

# GENETIC DIVERGENCE BETWEEN SWEET SORGHUM GENOTYPES BY THE WARD-MLM PROCEDURE

## *DIVERGÊNCIA GENÉTICA ENTRE GENÓTIPOS DE SORGO SACARINO PELO PROCEDIMENTO WARD-MLM*

**Adriano dos SANTOS<sup>1</sup>; Gessi CECCON<sup>2</sup>; Erina Vitória RODRIGUES<sup>3</sup>; Larissa Pereira RIBEIRO<sup>4</sup>; Katiane Secco CASTRO<sup>5</sup>; Paulo Eduardo TEODORO<sup>5</sup>**

1. Universidade Estadual do Norte Fluminense "Darcy Ribeiro", Campos dos Goytacazes, RJ, Brasil; 2. Embrapa Agropecuária Oeste, Dourados, MS, Brasil; 3. Universidade de Brasília, Brasília, DF, Brasil; 4. Universidade Federal de Viçosa, Viçosa, MG, Brasil; 5. Universidade Federal de Mato Grosso do Sul, Chapadão do Sul, MS, Brasil. eduteodoro@hotmail.com

**ABSTRACT:** Sweet sorghum has the potential for use in bioethanol production as a complement to sugarcane. This work aimed to identify sweet sorghum genotypes by mixed models, using the Ward-MLM method. Trials were carried out in the 2013/2014 and 2014/2015 crop years at the experimental area of Embrapa Agropecuária Oeste, in Dourados, Mato Grosso do Sul. The experiment consisted of a randomized blocks design, with three replications, using 16 sweet sorghum genotypes from the breeding program of *Embrapa Milho e Sorgo*. The genotype effect was significant for all the variables analyzed, and the genotypes x environments interaction was significant for most of these variables. Estimates of genetic parameters indicate that gains with selection can be obtained for the dry mass yield of panicles and °Brix. The Ward-MLM procedure is useful for detecting genetic divergence and clustering genotypes by simultaneously using morphological, agronomic, and molecular descriptors. In addition, this method showed that three was the ideal number of groups, according to the pseudo-F and pseudo-t<sup>2</sup> criteria, identifying the most divergent groups. When crossed, these groups have a higher probability of generating genotypes with high yield and variability.

**KEYWORDS:** *Sorghum bicolor* L. Genetic diversity. Mixed models. Canonical variables.

### INTRODUCTION

The composition of the world energy matrix over the history was predominantly marked by non-renewable energy sources. However, given the large amount of pollutant gases emitted by fossil fuels and constant pressures on the environmental sustainability, countries have been engaged in searching and adopting new renewable energy sources (CREMONEZ et al., 2015; KUMAR et al., 2016). In this context, biofuel production emerges as an alternative to fossil fuel, constituting a promising world energy matrix, especially in the Brazilian scenario.

The importance of diversifying the raw material used in the biofuel subgroups (ethanol and biodiesel) has been addressed in several studies. Ethanol productions have a marked presence of sugarcane as the primary raw material, which has made Brazil the world's largest bioethanol producer (ANP, 2016). However, alternative crops must be developed to meet the country's high energy demand. Sweet sorghum (*Sorghum bicolor* L.) has the potential for use in bioethanol production, besides sugarcane. *S. bicolor* is considered as a "smart crop" for it produces fuel from biomass and fermentable sugars, which, when industrialized, are transformed into ethanol and/or food (PEREIRA FILHO et al., 2013).

Despite the great relevance of sweet sorghum in the agro-energetic scenario, only a few studies have addressed its genetic improvement. Thus, genotypes must be accurately selected, and the genetic variability in the crop must be characterized, aiming at the composition of breeding programs. One of the primary objectives of breeding programs is to identify plants with a gene of interest for the desired trait. Therefore, the study of genetic diversity is fundamental in developing the potential of sweet sorghum cultivars for ethanol production.

Among the different multivariate methods to quantify the genetic variability of a species, the Ward-MLM (Modified Location Model), proposed by Franco et al. (1998), stands out. This method is performed in two steps: first, groups are defined by the Ward clustering method (WARD, 1963), using the Gower dissimilarity matrix (GOWER, 1971); subsequently, the mean vector of each subgroup is estimated by the MLM procedure. This method also enables defining the optimum number of groups and is reliable in the identification of the best probability of each genotype to be allocated into a given group (GONÇALVES et al., 2009).

