Selection of sugar cane families by using BLUP and multi-diverse analyses for planting in the Brazilian savannah

M.H.P. Barbosa[1], A. Ferreira[2], L.A. Peixoto[3], M.D.V. Resende[4], M. Nascimento[5] and F.F. Silva[6]

[1]Departamento de Fitotecnia, Universidade Federal de Viçosa,
Viçosa, MG, Brasil
[2]Universidade Federal do Espírito Santo, Centro Agropecuário,
Alegre, ES, Brasil
[3]Departamento de Biologia Geral, Universidade Federal de Viçosa,
Viçosa, MG, Brasil
[4]Empresa Brasileira de Pesquisa Agropecuária,
Centro Nacional de Pesquisa de Florestas, Colombo, PR, Brasil
[5]Departamento de Estatística, Universidade Federal de Viçosa,
Viçosa, MG, Brasil
[6]Departamento de Zootecnia, Universidade Federal de Viçosa, Viçosa,
MG, Brasil

Corresponding author: M.H.P. Barbosa
E-mail: marciobarbosaufv@gmail.com

ABSTRACT. This study evaluated different strategies to select sugar cane families and obtain clones adapted to the conditions of the Brazilian savannah. Specifically, 7 experiments were conducted, with 10 full sib families, and 2 witnesses in common to all experiments, in each experiment. The plants were grown in random blocks, with witnesses in common (incomplete blocks), and 6 repetitions of each experiment. The data were analyzed through the methodology of mixed patterns, in which

the matrices of kinship between the families were identified by the method of restricted maximum likelihood. The characteristics that were evaluated included soluble solids content (BRIX), BRIX ton/ha, average mass of a culm, number of culms/m, and tons of culms/ha. A multi-diverse alternative based on the analysis of groupings by using the UPGMA method was used to identify the most viable families for selection, when considering the genotypic effects on all characteristics. This method appeared suitable for the selection of families, with 5 family groups being formed. The families that formed Group 2 appeared superior to all other families for all the evaluated characteristics. It is recommended that the families in Group 2 are preferentially used in sugar cane improvement programs to obtain varieties optimally adapted to the conditions of the Brazilian savannah.

Key words: *Saccharum*; Multi-diverse analysis; Genetic improvement

INTRODUCTION

Brazil is one of the biggest sugar cane producers worldwide, with this crop being farmed on 8.5 millions of hectares (Conab, 2012). The outstanding performance of this crop in the country is mostly because of the number of projects focused on the development of genetically improved crops manipulated to optimally grow in several sugar cane areas in Brazil (Rosse et al., 2002).

Family selection is based in the phenotype mean (means of the individuals that compose the block). Some authors have already showed that the use of family selection contributes to parental selection, in addition to mating, which aim to obtain improved populations (Kimbeng and Cox, 2003; Shanthi et al., 2008). The strategy of family selection has been previously used in some sugar cane improvement projects in Brazil (Leite et al., 2009; Pedroso et al., 2011).

According to Kimbeng and Cox (2003), adoption of family selection produced positive results in sugar cane genetic improvement projects, because it reveals important information that helps to determine the genotypic value of mating. This is achieved by identifying genotypes and elite clones with the potential for new crossbreeding programs. High influence in the environment in sugar cane genotypes evaluate in the early stages of breeding program is responsible for the low heritability estimates found in these stages, difficult individual selection (Oliveira et al., 2008). For this reason, selection among and within families represents an important tool for the improvement of sugar cane cultivation. This strategy enhances the genetic gain compared to the individual selection, particularly during the first stages of sugar cane improvement (Shanthi et al., 2008).

The sugar cane cultivation areas in Brazil are being continuously expanded into regions where this type of cultivation did not previously exist, such as the Brazilian savannah region. Because of interest in reducing human dependence on the fossil fuels, along with concerns about the global environment, the use of the ethanol is increasing, stimulating the expansion of sugar cane cultivation. It is predicted that Brazilian mills will make 602 million of tons of sugar cane in the 2012/13 harvest, producing 23.96 billion liters of ethanol and 38.85 million tons of sugar (Conab, 2012). To produce satisfactory sugar cane products, it is necessary to obtain plants adapted to the region.

