting the  $m$  and  $[h']$  ( $= [h]_1 + [h]_2$ ; see Pooni *et al.*, 1992 for notation and definitions of various parameters) model to  $P' = 0.5(\bar{P}_1 + \bar{P}_2)$ ,  $F'_1 = 0.5(\bar{F}_1 + \overline{RF}_1)$ ,  $F'_2 = 0.5(\bar{F}_2 + \overline{RF}_2)$  and  $B' = 0.25(\bar{B}_1 + \overline{RB}_1 + \bar{B}_2 + \overline{RB}_2)$  where the cytoplasmic effects and their interactions with the progeny genotype are averaged out. The  $\chi^2$  values for two degrees of freedom determining the goodness of fit of this model are presented in Table 3. These  $\chi^2$  are significant for all the crosses except  $IR8 \times IR24$ , suggesting that the additive dominance model is adequate only in the latter case.

Whenever the  $m$  and  $[h']$  model failed to account for all the variability among the above statistics, we attempted to identify the most likely source of the failure by fitting a three-parameter model ( $m$ ,  $[H']$ ,  $[i]$  or  $m$ ,  $[h']$ ,  $[j']$  where  $[j'] = [j]_1 - [j]_2$ ) and determining its goodness of fit as  $\chi^2$  for one degree of freedom. The expectations of the various generations on this model are given below and the results are summarized in Table 3.

Generation	$m$	$[h']$	$[i]$	$[j']$
$P'$	1	0	1	0
$F'_1$	1	0.5	0.111	0.167
$F'_2$	1	0.25	0	0
$B'$	1	0.25	0.111	-0.063

Clearly the three-parameter model is still inadequate in six crosses suggesting that other epistatic component(s) must also be significant in these cases. On the other hand, the  $m$ ,  $[h']$  and  $[i]$  model is adequate for crosses  $IR307 \times IR29$  and  $BPI \times IR307$  and the  $m$ ,  $[h']$  and  $[j']$  account for all the significant variability among the four means of  $BPI \times IR24$  cross.

**Table 1** Generation means and Student's *t*-values determining the significance of differences between P<sub>1</sub> and P<sub>2</sub> and the reciprocals of the F<sub>1</sub>, F<sub>2</sub>, B<sub>1</sub> and B<sub>2</sub> families of various crosses

		Crosses									
		IR8 × IR29	BPI × IR29	IR24 × IR29	IR307 × IR29	IR8 × IR307	BPI × IR307	IR24 × IR307	IR8 × IR24	BPI × IR24	IR8 × BPI
P <sub>1</sub>	$\bar{x}$	27.26	24.89	15.11	7.41	27.26	24.89	15.11	27.26	24.89	27.26
	± S.E.	0.21	0.08	0.20	0.41	0.21	0.08	0.20	0.21	0.08	0.21
P <sub>2</sub>	$\bar{x}$	0.01	0.01	0.01	0.01	7.41	7.41	7.41	15.11	15.11	24.89
	± S.E.	0.01	0.01	0.01	0.01	0.41	0.41	0.41	0.20	0.20	0.08
<i>t</i>		130**	311**	76**	18**	43**	43**	17**	42**	46**	11**
F <sub>1</sub>	$\bar{x}$	27.54	22.40	12.22	3.09	26.88	18.65	12.88	25.24	24.47	29.27
	± S.E.	0.33	0.55	0.26	0.16	0.27	0.24	0.33	0.33	0.47	0.29
RF <sub>1</sub>	$\bar{x}$	23.88	22.92	5.59	1.51	23.47	13.84	6.86	27.96	27.21	26.43
	± S.E.	0.60	0.30	0.23	0.14	0.29	0.39	0.36	0.21	0.22	0.28
<i>t</i>		5.4**	0.8ns	19**	7.4**	8.6**	11**	12**	6.9**	5.3**	7.1**
F <sub>2</sub>	$\bar{x}$	19.71	18.51	9.60	3.09	20.29	13.06	10.65	23.80	21.28	25.19
	± S.E.	0.55	0.53	0.31	0.12	0.46	0.31	0.19	0.26	0.21	0.17
RF <sub>2</sub>	$\bar{x}$	20.50	17.60	9.57	4.15	22.21	14.84	9.46	24.60	22.72	27.21
	± S.E.	0.62	0.51	0.28	0.13	0.41	0.32	0.23	0.27	0.20	0.17
<i>t</i>		1.0ns	1.4ns	0.1ns	5.9**	3.1**	4.1**	4.0**	2.1*	5.0**	8.5**
B <sub>1</sub>	$\bar{x}$	26.46	22.50	6.33	4.05	24.03	19.77	9.25	27.42	21.22	26.74
	± S.E.	0.05	0.19	0.46	0.31	0.34	0.43	0.32	0.26	0.29	0.29
RB <sub>1</sub>	$\bar{x}$	23.33	21.96	7.18	5.10	21.78	17.31	8.33	29.47	22.28	30.28
	± S.E.	0.37	0.22	0.54	0.27	0.43	0.66	0.33	0.34	0.24	0.55
<i>t</i>		8.5**	1.8ns	1.2ns	2.6*	4.1**	3.1**	2.0*	4.8**	2.8**	8.4**
B <sub>2</sub>	$\bar{x}$	12.42	11.63	3.62	2.04	14.06	12.77	6.74	18.93	20.57	25.72
	± S.E.	1.50	1.27	0.41	0.18	1.06	0.53	0.26	0.75	0.47	0.30
RB <sub>2</sub>	$\bar{x}$	11.27	9.11	4.41	1.90	16.23	9.73	7.10	19.58	19.90	26.93
	± S.E.	1.40	1.20	0.53	0.18	0.91	0.95	0.16	0.72	0.46	0.42
<i>t</i>		0.6ns	1.5ns	1.2ns	0.9ns	1.6ns	2.8**	1.1ns	0.6ns	1.5ns	2.4*

