

GENETICAL ANALYSIS OF A SOYBEAN BIPARENTAL CROSS AND COMPARATIVE MODEL FITTING TO MEANS AND VARIANCES OF TWO SETS OF GENERATIONS

José Francisco Ferraz de Toledo, Maria Clarete Rossini,
Rogério Fernandes de Souza and Fernando Ferreira Leão

ABSTRACT

Genetical analysis of self-pollinating species usually involves the six basic generations (P_1 , P_2 , F_1 , F_2 , BC_1 and BC_2). The F_1 , BC_1 and BC_2 are obtained by hand pollination and for some species like soybeans, rice and wheat this process can be time and labour consuming. Greater advances in the knowledge of the genetical architecture of important characteristics would be possible if reliable analysis and predictions could be obtained from a more restricted set of generations, P_1 , P_2 , F_2 and F_3 , that are readily obtainable from most of the strictly self-pollinating plant species. In spite of a loss of sensitivity, accurate results are obtainable from the restricted set, especially regarding detection and estimation of additive genetic, dominance and additive environmental effects. It allows a reliable test for $G \times E$ interaction and linkage. However, it is inherently weak in detecting epistatic effects. The predictions made from the restricted set are at least as good as those from the six basic generations.

INTRODUCTION

Genetical analysis of self-pollinating species usually involves the six basic generations P_1 , P_2 , F_1 , F_2 , BC_1 and BC_2 (Jinks and Jones, 1958; Hayman, 1960; Mather and Jinks, 1982). For certain species such as soybeans, however, the number of hand pollinations necessary to obtain the required amount of F_1 and backcross seeds is very high. For others, such as certain cereals, artificial crossing may involve cutting of part of the flower structure thus generating F_1 and backcross seeds that result in seedlings weaker

than those obtained through natural pollination. This usually introduces a severe bias in the analysis that can be avoided only by cutting and hand pollinating all generations involved in the study. The process of obtaining the necessary generations for the experiment can be very demanding both in time and labour, because of the limitations imposed to artificial crossing by the flower reproductive structure and/or by the limited number of seeds set in each pollination.

Therefore, if satisfactory genetical analysis and predictions could be performed based on the restricted set of generations P_1 , P_2 , F_2 and F_3 , which are usually easily obtained in strictly self-pollinating species, advances in the knowledge of the genetic control of the characters could be more readily achieved.

We first conducted an extensive analysis of several genetically different characters of a cross between two soybean lines using all available data (P_1 , P_2 , F_1 , F_2 , BC_1 , BC_2 and F_3). We then used the restricted set and the six basic generations to illustrate our procedures. A comparative analysis of the data, using these two sets of generations was performed. We used several alternative genetic models that can be fitted to the means and variances of the parentals plus F_2 and F_3 generations and estimated the appropriate genetic parameters. Finally, we checked the reliability of these estimators in making predictions.

THEORY

Most typical genetical analyses of self-pollinating species involve scaling tests (Mather, 1949), comparisons of the within family variance of the non-segregating generations (Mather, 1949) and genetic model fitting to the means and variances of the available generations (Hayman, 1960; Perkins and Jinks, 1970; Jinks, 1978; Pooni and Jinks, 1982). Mather and Jinks (1982) showed several alternative models fitted to the six basic generations, using the F_∞ metric of Van der Veen (1959) and involving parameters describing additive genetic and dominance effects, epistasis, linkage, additive and interacting environmental effects.

Under the restricted set of generations a maximum of four mean parameters or five variance parameters can be estimated. The additive and dominance genetical model involving m , $[d]$ and $[h]$ can be adjusted to the means, and one degree of freedom is left for testing the model goodness of fit (Cavalli, 1952). If the model is not adequate one can still try to fit $[i]$ or $[I]$ and get a perfect fit solution for the four generations (Table I). A test for genotype \times microenvironmental effects is provided by the ratio of the parental within family variances. In the absence of $G \times E$ and non-allelic interaction a unambiguous test for linkage is provided by the variance model fitting. For the variances, the restricted set allows the fitting of several alternative models, starting with the simplest, Model I: $Y_i = f(D, H, E)$ where D , H and E stand for additive genetic, dominance and

additive environmental; Model II: $Y_i = f(D, H, E_1, E_2)$, which takes into consideration the genotype x microenvironment interaction; Model III: $Y_i = f(D_1, D_2, E)$, which takes linkage into consideration; Model IV: $Y_i = f(D, I, E)$, which concerns non-allelic interaction between homozygous loci (Table II). In all cases at least one degree of freedom is left for testing the model goodness of fit.

Table I - Genetic parameter components of the means of the generations.