Some methods have been regularly used for sugar cane improvement, such as mass

selection (Oliveira et al., 2011), sequence selection (selection between families followed by mass selection (Kimbeng and Cox, 2003), and simulated individual BLUP (BLUPIS) (Resende and Barbosa, 2006). However, none of the projects focused on cultivation of economic interest provides evidence about group formations (grouping analysis), which would facilitate the selection of families when considering many characteristics simultaneously. Therefore, within the context related to the project for genetic sugar cane improvement of UFV (Rede Interuniversitária para o Desenvolvimento do Setor Sugroenergético - RIDESA), this study aimed to select groups of families and clones that are adapted to Brazilian savannah conditions.

MATERIAL AND METHODS

The experiments were conducted in the experimental area of the "Destilaria WD Ltda" company, in the municipality of João Pinheiro (17°44'34"S latitude; 46°10'31"W), Minas Gerais, during the 2008/2009 harvest. The experimental groups were organized in random blocks. Sugar cane was planted in 7 experiments, with each experiment containing 10 ful sib families and 2 witnesses (these witnesses were tested in 7 experiments) (Table 1). Each experiment was repeated 6 times. Each block had 2 grooves, containing 10 plants each. Each family was represented by 120 genotypes (20 genotypes per block), of which the seedlings were obtained from seeds previously produced in the Sugarcane Flowering and Crossing Station of Serra do Ouro of Rede Interuniversitária para o Desenvolvimento do Setor Sugroenergético - RIDESA (Barbosa et al., 2002). Witnesses used in this experiment were 2 sugar cane varieties, RB867515 and RB855156. In this instance, in each block, 2 buds were planted together to form 1 plant. The spacing between the grooves and plants was 1.40 m and 0.6 m, respectively.

Several characteristics were evaluated and soluble solid content (BRIX) was quantified at the 5th internode from the bottom up to 10 culms, and sampled at random, following the methodology defined by Barbosa (2001). Other measures included tons of BRIX/ha (TBH), average mass of a culm (MMC), number of culms/m (NCM), and tons of culm/ha (TCH). The MMC was obtained by weighing 20 culms randomly collected within the plot.

The following mixed model was adjusted for each characteristic:

$$y = Xr + Za + Wc + Sb + e \qquad \text{(Equation 1)}$$

where y is the vector of phenotype observations; r is the vector of the experimental effects (considered as fixed), including the general average; a is the vector of the individual additive genotype random effects, $\sim N(0, A\sigma^2_a)$; c is the vector of dominant random effects of the full sib family (specific capacity of combination – cec), $c \sim N(0, I\sigma^2_c)$; b is the vector of random effects of blocks in all of the experiments, $b \sim N(0, I\sigma^2_b)$; and e is the vector of random errors, and $\sim N(0, I\sigma^2_e)$.

The matrices of incidence for the effects r, a, c, and b were represented by X, Z, W, and S, respectively. Hence, the equation system of the mixed models is presented as:

$$\begin{bmatrix} X'X & X'Z & X'W & X'S \\ Z'X & Z'Z + A^{-1}\dfrac{\sigma^2_e}{\sigma^2_a} & Z'W & Z'S \\ W'X & W'Z & W'W + I\dfrac{\sigma^2_e}{\sigma^2_c} & W'S \\ S'X & S'Z & S'W & S'S + I\dfrac{\sigma^2_e}{\sigma^2_b} \end{bmatrix} \begin{bmatrix} \hat{r} \\ \hat{a} \\ \hat{c} \\ \hat{b} \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \\ W'y \\ S'y \end{bmatrix}$$

Table 1. Family grouping (G) by the UPMGA method.

