Genetic variability and selection of soybean genotypes under shading[1]

Anderson Barbosa Evaristo[2], Amanda Gonçalves Guimarães[3], João Paulo Ribeiro Leite[4], Reimario de Castro Rodrigues[5], Leonardo Barros Dobbss[6], Eder Matsuo[7], José Barbosa dos Santos[8]

ABSTRACT - Shade-tolerant soybean genotypes will allow their greater adoption for use in agroforestry or intercropped systems in tropical and subtropical regions. These genotypes can be obtained from the selection of existing superior genotypes or through breeding programs. This study, was verified such genetic variability and select those genotypes with cultivation potential at reduced levels of photosynthetically active radiation (PAR). Morphophysiological and yield characteristics of 16 soybean genotypes were evaluated in two experiments conducted in a controlled environment and laid in a randomized block design with three replications. Genetic parameters, correlations and genetic diversity were estimated and path analysis was performed. The soybean cultivars used showed genetic variability with a high variation index and heritability, which is advantageous for the selection of superior cultivars. Leaf area showed genotypic correlation and had a direct effect on seed weight per plant, regardless of shading. Additionally, the number of pods, hypocotyl diameter, and leaf area, all indirectly contributed to seed weight per plant regardless of shading. Seven genotypes (NS7780, NS8338, NS7667, NS7901, RK8115, CZ37B43 and 8579RSF) were selected based on genetic gain obtained by the Mulamba and Mock index at the two levels of PAR reduction tested. Furthermore, five of these genotypes (CZ37B43, RK8115, NS7901, NS8338 and 8579RSF) were identified by the Unweighted Pair-Group method Arithmetic Mean, as suitable either as a shaded crop or, as parental materials in breeding programs.

Key words: *Glycine max*. Genetic parameters. Photosynthetically active radiation. Genetic gain.

[2]Institute Agricultural Sciences, Federal University of Jequitinhonha and Mucuri Valeys, Unaí-MG, Brazil, anderson.evaristo@ufvjm.edu.br (ORICD ID: 0000-0003-1236-0196)
[3]Rural Science Center, Special Coordination of Biological and Agronomic Sciences, Federal University of Santa Catarina, Curitibanos-SC, Brazil, amanda.goncalves@ufsc.br (ORICD ID: 0000-0003-0761-8287)
[4]KWS Seeds, Patos de Minas -MG, Brazil, joaopaulosaga1@gmail.com (ORICD ID: 0000-0002-8155-152X)
[5]Postgraduate Program in Plant Production, Federal University of Jequitinhonha and Mucuri Valeys (UFVJM), Diamantina-MG, Brazil, reimario.c.rodrigues@gmail.com (ORICD ID: 0000-0003-0829-8617)
[6]Institute Agricultural Sciences, Federal University of Jequitinhonha and Mucuri Valeys, Unaí-MG, Brasil, leonardo.dobbss@ufvjm.edu.br (ORICD ID: 0000-0003-0465-4186)
[7]Institute of Exact and Technological Sciences, Federal University of Viçosa (UFV), Rio Paranaíba-MG, Brasil, edermatsuo@ufv.br (ORICD ID: 0000-0002-2643-9367)
[8]Department of Agronomy, Federal University of Jequitinhonha and Mucuri Valeys (UFVJM), Diamantina-MG, Brasil, jbarbosa@ufvjm.edu.br (ORICD ID: 0000-0002-5746-7248)

INTRODUCTION

Soybean (*Glycine max* L.) is an important commercial crop worldwide, with such high grain protein and oil levels, that it is extensively used in the food, animal feed, bioenergy, and chemical industries. Brazil is currently the largest producer and exporter of soybean globally (FAO 2021). Indeed, the expansion of soybean cultivation has begun to occupy pasture areas, with great intensification of agricultural production as alley cropping, one important common agroforestry practices.

However, the viability of integrated systems depends on factors such as soil water and nutrient availability for the different species growing together, the level of understory shading, and the degree of competition among plants for resources. Although soybean expresses phenotypic plasticity during the vegetative and reproductive growth stages (CARPENTER; BOARD, 1997), these factors can influence the pattern of crop growth and development. As a case in point, shading can induce physiological and/or morphoagronomic changes, thus underlining the utmost importance of searching for high-yielding, shade-tolerant soybean cultivars (FENG *et al.*, 2019; TIBOLLA *et al.*, 2019; YANG *et al.*, 2014). Soybean grown in the understory of agroforestry or intercropping systems, such as the soybean-corn system, may exhibit low productivity (CRISTO *et al.*, 2020; RAZA *et al.*, 2019; WERNER *et al.*, 2017; WU *et al.*, 2016).

Current soybean varieties were developed by improvement programs focused on monoculture in restriction-free environments. On the other hand, light restriction can alter cultivar performance significantly, whereby, its effects must be evaluated to avoid yield losses. However, the phenotype of superior cultivars depends on important quantitative traits that may be strongly influenced by the environment. The measuring genetic parameters at different levels of restriction of photosynthetically active radiation (PAR) is important for the selection of superior cultivars, as successful selection will depend on the quality of data, the extent of environmental influence, and the degree of genetic variability within the collection of available soybean genotypes (CRUZ; REGAZZI; CARNEIRO, 2014; MONTEIRO *et al.*, 2021). Furthermore, the process of cultivar selection may take advantage of associated traits. Therefore, it is necessary to determine any existing associations and to estimate the extent of association to accelerate progress by indirect selection of a desirable trait through direct selection of another desirable trait or traits (BISINOTTO *et al.*, 2017; FERRARI *et al.*, 2018; SOUSA *et al.*, 2015). When several traits are correlated with each other, it is still necessary to use path analysis, which is a method that seeks to analyze these direct and indirect effects between traits under a main variable (MACHADO *et al.*, 2017; TEODORO *et al.*, 2015).

For the selection of soybean genotypes at different levels of PAR reduction, a selection index is adopted to allow genetic gain for the traits of interest through the selection for other traits prioritized by the breeder in the process (SOARES *et al.*, 2015). Further, the estimation of genetic divergence enables the selection of genotypes through measures of dissimilarity and graph grouping of similar and divergent cultivars (CRUZ; CARNEIRO; REGAZZI, 2014).

The hypothesized is that there is ample genetic variability among available soybean genotypes for agronomic performance in shaded environments. Therefore, was evaluated genetic variability among various soybean genotypes at different levels of PAR reduction and estimated genetic parameters, genetic correlations, path analysis, genetic gains, and genetic diversity.

MATERIAL AND METHODS

Study site, plant materials and growth conditions

This study was conducted in Unaí-MG, Brazil (16°26'10.48" S, 46°54'2.28" W, at 634 m above sea level), Savanna biome, between October and February 2020. Sixteen commercial soybean cultivars (supplementary material) of different maturity from the Cerrado region (POEHLMAN, 1987) were tested in two experiments, each with a different level of light restriction, namely, experiments I and II, with plants subjected to 25% and 48% PAR reduction, respectively. To provide the desired photosynthetically active radiation (PAR) reduction level, in experiment I, plants were grown in a greenhouse under black shade nets that allowed passage of 18% of the incident light, thus providing 25% PAR reduction. Meanwhile, in experiment II, plants were grown in a greenhouse under black shade nets that allowed 30% of the light to pass through, thus providing 48% PAR reduction. To determine PAR reduction, photosynthetic photon flux density (PPFD) was measured throughout the day using a PAR meter (Apogee quantum meters, model MQ-200) during the entire duration of the experiment. Average PAR reduction values were determined and compared with measurements made under full light conditions (i.e., no shade). Soybean plants were grown in 10 L plastic pots. The substrate was composed of 50% soil (69.5% clay content) and 50% sand mixed with 36 g of dolomitic limestone, 8 g of simple superphosphate, and 1.60 g of potassium chloride per pot. Soybean seeds were treated with a commercial mixture of pyraclostrobin, thiophanate-methyl, and fipronil (Standak®Top, 2 mL/kg) and inoculated with *Bradyrhizobium japonicum* at the time of planting. Five seeds were sown per pot and thinning to one plant per pot was performed after seedling emergence.