Generation	m	[d]	[h]	[i]	[l]
P ₁	1	1	0	1	0
P ₂	1	-1	0	1	0
F ₂	1	0	1/2	0	1/4
F ₃	1	0	1/4	0	1/16

Perfect fit solutions:

a) Considering m, [d], [h] and [i] in the model:

$$m = 2\bar{F}_3 - \bar{F}_2$$

$$[d] = 1/2 (\bar{P}_1 - \bar{P}_2)$$

$$[h] = 4\bar{F}_2 - 4\bar{F}_3$$

$$[i] = 1/2 (\bar{P}_1 + \bar{P}_2) - 2\bar{F}_3 + \bar{F}_2$$

b) Considering m, [d], [h] and [l] present in the model:

$$m = 1/2 (\bar{P}_1 + \bar{P}_2)$$

$$[d] = 1/2 (\bar{P}_1 - \bar{P}_2)$$

$$[h] = 8\bar{F}_3 - 2\bar{F}_2 - 3\bar{P}_1 - 3\bar{P}_2$$

$$[l] = 8\bar{F}_2 - 16\bar{F}_3 + 4\bar{P}_1 + 4\bar{P}_2$$

Whenever a model more complex than the additive genetic, dominance (linked or unlinked) plus additive environmental effects has to be considered, i.e., involving G x E or non-allelic interaction, the estimates of D, H, D₁, D₂, I, E₁, E₂, etc., will be biased estimates of the true parameters. They would also be biased estimates even if obtained from the six basic generations (Mather and Jinks, 1982). The designs that minimize the biases of D and H are the North Carolina III (Comstock and Robinson, 1952) and the triple test cross (Kearsey and Jinks, 1968). Their use with soybeans, however, is restricted because of the number of hand pollinations required.

Table II - Genetic parameters components of the variances of the generations.

Generation	Model I			Model II					Model III				Model IV				
	D	H	E	D	H	E1	E2	W	D1	D2	E	W	D	I	E	W	
P ₁	0	0	1	0	0	1	0	0	0	0	1	0	0	0	0	1	0
P ₂	0	0	1	0	0	0	1	0	0	0	1	0	0	0	0	1	0
F ₂	1/2	1/4	1	0	1/2	1/4	1/2	1/2	0	1/2	0	1	0	1/2	1/4	1	0
F _{3b}	1/2	1/16	0	1/n	1/2	1/16	0	1/n	1/2	0	0	1/n	1/4	1/4	0	1/n	
F _{3w}	1/4	1/8	1	0	1/4	1/8	1/2	1/2	0	1/4	1	0	1/4	5/16	1	0	

D: additive genetic; H: dominance; E: additive environmental; E1, E2: interactive environmental; D1, D2: additive with linkage; I: additive x additive epistasis; W: sampling (phenotypic) variance within F₃ families; n: harmonic mean of the number of sibs in the F₃ families.

MATERIAL AND METHODS

The experiment was conducted in a glass-house, with three sowing dates: 26/Sept/88, 17/Oct/88 and 16/Nov/88 in Londrina, Paraná state, Brazil, located at latitude 22°30'.

The material involved the parental lines BR 85-29009 (União (6) x Industrial), OC-8 (mutant of Paraná) and the F₁, F₂, F₃, BC₁ and BC₂ generations, and included reciprocals. The varieties FT-2 and BR-13 were included as references (Table III).

Table III - References, parentals and segregating generations with respective family number and size.

Generation	Identification	Number of	
		Families	Individuals
01	BR-13	1	20
02	FT-2	1	20
03	BR 85-29009	1	20
04	OC-8	1	20
05	F ₁ (BR 85-29009 x OC-8)	1	15
06	RF ₁ (OC-8 x BR 85-29009)	1	15
07	F ₂ (F ₁ x F ₁)	1	30
08	RF ₂ (RF ₁ x RF ₁)	1	30
09	BC ₁ (P ₁ x F ₁)	1	10
10	BC ₁ (P ₁ x RF ₁)	1	10
11	BC ₁ (F ₁ x P ₁)	1	10
12	BC ₁ (RF ₁ x P ₁)	1	10
13	BC ₂ (P ₂ x F ₁)	1	10
14	BC ₂ (P ₂ x RF ₁)	1	10
15	BC ₂ (F ₁ x P ₂)	1	10
16	BC ₂ (RF ₁ x P ₂)	1	10
17	F ₃ (F ₂ x F ₂)	15	02
18	RF ₃ (RF ₂ x RF ₂)	15	02

The parental line BR 85-29009 was developed to introduce frog-eye leaf spot resistance into União, a stable and productive variety. The parental line OC-8 is a late flowering mutant of Paraná (Hill x D 52-810), showing a long juvenile period under short day conditions. Both belong to the same maturity group, taking about 125-130 days from

emergence to 95% ripened pods, at 22° latitude and are adapted for planting at latitudes ranging from 22° to 35°.

The means and variances of days from emergence to the first fully developed trifoliolate leaf and days to flowering, number of trifoliolate nodes, final plant height and grain weight were obtained for each generation and reciprocals. Averaging over reciprocals was performed.

Genetic and environmental model fitting to the means and variances of the generations used the models given under Theory and followed the methodology described by Cavalli, (1952) and Hayman (1960). A computer program specifically developed for this purpose was used (Toledo, 1991). The parameters estimated from the restricted set and from the six basic generations were compared and used to predict the potencial of the cross to generate superior inbred lines (Jinks and Pooni, 1976; Toledo, 1987). The accuracy of the predictions (16/Nov/88 sowing) was checked by comparing them with the observed proportions of inbred lines (F₄ lines) falling in each specified category in data collected from an experiment showed in the field at Londrina on November 17, 1989.

RESULTS AND DISCUSSION

The means and variances of all parental and segregating generations, after averaging over reciprocals, are presented in Tables IV to VI. The parental materials were, as expected, best adapted for the 17th of October and 16th of November sowing dates, for which they presented satisfactory plant height. The first sowing date was near normal for OC-8 (plant height 64.55 cm and 10.55 leaf nodes) but was too early for the BR 85-29009 that showed a non-satisfactory growth (plant height 47.98 cm and 9.21 leaf nodes). Yield was higher at the latter two sowing dates for both parental materials.

