Theor. Appl. Genet. 56, 233-239 (1980)

Inheritance of Seed Protein Content and Other Agronomic Characters in Long Bean (*Vigna sesquipedalis* Fruw.)

C. Mak
Department of Genetics and Cellular Biology, University of Malaya, Kuala Lumpur (Malaysia)

T.C. Yap
Department of Agronomy and Horticulture, University of Agriculture, Serdang (Malaysia)

Summary. Seven varieties of long bean, which included three local and four exotic, were crossed in a complete diallel. This was an attempt to study the inheritance of crude protein content, protein yield, flowering date, pod yield and yield components.

Both additive and non-additive gene effects were responsible for the genetic variation in the diallel population. However, dominance variance was more important than additive variance in crude protein content, number of pods per plant and number of seeds per pod. For seed weight and pod length, additive variance was more important.

The crude protein content, protein yield and number of pods per plant appeared to be controlled by overdominance effects. Partial dominance seemed to be the case for flowering date, pod length and seed weight; complete to overdominance for pod yield. High protein appeared to be associated with recessive genes whereas there was a general trend of high yielding parents carrying more dominant genes.

Key words: *Vigna sesquipedalis* – Seed protein – Yield– Yield components

Introduction

Long bean is a tropical, short term annual legume grown mainly for its tender pods. It has a crude protein content of about 28% (Mak and Yap 1977a) and is, therefore, an important source of cheap vegetable protein for many people in the tropics. Breeding and selection have been carried out for many years in this crop (Mak and Yap 1977a, b; Yap et al. 1977b). This paper reports studies of the inheritance of seed protein content, yield and yield components and flowering date in a 7 × 7 diallel cross.

Materials and Methods

In 1973, a 7 × 7 diallel cross, including the reciprocals, was made among seven varieties which included three local and four exotic ones. The seven parents used were (1) 'PI 286439', (2) 'Taiwan Stripe seed', (3) 'Taiwan Purple Pod', (4) 'Yates', (5) 'Kapok', (6) 'Kim San' and (7) 'Local Long' (hereafter the parents are referred to as parents 1 to 7, respectively). On the basis of an evaluation trial involving 22 cultivars (Mak and Yap 1977a), these parents were chosen on three criteria:

(a) genetic diversity as indicated by their sources,
(b) differences in seed protein content and
(c) differences in yielding ability.

The local varieties are relatively high-yielding with medium protein content: 'Kapok' being the most popular local variety, 'Kim San' having the largest seed size and 'Local Long' the longest pod length. 'Yates' was chosen for being a good yielder and 'Taiwan Stripe Seed', 'Taiwan Purple Pod' and 'PI 286439' for being high protein contein varieties.

The F_1 and parent seeds were sown in a sandy loam soil at the University Farm University of Malaya. A partially balanced lattice design with three replications was used. Ten seeds per plot were planted in raised beds 20 cm high and 75 cm wide. Distance between planting points was 30 cm. Since long bean is a climber, a trellis was constructed for support. The trellis consisted of heavy posts set up at the end of each bed and with two strands of wire fastened to the posts, one at the top and the other one near the base. The top wire was 200 cm and bottom wire was 30 cm above the ground. Raffia was strung between the two wires in a zigzag manner and each strand of the raffia joining the top and the bottom wires accomodated one plant. Ammonium sulfate was applied to the soil at the rate of 600 kg/ha.

Seeds of mature pods from the first three harvests were used for the determination of total N. Total N was determined by means of the Kjeldahl method and multiplied by 6.25 to give the percentage of seed protein. Other characters studied were (a) pod yield per plant, g (plants in the whole plot were harvested at 2 to 3 day intervals for a period of 4 to 5 weeks and data from all harvests was accumulated and divided by the total number of plants to obtain seasonal estimates of yield); (b) pods per plant

0040-5752/80/0056/0233/$ 1.40

(total number of pods per plot divided by the total number of plants per plot); (c) seeds per pod (mean of a total sample of 50 pods over the harvesting period); (d) seed weight, g (mean a four 100-seed samples); (e) pod length, cm (mean of a total sample of 50 pods over the harvesting period); (f) protein yield, g (total seed yield per plant X percent crude protein content), (d) flowering date, days (number of days from planting to the time when more than 50% of the plants in the plot with at least two fully open flowers).

