The effect of data transformation on the genetic parameter estimates of resistance to *Cercospora Sojina* Hara in soybean

Geraldo de Amaral Gravina*[1]; Paulo Roberto Cecon[2]; Sebastião Martins Filho[3] and Carlos Sigueyuki Sediyama[1]

[1]Departamento de Fitotecnia, Universidade Federal de Viçosa (UFV), CEP 36571-000, Minas Gerais, Brazil; [2]Departamento de Informática, UFV; [3]Departamento de Engenharia Rural, CAUFES, Universidade Federal do Espírito Santo (UFES), CEP 29500-000, Alegre, ES, Brazil. (* Corresponding Author. E-mail: gravina@escola24horas.com.br)

ABSTRACT

The objective of this work was to study the effect of data transformation on the estimates of genetic parameters of soybean resistance to *Cercospora sojina* Hara inheritance. The means and variances of populations (P_1 and P_2, and the F_1, F_2, RC_1 and RC_2 generations) derived from a cross between a resistant ('Paraná') and a susceptible ('Bossier') cultivar were evaluated for five characteristics associated with the disease: infection degree (ID) – visually evaluated; number of lesions per leaflet (NLF); lesion mean diameter (LMD); percentage of lesioned foliar area (PLFA) and disease index (DI). The studied transformations were: square root of (x), (x + 1), (x + 0,5) and (x + 3/8) and logarithm (x+1). Results showed that the transformations did not alter the significance of the estimated genetic parameters. The additive genetic effect was the most important, for all the evaluated traits. The additive-dominant model was sufficient to explain the variations only for PLFA; for the other traits, the non-allelic interactions were significant. The influence of the effects of the epistatic interactions varied from 5.89% for PLFA up to 35.93% for NLF. For transformed or non-transformed data, the additive-dominant genetic model was satisfactory to explain the behavior of the generation means for the PLFA character. However, for ID, NLF, LMD and DI, the additive-dominant-epistatic model was more appropriate. The ID, NLF and DI traits, which presented higher additive genetic variability than dominance variability, can be utilized by breeders without transformation as selection criteria, to obtain superior homozygous genotypes with selection beginning at the F_2 generation.

KEY WORDS: *Glycine max*, scale change, generation analysis, frogeye leaf spot.

INTRODUCTION

Among the main factors that limit high economic returns in soybean cropping are the diseases that, in general, are difficult to control (Yorinori, 2001). Frogeye leaf spot or Cercospora leaf spot, one of the main foliar diseases, was responsible for great losses of soybean in Brazil. Currently, this disease is under control through genetic resistance and its occurrence is sporadic, but new physiological races are likely to appear (Yorinori and Klingelfuss, 1999).

The use of resistant cultivars and the incorporation of genes of resistance into commercial susceptible cultivars are the most economical and efficient means of controlling frogeye leaf spot (Cordeiro et al., 1992; Martins Filho, 1999; Yorinori, 2001). Therefore, it is necessary to know the mechanism of character inheritance.

The evaluation of the nature and magnitude of the genetic effects that control the resistance to *C. sojina* Hara, is of great importance for any plant breeder who aims at incorporating genes of resistance into cultivated soybean. It is paramount to investigate, in the genetic fraction, which proportion can be attributed to the additive, dominance and epistatic genetic factors (Cruz and Regazzi, 2001).

Estimates of the relative proportion of the additive, dominance and epistatic genic effects have been reported for a number of crops, using several types of hybrids and populations and varied statistical-genetic procedures. Among the statistical-genetic procedures, the generation analysis, which deals with the averages and variances of different kinds of populations, is a useful tool. Thus the sampling distribution of the experimental data can pose a problem.

To estimate genetic parameters, there is no need of data normality, but to test the significance of the effect of the model, this assumption becomes fundamental. In experiments of this nature, in which many variables are worked on, it may be necessary to transform the data according to their distribution.

Bartlett (1936) states that the logarithmic transformation can be used to stabilize the variance whenever the standard deviation is proportional to the mean, and when the variance is proportional to the mean, the square-root transformation may be more adequate.

According to Demétrio (1978), the following transformations can be used to stabilize variances: arcsine, for binomial variables; square-root, for Poisson distribution with variance proportional to the mean; logarithmic transformation, for distribution with variance proportional to the mean square, and hyperbolic arcsine for negative binomial variables.

The objective of this work was to verify the possible alterations data transformations can bring about in the estimates of genetic parameters of the inheritance of resistance to *C. sojina* Hara in soybean, using the generation analysis, also known as the joint scaling test.