Traits evaluated

Growth characteristics: Leaf area (LA, cm²) was determined with a leaf area meter (LI-COR, model LI-3100), using the central leaflet of the third leaf from the apex, in full bloom at 62 days after sowing (DAS). Plant height (PH, cm) measured from the ground to the plant apex, and hypocotyl diameter (D, mm) were determined at harvest, when plants were in the R8 phenological stage, as per the Fehr and Caviness scale (1977).

Chlorophyll: Leaf discs were cut with a cork borer from the central leaflet of the third leaf from the apex to determine chlorophyll a content (Chl a), at 62 DAS. The discs were placed in dimethyl sulfoxide in the dark for 12 h in a water bath at 60 °C to extract chlorophyll pigments. Chl a content was determined according to the method described by Wellburn (1994).

Yield components: At phenological stage R8, i.e., full maturity, plants were harvested and the number of branches (NB) containing pods, the number of pods per plant (NPP), and the number of seeds per pod (NSP) from a sample of 30 pods per plant, were measured. Hundred seed weight (HSW, g) was determined by counting three subsamples of grain. Seed weight per plant (SWP, g) was obtained from the total grain mass of each plant. Grain moisture content was determined using a digital moisture meter and the grain mass was quantified in an analytical balance. HSW and SWP data were adjusted to 13% grain-moisture content.

Experimental design and statistical analysis

The two experiments reported herein were laid in a randomized block design with 16 treatments (soybean genotypes) and three replications, one plant per plot, according to the following statistical model: $Y_{ij} = \mu + g_i + b_j + e_{ij}$ where Y_{ij} = effect of the ith soybean genotype on the jth block; μ = mean; b_j = j-th block effect (j=1, 2, 3); g_i = effect of the ith soybean genotype (i= 1, 2,...16); e_{ij} = effect of the random experimental error associated with Y_{ij}, supposedly independent and normally distributed (NID), with mean zero and constant variance, considering $e_{ij} \sim$ NID (0, σ^2). All measured traits were statistically analyzed by calculating position measurements (10th, 25th, 75th and 90th), average, and median; standard error is represented by boxplot graphs. The following genetic parameters were determined at each level of PAR reduction: (i) coefficient of experimental variation $(CVe): CVe(\%) = 100 \left(\frac{\sqrt{\sigma_e^2}}{\mu} \right)$, where σ_e^2 = residual variance; μ = mean. (ii) coefficient of genetic variation $(CVg): CVg(\%) = 100 \left(\frac{\sqrt{\sigma_g^2}}{\mu} \right)$, where σ_g^2 = genotypic variance among soybean genotypes. (iii) variation index (Iv) $Iv = \left(\frac{CVg}{CVe} \right)$. (iv) broad-sense heritability (h_a^2) $h_a^2 = \frac{\sigma_g^2}{\sigma_f^2}$ where σ_g^2 = genotypic variance among soybean genotypes; σ_f^2 = phenotypic variance among soybean genotypes.

For phenotypic (rF) and genotypic correlations (rG) at each level of PAR reduction, the t-test was used test to determine significance. These parameters were estimated according to the following expressions: $rF = \frac{COVF(X,Y)}{\sqrt{\sigma_{FX}^2 . \sigma_{FY}^2}}$, $rG = \frac{COVG(X,Y)}{\sqrt{\sigma_{GX}^2 . \sigma_{GY}^2}}$ where $COVF(X,Y)$ and $COVG(X,Y)$ correspond to the estimates of phenotypic and genotypic covariances between traits represented by x and y, respectively; σ_{FX}^2 and σ_{GX}^2, correspond to the estimates of phenotypic and genotypic variances for trait x, respectively, while σ_{FY}^2 and σ_{GY}^2 correspond to phenotypic and genotypic variances of trait y, respectively.

Before the path analysis was carried out, the degree of multicollinearity of the X′X matrix was estimated based on its number of conditions (NC), which is the ratio between the highest and lowest eigenvalues of the X′X correlation matrix (MONTGOMERY; PECK; VINING, 2012). This criterion considers that multicollinearity will be weak only among the explanatory variables when the relationship between the highest and the lowest values is equal to or below 100. When the value resulting from this division is 100 < NC < 1,000, multicollinearity is considered moderate to severe, and when it is NC ≥ 1,000 multicollinearity is considered severe. The correlation matrix at the level of 25 and 48% PAR the NC was 72.46 and 186.70 considering low to moderate severity respectively. To overcome this multicollinearity, the method proposed by Carvalho and Cruz (1996) was adopted, which consists of applying a constant k to the diagonal of the matrix X′X of the least squares estimator. It was chosen to make the constant at the level of 25% of PAR for the NC is as small as possible. The value of k applied at the level of 25% PAR was 0.49, making the NC to decrease to 22.28 and at the level of 48% PAR the value of k applied was 0.12 with the NC to decrease to 11.84.

For the path analysis, the resolution in the form of a matrix was obtained according to the equation: X′Xβ = X′Y, where X′X is a nonsingular matrix of the correlations between the explanatory variables, β is the column vector of the path coefficients, and X′Y is the column vector of the correlations between the explanatory variables and the main variable. Then, path analysis was performed using the genotypic correlation between the dependent variable SWP and the explanatory variables (i.e., the other measured characteristics). Path analysis was performed according to the following equation: $r_{ix} = P_{ix} + \sum_{j \neq i}^{n} r_j P_{jx}$ where: rix is the correlation between the dependent variable and the ith explanatory variable, Pix is the direct effect of variable i on the dependent variable, $rjPjx$ is the indirect effect of variable i on the dependent variable, through variable j.

The selection index of Mulamba and Mock (1978) was used to select superior soybean genotypes at each level of PAR reduction. Genetic gain was estimated based on the ranking of genotypic means for each trait on a scale from least to most favorable for breeding, lower chlorophyll a content and other superior trait. The choice of direction was based on the direction of correlations between seed weight per plant and the other traits. After the classification, cultivars were analyzed for economic weight, i.e., the coefficient of genetic variation (CVg) of each trait, and the rankings of each trait were added, such that an additional average was obtained, which in turn was considered as a selection index: $I = p_1 r_1 + p_2 r_2 + \cdots + p_n r_n$ where $I =$ index value for a given genotype, p_j represents the economic weight (CVg) for the jth trait; $rj =$ Classification (rank) of a genotype in relation to the jth trait; $n =$ number of traits considered in the index. The nine genotypes showing the lowest total I values were selected for each level of shading.

For the dissimilarity between soybean genotypes, the average values for each trait of each soybean genotype at each level of PAR reduction were used, and the matrices of genetic distances were extracted through the generalized Mahalanobis distance (MAHALANOBIS, 1936), expressed as follows: $D_{ii'}^2 = \delta'_{ii' \varphi^{-1} \delta_{ii'}}$, where $D_{ii'}^2$ is the Mahalanobis generalized distance between accessions i and i', and i = 1, 2, ..., 16; $\delta'_{ii'} = [d_1 d_2 \ldots d_v]$, and $d_j = Y_{ij} - Y_{i'j}$; $Yij =$ is the mean of the ith accession in relation to the jth variable, where j = 1, 2, ..., p; $\varphi^{-1} =$ the inverse of the matrix of residual variances and covariances.

