Rev. Brasil. Genet. 17, 3, 291-297 (1994)
(Brazil. J. Genetics)

AN IMPROVED PROCEDURE FOR TESTING THEORETICAL SEGREGATION MODELS IN QUALITATIVE GENETIC STUDIES OF SOYBEANS

Carlos A. Arrabal Arias, José F. Ferraz de Toledo and José Tadashi Yorinori

ABSTRACT

A procedure to study simply inherited traits is proposed. It consists of a chi-square test, where the theoretical expected frequencies for the F2 generation are calculated based on the observed distribution of parents and F1, together with the theoretical hypothesis for segregation ratios, taking into consideration possible overlapping between phenotypic classes. A single splitting point is used to limit the phenotypic classes, regardless of the cross being studied. The procedure is shown to complement well the mean and variance analyses inherent of a quantitative genetics approach.

Three crosses of a study on the inheritance of resistance of soybeans (*Glycine max* L. Merrill) to race Cs-15 of *Cercospora sojina* (frog-eye leaf spot) were chosen to illustrate our procedure. Qualitative and quantitative analyses were performed on mean and variance data. Estimates of mean components and scaling tests helped to interpret qualitative analyses made by the modified chi-square test.

Comparisons with conventional analysis were made. The improved method is useful for research involving simple genetic inheritance of environmentally susceptible characters. In spite of environmental interference and the absence of dominance in some crosses, which makes interpretation of segregation patterns more difficult, one main locus for resistance was observed among the cultivars. The results suggest that the study of inheritance of resistance to *C. sojina* must be done near or at the V5 growth stages of soybean.

INTRODUCTION

Resistance to *Cercospora sojina* (frog-eye leaf spot) in soybeans is reported to have a simple genetic control but is highly affected by environmental variations (Yorinori, 1980). The environment has a comparatively greater influence at the earlier than at the later stages of plant development (Arias, 1991). In these cases, the limit between resistant and susceptible classes is not very clearly defined and the application of the goodness-of-fit chi-square tests on the theoretical segregations (such as 3:1, 1:2:1 or 1:3) is usually performed, neglecting phenotypic overlapping between the parentals and F1 generations. As a consequence, the classical chi-square test on the segregation of the F2 generation is not completely satisfactory. Another inconvenience is the need for different splitting points to separate the phenotypic classes for the distinct crosses.

A method of obtaining the expected frequencies in a goodness-of-fit chi-square test that takes into consideration the environmental effects in the parental and F1 distributions using a single splitting point to limit the phenotypic classes is presented.

THEORY

The observed phenotypic frequencies of the parents and their derived F1 are used to calculate the expected numbers in the F2 generation distribution

Suppose a 1:2:1 ratio, resulting from the segregation between one pair of alleles showing incomplete dominance towards resistance, is to be tested. In the classical chi-square test, among 80 F2 plants, 20 are expected to have the same genotype as P1 (resistant), 40 the same genotype as the F1 hybrid (intermediate response) and the remaining 20 the genotype of P2 (susceptible). According to the proposed methodology, the theoretical

Empresa Brasileira de Pesquisa Agropecuária, EMBRAPA, CNPSo, Rodovia Carlos João Strass, Acesso Orlando Amaral, Distrito de Warta, Caixa Postal 1061, 86001-970 Londrina, PR, Brasil. Send correspondence to C.A.A.A.

ratio and the available parental and F1 distributions provide the necessary information to obtain the expected numbers for the chi-square. The method takes into account the environmental effects on the frequency distribution of the parents and F1 generations to obtain the expected numbers of the F2 distribution. If a resistant or susceptible parent presents in its frequency distribution some plants in the susceptible or resistant classes, respectely, the method accounts for them in the expected number of the F2 generation. The same happens with the F1 plants, i.e. their contribution to the number of expected plants in each class of the F2 generation is calculated proportionally to the observed numbers in the F1. In summary, the method accounts for the fact that a genotype in homozygous or heterozygous condition may not show all plants scoring in a determined class.