G	Genitors	G	Genitors
1	TUC71-70 x RB92579	3	RB952692 x RB867515
1	SP81-3250 x RB92579	4	SP80-1816 x RB92579
1	RB92606 x RB92579	4	RB946022 x RB92579
1	RB941019 x RB93509	4	RB941570 x RB945961
1	RB92579 x RB9371	4	RB952973 x RB957689
1	SP933073 x RB955970	4	RB92579 x RB961
1	RB943161 x RB925211	4	RB953265 x RB931011
1	RB925211 x LAICA9611	4	RB863129 x RB99710
1	RB92606 x RB92579	4	RB98710 x SP932847
2	RB92579 x RB93509	4	RB815181 x F150
2	RB931556 x IAC87-3396	4	RB92579 x SP801816
2	RB951536 x RB896342	4	RB946900 x RB977619
2	H64-1881 x RB9371	4	RB941570 x RB925345
2	RB92579 x RB98710	4	RB92579 x SP813250
2	RB92579 x RB925211	4	RB966920 x RB955970
2	SP79-1011 x RB92579	4	SP89-1115 x RB925211
2	RB981810 x RB925211	5	RB855511 x RB72199
2	RB92606 x RB855511	5	IAC86-2210 x RB93509
2	RB9371 x RB92579	5	UFAL011520 x RB92579
2	RB92579 x SP86-155	5	RB92579 x TUC717
2	UFAL011535 x RB92579	5	RB946016 x RB957610
2	RB9629 x RB93509	5	RB931013 x LAICA01-615
2	RB931003 x RB92579	5	IAC86-2210 x RB867515
2	RB736018 x RB845210	5	RB946022 x RB925345
2	RB951557 x RB965911	5	RB949 x RB955970
2	RB931530 x RB92579	5	RB952940 x RB957689
2	RB92606 x SP832847	5	H64-1881 x RB867515
2	SP832847 x RB9344	5	RB966920 x RB945961
2	RB92579 x CO62175	5	RB971509 x RB835486
2	RB965911 x IAC87-3396	5	RB945955 x RB955970
2	RB9536 x RB867515	5	RB855156 x RB997751
3	RB946016 x RB955970	5	RB92579 x Q107
3	RB75126 x RB98710	5	RB92606 x UFAL01-1520
3	RB971537 x RB965911	5	RB98710 x RB92579
3	RB92579 x RB845210	5	RB941570 x RB92579

The adjustment of the model presented was completed by the Restricted Maximum Likelihood (REML) with the Selegen-Reml/Blup software (Resende, 2007). From the genotypic values of the families was estimated BLUPIS (BLUP individual simulated) that determined the number of individuals that were selected for family (Resende and Barbosa, 2006). The clones selected were planted in experiments for the sequence of clonal selection of the genetic improvement project.

The estimator of the variance component of dominance between families is represented by:

$$\hat{\sigma}_c^2 = \hat{V}_{cec} = (1/4)\hat{\sigma}_d^2 , \qquad \text{(Equation 2)}$$

This equation is the same as the variance component associated to the specific capacity of combination. In this case, we overlooked that epistasis $\hat{V}_{cec}$ is equivalent to (1/4) for the genetic variance of total dominance present in the population.

The genotypic effects of the families were predicted by:

$$(1/2)(\hat{a}_i + \hat{a}_j) + \hat{c}_{ij} , \qquad \text{(Equation 3)}$$

where $\hat{a}_i$ and $\hat{a}_j$ are the additive genetic values predicted by the genitor i and j, respectively; and $\hat{c}_{ij}$ is the specific capacity of the mating combination between genitors i and j.

Aiming at the joint recommendation of families, a cluster analysis was performed using the family genotypic values of all traits (BRIX, TBH, MMC, NCM, and TCH). To group the families, unweighted pair-group method using arithmetic averages (UPGMA) was used. In this method, the distance between an individual k in a group, formed by the individuals i and j, is presented as:

$$d_{(ij)k} = \frac{d_{ik} + d_{jk}}{2}$$ (Equation 4)

i.e., the average of the set distance of the pair of the individuals.

