MELISSA PISAROGLO DE CARVALHO

EFFICIENCY OF AMONG AND WITHIN FAMILY SELECTION IN PLANT BREEDING THROUGH SIMULATION

Thesis presented to the Federal University of Viçosa as partof the requirements for Graduation Program in Genetics and Breeding for obtaining the title of *Doctor Scientiae.*

VIÇOSA

MINAS GERAIS – BRAZIL

2012

Ficha catalográfica preparada pela Seção de Catalogação e Classificação da Biblioteca Central da UFV

T
C331e
2012

Carvalho, Melissa Pisaroglo de, 1975-
Efficiency of among and within family selection in plant breeding through simulation / Melissa Pisaroglo de Carvalho. – Viçosa, MG, 2012.
x, 55f. : il. (algumas col.) ; 29cm.

Inclui apêndice.
Orientador: Luiz Alexandre Peternelli.
Tese (doutorado) - Universidade Federal de Viçosa.
Referências bibliográficas: f. 50-54

1. Cana-de-açúcar – Melhoramento genético. 2. Plantas – Melhoramento genético. 3. *Sccharum afficinarum*. I. Universidade Federal de Viçosa. II. Título.

CDD 22. ed. 633.612

MELISSA PISAROGLO DE CARVALHO

EFFICIENCY OF AMONG AND WITHIN FAMILY SELECTION IN PLANT BREEDING THROUGH SIMULATION

Thesis presented to the Federal University of Viçosa as part of the requirements for Graduation Program in Genetics and Breeding for obtaining the title of Doctor Scientiae.

APPROVED: February, 27th, 2012

Prof. Lindolfo Storck

Pesq. Marcos Deon Vilela Resende
(Co-orientador)

Prof. Américo José dos Santos Reis

Prof. Nerinéia Dalfollo Ribeiro

Prof. Luiz Alexandre Peternelli
(Orientador)

DEDICATION

God for always being by my side regardless of the mistakes or successes;

And to my Father,**Carlos Alberto Marin de Carvalho** (in memoriam) for always helping me along my way.

ACKNOWLEDGMENTS

It is a pleasure to thank the many people who made this thesis possible.

I wish to thank GOD, for helping and supporting me every single day and to give me strength in the best and worst moments.

It is difficult to overstate my gratitude to my PhD Supervisor Dr. Luiz Alexandre Peternelli, with his enthusiasm, his inspiration, his time, his intelligence and his great effort to explain things clearly and simply. Throughout my Doctor Degree period, he provided encouragement, sound advice, good teaching and lots of good ideas. For trust me and give me challenges, also for providing a stimulating and fun environment in that I learned and grew.

I wish to express my warm and sincerely thanks to Professor PhD Salvador A. Gezan, who gave me the opportunity to work with him in the School of Forest Resorce and Conservation at University of Florida. My sincerely thanks for his helping, patience, ideas and good teaching. I would have been lost without him.

I would like to thank Professor Dr. Márcio Henrique Perreira Barbosa, for his help and for giving me all field data to development my thesis.

I am especially grateful to my entire-extend family for providing a loving environmental for me, my sisters Bianca de Carvalho Hahn and her husband Silvio Cesar Hahn, Andrezza Pisaroglo de Carvalho and my brothers Tiago Pisaroglo de Carvalho and Maximiliano Pisaroglo de Carvalho, my cousin Daniela Barakat and her family, my aunt Maria de Lourdes Groth and my uncle Almiro Luis Groth for all them help, support and love.

My special thanks for my sister Andrezza Pisaroglo de Carvalho for all helping and for getting that I did not give up of my dreams.

I owe my loving thanks to Jeffrey Davis, without his encouragement, understanding and for making my life easy, special and funny.

I wish to thanks all my friends that I made in Gainesville, especially to Liz Gregg during a hard and great year.

I am grateful to my friends and my Capoeira group (Cordão de Ouro – CDO, Viçosa) for helping me get through the difficult time and all the emotional support and entertainment.

The financial support of the CNPq is gratefully acknowledged, for supporting my studies with great opportunities through of scholarship to my Doctor degree, visiting scholar and now my future Post- doc.

I am grateful to the secretary Édna and the Programa de Pós-Graduação em Genética e Melhoramento at Universidade Federal de Viçosa, for helping the departments to run smoothly and for assisting me in many different ways

Lastly, and most importantly, I wish to thank my parents, Carlos Alberto Marin de Carvalho (in memorian) and Elida Maria Pisaroglo de Carvalho. They bore me, raised me, supported me, taught me, and loved me. To them I dedicate this thesis

BIOGRAPHY

MELISSA PISAROGLO DE CARVALHO, daughter of Carlos Alberto Marin de Carvalho (in memorian) and Elida Maria Pisaroglo de Carvalho. She has two sisters, Bianca de Carvalho Hahn, Andrezza Pisaroglo de Carvalho and two brothers Tiago Pisaroglo de Carvalho and Maximiliano Pisaroglo de Carvalho.

Born in São Leopoldo, Rio Grande do Sul, on May 11st,1975, left her hometown in September, 2000, when she began to study in Agronomy at Universidade Federal de Santa Maria (UFSM), graduating in July, 2005.

On March, 2006, she began the Programa de Pós Graduação em Produção Vegetal na Universidade Federal de Santa Maria, in Master degree, with dissertation done in February, 2008.

On March of the same year, she began the Programa de Pós Graduação em Genética e Melhoramento na Universidade Federal de Viçosa, in Doctor degree. In September, 15th, 2010, she was realized her Visiting Shcolar at University of Florida, USA, she finished in September 14th, 2011.

After four years of lot of learning, studies and dedication, she finished her Doctor degree in February, 27th, 2012.

ÍNDICE

RESUMO

CARVALHO, Melissa Pisaroglo, D.Sc.,Universidade Federal de Viçosa, fevereiro de 2012. **Eficiência da seleção entre e dentro de famílias no melhoramento de plantas por meio de simulação.** Orientador: Luiz Alexandre Peternelli. Co-orientadores: Márcio Henrique Pereira Barbosa e Marcos Deon Vilela de Resende.