After estimating the genetic distance matrices, clustering was performed using the hierarchical Unweighted Pair-Group Method Average (UPGMA). The cutoff point for determining the number of groups at each level of PAR reduction was defined using the expression proposed by Mojena (1977). The consistency of the clustering methods was assessed by cophenetic correlation coefficients (CCC), where the significance of CCCs was examined using the Mantel test (MANTEL, 1967). The Singh (1981) criterion was also used to quantify the relative contribution of these characteristics to genetic divergence. Analysis of variance, estimates of genetic parameters, path analysis, genetic gains, and dissimilarity were all performed using the Genes statistical program (CRUZ, 2016). Correlation analyses were performed using the corrplot package and box plot graphs were elaborated the ggplot2 package in the statistical program R (R CORE TEAM, 2020).

RESULTS AND DISCUSSION

Analysis of variance, genetic parameters and phenotypic variation under both levels of PAR reduction

Analysis of variance (ANOVA) and the estimation of genetic parameters at each level of PAR reduction are shown in Table 1. Significant ($p \leq 0.05$) differences were observed among the soybean genotypes (G) tested for all the traits under study in both shaded environments. These results are essential for the successful selection of superior cultivars and to obtain greater gains exploiting the existing genetic variability (GUIMARÃES *et al.*, 2018; MONTEIRO *et al.*, 2021).

Table 1 - Analysis of variance (mean squares) and estimation of genetic parameters of soybean cultivars under two levels of photosynthetically active radiation (PAR) reduction. Nine morphophysiological and yield characteristics are presented

Source of variation		[1]Chl a	LA	PH	D	NB	NPP	NSP	HSW	SWP
Genotypes	25% PAR reduction	6.49**	76.87*	367.75**	17.24**	3.72*	13378.59**	0.11**	17.53**	2709.29**
	48% PAR reduction	10.34*	205.05**	178.40**	10.01**	6.61**	14603.54**	0.18**	24.06**	2279.44**
[2] Genetic parameters		25% PAR reduction								
CVg (%)		12.32	8.57	19.31	15.58	11.82	33.02	6.54	10.73	37.95
CVe (%)		12.96	13.82	13.32	12.36	19.06	17.56	7.34	10.83	18.07
Iv		0.95	0.62	1.45	1.26	0.62	1.88	0.89	0.99	2.10
h_a^2 (%)		73.07	53.37	86.32	82.71	53.53	91.39	70.54	74.69	92.94
		48% PAR reduction								
CVg (%)		12.31	13.49	11.74	13.77	16.33	35.42	9.10	14.32	36.63
CVe (%)		20.52	12.85	6.95	12.87	17.19	15.95	7.28	6.39	18.69
Iv		0.60	1.05	1.69	1.07	0.95	2.22	1.25	2.24	1.96
h_a^2 (%)		52.29	76.93	89.56	77.55	73.12	93.68	82.35	93.76	92.00

[1]Chl a = chlorophyll a; LA = leaf area (cm²); PH = plant height (cm); D = hypocotyl diameter (mm); NB = number of branches; NPP = number of pod per plant; NSP = number of seed per pod; HSW = hundred seed weight (g); SWP = seed weight per plant (g). [2]CVg = coefficient of genetic variation (%); CVe = coefficient of experimental variation; Iv = variation index; h_a^2 = heritability in the broad sense. * and **, significant at p ≤ 0.05 e p ≤ 0.01, respectively, by the F test

Estimates of the coefficients of experimental variation (CVe) for the genetic parameters under study revealed a good level of precision (below 20%), indicating environmental effects on all evaluated characteristics (GOMES, 2000), except for Chl a under 48% PAR reduction, which showed a precision of 20.52%. In another study on soybean, Carvalho et al. (2003) reported a 12% upper limit for the coefficient of variation for PH and 16% for SWP. However, in this study, the observed estimate of CVe for PH at 25% PAR reduction was close to the suggested limit (13.32%) and at 48% shading, the observed value was below the suggested value (6.94%). Genotype, location, and shading level are likely the factors that might have contributed to these differences relative to our results.

Estimates of the coefficient of genetic variation (CVg) were greater than those of CVe for PH, D, NPP, NSP, and seed weight per plant, at both levels of shading. LA and HSW also showed CVg values greater than those of CVe at 48% PAR reduction, suggesting that, for the aforementioned characteristics, genetic variance is higher than environmental variance (Iv > 1.0); that is, the magnitude of the observed genetic variability may directly influence the genetic gain for selection among cultivars.

Estimated broad-sense heritability (h_a^2) was above 50% for all the studied traits, at both levels of shading, and ranged from medium to high magnitude. This estimate was applied to the genotypes under study. Being specific to the environment in which they were studied, h_a^2 determines the confidence in using the phenotypic value to estimate the genotypic variance (RAMALHO et al., 2012), since phenotype is shaped by both genetic and environmental factors (FALCONER; MACKAY, 1996). It is noteworthy that these characteristics had varying values of h_a^2, ranging from 53.37% (LA) to 92.94% (SWP) at the 25% level of PAR reduction, and from 52.29% (Chl a) to 93.76% (HSW) at the 48% level of PAR reduction. High values of h_a^2 indicate that phenotypic variability is primarily affected by the genotype and not by the environment. For SWP at both shading levels, h_a^2 was above 90%, and Iv values were close to 2; thereby, confirming that a high level of environmental control is likely to improve the efficiency in the selection of superior soybean cultivars.

The values shown in the boxplot indicate the presence of phenotypic variation for morphophysiological and yield characteristics at different levels of PAR reduction (Fig. 1). Plants grown in the environment with greater (48%) light restriction showed higher mean values for PH, LA, and NB, and smaller values for D. Tolerance and avoidance are two strategies that plants use to reduce light stress. To avoid the effects of shade, plants often invest in increasing leaf Chl a content, PH, and LA, while reducing D. These changes suggest that plants escape shade as a response to light stress (TAIZ et al., 2017). Several studies have shown that increasing shading of soybean plants, whether by other plants, as in intercropping systems, or under artificial conditions of light restriction, promotes plant etiolation and increases the content of chloroplast pigments (FAN et al., 2019; GONG et al., 2015; YANG et al., 2014).

Soybean yield components including NPP, NSP, HSW, and SWP showed a similar behavior under both PAR reduced environments. However, large phenotypic variability for these parameters was indicated by the amplitude of the standard error bars. Such phenotypic variation for these traits is due to the genetic variation among cultivars within each shaded environment (Table 1), which allows for the selection of cultivars showing superior agronomic performance. Consistently with our results, Wu et al. (2017) evaluated 131 soybean genotypes grown in shaded environments and reported high phenotypic variability for NPP, NSP, HSM, and SWP.

Phenotypic and genotypic correlations and path analysis

Correlation analysis has been extensively used as a strategy in the selection of favorable genotypes indicating highlighting the indirect influence exerted by one trait on another according to magnitude (0 to 1), direction (positive or negative), and statistical significance of the correlation, without the need to evaluate both features, thereby, making it more efficient and faster than direct selection (CRUZ; REGAZZI; CARNEIRO, 2014).

Genotypic correlations were either higher than phenotypic correlations and equal sign for all the traits studied here (Fig. 2b and 2d), thus suggesting that the environment had less influence on the expression of these traits (ANDRADE JUNIOR et al., 2019). This result is very promising for breeding programs whose main objective is to perform indirect selection, as the desirable traits are genetically determined to a larger extent than environmentally, whereby the genetic fraction is enhanced regardless of shading level.