A large number of the groups (divisions) used for the hierarchical grouping method was obtained through the method proposed by Mojena (1977), in which the number of groups is provided at the first stage in the dendrogram, where $\alpha_j > \alpha + \Phi S_\alpha$, in which j = 1,2, ..., n, α_j is the value of the distance for the joint stage corresponding n-j+1 groups; α and S_α are the average and the standard deviation of α; and Φ is the constant. In this study, the value of Φ was 1.25, based on the suggestion of Milligan and Cooper (1985).

RESULTS AND DISCUSSION

The high values of the estimative of additive variance were verified for the characteristics of NCM and TCH (Table 2). The restricted heritability estimate and the family heritability for the characteristics of TBH and MMC was high, producing similar values (Table 2). This result shows that a large portion of the variance was associated to the additive action of the genes. To assess the genetic patterns of sugar cane, Barbosa et al. (2005), aimed to predict the additive of genes value for the selection of the parents and the families based on REML/BLUP analysis in 113 full sib families. The authors showed that the additive genetic effects were predominant for certain characteristics, including the number of culms, average weight of a culm, and ton of sugar cane per hectare (TCH).

In general, the heritability values ranged from 0.3197 for TCH to 0.9389 for MMC; however, that of BRIX was much lower compared to all other values, at 0.1746 (Table 2). These values corroborate the study of Leite et al. (2009), who obtained heritage values of 0.88 for TBH and 0.66 for BRIX. In contrast, Pedroso et al. (2011) obtained 0.48 and 0.30 for TBH and BRIX, respectively. The observed range in heritability obtained by these studies might have been caused by several factors. Examples include differences in the study populations, environmental variance, and experimental delineation. The values obtained in the current study indicate that the selection between superior families is efficient during the initial phases of a sugar cane improvement project, because of the high heritage that the characteristics studied showed.

Families with high genotypic values, superior to the controls, were found for all traits. In addition, the exploitation of the genetic variability within crosses is likely to provide very promising new clones.

**Table 2.** Variance components for characteristics.

Components	BRIX	TBH	MMC	NCM	TCH
V_a	0.3134	0.1686	0.0331	1.5438	125.1891
CVa%	2.9350	18.7921	19.0095	12.0534	15.7722
V_b	0.2919	0.0112	0.0010	0.0252	6.0703
V_{cec}	0.2157	0.0379	0.0044	0.7034	81.8623
V_e	0.9743	0.0107	-0.0033	2.2653	178.5050
V_f	1.7953	0.2284	0.0353	4.5376	391.6267
h^2_a	0.1746	0.7383	0.9389	0.3402	0.3197
R^2_b	0.1626	0.0489	0.0282	0.0055	0.0155
R^2_{cec}	0.1202	0.1659	0.1255	0.1550	0.2090
h^2_{fam}	0.6640	0.8853	0.9045	0.7445	0.7824
General Average	19.0749	2.1853	0.9576	10.3081	70.9401

TBH = tons of brix by hectare; MMC = average mass of a culm; NCM = number of culms by meter; and TCH = tons of culms by hectare. V_a = additive genetic variance; V_b = variance between blocks; CVa = coefficient of additive genetic variance; V_{cec} = variance of the specific capacity of combination; V_e = residual variance; V_f = phenotype variance; h^2_a = heritage on the restrict sense of the block, that is, of the additive effects; h^2_{fam} = heritage at the level of the averages of families on the broad sense; R^2_b = coefficient of determination of the block effects; R^2_{cec} = coefficient of determination of the effects of the specific capacity of combination.

The exploration of the superior families might facilitate the use of a greater number of superior clones, which would help advance the phases of sugar cane genetic improvement (Barbosa et al., 2005; Resende and Barbosa, 2005). This fact might contribute toward increasing the efficiency of genetic improvement, through the selection of a small number of clones, which are already known to be superior; thus, reducing the costs of experimentation, in parallel to accelerating the selection phases, because the superior clones may be selected more easily.

Five family groups were formed by the UPGMA method. Group 1 contained 8 families, Group 2 contained 23 families, Group 3 contained 5 families, Group 4 contained 16 families, and Group 5 contained 20 families (Table 1). The way in which families are grouped is important for an improvement project, because it helps define which families are selected. If the goal is to create genetic variability to increase the genetic variance, and consequently increase the gain selection in a new selection cycle, it is advisable to select families from different groups. This procedure is also suitable to obtain heterotic clones for the characteristics that show dominance (like TCH, as shown in Table 2), because of mating between divergent genetic genitors.