ns =  $P > 0.05$ ; \* =  $0.05 \geq P > 0.01$ ; \*\* =  $P \leq 0.01$ .  
See text for other symbols.

*(iv) Components of generation means*

Results of the above tests were used to determine the parameters of the model which was most likely to fit the data of a particular cross. Pooni *et al.* (1992) have already defined all the genetic/maternal/cytoplasmic parameters of the model and given their contributions to the means of various generations (in their table 3). Contributions to the F<sub>2</sub>, B<sub>1</sub> and B<sub>2</sub> generations of the additional parameters [z]F<sub>2</sub>, [z]B<sub>1</sub> and [z]B<sub>2</sub>, which were used to account for the differential interactions between the cytoplasmic effects and the nuclear effects of these generations, are described below.

Generation	[c]	[z]F <sub>2</sub>	[z]B <sub>1</sub>	[z]B <sub>2</sub>
F <sub>2</sub>	1	1	0	0
RF <sub>2</sub>	-1	-1	0	0
B <sub>1</sub>	1	0	1	0
RB <sub>1</sub>	-1	0	-1	0
B <sub>2</sub>	1	0	0	1
RB <sub>2</sub>	-1	0	0	-1

The models were fitted by the weighted least-squares procedures of Cavalli (1952) and Mather & Jinks (1982). Models that had (a) all the parameters signifi-

**Table 2** Tests of cytoplasmic effects ( $[c]$ ) and their interactions with the nuclear genes

Cross	Maximum likelihood estimate of $[c]$	$\chi^2_{(2)}$	Possible relationships between $[c]$ values
IR8 × IR29	0.44 ± 0.17	6.0*	$[c]F_2 < [c]B_1 = [c]B_2$
BPI × IR29	1.13 ± 0.13	5.1ns	$[c]F_2 = [c]B_1 = [c]B_2$
IR24 × IR29	-0.17 ± 0.16ns	1.8ns	$[c] = 0.0$
IR307 × IR29	-0.45 ± 0.07	9.4**	$[c]B_1 = [c]F_2 < [c]B_2$
IR8 × IR307	-0.86 ± 0.19	8.9*	$[c]B_1 > [c]F_2 = [c]B_2$
BPI × IR307	-0.13 ± 0.18ns	35.5***	$[c]F_2 < [c]B_1 = [c]B_2$
IR24 × IR307	0.40 ± 0.10	8.1**	$[c]B_1 = [c]F_2 > [c]B_2$
IR8 × IR24	-0.41 ± 0.14	1.6ns	$[c]F_2 = [c]B_1 = [c]B_2$
BPI × IR24	-0.35 ± 0.11	20.1***	$[c]B_1 = [c]F_2 < [c]B_2$
IR8 × BPI	-1.09 ± 0.10	9.8**	$[c]B_1 < [c]F_2 = [c]B_2$

$[c]$  is highly significant (\*\*) except when marked ns.  
 $P > 0.05$ ; \* =  $0.05 \geq P > 0.01$ ; \*\* =  $0.01 \geq P > 0.001$ ; \*\*\* =  $P \leq 0.001$ .

**Table 3**  $\chi^2_{(2)}$  values determining the presence of non-allelic interactions and the components of the parental (P),  $F_1$ ,  $F_2$  and backcross (B) means of various crosses

Cross	Test of epistasis ( $\chi^2_{(2)}$ )	Estimates of components				Goodness of fit ( $\chi^2_{(1)}$ )
		$m$ †	$[h]'$	$[i]$	$[j]'$	
IR8 × IR29	7.20*	13.65 ± 0.10	22.82‡ ± 0.98	—	4.64 ± 2.98	4.8*
BPI × IR29	23.67***	12.45 ± 0.04	18.50 ± 0.55	—	6.78 ± 1.72	8.2***
IR24 × IR29	176.51***	Complex model				2.3ns
IR307 × IR29	18.54***	4.70 ± 0.17	-4.60 ± 0.47	-1.04 ± 0.26	—	
IR8 × IR307	31.14***	17.45 ± 0.23	12.11 ± 1.00	—	10.50 ± 2.23	9.0***
BPI × IR307	71.27***	12.33 ± 0.42	7.02 ± 1.05	3.83 ± 0.46	—	1.7ns
IR24 × IR307	197.00***	11.53 ± 0.20	-8.16 ± 0.72	—	16.74 ± 1.26	21.7***
IR8 × IR24	2.10ns	21.22 ± 0.10	10.96 ± 0.43	—	—	—
BPI × IR24	71.67***	20.00 ± 0.51	7.89 ± 0.40	—	11.46 ± 1.35	0.1ns
IR8 × BPI	28.62***	24.99 ± 0.27	5.60 ± 0.82	1.11 ± 0.29	—	14.2***

†See text for the definitions of various components and Table 2 for probability levels.