In turn, phenotypic correlations take into account both the genotype and the phenotype that is, genotypes can be physiologically and morphologically modified according to the conditions imposed by the cultivation environment (CRUZ; REGAZZI; CARNEIRO, 2014). In this study, we observed that significant phenotypic correlations were different for some pairs of traits, indicating the influence of shading on plant growth and development.

A genotypic correlation can be explained by the pleiotropy, that is, when a gene influences the expression of more than one trait, it can be inferred that when selecting one trait, the other is simultaneously selected (FALCONER; MACKAY, 1996).

Figure 1 - Box-plot: plots the average (line red), median, 25th and 75th percentiles with error bars. A) plant height (cm); B) hypocotyl diameter (cm); C) leaf area (cm²); D) chlorophyll a (mg/g leaf dry mass); E) number of branches; F) number of pod per plant; G) number seed per pod; H) hundred seed weight (g); and I) seed weight per plant (g). 25% and 48% photosynthetically active radiation (PAR) reduction. * significant at 5% probability by test F anova. ns not significant by test F anova
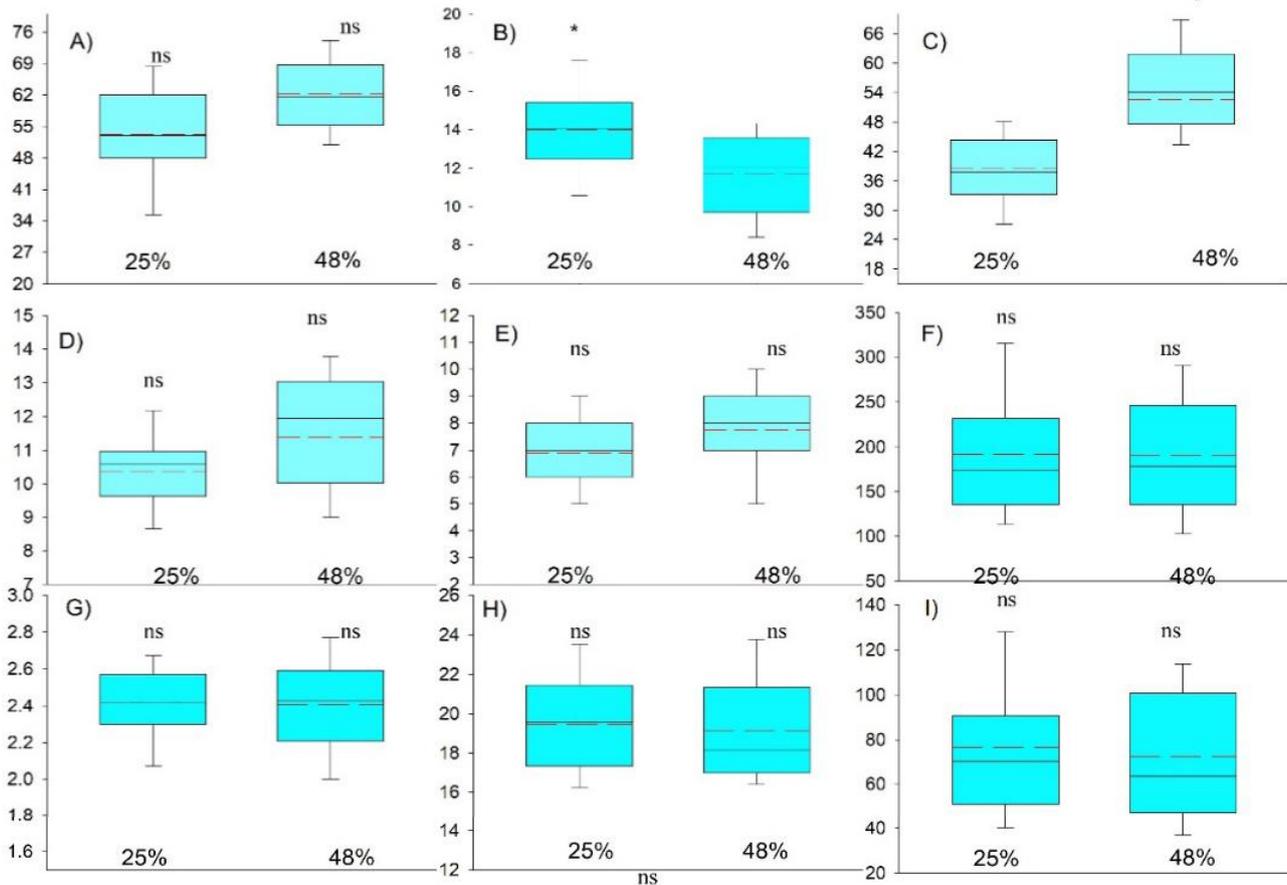
Seed weight per plant showed a significant positive correlation with NPP, D, HSW, and LA, and a negative correlation with Chl a both shading levels (Fig. 2). This finding suggests that the indirect selection of plants for larger D, NPP, LA, and HSW, and a lower leaf Chl a content may concomitantly aid the selection for higher yield under shaded environments. Similarly, Wu *et al*. (2017) studied different soybean genotypes intercropped with corn and reported correlations between grain yield and D, NPP, and NSP under light-restricted conditions. Other studies have also reported correlations between SWP and NPP (FERRARI *et al*., 2018; MACHADO *et al*., 2017; SOUSA *et al*., 2015; TEODORO *et al*., 2015) and HSW (BISINOTTO *et al*., 2017).

Taller plants showed increased SWP (Fig. 2c, 2d) under the higher level of shading (48% PAR reduction) tested. The increase in PH in shaded environments is induced by the quality of light, which may favor some genotypes under shade conditions (SCHMITT, 1997).

Furthermore, soybean plants with longer branches have larger numbers of axillary nodes for the development of pods, thus resulting in greater SWP (TEODORO *et al*., 2015).

In the 25% PAR reduction treatment, NPP was significantly and positively correlated with HSW, LA, and D (Fig.s 2a and 2b), thus demonstrating that soybean genotypes cultivated in shaded areas and showing greater seed weight, and higher D and LA, produce larger NPP and, ultimately, greater SWP.

Additional phenotypic and genotypic correlations for morphophysiological and yield characteristics in both shaded environments were observed in this study (Fig. 2), highlighting the importance of selecting and evaluating morphophysiological and yield traits that are essential for the selection of genotypes that are more adapted to shaded environments and for soybean breeding programs aimed at developing new high-yielding cultivars for cultivation in shaded environments.