MATERIAL AND METHODS

Resistance of parental (P_1 and P_2) and F_1, F_2, BC_1 and BC_2 generations from a cross between a resistant (Paraná) and a susceptible (Bossier) soybean cultivar, and artificially inoculated with the *Cercospora sojina* Hara fungus, were evaluated in a trial carried out in a greenhouse at the Federal University of Viçosa Plant Science Department, Viçosa, Minas Gerais. The number of plants evaluated per population was as follows: P_1 26, P_2 23, F_1 26, F_2 219, BC_1 56, and BC_2 60 plants. A completely randomized experimental design was utilized.

Race 04 of *C. sojina*, obtained from the EMBRAPA (CNPSo) National Center for Soybean Research at Londrina-PR, was used. Multiplication was made in Petri dishes, containing tomato-agar as culture medium, prepared similarly to the V8-agar medium described by Veiga (1973), but substituting the V8 by a spiceless tomato juice (Superbom brand).

The pathogen was inoculated when the plants presented the third trifoliate leaf fully developed, using 10 mL per plant of a 40.000 conidia per mL suspension. Immediately after the inoculation, the pots were taken to a humid chamber where they remained for 3 days under 100% relative humidity.

Twenty days after inoculation, when the symptoms of the disease had already been established, the following traits were evaluated: a) infection degree (ID) – symptoms evaluated visually using a score scale (1.0 – without apparent infection and 5.0 - the maximum infection degree); b) number of lesions per foliole (NLF) – taken from the most infected plant leaflet; c) lesion mean diameter (LMD) – an average of the 10 largest lesions in the most infected foliole, in millimeters; d) percentage of lesioned foliar area (PLFA) - lesioned foliar area divided by the area of the foliole and multiplied by one hundred; and e) disease index (DI) - the number of lesions per foliole multiplied by the square of the lesion mean diameter, $DI = NLF \times LMD^2$. Items from a to e were evaluated according to the methodology used by Cordeiro et al. (1992), Veiga (1973) and Martins Filho (1999).

Data were analyzed with and without the transformations: square root of (x), (x+1), (x+0,5), or (x+3/8) and to log of (x+1), were x is the original value of the trait.

The genetic analysis of the inheritance of soybean resistance to *C. sojina* based on generation means and variances were made by the method of Mather and Jinks (1984), using the GENES software (Cruz, 2001).

RESULTS AND DISCUSSION

The results presented here are related to the data with and without square root of (x + 1) and to log (x + 1) transformations. Square root of (x), (x+0.5) or (x+3/8) transformations presented the same results as the square root of (x + 1) transformation.

The means, variances and variances of the mean for each of the studied trait, without and with transformation of the data, are presented in Table 1. F_1 and BC_2 character means were closer to the means of the resistant parent (Paraná), indicating dominance for resistance to the fungus. F_2 and BC_1 generation means were a little more distant from the means of the resistant parent, as expected, due to the effect of gene segregation in F2, and the expression of the recessive alleles in BC_1. Some traits, such as the NLF, presented high mean variance. However, either the square root or the log transformation reduced the variance of the mean, mainly of the parent, showing higher expression of the disease symptoms (P_1). After these transformations, the generation that presented greater variance was F_2, as expected. A possible explanation for this high variance (data with no transformation), in the case of NLF, lies on the fact that the lesions coalesce, making the evaluation of the actual number of lesioned points in some leaves difficult. For other analyzed characteristics, such as the ID (one of the traits considered by some authors as the most reliable, although subjectively), lower

variances were obtained. The F_2 population presented the largest variance since this generation presents the largest gene segregation. For ID, the used transformations promoted small alterations in variance behavior.

The estimates of phenotypic, genotypic, additive, dominance and environmental variances, of broad as well as narrow sense heritabilities, average degree of dominance and number of genes of each studied trait, with and without transformation, are presented in

Table 1. Means $(\hat{m})$, variances $(\hat{\sigma}^2)$ and variances of the mean $\left[\hat{v}(\hat{m})\right]$ of parents and their derived populations for the ID, NLF, LMD, PLFA and DI traits evaluated of soybean reaction to *C. sojina* Hara trial.

Population[2]	Character[1]								
	ID			Square root (ID+1)			Log (ID+1)		
	$(\hat{m})$	$\hat{\sigma}^2$	$\hat{v}(\hat{m})$	$(\hat{m})$	$\hat{\sigma}^2$	$\hat{v}(\hat{m})$	$(\hat{m})$	$\hat{\sigma}^2$	$\hat{v}(\hat{m})$
P1	4.60	0.1604	0.00617	2.36	0.0073	0.0003	0.7447	0.0010	0.00004
P2	2.02	0.3236	0.01407	1.75	0.0316	0.0013	0.4805	0.0077	0.00032
F1	1.93	0.3886	0.01495	1.70	0.0318	0.0012	0.4579	0.0080	0.00031
F2	2.36	1.1585	0.00919	0.81	0.0775	0.0006	0.5073	0.0163	0.00013
RC1	2.57	1.0815	0.01803	1.87	0.0787	0.0013	0.5330	0.0178	0.00030
RC2	1.47	0.2895	0.00491	1.56	0.0266	0.0005	0.3845	0.0076	0.00013