The families belonging to Group 2 obtained averages that were superior to the general average for all of the evaluated characteristics (Figure 1). This result confirms the superiority of this group compared to the other groups if a program intends to jointly select characteristics; that is, to mate individuals in which most of the characteristics are above the population average. With this type of selection, superior families selection is guaranteed for one trait, but do not get to select superior families for other traits. For example, this phenomenon may be observed in Group 1, where the average BRIX is above average, but the NCM is much lower than the general average across families. As a further example, if we select the characteristics separately, Group 1 families would be selected for BRIX, but are likely to have unfavorable genes for the NCM compared to other families. By using joint selection, it is easier to obtain superior individuals for all of the evaluated characteristics to compose the desired variety in an improvement project.
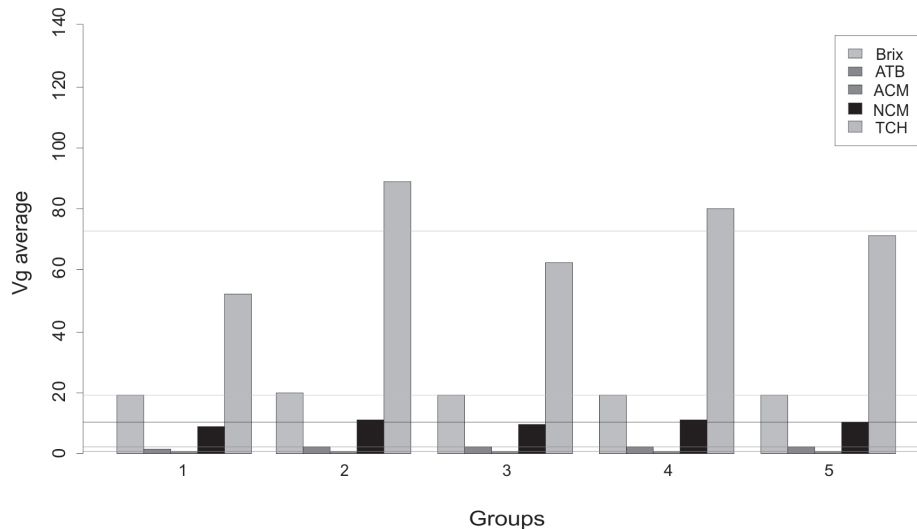
Figure 1. Average genotype value (Vg) of each variance within each group classified by the UPGMA method. TBH = ton of BRIX by hectare; MMC = average mass of the culm; NCM = number of culms by meter; TCH = tons of culms by hectare. The horizontal lines represent average genotype values for each character.

Landell and Silva (2004) observed that the characteristic of number of culms is the most responsible characteristic for the significant effects of TCH. The important of this characteristic was corroborated in this study; whereby, in 30% of the superior families that were evaluated (21 families) for NCM and TCH, there was a coincidence rate of 47.62% for the families that had characteristics.

It is known that phenotypic expression of a particular genotype, subjected to a specific environment, will result in genes responsible for the trait expression and action of genes related with genotype adaptation in environments to which it was subjected. Hence, for these characteristics, the selection of specific materials to the differential environment, such as the Brazilian savannah, is a way of finding the solution to the problem. Therefore, the selection of superior genotypes for specific environment requires efficient strategies to obtain superior varieties.

CONCLUSIONS

All of the characteristics showed additive genetic variability that was sufficient for the continuity of an improvement project for planting sugarcane under savannah conditions. The characteristics also expressed suitable levels of genetic variance of dominance, which allow to explore the specific capacity of combination and the heterose through crossing between divergent parents. It is recommended that multi-variance selection is used to select families, whenever the goal of the improvement project is to gain multiple characteristics simultaneously.