Os objetivos desse trabalho foram realizar um estudo aprofundado dos experimentos do programa de melhoramento da cana-de-açúcar da UFV, estimando os componentes de variância e os parâmetros genéticos. Foram selecionadas as 20 melhores famílias e foi realizada a análise de componentes principais. Foram utilizados os dados de cinco experimentos no delineamento de blocos ao acaso com cinco repetições, 22 famílias de irmãos completos e dois controles. As variáveis analisadas foram: número, altura, diâmetro e peso de colmos e BRIX. Com os resultados desse estudo foram geradas as simulações para obter uma comparação entre os métodos BLUPI e BLUPIS. Os dados foram simulados com base em seis cenários genéticos, posteriormente esses dados foram analisados no contexto de modelo linear misto, via modelo individual para estimar o BLUPI e via modelo em nível de parcela para estimar o BLUPIS. Os valores de coincidência entre o número de indivíduos selecionados por famílias foram comparadas através de valores de correlação via ambos os métodos. O método BLUPIS mostrou se eficiente em todos os cenários, sendo que os melhores valores foram obtidos para os cenários com altos valores de variância aditiva e herdabilidade.Altos valores de correlação foram encontrados entre os métodos BLUPI e BLUPIS, principalmente para os cenários com altos valores de variância aditiva e herdabilidade.

ABSTRACT

CARVALHO, Melissa Pisaroglo, D.Sc.,Universidade Federal de Viçosa, February, 2012. **Efficiency of among and within family selection in plant breeding through simulation.** Adviser: Luiz Alexandre Peternelli. Co-Advisers: Márcio Henrique Pereira Barbosa and Marcos Deon Vilela de Resende.

The objective of this work was to conduct in depth study of the experiments in the breeding program of sugarcane, estimating the variance components and genetic parameters. It was selected 20 families and with this was made the principal component analysis. The study used data from five experiments in a randomized block design with five replicates, 22 full-sib families and two controls. The variables analyzed were: number, height, diameter and weight of stems and BRIX. The results of this study were used to generate simulations for a comparison between the methods BLUPI and BLUPIS. The data were simulated based on six genetic scenarios, the data were subsequently analyzed in the context of the linear mixed model, via individual model to estimate the BLUPI and plot mean model to estimate the proportion of selected by BLUPIS. The values of coincidence between the numbers of individuals selected by families via the two methods were compared by correlation values. The BLUPIS method showed to be efficient in all scenarios, but the best values were obtained for scenarios with high values of additive variance and heritability. Higher values of correlation between the BLUPI and BLUPIS selection methods were found, mainly for the scenarios with high additive variance component and heritability.

INTRODUCTION

Sugarcane is a preferably alogamous plant and when grown commercially is asexually propagated (Matsuoka et al., 1999). It is a C_4 plant with high photosynthetic efficiency and capacity for growing in regions of high temperatures (Machado et al., 1982; Taiz and Zeiger, 2004).

For sugarcane, the common breeding strategy has been the production of cultivars that are asexually propagated based on selection of offspring from crosses among superior parents.

Selection aims at inducing a change in the genotypic average of the population in the desired direction, through a change in allele frequencies in loci that control the traits in the selection process seeking to meet the main objectives of genetic improvement.

All breeding program want to selected families and elites clones in a short time, sometimes it can occur death of plants and wrong measurements, a tool that can help in all situation above are stochastic simulation. Other toll that also can help the sugarcane and other breeding programs are mixed models via REML/BLUP (maximum likelihood/ best linear unbias prediction with two selection methods, Individual BLUP (BLUPI) and Simulated Individual BLUP (BLUPIS). The BLUPI selected individuals inside the family, for this it need to have individual information per plot, in other hand, BLUPIS estimate the number of individuals within the family that to have to selection, all the measured for BLUPIS are in family level (plot means).

Therefore, the chapter one estimate the variance components and their genetics parameters and the chapter two compared two selection methodologies (BLUPI and BLUPIS) using simulation through linear mixed models (REML/BLUP).

CHAPTER 1

GENETIC ANALYSES OF SUGARCANE TRIALS IN A BREEDING PROGRAM

RESUMO

As variedades comerciais de cana de açúcar são um complexo poliplóide. A heterogeneidade e a poliploidia são grandes fontes de variabilidade genética. Os objetivos desse trabalho são analisar dados de cinco experimentos de cana-de-açúcar, através da analise individual e também múltipla usando a análise de modelos lineares mistos (REML/BLUP); estimar e interpretar os parâmetros genéticos, como os componentes de variância aditiva e dominância, herdabilidades; e realizar a seleção das 20 famílias mais produtivas usando a melhor predição linear não viciada (BLUPs) e para essas famílias realizar a analise de componentes. Cinco experimentos foram instalados, no delineamento de blocos ao acaso com cinco repetições e 24 tratamentos, sendo 22 famílias de irmãos completos e dois controles (Variedades comerciais). Os controles foram desconsiderados nessas analises. As variáveis analisadas foram: número, peso, diâmetro e altura de colmos e BRIX. Os modelos estatísticos utilizados foram a nível de família e a nível de pais. O modelo de família foi criado para as análises individuais e o a nível de pais para a analise múltipla, o software estatístico utilizado foi o ASReml v.3 e o SELEGEN – REML/BLUP. Com base nos resultados encontrados nas análises individuais podemos considerar as variáveis altura de colmos, BRIX e diâmetro de colmos as mais importantes para o programa de melhoramento, visto que podem apresentar os maiores ganhos. Nos resultados da análise múltipla, interessante efeito significativo não aditivo foi encontrado para peso, diâmetro e altura de colmos. A variável altura de colmos pode ser recomendada para a seleção de famílias e de indivíduos devido aos seus valores de herdabilidade. A seleção das melhores 20

famílias separou-as em dois grupos de acordo com os valores médios das famílias e análise de componentes principais comprovou esses resultados de acordo com as estruturas morfológicas e fisiológicas das famílias.