‡All estimates are highly significant (\*\*).

cant, (b) allowed maximum degrees of freedom for the  $\chi^2$  of goodness of fit and (c) had the  $\chi^2$  values non-significant were considered adequate. The estimates of various components obtained from such models and the corresponding  $\chi^2$  values are tabulated for various crosses in Table 4.

#### (v) Comparisons of within family variances

According to Perkins & Jinks (1970), the within family variances of  $P_1$  and  $P_2$  generations can differ either due to the differential effects of the scale or due to significant differences in the micro-environmental sensitivities of the parental genotypes. Variances of  $F_1$ ,  $F_2$ ,  $B_1$

**Table 4** Components of generation means and goodness of fit of the model for various crosses

Parameter	Crosses									
	IR8	BPI	IR24	IR307	IR8	BPI	IR24	IR8†	BPI	IR8
	× IR29	× IR29	× IR29	× IR29	× IR307	× IR307	× IR307	× IR24	× IR24	× BPI
$m‡$	13.64	12.45	7.55	4.70	17.34	12.35	11.26	21.18	20.03	23.60
± S.E.	0.10	0.41	0.10	0.17	0.23	0.41	0.23	0.14	0.10	0.33
$[d]$	12.12	12.12	7.54	4.13	10.91	7.31	3.30	6.75	5.51	2.08
± S.E.	0.21	0.14	0.10	0.18	0.36	0.35	0.26	0.20	0.16	0.15
$[h]_1$	8.34	5.43	—	-2.34	4.91	1.88	-3.25	2.39	3.61	5.75
± S.E.	0.37	0.53	—	0.27	0.80	0.50	0.44	0.36	0.38	0.52
$[h]_2$	15.24	13.15	-3.10	-2.23	6.34	5.25	-4.94	9.16	4.07	5.01
± S.E.	0.35	0.23	0.77	0.25	0.48	0.68	0.44	0.32	0.28	0.60
$[i]$	—	—	—	-1.08	—	3.73	—	—	—	2.47
± S.E.	—	—	—	0.22	—	0.45	—	—	—	0.34
$[j]_1$	—	—	6.50	—	5.96	—	9.68	—	—	—
± S.E.	—	—	0.77	—	2.19	—	0.91	—	—	—
$[j]_2$	-2.09	-5.03	-10.96	—	-7.36	—	-6.59	2.30	-12.91	8.22
± S.E.	1.08	0.64	1.14	—	0.97	—	0.93	0.76	0.66	0.96
$[l]$	—	15.25	44.86	—	17.63	—	—	—	—	—
± S.E.	—	6.27	3.77	—	5.88	—	—	—	—	—
$[c]$	1.51	0.32	—	-0.52	-0.98	1.52	0.55	-0.66	-0.65	-0.96
± S.E.	0.18	0.13	—	0.08	0.28	0.29	0.13	0.14	0.12	0.11
$[z]F_2$	-1.85	—	—	—	—	-2.41	—	—	—	—
± S.E.	0.45	—	—	—	—	0.37	—	—	—	—
$[z]B_1$	—	—	—	—	2.11	—	—	—	—	-0.85
± S.E.	—	—	—	—	0.39	—	—	—	—	0.24
$[z]B_2$	—	—	—	0.59	—	—	-0.73	—	0.98	—
± S.E.	—	—	—	0.15	—	—	0.20	—	0.35	—
$\chi^2$	6.2	2.4	2.9	2.6	0.0	4.6	0.6	8.6	1.9	5.3
d.f.	3	3	4	3	1	3	2	4	3	2

†See Table 7.

‡All components are significant ( $P < 0.05$ ) and all  $\chi^2$  values are non-significant ( $P > 0.05$ ).

and  $B_2$  generations, on the other hand, are not expected to display any reciprocal differences except when they are differentially affected by the maternal factors (Pooni *et al.*, 1992).

One-tailed variance ratios presented in Table 5 show that the parental variances differ significantly from one another in all cases except in cross  $IR8 \times IR24$ . Furthermore, this heterogeneity is mostly caused by the differential reactions of various genotypes to the micro-environmental variation and not by the scale [the

rank correlation between the means and variances of the parental genotypes ( $r = 0.4$ ) is rather low].

Differences between the variances of reciprocal families are also significant on 11 out of 40 occasions. Clearly, the frequency of these significances is much higher than the mandatory 5 per cent which is expected in theory. However, all except three of these differences attain a marginal significance (at the 5 per cent level) and therefore need not be considered important. Thus the variances of reciprocal families were treated as

homogeneous for further analysis and their pooled values are presented in Table 6 with the variances of the parental families.