The Ward-MLM procedure has been effective in discriminating genotypes of several crops, such as maize (COIMBRA et al., 2009; RIBEIRO et al., 2015); common bean (CABRAL et al., 2011; OLIVEIRA et al., 2013); soybean

(TEODORO et al., 2016); coffee (RODRIGUES et al., 2016); sugarcane (BRASILEIRO et al., 2014); sour passion fruit (SILVA et al., 2014); guava (CAMPOS et al., 2013); and castor (OLIVEIRA et al., 2013). However, despite the great importance of using accurate techniques in the quantification of genetic diversity (such as the Ward-MLM method), to date, no studies have reported the use of this procedure in sweet sorghum. Thus, this work aimed to identify sweet sorghum genotypes by mixed models, using the Ward-MLM method.

## MATERIAL AND METHODS

Trials were carried out in the 2013/2014 and 2014/2015 crop seasons at experimental area of Embrapa Agropecuária Oeste, in Dourados, MS (22°17'S and 54°48'W, with 380 m of altitude). The soil of the experimental area was identified as Oxisol dystroferric clay texture. The climate of the region is Cwa, according to the Köppen classification, with hot summers and dry winters.

The experiment consisted of a randomized blocks design, with three replications. The following 16 sweet sorghum genotypes from the breeding program of *Embrapa Milho e Sorgo* were evaluated: CMSXS629 (1), CMSXS630 (2), CMSXS643 (3), CMSXS644 (4), CMSXS646 (5), CMSXS647 (6), BRS 506 (7), BRS 508 (8), BRS509 (9), BRS 511 (10), CV 198 (11), CV 568 (12), Sugargraze (13), V82391 (14), V82392 (15), and V82393 (16). Plots contained four 5-m rows, with 0.50 m of spacing.

Fertilization consisted of applying 200 kg ha<sup>-1</sup> (08-16-16) at sowing and 50 kg ha<sup>-1</sup> of N as topdressing. Weed control was performed with pre-sowing desiccation with 1.44 L ha<sup>-1</sup> of glyphosate, plus one application of atrazine at 1.2 kg ha<sup>-1</sup>, between 20 and 25 days after emergence. Insect pests were controlled by applying Tiametoxam+ Lambda-Cialotrina (Engeo-pleno<sup>®</sup>) insecticide at 15 days after emergence, using 0.005 L ha<sup>-1</sup>.

The following agronomic traits were evaluated: percentage of panicles (PP, %); percentage of stems (PS, %); dry mass content in the stems (DMCS, kg ha<sup>-1</sup>); dry mass content in the panicles (DMCP, kg ha<sup>-1</sup>); fresh mass yield of stems (FMYS, kg ha<sup>-1</sup>); fresh mass yield of leaves (FMYL, kg ha<sup>-1</sup>); fresh mass yield of panicles (FMYP, kg ha<sup>-1</sup>); dry mass yield of panicles (DMYP, kg ha<sup>-1</sup>); total fresh mass yield (TFMY, kg ha<sup>-1</sup>); total dry mass yield (TDMY, kg ha<sup>-1</sup>); mean dry mass (MDM, kg ha<sup>-1</sup>); broth yield (BY, l ha<sup>-1</sup>); and Brix (°Brix). A subsample of five plants was collected to separate stems, leaves, and panicles, aiming at determining the PS and PP. Subsequently, this

sample was dried in an oven at 65 °C for 72 h to determine MDM. The BY was calculated by the difference between TFMY and MDM. Brix was evaluated using a Portable Digital Refractometer, with a range of 0 - 45% of Brix.

The statistical analysis consisted of joint analysis of variance, considering the two years, to obtain phenotypic means. Afterward, data were analyzed simultaneously, by the Ward-MLM procedure, to create the groups. For the Ward clustering method, the distance matrix was generated by the Gower's algorithm (GOWER, 1971). The ideal number of groups was determined according to the pseudo-F and pseudo-t<sup>2</sup> criteria. Differences between groups, the correlation between the variables, and the canonical (CAN) variables were examined graphically. The distance for the distribution of the trait proposed by Franco et al. (1998) was used to determine the dissimilarity among groups. All analyses were carried out in the statistical SAS software, using the procedures PROC-MLM and PROC-CANDISC (SAS INSTITUTE, 2009).