Results

(i) Analysis of Variance

Significant mean square values of 'a' and 'b' were found for yield, seed protein content, protein yield, pod length, 100 seed weight, number of seeds per pod, number of pods per plant and flowering date. This indicated that both additive and dominance effects were responsible for the genetic variation (Table 1).

Unidirectional gene action was the cause of significant dominance effects in number of pods per plant, protein yield, yield and pod length as indicated by the significant mean square values of 'b_1' for these traits. Significant mean squares of 'b_2' were found for seed protein content, pod length and number of seeds per pod, suggesting the dominance effects of these traits were probably caused by the asymmetrical gene distribution. Dominance effects not due to the above two causes were found in seed weight and flowering date.

Highly significant maternal effects were detected for pod length and number of seeds per pod. Reciprocal differences were also found for number of pods per plant and protein yield. However, the reciprocal differences which were not significant might be due to non-heritable variations since any variations not ascribable to 'a' and 'b' would be included in either 'c' or 'd'.

(ii) Genetic Components of Variation

The homogeneity tests of $W_r - V_r$ indicated that Hayman's assumptions were valid for all traits except pod length. All the regression coefficients of W_r on V_r for the traits were not significantly different from unity except

Table 1. Mean square values for the items of diallel analysis for the eight agronomic characters

Item	df	Pods/ plant	Seed weight	Protein yield	Flowering date	Pod yield	Protein content	Pod length	Seeds/ pod
a	6	151.16^b	52.79^b	82.15^b	35.34^b	12412.34^a	9.77^b	358.33^b	4.04^b
b	21	46.83^b	1.73^b	21.51^a	3.19^a	4012.99^a	3.18^b	28.77^b	2.53^b
'b_1'	1	283.27^b	0.88	133.51^b	0.63	26570.88^b	3.85	30.19^b	1.79
'b_2'	6	33.74	1.77	9.12	2.89	898.09	3.10^b	25.18^b	2.54^b
'b_3'	14	35.55	1.77^a	18.82	3.50^a	3736.27	3.18^b	30.20^b	2.58^b
c	6	20.03	1.87	8.03	1.48	2520.43	1.55	14.77^b	2.41^b
d	15	47.65^a	1.55	21.73^a	1.57	4487.91	1.29	4.02	1.69^b
Error	96	18.54	0.83	10.13	1.76	2740.61	0.87	3.43	0.58

$^{a, b}$ $0.05 > p > 0.01$

Table 2. Components of variation and other statistics for the eight agronomic characters

Item	Pods/ plant	Seed weight	Protein yield	Flowering date	Pod yield	Protein content	Pod length	Seeds/ pod
D	7.48^b	7.02^b	6.98	0.97^b	444.24^a	0.66	38.79^b	0.57^b
H_1	22.47^b	0.82	7.35^b	1.23^a	409.73	2.25^a	22.08^a	1.77^b
H_2	18.85^b	0.59	7.58^b	0.95^a	848.31^a	1.54^a	16.89^a	1.30^b
F	−1.53	2.29^b	−0.11	-1.95^b	-915.59^a	0.52	10.17	0.71
$D - H_1$	-14.99^b	6.20^b	−0.36	−0.26	34.50	-1.58^a	16.71^a	-1.19^b
$(H_1/D)^{\frac{1}{2}}$	1.73	0.34	1.02	1.12	0.96	1.83	0.75	1.74
$H_2/4H_1$	0.20	0.18	0.25	0.19	0.51	0.17	0.19	0.18
K_D/K_R	0.88	2.83	0.98	0.05	−0.03	1.54	1.42	2.09
h^2	43.22^b	0.01	20.14^b	−0.18	3890.68^b	−0.08	4.37	0.19
$K = h^2/H_2$	2.29	0.01	2.65	−0.19	4.58	−0.05	0.25	0.15
h^2N	0.13	1.05	0.24	0.14	0.08	0.18	0.70	0.24
E	6.18	0.27	3.37	0.58	913.44	0.29	1.14	0.19

$^{a, b}$ $0.05 > p > 0.01$

the one for seed protein content. However, the regression coefficients for number of seeds per pod and pod length were also not significantly different from slop zero (b = 0). These results revealed that some of the traits showed at least partial failure of the assumptions. Even though the assumptions might not be strictly fulfilled, as suggested by Hayman (1954a), the components of variation could be estimated by using the mean value of cross and its reciprocal for the analysis.