ABSTRACT

The commercial varieties of sugarcane are a polyploidy complex. The heterozygous and polyploidy are great source of genetic variability. The objectives of this work are to analyze field data from five experiments in sugarcane, through single and multiple-sites analyses using linear mixed models (REML/BLUP), estimate and interpreted genetic parameters, such as additive, dominance variance component, heritabilities and to select the top 20 families using the best linear predictions (BLUPs) and for these families make the principal component analysis (PCA), to better understand the physiological and morphological characteristics. Five experiments were installed, on randomized complete block design with five repetitions and 24 treatments, 22 full-sib families and two controls. The controls were not considered in this work. These traits evaluated were: number (NS), weight (WS), diameter (DS) and height (HS) of stalks and juice percentage of solid soluble (BRIX). Two statistical models were fitted for this analyses, the family level and the parental level. The family model was used for the single site analyses and the parental model for the multiple site analyses. The ASReml v.3 and SELEGEN REML/BLUP statistical package were used. With the results of single site analyses might consider the variables height of stalks, BRIX and diameter of stalks the most important variable for the breeding program since they show higher gains. In the multiple site results interesting significant non additive effect was found for weight, diameter and height of stalks. The variable height of stalks might be recommending for family and individual selection due it heritability values. The selected top 20 families

separated in two groups according with the family average and the principal component analyses proven this group through the morphological and physiological structure.

1. INTRODUCTION

Sugarcane (*Saccharum officinarum* L.) has an important role in the economy due to its multiple uses. This crop can be used fresh, in the form of fodder for animal feed, or as a raw material, for manufacture of brown sugar, molasses, sugar and ethanol.

Sugarcane is grown in the tropics and subtropics, for example, Indian, China, Australia, Mexico, United States and Brazil. In the United States, sugarcane is grown commercially in Florida, Louisiana, Texas and Hawaii. In 2010, sugarcane crop totaled nearly 27.9 million tons with a value of more than US$ 991 million (NASS, 2011). In Brazil, the sugarcane production was 623 million ton per hectare. The total of 27 Brazilian states, only five states do not produce sugarcane, where São Paulo is the main producer with 54% of production (CONAB, 2011).

The commercial varieties of sugarcane are a polyploidy complex. The heterozygous and polyploidy of this crop is a source of great present and future genetic variability. The knowledge of the nature and magnitude of this variability is of relevant importance for breeding programs. Parameters, such as phenotypic and genetic variation and its partition into additive, dominance, and epistasis components, narrow and broad sense heritability, and magnitudes of genotype by environment interactions are relevant for some breeding programs, such as of sugarcane (Tyagi and Singh, 1998; Oliveira, 2007), as they allow for the distribution of the crossing and propagation efforts, and planning for given expected genetic gains. Whenever sufficient knowledge is available from a crop, including germplasm, experimental plan, measurements, and selection

methods, it is possible to obtain good estimates of the genetic variance and covariance and heritability (Dudley and Moll, 1969).

The success in improving sugarcane is related to the knowledge of the genetic properties necessary for selection and genetic aspects from several traits. Knowledge of genetic variability, that can represent the potential of the population for the selection (Ramalho et al., 2004), and heritability are of great importance for the estimation of the genetic gain, success in the selection, development and deployment of new cultivars (Oliveira et al., 2010). For example, the simplest selection method can be used when the heritability is high, or more sophisticated methods are necessary when it is low, (Vencovsky e Barriga, 1992).

The most important genetic parameters to help the choice of the population and its method of selection are: the additive and dominance variance, heritability and genetic correlations. The knowledge of additive variance values allows the breeder to estimate the expected progress per cycle of selection. If the additive and dominance variances are known, one can obtain the relationship between them, which reflects the knowledge of the type of allelic interaction in controlling the character (Ramalho,1977).

The genetic value of each parent is measured through the additive effects or breeding values by evaluating their offspring (Falconer & Mackay, 1996). The non-additive effects correspond to the dominance and epistasis. The dominance corresponds to the interaction of alleles at a particular locus. In contrast, epistasis effects result from the interaction between alleles at different loci.

An important element to estimate the above genetic components is the use, and knowledge, of the pedigree. In addition, the pedigree information of a cultivar may be useful for the identification and understanding of source of genes of interest and its role in genetic variability (Martin et al., 2007), in aspects such as control of relatedness and inbreeding.

In most of trials that used family selection, different number of seedlings per family can occur, generating an unbalanced experiment. In this situation, the best random prediction of variables, which assumed that the variance component should be estimate with high accuracy, employing the standard procedure in the context of mixed models, is restricted maximum likelihood. (Henderson, 1984; Searle et al., 1992) (REML). According with Searle et al (1992), this procedure allow the selection of individual with higher genetic values, independent of it provenance (Simeão et al.,2002). Some package that use mixed models via REML/BLUP are ASReml (Gilmour et al., 2009), Selegen – Reml/Blup (Resende, 2007), R (R Development Core Team, 2010) and SAS (Littell et al., 2006).

In sugarcane, some researchers, referring to the sexual phase, reported an individual narrow-sense heritability of 0.57 (Bressiani et al, 2007) for disease resistance, and for brix these values were 0.17 (Babu, 1990) and 0.18 (Wu et al., 1989).Results found by Hogarth et al. (1981) showed that the additive genetic variance is superior for brix, and for number, diameter and height of stalks. Others (Ferreira, 2007; Chaudhary, 2001; Bressiani, 2001; Oliveira, 2007) reported that for the variable brix the gene action is mainly additive, and according to Barbosa and Silveira (2012) and Oliveira (2007) this is the case for most economically relevant traits such as content of sugar, fiber and disease resistance. An exception is sugarcane yield where the non-additive variance is more relevant than additive variance (Bastos et al.,2003). They related also the no-additive genetic variance is significant for all character, except for brix and number of stalks.

The objectives of this chapter are to: 1) analyze field data from five experiments in sugarcane through single and multiple-sites analyses using linear mixed models (REML/BLUP); 2) estimate and interpret genetic parameters, such as additive, dominance variance components, heritabilities, etc.; and 3) to select the top 20 families

using the best linear unbiased predictions (BLUPs) and for these families make the principal component analysis (PCA) to better understand the physiological and morphological characteristics.