Genetic parameters from each trial were estimated for each variable, according to Cruz et al. (2014). Selective accuracy (SA) and the relative coefficient of variation (CV<sub>r</sub>) were estimated using the method proposed by Resende (2002) and Resende and Duarte (2007), according to the following equation:  $SA = \sqrt{1 - \frac{1}{F_c}}$  and  $CV_r = \sqrt{F_c - \frac{1}{b}}$ , where F<sub>c</sub> is the F-test value of the genotype effect associated with the analysis of variance, and b is the number of replications, in this case, blocks.

## RESULTS AND DISCUSSION

Effect of genotype was significant for all analyzed variables, which allows inferring about the presence of genetic variability among the studied genotypes (Table 1). The variables TFMY, BY, and °Brix should be highlighted for ethanol production since cultivars with high biomass production, associated with high BY and °Brix are more promising for this purpose.

Significant Genotypes vs. Environments interactions (G x E) were observed. This fact is related to the differences in macroenvironmental aspects of the sites, such as the climatic variables (temperature and rainfall) that influence traits expression. For variables FMYS, FMYL, TDMY, BY, and °Brix, genotypes were not influenced by the environments, that is, genotypes x environments interaction was not significant for this set of

variables. The absence of genotype x environment interaction allows recommending genotypes based

only on the mean of environments.

**Table 1.** Summary of joint analysis of variance for 13 agro-industrial variables, evaluated in 16 sweet sorghum genotypes, in the 2013/14 and 2014/15 crop seasons, in Dourados, MS.

SV	Blocks/Env	Genotypes (G)	Environments (E)	G x E	Error
TFMY	43,790,730.2	291331404.1**	506736600.0*	104834122.22	83296924.65
PS	19.78	22.90**	1210.05**	20.08**	7.03
PP	3.97	18.86**	108.63**	23.92**	5.29
DMCS	37.78	19.17*	36.82	21.08*	9.25
DMYP	192.48	558.01**	13658.90**	365.77**	106.06
FMYS	49183560.59	193555620.3**	1670883830.9**	68168669.14	63885215.78
FMYL	369468.08	7134919.62**	227516793.41**	4524083.62	2468073.26
FMYP	2103450.80	7794017.28**	35618365.09*	11288281.1**	1969146.31
DMCP	799749.43	1208210.01**	27169037.85**	1916851.26**	461604.67
TDMY	9010358.12	21255190.69**	27994280.40	12610344.65	8690812.48
MDM	37.31	17.81*	292.22*	18.11*	7.91
BY	57064779.75	139560827.1**	1237571314.1**	52875917.27	40263426.05
°Brix	10.70	16.07**	48.86	6.52	3.60
DF	4	15	1	15	60

\*\* and \*: significant at ( $P < 0.01$ ) and significant at ( $P < 0.05$ ), respectively, by the F test. SV: sources of variation; DF: degrees of freedom; percentage of panicles (PP, %); percentage of stems (PS, %); dry mass content in the stems (DMCS,  $\text{kg ha}^{-1}$ ); dry mass content in panicles (DMCP,  $\text{kg ha}^{-1}$ ); fresh mass yield of stems (FMYS,  $\text{kg ha}^{-1}$ ); fresh mass yield of leaves (FMYL,  $\text{kg ha}^{-1}$ ); fresh mass yield of panicles (FMYP,  $\text{kg ha}^{-1}$ ); dry mass yield of panicles (DMYP,  $\text{kg ha}^{-1}$ ); total fresh mass yield (TFMY,  $\text{kg ha}^{-1}$ ); total dry mass yield (TDMY,  $\text{kg ha}^{-1}$ ); mean dry mass (MDM,  $\text{kg ha}^{-1}$ ); broth yield (BY,  $\text{l ha}^{-1}$ ); and Brix (°Brix).

In order to infer more details on the genetic variability in the set of genotypes analyzed, as well as predict genetic gains with selection, genetic parameters associated with variables were estimated (Table 2). A marked interaction with the environment was detected for all variables, which is confirmed by the higher magnitudes of the phenotypic coefficients of variation ( $\hat{C}V_F$ ), in detriment to the genotypic coefficients of variation ( $\hat{C}V_G$ ) for all variables.

Estimates of genotypic coefficient of determination ( $H^2$ ) were higher than 70% for variables TFMY, PP, DMYP, FMYP, BY, and °Brix. These estimated values suggest the possibility of obtaining genetic gains and allow the phenotypic selection of superior individuals for these variables. However, since the genotypic coefficient of determination is not the property of a trait, but of the population and environmental conditions, the values obtained can be considered only as indicative of favorable heritability for the development of sweet sorghum genotypes with high °Brix contents and high BY.