In the case:

Generation	Class					N
	C1	C2	...	...	Cj	
P1	p11	p12	...	...	p1j	Np1
P2	p21	p22	...	...	p2j	Np2
F1	f11	f12	...	...	f1j	Nf1
F2	f21	f22	...	...	f2j	Nf2

where:

- p1j is the observed relative frequency of the j^{th} class in the P1 generation.
- p2j is the observed relative frequency of the j^{th} class in the P2 generation.
- f1j and f2j are the same for the F1 and F2 generations, respectively.

With only one locus involved in the expression of the character, the expected frequency in each class of the F2 distribution, namely f2j', can be predicted as follows:

f2j' = expected relative frequency to j class of F2.

$$f2j' = 1/4\ p1j + 1/4\ p2j + 1/2\ f1j \quad (1)$$

And:

F2 generation	Class				
	C1	C2	...	...	Cj
Observed	f21	f22	...	...	f2j
Expcted	f21'	f22'	...	...	f2j'

The next step is to multiply the relative frequencies by the number of observations of F2 (Nf2):

F2 generation	Class					N
	C1	C2	...	...	Cj	
Observed	N1	N2	...	...	Nj	Nf2
Expected	N1'	N2'	...	...	Nj'	Nf2

With Nj and Nj' values a chi-square test can be applied:

$$\chi^2 = \sum_j \frac{(Nj - Nj')^2}{Nj'} \quad (2)$$

A theoretical example to calculate the expected numbers of resistant and susceptible plants in the F2 generation based on the parental and F1 observed distribution is given to clarify the procedures.

Consider a case where classes from 0 to 5 are present (Table I).

Table I - Observed distribution of classes for the parental, F1 and F2 generations.

Generation	Classes						N
	0	1	2	3	4	5	
Parent 1	5	5	5	5	0	0	20
Parent 2	0	0	5	5	5	5	20
F1	2	5	5	5	5	3	25
F2	4	16	20	20	16	4	80

The relative frequency for each class is shown in Table II.

Table II - Relative frequencies for classes.

Generation	Classes						N
	0	1	2	3	4	5	
Parent 1	0.25	0.25	0.25	0.25	0	0	20
Parent 2	0	0	0.25	0.25	0.25	0.25	20
F1	0.08	0.20	0.20	0.20	0.20	0.12	25
F2	0.05	0.20	0.25	0.025	0.20	0.05	80

According to a 1:2:1 ratio of a single gene segregation no dominance model, it is expected that 1/4 of plants will have the P1 distribution, 1/4 the P2 distribution and 1/2 and F1 distribution, and from (1) the observed and expected relative frequencies for F2, (Table III) can be obtained.

Table III - Observed and expected relative frequencies of individuals in each class for the F2 generation, according to a 1:2:1 genetic ratio.

F2 generation	Class						N
	0	1	2	3	4	5	
Observed	0.05	0.20	0.25	0.25	0.20	0.05	80
Expected	.1025	.1625	.225	.225	.1625	.1225	80

The frequency value of each class is multiplied by N before applying the chi-square test. If a decision is taken that grades 2 and 3 denote the splitting point between resistant and suceptible phenotypes, the situation of Table IV is configured.

Table IV - Observed and expected numbers in group 1 (0, 1 and 2 classes) and group 2 (3, 4 and 5 classes).

F2 generation	Class							
	Group 1				Group 2			
	0	1	2	Total	3	4	5	Total
Observed	4	16	20	40	20	16	4	40
Expected	8.2	13	18	39.2	18	13	9.8	40.8

Goodness-of-fit chi-square tests are performed between the expected numbers obtained from the parentals and F1 distributions and the observed numbers obtained from the F2 distribution using the expression shown in (2). If the single gene model is rejected, more elaborate models can be tested. The results of quantitative analysis are helpful in the process of deciding which of these models to use. The quantitative approach, used to support the prior analysis, resorted to the genetic modeling of means and variances of the available generations, using the methods described in Mather and Jinks (1982).

In comparison with the goodness-to-fit chi-square test this procedure is slightly more laborious. Its advantages, however, are three fold: a) it makes full use of the information provided by the parental and F1 data; b) it gives the opportunity to define a single splitting point for all crosses before evaluation; c) it correlates well with quantitative tests.