With this in mind, the genetic parameters, D, H_1, H_2 and F and their standard errors were estimated for each trait. These together with the ratios of genetic component and heritability estimates are presented in Table 2.

Both D and H were significant in number of pods per plant, protein yield, flowering date, pod length and number of seeds per pod, suggesting that both additive and dominance effects were responsible for the variation observed in the diallel cross population. The value D-H_1 shows the relative importance of additive and dominant gene action. The values were negative and highly significant for number of pods per plant and number of seeds per pod. Therefore, the dominance effects were more important than additive effects for these traits. Negative but non-significant values of D-H_1 were also observed for protein yield and flowering date. In the case of pod length, additive gene effect was more important than dominant gene effect since the value of D-H_1 was positive and significant, statistically.

For the seed weight, the variation was mainly due to additive gene action since D was highly significant and H_1 and H_2 were non-significant. A different situation was found for seed protein content where H_1 and H_2 were significant but D was not significant.

The average degree of dominance over all loci was estimated by the value $(H_1/D)^{\frac{1}{2}}$. Overdominance was apparent for number of pods per plant, protein content and number of seeds per pod as the values of $(H_1D)^{\frac{1}{2}}$ were greater than unity. Complete dominance or over-dominance was found for protein yield and flowering date since the values of $(H_1D)^{\frac{1}{2}}$ were slightly greater than unity. Dominance was incomplete for seed weight and pod length. For pod yield, partial to complete dominance was indicated and $(H_1/D)^{\frac{1}{2}}$ was 0.96, which is very close to unity

An excess of dominant genes was found for seed weight, protein content, pod length and number of seeds per pod since the values of K_D/K_H were greater than unity and the F values were positive. The values of K_D/K_H were less than unity for number of pods per plant, protein yield, flowering date and yield, indicating that the parents carried more recessive genes than dominant genes for these traits. The negative F values found for these traits also confirms the finding. According to the theory of diallel analysis, when positive and negative genes are equally distributed in the parents, the value of $H_2/(4H_1)$ is 0.25. This value was not found in all the traits except protein yield, indicating that positive and negative genes were not equally distributed in the parents.

The quantity h^2/H_2 or K estimates the number of groups of genes exhibiting dominance to some degree. It appeared that at least three groups of genes were involved for number of pods per plant and protein yield and at least five groups for yield.

The narrow sense heritability was very high for seed weight and pod length. For the other agronomic characters, the heritability values were relatively very low (Table 2).

(iii) Graphical Analysis

The seven pairs of W_r and V_r statistics obtained for each array for the eight agronomic characters in the diallel cross population were plotted in the form of a variance-covariance diagram (Figs. 1-9).

For each trait, the points for the parental arrays were widely separated along the regression line indicating that the genetic background of the parents was quite diverse.

A general tendency for parents having the greater number of recessive genes to be low yielding and those possessing dominant genes to be high yielding was shown for pod yield (Fig. 1 and Table 3). Overdominance for pod yield was indicated by the regression line which cuts the ordinate below the origin. Similar results were also obtained for protein yield (Fig. 2). This was expected as protein yield was highly correlated with pod yield (r = 0.87**).

In the case of crude protein content, Hayman's assumptions were not strictly fulfilled since the regression

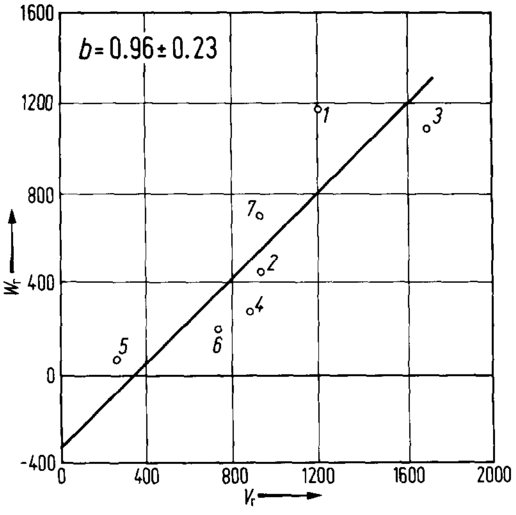

Fig. 1. The regression of W_r on V_r for pod yield

Table 3. Mean performance of parents and correlation coefficient of parents with ($W_r + V_r$) values