The genotypic coefficient of variation ( $\hat{C}V_G$ ) is a more accurate choice of breeding strategy since it allows understanding the genotypic variability of

the different traits. However, for a real inference on each variable aiming at breeding, the genotypic coefficient of variation and the residual coefficient of variation must be analyzed using the variation index ( $I_v$ ). Thus, all variables presented values lower than the unit, indicating that the greatest variation between genotypes is due to the environment.

Only DMYP and °Brix showed values of relative coefficient of variation (CVr) higher than the unit, suggesting that selection gains can be obtained (VENCOVSKY; BARRIGA, 1992). Thus, the values observed for the variables cited are sufficient to provide accurate inference about the genetic value of the progenies.

The trial's precision was measured by the estimate of the selective accuracy ( $A_s$ ) since it reflects the quality of the information and procedures used in predicting breeding values. This measure is associated with the selection accuracy and refers to the correlation between predicted genetic values and true genetic values of individuals (PIMENTEL et al., 2014). In this context, the results represent high experimental accuracy since the estimates of selective accuracy were higher than 0.70 (RESENDE; DUARTE, 2007).

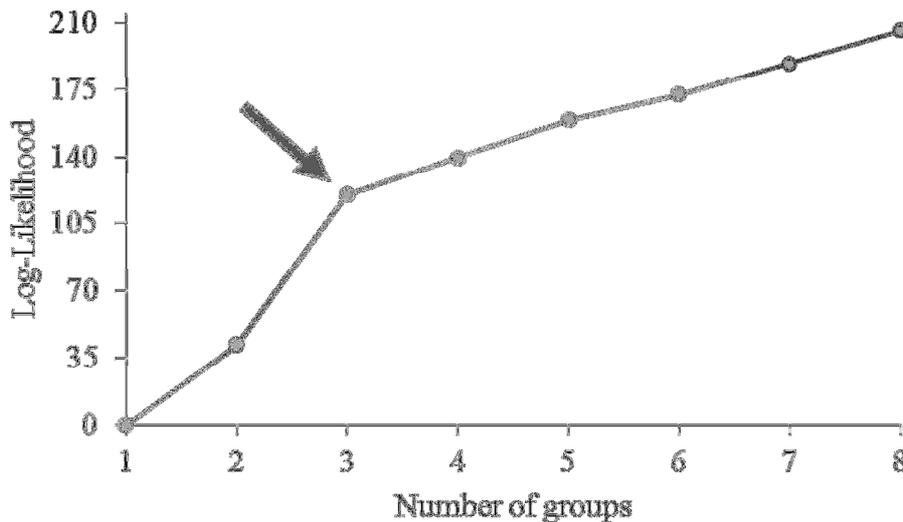
**Table 2.** Estimates of genotypic coefficient of determination ( $H^2$ ), genotypic coefficient of variation ( $\hat{C}V_g$ ), phenotypic coefficient of variation ( $\hat{C}V_f$ ), variation index ( $I_v$ ), relative coefficient of variation ( $CV_r$ ), and selective accuracy ( $AS$ ) for 13 variables analyzed in sweet sorghum.

Variables	$H^2$	$CV_g$	$CV_f$	$I_v$	$CV_r$	As	Mean
TFMY	71.41	9.28	10.98	0.65	0.91	0.85	63458.96
PS	69.28	2.00	2.40	0.61	0.87	0.83	81.26
PP	71.92	28.27	33.34	0.65	0.92	0.85	5.32
DMCS	51.73	5.89	8.20	0.42	0.60	0.72	21.81
DMYP	80.99	16.03	17.81	0.84	1.19	0.90	54.15
FMYS	66.99	9.00	10.99	0.58	0.82	0.82	51658.04
FMYL	65.41	10.80	13.35	0.56	0.79	0.81	8168.13
FMYP	74.74	29.34	33.93	0.70	0.99	0.86	3358.64
DMCP	61.79	20.92	26.61	0.52	0.73	0.79	1686.09
TDMY	59.11	9.09	11.82	0.49	0.69	0.77	15919.27
MDM	55.60	5.10	6.84	0.46	0.65	0.75	25.18
BY	71.15	10.05	11.92	0.64	0.91	0.84	40466.18
°Brix	77.57	10.78	12.24	0.76	1.07	0.88	13.38