The mean and variance genetic and environmental models fitted to all generations, to the six basic generations and the restricted set are shown in Tables VII to XV. The whole set of generations (P₁, P₂, F₁, F₂, BC₁, BC₂ and F₃) was used to estimate genetic parameters and to obtain the most complete genetic analysis. Genetic variability was always detected. A diverse genetic framework was found to control the analysed characteristics. This would represent the different genetic situations a breeder might face. Dispersion of genes in both parents was observed for practically all characters and sowing dates, as indicated by the small value for [d] and the transgressive distributions of the F₂ and F₃ generations. Heterotic F₁'s for yield were detected consistently in all sowing dates and in the third sowing for number of trifoliolate nodes and final height. Simple models involving additive genetic and dominance effects were found to control the number of trifoliolate leaf nodes. Final height was also found to be controlled by additive and dominance gene effects but linkage between repulsion fase loci was detected. Digenic

Table IV - Means, variances and degrees of freedom of the generations. September 26 sowing.

Genotype	First trifoliolate leaf			Days to flowering			Number of trifoliolate nodes			Final height (cm)			Grain weight (g)		
	df	mean	var	df	mean	var	df	mean	var	df	mean	var	df	mean	var
FT-2 (R1)	19	10.80	0.27	19	32.60	1.20	18	7.89	0.21	18	35.76	28.70	19	35.00	54.51
BR-13 (R2)	19	10.30	0.96	19	31.25	0.93	19	8.05	0.16	19	47.75	26.14	19	33.68	72.38
BR 85-29009 (Pat.)	18	10.32	0.67	18	38.05	1.94	18	9.21	0.95	18	47.89	110.77	18	49.85	70.47
OC-8 (Pat.)	19	11.15	0.56	19	38.45	5.21	19	10.55	1.31	19	64.55	109.71	19	61.66	87.74
F ₁	27	10.52	1.16	27	38.62	4.36	27	10.34	1.33	27	61.08	126.62	26	67.85	79.23
F ₂	65	10.24	0.57	65	38.40	9.02	65	10.35	2.51	65	60.85	244.32	62	60.46	169.88
BC ₁	32	11.42	0.73	32	39.44	6.07	32	10.66	0.97	32	65.67	118.22	32	67.29	122.43
BC ₂	38	10.99	1.14	38	40.05	7.57	38	10.38	1.74	38	59.18	147.34	38	63.14	83.53
F ₃ Total	55	10.55	0.82	56	39.02	15.04	56	10.26	2.72	56	61.76	289.77	56	54.19	173.01
Between (families)	28		1.18	28		21.75	28		3.62	28		435.65	28		181.77
Within (families)	27		0.46	28		8.32	28		1.82	28		143.88	28		164.24

Table V - Means, variances and degrees of freedom of the generations. October 17 sowing.

Genotype	First trifoliolate leaf			Days to flowering			Number of trifoliolate nodes			Final height (cm)			Grain weight (g)		
	df	mean	var	df	mean	var	df	mean	var	df	mean	var	df	mean	var
FT-2 (R1)	19	13.00	0.63	19	37.40	1.62	19	9.10	0.20	19	55.85	61.19	19	34.25	131.86
BR-13 (R2)	19	12.70	0.85	19	34.45	1.52	19	9.75	0.20	19	61.08	43.88	18	31.22	47.92
BR 85-29009 (Par.)	18	12.79	0.40	18	43.47	3.60	18	12.89	2.21	18	79.60	284.13	18	58.77	370.92
OC-8 (Par.)	19	12.80	0.69	19	42.00	1.79	19	13.60	1.41	19	90.85	120.82	19	68.22	152.24
F ₁	28	12.26	0.70	28	42.50	4.16	28	13.53	1.61	28	88.65	157.28	28	76.08	229.12
F ₂	68	12.06	1.44	68	42.63	10.09	68	13.21	2.87	68	85.48	300.60	68	65.61	282.53
BC ₁	32	13.61	0.70	32	43.90	11.24	36	14.00	1.48	31	95.54	182.38	32	73.57	225.12
BC ₂	32	13.01	0.69	32	43.97	12.22	36	13.39	2.08	32	88.07	263.69	31	67.52	270.44
F ₃ Total	58	12.57	1.00	58	42.48	11.17	56	12.94	4.01	56	61.76	289.77	55	61.10	418.59
Between (families)	28		1.33	28		14.60	28		4.72	28		435.65	28		507.91
Within (families)	30		0.70	30		7.75	28		3.30	28		143.88	27		329.27

Table VI - Means, variances and degrees of freedom of the generations. November 16 sowing.