(vi) Components of variances

With six variances available from each cross it is possible to estimate six components. According to

Pooni *et al.* (1992) and Perkins & Jinks (1970) these components must be  $D$ ,  $F'$ ,  $F''$ ,  $E_1$ ,  $E_2$  and  $E_3$ . When each genetic component is significant and  $E_1 (= VP_1)E_2/(= VP_2)E_3(= VF_1)$ , these components are estimated by a perfect fit solution. Otherwise, the model can be fitted by the iterative weighted least squares procedure of Hayman (1960). Contributions of these parameters to various generations are provided

**Table 5** Comparisons of the parental variances and those of reciprocal families of each generation

Cross	$P_1$ vs $P_2$	$F_1$ vs $RF_1$	$F_2$ vs $RF_2$	$B_1$ vs $RB_1$	$B_2$ vs $RB_2$
IR8 × IR29	405.00***	3.09*	1.20ns	48.13***	1.13ns
BPI × IR29	65.00***	3.35**	1.12ns	1.74*	1.00ns
IR24 × IR29	410.00***	1.31ns	1.21ns	1.12ns	1.85*
IR307 × IR29	1655.00***	1.30ns	1.27*	1.07ns	1.02ns
IR8 × IR307	4.08**	1.13ns	1.14ns	1.69*	1.21ns
BPI × IR307	25.46***	2.36ns	1.09ns	1.17ns	1.32ns
IR24 × IR307	4.04**	1.19ns	1.18ns	1.06ns	3.01*
IR8 × IR24	1.01ns	2.47ns	1.09ns	1.95*	1.09ns
BPI × IR24	6.31***	4.71**	1.19ns	1.43ns	1.03ns
IR8 × BPI	6.32***	1.06ns	1.02ns	1.53*	1.19ns

See Table 2 for probability levels.

**Table 6** Within family variances of various crosses

Cross		$VP_1$	$VP_2$	$VF_1$	$VF_2$	$VB_1$	$VB_2$
IR8 × IR29	Var.	0.81	0.002	4.67	136.18	5.64	160.03
	d.f.	19	19	38	792	178	150
BPI × IR29	Var.	0.13	0.002	3.90	106.46	3.53	134.62
	d.f.	19	19	38	792	154	176
IR24 × IR29	Var.	0.82	0.002	1.20	38.17	17.65	19.25
	d.f.	19	19	39	855	143	167
IR307 × IR29	Var.	3.31	0.002	0.46	6.26	6.31	2.92
	d.f.	19	19	38	782	149	177
IR8 × IR307	Var.	0.81	3.31	1.57	73.02	14.11	101.85
	d.f.	19	19	38	767	183	211
BPI × IR307	Var.	0.13	3.31	1.94	38.33	19.64	32.46
	d.f.	19	19	36	788	148	145
IR24 × IR307	Var.	0.82	3.31	2.37	15.80	8.48	4.65
	d.f.	19	19	38	720	157	179
IR8 × IR24	Var.	0.81	0.82	1.56	29.03	8.92	53.48
	d.f.	19	19	38	809	184	197
BPI × IR24	Var.	0.13	0.82	2.66	14.50	7.15	21.44
	d.f.	19	19	38	678	195	198
IR8 × BPI	Var.	0.81	0.13	1.60	11.36	8.66	13.93
	d.f.	19	19	38	793	192	221

**Table 7** Maximum likelihood estimates of the components of variances and the  $\chi^2$  values testing the adequacy of the model

Cross	Components							$\chi^2$ (d.f.)
	<i>D</i>	<i>F'</i>	<i>F''</i>	<i>E</i> <sub>1</sub>	<i>E</i> <sub>2</sub>	<i>E</i> <sub>3</sub>		
IR8 × IR29	est.	170.89	122.82	109.25	0.81	0.002	4.67	—
	± S.E.	14.89	34.24	9.92	0.26	0.0006	1.07	
BPI × IR29	est.	128.84	113.13	83.50	0.13	0.002	3.90	—
	± S.E.	11.62	26.60	7.73	0.04	0.0006	0.89	
IR24 × IR29	est.	66.24	-16.04	19.24	1.07	0.002	as <i>E</i> <sub>1</sub>	0.8(1)†
	± S.E.	3.59	4.43	3.40	0.20	0.0006		
IR307 × IR29	est.	10.17	-2.73	—	3.27	0.002	0.46	0.0(1)
	± S.E.	0.94	0.84		1.01	0.0006	0.11	
IR8 × IR307	est.	89.00	89.98	40.07	1.32	3.31	as <i>E</i> <sub>1</sub>	2.1(1)
	± S.E.	8.10	18.41	5.65	0.25	1.07		
BPI × IR307	est.	62.37	—	13.73	0.13	2.42	as <i>E</i> <sub>2</sub>	2.6(2)
	± S.E.	2.92		3.46	0.04	0.45		
IR24 × IR307	est.	26.57	-14.74	7.60	0.82	2.68	as <i>E</i> <sub>2</sub>	0.8(1)
	± S.E.	1.68	1.42	1.53	0.27	0.50		
IR8 × IR24‡	est.	30.07	58.42	8.38	1.19	as <i>E</i> <sub>1</sub>	as <i>E</i> <sub>1</sub>	3.7(2)
	± S.E.	3.60	9.85	2.62	0.19			
BPI × IR24	est.	15.17	22.40	—	0.13	0.83	2.60	3.2(1)
	± S.E.	1.52	3.64		0.04	0.27	0.59	
IR8 × BPI	est.	16.03	8.94	—	1.34	0.13	as <i>E</i> <sub>1</sub>	2.3(1)
	± S.E.	1.26	2.32		0.25	0.04		

†All  $\chi^2$  values are non-significant ( $P > 0.05$ ) and all estimates are significant ( $P \leq 0.05$ ).