Percentage of panicles (PP, %), percentage of stems (PS, %); dry mass content in the stems (DMCS,  $kg\ ha^{-1}$ ); dry mass content in panicles (DMCP,  $kg\ ha^{-1}$ ); fresh mass yield of stems (FMYS,  $kg\ ha^{-1}$ ); fresh mass yield of leaves (FMYL,  $kg\ ha^{-1}$ ); fresh mass yield of panicles (FMYP,  $kg\ ha^{-1}$ ); dry mass yield of panicles (DMYP,  $kg\ ha^{-1}$ ); total fresh mass yield (TFMY,  $kg\ ha^{-1}$ ); total dry mass yield (TDMY,  $kg\ ha^{-1}$ ); mean dry mass (MDM,  $kg\ ha^{-1}$ ); broth yield (BY,  $l\ ha^{-1}$ ); and Brix (°Brix).

The logarithmic likelihood function was applied according to the pseudo-F and pseudo- $t^2$  criteria, associated with the likelihood ratio test, to set the optimal number of groups required. Thus, three groups were formed, with a maximum increase

of 78.15 (Figure 1). The indication of the number of groups is an innovative aspect of the Ward-MLM procedure, comparing to other hierarchical methods, and it results in a more accurate and less subjective clustering formation.



**Figure 1.** Graph of the logarithmic likelihood function (Log-likelihood) in relation to the number of groups formed by the Ward MLM strategy in sweet sorghum genotypes.

The major contributors to the diversity analysis in the first canonical variable were variables PS, FMYS, MDM, and BY. In the second

canonical variable, which explained 21.17% of the variation of the greatest contributors, were variables PS, PP, DMYP, FMYP, and DMCP (Table 3).

Assessing traits associated with BY and quality is essential to increase ethanol production. These results demonstrate the relevance of these traits for studies on genetic diversity, and consequently for

the choice of parents to be used in crosses aimed at optimizing the development of sweet sorghum cultivars.

**Table 3.** Means of quantitative and qualitative variables for each three groups formed by the *Ward-MLM* procedure and the coefficients of quantitative variables in the first two canonical variables.

Variables	G I (4)	G II (9)	G III (3)	CAN1	CAN2
TFMY	70987.5	61886.3	58138.89	0.421	0.543
PS	79.8975	82.647	78.935	0.548	-0.660
PP	6.9845	3.9973	7.056	-0.434	0.773
DMCS	20.8043	21.7682	23.2713	-0.424	-0.205
DMYP	45.3643	60.6009	46.4923	0.345	-0.720
FMYS	56866.68	51273.66	45866.32	0.539	0.393
FMYL	9061.761	7882.033	7834.919	0.186	0.462
FMYP	4785.974	2497.981	4037.486	-0.233	0.900
DMCP	2280.563	1365.648	1854.772	-0.122	0.905
TDMY	17637.94	15182.73	15837.36	0.063	0.569
MDM	24.938	24.6818	27.0157	-0.524	0.105
BY	45024.02	40188.09	35223.3	0.573	0.398
°Brix	11.846	13.8682	13.936	-0.210	-0.528

total fresh mass yield (TFMY, kg ha<sup>-1</sup>); percentage of panicles (PP, %), percentage of stems (PS, %); dry mass content in the stems (DMCS, kg ha<sup>-1</sup>); dry mass content in panicles (DMCP, kg ha<sup>-1</sup>); fresh mass yield of stems (FMYS, kg ha<sup>-1</sup>); fresh mass yield of leaves (FMYL, kg ha<sup>-1</sup>); fresh mass yield of panicles (FMYP, kg ha<sup>-1</sup>); dry mass yield of panicles (DMYP, kg ha<sup>-1</sup>); total dry mass yield (TDMY, kg ha<sup>-1</sup>); mean dry mass (MDM, kg ha<sup>-1</sup>); broth yield (BY, l ha<sup>-1</sup>); and Brix (°Brix).

The largest distances of dissimilarity between the groups, based on the Gower algorithm by the *Ward-MLM* strategy, were detected between groups I and III (573.27) and between groups II and III (479.62), while the shortest distance was observed between groups I and II (116.46) (Table 4). In this case, to avoid exhausting genetic variability, as well as restrict possible gains obtained with selection, hybridization between genotypes

from these groups is not recommended. This fact is because genetically related parents tend to have many common alleles, and therefore they provide little advantage due to the low level of heterozygosity (CRUZ et al., 2014). Crosses between genotypes from the other groups with those from group III would be interesting for the exploitation of the heterosis in the obtainment of transgressive individuals for the traits of interest.