2. MATERIALS AND METHODS

2.1. Data set

The dataset used for this work originates from the Programa de Melhoramento Genético da Cana de Açúcar (PMGCA) directed by the Federal University of Viçosa, Minas Gerais, Brazil. Data from a total of five trials were available. These trials were carried out at a breeding center at Oratórios, Minas Gerais, Brazil, located at latitude 20°25'S and longitude 42°48'W, with an altitude of 494m, an average annual rainfall of 1.250mm, an UR 64.7%, and an average minimum and maximum annual temperature of19.5 ºC and 21.8 ºC, respectively (Silva, 2009).

The trials were installed on May 2007; each consisted on a randomized complete block design (RCBD) with five blocks and 24 treatments, corresponding to a grid of 20 by 30 plots (APPENDIX – Figure I).Each experimental plot consisted of two lines spaced 1.40m with 10 plants each, equidistant at 0.5m. Each trial was composed of 22diferentfull-sib families and two controls (RB72454 and RB855046, two of the commercial varieties used in Brazil). The majority of the parents are represented in all five trials (Table 3); however, full-sib families are represented only in a single trial. The details of the collection and seedling planting procedures are described by Barbosa and Silveira (2000).Fertilization was conducted in accordance with the recommendation for the crop in the region given by Macedo et al. (2010).

The phenotypic traits evaluated were: number (NS) and weight (WS, $kg.h^{-1}$) of stalks, juice percentage of soluble solids (BRIX, %), and diameter (DS, mm), and height (HS, m) of stalks. For each plant, NS was measured by counting the number of stalks

longer than 1 m, and WS was obtained by weighting the complete plant and dividing by its NS only on those plants with more than five leaves. For BRIX, a single randomly selected stalk for each of the 20 plants in the plot was determined by refractometry, DS was measured using a digital caliper at ground level, and HS correspond to the height from ground level until dewlap. A summary of the trials and their measurements are presented in Table 1 and 2.

Table 1: Summary of genotypes by trial, considering a total of five trials, Oratórios, 2009.

	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5	Total
Number of individuals	1390	1574	1790	1213	1464	7431
Number of mothers	19	21	19	18	18	95
Number of fathers	22	21	19	17	19	98
Number of families	22	22	22	22	22	110

Table 2: Average of trial for traits: number (NS), weight (WS, kg) of stalks, juice percentage of soluble solids (BRIX,%), diameter (DS, mm) and height (HS, m) of stalks of 22 families per trial, Oratórios, 2009.

	Average				
Trial	NS	WS	BRIX	DS	HS
1	8.60	12.12	19.42	25.18	2.49
2	7.52	10.89	19.58	25.40	2.48
3	7.86	10.61	19.98	25.26	2.54
4	8.52	11.89	19.96	25.40	2.58
5	8.30	11.77	20.07	25.18	2.62

2.2. Single-site analyses

After inspecting the data to detect inconsistencies and outliers, each trials was individually (i.e. single-site) analyzed using the statistical program ASReml v.3 (Gilmour et al., 2009) based on the following linear mixed model:

$$y_{ijk} = \mu + fam_i + b_j + p_{ij} + e_{ijk} \quad (1)$$

where, y_{ijk} corresponds to the observation belonging to the j^{th} block, ij^{th} plot, and i^{th} family; μ is the fixed overall mean; fam_i is the random effect of the i^{th} family, i=1,…,22,$fam_i \sim N(0,\sigma^2_{fam})$; b_j is the fixed effect of the j^{th} block, j = 1,…,5; p_{ij} is the random effect of the ij^{th} plot within the j^{th} block, ij=1,…,22, $p_{ij} \sim N(0,\sigma^2_p)$; and e_{ijk} is the random error effect, $e_{ijk} \sim N(0, \sigma^2)$. All random effects, including the error term, were considered independent and identically distributed. For this model, parental information (and pedigree) is ignored and controls were omitted from the dataset.

The total variance and cross heritability were obtained using the following expressions:

Total variance: $V_T = \sigma^2_p + \sigma^2_{fam} + \sigma^2$

Heritability of cross: $h^2_{cross} = \frac{\sigma^2_{fam}}{V_T}$

Where σ^2_{fam}corresponds to the variability of family means, which corresponds to the following expression of the genetic variance components:

$$\sigma^2_{fam} = \frac{V_a}{2} + \frac{V_d}{4}$$

Here, V_a and V_d correspond to the additive and dominance variance, respectively. Therefore, for the above model it is not possible to partition the additive from the dominance components.

2.3. Multiple-site analyses

For the combined analyses considering all trials for a given trait, two linear mixed models were fitted. The first model called *family model*, as it only considers the family random effect, is equivalent to the one fitted for single-site analyses. On the other hand, the second model fitted partitioned the family genetic components into additive and dominance, and will be called *parental model*.

The family model was fitted using the following linear mixed models:

$$\boldsymbol{y = 1\mu + X\,s + Z_1\,b + Z_2\,p + Z_3\,fam + e} \quad (2)$$

where, $\boldsymbol{y}$ is the vector of observation for all trials; $\boldsymbol{\mu}$ is a constant representing the overall mean, $\boldsymbol{s}$ is the vector of trials fixed effects; $\boldsymbol{b}$ is the vector of random block effects within trial, with $\boldsymbol{b}$ ~ MVN ($\mathbf{0}$,$\sigma_b^2\mathbf{I}_b$); $\boldsymbol{p}$ is the vector of random plot effects, with $\boldsymbol{p}$ ~ MVN ($\mathbf{0}$,$\sigma_p^2\mathbf{I}_p$); $\boldsymbol{fam}$ is the vector of random family effects, with $\boldsymbol{fam}$ ~ MVN($\mathbf{0}$,$\sigma_{fam}^2\mathbf{I}_{mf}$) and $\boldsymbol{e}$ is the vector of random residual effects, with $\boldsymbol{e}$ ~ MVN ($\mathbf{0}$,$\mathbf{R}$), with $\mathbf{R}$ corresponding to a block diagonal matrix with an independent and identically distributed errors. The matrices $\boldsymbol{X}$, $\boldsymbol{Z_1}$, $\boldsymbol{Z_2}$ and $\boldsymbol{Z_3}$ are incidence matrices that relate the effects to the data. Also, $\boldsymbol{I}_x$ are identity matrices of the proper size. The family-by-trial interaction term (i.e. $\boldsymbol{fam \times s}$) is not included in this model as it is confounded with $\boldsymbol{fam}$, because none of the full-sib families were planted on more than one trial.