‡Component estimates differ from Pooni *et al.* (1992) due to the misreading of *RB*<sub>1</sub> variance in the latter case.

below (see Pooni *et al.*, 1992 for definitions of *D*, *F'* and *F''* etc.).

	<i>D</i>	<i>F'</i>	<i>F''</i>	<i>E</i> <sub>1</sub>	<i>E</i> <sub>2</sub>	<i>E</i> <sub>3</sub>
VP <sub>1</sub>	0	0	0	1	0	0
VP <sub>2</sub>	0	0	0	0	1	0
VF <sub>1</sub>	0	0	0	0	0	1
VF <sub>2</sub>	0.556	0.167	0.167	0.25	0.25	0.50
VB <sub>1</sub>	0.444	0	-0.667	0.50	0	0.50
VB <sub>2</sub>	0.444	0.667	0	0	0.50	0.50

It is apparent from the comparisons of within family variances that *E*<sub>1</sub> = *E*<sub>2</sub> in one cross (cross IR8 × IR24) only. Comparisons of the *F*<sub>1</sub> and parental variances further revealed that *E*<sub>3</sub> was also equal to *E*<sub>1</sub> and *E*<sub>2</sub> in this case. *E*<sub>3</sub> was also found to be equal to *E*<sub>1</sub> or *E*<sub>2</sub> in five other crosses. In these cases therefore there was no need to allocate three separate environmental parameters and their numbers were reduced to 1 and 2 respectively for model fitting. Furthermore, contribu-

tions of the new parameters to various generations were determined by adding the coefficients of those parameters which they replaced. For example, *E*<sub>1</sub>, *E*<sub>2</sub> and *E*<sub>3</sub> were replaced by a single *E* in IR8 × IR24 and it was given a coefficient of 1 throughout, which is the sum of the coefficients of the former parameters for each variance.

Estimates of the components of variances and the  $\chi^2$  values determining the goodness of fit of the models for various crosses are presented in Table 7.

*(vii) Comparison of the new model with the disomic model*

The efficiency of the new model can be easily compared with the already existing additive/dominance or digenic interaction model based on disomic inheritance/expression by applying both models to the same data and comparing their  $\chi^2$  of goodness of fit. When

**Table 8** Goodness of fit of the models based on disomic inheritance/expression when fitted to generation means and the level of improvement achieved by fitting the new model

Cross	Additive/ dominance model		Model with epistasis		Improvement due to the new model	
	$\chi^2$	d.f.	$\chi^2$	d.f.	$\chi^2$	d.f.
IR8 × IR29	30.9***‡	5	8.9ns	4	2.7ns	1
BPI × IR29	73.3***	6	13.9**	4	11.5***	1
IR24 × IR29	635.1***	6	551.8***	5	548.9***	1
IR307 × IR29	153.8***	5	112.4***	4	109.8***	1
IR8 × IR307	151.5***	5	83.1***	3	83.1***	2
BPI × IR307	95.0***	5	10.3*	4	5.8*	1
IR24 × IR307	278.5***	5	276.7***	4	276.1***	2
IR8 × IR24	34.7***	6	15.4**	5	7.0**	1
BPI × IR24	384.7***	5	108.6***	4	106.8***	1
IR8 × BPI	156.9***	5	130.2***	3	124.9***	1

†Calculated as the difference between column 4 above and the  $\chi^2$  values given in Table 4.

‡See Table 2 for probability levels.

the  $\chi^2$  take different degrees of freedom then the difference can be calculated and its significance tested in the usual manner. Otherwise, the model with a lower  $\chi^2$  value must be considered the more efficient.

The models were compared both at means and variance levels. Initially, the  $m$ ,  $[d]$  and  $[h]$  parameters of an additive/dominance model (see Mather & Jinks, 1982 for definitions etc.) and the requisite  $[c]$  and  $[z]$  parameters of the maternal effects (based on the results in Table 4) were fitted to generation means of each cross with a view to determining the presence of epistasis on the disomic model. Later the model was extended to digenic interactions by including  $[i]$ ,  $[j]$  and  $[l]$  parameters. However, to keep parity between the two models these parameters were added to the model only when the corresponding  $[i]$ ,  $[j]_1$  and/or  $[j]_2$  and  $[l]$  parameters of the new model were observed to be significant (in Table 4). The results presented in Table 8 show that the new model explains a significantly higher level of genetic variability than the disomic model in nine out of 10 crosses.

A similar procedure was followed to compare the models at variance level. Parameters  $D$ ,  $F'$  and  $F''$  of the new model were replaced by  $D$  and  $F$  of the disomic additive/dominance model of Mather & Jinks (1982) while  $E_1$ ,  $E_2$  and  $E_3$  were retained as such.  $\chi^2$  values determining the goodness of fit of the disomic model and differences between these and the corresponding values given in Table 7 are tabulated in Table 9.

Once again the new model seems to explain a significantly higher level of genetic variation in eight crosses. Furthermore, in four cases this improvement has been largely due to a change of model as the two models have the same number of parameters.