Population[2]	NLF			Square root (NLF+1)			Log (NLF+1)		
	$(\hat{m})$	$\hat{\sigma}^2$	$\hat{v}(\hat{m})$	$(\hat{m})$	$\hat{\sigma}^2$	$\hat{v}(\hat{m})$	$(\hat{m})$	$\hat{\sigma}^2$	$\hat{v}(\hat{m})$
P1	139.69	1872.78	72.03	11.73	3.1064	0.1195	2.13	0.0160	0.0006
P2	20.96	215.32	9.36	4.62	4.0239	0.1677	1.24	0.1662	0.0069
F1	16.92	147.59	5.68	3.98	2.1684	0.0834	1.13	0.1317	0.0051
F2	27.47	820.68	6.51	4.72	6.2040	0.0492	1.22	0.2469	0.0020
RC1	19.73	279.62	4.66	4.18	3.2961	0.0549	1.14	0.2155	0.0036
RC2	7.78	79.83	1.35	2.57	2.2304	0.0378	0.65	0.3130	0.0053

Population[2]	LMD			Square root (LMD+1)			Log (LMD+1)		
	$(\hat{m})$	$\hat{\sigma}^2$	$\hat{v}(\hat{m})$	$(\hat{m})$	$\hat{\sigma}^2$	$\hat{v}(\hat{m})$	$(\hat{m})$	$\hat{\sigma}^2$	$\hat{v}(\hat{m})$
P1	0.3481	0.0041	0.00016	1.16	0.0007	0.00003	0.129	0.0004	0.000016
P2	0.2196	0.0068	0.00029	1.10	0.0013	0.00005	0.086	0.0008	0.000034
F1	0.1154	0.0006	0.00002	1.06	0.0001	0.00001	0.047	0.0001	0.000003
F2	0.2313	0.0105	0.00008	1.11	0.0021	0.00002	0.089	0.0013	0.000010
RC1	0.3133	0.0235	0.00039	1.14	0.0046	0.00008	0.115	0.0027	0.000045
RC2	0.1322	0.0146	0.00025	1.06	0.0032	0.00005	0.051	0.0021	0.000036

Population[2]	PLFA			Square root (PLFA+1)			Log (PLFA+1)		
	$(\hat{m})$	$\hat{\sigma}^2$	$\hat{v}(\hat{m})$	$(\hat{m})$	$\hat{\sigma}^2$	$\hat{v}(\hat{m})$	$(\hat{m})$	$\hat{\sigma}^2$	$\hat{v}(\hat{m})$
P1	23.08	148.6081	5.7157	4.7892	1.1931	0.0459	1.3411	0.0338	0.0013
P2	2.65	9.0340	0.3928	1.8179	0.6083	0.0253	0.4631	0.1248	0.0052
F1	0.56	0.1478	0.0057	1.2404	0.0230	0.0009	0.1809	0.0111	0.0004
F_2	4.69	63.3413	0.5027	2.0158	1.6335	0.0130	0.4787	0.1990	0.0016
RC_1	12.06	228.8764	3.8146	3.1463	3.2171	0.0536	0.8581	0.2516	0.0042
RC_2	1.66	4.9009	0.0831	1.5134	0.3719	0.0063	0.2981	0.1031	0.0017

Population[2]	DI			Square root (DI+1)			Log (DI+1)		
	$(\hat{m})$	$\hat{\sigma}^2$	$\hat{v}(\hat{m})$	$(\hat{m})$	$\hat{\sigma}^2$	$\hat{v}(\hat{m})$	$(\hat{m})$	$\hat{\sigma}^2$	$\hat{v}(\hat{m})$
P_1	16.88	47.9776	1.8453	15.78	4.96	0.1907	0.0362	0.000130	0.000005
P_2	1.53	3.0633	0.1332	5.73	7.21	0.3006	0.0114	0.000080	0.000003
F_1	0.25	0.0486	0.0019	4.45	2.77	0.1055	0.0027	0.000002	0.000000
F_2	2.98	24.8413	0.1972	6.03	13.04	0.1035	0.0128	0.000180	0.000001
RC_1	3.25	20.2005	0.3367	5.68	7.93	0.1322	0.01214	0.000320	0.000005
RC_2	0.46	0.4538	0.0077	3.06	3.89	0.0660	0.0054	0.000040	0.000001

[1] ID – infection degree; NLF – number of lesions per foliole; LMD – Lesion mean diameter; PLFA – percentage of lesioned foliar area; and DI – disease index; [2] P1 (Bossier); P2 (Paraná); BC1 (F1 x Bossier) and BC2 (F1 x Paraná). P_1 (N=26 plants); P_2 (N=23); F_1 (N=26); F_2 (N=219); BC_1 (N=56) and BC_2 (N=60).