Genotype	First trifoliolate leaf		Days to flowering		Number of trifoliolate nodes		Final height (cm)		Grain weight (g)						
	df	mean	var	df	mean	var	df	mean	var	df	mean	var			
FT-2 (R1)	19	12.65	0.45	19	39.35	2.87	19	12.00	0.53	19	75.03	98.33	19	56.03	301.13
BR-13 (R2)	19	12.20	0.96	19	36.75	0.51	19	12.00	0.42	19	72.15	90.82	19	46.02	251.43
BR 85-29009 (Par.)	19	12.15	0.45	18	45.26	3.76	18	14.32	1.34	18	95.89	193.10	18	55.54	310.59
OC-8 (Par.)	19	12.75	0.20	19	42.05	0.68	19	14.40	1.62	19	94.45	422.73	19	60.07	308.12
F ₁	28	12.23	0.59	28	44.50	2.40	27	15.20	0.77	27	110.03	225.03	28	65.53	303.76
F ₂	62	12.29	0.28	62	44.66	8.71	62	15.05	2.13	62	104.90	457.40	62	61.90	436.36
BC ₁	36	13.10	0.60	33	44.78	8.76	35	15.12	1.28	33	95.61	181.82	36	66.06	238.40
BC ₂	34	12.43	0.56	36	44.07	3.31	33	14.59	1.28	35	104.15	375.71	34	62.98	226.07
F ₃ Total	58	12.43	0.52	57	43.64	10.25	58	14.45	2.33	58	99.37	543.55	58	55.64	684.34
Between (families)	28		0.70	28		13.15	28		3.02	28		731.93	28		502.36
Within (families)	30		0.33	28		7.34	30		1.65	30		355.18	30		363.97

epistasis was significant in the control of days to flower and grain weight. For days to flower an additional complication in the form of genotype x microenvironmental interaction was detected: the ratios between the within generation variance of P_1 and P_2 (one tail F test) indicated that they were significantly different at the 5% and 1% levels for the 1st and 3rd sowings, respectively. The model fitting procedure detected G x E interaction for days to flower in the 3rd sowing. Since the combination of the results of both types of analysis brings the highest chance of obtaining more accurate results (Perkins and Jinks, 1970) we should accept the evidence of G x E for flowering time in the 3rd sowing. The genetic control of days to first trifoliolate leaf was the most complicated and involved digenic and trigenic epistasis, as detected by the means analysis. Alternatively, a model including the presence of linkage between digenic epistatic loci would also fit the means of days to first trifoliolate leaf. The variance analysis, on the other hand, confirmed that some of the loci are linked. Detailed study of this character is needed if consistent and unequivocal information is required. The genetic models and the performance of the F_3 families at each sowing date indicated that the genetic mechanism controlling the characters changed from the first to the last sowing. The [d], [h], [i] and [l] estimates changed significantly from the first to the third sowing dates and the best families for given characters for the September 26 sowing were not the same for the October 17 or the November 16 sowings.

The other two ways of grouping the generations, $P_1, F_1, F_2, BC_1, BC_2$ and P_1, P_2, F_2 and F_3 , were used to illustrate our procedures and to compare estimates and predictions. We tested our proposition of using the restricted set safely under all genetic situations studied above. When comparing the estimates obtained with the six generations and the restricted set we must take into considerations that some of the generations (P_1, P_2 and F_2) are common to both sets of data and, therefore, not independent. This fact, however, should not invalidate our results or preclude our discussions.

With the restricted set a loss of sensitivity to detect genetic effects occurred. Using the six basic generations we detected the presence of significant m, [d], [h], [i] and [l] in 15, 10, 11, 5 and 5 cases out of 15 possible, respectively. Using the restricted set we detected these significant estimates in 15, 8, 6, 0 and 0 cases out of the total of 15 opportunities. A significant [j] was detected only once from the whole group of seven generations or the six basic generations and could not be estimated from the restricted set. For variance effects estimation, however, the restricted set was slightly superior to the six basic generations. It was more sensitive to detect D in 12 out of 15 cases against 10 out of 15 cases, linkage in 2 out of 15 against 0 out of 15 (the variances of the six basic generations are all of rank 1 and cannot detect linkage) and equally efficient to detect genotype x microenvironmental interaction in one out of 15 cases.

Table VII - Genetic models adjusted to the means and variances of the P₁, P₂, F₁, F₂, BC₁, BC₂ and F₃. September 26 sowing.

	First trifoliolate leaf	Days to flowering	Number of trifoliolate nodes	Final height	Grain weight
(i) Mean estimates					
m	6.87 ± 0.60	38.18 ± 0.27	10.05 ± 0.14	58.31 ± 1.37	42.33 ± 3.48
[d]	0.42 ± 0.11	-	0.60 ± 0.14	7.80 ± 1.43	5.40 ± 1.22
[h]	26.99 ± 4.60	3.70 ± 1.29	0.55 ± 0.26	5.12 ± 2.60	51.68 ± 10.52
[i]	3.86 ± 0.58	-	-	-	13.52 ± 3.55
[j]	-	-	-	-	-
[l]	-57.67 ± 10.38	-3.27 ± 1.32	-	-	-26.23 ± 8.07
[ll] ¹	34.33 ± 6.39	-	-	-	-
χ ²	0.00 ^{n.s.}	9.22 ^{n.s.}	5.72 ^{n.s.}	5.29 ^{n.s.}	1.12 ^{n.s.}
df	1	4	4	4	2
(ii) Variance estimates					
D	-	12.38 ± 2.68	2.10 ± 0.61	229.90 ± 60.77	125.04 ± 41.75
H	-	-	-	-	-
F	-	-	2.26 ± 1.04	-	-
E	0.80 ± 0.07	3.93 ± 0.65	1.16 ± 0.18	106.41 ± 16.96	85.38 ± 13.18
χ ²	15.57 ^{/2}	4.95 ^{n.s.}	2.72 ^{n.s.}	2.66 ^{n.s.}	5.79 ^{n.s.}
df	7	6	5	6	6

/1 - Dominance type trigenic epistasis.