#### (viii) Parental diversity and genetic variability

When parents differ at several loci, the gene effects are more or less equal and the allelic frequencies are in equilibrium, genetic variability displayed by a random set of crosses is not expected to correlate critically with the deviations between the parental lines. Failure of one or more of the above conditions, however, can lead to a significant association between parental diversity and genetic variability in the crosses. In the present study the parental lines stratify virtually the whole range of amylose levels from 0 to 28 per cent. Furthermore, the inheritance of amylose is known to be controlled by genes with unequal effects and a single QTL located on chromosome 6 is assumed to control most of the variability displayed by the trait (Kumar & Khush, 1986, 1988). Therefore, a correlation between parental diversity and genetic variability is expected to be high among these crosses.

Component  $[d]$  of various crosses was used as a measure of parental diversity and the variances of  $F_2$ ,  $B_1$  and  $B_2$  generations provided alternative estimates of genetic variability. Correlations between these statistics were calculated and their significance tested using



**Table 9** Goodness of fit of the model based on disomic inheritance/expression when fitted to the within family variances and the level of improvement achieved by applying the new model

Cross	Models based on diploid inheritance/expression		Improvement due to the new model	
	$\chi^2$	d.f.	$\chi^2$	d.f.
IR8 × IR29	2.4ns	1	2.4ns	1
BPI × IR29	5.2*	1	5.2*	1
IR24 × IR29	1.1ns	2	0.3ns	1
IR307 × IR29	4.9*	1	4.9‡	—
IR8 × IR307	26.1***	2	24.0***	1
BPI × IR307	9.9**	2	7.3‡	—
IR24 × IR307	9.8**	2	9.0**	1
IR8 × IR24	78.9***	3	75.2***	1
BPI × IR24	43.4***	1	40.4‡	—
IR8 × BPI	59.1***	2	56.8***	1

See Table 2 for probability levels.

†Calculated as difference between column 2 and the  $\chi^2$  values given in Table 7.

‡These  $\chi^2$  have 0 d.f.

standard procedures. These correlations ( $r$ ) and those between  $[d]$  and  $\sqrt{D}$  are listed below.

Statistics	$r$	Significance
$[d]$ and $VF_2$	0.92	***
$[d]$ and $VB_1 + VB_2$	0.94	***
$[d]$ and $\sqrt{D}$	0.93	***
Rank correlation of $[d]$ and $\sqrt{D}$	0.89	***

*(ix) Transgression among crosses*

Individuals with amylose content either  $> \bar{P}_1$  or  $< \bar{P}_2$  were counted and their highest and lowest scores noted for the  $F_2$ ,  $B_1$  and  $B_2$  generations of various crosses. The results of Table 10 show that the numbers of these individuals differ not only between crosses but also between generations and between categories ( $> \bar{P}_1$  and  $< \bar{P}_2$ ). While the observed transgression in the crosses between IR8, BPI, IR24 and IR307 can be explained by invoking gene dispersion between the parental lines, interpretation of the results of the crosses involving IR29, however, is not so straightforward. These results are discussed in detail in the next section.

**Interpretation and conclusion**

*(i) Gene action*

The  $F_1$  of various crosses show significant heterosis on three occasions (see Table 1). However, on each

occasion only one of the reciprocal  $F_1$ s outperforms the  $P_1$  while the other does not. This suggests that reciprocal effects play an important role in determining the overall level of hybrid vigour displayed by these crosses. Reciprocal differences are also significant in nine crosses and cytoplasm is the main source of these deviations. It is further possible that some of these differences are caused by the delayed maternal effects and/or differential interaction between the cytoplasm and the nuclear genes (see Table 2 and Pooni *et al.*, 1987).

Table 3 confirms the widespread occurrence of epistasis. All three types, namely additive × additive, additive × dominance and dominance × dominance interactions are detected and each cross shows some significant interaction. Nevertheless, the inheritance of amylose content seems to be rather simple in the IR8 × IR29 cross as only one epistatic component ( $[j]_2$ ) is detected to be significant. At the other end, the IR24 × IR29 cross easily qualifies as the most interactive because it has three epistatic components, ( $[j]_1$ ,  $[j]_2$  and  $[l]$ ) that are significant.

Table 4 reveals the truly complex nature of the genetic control of amylose content in the materials under investigation. Altogether 12 components explain all the significant variability among 10 sets of basic generations and the simple additive/dominance model is not adequate in any case. However, the additive genetic and dominance components still seem to be the most important sources of variation in all the crosses. The magnitude of  $[d]$  differs considerably between

**Table 10** Numbers of transgressive segregants and the ranges of phenotypic scores observed in the F<sub>2</sub>, B<sub>1</sub> and B<sub>2</sub> generations of various crosses

Cross	F <sub>2</sub> generation				B <sub>1</sub> generation				B <sub>2</sub> generation			
	>P <sub>1</sub>	<P <sub>2</sub>	Range	<i>n</i>	>P <sub>1</sub>	<P <sub>2</sub>	Range	<i>n</i>	>P <sub>1</sub>	<P <sub>2</sub>	Range	<i>n</i>
IR8 × IR29	244	0	0-34	794	12	0	0-28	180	29	0	0-30	152
BPI × IR29	88	0	0-29	794	4	0	0-26	156	3	0	0-26	178
IR28 × IR29	164	0	0-21	857	5	0	0-17	145	0	0	0-15	169
IR307 × IR29	64	0	0-16	784	14	0	0-12	151	0	0	0-8	179
IR8 × IR307	181	97	2-34	769	18	0	13-30	185	11	1	1-30	213
BPI × IR307	0	135	1-26	790	10	0	12-28	150	0	22	1-28	147
IR24 × IR307	61	138	1-18	722	0	39	1-15	159	0	86	1-12	181
IR8 × IR24	244	40	10-34	811	104	0	23-35	186	15	90	8-30	199
BPI × IR24	96	39	12-29	680	2	8	10-26	197	10	54	9-27	200
IR8 × BPI	277	270	16-35	795	94	22	19-36	194	77	90	18-34	223