**Table 4.** Distance between groups formed by the Gower algorithm, proposed by Franco et al. (1998).

	G I	G II	G III
G I	0	116.46	573.27
G II		0	479.62
G III			0

Sorghum breeding programs aim to increase the frequency of favorable alleles to maximize ethanol production. In this sense, using genes with larger genetic distance increases the probability of occurrence of superior segregants in advanced generations. However, genotypes should be chosen considering their *per se* performance. Thus, to produce commercial hybrids with high heterotic

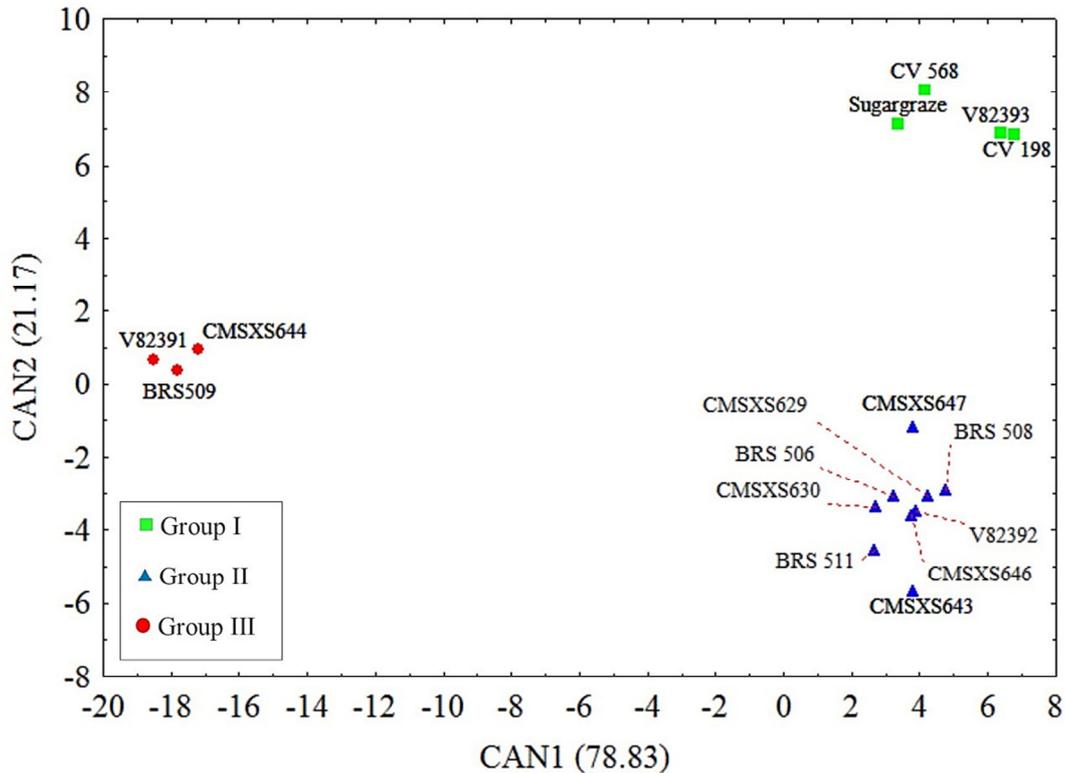
vigor, the focus should be on divergent genotypes that have superior performance for the main traits of agronomic importance and meet the breeding program goals for which these hybrids are being developed.

Due to the shorter distance between the genotypes belonging to groups I and II, the cross involving genotypes from these two groups is not

recommended since it can generate progenies with high similarity. However, depending on the breeding program's strategy, this cross, considered as convergent, could facilitate the breeder's works in selecting superior lines in less time since both groups have higher means for most of the analyzed variables.

The first two canonical variables (CAN) obtained using the Ward-MLM procedure explained 100% of the total variation (Figure 2). According to Cruz et al. (2012), when the first two CAN explain

more than 80%, the genetic variability between genotypes can be observed by dispersing the scores in a two-dimensional plot. Thus, this high value indicates that the graphical representation of the first two canonical variables is adequate to verify the relation between groups and individuals within a group, increasing the reliability of the results. The distance between the groups agreed with the canonical variables. Groups 1 and 2 were the closest, with a distance of 116.46 (Table 5). Group 2 allocated the largest number of genotypes.