MATERIAL AND METHODS

The materials include the cultivars Bragg, Davis and Parana and their derived F1 and F2 generations. Davis and Parana are resistant and Bragg is susceptible to race Cs-15 of *Cercospora sojina*. The experimental data used in this work were obtained in a project developed to study the inheritance of soybean resistance to Brazilian races Cs-4 and Cs-15 of *Cercospora sp.* (Arias, 1991). Race Cs-4 is the prevalent race in Brazil (Yorinori, 1989). Race Cs-15 is relatively recent and broke the resistance of cultivar Santa Rosa, which had been widely used as source of resistance against frog-eye leaf spot in Brazil.

The experiment was carried out in a green house in a complete randomized design. Three hundred and seventy-five pots were individually sown with five seeds of the pertinent generation and later thinned to a single plant. Twenty plants of each parental plus 25 F1 plants and 80 F2 plants of each cross formed the experimental material. Inoculation was carried out uniformly for all plants at stages V3 and V5 (Fehr *et al.*, 1971), with innoculum prepared from mycelium and spores from ten plates totally colonized suspended in one liter of water. Five drops of Tween blender 80% was added to the solution.

The evaluation was done on the most severely infected leaf, 12 to 15 days after inoculation. The infection level was graded based on the proportion of the leaf area showing typical disease symptoms. Infection levels varied from 0 to 5, where: 0 = leaf without lesions; 1 = 1 to 10% of leaf area infected (l.a.i.); 2 = 11 to 25% l.a.i.; 3 = 26 to 50% l.a.i.; 4 = 51 to 75% l.a.i.; and 5 = more than 75% l.a.i. (Yorinori, 1982). The splitting point was fixed between grades 2 and 3 to separate resistant (0, 1 and 2) and susceptible (3, 4 and 5) groups.

The individual scores of the parentals and their derived F1 and F2 populations were obtained for calculation of their respective means and variances and other distribution data.

RESULTS AND DISCUSSION

Means, variances and frequency distributions of the leaf infection level caused by inoculation with race Cs-15 of *Cercospora sp.* at stages V3 and V5 of soybean development are presented in Table V. A general tendency of higher susceptibility of plants when inoculated at the V3 stage was observed in the F1 and F2 generations. The cultivars presented a well defined reaction for both inoculation dates. Their distribution, however, showed some plants in classes distinct from that of the predominant reaction (Table V). The F1 from Bragg x Parana showed a wide distribution around the intermediate grades for inoculation at both stages. The F2 distributions of the crosses Davis x Bragg and Parana x Bragg showed wide segregation and several plants scored as resistant and susceptible for inoculation at both stages.

Table V - Means, variances and frequency distribution of the parental, F1 and F2 generations for leaf infection level, after inoculation with race Cs-15 of *Cercospora sojina* at soybean stages V3 and V5.

Generation	Leaf infection level								
	N	Mean	Var.	0	1	2	3	4	5
				Stage V3					
Bragg	20	3.65	1.08		1	1	6	8	4
Parana	20	1.85	0.45		6	11	3		
Davis	20	0.00	0.00	20					
Davis x Parana									
F1	25	0.96	0.46	6	14	5			
F2	80	0.97	0.94	30	30	12	8		
Davis x Bragg									
F1	25	1.24	0.94	7	7	9	2		
F2	80	1.11	1.47	32	24	12	7	5	
Parana x Bragg									
F1	25	2.60	1.17		5	6	8	6	
F2	80	3.00	1.04	1	5	17	30	24	3
				Stage V5					
Bragg	20	3.50	0.47				12	6	2
Parana	20	1.50	0.58	1	10	7	2		
Davis	20	0.00	0.00	20					
Davis x Parana									
F1	25	0.52	0.26	12	13				
F2	80	0.63	0.64	44	24	10	2		
Davis x Bragg									
F1	25	0.76	0.77	12	8	4	1		
F2	80	1.06	1.96	41	19	2	10	8	
Parana x Bragg									
F1	25	2.04	1.12	2	6	7	9	1	
F2	80	2.35	1.24	3	17	23	23	14	

The test of the model for single gene segregation was performed on the expected and observed numbers of resistant and susceptible plants in the F2 distribution of the three crosses inoculated at both growth stages (Table VI and VII). The conventional analysis is presented in Table VIII.