crosses confirming a high degree of gene association in the parental lines. Absolute values of  $[h]_1$  are also comparatively smaller than those of  $[h]_2$  on eight occasions, indicating that a single dose of the dominant allele is normally not sufficient to compensate fully for the two recessive alleles, at least at some of the loci. The directional impact of  $[h]_2$  on the F<sub>1</sub> performance, on the other hand, is reinforced by  $[j]_2$  in four crosses (where  $[h]_2$  and  $[j]_2$  have opposite sign) while  $[j]_1$  and  $[h]_1$  take the same sign in only one cross.

Analyses of second-degree statistics further confirm the predominantly additive/dominance nature of the genetic variability for amylose content. Component *D* is highly significant throughout and components *F'* and/or *F''* are detected in all the crosses (see Table 7). As the latter components represent the cross products between the additive ( $d_i$ ) and the dominance effects ( $h_{i1}$  and  $h_{i2}$ ) of various loci, their significance clearly indicates the presence of dominance variation. Strong association between the variance components (*D*, *F'* and *F''* etc.) and phenotypic diversity among the parents further suggests that the genes controlling amylose content are either linked tightly in coupling phase or have markedly unequal effects.

The amylose content is also observed to be highly affected by genotype × micro-environmental interactions like any other metric trait (see Tables 5 and 7).

### (iii) Type of epistasis

Mather (1967) classified digenic interactions (e.g. between *A/a* and *B/b* loci) into predominantly complementary and duplicate types and obtained the classical 9:7 and 15:1 ratios in the F<sub>2</sub> of a dihybrid assuming

$d_a = d_b = h_a = i_{ab} = j_{ab} = j_{ba} = l_{ab}$  and  $d_a = d_b = h_a = h_b = -i_{ab} = -j_{ab} = -j_{ba} = -l_{ab}$  respectively (see Mather & Jinks, 1982 for symbols and definitions). He further showed that the true nature of epistasis can only be determined from the relative signs of the dominance and the dominance × dominance interaction ( $[h]$  and  $[l]$ ) components because the magnitudes and signs of the remaining components are often distorted by gene dispersion. Complementary epistasis is assumed to predominate when  $[h]$  and  $[l]$  take the same sign (i.e. either + or -) and duplicate epistasis is identified when these components have opposing signs.

In theory, the same relationships also hold for the trisomic model. For example, the 9:7 ratio of complementary epistasis is obtained when  $d_a = d_b = 3/2h_{a1} = 3/4h_{a2} = 3/2h_{b1} = 3/4h_{b2} = i_{ab} = 3/2j_{ab1} = 3/4j_{ab2} = 3/2j_{ba1} = 3/4j_{ba2} = 9/4l_{a1b1} = 9/8l_{a1b2} = 9/8l_{a2b1} = 9/16l_{a2b2}$  and this ratio is modified to 15:1 when all the epistatic components take a negative sign (see Pooni *et al.*, 1992 for symbols and definitions). Therefore, the type of epistasis prevailing in the present materials can be determined from the signs of  $[h]_1$  and  $[l]_1$  and  $[h]_2$  and  $[l]_2$  respectively, whenever these components are detected to be significant. Alternatively, we can compare the signs of  $[l]$  and  $[h]_1$  or  $[h]_2$ , particularly when the latter components do not take opposing signs.

On these criteria, we identify complementary gene interaction in BPI × IR29 and IR8 × IR307 crosses and duplicate epistasis in the IR24 × IR29 cross. Non-allelic interactions, however, cannot be classified in the remaining crosses because all the dominance × dominance interaction components are non-significant throughout.

*(iii) Dominance and non-additivity*

In general, the relative importance of the dominance variance is determined from the dominance ratio which is calculated (for disomically inherited/expressed traits) as  $\sqrt{H/D}$  (see Mather & Jinks, 1982) and interpreted as partial (d.r. < 1), complete (d.r. = 1) or super (d.r. > 1). In the present case it is possible to calculate two dominance ratios and they are  $\sqrt{H_1/D}$  and  $\sqrt{H_2/D}$  respectively. The interpretation of these ratios, however, is not as straightforward as that of  $\sqrt{H/D}$  because their expected values for complete dominance are not unity but 2/3 and 4/3 (these ratios change to 4/3 and 2/3 when dominance is negative). However, simple modifications of the above ratios to  $\sqrt{H_1/0.4444D}$  and  $\sqrt{H_2/1.7777D}$ , respectively, not only allow a standard interpretation of the dominance components of the new model but also facilitate an examination of the extent of dosage compensation that can be achieved by substituting the recessive alleles by one, two or three dominant alleles at various loci.