	Pods/ plant (gm)	Seed weight (gm)	Protein yield/ plant (gm)	Flowering date (days)	Pod yield/ plant (gm)	Protein content (%)	Pod length (cm)	Seeds/ pod
1.	12.58	9.23	5.24	44.50	86.15	25.79	34.94	17.50
2.	14.68	11.99	8.18	46.00	147.69	28.33	35.58	16.40
3.	11.17	11.78	5.39	49.00	105.72	26.82	33.88	15.27
4.	22.53	13.46	14.21	45.30	181.40	26.21	37.60	17.88
5.	15.98	14.14	10.48	44.50	178.13	26.46	47.80	17.52
6.	14.59	18.05	11.77	45.20	150.49	25.20	42.85	17.74
7.	17.41	12.64	9.71	45.00	171.15	26.62	50.44	16.58
r.	0.22	0.62	0.53	0.85	0.82	0.68	0.69	0.67

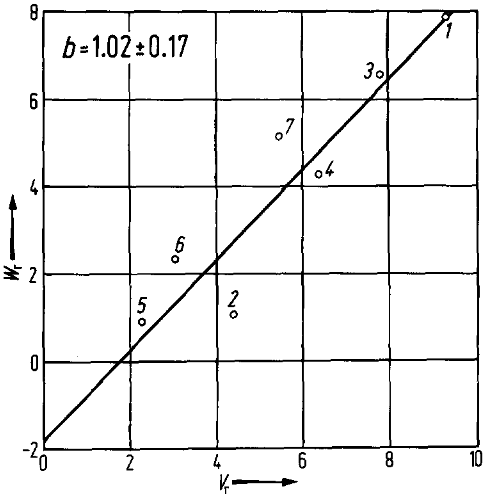

Fig. 2. The regression of W_r on V_r for protein yield

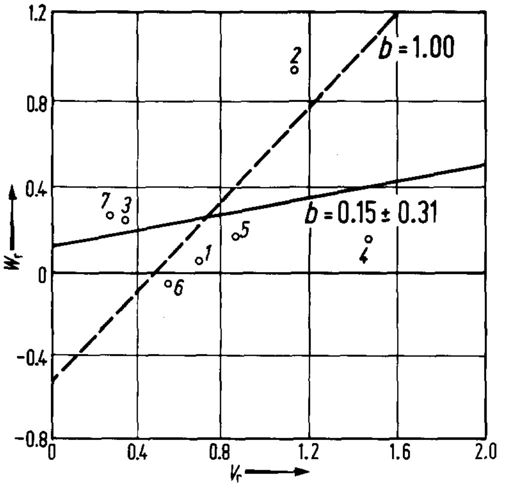

Fig. 3. The regression of W_r on V_r for crude protein content. Solid line is the best fitting regression line while broken line indicates theoretical slope of unity

coefficient was significantly different from unity and not significantly different from zero. Removal of progenies of parental array 4 resulted in the regression line to be not significantly different from unity, but it was also not significantly different from zero. Similar results were obtained by removing progenies of each parental array one at a time. Removal of two or three parental arrays could improve the regression line, but non-allelic gene interaction was still apparent. This finding suggests that the inheritance of protein content in long bean is very complicated. In consequence, the conclusion drawn from the interpretation of (V_r,W_r) graph might not be reliable for this trait. Nevertheless, in this analysis it appeared that high protein content was associated with recessive genes since parent 2 which carried the greatest number of recessive genes had the highest crude protein content (Fig. 3 and Table 3). For number of pods per plant, there appeared to be a tendency for a greater number of recessive genes to be associated with a greater number of pods per plant. However, an exception was observed for parent 1 which had the greatest number of recessive genes but carried a small number of pods per plant. This discrepancy might suggest that gene controlling this trait are complicated and their expression may be greatly affected by environment. Overdominance for this trait was indicated by the regression line which passed below the origin (Fig. 4).

Partial dominance apparently influenced seed weight since the regression lines intersected the W_r axis above the origin (Fig. 5). Variety 6 had the largest seed size and contained the greatest number of recessive genes. Dominance was directed towards small seed size. This was also indicated by the positive correlation (r = 0.62) between $V_r + W_r$ values and the parental values although the correlation coefficient is not significant statistically (Table 3).