/2 - Best model found. χ₇² is significant at 3% level of probability.

Table VIII - Genetic models adjusted to the means and variances of the P₁, P₂, F₁, F₂, BC₁, BC₂ and F₃. October 17 sowing.

	First trifoliolate leaf	Days to flowering	Number of trifoliolate nodes	Final height	Grain weight
(i) Mean estimates					
m	7.79 ± 0.72	42.80 ± 0.16	13.41 ± 0.09	88.37 ± 1.00	61.06 ± 1.93
[d]	-	0.70 ± 0.24	0.44 ± 0.17	5.39 ± 1.83	6.00 ± 2.11
[h]	35.38 ± 5.67	-	-	-	14.17 ± 3.61
[i]	5.00 ± 0.71	-	-	-	-
[j]	1.20 ± 0.41	-	-	-	-
[l]	-76.49 ± 13.17	-	-	-	-
[ll] ¹	45.58 ± 8.22	-	-	-	-
χ ²	0.00 ^{n.s.}	9.35 ^{n.s.}	9.34 ^{n.s.}	9.00 ^{n.s.}	5.72 ^{n.s.}
df	1	5	5	4	4
(ii) Variance estimates					
D	0.89 ± 0.30	15.74 ± 3.02	2.32 ± 0.79	-	-
D1	-	-	-	228.62 ± 102.36	-
D2	-	-	-	1216.96 ± 435.48	-
H	-	-	-	-	-
F	-	-	-	-	-
E	0.64 ± 0.10	3.94 ± 0.65	1.69 ± 0.25	179.13 ± 28.41	294.41 ± 26.28
χ ²	6.35 ^{n.s.}	13.51 ²	6.40 ^{n.s.}	5.53 ^{n.s.}	12.11 ^{n.s.}
df	6	6	6	5	7

/1 - Dominance type trigenic epistasis.

/2 - Best model found. χ₆² is significant at 4% level of probability.

Table IX - Genetic models adjusted to the means and variances of the P₁, P₂, F₁, F₂, BC₁, BC₂ and F₃. November 16 sowing.

	First trifoliolate leaf	Days to flowering	Number of trifoliolate nodes	Final height	Grain weight
(i) Mean estimates					
m	10.54 ± 0.46	43.68 ± 0.21	14.39 ± 0.14	95.08 ± 2.03	57.12 ± 2.10
[d]	0.37 ± 0.08	1.51 ± 0.21	-	-	-
[h]	13.74 ± 3.51	1.14 ± 0.36	0.90 ± 0.24	14.23 ± 3.71	10.31 ± 4.05
[i]	1.88 ± 0.45	-	-	-	-
[j]	-	-	-	-20.07 ± 7.67	-
[l]	-28.90 ± 7.86	-	-	-	-
[ll] ¹	16.84 ± 4.82	-	-	-	-
χ ²	3.33 ^{n.s.}	7.45 ^{n.s.}	6.14 ^{n.s.}	2.57 ^{n.s.}	5.62 ^{n.s.}
df	1	4	5	4	5
(ii) Variance estimates					
D	-	12.59 ± 2.27	1.58 ± 0.54	349.02 ± 121.53	-
H	-	-	-	-	-
F	-	-7.45 ± 3.23	-	565.82 ± 201.97	-
E	0.45 ± 0.04	3.19 ± 0.64 ³	1.13 ± 0.17	256.09 ± 39.29	346.11 ± 30.65
χ ²	16.62 ²	4.00 ^{n.s.}	4.77 ^{n.s.}	6.00 ^{n.s.}	9.21 ^{n.s.}
df	7	4	6	5	7

/1 - Dominance type trigenic epistasis.

/2 - Best model found. χ² is significant at 3% level of probability.

/3 - E1; E2 - 0.70 ± 0.23 indicating G x E interaction.

Table X - Models adjusted to the means of the six basic generations and the restricted set. September 26 sowing.

	First trifoliolate leaf	Days to flowering	Number of trifoliolate nodes	Final height	Grain weight
(i) Six basic generations					
m	6.87 ± 0.60	32.81 ± 1.94	10.02 ± 0.15	57.73 ± 1.47	36.00 ± 8.20
[d]	0.42 ± 0.11	-	0.60 ± 0.14	7.81 ± 1.43	5.45 ± 1.23
[h]	9.83 ± 1.60	16.53 ± 4.81	0.57 ± 0.27	5.74 ± 2.67	65.99 ± 19.80
[i]	3.86 ± 0.58	5.35 ± 0.27	-	-	19.72 ± 8.08
[j]	-	-	-	-	-
[l]	-6.18 ± 1.08	-10.73 ± 3.02	-	-	-34.14 ± 12.29
χ^2	0.00 ^{n.s.}	1.42 ^{n.s.}	5.47 ^{n.s.}	4.18 ^{n.s.}	0.39 ^{n.s.}
df	1	2	3	3	1
(ii) Restricted set					
m	10.76 ± 0.11	38.36 ± 0.20	10.13 ± 0.11	59.07 ± 1.10	58.81 ± 1.33
[d]	0.41 ± 0.12	-	0.70 ± 0.17	8.24 ± 1.68	5.83 ± 1.42
[h]	-1.01 ± 0.31	-	-	-	8.78 ± 4.33
[i]	-	-	-	-	-
[l]	-	-	-	-	-
χ^2	0.19 ^{n.s.}	2.63 ^{n.s.}	3.79 ^{n.s.}	5.15 ^{n.s.}	3.64 ^{n.s.}
df	1	3	2	2	1

Table XI - Models adjusted to the means of the six basic generations and the restricted set. October 17 sowing.

