GENETIC STUDIES ON QUANTITATIVE CHARACTERS IN SOYBEAN

I: Genetic analysis of first internode length by partitioning method

Fu-Sheng THSENG and Sadaji HOSOKAWA

(Laboratory of Industrial Crops, Faculty of Agriculture, Hokkaido University, Sapporo, Japan)

Received May 2, 1970

The purpose of this report is to estimate the number of effective factors controlling the length of first internode of soybean.

Partitioning method of genetic analysis developed by L. Powers was applied for this purpose. Methods to estimate the number of effective factors and their actions controlling quantitative characters for instance such as length of internode have been studied by W. E. CASTLE (1921), L. POWERS (1939, '42, '50, '63), K. MATHER (1949a, 1949b), K. MATHER and A. VINES (1952), W. H. LEONARD *et al.* (1957) and others.

Partitioning method is pertinent for genetic analysis on the assumption that genetic and environmental variabilities follow the same scale and the effects of factors are additive.

The scope of this report is limited to estimate the number of effective factors controlling the length of the first internode in cultivated soybean. This character has been regarded of great importance for the improvement of lodging resistance. Perhaps more detailed research will be necessary to explain the mode of inheritance on this character, but this is a preliminary report because this type of studies has been published very few.

Materials and Methods

Two cultivated soybean varieties, Harosoy (indeterminate growth habit) and Tokachi-nagaha (determinate growth habit), and their F_1 and F_2 generations were used in this study. The experimental design followed a modified randomized complete block method, with six replications. Each block contained two plots of Harosoy (P_1), Tokachi-nagaha (P_2), the F_1, and 10 plots of F_2 plants. Each plot contained 12 plants. This experiment was made at

the Tokachi Agriculture Experimental Station, Hokkaido, Japan in 1969.

To minimize the effect of competition, the parents, F_1 and F_2 were space-planted at 60 cm intervals in rows which were spaced 60 cm apart.

The length of the first internode from the cotyledon node to the node of primary leaf was measured in cm.

The number of effective factors controlling the first internode length was estimated by the partitioning method as described below.

Results

1. Normality of frequency distributions

The mean values for the first internode length for the P_1 (Harosoy), P_2 (Tokachi-nagaha) and their F_1 were 3.51 cm, 4.96 cm and 3.58 cm, respectively (Table 1). If the data follows an arithmetic scale, the $\bar{F}_1$ should approximate $(\bar{P}_1+\bar{P}_2)/2$. But this was not the case. On a logarithmic scale, the $\bar{F}_1$ value must be $\sqrt{(\bar{P}_1)(\bar{P}_2)}$. This also was not the case. Clearly, the effects of genes are not additive on either an arithmetic or a logarithmic scale. Since the $\bar{F}_1$ value was closer to the $\bar{P}_1$ than to the $\bar{P}_2$ value, it was assumed that shorter internode length was dominant over longer one. Possibly, there may be intra-allelic as well as inter-allelic interaction of genes.

Since an arithmetic scale was used for calculation, it should be tested first whether the variability followed the normal probability integral (LEONALD *et al.* 1957). For testing normality, the observed and expected frequency distributions of first internode length for P_1, P_2, F_1 and F_2 population are given in Table 1. The chi-square value was 179.5160 (d.f=9), with P-value less than 0.001 for the F_2 population indicates that the observed distribution does not fit a normal distribution. This suggests that the F_2 population is composed of different genotypes, each of which may fluctuates normally around its own mean. The partitioning of the F_2 population into its component genotypes is made on this assumption. The P-values (0.02-0.05) for non-segregating populations indicate that the environmental variations follow the normal distribution.