The parental model aims at partitioning the $\boldsymbol{fam}$ and $\boldsymbol{fam \times s}$ terms into several components, based on the following expression:

$$\sigma_{fam}^2 + \sigma_{fam\times s}^2 = (\sigma_m^2 + \sigma_f^2 + \sigma_{mf}^2) + (\sigma_{m\times s}^2 + \sigma_{f\times s}^2)$$

where σ_{fam}^2 is the family variance component, $\sigma_{fam\times s}^2$ is the interaction family-by-trial,σ_m^2 is the male variance component,σ_f^2 is the female variance component,σ_{mf}^2 is a variance component of the mother and father interaction, $\sigma_{m\times s}^2$ is the male-by-trial interaction variance component, $\sigma_{f\times s}^2$ is the female-by-trial interaction variance

component, and $\sigma^2_{mf\times s}$ is the interaction of a given cross with site. The above equation can also be expressed in terms of genetic components, as

$$\sigma^2_{fam} + \sigma^2_{fam\times s} = \left(\frac{V_a}{2} + \frac{V_d}{4}\right) + \left(\frac{V_{a\times s}}{2}\right)$$

where, V_{axs} is the additive-by-trial interaction variance components, respectively.

The following linear mixed model was fitted for the parental model:

$$\boldsymbol{y = 1\mu + X\,s + Z_1 b + Z_2 p + Z_3 m + Z_4 f + Z_5\, m\times s + Z_6\, f\times s + Z_7 mf + e} \qquad (3)$$

where, $\boldsymbol{y}$ is the vector of observation for all trials; $\boldsymbol{\mu}$ is a constant representing the overall mean, $\boldsymbol{s}$ is the vector of trials fixed effects; $\boldsymbol{b}$ is the vector of random block effects within trial, with $\boldsymbol{b}$ ~ MVN ($\mathbf{0}$,$\sigma^2_b \boldsymbol{I}_b$); $\boldsymbol{p}$ is the vector of random plot effects, with $\boldsymbol{p}$ ~ MVN ($\mathbf{0}$,$\sigma^2_p \boldsymbol{I}_p$); $\boldsymbol{m}$ and $\boldsymbol{f}$ is the random effect of male ($\boldsymbol{m}$) and female ($\boldsymbol{f}$), with $\boldsymbol{m}$ ~ MVN ($\mathbf{0}, \sigma^2_m \boldsymbol{I_m}$) and $\boldsymbol{f}$ ~ MVN ($\mathbf{0}, \sigma^2_m \boldsymbol{I_m}$); $\boldsymbol{m\times s}$ and $\boldsymbol{f\times s}$ are the vector of random interaction effects between parents and trials (overlaid), with $\boldsymbol{m\times s}$ ~MVN (0,$\sigma^2_{ms} \boldsymbol{I}_s$) and $\boldsymbol{f\times s}$ ~ MVN (0,$\sigma^2_{ms} \boldsymbol{I}_s$); $\boldsymbol{mf}$ is the random effect of families, with $\boldsymbol{mf}$ ~ MVN($\mathbf{0}$,$\sigma^2_{,mf} \boldsymbol{I_{mf}}$) and $\boldsymbol{e}$ is the vector of random residual effects, with $\boldsymbol{e}$ ~ MVN ($\mathbf{0}$,$\boldsymbol{R}$). The matrices $\boldsymbol{X}$, $\boldsymbol{Z_1}$, $\boldsymbol{Z_2}$, $\boldsymbol{Z_3}$, $\boldsymbol{Z_4}$, $\boldsymbol{Z_5}$, $\boldsymbol{Z_6}$and $\boldsymbol{Z_7}$ are incidence matrices that relate the effects to the data, and $\boldsymbol{I}_x$ are identity matrices of the proper size. The $\boldsymbol{R}$ matrices correspond to a block diagonal matrix with independent and identically distributed errors, respectively.

Approximately 33% of the parents are represented in two or more trials (Table 3) for which it is possible to estimate the parental-by-site interaction (i.e. additive-by-site interaction). However, as before, the variance associated with $\boldsymbol{mf}$ contains the confounded terms of family-by-trial interaction, i.e. $\boldsymbol{mf\times s}$. Therefore, σ^2_{mf} will be upwardly biased.

Table 3: Number of parents (mother and father), for each of the trials (diagonal) and number of common parents among a pair of trials (off-diagonal).

Trials	1	2	3	4	5
1	**39**	11	11	5	5
2		**39**	9	2	5
3			**36**	7	5
4				**30**	12
5					**35**

Both, the family and parental models were fitted using ASReml (Gilmour et al., 2009) and the following parameters were calculated based on the causal components:

Additive variance: $V_a = 4 * \sigma^2_m$

Dominance variance: $V_d = 4 * \sigma^2_{mf}$

Additive-by-site variance: $V_{a\times s} = 4 * \sigma^2_{ms}$

Total variance: $V_T = \sigma^2_b + \sigma^2_p + \sigma^2_m + \sigma^2_{mf} + \sigma^2$

Narrow sense heritability: $h^2_a = \frac{V_a}{V_T}$

Dominance ratio: $d^2 = \frac{V_d}{V_T}$

Heritability of cross: $h^2_{cross} = \frac{\sigma^2_{fam}}{V_T}$

2.4. Comparison of statistical packages

In order to compare the use of different software to estimate the variance components, in addition to the use of ASReml, the statistical package Selegen – Reml/Blup (Resende, 2007) was used. Both packages estimate the variance components by restricted maximum likelihood (REML); however, ASReml uses the average information algorithm (Gilmour et al., 1995) and Selegen uses the SB-EMREML algorithm that combines the Takahashi and Sparse bifactorization presented by Zollenkoff (1973)(Resende, 2007).