For pod length, the slope of the regression line was not significantly different from zero or unity, suggesting that epistatis existed for this trait. Variety 4 appeared to be responsible for this non-allelic gene interaction since removal of the progenies of array 4 made the regression significantly different from zero but not from unity (Fig. 6, 7). Partial dominance was indicated since the regression line cut the W_r axis above the origin. Variety 5 carried the greatest number of recessive genes while variety 1 carried the greatest number of dominant genes. Though the correlation between $V_r + W_r$ values and the parents was not significant, longer pod length seemed to be associated with recessive genes.

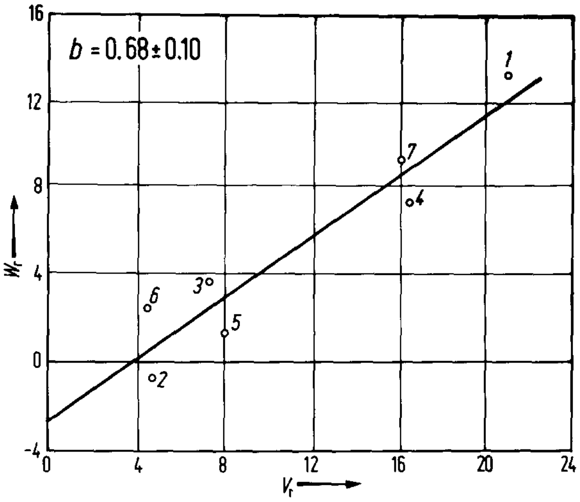

Fig. 4. The regression of W_r on V_r for number of pods per plant

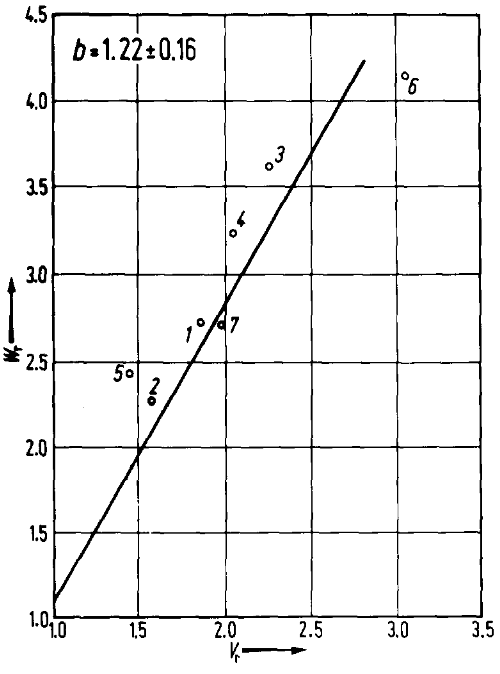

Fig. 5. The regression of W_r on V_r for seed weight

As in the case of pod length, the slope of the regression line for number of seeds per pod was not significantly different from either zero or unity, suggesting that non-allelic gene interaction also existed for this trait. There was no definite trend between the magnitude of this trait and order of dominance. Overdominance apparently was the case since the regression line cut below the origin. (Fig. 8).

For flowering date, partial dominance was indicated since the line cut the W_r axis above the origin (Fig. 9). A high correlation (r = 0.85) was found between parental values and $W_r + V_r$ suggesting that a tendency for dominance to be associated with lateness and recessiveness to be associated with earliness.