Figure 2 - Correlations between nine morphophysiological and yield characteristics evaluated in sixteen soybean genotypes: a) phenotypic at 25% photosynthetically active radiation (PAR) reduction; b) genotypic at 25% PAR reduction; c) phenotypic at 48% PAR reduction; and, d) genotypic at 48% PAR reduction
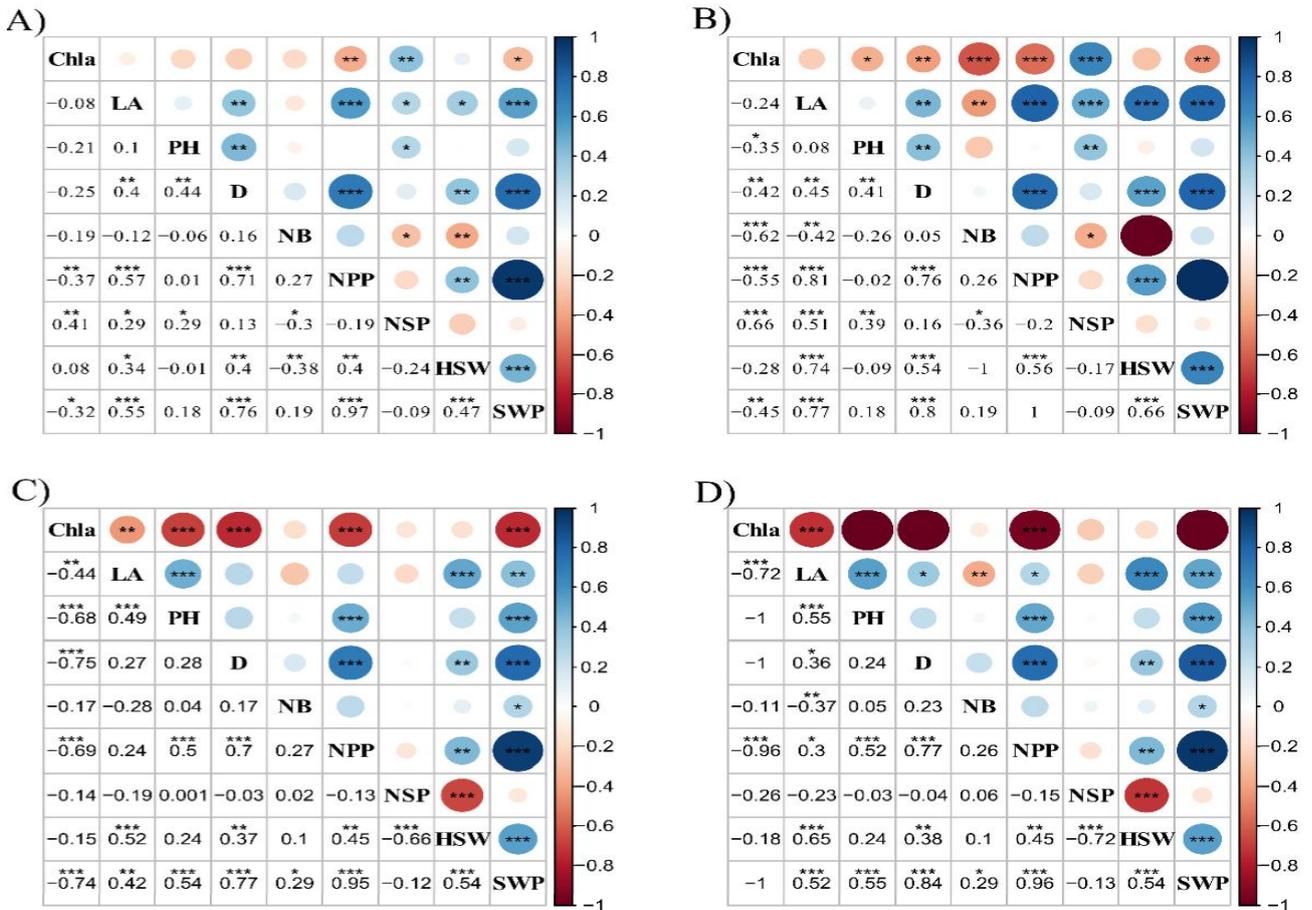
Main diagonal: variable name. Upper diagonal: circles with blue colors are positively correlated and red colors are negatively correlated. Bottom diagonal: correlation values. ***, ** and * = significant at p < 0.001, p < 0.01, p < 0.05, respectively, by the T test. Chl a = chlorophyll a; LA = leaf area; PH= plant height; D= hypocotyl diameter; NB = number of branches; NPP = number of pod per plant; NSP = number of seed per pod; HSW = hundred seed weight; SWP = seed weight per plant

However, a strong correlation between the two types of traits may hide indirect effects from other traits; therefore, these effects need to be assessed. As in addition to being heritable, genotypic correlations were higher in modulus and more significant than phenotypic correlations, we used the former for path analysis (Table 2).

Path analysis models explained 93% and 95% of SWP in the lower and the higher levels of PAR reduction, respectively, demonstrating a direct effect of the explanatory variables (Table 2). At 25% reduction in PAR, the evaluated traits that showed a greater direct effect on SWP were Chl a (0.35), LA (0.47), and NB (0.43), all with values greater than that of the residual effect (0.27). Meanwhile, at 48% reduction in PAR, LA

(0.26) and NPP (0.58) showed a greater direct effect on SWP than the residual effect (0.22). These findings suggest that these characteristics have a more significant effect on SWP, with LA at both shading levels.

Characteristics with significant genotypic correlations that have a direct effect on SWP, as well as traits that have indirect effects on SWP, should be analyzed to determine which of them might contribute most to the selection of superior cultivars. Thus, the traits with significant genotypic correlations and a direct effect on SWP were LA, NB, and Chl a at 25% reduction in PAR, and LA and NPP at 48% reduction in PAR. Significant genotypic correlations among characteristics with no direct effect on SWP might be due to indirect effects of other characteristics.

Table 2 - Partitioning of genotypic correlations into direct (bolded and underlined) and indirect effects (in column) of eight traits with the main variable dependent on seed weight per plant by path analysis of soybeans in two levels of photosynthetically active radiation (PAR) reduction

25% PAR reduction								
/1Traits	Chl a	LA	PH	D	NB	NPP	NSP	HSW	Genotypic correlation with SWP
Chl a	0.35	-0.09	-0.12	-0.15	-0.22	-0.19	0.23	-0.14	-0.45**
LA	-0.12	0.47	0.04	0.21	-0.20	0.38	0.24	0.30	0.77***
PH	-0.09	0.02	0.26	0.11	-0.07	-0.01	0.1	0	0.18
D	-0.11	0.12	0.11	0.26	0.01	0.2	0.04	0.09	0.80***
NB	-0.27	-0.18	-0.11	0.02	0.43	0.11	-0.16	-0.27	0.19
NPP	-0.13	0.18	-0.01	0.17	0.06	0.23	-0.05	0.13	1.00***
NSP	-0.17	-0.14	-0.1	-0.04	0.1	0.05	-0.27	0.13	-0.09
HSW	-0.09	0.14	0	0.08	-0.14	0.12	-0.11	0.22	0.56***
Coefficient of determination	0.93	Residual effect	0.27	K	0.49	Number of condition		22.28
48% PAR reduction								
Traits	Chl a	LA	PH	D	NB	NPP	NSP	HSW	Genotypic correlation with SWP
Chl a	-0.07	0.05	0.07	0.08	0.01	0.07	0.02	0.02	-1.00***
LA	-0.19	0.26	0.14	0.10	-0.1	0.08	-0.06	0.17	0.52***
PH	0.02	-0.01	-0.02	0	0	-0.01	0	0	0.55***
D	-0.16	0.05	0.04	0.15	0.03	0.11	-0.01	0.06	0.84***
NB	-0.02	-0.07	0.01	0.04	0.18	0.05	0.01	0.02	0.29*
NPP	-0.55	0.17	0.30	0.44	0.15	0.58	-0.09	0.26	0.96***
NSP	0	0	0	0	0	0	0.01	-0.01	-0.13
HSW	-0.01	0.02	0.01	0.01	0	0.02	-0.03	0.04	0.54***
Coefficient of determination	0.95	Residual effect	0.22	K	0.12	Number of condition		11.84

/1Chl a = chlorophyll a; LA = leaf area (cm²); PH = plant height (cm); D = hypocotyl diameter (mm); NB = number of branches; NPP = number of pod per plant; NSP = number of seed per pod; HSW = hundred seed weight (g); SWP = seed weight per plant (g). ***, ** and * = significant at a $p < 0.001$, $p < 0.01$, $p < 0.05$, respectively, by the T test

The characteristics that indirectly contributed to greater SWP at 25% reduction in PAR were NPP, D, and LA. Furthermore, the indirect effect of NPP enabled the identification of an association of SWP with D (0.17), and with LA (0.18). In turn, D had a high indirect effect on SWP when the effect of NPP (0.20) was verified. Similarly, LA had a high indirect effect on SWP when effects of D (0.21), NPP (0.38), and HSW (0.30) were verified.