**Figure 2.** Graph of the first two canonical variables for the three groups formed by the Ward-MLM procedure.

**CONCLUSIONS**

The Ward-MLM procedure is useful to detect genetic divergence and cluster genotypes by simultaneously using morphological and agronomic traits. In addition, this method showed that three was the ideal number of groups, according to pseudo-F and pseudo-t<sup>2</sup> criteria, identifying the most divergent groups.

The cross between groups I and III has a higher probability of generating genotypes with high yield and variability.

**ACKNOWLEDGMENTS**

The authors would like to thank the Federal University of Mato Grosso do Sul for their financial support (PROPP Edict 111/2018). This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

---

**RESUMO:** O sorgo sacarino apresenta potencial para utilização na produção de bioetanol em complementação à cana-de-açúcar. O objetivo deste trabalho foi identificar genótipos de sorgo sacarino via modelos mistos, usando o método Ward-MLM. O experimento foi conduzido nas safras 2013/2014 e 2014/2015, na área experimental da Embrapa Agropecuária Oeste, em Dourados, MS. O delineamento experimental utilizado foi o de blocos casualizados com três repetições. Foram avaliados 16 genótipos de sorgo sacarino oriundos do programa de melhoramento da Embrapa Milho e

Sorgo. O efeito de genótipo foi significativo para todas as variáveis analisadas, e a interação genótipos x ambientes foi significativa para a maioria destas variáveis. As estimativas de parâmetros genéticos obtidas indicam que ganhos com a seleção podem ser obtidos para as variáveis rendimento da massa seca de panículas e °Brix. O procedimento Ward-MLM é útil para detectar divergência genética e agrupar genótipos pelo uso simultâneo de descritores morfológicos, agrônomicos e moleculares. Além disso, este método mostrou que três foi o número ideal de grupos, conforme os critérios do pseudo-F e pseudo-t<sup>2</sup>, identificando os grupos mais divergentes. Estes grupos, quando cruzados, apresentam maior probabilidade gerar genótipos com alta produtividade e variabilidade.

**PALAVRAS-CHAVE:** *Sorghum bicolor* L. Diversidade genética. Modelos mistos. Variáveis canônicas.

---

## REFERENCES

- ANP - Agência Nacional do Petróleo, Gás Natural e Biocombustíveis. **Anuário estatístico brasileiro do petróleo, gás natural e biocombustíveis 2016**. Disponível em: [http://www.anp.gov.br/wwwanp/images/publicacoes/Anuario\\_Estatistico\\_ANP\\_2016.pdf](http://www.anp.gov.br/wwwanp/images/publicacoes/Anuario_Estatistico_ANP_2016.pdf). Acesso em: 14/02/2017.
- BRASILEIRO, B. P. MARINHO, C. D.; COSTA, P. M.; MOREIRA, E. F.; PETERNELLI, L. A.; BARBOSA, M. H. Genetic diversity in sugarcane varieties in Brazil based on the Ward-Modified Location Model clustering strategy. **Genetics and Molecular Research**, Ribeirão Preto, v. 13, n. 1, p. 1650-1660, 2014. <https://doi.org/10.4238/2014.January.17.2>
- CABRAL, P. D. S.; SOARES, T. C. B.; GONÇALVES, L. S. A.; AMARAL JÚNIOR, A. T.; LIMA, A. B. P.; RODRIGUES, R.; MATTA, F. P. Quantification of the diversity among common bean accessions using Ward-MLM strategy. **Pesquisa Agropecuária Brasileira**, Brasília, v. 45, n. 10, p. 1124-1132, 2010. <https://doi.org/10.1590/S0100-204X2010001000011>
- COIMBRA, R. R.; MIRANDA, G. V.; CRUZ, C. D.; SILVA, D. J. H.; VILELA, R. A. Development of a Brazilian maize core collection. **Genetics and Molecular Biology**, São Paulo, v. 32, n. 3, p. 538-545, 2009. <https://doi.org/10.1590/S1415-47572009005000059>
- CREMONEZ, P. A.; FEROLDI, M.; NADALETI, W. C.; ROSSI, E.; FEIDEN, A.; CAMARGO, M. P.; CREMONEZ, F. E.; KLAJN, F. F. Biodiesel production in Brazil: current scenario and perspectives. **Renewable and Sustainable Energy Reviews**, Amsterdã, v. 42, p. 415-428, 2015.
- CRUZ, C. D.; REGAZZI, A. J., CARNEIRO, P. C. S. (2012) **Métodos biométricos aplicados ao melhoramento genético**. 3 ed. Viçosa: Editora UFV, 2014.
- FRANCO, J.; CROSSA, J.; VILLASEÑOR, J.; TABA, S.; EBERHART, S.A. Classifying genetic resources by categorical and continuous variables. **Crop Science**, Madison, v.38, n.6, p.1688-1696, 1998. <https://doi.org/10.2135/cropsci1998.0011183X003800060045x>
- KUMAR, Y.; RINGENBERG, J.; DEPURU, S. S.; DEVABHAKTUNI, V. K.; LEE, J. W.; NIKOLAIDIS, E.; ANDERSEN, B.; AFJEH, A. Wind energy: Trends and enabling technologies. **Renewable and Sustainable Energy Reviews**, Amsterdã, v. 53, p. 209-224, 2016.
- OLIVEIRA, R. S.; SILVA, S. A.; BRASILEIRO, B. P.; MEDEIROS, E. P.; ANJOS, E. V. A. Genetic divergence on castor bean using the ward-mlm strategy. **Revista Ciência Agronômica**, Fortaleza, v. 44, n. 3, p. 564-570, 2013. <https://doi.org/10.1590/S1806-66902013000300019>
- PEREIRA FILHO, I. A.; PARRELLA, R. A. C.; ALOÍSIO, J. Avaliação de cultivares de sorgo sacarino [*Sorghum bicolor* (L.) MOENCH] em diferentes densidades de semeadura visando a características importantes na produção de etanol. **Revista Brasileira de Milho e Sorgo**, Sete Lagoas, v. 12, n. 2, p. 118-127, 2013.