Table VI - Observed and expected numbers for infection level on leaves in the F2 generation, obtained after inoculation with race Cs-15 of *Cercospora sojina* at soybean stages V3 and V5.

Grade	Davis x Parana		Davis x Bragg		Parana x Bragg	
	Exp.	Obs.	Exp.	Obs.	Exp.	Obs.
			Stage V3			
0	29.6	30	31.2	32	0.0	1
1	28.4	30	12.2	24	15.0	5
2	19.0	12	15.4	12	21.6	17
3	3.0	8	9.2	7	21.8	30
4	0.0	0	8.0	5	17.6	24
5	0.0	0	4.0	0	4.0	3
			Stage V5			
0	40.2	44	39.2	41	4.2	3
1	30.8	24	12.8	19	19.6	17
2	7.0	10	6.4	2	18.2	23
3	2.0	2	13.6	10	28.4	23
4	0.0	0	6.0	8	7.6	14
5	0.0	0	2.0	0	2.0	0

Table VII - Proposed method for goodness-of-fit chi-square test of the 1:2:1 segregation ratio in the F2 generation for the infection level on soybean leaves. Inoculation with race Cs-15 of *Cercospora sojina* at soybean stages V3 and V5.

Cross	Resistant		Susceptible				
	Obs.	Exp.	Obs.	Exp.	χ^2	d.f.	P
			Stage V3				
Davis x Parana	72	77.0	8	3.0	a	-	-
Davis x Bragg	68	58.8	12	21.2	5.43	1	0.02
Parana x Bragg	23	36.6	57	43.4	9.31	1	0.00
			Stage V5				
Davis x Parana	78	78.0	2	2.0	a	-	-
Davis x Bragg	62	58.2	18	21.6	0.82	1	0.36
Parana x Bragg	43	42.0	37	38.0	0.05	1	0.82

a The test was not applied because too few plants were expected to score in the susceptible class.

The cross between cultivars Davis and Parana is a resistant x resistant type cross. The expected theoretical proportion when the resistant alleles were in the same locus is 80 resistant plants to zero susceptible plants. The proportions estimated by the proposed method were very close to the theoretical expectations, indicating that the

Table VIII - Conventional goodness-of-fit chi-square test of the 1:2:1 segregation ratio in the F2 generation for the infection level on soybean leaves. Inoculation with race Cs-15 of *Cercospora sojina* at soybean stages V3 and V5.

Cross	Resistant Obs.	Resistant Exp.	Susceptible Obs.	Susceptible Exp.	χ^2	d.f.	P
			Stage V3				
Davis x Parana	72	80	8	0	a	-	-
Davis x Bragg	68	60	12	20	4.27	1	0.04
Parana x Bragg	23	20	57	60	0.60	1	0.44
			Stage V5				
Davis x Parana	78	80	2	0	a	-	-
Davis x Bragg	62	60	18	20	0.27	1	0.60
Parana x Bragg[b]	43	60	37	20	19.27	1	0.00
Parana x Bragg[c]	66	60	14	20	2.40	1	0.12

a The test was not applied because too few plants were expected to score in the susceptible class.

b and c use splitting points between 2-3 and 3-4, respectively.

single gene hypothesis is probably true for this cross (Table VII). Also, the chosen splitting point between the susceptible and resistant classes (grades 2 and 3), seemed adequate. No goodness-of-fit test was carried out because too few plants were expected to fall in the susceptible class for inoculations at either stage. The mean analysis (Table IX) indicated the presence of significant additive genetic effects but no dominance or non-allelic interaction effects for inoculation at both stages. The variance analysis (Table X) detected significant additive and genotype x microenvironment interaction effects. The significant g x micro-environment interaction indicates a differential response of the parental cultivars to the environment. The quantitative analysis did not detect the presence of more than one locus controlling the character.