As for the 48% PAR reduction treatment, the characteristics that most indirectly influenced SWP were the same as those in the 25% PAR reduction treatment, that is, NPP, D, and LA. For NPP, an indirect effect was observed with LA (0.17), PH (0.30), D (0.44), and NB (0.15)). As for D, the indirect effect was associated with NPP (0.11). LA showed indirect effect of 0.14 (associated with PH), 0.10 (associated with D), and 0.08 (associated with NPP).

Indirect selection of high-yielding soybean cultivars may be achieved by the selection for LA, as a larger LA implies that the plant will be able to produce more photoassimilate, which in turn will result in the growth of more leaves, one of the morphological traits that most contribute to SWP and, ultimately, to final yield. Furthermore, SWP can also be increased by increasing NB, as the increase in NB favors the development of reproductive structures from which pods develop and, consequently, increase NPP (FERRARI *et al*., 2018). In most cases, high-yielding plants produce a greater number of fruits and seeds, thereby increasing productivity (SOUZA *et al*., 2013).

The number of pods per plant has not only one of the strongest direct effects on seed yield but reportedly exerts strong indirect effects as well (MACHIKOWA; LAOSUWAN, 2011). Furthermore, Teodoro *et al*. (2015) observed an indirect effect of NB on grains yield due to its association with NPP. The differences between our own results and those previously reported for direct and indirect effects and correlations are likely due to the environments and genotypes tested in each case.

In summary, LA showed a genotypic correlation and a direct effect on SWP at both shading levels. Furthermore, NPP, D, and LA indirectly contributed to SWP. In other words, breeding programs should aim for selecting soybean cultivars with plants that grow larger LA, greater NPP, and larger D, as these, will guarantee greater SWP.

Genetic gains, selection and dissimilarity between soybean cultivars

Simultaneous selection of traits is the most appropriate strategy for those that show significant correlations but with a low direct effect, especially those in which indirect effects are significant (CRUZ; REGAZZI; CARNEIRO, 2014). Thus, all the characteristics of the path analysis were used, and genetic gains attributed to each characteristic at each shading level were obtained through the rank-sum index of Mulamba and Mock (1978) (Table 3), which has already been shown to be effective in selecting soybean genotypes (SOARES et al., 2015).

At the lower level of shading, group I was composed of 13 (over 81%) of the soybean cultivars included for analysis, with a cut similarity close to 50%. Thus, these cultivars showed some similarities that remained in the same group of the evaluated characteristics, while the most divergent cultivars went into separate groups. Groups II, III, and IV were represented by cultivars NS7667, RK7518, and NS7780, respectively (Fig. 3a).

Meanwhile, at the higher level of shading, group I included the same number of cultivars as in the lower level of shading, i.e., 13 cultivars, 10 of which were the same at both levels of shading but with similarity close to 35%. In turn, group II was formed by cultivars NS8338 and 8579RSF, and one of the characteristics that may have contributed to the formation of this group was SWP, whose values were some of the largest (data not shown). Lastly, group III was represented by one cultivar only, NS7901(Fig. 3b).

Table 3 - Estimates of genetic gain (%) and of the nine soybean genotypes selected by the Mulamba and Mock (1978) selection index using the economic weight coefficient of genetic variation in the analysis of variance for each level of photosynthetically active radiation (PAR) reduction

[1]Traits	25% PAR reduction				
	Genetic gain	Genotypes	Rank	Genotypes	Rank
Chl a	-3.15	NS 7780	1	AS 3680	10
LA	2.48	NS 8338	2	M 6210	11
PH	1.13	NS 7901	3	8473 RSF	12
D	6.52	RK 8115	4	RK 6316	13
NB	3.4	NS 7667	5	RK 7518	14
NPP	19.64	8579 RSF	6	CD 2728	15
NSP	-0.9	CZ 37B43	7	M 7110	16
HSW	2.86	74177 RSF	8	-	-
SWP	23.82	RK 6719	9	-	-
Traits	48% PAR reduction				
	Genetic gain	Genotypes	Rank	Genotypes	Rank
Chl a	-5.93	8579 RSF	1	74177 RSF	10
LA	3.29	NS 7901	2	RK 6719	11
PH	4.3	RK 7518	3	M 7110	12
D	9.03	NS 7667	4	RK 6316	13
NB	3.2	RK 8115	5	M 6210	14
NPP	23.63	NS 7780	6	AS 3680	15
NSP	0.34	NS 8338	7	8473 RSF	16
HSW	5.38	CD 2728	8	-	-
SWP	25.94	CZ 37B43	9	-	-

[1]Chl a = chlorophyll a; LA = leaf area; PH = plant height; D = hypocotyl diameter; NB = number of branches; NPP = number of pod per plant; NSP = number of seed per pod; HSW = hundred seed weight; SWP = seed weight per plant

Group separation is important for the design of artificial crossing strategies to obtain plants that have the best characteristics at each shading level (DELLAGOSTIN *et al*., 2011). To this purpose, Cruz, Regazzi and Carneiro (2014) suggested the no crossing of cultivars in the same group, such as not to restrict genetic variability and thus, avoid negative reflections on the gains to be obtained by selection. Therefore, we recommend that cultivars be clustered in more distant groups that have parents with a high average for the traits under improvement.

Thus, at the 25% PAR reduction level, group I included seven cultivars (CZ37B43, 74177RSF, RK8115, RK6719, NS7901, 8579RSF, and NS8338) selected by genetic gain through the Mulamba and Mock index. When crossed with group II (NS7667) or group IV (NS7780), which were also selected by the index, these seven cultivars in group I may provide a high heterotic effect upon hybridization. Similarly, at the 48% PAR

reduction level, group I included six cultivars (CZ37B43, NS7780, RK8115, CD2728, NS7667, and RK7518) again, selected according to the Mulamba and Mock selection index, whose crossings with cultivars in groups II (NS8338 and 8579RSF) or III (NS7901), which are also cultivars selected by the index, may be successful in future generations. Thus, at the two levels of shading, five cultivars (CZ37B43, RK8115, NS7901, NS8338, and 8579RSF) were identified as superior by the Mulamba and Mock selection index, and can be used as parents together with cultivars from other groups.

The contribution of the various characteristics under study to the divergence among the soybean cultivars tested were different at the two levels of shading, with SWP (27%) and PH (21.16%) discriminating more at the 25% PAR reduction level, and NPP (25.28%) and HSW (25.83%) discriminating more at 48% PAR reduction (Fig. 4).