PIMENTEL, A. J. B.; GUIMARÃES, J. F. R.; SOUZA, M. A.; RESENDE, M. D. V.; MOURA, L. M.; ROCHA, J. R. A. S. C.; RIBEIRO, G. Estimação de parâmetros genéticos e predição de valor genético aditivo de trigo utilizando modelos mistos. **Pesquisa Agropecuária Brasileira**, Brasília, v. 49, p. 882-890, 2014. <https://doi.org/10.1590/S0100-204X2014001100007>

RESENDE, M. D. V.; DUARTE, J. B. Precisão e controle de qualidade em experimentos de avaliação de cultivares. **Pesquisa Agropecuária Tropical**, Goiânia, v. 37, n. 3, p. 182-194, 2007.

RIBEIRO, L. P. TEODORO, P. E.; LOPES, M. C.; SANTOS, A.; CORRÊA, C. C.; TORRES, F. E.; SILVA, F. A. Quantifying the genetic divergence among maize hybrids using Ward-MLM strategy. **Revista Brasileira de Milho e Sorgo**, Sete Lagoas, v. 14, n. 3, p. 400-408, 2015.

RODRIGUES, W. P.; VIEIRA, H. D.; TEODORO, P. E.; PARTELLI, F. L.; BARBOSA, D. H. **Genetics and molecular Research**, Ribeirão Preto, v. 15, n. 2, 2016.

SAS INSTITUTE. **SAS/STAT**: user's Guide. Version 9.2. Cary: SAS Institute, 2009. 240p.

SILVA, F. H. L.; VIANA A. P.; FERREIRA, R. T.; FREITAS, G. C. O.; SANTOS, J. O.; RODRIGUES, D. L. Measurement of genetic diversity in progenies of sour passion fruit by ward-mlm methodology: a strategy for heterotic group formation. **Ciência e Agrotecnologia**, Lavras, v. 38, n. 3, p. 240-246, 2014. <https://doi.org/10.1590/S1413-70542014000300003>

TEODORO, P. E.; TORRES, F. E.; CORRÊA, A. M. Clustering of soybean genotypes via Ward-MLM and ANNs associated with mixed models. **Genetics and molecular Research**, Ribeirão Preto, v. 15, n. 3, 2016.

VENCOVSKY, R.; BARRIGA, P. **Genética biométrica no fitomelhoramento**. Ribeirão Preto: Sociedade Brasileira de Genética, 1992. 496p.