In Selegen the following linear mixed model, identified by Resende (2007) as 147, was fitted:

$$\boldsymbol{y = 1\mu + X\, s\times b + Z_1\, p + Z_2\, fam + e} \quad (4)$$

where, $\boldsymbol{y}$ is the vector of observation for all trials; $\boldsymbol{\mu}$ is a constant representing the overall mean, $\boldsymbol{s\times b}$ is the vector of fixed effects corresponding to blocks across trials; $\boldsymbol{p}$ is the vector of random plot effects across trials; $\boldsymbol{fam}$ is the vector of random family effects, and $\boldsymbol{e}$ is the vector of random residual effects. As before, $\boldsymbol{X}$, $\boldsymbol{Z_1}$ and $\boldsymbol{Z_2}$ are incidence matrices that relate the effects to the data. This model is equivalent to the family model (2), but here, block effects are considered fixed. A common residual term for all trials was assumed as in (2).

2.5. Genotype selection

After performing multiple-site analyses it is possible to selects parents, crosses (i.e. families) or individual genotypes by using the family prediction according with family model (2). In this study, the top 20 families were selected, corresponding to 18.2% of 110 families evaluated. These were selected using an index with equal weight on BRIX and WS. In order to understand the physiological and morphological characteristic of these 20 selected families, principal components analysis (PCA) were done using the correlation matrix among the traits NS, HS, DS, BRIX, and WS. The first PCA analysis was done with all variables and the second one was done with the same variables as before, with the exception of HS. The exclusion of HS was to see the comportment of all others variables in terms of direction and distribution of families.

Recall that PCA is a statistical method used to analyze multivariate data. The results of PCA are a linear combination of the original set of variables. The principal components are calculated in a decreasing order of importance, i.e., the first principal

component explains the largest proportion of total variation associated with the variables of interest (Reis, 2001).

3. RESULTS AND DISCUSSION

3.1. Single-site analyses

The phenotypic and genotypic variance, together with the heritability of a cross for a single-site analyses are presented in Table4.For most trials and variables, the plot variance component (Table 4) was small or zero. The a little bit higher proportions of the total variability were found for HS with values reaching 17.3% (trial 3) and an average of 5.5%. BRIX also presented some plot variability, with an average of 2.8% across all trials. In relation of the magnitudes of the heritability of a cross, h^2_{cross}, HS presented consistently the highest values with an average value of 14.1%, followed by BRIX and DS, with values of 7.3% and 6.6%, respectively. Almost null values of heritability were found for NS and WS for the majority of the trials. However, heritability values differed considerable from trail to trial.

Table 4: Variance component and genetics parameters of single site analyses via ASReml for each trial and for the variables number (NS)and weight (WS) of stalks, juice percentage of soluble solids (BRIX), diameter (DS) and height (HS) of stalks of 22 families per trial, Oratórios, MG, 2009.

			Trial 1		
	NS	WS	BRIX	DS	HS
σ^{2*}_{fam}	1.353	3.590	0.802	0.641	0.027
σ^2_p	0.000	0.000	0.107	0.60	0.0025
σ^2	40.91	106.62	6.99	18.22	0.236
V_T	42.29	110.23	7.899	19.461	0.266
h^2_{cross}	0.032	0.033	0.101	0.033	0.101
			Trial 2		
	NS	WS	BRIX	DS	HS
σ^2_{fam}	0.919	4.770	0.352	1.76	0.051
σ^2_p	0.000	0.000	0.097	0.22	0.004
σ^2	30.79	91.42	6.810	18.18	0.222
V_T	31.71	96.42	7.256	20.16	0.275
h^2_{cross}	0.029	0.049	0.048	0.007	0.185

	Trial 3				
	NS	WS	BRIX	DS	HS
σ^2_{fam}	0.451	1.050	0.240	1.560	0.019
σ^2_p	0.346	0.280	0.202	0.000	0.050
σ^2	32.96	85.26	6.56	17.81	0.220
V_T	33.75	86.59	7.00	19.37	0.289
h^2_{cross}	0.013	0.012	0.034	0.081	0.066
	Trial 4				
	NS	WS	BRIX	DS	HS
σ^2_{fam}	0.847	1.158	0.655	2.710	0.050
σ^2_p	0.000	0.000	0.228	0.074	0.018
σ^2	40.01	99.47	5.830	17.05	0.199
V_T	40.85	100.63	6.713	19.83	0.267
h^2_{cross}	0.021	0.011	0.097	0.136	0.187
	Trial 5				
	NS	WS	BRIX	DS	HS
σ^2_{fam}	0.552	0.686	0.671	1.410	0.045
σ^2_p	0.000	0.000	0.391	0.000	0.003
σ^2	38.03	97.06	6.606	18.53	0.220
V_T	38.58	97.74	7.667	19.94	0.268
h^2_{cross}	0.014	0.007	0.087	0.071	0.168

*σ^2_{fam} is variance component of family; σ^2_p is variance component of plot; σ^2 is variance component of error; V_T is total variance and h^2_{corss} is a heritability of crosses.

3.2. Multiple-site analyses

For multiple-sites analyses, considering the results from the family model (2) (Table 5), it was found that the design terms, i.e. block and plot, were small. When expressed as proportion of the total variability, these ranged from 0.4% to 2.3%, for blocks, and 0.0% to 2.6% for plots. The highest proportions were found for BRIX. For all traits, the residual variance components corresponded to a large portion of the total variability ranging from 83.5% to 97.3%. Therefore, in these trials there is a large within-plot variability that affects the prediction of genetic components, and these values were similar for all trials.

The proportion of family variation, which is equivalent to the heritability of a cross, was largest for HS with a value of 14.1%, followed by DS and BRIX, with values

of 8.0% and 7.0%, respectively. In contrast, almost no genetic control was found for the traits NS and WS.