Table 2. Dominance variance for ID, NLF and DI presented negative or near zero estimates, and were considered null. With the log transformation, the problem of these negative estimates was resolved for these variables; however, it generated negative additive variance for NLF. For LMD and PLFA, the additive variance estimates were considered null, generating estimates of narrow sense heritabilities equal to zero. In these traits, transformations did not modify the sign of the additive variance, except the log transformation for PLFA. It is important to emphasize that for the ID, NLF and ID variables, high estimates of additive genetic variance were obtained, indicating the additive genetic effect as the most important in breeding programs aiming at the resistance of soybean to *C. sojina*.

In selected materials, the additive genetic effect is reduced specially in characters of less complex inheritance. In quantitative characters of more complex inheritance, the contribution of the genetic effect resulting from dominance becomes larger. The trait dominant gene action cannot be fixed and the exploitation of heterozygosis is the only solution (Gamble, 1962; Arias, 1986).

The dominance effect and its epistatic interactions are of lesser significance for the autogamous species. Thus, breeding techniques that take advantage of additive variance for the attainment of higher gains will be more important for the improvement of soybean resistance to frogeye leaf spot. Amongst the evaluated traits of this study, those that presented the most important additive genic effect, such as ID, PLFA and ID, must be prioritized in the breeding program.

The average degree of dominance, calculated on the basis of variance, could not be computed for the majority of the studied traits, because the estimates of the variance due to dominance were considered null for characters ID, NLF and ID. On the other hand, the LMD and PLFA characters did not present null estimates for dominance variance, but did for additive genetic variance. Thus, in an analogous way, it was not possible to obtain estimates for degree of dominance based on variances. Estimates in Table 2, which were calculated on the basis of means, do not reflect the effect of the bi-directional dominance.

Estimates for the number of genes that control each character associated with the disease resistance may not reflect the true parametric value, since the method for their estimation assumes absence of genetic linkage, equal loci effect and contrasting parents (Cruz and Regazzi, 2001). However, for characters where the estimates could be obtained, it may be observed that the gene set involved in the determination of the resistance of soybean to frogeye leaf spot has more than three genes.

The estimates and the significance of the null-hypothesis for each parameter in the complete model for the studied traits, based on transformed and not transformed data, are shown in Table 3. The non-allelic interactions were significant for ID, NLF, LMD and ID traits, but were not for PLFA, at 5% level of probability. Therefore, the PLFA variable could be analyzed through the simple (additive-dominant) model, for the original or for the transformed data.

The genetic effect associated to the dominance presented the greatest estimate (with negative sign towards resistance) for ID, NLF and DI and with the largest variance for all traits.

Data transformations did not modify the tests of

Table 2 – Estimates of phenotypic, environmental, genotypic, additive and dominance variances, broad and narrow sense heritabilities, average degree of dominance and number of genes, for traits evaluated on soybean resistance to *C. sojina* Hara trial, obtained from the parents and their derived populations.

Parameter	Character [3/]														
	ID	Square root (ID+1)	Log (ID+1)	NLF	Square root (NLF+1)	Log (NLF+1)	LMD	Square root (LMD+1)	Log (LMD+1)	PLFA	Square root (PLFA+1)	Log PLFA+1)	DI	Square root (DI+1)	Log (DI+1)
Phenotypic var.	1.1585	0.0775	0.0163	820.68	6.2040	0.2469	0.0105	0.0021	0.0013	63.34	1.6355	0.1990	24.841	13.05	0.00018
Environmental var.	0.3153	0.0256	0.0062	655.31	2.8668	0.1114	0.0030	0.0006	0.0004	66.31	0.4619	0.0452	12.785	4.43	0.00005
Genotypic var.	0.8432	0.0518	0.0101	165.37	3.3372	0.1355	0.0075	0.0016	0.0010	23.86	1.1736	0.1538	12.057	8.62	0.00013
Additive var.	0.8432	0.0496	0.0073	1281.9	6.8814	0.00	0.00	0.00	0.00	0.00	0.00	0.0433	12.057	8.62	0.000001
Dominance var.	0.00	0.0022	0.0028	0.00	0.00	0.1355	0.0075	0.0015	0.0010	23.86	1.1736	0.1105	0.00	0.00	0.00012
Broad heritability	0.7278	0.67	0.62	0.20	0.53	0.54	0.7155	0.73	0.73	0.377	0.717	0.77	0.485	0.66	0.70
Narrow heritability	0.7278	0.64	0.45	0.20	0.53	0.00	0.00	0.00	0.00	0.00	0.00	0.22	0.485	0.66	0.49
Aver. d. dominance[1/]	-1.07	-1.14	-1.17	-1.12	-1.18	-1.24	-2.62	-2.71	-2.76	-1.21	-1.41	-1.64	-1.17	-1.26	-1.70
Aver. d. dominance[2/]	(-)	0.30	0.88	(-)	(-)	(-)	(-)	(-)	(-)	(-)	(-)	2.26	(-)	(-)	5.18
Number of genes[2/]	2.11	2.69	3.89	1.29	1.73	(-)	(-)	(-)	(-)	(-)	(-)	7.12	2.97	1.82	(-)