2. Number of effective factors differentiating between the parents

As shown in Table 2, the F_2 distribution, expressed by percentages, was unimodal. The mode occurred in the 3.6-class with a frequency of 16.5 percent. This mode corresponded to that of P_1. The mode of P_2 fell in the 5.1-class with a frequency of 29.17 percent. When all percentages in the 3.9 and lower classes in the F_2 distribution are summed up, a value of 73.84 percent is obtained. Similarly, when the frequencies for 4.2 and higher classes

TABLE 1. Observed and expected frequency distributions for first internode length in parental, F_1 and F_2 populations

Population	Upper limit of class																Mean	σ^2	σ	No. of plants
	1.5	1.8	2.1	2.4	2.7	3.0	3.3	3.6	3.9	4.2	4.5	4.8	5.1	5.4	5.7	1.0				
Harosoy (P_1)																				
Observed					9	17	25	35	21	13							3.51	0.1885	0.4342	120
Expected					5	18	23	32	28	16	4									
F_1																				
Obs.				2	4	11	18	25	17	11	4						3.58	0.2285	0.4780	92
Exp.				2	3	8	15	22	21	13	6	2								
F_2																				
Obs.		10	14	23	57	74	78	99	54	88	50	26	15	9	3		3.66	0.6134	0.7832	600
Exp.		1	3	8	23	50	82	109	115	95	62	32	13	4	3					
Tokachi-nagaha (P_2)																				
Obs.										13	21	23	35	15	13		4.96	0.2043	0.4520	120
Exp.									1	5	13	25	31	25	15	5				

TABLE 2. Frequency distributions in parental, F_1 and F_2 populations shown in percent

Population	Upper limit of class														N
	1.8	2.1	2.4	2.7	3.0	3.3	3.6	3.9	4.2	4.5	4.8	5.1	5.4	5.7	
Harosoy				7.50	14.17	20.83	29.17	17.50	10.83						120
F_1			2.14	4.35	11.96	19.57	27.17	18.48	11.96	4.35					92
F_2	1.67	2.34	3.83	9.50	12.33	13.00	16.50	9.00	14.67	8.33	4.33	2.50	1.50	0.50	600
Tokachi-nagaha									10.83	17.50	19.17	29.17	12.50	10.83	120

are summed, a value of 26.16 percent is obtained. This indicates that in the F_2 one effective factor conditions the first internode length. The mode for 4.2 and higher classes did not correspond to that of the P_2. This indicates that another factor may be involved. The simplest hypothesis is thus obtained for the effective factors differentiating the parents with respect to the first internode length.

This hypothesis can also be confirmed by the value of F_2/P_2 (POWERS 1955). It can be seen in Table 3 that the first three estimates appeared to

TABLE 3. Calculated percentage values used for estimating the number of effective factors

Estimate	Percentages in F_2/P_2
1	4.62
2	8.57
3	8.57
4	12.32
5	20.37
6	32.83

fluctuate around a value of 6.25 ($=1/16\times 100$), indicating that the parents may be differentiated by two effective factors (LEONARD *et al.* 1957). These three classes would represent the double recessive genotype. According to LEONARD *et al.* (1957), the sudden rise in the fourth estimate 12.32 may be regarded as indicating that the plants with genotypes other than the double recessive one occur in the 4.8-class. This information makes it possible to formulate a genetic hypothesis of two gene pairs as the basis for partitioning the F_2 frequency distribution. Accordingly, the genotypes of parents, the Harosoy (P_1) and Tokachi-nagaha (P_2) were assumed to be *AABB* and *aabb*, respectively.

Further, the data for the F_1 plants, which should be of *AaBb* genotype, indicated that they were similar to *AABB* plants; the first internode length of Harosoy was 3.51 cm, and that of the F_1 was 3.58 cm (Table 1).

When *AaBb* plants fall in the same class as *AABB* plants, the *AABb* and *AaBB* plants also will fall in the same class. Consequently, the percentage of F_2 plants in the same class as of Harosoy (P_1) is expected to be $9/16=56.25$. This corresponds to the expected frequency of the double dominant class in a digenic F_2 population.