	First trifoliolate leaf	Days to flowering	Number of trifoliolate nodes	Final height	Grain weight
(i) Six basic generations					
m	7.79 ± 0.72	42.85 ± 0.17	13.47 ± 0.10	88.11 ± 1.06	62.41 ± 2.14
[d]	-	0.72 ± 0.24	0.42 ± 0.17	5.45 ± 1.85	5.53 ± 2.13
[h]	12.59 ± 1.73	-	-	-	12.62 ± 3.76
[i]	5.00 ± 0.71	-	-	-	-
[j]	1.20 ± 0.41	-	-	-	-
[l]	-8.12 ± 1.07	-	-	-	-
χ^2	0.00 ^{n.s.}	8.71 ^{n.s.}	5.91 ^{n.s.}	8.27 ^{n.s.}	3.63 ^{n.s.}
df	1	4	4	4	3
(ii) Restricted set					
m	12.83 ± 0.11	42.66 ± 0.19	13.19 ± 0.13	86.40 ± 1.39	64.52 ± 1.34
[d]	-	0.71 ± 0.26	-	5.12 ± 2.16	-
[h]	-1.42 ± 0.36	-	-	-	-
[i]	-	-	-	-	-
[l]	-	-	-	-	-
χ^2	0.81 ^{n.s.}	0.26 ^{n.s.}	4.06 ^{n.s.}	2.44 ^{n.s.}	5.35 ^{n.s.}
df	2	2	3	2	3

Table XII - Models adjusted to the means of the six basic generations and the restricted set. November 16 sowing.

	First trifoliolate leaf	Days to flowering	Number of trifoliolate nodes	Final height	Grain weight
(i) Six basic generations					
m	10.53 ± 0.46	43.75 ± 0.22	14.46 ± 0.16	94.22 ± 2.22	62.86 ± 1.22
[d]	0.37 ± 0.08	1.57 ± 0.22	-	-	-
[h]	5.35 ± 1.24	1.09 ± 0.36	0.82 ± 0.25	14.40 ± 3.85	-
[i]	1.89 ± 0.45	-	-	-	-
[j]	-	-	-	-	-
[l]	-3.64 ± -0.84	-	-	-	-
χ^2	3.38 ^{n.s.}	6.61 ^{n.s.}	5.28 ^{n.s.}	9.22 ^{n.s.}	6.19 ^{n.s.}
df	1	3	4	4	5
(ii) Restricted set					
m	12.37 ± 0.05	43.59 ± 0.23	14.27 ± 0.18	95.21 ± 2.44	58.55 ± 1.57
[d]	0.33 ± 0.08	1.56 ± 0.24	-	-	-
[h]	-	1.83 ± 0.87	1.40 ± 0.53	18.93 ± 7.51	-
[i]	-	-	-	-	-
[l]	-	-	-	-	-
χ^2	2.53 ^{n.s.}	1.19 ^{n.s.}	1.16 ^{n.s.}	0.11 ^{n.s.}	3.48 ^{n.s.}
df	2	1	2	2	3

Table XIII - Models adjusted to the variances of the six basic generations and the restricted set. September 26 sowing.

	First trifoliolate leaf	Days to flowering	Number of trifoliolate nodes	Final height	Grain weight
(i) Six basic generations					
D	-	10.61 ± 3.21	2.16 ± 0.82	213.42 ± 79.69	158.90 ± 57.22
H	-	-	-	-	-
F	-	-	2.25 ± 1.05	-	-
E	0.79 ± 0.08	3.98 ± 0.68	1.14 ± 0.19	109.90 ± 18.66	75.87 ± 13.01
χ^2	10.24 ^{n.s.}	3.91 ^{n.s.}	2.62 ^{n.s.}	2.24 ^{n.s.}	2.09 ^{n.s.}
df	5	4	3	4	4
(ii) Restricted set					
D	-	12.97 ± 2.96	2.33 ± 0.71	258.88 ± 70.23	112.16 ± 47.63
H	-	-	-	-	-
F	-	-	-	-	-
E	0.67 ± 0.07	3.70 ± 0.82	1.19 ± 0.25	105.90 ± 22.93	99.04 ± 20.03
χ^2	10.32 ^{/1}	4.73 ^{n.s.}	0.81 ^{n.s.}	0.42 ^{n.s.}	3.16 ^{n.s.}
df	4	3	3	3	3

/1 - Best model found; χ^2_4 - is significant at 3% level of probability.

Table XIV - Models adjusted to the variances of the six basic generations and the restricted set. October 17 sowing.

	First trifoliolate leaf	Days to flowering	Number of trifoliolate nodes	Final height	Grain weight
(i) Six basic generations					
D	1.34 ± 0.44	20.34 ± 4.10	-	227.52 ± 114.36	-
H	-	-	-	-	-
F	-	-	-	-	-
E	0.57 ± 0.10	3.57 ± 0.62	2.12 ± 0.32	178.49 ± 29.82	259.10 ± 26.17
χ^2	4.53 ^{n.s.}	9.29 ^{n.s.}	7.34 ^{n.s.}	5.39 ^{n.s.}	4.06 ^{n.s.}
df	4	4	5	4	5
(ii) Restricted set					
D	1.13 ± 0.34	13.37 ± 2.70	2.21 ± 0.94	213.47 ± 118.57 ¹	-
H	-	-	-	-	-
F	-	-	-	-	-
E	0.58 ± 0.12	2.93 ± 0.65	1.99 ± 0.40	197.47 ± 43.10	324.33 ± 36.26
χ^2	5.47 ^{n.s.}	3.75 ^{n.s.}	2.36 ^{n.s.}	3.19 ^{n.s.}	7.91 ^{n.s.}
df	3	3	3	2	4

¹ - D1; D2 - 1163.73 ± 447.92 indicating repulsion linkage.