Figure 3 - Dendrograms obtained by the UPGMA method, based on the generalized Mahalanobis distance between sixteen soybean genotypes based on nine morphophysiological and yield characteristics, a) 25% photosynthetically active radiation (PAR) reduction, cophenetic correlation: 0.75; e, b) 48% PAR reduction, cophenetic correlation: 0.82. x axis: distance in percentage (%)
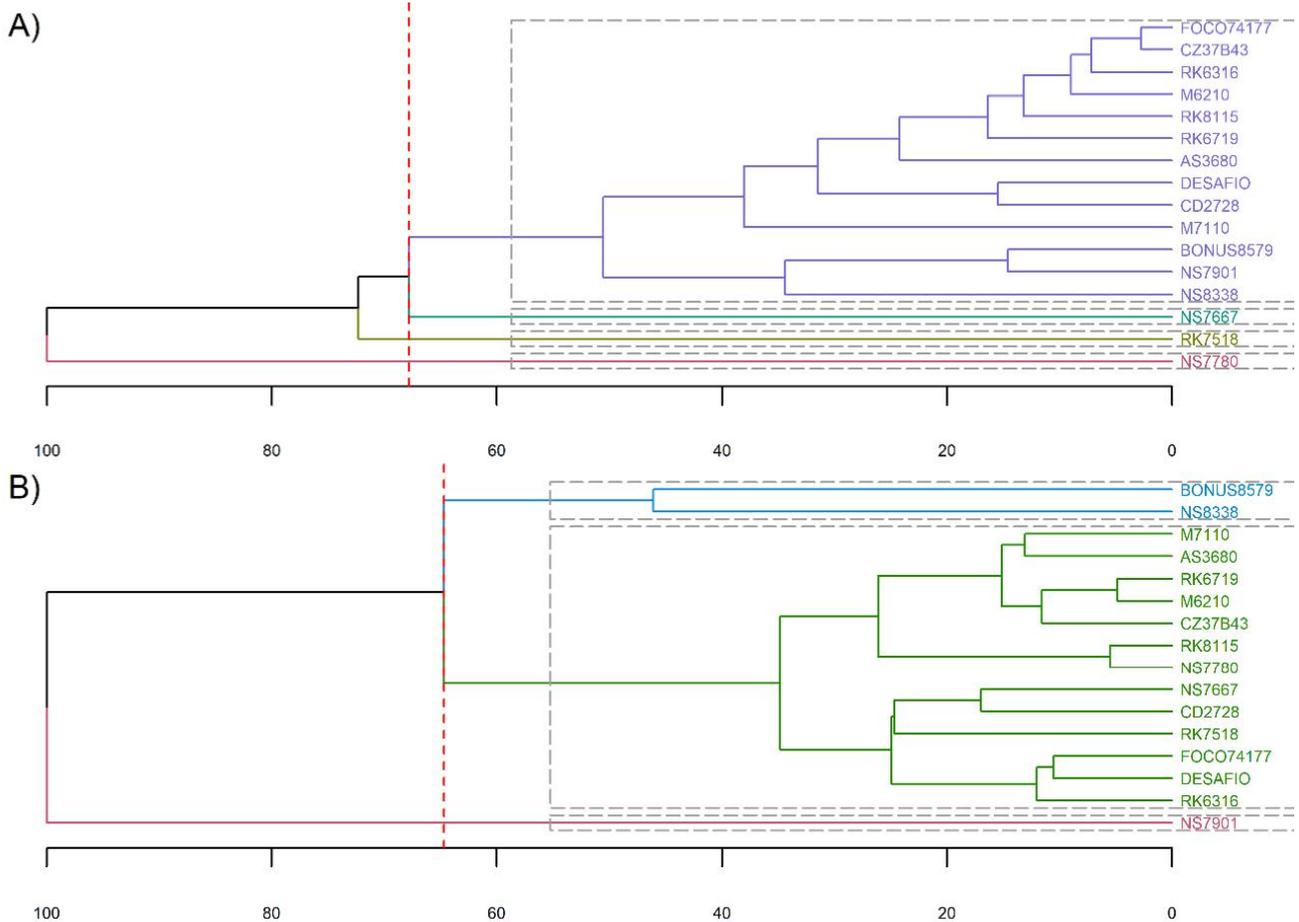
Red line vertically means the cut made according to the method of Mojena (1977) for the formation of groups identified by colors within the dendrogram

Figure 4 - Relative contribution (%) of traits to divergence (Singh, 1981) using the Mahalanobis distance, in sixteen soybean genotypes of two levels of photosynthetically active radiation (PAR) reduction
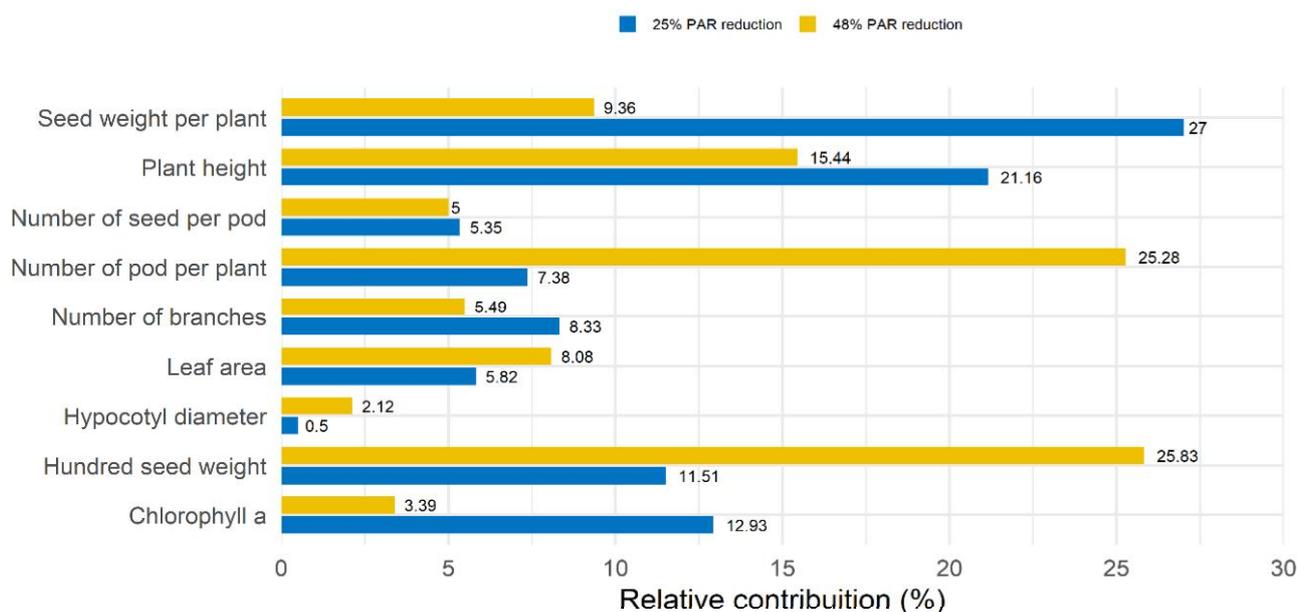
The smallest contributions were from D at both shading levels, whereby, this variable can be disregarded in future analyses. This demonstrates that shading interferes with attempts to explain the dissimilarity among cultivars and that, at each level of shade, the characteristics that contributed the most should be prioritized when choosing cultivars and implementing integrated systems and/or in choosing parent materials in future breeding programs. Th, our results are important because they support soybean genetic-improvement programs for integrated cropping systems.

CONCLUSIONS

1. Soybean genotypes showed genetic variability for all the traits under study, with enhanced genetic expression, and high variation index and heritability;

2. At both levels of photosynthetically active radiation (PAR) reduction tested, leaf area showed genotypic correlation and a direct effect on seed weight per plant, while the number of pods per plant, hypocotyl diameter and leaf area had indirect effects on seed weight per plant;

3. Seed weight per plant and plant height at 25% PAR reduction, and number of pods per plant and hundred seed weight at the 48% PAR reduction allowed the best discrimination criteria among cultivars in shaded environments;

4. Genoytpes NS7780, NS7667, CZ37B43, RK8115, NS7901, NS8338 and 8579RSF are potential genetic

resources for use as parents in breeding programs aimed at developing shade-tolerant soybean.

REFERENCES

ANDRADE JÚNIOR, V. C. *et al*. Associations between morphological and agronomic characteristics in garlic crop. **Horticultura Brasileira**, v. 37, p. 204-209, 2019.

BISINOTTO, F. F. *et al*. Path analysis and traits correlation in soybean. **Communications in Plant Sciences**, v. 7, n. 1/2, p. 27-33, 2017.