TABLE 4. Partitioning of the F_2 distributions on the basis of parental distributions

Population	Upper limit of class														Theoretical percentages in F_2
	1.8	2.1	2.4	2.7	3.0	3.3	3.6	3.9	4.2	4.5	4.8	5.1	5.4	5.7	
F_2 Observed ($AABB \cdots aabb$)	1.67	2.34	3.83	9.50	12.33	13.00	16.50	9.00	14.67	8.33	4.33	2.50	1.50	0.50	100.00
Harosoy ($AABB$, $AaBB$ / $AABb$, $AaBb$)				4.22	7.79	11.72	16.41	9.84	0.68						56.25
Tokachi-nagaha ($aabb$)									0.68	1.09	1.20	1.82	0.78	0.68	6.25
Balance ($aaBB$, $aaBb$ / $AAbb$, $Aabb$)	1.67	2.34	3.83	5.28	4.36	1.28	0.09	−0.84	7.90	7.24	3.13	0.68	0.72	−0.18	37.50

TABLE 5. Frequency distributions of first internode length for $AAbb+Aabb$ and $aaBB+aaBb$ genotype

Genotype	Class center															Total percent
	1.65	1.95	2.25	2.55	2.85	3.15	3.45	3.75	4.05	4.35	4.65	4.95	5.25	5.55	$\sum x$	$\bar{x}$	
$AAbb+Aabb$	1.67	2.34	3.83	5.28	4.36	1.28	0.09								46.17	2.45	18.85
$aaBB+aaBb$								−0.84	7.90	7.24	3.13	0.68	0.72	−0.18	81.04	4.35	18.65

3. Partitioning of F_2 population and the χ^2 test for homogeneity

The F_2 distribution was partitioned on the basis of the parental distributions of Table 4. The expected percentages given in the last column of Table 4, as compared with the observed frequency distributions of P_1 and P_2, were used to obtain the theoretical frequeney distributions of plants having *AABB*, *AaBB*, *AABb* and *AaBb*, and *aabb*, respectively, which are shown in the second and third rows of Table 4. After subtracting these values from the observed percentages of the F_2 population, the distribution for the balance of the plants with *AAbb*, *Aabb*, *aaBb* and *aaBB* was obtained.

This distribution (in the bottom row of Table 4) had two peaks, one at 5.28 (2.7-class) and the other at 7.24 (4.2-class), indicating that the $A:a$ and $B:b$ loci had different magnitudes of effect. When all the frequencies in the 3.6 and lower classes are summed, a value of 18.85 is obtained. The sum of values in the 3.9 and higher classes is 18.65. The sum of these two values is 37.50 ($=18.85+18.65$), and their average is 18.75 (3/16).

To estimate the means and standard errors for genotype groups 2 and 3, the method described by LEONARD *et al.* (1958) was used. Since the parental distributions followed the normal distribution, the distributions of plants with *AAbb*+*Aabb* and *aaBB*+*aaBb* (Table 5) were computed on the same basis from the balance given in Table 4, as follows: The means and the total variance of parents as given in the Table 1 indicated that they were positively correlated. Assuming that their relationship is linear, the total variance of plants *AAbb*+*Aabb*, as well as of *aaBB*+*aaBb* plants, can be estimated from $y=mx+b$ (POWERS 1942). The results of this computation are given in Table 6.

TABLE 6. Mean and variation parameters expected for different genotype

Genotype group	Mean	σ^2	σ	Theoretical No. individuals
AABB+*AABb*+ *AaBB*+*AaBb* (Harosoy)	3.51±0.0245	0.1885	0.4342	337.5
AAbb +*Aabb*	2.45±0.0400	0.1769	0.4199	112.5
aaBB+*aaBb*	4.35±0.0424	0.1976	0.4440	112.5
aabb (Tokachi-nagaha)	4.96±0.0735	0.2043	0.4520	37.5

The expected standard errors of genotype means supported the assumption that the F_2 population was made up of different genotypes, each of

TABLE 7. Expected F_2 distribution as shown by percentages

Genotype	Upper limit of class								
	1.2	1.5	1.8	2.1	2.4	2.7	3.0	3.3	3.6
AABB+*AABb* *AaBB*+*AaBb* (Harosoy)						7.50	14.17	20.83	29.17
aabb (Tokachi-nagaha)									
AAbb+*Aabb*	0.2	1.5	6.2	16.6	26.3	26.2	15.7	5.8	1.3
aaBB+*aaBb*							0.2	1.0	4.5
Theoretical	0.04	0.28	1.16	3.12	4.83	9.14	10.95	13.00	17.48