Table XV - Models adjusted to the variances of the six basic generations and the restricted set. November 16 sowing.

	First trifoliolate leaf	Days to flowering	Number of trifoliolate nodes	Final height	Grain weight
(i) Six basic generations					
D	-	13.81 ± 2.87	1.62 ± 0.76	312.48 ± 168.81	-
H	-	-	-	-	-
F	-	-8.37 ± 3.42	-	546.04 ± 205.32	-
E	0.44 ± 0.04	-	1.12 ± 0.19	258.78 ± 43.60	321.18 ± 32.36
E1	-	2.97 ± 0.61	-	-	-
E2	-	0.68 ± 0.22	-	-	-
χ^2	12.27 ^{/1}	1.47 ^{n.s.}	4.78 ^{n.s.}	5.83 ^{n.s.}	6.74 ^{n.s.}
df	5	2	4	3	5
(ii) Restricted set					
D	0.86 ± 0.41 ^{/2}	11.78 ± 2.57	1.26 ± 0.64	324.61 ± 141.64	-
H	-	-	-	-	-
F	-	-	-	-	-
E	0.30 ± 0.04	-	1.45 ± 0.29	299.19 ± 60.01	404.36 ± 45.64
E1	-	4.58 ± 1.44	-	-	-
E2	-	0.70 ± 0.27	-	-	-
χ^2	7.32 ^{n.s.}	2.57 ^{n.s.}	0.26 ^{n.s.}	2.84 ^{n.s.}	2.19 ^{n.s.}
df	3	2	3	3	4

/1 - Best model found; χ^2 is significant at 3% level of probability.

/2 - D2; D1 is not significant indicating repulsion linkage.

In general, the restricted set of generations allowed a partial and reasonably accurate analysis of the genetic effects controlling the experimental material. The method was sufficiently accurate to detect main mean or variance genetical effects, the presence of G x E interaction and linkage. It was inherently weak in detecting epistasis. Therefore, when estimation of main genetical and environmental effects suffices, or when only predictions are required or when preliminary information is necessary for planning large and detailed experiments, the genetical information provided by the restricted set of generations can be very useful.

The mean parameters estimated by the six generations set are highly correlated whenever digenic epistasis is present. J.L. Jinks (unpublished data) showed that in fourteen environments and for seven characters the correlation between m and $[i]$ and also $[h]$ and $[l]$ can be as high as 95%. The correlation make those estimates very difficult to interpret, especially when additional sources of complications like G x E interaction or linkage between interacting loci are present. In our models for instance, it is difficult to accept that for days to flowering in the first sowing and first trifoliolate leaf in the second sowing the $[i]$ estimate, which is highly affected by gene dispersion, is significant while that of $[d]$ which is also sensitive for gene dispersion is non-significant. Those correlations, because of their influence on the estimate of the m parameter, may sometimes lead to non-reliable predictions (days to flowering, first sowing and first trifoliolate leaf, second sowing). In those cases, perhaps, $m + [i]$ would be an estimate of m , but that could only be ascertained by further study involving other generations or designs.

The predictions made from both sets of generations for the three sowing dates are given in Table XVI. For most characters and categories they are close. Their degree of correlation must be high. We do not present the analysis here because of the lack of independence of the parameter estimates. We can, however, correlate the predictions made for the November 16 sowing with the observations made from the November 17 sowing, in the experimental field. These predictions and observations are shown in Table XVII. Even though the predictions and observations were made under different conditions, glass house and experimental field, respectively, in different years, the correlations were very high (0.85 for the six basic generations and for the restricted set) clearly demonstrating the power of the approach. The predictions obtained from the restricted set were as good as those obtained from the more complete set of generations. It became clear, however, as described by Toledo (1989), that predictions are qualitative rather than quantitative, i.e., they can screen between materials better than they can accurately predict the expected number of inbreds falling in each category. Except for two cases our predictions consistently overestimated the proportions of inbreds scoring above P_1 and R_1 .

Table XVI - Expected proportions of superior inbred lines falling in each category. Expectations obtained using the six basic generations and the restricted set for the three sowing dates (I, II and III).