Table 5: Variance component and proportions to total variance from multiple-site analyses of family model (2) for number (NS) and weight (WS) of stalks, juice percentage of soluble solids (BRIX), diameters (DS) and height (HS) of stalks of 110 families via ASReml, Oratorios, MG, 2009

Variance component					
	NS	WS	BRIX	DS	HS
Block*	0.168	0.364	0.171	0.070	0.003
Plot	0.059	0.032	0.198	0.154	0.003
Family	0.799	2.25	0.533	1.593	0.037
Error	36.56	95.98	6.564	17.96	0.219
Average	8.127	11.40	19.81	25.30	2.541
Proportions of total variance					
Block	0.004	0.004	0.023	0.003	0.011
Plot	0.002	0.000	0.026	0.008	0.011
Family	0.021	0.023	0.070	0.080	0.141
Error	0.970	0.973	0.880	0.907	0.835

*Block = Variance component of block; Plot = variance component of plot of each trial; Family = variance component of family effect and Error = variance component of error.

The results from fitting the parental model are presented in Table 6. The block and plot components are very similar to the ones obtained from the family model, with small variations due to the unbalanced structure of the data. For all traits, the residual variance components corresponded to a large portion of the total variability ranging from 85.8% to 98.2%. In addition, all traits presented a limited amount of genetic control (additive and dominance), which is due, in part, to large environmental noise. Chaudhary (2001) compared the phenotypic and genotypic proportions through of coefficient of variation and found that all traits (e.g. weight, diameter and length of stalks and BRIX) indicated great influence of environment (phenotypic) on genetic variation.

Table 6: Variance component and proportions of total variance from multiple-site analyses of parental model (3), for number (NS) and weight (WS) of stalks, juice percentage of soluble solids (BRIX), diameters (DS) and height (HS) of stalks of 110 families via ASReml v.3, Oratórios, MG, 2009.

Variance component					
	NS	WS	BRIX	DS	HS
Block *	0.167	0.363	0.173	0.069	0.003
Plot	0.069	0.036	0.192	0.144	0.003
Error	36.56	95.97	6.57	17.97	0.219
Parental	0.341	0.216	0.199	0.501	0.008
Family	0.093	1.835	0.072	0.556	0.022
Parental x sites	0.000	0.000	0.041	0.000	0.000
Proportions of total variance					
Block	0.004	0.004	0.024	0.004	0.012
Plot	0.002	0.000	0.027	0.008	0.012
Error	0.982	0.975	0.907	0.934	0.858
Parental	0.009	0.002	0.027	0.026	0.031
Family	0.003	0.018	0.010	0.028	0.086
Parental x sites	0.000	0.000	0.005	0.000	0.000

*Block = Variance component of block; Plot = variance component of plot of each trial; Mother = variance component of additive effect; Family = variance component of dominance effect; Mother x sites = interaction between parents and trials, and Error = variance component of error.

The genetic parameters estimates for all traits are presented in Table 7. The narrow-sense heritability estimates, h^2_a, for all variables were found to be relatively low, ranging from 0.9% to 12.5%, with the largest values for HS, BRIX and DS and very small values for WS and NS. Therefore, it is recommended to perform parental selection on HS, BRIX and DS.

No prior studies have reported results for multiple-site analyses. However, for single site analyses, both of the heritability values (narrow-sense and cross) found in this study was lower than those reported in the literature (Hogarth, 1971; Skinner et al. (1987); Oliveira, 2007; Martin et al. (2007) and Silva, 2009). For example, Hogarth, 1971 reported a considerable range of narrow-sense of 0.21 to 0.84 for HS, 0.30 to 0.94 for DS and 0.13 to 0.82 for NS. Skinner et al. (1987) reported the narrow-sense and cross heritability of 0.17 and 0.75 for WS, respectively. Oliveira, 2007 found for narrow-sense heritability of 0.45, 0.31, 0.23 and 0.09 for BRIX, height, diameter and

number of stalks, respectively. Martin et al (2007), estimated the heritability of cross of 0.7244 and 0.6201 for HS and DS, respectively. Silva, 2009 reported for BRIX, 0.076 of narrow-sense individual heritability and 0.63 for cross heritability.

With respect to the proportion of dominance, d^2 (Table 7), which contain the dominance and dominance-by-trial interaction confounded, presented interesting values for HS, DS and WS, with a proportion as large as 34.5% for HS. For HS, DS and WS further insight through proper experiments is required in order to determine if this d^2 corresponds to dominance or their interactions with site in order to determine the best selection strategy. The proportion of dominance has been reported in several studies for sugarcane. Liu et al. (2007) found a significant and higher dominance effect for WS and also a significant effect of interaction between dominance and environment.

Table 7: The average values and estimates of genetic parameters considering the five trials for the variables number (NS) and weight (WS) of stalks, juice percentage of soluble solids (BRIX), diameter (DS) and height (HS) of stalks of 22 families via ASReml v.3, Oratórios, MG, 2009.

	NS	WS	BRIX	DS	HS
V_a*	1.364	0.864	0.796	2.000	0.032
V_d	0.372	7.340	0.288	2.230	0.088
V_{axs}	0.000	0.000	0.164	0.000	0.000
V_T	37.23	98.420	7.250	19.240	0.255
h^2_a	0.036	0.009	0.110	0.104	0.125
d^2	0.010	0.074	0.039	0.115	0.345
$V_a/(V_a+V_d)$	0.786	0.105	0.734	0.474	0.266
h^2_{cross}	0.002	0.018	0.009	0.029	0.086

*V_a= additive variance component; V_d = dominance variance component (contains dominance-by-site); $V_{a\times s}$ = interaction variance component additive-by-site; V_T = total variance; h^2_a = Narrow-sense heritability; d^2 = dominance index (contains dominance-by-site) and h^2_{cross} = cross heritability.

The source of variation additive-by-environment, $V_{a\times s}$, is an important component as it can affect the parental stability across different locations and years (Yan and Hunt, 1998) and therefore, it has a strong effect on selection strategies. In this study, non-existent to small proportions of this component were found for all traits with the exception of BRIX, which presented a small proportion of the total variability

of2.26 %. Therefore, this interaction appears to be not relevant in these analyses, and it is expected that most parents will responds similarly across different environments.