	3.9	4.2	4.5	4.8	5.1	5.4	5.7	Theoretical percentage in F_2 population
AABB+*AABb* *AaBB*+*AaBb* (Harosoy)	17.50	10.83						56.25
aabb (Tokachi-nagaha)		10.83	17.50	19.17	29.17	12.50	10.83	6.25
AAbb+*Aabb*	0.2							18.75
aaBB+*aaBb*	12.7	22.5	26.5	19.7	9.3	2.9	0.7	18.75
Theoretical	12.26	11.00	6.06	4.89	3.56	1.32	0.81	100.00

them fluctuating around its mean.

Theoretical frequency distributions for the nine F_2 genotypes (Table 7) were computed by the method described by POWERS *et al.* (1950).

The chi-square value between observed and calculated F_2 distributions was 12.4562 (d.f=9), with a P-value between 0.20-0.10, which indicated a good fit and supported the hypothesis that the two parents were differentiated by two effective factor pairs with respect to the first internode length.

4. Magnitude of genic effects and interaction

From the F_2 populations genotypes are sampled at random. On this basis, the theoretical standard error for a genotype can be represented by the standard error of a single determination divided by the expected number of plants having the given genotype (POWERS 1951). The standard errors thus are computed compiled in Table 6.

The magnitude of genic effects as well as their interaction were determined through comparisons between the means and standard errors (LEONARD *et al.* 1957).

(a). Effect of factor $A : a$:

First estimate: $2.45 \pm 0.0400 - 4.96 \pm 0.0735$
$= -2.51 \pm 0.0836$

Second estimate: $3.51 \pm 0.0245 - 4.35 \pm 0.0424$
$= -0.84 \pm 0.0490$

(b). Effect of factor $B : b$:

First estimate: $4.35 \pm 0.0424 - 4.96 \pm 0.0735$
$= -0.61 \pm 0.0849$

Second estimate: $3.51 \pm 0.0245 - 2.45 \pm 0.0400$
$= 1.06 \pm 0.0469$

It seemed from these estimates that the $A : a$ factor was approximately from two to four times as large as the effect of the $B : b$ factor.

For testing their interaction, estimates were obtained as follow:

(a). Effective factor $(A-a)$:

$-2.51 \pm 0.0836 - (-0.84 \pm 0.0490)$
$= -1.67 \pm 0.0969$

(b). Effective factor $(B-b)$:

$-0.61 \pm 0.0849 - 1.06 \pm 0.0496$
$= -1.67 \pm 0.0969$

The data given above shows that the two values for effective interaction are identical. As a test of the interaction, the P-value was very large ($1.78/0.0969 = 18.3694$ from Table 1 of FISHER 1950). It can be concluded that there should be gene interaction.

Discussion

The partitioning method of genetic analysis is based on analysis of frequency distribution by segregating populations into their component genotypes. Tests for the validity of an obtained hypothesis may be accomplished through comparisons between the obtained and theoretical means, their distributions and variances.

In the application of this method, its limitation should be realized (POWERS 1963). As the number of effective factor pairs differentiating the parents increases, the reliability of the POWERS' method decreases. When the number of individuals in a population is small, the validity of chi square tests is reduced, disturbing the conclusion. There should be a fairly large number of individuals for all populations and an extensive genetic design is desirable to arrive in a final conclusion. Such an adequate design contains backcrosses and advanced generations (POWERS 1939, '42, '55, '63). In this experiment,

the available data was from only the parents, F_1 and F_2, so the genetic analysis on the mode of inheritance could not be illustrated adequately.