The cross between cultivars Davis and Bragg is a resistant x susceptible cross. The frequency distributions of the parents and F1 were well defined and distinct, especially at stage V5, probably due to the stability of the resistance genes of Davis. The conventional and the proposed method gave similar results as expected in absence of large environmental effects (Tables VII and VIII). For inoculation at stage V3 the single gene model was rejected and the genetic model adjusted to the means of the generations, indicated the presence of additive and additive x additive genetic effects (Table IX). There is a clear indication that more than one factor is involved in the resistance reaction. The hypothesis of a two loci model controlling the resistance reaction was tested. Considering a duplicate type of gene interaction, a 15:1 ratio was expected. This hypothesis was not rejected and seems adequate to explain the results obtained with inoculation at stage V3. The goodness-of-fit chi-square value made by the proposed method was 1.22 ($0.50 > P > 0.25$). Progeny testing would be required to further validate this hypothesis. For inoculation at stage V5 the single gene model was not rejected by the chi-square test. The analysis of the means indicated the presence of significant additive and dominance effects but no non-allelic interaction was observed (Table IX). Therefore, the one locus hypothesis seems adequate to explain the results.

The cross between cultivars Parana and Bragg is a resistant x susceptible cross. The F1 distribution

Table IX - Mean components for infection level on leaves obtained from inoculation with race Cs-15 of *Cercospora sojina* at soybean stages V3 and V5.

Mean components	Davis x Parana Stage V3	Davis x Parana Stage V5	Davis x Bragg Stage V3	Davis x Bragg Stage V5	Parana x Bragg Stage V3	Parana x Bragg Stage V5
m	0.95 ± 0.06	0.65 ± 0.05	1.15 ± 0.11	1.74 ± 0.08	2.85 ± 0.08	2.38 ± 0.08
[d]	0.94 ± 0.06	0.64 ± 0.06	1.82 ± 0.12	1.73 ± 0.08	0.94 ± 0.13	1.01 ± 0.11
[h]	-	-	-	-1.07 ± 0.18	-	-
[i]	-	-	0.68 ± 0.16	-	-	-
[l]	-	-	-	-	-	-
χ^2	0.10	3.17	0.30	1.16	3.58	3.73
d.f.	2	2	1	1	2	2
P	0.9492	0.2051	0.5827	0.2820	0.1667	0.1545

Table X - Variance components for infection level on leaves inoculated with race Cs-15 of *Cercospora sojina* at soybean stages V3 and V5.

Variance components	Cross					
	Davis x Parana		Davis x Bragg		Parana x Bragg	
	Stage V3	Stage V5	Stage V3	Stage V5	Stage V3	Stage V5
D	1.19 ± 0.33	0.73 ± 0.24	1.43 ± 0.57	2.96 ± 0.66	-	1.13 ± 0.46
H	-	-	-	-	-	-
E1	0.46 ± 0.09	0.54 ± 0.12	1.00 ± 0.22	0.64 ± 0.14	1.19 ± 0.17	0.53 ± 0.12
E2	0.01 ± 0.003	0.01 ± 0.003	0.01 ± 0.003	0.01 ± 0.003	0.46 ± 0.15	1.12 ± 0.32
χ^2	0.003	0.09	0.10	1.17	0.13	0.21
d.f.	1	1	1	1	2	1
P	0.9597	0.7679	0.7476	0.2785	0.9373	0.6479

overlapped with that of the two parents (dominance was absent, Table IX). The new procedure detected significant deviation of the single gene segregation pattern for inoculation at stage V3 but not for inoculation at stage V5 (Table VII). The mean component analysis indicated the presence of additive effects at both stages. For stage V3 a model with two genes determining the resistance in cultivar Parana could be tested but the absence of non-allelic interaction, plus the non significance of additive effect and the adequacy of a model with genotype by micro-environment (E1 and E2 in variance components, Table X) make this hypothesis less likely. Since no significant deviation from the single gene segregation pattern was observed for inoculation at stage V5, it seems that a single locus determines resistance at advanced stages of the plant development. Conventional tests detected significant deviation from simple inheritance at stage V5, using a splitting point between grades 2 and 3. The change of the splitting point to between grades 3 and 4 would cause the hypothesis of simple inheritance to be accepted (Table VIII). However, this adjustment on the point that determines the classification groups (resistant or susceptible) would be specific for each cross. Also the model with complete dominance to susceptibility (1:3) was accepted at stage V3 (Table VIII), but the dominance [h] was absent in the mean components analysis (Table IX).