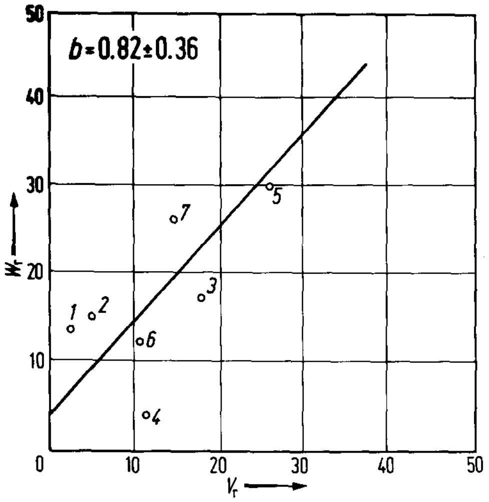

Fig. 6. The regression of W_r on V_r for pod length

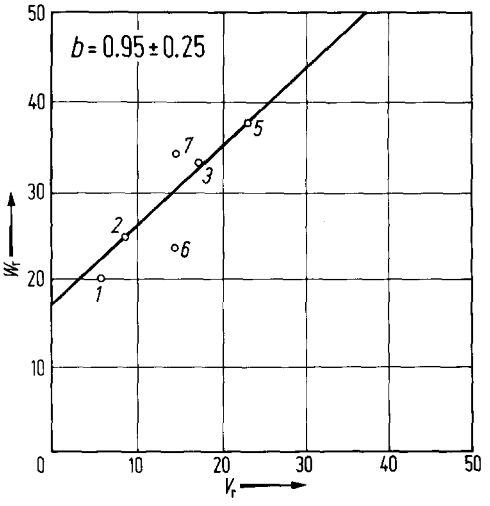

Fig. 7. The regression of W_r on V_r for pod length after removing the progenies of parental array 4

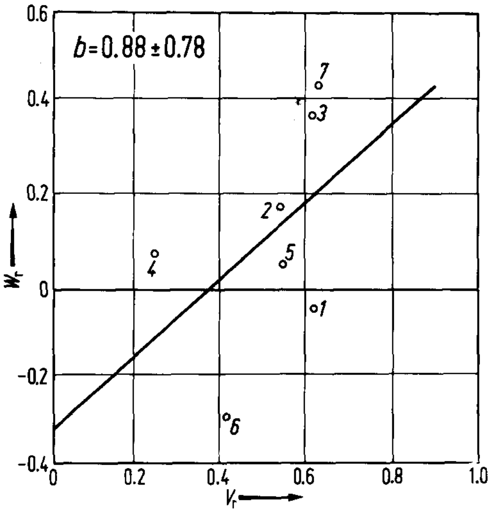

Fig. 8. The regression of W_r on V_r for number of seeds per pod

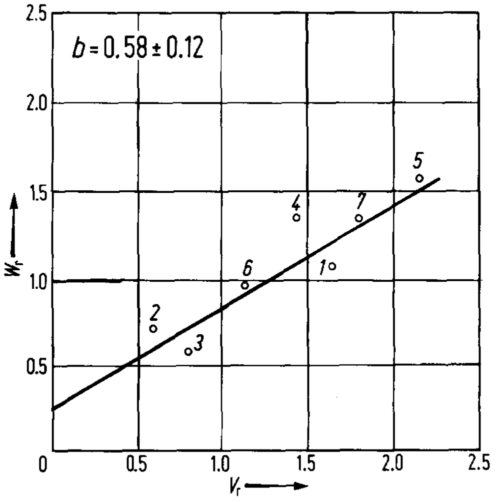

Fig. 9. The regression of W_r on V_r for flowering date

Discussion

There are six assumptions underlying the diallel analysis as proposed by Jinks and Hayman (Jinks 1954; Hayman 1954a, b). Since long bean, *Vigna sesquipedalis* Fruw. is a diploid self-pollinating species, the validity of parental homozygosity and diploid segregation may not be a problem. With the exception of seed weight, the reciprocal effects detected in number of pods per plant, protein yield, pod length and crude protein content were small when compared to the additive or dominance variances. The reciprocal effects might not cause serious bias in the analysis when the mean value for the cross and its reciprocal was used in the diallel analysis (Hayman 1954a). The assumptions of no linkage, no non-allelic gene interaction and no multiple allelism are difficult to evaluate independently. However, according to the diallel cross theory, if the assumptions are valid, the values of $W_r - V_r$ should be constant. Moreover, the regression of W_r on V_r should have a unit slope. Thus, if the values of $W_r - V_r$ of all arrays were not constant or the regression coefficient deviated significantly different from unit slope, it is a good indication of at least partial failure of the assumptions. However, Allard (1956) working on lima bean, Crumpacker and Allard (1962) on wheat and Johnson (1963) on barley found that the invalidity of the assumptions did not create a significant source of bias in the genetic analysis of their diallel cross data.