The two parental lines, Harosoy and Tokachi-nagaha, were found to be differentiated by two effective factor pairs with respect to the first internode length. Their genotypes were *AABB* and *aabb*, respectively. The $A:a$ factor appeared to exert from about two to four times as great an effect on the first internode length as the $B:b$ factor. From Tables 5 and 6, it can be found that the mean of plants having the genotype (*AAbb*, *Aabb*) was smaller than that of the plants having the genotype of dominant alleles (*AABB*, *AABb*, *AaBB* and *AaBb*). Therefore, There may be intra-allelie and inter-allelic interaction between $A:$a and $B:b$, though further studies should be required.

This finding may be of interest in soybean breeding for lodging resistance.

Acknowledgement

The authors wish to express their hearty thanks to Dr. T. KUSUNOKI, Director, Mr. M. SAITO, Head of the Soybean Laboratory, Mr. T. SANBUICHI, and other staff-member of the Tokachi Agriculture Experimental Station, who encouraged them to carry out this work and helped them with valuable suggestions. Likewise, the authors are also grateful to Dr. C. TSUDA, Dr. Y. SHIMAMOTO and Mr. H. NAKAZIMA for their interest in this work and discussions.

Summary

1. The F_1 and F_2 generations obtained from a cross between two soybean varieties, Harosoy and Tokachi-nagaha, were used for genetic analysis of first internode length by the partitioning method (POWERS 1955, LEONARD *et al.* 1957).

2. The parental varieties were found to be differentiated by two effective factor pairs, $A:a$ and $B:b$.

3. Short first internode showed complete phenotypic as well as genic dominance. The $A:a$ locus was found to have a 2.4-3.9 times as strong effect as the $B:b$ locus.

4. There were also found intra-allelic and inter-allelic interactions between $A:a$ and $B:b$.

Literature cited

1. CASTLE, W. E., 1921. On a method of estimating the number of genetic factors concerned in cases of blending inheritance. Science 54: 93-96.
2. FISHER, R. A., 1950. Statistical methods for research workers. 11th edition. Oliver and Boyd, Edinburgh and London.
3. LEONARD, W. H., H. O. MANN and L. POWERS, 1957. Partitioning method of genetic analysis applied to plant height inheritance in barley. Colorado Agr. Exp. Sta. Tech. Bull. 60.
4. MATHER, K., 1949a. Biometrical genetics. Methuen, London.
5. MATHER, K., 1949b. The genetical theory of continuous variation. Proc. 8th Intern. Congr. Genet. (Hereditas suppl. vol.): 376-401.
6. MATHER, K. and A. VINES, 1952. The inheritance of height and flowering time in a cross of *Nicotiana rustica*. Pp. 49-80. Quantitative inheritance. His majesty's Stationory Office, London.
7. PEARSON, E. S. and H. O. HARTLEY, 1958. Biometrika Tables for statisticians. Vol. 1: 104-110. Second edition, Cambridge Univ. Press.
8. POWERS, L., 1939. Studies on the nature of the interactions of the genes differentiating quantitative characters in a cross between *Lycopersicon esculentum* and *L. pimpinellifollium*. J. Genet. 39: 139-190.
9. POWERS, L., 1942. The nature of the series of environmental variance and the estimation of the genetic variances and the geometric means in crosses involving species of *Lycopersicon*. Genetics 27: 561-575.
10. POWERS, L., 1950. Gene analysis of weight per locule in tomato hybrids. Botan. Gaz. 112: 163-174.
11. POWERS, L., 1951. Gene analysis by the partitioning method when interactions gene are involved. Botan. Gaz. 113: 1-23.
12. POWERS, L., 1955. Components of variance method and partitioning method of genetic analysis applied to weight per fruit of tomato hybrids and parental populations. U. S. Dept. Agr. Tech. Bull. 1131.
13. POWERS, L., 1963. The partitioning method of genetic analysis and some aspects of its application to plant breeding. Natl. Acad. Sci. Natl. Res. Council. Publ. 982: 280-318.
14. POWERS, L., F. L. LOCKE and J. C. GARRETT, 1950. Partitioning method of genetic analysis applied to quantitative characters of tomato crosses. U.S. Dept. Agr. Tech. Bull. 998.