Category	Days to first trifoliate leaf			Days to flowering			Number of trifoliate nodes			Final height			Grain weight		
	I	II	III	I	II	III	I	II	III	I	II	III	I	II	III
	(i) Six basic generations														
≥ P1	-	0.00	-	.04	.47	.34	.36	-	.52	.32	.43	.46	.02	-	-
≤ P2	-	1.00	-	.95	.42	.32	.29	-	.46	.24	.24	.50	.86	-	-
≥ R1	-	0.00	-	.52	.89	.88	.93	-	.97	.93	.98	.86	.53	-	-
≥ R2	-	0.00	-	.68	.97	.97	.91	-	.97	.76	.96	.89	.57	-	-
(ii) Restricted set															
≥ P1	-	.47	.34	.49	.45	.31	.39	.39	.45	.37	-	.48	.26	-	.45
≤ P2	-	.49	.40	.46	.43	.33	.27	.34	.52	.24	-	.48	.32	-	.41
≥ R1	-	.44	.38	.94	.92	.89	.93	1.00	.98	.93	-	.87	.97	-	.57
≥ R2	-	.55	.57	.98	.99	.98	.91	.99	.98	.75	-	.90	.98	-	.83

Table XVII - Predicted (F_∞) (16/Nov/88 sowing) and observed (F₄) (17/Nov/89 sowing) number of inbred lines falling in each specified category.

Category	First trifoliolate leaf		Days to flowering		Number of trifoliolate nodes		Final height		Grain weight	
	E	O	E	O	E	O	E	O	E	O
(i) Sic basic generations										
≥ P1	-	5	10	9	14	7	16	9	-	9
≤ P2	-	8	10	14	15	19	14	14	-	21
≥ R1	-	10	26	18	26	19	29	25	-	16
≥ R2	-	21	29	30	27	28	29	29	-	27
(ii) Restricted set										
≥ P1	10	5	9	9	14	7	13	9	-	9
≤ P2	12	8	10	14	14	19	16	14	-	21
≥ R1	11	10	27	18	26	19	29	25	-	16
≥ R2	17	21	29	30	27	28	29	29	-	27

RESUMO

As análises genéticas de plantas que se reproduzem por autofecundação são usualmente realizadas através do estudo das médias e variâncias das gerações P₁, P₂, F₁, F₂, RC₁ e RC₂. As gerações F₁, RC₁ e RC₂ são obtidas por polinização artificial e para algumas espécies como a soja, o arroz e o trigo podem demandar bastante tempo e trabalho. Se análises genéticas confiáveis puderem ser obtidas de um conjunto restrito de gerações, envolvendo os parentais, F₂ e F₃, facilmente conseguidas em praticamente todas as plantas de autopolinização, o conhecimento da arquitetura genética de importantes características seria mais facilmente obtido. A análise dos dados deste trabalho, que envolve uma grande gama de diferentes situações e modelos genéticos, indicou que quando o maior interesse do experimentador é detectar e estimar efeitos aditivos genéticos e dominância e efeitos aditivos do ambiente o conjunto restrito é adequado. O conjunto restrito de gerações também permite detectar interações do tipo genótipo x ambiente e ligações gênicas. Entretanto, se o interesse reside em estimar efeitos epistáticos outras gerações são necessárias. As previsões realizadas com o conjunto restrito são tão precisas quanto aquelas obtidas com P₁, P₂, F₁, F₂, RC₁ e RC₂.

REFERENCES

- Cavalli, L.L. (1952). An analysis of linkage in quantitative inheritance. In: *Quantitative Inheritance* (E.C.R. Reeve and C.H. Waddington, eds.). pp. 135-144, HMSO, London.
- Comstock, R.E. and Robinson, H.F. (1952). Estimation of average dominance of genes. In: *Heterosis* (Gowen, J.W., ed.). pp. 494-516, Iowa State College Press, Ames, Iowa.
- Hayman, B.I. (1960). Maximum likelihood estimation of genetic components of variation. *Biometrics* 16: 369-381.
- Jinks, J.L. (1978). Unambiguous test for linkage of genes displaying non-allelic interactions for a metrical trait. *Heredity* 40: 171-173.
- Jinks, J.L. and Jones, R.M. (1958). Estimation of the components of heterosis. *Genetics* 43: 223-234.
- Jinks, J.L. and Pooni, H.S. (1976). Predicting the properties of recombinant inbred lines derived by single seed descent. *Heredity* 36: 253-266.
- Kearsey, M.J. and Jinks, J.L. (1968). A general method of detecting additive, dominance and epistatic variation for metrical traits: I Theory. *Heredity* 23: 403-409.
- Mather, K. (1949). *Biometrical Genetics* (1st Edn.). Methuen, London.
- Mather, K. and Jinks, J.L. (1982). *Biometrical Genetics* (3rd Edn.), Chapman and Hall, London.
- Perkins, J.M. and Jinks, J.L. (1970). Detection and estimation of genotype-environmental, linkage and epistatic components of variation for a metrical trait. *Heredity* 25: 157-177.
- Pooni, H.S. and Jinks, J.L. (1982). Comparative analysis of association and dispersion crosses to detect linkage and epistatic components of variation. *Heredity* 49: 211-220.
- Toledo, J.F.F. de. (1987). Predicting the inbreeding and the outcrossing potential of soybean (*Glycine max* (L.) Merrill) varieties. *Rev. Brasil. Genet.* 10: 543-558.

- Toledo, J.F.F. de. (1989). Quantitative Genetics in Soybean Breeding. *Proceedings of the World Soybean Conference IV* (ed. Pascall, A.J.). pp. 903-914, Buenos Aires.
- Toledo, J.F.F. de. (1991). Programa de Computador para Estimar Parâmetros Genéticos pelo Método dos Quadrados Mínimos Ponderados. *Pesq. Agropec. Bras.* 26: 1023-1039.
- Van der Veen, J.H. (1959). Test of non-allelic interaction and linkage for quantitative characters in generations derived from two diploid pure lines. *Genetica* 30: 201-232.

(Received October 5, 1990)