In the present studies, the results of the test of homogeneity of $W_r - V_r$ by the analysis of variance (Hayman 1954a) and deviation of the slope of unity of the regression of W_r on V_r indicated that crude protein content, number of seeds per pod and pod length showed some failure of the assumptions. Parent array 4 appeared to be the cause of the failure of the assumptions in pod length since removal of the progenies of this array resulted in a regression line significantly different from zero but not significantly different from unity. Removal of progenies of parent 4 improved the regression line for crude protein content and number of seeds per pod. However, partial failure of the assumptions was still apparent since the regression line was significantly different from either zero or one. Thus, for the number of seeds per pod and crude protein content, the genetic information obtained might not be very reliable.

For all the traits studied in this diallel population, both additive and non-additive gene effects were shown to be involved in the genetic variation. However, dominance variance was found to be more important than additive variance in crude protein content, number of pods per plant and number of seeds per pod. For seed weight and pod length, additive variance is more important. High heritability estimates for these two traits confirmed the importance of additive variance for the genetic variation.

The crude protein content, protein yield and number of pods per plant seemed to be controlled by overdominance effects. However, the degree of dominance might be inflated for crude protein content and number of seeds per pod since these two traits showed at least partial failure of the assumptions. Partial dominance seems to be the case for flowering date, pod length and seed weight, and complete to overdominance for pod yield.

In *Phaseolus vulgaris*, Leleji et al. (1972) found low protein content was controlled by partial dominance effects whereas in soybean, high protein was completely dominant over low protein (William 1948). In the present studies, dominance to overdominance seems to be the case for protein content in *Vigna sesquipedalis*, since the value of $(H_1/D)^{\frac{1}{2}}$ for this trait was greater than unity. But the degree of dominance might be overestimated here because this trait showed some failure of the assumptions. The value of D-H_1 for this trait was negative and significant. This might indicate that dominance effect is more important than additive effect for this trait. However, seeds of high protein appeared to be associated with recessive genes since parent 2 carrying the greatest number of recessive genes had also the highest crude protein content (Fig. 3). On the other hand, there was a general trend to show that low yielding parents carried more recessive genes and high yielding parents carried more dominant genes.

Literature

Allard, R.W. (1956): Biometrical approach to plant breeding. In: Genetics in Plant Breeding. Brookhaven Symposia in Biology. 9, 69-88 U.S. Brookhaven National Laboratory, Upton, N.Y.

Crumpacker, D.W.; Allard, R.W. (1962): A diallel cross analysis of heading date in wheat. Hilgardia 32, 275-318

Hayman, B.I. (1954a): The theory and analysis of diallel crosses. Genetics 39, 789-809

Hayman, B.I. (1954b): The analysis of variance of diallel tables. Biometrics 10, 235-244

Jinks, J.L.: (1954): The analysis of continuous variation in a diallel cross of *Nocotiana rustica* varieties. Genetics 39, 223-238

Johnson, L.P.V. (1963): Application of the diallel cross techniques to plant breeding. In: Statistical genetics and plant breeding eds. Hanson, W.D.; Robinson, H.F. Washington, D.C. National Academy of Science, National Research Council 982, 561-570

Leleji, O.I.; Dickson, M.H.; Crowder, L.V.; Bourke, J.B. (1972): Inheritance of crude protein percentage and its correlation with seed yield in beans, *Phaseolus vulgaris* L. Crop Sci. 12, 168-171

Mak, C.; Yap, T.C. (1977a): Variability and affects of N-fertilizer on seed protein contents in long bean. Malaysian Appl. Biol. 6, 39-44

Mak, C.; Yap, T.C. (1977b): Heterosis and combining ability of seed protein, yield and yield components in long bean. Crops. Sci. 17 (2), 339-41

William, L.F. (1948): Inheritance in a species cross in soybean. Genetics 33, 131-132

Yap, T.C.; Mak, C.; Poh, T.C. (1977a): Genetic studies and breeding of long beans in Malaysia. Malaysian Appl. Biol. 6, 155-162

Yap, T.C.; Poh, C.T.; Mak, C. (1977b): Comparative studies on selection methods in long beans. Proc. Third Intern. Congr. SABRAO, Canberra vol. 1, 2c(1), 23-27

Received August 22, 1979
Communicated by D. von Wettstein

Dr. C. Mak
Department of Genetics and Cellular Biology
University of Malaya
Kuala Lumpur (Malaysia)

Prof. Dr. T.C. Yap
Dept. of Agronomy and Horticulture
University of Agriculture
Serdang (Malaysia)