[1/] based on means; [2/] based on variances. [3/] ID: infection degree; NLF: number of lesions per foliole; LMD: lesion mean diameter; PLFA: percentage of lesioned foliar area and DI: disease index; (-) negative estimate.

significance of the parameters for all studied traits.

The coefficient of multiple, partial and relative determination (R^2), or relative contribution, highlights the importance of a particular genetic effect on the available variability of the studied trait, although the partitions presented in Table 4 are not orthogonal.

Thus, the most important genetic effect on the determination of the ID character was the additive genetic effect (68.35%), for the original and for the transformed data, whenever the mean effect is not considered, while the effect due to dominance was of lesser importance (6.01%). This finding evidences

Table 3. Significance of genetic parameters [1] of the complete model, adjusted to the ID, NLF, LMD, DI and PLFA means of the six generations (P_1, P_2, F_1, F_2, BC_1, BC_2) included in the trial for soybean reaction to C sojina Hara.

				Character [3]					
Parameter		ID			Square root (ID+1)			Log (ID+1)	
	Estimate	Variance	t	Estimate	Variance	t	Estimate	Variance	t
m	4.71	0.2445	$9.53^{2/}$	2.44	0.0173	$18.57^{2/}$	0.8077	0.0039	$13.00^{2/}$
a	1.26	0.0057	$16.62^{2/}$	0.31	0.0004	$15.44^{2/}$	0.1332	0.0001	$14.08^{2/}$
d	-6.61	1.4804	$-5.43^{2/}$	-1.78	0.1076	$-5.41^{2/}$	-0.8520	0.0247	$-5.42^{2/}$
aa	-1.37	0.2388	$-2.81^{2/}$	-0.39	0.0169	$-2.97^{2/}$	-0.1941	0.0038	$-3.16^{2/}$
ad	-0.33	0.1146	-0.96^{ns}	-0.01	0.0086	-0.10ns	0.0307	0.0021	0.68ns
dd	3.83	0.5967	$4.95^{2/}$	1.04	0.0445	$4.92^{2/}$	0.5022	0.0105	$4.91^{2/}$
Parameter		NLF			Square root (NLF+1)			Log (NLF+1)	
	Estimate	Variance	t	Estimate	Variance	t	Estimate	Variance	t
m	135.17	148.62	$11.08^{2/}$	13.57	1.2306	$12.24^{2/}$	2.98	0.0688	$11.35^{2/}$
a	59.37	20.35	$13.16^{2/}$	3.56	0.0718	$13.28^{2/}$	0.44	0.0019	$10.21^{2/}$
d	-312.57	822.15	$-10.90^{2/}$	-25.08	7.2193	$-9.60^{2/}$	-5.18	0.1677	$-7.58^{2/}$
aa	-54.85	128.27	$-4.84^{2/}$	-5.40	1.1588	$-5.01^{2/}$	-1.29	0.0670	$-4.99^{2/}$
ad	-94.82	105.45	$-9.23^{2/}$	-3.89	0.6581	$-4.79^{2/}$	0.09	0.0431	0.42ns
dd	194.32	304.53	$11.14^{2/}$	16.21	2.8924	$9.53^{2/}$	3.34	0.2015	$7.44^{2/}$
Parameter		LMD			Square root (LMD+1)			Log (LMD+1)	
	Estimate	Variance	t	Estimate	Variance	t	Estimate	Variance	t
m	0.3181	0.0040	$5.029^{2/}$	1.15	0.0008	$40.40^{2/}$	0.13	0.00050	$5.74^{2/}$
a	0.0643	0.0001	$6.05^{2/}$	0.03	0.00002	$6.15^{2/}$	0.02	0.00001	$6.15^{2/}$
d	-0.1444	0.0294	-0.84^{ns}	-0.08	0.0060	-1.08^{ns}	-0.08	0.00370	-1.31^{ns}
aa	-0.0343	0.0039	-0.55^{ns}	-0.02	0.0008	-0.77^{ns}	-0.02	0.00050	-0.98^{ns}
ad	0.2337	0.0030	$4.26^{2/}$	-0.11	0.0006	$4.32^{2/}$	0.08	0.00040	$4.34^{2/}$
dd	-0.0583	0.0121	-0.53^{ns}	-0.01	0.0025	$-.029^{ns}$	-0.002	0.00150	-0.06^{ns}
Parameter		DI			Square root (DI+1)			Log (DI+1)	
	Estimate	Variance	t	Estimate	Variance	t	Estimate	Variance	t
m	13.70	5.0265	$6.11^{2/}$	17.41	2.5723	$10.86^{2/}$	0.0256	0.000050	$3.63^{2/}$
a	7.68	0.4946	$10.91^{2/}$	5.02	0.1228	$14.34^{2/}$	0.0124	0.000002	$8.53^{2/}$
d	-29.44	29.4685	$-5.42^{2/}$	-32.56	14.9747	$-8.41^{2/}$	-0.0242	0.000330	-1.34ns
aa	-4.49	4.5319	$-2.11^{2/}$	-6.65	2.4494	$-4.25^{2/}$	-0.0018	0.000047	-0.25ns
ad	-9.77	3.3559	$-5.33^{2/}$	-4.81	1.2843	$-4.25^{2/}$	0.0072	0.000032	1.26ns
dd	15.99	10.6503	$4.90^{2/}$	19.60	5.7457	$8.18^{2/}$	0.0014	0.000128	0.12ns
Parameter		PLFA			Square root (PLFA+1)			Log (PLFA+1)	
	Estimate	Variance	t	Estimate	Variance	t	Estimate	Variance	t
m	4.18	25.16	0.83^{ns}	2.0625	0.4652	$3.02^{2/}$	0.50	0.0507	$2.24^{2/}$
a	10.21	1.53	$8.27^{2/}$	1.4707	0.0178	$11.02^{2/}$	0.44	0.0016	$10.89^{2/}$
d	5.65	186.24	0.41^{ns}	0.6353	3.1491	0.36ns	0.22	0.3300	0.38^{ns}
aa	8.69	23.63	1.79^{ns}	1.2561	0.4474	1.87ns	0.40	0.0490	1.80^{ns}
ad	0.38	21.70	0.08^{ns}	0.3244	0.3109	0.58ns	0.24	0.0303	1.39^{ns}
dd	-9.27	76.54	-1.06^{ns}	-1.4574	1.2412	-1.31ns	-0.54	0.12285	-1.52^{ns}