In the present case, it is not possible, however, to estimate the dominance ratio by the above methods because neither are independent estimates of  $H_1$  and  $H_2$  available nor is the additive/dominance model adequate in any cross. Nevertheless, ratios  $\{(\bar{F}_1 - m - 1/3[d] - [c]) / (2/3(\bar{P}_1 - m - [c]))\}$  and  $\{(\bar{R}\bar{F}_1 - m + 1/3[d] + [c]) / \{4/3(\bar{P}_1 - m - [c])\}\}$  [or  $\{(\bar{F}_1 - m + 1/3[d] - [c]) / \{4/3(\bar{P}_2 - m - [c])\}\}$  and  $\{(\bar{R}\bar{F}_1 - m - 1/3[d] + [c]) / \{2/3(\bar{P}_2 - m - [c])\}\}$  where dominance is towards the lower score] can provide some indication of the level of non-additivity displayed by the families of various crosses, particularly of the first four where the genes are highly associated in the parental lines. These ratios are 1.03 and 0.98, 0.69 and 0.92, 0.55 and 0.00 and 0.38 and 0.98, respectively, indicating that non-additive variation becomes more important as differences between the parental lines increase. Furthermore, its direction also seems to change with the parental diversity as the least diverse crosses show partial dominance for low amylose and most diverse crosses exhibit complete dominance for high amylose. It is also apparent that a single dominant allele is often insufficient for the fuller expression of the non-additive effects and the presence of a second allele enhances these effects by up to 1.6 times its additive value (= 1.6/3[d]). The averaged dosage effect of the second allele, however, remains low (about 0.46/3[d]) in BPI × IR29 cross.

*(iv) Transgression*

The 10 crosses investigated in this study fall distinctly into two groups, namely all those involving IR29 and

the rest. Kernels of IR29 do not have any measurable (0 per cent) amylose which is the lowest possible score for this trait. Consequently it must be considered the extreme homozygote for low score, which implies that it has decreasing alleles fixed at all the loci controlling amylose content. In comparison, all the remaining parents have significantly higher scores than IR29 and therefore must possess the alternative alleles that increase amylose content. This means that the crosses involving IR29 must be associated at all the loci for which they are segregating and the recombinant genotypes are not expected to transgress their parental scores unless there is super dominance and/or super epistasis for high amylose content.

Parents of the remaining crosses, on the other hand, may not subscribe to the above conditions and consequently gene dispersion cannot be ruled out in these crosses.

Transgression for high amylose in the first four crosses (see Table 10) confirms that dominance is high and epistasis has a marked effect on the phenotypic performance of the transgressive genotypes in these crosses. The combined rate of transgression (for high amylose) in these crosses is also significantly less than that in the remaining crosses ( $\chi^2_{(1)}$  is 4.36\* for  $F_2$ , 75.5\*\*\* for  $B_1$  and 14.74\*\*\* for  $B_2$ ) suggesting that the alleles are indeed dispersed in the latter crosses.

*(v) Practical implications*

The present investigation has two important implications for rice breeding. The first concerns the inheritance of amylose content *vis à vis* its manipulation for breeding purposes. It is apparent that the genetic control of amylose content is indeed complex. Not only is it controlled by two or more genes showing epistasis but it is also highly influenced by cytoplasmic effects and the interaction between cytoplasm and nuclear genes. Consequently, amylose content will be better manipulated if it is treated as a quantitative trait and due consideration is given to the cytoplasmic effects when choosing the parental sources.

The second implication concerns the exploitation of heterosis for which there is currently an enormous interest (see Virmani *et al.*, 1982). The predominant characteristics of the non-additive variation revealed by the present study (e.g. partial to complete dominance ratio, partial- to over-dosage compensation and changing direction of dominance) suggest that  $F_1$  hybrids with desired levels of amylose can be produced without much difficulty. This implies that cooking quality will not be an impediment to hybrid breeding and that other agronomic traits will determine the commercial worth of the hybrid varieties.

## Acknowledgements

Dr Ish Kumar was supported by an Association of Commonwealth Universities (ACU) Fellowship.

## References

- 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.
- HAYMAN, B. I. 1960. Maximum likelihood estimation of genetic components of variation. *Biometrics*, **16**, 369-381.
- JULIANO, B. O. 1971. A simplified assay for milled rice amylose. *Cereal Sci. Today*, **16**, 334-338.
- KUMAR, I. AND KHUSH, G. S. 1986. Genetics of amylose content in rice (*Oryza sativa* L.). *J. Genet.*, **65**, 1-11.
- KUMAR, I. AND KHUSH, G. S. 1987. Genetic analysis of different amylose levels in rice. *Crop Sci.*, **27**, 1167-1172.
- KUMAR, I. AND KHUSH, G. S. 1988. Inheritance of amylose content in rice (*Oryza sativa* L.). *Euphytica*, **38**, 261-269.
- MATHER, K. 1967. Complementary and duplicate gene interactions in biometrical genetics. *Heredity*, **22**, 97-103.
- 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.
- POONI, H. S., KUMAR, I. AND KHUSH, G. S. 1992. A comprehensive model for disomically inherited metrical traits expressed in triploid tissues. *Heredity*, **69**, 166-174.
- VIRMANI, S. S., AQUINO, R. C. AND KHUSH, G. S. 1982. Heterosis breeding in rice (*Oryza sativa* L.). *Theor. Appl. Genet.*, **63**, 373-380.