The results demonstrated that resistance to race Cs-15 of *C. sojina* in soybeans is simply inherited and controlled by one or two loci. The variable degree of resistance or susceptibility of the parents and their distinct interactions with the micro-environment suggests that other minor genes are involved.

The proposed method seemed capable of correctly interpreting the results in all cases presented. It seems to be equal to the conventional method in the absence of environmental interference. It is most useful to validate simple segregation models, eliminating the need of changes of the limiting point between classes, thus decreasing the possibility of errors of interpretation due to subjective interference and environmental influence on character performance. Since parental and F1 generations are required to calculate the expected frequencies by this method, the number of plants in the parental and F1 generations must be large enough to represent their expected frequency distribution.

ACKNOWLEDGMENTS

The authors are grateful to the Coordenadoria de Apoio a Pesquisa e Ensino Superior (CAPES) and Centro Nacional de Pesquisa de Soja-EMBRAPA (CNPSo-EMBRAPA), for financial support necessary to carry out this work.

RESUMO

Um novo método aplicável em estudos de herança genética simples foi proposto. Sua diferença em relação ao método convencional está no cálculo das freqüências esperadas para a geração F2, que é feita com base nas distribuições observadas dos parentais e geração F1 mais a hipótese de segregação teórica. Ele considera possíveis sobreposições entre classes fenotípicas e evita a mudança do ponto limite entre classes para os vários cruzamentos.

Para ilustrar a metodologia apresentada foram obtidas informações de três cruzamentos de soja programados para o estudo da herança da resistência da soja à raça Cs-15 de *Cercospora sojina*. A distribuição de cada cruzamento foi analisada qualitativamente através do método proposto utilizando um teste de qui-quadrado modificado. As médias e variâncias foram analisadas quantitativamente. As estimativas dos componentes de médias e variâncias auxiliaram na interpretação da análise qualitativa. Apesar da interferência significativa do ambiente sobre a expressão do caráter e da ausência de dominância em alguns cruzamentos que tendem a dificultar a determinação dos

modelos de segregação teóricos, observou-se a presença de um ou dois pares de alelos principais paa a resistência. De modo geral os resultados sugerem que estudos de herança genética para este caráter sejam feitos à partir do estádio V5 de desenvolvimento da soja. Em comparações com o método convencional o método proposto demonstrou-se útil principalmente em pesquisas de herança genética em que o caráter estudado tem sua expressão muito influenciada por variações ambientais.

REFERENCES

Arias, C.A.A. (1991). Herança da resistência da soja (*Glycine max*) às raças 4 e 15 de *Cercospora sojina*. Master's Thesis, Universidade Estaduald e Londrina, Londrina, PR, Brasil.

Fehr, W.R., Caviness, C.E., Burmood, D.T. and **Pennington, J.S.** (1971). Stage of development descriptions for soybeans (*Glycine max*). *Crop. Sci. 11*: 929-931.

Mather, K. and **Jinks, J.L.** (1982). *Biometrical Genetics: the Study of Continuous Variation*. 2nd edn. Cornell University Press, Ithaca, pp. 382.

Yorinori, J.T. (1980). *Cercospora sojina*: pathogenicity, new races and seed transmission in soybean. Doctoral Thesis, University of Illinois, Urbana-Champaign, U.S.A.

Yorinori, J.T. (1982). Identificação de raças de *Cercospora sojina* Hara a nível nacional. In: *Resultados de Pesquisa de Soja 1981/82*. Empresa Brasileira de Pesquisa Agropecuária; Centro Nacional de Pesquisa de Soja, Londrina, pp. 234-235.

Yorinori, J.T. (1989). Frogeye leaf spot of soybean (*Cercospora sojina* Hara). In: *World Soybean Research Conference - IV*. Buenos Aires, Argentina, March 5-9. Vol. III.

(Received September 21, 1993)