[1] m: mean of homozygous lines derived from F_2 without selection; a: additive genic effect; d: dominance, aa: additive x additive interaction measure; ad: additive x dominance interaction and dd: dominance x dominance interaction; [2] Significant at 5%; ns: non-significant; [3] ID: infection degree; NLF: number of lesions per foliole; LMD: lesion mean diameter; PLFA: percentage of lesioned foliar area and DI: disease index.

Table 4. Non-orthogonal partition of sum of squares by Gauss elimination method for m, a, d, aa, ad, dd parameters of the ID, NLF, LMD, PLFA and DI traits evaluated on the parents and their derived populations in the trial for soybean resistance reaction to *C. sojina* Hara.

Source of variation[1]	Character[2]					
	ID		Square root (ID+1)		Log (ID+1)	
	SS	"R²"(%)	SS	'R²'(%)	SS	'R²'(%)
m/a, d, aa, ad, dd	89.81	18.74	343.40	50.15	167.94	36.39
a/m, d, aa, ad, dd	327.48	68.35	280.35	40.94	231.05	50.06
d/m, a, aa, ad, dd	28.78	6.01	28.61	4.18	28.76	6.23
aa/m, a, d, ad, dd	7.87	1.64	8.83	1.29	10.00	2.17
aa/m, a, d, aa, dd	1.34	0.28	0.08	0.01	0.27	0.06
dd/m, a, d, aa, ad	23.86	4.98	23.49	3.43	23.51	5.09
Sum	479.15	100	684.76	100	461.52	100

Source of variation[1]	NLF		Square root (NLF+1)		Log (NLF+1)	
	SS	'R²'(%)	SS	'R²'(%)	SS	'R²'(%)
m/a, d, aa, ad, dd	122.94	18.98	148.47	24.60	127.57	33.20
a/m, d, aa, ad, dd	173.21	26.74	221.68	36.73	120.79	31.43
d/m, a, aa, ad, dd	118.83	18.35	90.98	15.08	56.52	14.71
aa/m, a, d, ad, dd	23.45	3.62	25.14	4.17	24.86	6.47
ad/m, a, d, aa, dd	85.28	13.17	27.60	4.57	0.07	0.02
dd/m, a, d, aa, ad	123.99	19.14	89.61	14.85	54.46	14.17
Sum	647.71	100	603.48	100	384.28	100