The proposed was proven to be suitable for the identification of an adapted variety, in which 5 family groups were formed. Family group 2 was superior to all other groups for all of the

characteristics evaluated (Group 2 contained the families #RB92579 x RB93509, RB931556 x IAC87-3396, RB951536 x RB896342, H64-1881 x RB9371, RB92579 x RB98710, RB92579 x RB925211, SP79-1011 x RB92579, RB981810 x RB925211, RB92606 x RB855511, RB9371 x RB92579, RB92579 x SP86-155, UFAL011535 x RB92579, RB9629 x RB93509, RB931003 x RB92579, RB736018 x RB845210, RB951557 x RB965911, RB931530 x RB92579, RB92606 x SP832847, SP832847 x RB9344, RB92579 x CO62175, RB965911 x IAC87-3396, RB9536 x RB867515). The families in Group 2 are recommended for use in improvement projects to obtain varieties that are adapted to the conditions of the Brazilian savannah.

REFERENCES

Barbosa MHP (2001). Study of genetic divergence in sugar cane varieties grown in Brazil using the parentage coefficient. *Int. Sugar J.* 103: 294-297.

Barbosa GVS, Cruz MM, Soares L, Rocha AMC, et al. (2002). A brief report on sugarcane breeding program in Alagoas, Brazil. *Crop Breed. Appl. Biotechnol.* 2: 613-616.

Barbosa MHP, Resende MDV, Bressiani JA, Silveira LCI, et al. (2005). Selection of sugarcane families and parents by REML/BLUP. *Crop Breed. Appl. Biotechol.* 5: 443-450.

Conab-Companhia Nacional de abastecimento (2012). Acompanhamento de Safra Brasileira: Cana-de-Açúcar. Primeiro Levantamento, Abril/2012. Conab, Brasília.

Kimbeng CA and Cox MC (2003). Early generation selection of sugarcane families and clones in Australia: a review. *J. Am. Soc. Sugarcane Technologists* 23: 20-39.

Landell MGA and Silva MA (2004). As estratégias de seleção da cana em desenvolvimento no Brasil. *Visão Agrícola* 1: 18-23.

Leite MSO, Peternelli LA, Barbosa MHP, Cecon PR, et al. (2009). Sample size for full-sib family evaluation in sugarcane. *Pesq. Agropec. Bras.* 44: 1562-1574.

Milligan GW and Cooper MC (1985). An Examination of Procedures for Determining the Number of Clusters in a Data Set. *Psychometrika* 50: 159-179.

Mojena R (1977). Hierarchical grouping methods and stopping rules: An evaluation. *Comp. J.* 20: 359-363.

Oliveira RA, Daros E, Bespalhok Filho JC and Zambom JLC (2008). Seleção de famílias de cana-de-açúcar via modelos mixtos. *Sci. Agraria* 9: 269-274.

Oliveira RA, Daros E, Resende MDV, Bespalhok Filho JC, et al. (2011). Procedimento Blupis e seleção massal em cana-de-açúcar. *Bragantia* 70: 796-800.

Pedroso CA, Barbosa MHP, Silva FL, Resende MDV, et al. (2011). Repeatability of full-sib sugarcane families across harvests and the efficiency of early selection. *Euphytica* 182: 423-430.

Resende MDV and Barbosa MHP (2005). Melhoramento Genético de Plantas de Propagação Assexuada. Colombo, Brasília.

Resende MDV and Barbosa MHP (2006). Selection via simulated individual BLUP based on family genotypic effects in sugarcane. *Pesq. Agrop. Bras.* 41: 421-429.

Resende MDV (2007). Selegen-Reml/Blup: Sistema Estatístico e Seleção Genética Computadorizada via Modelos Lineares Mistos. Colombo, Brasília.

Rosse LN, Vencovsky R and Ferreira A (2002). Comparação de métodos de regressão para avaliar a estabilidade fenotípica em cana-de-açúcar. *Pesq. Agrop. Bras.* 37: 25-32.

Shanthi RM, Bhagyalakshmi KV, Hemaprabha G, Alarmelu S, et al. (2008). Relative performance of the sugarcane families in early selection stages. *Sugar Tech.* 10: 114-118.