CARPENTER, A. C.; BOARD, J. E. Branch yield components controlling soybean yield stability across plant populations. **Crop Science**, v. 37, p. 885-891, 1997.

CARVALHO, C. G. P. *et al*. Proposta de classificação dos coeficientes de variação em relação à produtividade e altura da planta de soja. **Pesquisa Agropecuária Brasileira**, v. 38, n. 2, p. 187-193, 2003.

CARVALHO, S. P.; CRUZ, C. D. Diagnosis of multicollinearity: assessment of the condition of correlation matrices used in genetic studies. **Brazilian Journal of Genetics**, v. 19, p. 479-484, 1996.

CRISTO, E. *et al*. Growth and yield of soybean cultivated in agroforestry systems. **Revista Ceres**, v. 67, p. 165-175, 2020.

CRUZ, C. D. Genes Software-extended and integrated with the R, Matlab and Selegen. **Acta Scientiarum. Agronomy**, v. 38, p. 547-552, 2016.

CRUZ, C. D.; REGAZZI, A. J.; CARNEIRO, P. C. S. **Modelos biométricos aplicados ao melhoramento genético**. 4. ed. Viçosa, MG: Editora UFV, 2014. v. 1, 514 p.

DELLAGOSTIN, M. *et al*. Dissimilaridade genética em população segregante de soja com variabilidade para caracteres morfológicos de semente. **Revista Brasileira de Sementes**, v. 33, n. 4, p. 689-698, 2011.

FALCONER, D. S.; MACKAY, T. F. C. **Introduction to quantitative genetics**. England: Pearson Education India Longmans Green, 1996.

FAN, Y. *et al*. Soybean (*Glycine max* L. Merr.) seedlings response to shading: leaf structure, photosynthesis and proteomic analysis. **BMC Plant Biology**, v. 19, n. 1, p. 1-12, 2019.

FEHR, W.; CAVINESS, C. Stages of soybean development. **Special Report Cooperative Extension Service. Agriculture Home Economics Experiment Station**, v. 80, p. 1-12, 1977.

FENG, L. *et al*. The influence of light intensity and leaf movement on photosynthesis characteristics and carbon balance of soybean. **Frontiers in Plant Science**, v. 9, 1952, 2019.

FERRARI, M. *et al*. Path analysis and phenotypic correlation among yield components of soybean using environmental stratification methods. **Australian Journal of Crop Science**, v. 12, n. 2, p. 193-202, 2018.

FAO. FAOSTAT. **Statistical database**. Disponível em: https://www.fao.org/faostat/en/. Acesso em: 30 out. 2021.

GOMES, F. P. **Curso de estatística experimental**. 14. ed Piracicaba: Nobel, 2000. 468 p.

GONG, W. Z. *et al*. Tolerance vs. avoidance: two strategies of soybean (Glycine max) seedlings in response to shade in intercropping. **Photosynthetica**, v. 53, n. 2, p. 259-268, 2015.

GUIMARÃES, A. G. *et al*. Population structure and impact of recurrent selection on popcorn using EST-SSR markers. **Acta Scientiarum. Agronomy**, v. 40, 2018.

MACHADO, B. Q. V. *et al*. Phenotypic and genotypic correlations between soybean agronomic traits and path analysis. **Genetics and Molecular Research**, v. 16, n. 2, 2017.

MACHIKOWA, T.; LAOSUWAN, P. Path coefficient analysis for yield of early maturing soybean. **Songklanakarin Journal of Science & Technology**, v. 33, n. 4, 2011.

MAHALANOBIS, P. C. On the generalized distance in statistics. **National Instituto of Science of India**, v. 2, p. 49-55, 1936.

MANTEL, N. The detection of disease clustering and a generalized regression approach. **Cancer Research**, v. 27, n. 2, p. 209-220, 1967. Part 1.

MOJENA, R. Hierarchical grouping methods and stopping rules: an evaluation. **The Computer Journal**, v. 20, n. 4, p. 359-363, 1977.

MONTEIRO, F. F, *et al*. Breeding for yield and seed quality in soybean. **Euphytica**, v. 217, n. 12, p. 1-10, 2021.

MONTGOMERY, D. C; PECK, E. A; VINING, G. G. **Introduction to linear regression analysis**. New York: John Wiley, 2012.

MULAMBA, N. N.; MOCK, J. J. Improvement of yield potential of the ETO blanco maize (Zea mays L.) population by breeding for plant traits [Mexico]. **Egyptian Journal of Gnetics and Cytology**, 1978.

POEHLMAN, J. M. Breeding soybeans. *In*: POEHLMAN, J. M. (ed.) **Breeding field crops**. 3rd ed. New York: Van Nostrand Reinhold, 1987. p. 421-450.

R CORE TEAM. **R**: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing, 2020.

RAMALHO, M. A. P. *et al*. **Aplicações da genética quantitativa no melhoramento de plantas autógamas**. Lavras: Editora UFLA, 2012. 522 p.

RAZA, M. A. *et al*. Maize leaf-removal: a new agronomic approach to increase dry matter, flower number and seed-yield of soybean in maize soybean relay intercropping system. **Scientific Reports**, v. 9, n. 1, p. 1-13, 2019.

SCHMITT, J. Is photomorphogenic shade avoidance adaptive? Perspectives from population biology. **Plant, Cell & Environment**, v. 20, n. 6, p. 826-830, 1997.

SINGH, D. The relative importance of characters affecting genetic divergence. **The Indian Journal of Genetic and Plant Breeding**, v. 41, n. 2, p. 237-245, 1981.

SOARES, I. O *et al*. Adaptability of soybean cultivars in different crop years. **Genetics and Molecular Research**, v. 14, n. 3, p. 8995-9003, 2015.

SOUSA, L. B. D. *et al*. Correlation between yield components in F6 soybean progenies derived from seven biparental crosses. **Bioscience Journal (Online)**, p. 1692-1699, 2015.

SOUZA, C. A. *et al*. Plant architecture and productivity of soybean affected by plant growth retardants. **Bioscience Journal**, v. 29, n. 3, p. 634-643, 2013.

TAIZ, L. *et al*. **Fisiologia e desenvolvimento vegetal**. 6. ed. Porto Alegre: Artmed, 2017. 858 p.

TEODORO, P. E. *et al*. Path analysis in soybean genotypes as function of growth habit. **Bioscience Journal (Online)**, p. 794-799, 2015.

TIBOLLA, L. B. *et al*. Effect of artificial shading on soybean growth and yield. **Revista Brasileira de Ciências Agrárias**, v. 14, n. 4, p. 1-7, 2019.

WELLBURN, A. R. The spectral determination of chlorophylls a and b, as well as total carotenoids, using various solvents with spectrophotometers of different resolution. **Journal of Plant Physiology**, v. 144, p. 307-313, 1994.

WERNER, F. *et al*. Agronomic performance of soybean cultivars in an agroforestry system. **Pesquisa Agropecuária Tropical**, v. 47, p. 279-285, 2017.

WU, Y. *et al*. Responses to shade and subsequent recovery of soya bean in maize-soya bean relay strip intercropping. **Plant Production Science**, v. 19, n. 2, p. 206-214, 2016.

WU, Y. S. *et al*. Shade adaptive response and yield analysis of different soybean genotypes in relay intercropping systems. **Journal of Integrative Agriculture**, v. 16, n. 6, p. 1331-1340, 2017.

YANG, F. *et al*. Growth of soybean seedlings in relay strip intercropping systems in relation to light quantity and red: far-red ratio. **Field Crops Research**, v. 155, p. 245-253, 2014.