Source of variation[1]	LMD		Square root (LMD+1)		Log (LMD+1)	
	SS	'R²'(%)	SS	'R²'(%)	SS (x10⁻⁴)	'R²'(%)
m/a, d, aa, ad, dd	25.24	31.07	1624.00	96.65	32.67	36.34
a/m, d, aa, ad, dd	36.59	45.03	36.27	2.16	36.18	40.25
d/m, a, aa, ad, dd	0.71	0.87	1.14	0.07	1.67	1.86
aa/m, a, d, ad, dd	0.30	0.37	0.58	0.03	0.96	1.07
ad/m, a, d, aa, dd	18.13	22.31	18.23	1.08	18.40	20.47
dd/m, a, d, aa, ad	0.28	0.35	0.09	0.01	0.01	0.01
Sum	81.25	100	1680.30	100	89.88	100

Source of variation[1]	PLFA		Square root (PLFA+1)		Log (PLFA+1)	
	SS	'R²'(%)	SS	'R²'(%)	SS	'R²'(%)
m/a, d, aa, ad, dd	0.69	0.94	8.85	6.00	4.77	3.30
a/m, d, aa, ad, dd	68.32	92.94	132.80	90.07	132.01	91.48
d/m, a, aa, ad, dd	0.17	0.23	0.17	0.12	0.21	0.14
aa/m, a, d, ad, dd	3.20	4.35	3.53	2.39	3.23	2.24
ad/m, a, d, aa, dd	0.01	0.01	0.21	0.14	1.55	1.07
dd/m, a, d, aa, ad	1.12	1.53	1.88	1.28	2.54	1.76
Sum	73.52	100	147.45	100	144.30	100

Source of variation[1]	DI		Square root (DI+1)		Log (DI+1)	
	SS	'R²'(%)	SS	'R²'(%)	SS (x10⁻⁴)	'R²'(%)
m/a, d, aa, ad, dd	37.33	15.38	116.63	21.50	0.2700	7.77
a/m, d, aa, ad, dd	119.11	49.06	251.22	46.30	3.0700	89.47
d/m, a, aa, ad, dd	29.41	12.11	69.61	12.83	0.0400	1.21
aa/m, a, d, ad, dd	4.45	1.83	18.07	3.33	0.00001	0.04
ad/m, a, d, aa, dd	28.47	11.73	21.38	3.94	0.0500	1.50
dd/m, a, d, aa, ad	24.00	9.89	65.63	12.10	0.00001	0.01
Sum	242.77	100	542.54	100	3.4400	100

[1] m: mean of homozygous lines derived from F_2 without selection; a: additive genic effect; d: dominance deviation, aa: additive x additive interaction; ad: additive x dominance interaction and dd: dominance x dominance interaction; [2] ID: infection degree; NLF: number of lesions per foliole; LMD: lesion mean diameter; PLFA: percentage of lesioned foliar area and DI: disease index.

the possibility of obtaining superior homozygotic materials, by means of selection starting from the F_2 generation, and satisfactory gains in the selection cycles, since the component of additive nature and the heritability are of high magnitude. Table 4 also shows that, for the ID character, the epistatic additive x additive, additive x dominance and dominance x dominance interactions influenced the means of the studied generations in 6.90% of the cases. Therefore, the results suggested that the additive-dominant model can explain most of the variations in this trait. However, in spite of the low contribution of the epistatic interaction effects, the use of this reduced model is subject to a bias, since the aa and dd epistasis were significant and should not be disregarded. This same reasoning can be extended to the other traits.

The transformations influenced magnitude, but they did not modify the order of importance of the estimated genetic effects for all the evaluated traits. Therefore, the conclusions are the same for data with or without the transformations.

Similar results were obtained by Otsuk et al. (1991), in the estimation of variance components working with simulated Binomial or Poison distribution data. They found small differences between the components of variance estimated from the original data or from data submitted to square root, logarithmic or arcsine transformations. They concluded that, in genetic breeding programs, the estimates of variance components from non-transformed data can be used.

According to the literature review (Arias, 1986), the dominance effect and genetic interactions are of lesser importance to autogamous species. However, genetic models that neglect epistasis can sometimes be biased (Quinby, 1963, according to Scapim, 1997).

CONCLUSIONS

Data transformations neither modify the tests of significance of the estimates of genetic parameters studied nor the order of importance of the genetic effects.

The most important genetic effect in the determination of traits related to resistance is the additive genetic effect.

The infection degree (ID) can be used to select for resistance to *C. sojina* in soybean breeding programs, because, although subjective, it considers the set of all characteristics in its attribution.

Original, square root or log transformed ID, NLF, LMD and ID data presented significant epistatic

interactions. Therefore, they should be evaluated using the additive-dominant-epistatic complete model. The PLFA character can be analyzed using the simpler additive-dominance model, since it presented non-significant epistatic interactions.

RESUMO

Efeito da transformação de dados nas estimativas dos parâmetros genéticos da resistência à *Cercospora sojina* Hara em soja

O objetivo deste trabalho foi estudar os efeitos da transformação de dados nas estimativas dos parâmetros genéticos da herança da resistência à *Cercospora sojina* Hara em soja. As médias e variâncias de cinco características associadas à doença, das populações (P_1, P_2, F_1, F_2, RC_1 e RC_2), derivadas do cruzamento entre um cultivar resistente (Paraná) e um suscetível (Bossier), foram avaliadas: grau de infecção (ID), avaliado visualmente; número de lesões por folíolo (NLF); diâmetro médio da lesão (LMD); percentagem de área foliar lesionada (PLFA) e índice de doença (DI). As transformações estudadas foram: raiz quadrada de (x), (x +1), (x + 0,5) e (x + 3/8) e logaritmo de (x +1). Os resultados mostraram que as transformações não alteraram a significância dos parâmetros genéticos estimados. O efeito genético aditivo foi o mais importante, em todos os caracteres avaliados. O modelo aditivo-dominante foi suficiente para explicar as variações somente no caráter PLFA; para os demais caracteres, as epistasias foram significativas. A influência dos efeitos das interações epistáticas variaram de 5,89% no caráter PLFA até 35,93% no caráter NLF. Para os dados com ou sem transformação, o modelo genético aditivo-dominante foi satisfatório para explicar o comportamento da média das gerações do caráter PLFA. Entretanto, para ID, NLF, LMD e DI o modelo aditivo-dominante-epistático foi o mais apropriado. Os caracteres ID, NLF e DI, que apresentaram variabilidade genética aditiva maior que da dominância, poderão ser utilizados pelos melhoristas, sem transformação dos dados, como critério de seleção para obter genótipos homozigotos superiores, a partir da geração F_2.

REFERENCES

Arias, E.R.A. 1986. Teste de escala conjunta na estimação de parâmetros genéticos em soja (*Glycine max* (L.) Merrill). M.S. Thesis. Universidade Federal de Viçosa, Viçosa.

Bartlett, M. S. 1936. The square root-transformation in the analysis of variance. Journal of the Royal Statistical Society (Supplement). 3:68-78.

Cordeiro, A.C.C.; Sediyama, T.; Gomes, J.L.P.; Sediyama, C.S. and Reis, M.S. 1992. Herança da resistência da soja à *Cercospora sojina* Hara, isolado de São Gotardo, Minas Gerais. Pesquisa Agropecuária Brasileira. 27:1035-1042.

Cruz, C.D. 2001. Programa Genes: versão Windows: aplicativo computacional em genética e estatística. Ed. UFV. Viçosa.

Cruz, C.D. and Regazzi, A.J. 2001. Modelos biométricos aplicados ao melhoramento genético. 2.ed. Ed. UFV, Viçosa.

Demétrio, C.G.B. 1978. Transformação de dados: efeitos sobre a análise da variância. Piracicaba: M.S. Thesis. Escola Superior de Agricultura "Luiz de Queiroz", USP, São Paulo.

Gamble, E.E. 1962. Gene effects in corn (*Zea mays* L.*)* I. Separation and relative importance of gene effects for yield. Canadian Journal of Plant Science. 42:339-348.

Martins Fillho, S. 1999. Mancha olho de rã da soja: análise genética e identificação de marcadores moleculares. Ph.D. Diss. Universidade federal de Viçosa. Viçosa.

Mather, F.R.S. and Jinks, F.R.S. 1984. Introdução à genética biométrica. Sociedade Brasileira de Genética, Ribeirão Preto.

Otsuk, I.P.; Barbin, D. and Corrente, J.E. 1991. Influência de transformações de dados sobre a estimação dos componentes de variância. Boletim da Indústria Animal. 48:129-139.

Scapim, C.A. 1997. Efeitos gênicos, heterose e depressão endogâmica em caracteres de sorgo granífero e forrageiro (*Sorghum bicolor* L. Moench). Ph.D. Diss. Universidade Federal de Viçosa, Viçosa.

Veiga, P. 1973. *Cercospora sojina* Hara: obtenção de inóculo, inoculação e avaliação da resistência em soja (*Glycine max* (L.) Merrill). M.S.Thesis. Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba.

Yorinori, J.T. and Klingelfuss, L.H. 1999. Novas raças de *Cercospora sojina* em soja. Fitopatologia Brasileira. 24:509-512.

Yorinori, J.T. 2001. Mancha "olho de rã" (*Cercospora sojina*): Cs-15: reação à raça Cs-15, patogênica ao gene de resistência da cultivar Santa Rosa; Cs-23: reação à raça Cs-23;Cs-24 e Cs-25: novas raças presentes no MA e no PI, e Mist.: reação de seis raças de *C. sojina* mais prevalentes no Brasil. Resultados de Pesquisa de Soja. Embrapa Soja, Londrina.