The ratio $V_a/(V_a+V_d)$ (Table 7), which denotes the relative contribution of additive variance to total genetic variance (i.e. additive plus non-additive), was large for NS and BRIX, indicating that for these traits the genetic control is mainly additive. Equal contributions of additive and non-additive components were found for DS. In contrast, a very small ratio was found for WS, indicating almost exclusively dominance and dominance-by-site effects. The ratios obtained for NS and WS in this study agree with previous results. Matsuoka et al (1999), analyzing the relation between additive and dominance components also for sugarcane, found that for BRIX, the additive component is predominant, and for WS, the non-additive component is predominant. Also, according to Hogarth (1971), for sugarcane the variables NS and BRIX, the additive component is predominant.

Different selection strategies could be pursued for each trait according to their heritability of a cross, or the narrow-sense heritability. Hence, for a breeding program, the traits with mostly additive control could be based in early generation selection of parents, while traits with more dominance control could favor selection of families or specific crosses. In this study, the heritability of a cross, h^2_{cross}, is only relevant for HS; therefore, it is recommended to make selection of specific crosses for this trait. Also, this trait allows for some level of gain from selection of families and parents, and BRIX and DS could be favored by selecting individual parents rather than families. Cavalcanti (1990) performed a review of literature on selection of clones and also found a high range of heritability estimates for DS and WS in studies with sugarcane.

In the present study, epistasis could not be estimated as this component is confounded with the error term, which for this study ranged from 85.8% to 98.2% for the five traits analyzed. Several studies in sugarcane indicate that the amount of

epistasis depends on a given trait. Epistasis has been estimated in a sugarcane population for mosaic virus resistance (Xing et al., 2006), sucrose level (Richard and Henderson, 1981), and biomass traits (Hongkai et al., 2009). A study from southern China (Hongkai et al. 2009) found significant epistatic effects (additive-by-additive) for WS corresponding to a proportion of 32.1%, but no significant epistasis for the trait NS. In addition, these authors found a significant portion of epistasis-by-environment effects for WS and total biomass yield. Therefore, for the present breeding population, further studies, with appropriate experimental designs, are needed in order to understand the importance of epistasis in the traits of interest.

3.3. Comparison of statistical packages

The results from fitting a family model using Selegen are presented in Table 8. In comparison with the family model fitted using ASReml (Table 5), the variance component estimates are similar for all traits. The only exception is the plot variance that resulted in larger estimates. This is probably due to the use of a different model in Selegen with blocks assumed fixed and a almost the same residual variance values for all trials. Minimal changes were noted for h^2_{cross}, with the exception of BRIX which presented smaller values, but not relevant in practice. Therefore, according to these analyses, both statistical packages provide with equivalent results. The values from h^2_{cross} were much better than results from Table 7 and close from Table 5.

Table 8: Estimative of variance components and genetics parameters for traits: number and weight (WS) of stalks, juice percentage of solids soluble (BRIX), diameter (DS) and height (HS) of stalks of 110 families, via model 147 from Selegen – Reml/Blup, Oratórios, MG, 2009.

	NS	WS	BRIX	DS	HS
Family	0.813	2.210	0.450	1.510	0.036
Plot	0.097	0.232	0.492	0.440	0.009
Error	36.16	95.21	6.377	17.77	0.215
Average	8.215	11.47	19.75	25.24	2.520
V_T	37.07	97.65	7.289	19.72	0.261
h^2_{cross}	0.022	0.023	0.062	0.077	0.138

*Family = family variance component; *Plot* = environmental variance between plots; *Error* = variance component of error; V_T = Total variance component; h^2_{cross} = heritability of cross and *Average*= average of each trait.

3.4. Genotype selection

The PCA obtained from using family predictions from model (2) based on the top 20 families over all traits with the exception of HS, resulted in a total variation explained of 92.33%. The first and second components explained 69.49% and 22.84%, respectively. From the biplot (Figure 1), two opposite groups can be identified. These correspond to the families f2, f24, f50, and f98 (Group 1), and to f6, f70, f89, and f105 (Group 2). The average family trait values for different subset of families are shown in Table 9, together with their respective genetic gains in reference to the overall mean. From this table, it is clear that Group 1 has higher average than Group 2 for NS and WS, and lower for DS and BRIX. For HS, both groups show similar average family values. The difference between these groups indicates that the groups correspond to two distinctive physiological/morphological ideotypes. Group 1 corresponds to plants with larger weight distributed in several stalks of small diameter. In contrast, Group 2 is represented by plants with smaller weight, but distributed in fewer stalks with larger diameter. In addition, all the families from Group 2 have the same mother (RB92606), except for f89 but for Group 1 all parents were unique. In terms of family selection, the

best genotype depends on the trait of interest. For example, family f70 had larger DS and BRIX, and families f2 and f24 had the largest WS and NS.

Table 9: Average values of all variables of group 1 and 2 for their family phenotype and BLUP values.

Average of families					
	NS	WS	BRIX	DS	HS
Group 1	9.28	12.76	19.88	24.84	2.57
Group 2	7.37	11.74	20.90	26.66	2.65
Top 20 mean	8.50	12.36	20.34	25.61	2.63
Overall mean	8.20	11.44	19.76	25.23	2.52
Proportion of Gain (%)					
Group 1	13.27	11.55	0.57	-1.56	1.72
Group 2	-10.04	2.63	5.77	5.67	4.89
Top 20	3.64	8.07	2.94	1.53	4.14

* Group 1 consisting of families f2, f24, f50 and f98; Group 2 consisting of families f6, f70, f89, f105.

The PCA results obtained from analyzing all the traits simultaneously resulted in a total variation explained of 78.64%, with 57.30% and 21.354%, for the first and second principal component, respectively. The biplot for this analysis is presented in Figure 2, and is similar to the one presented earlier, with the same grouping. Also, a high, and positive correlation, between HS and DS is shown in this plot. One important difference is that family f98 now represents a more isolated point.

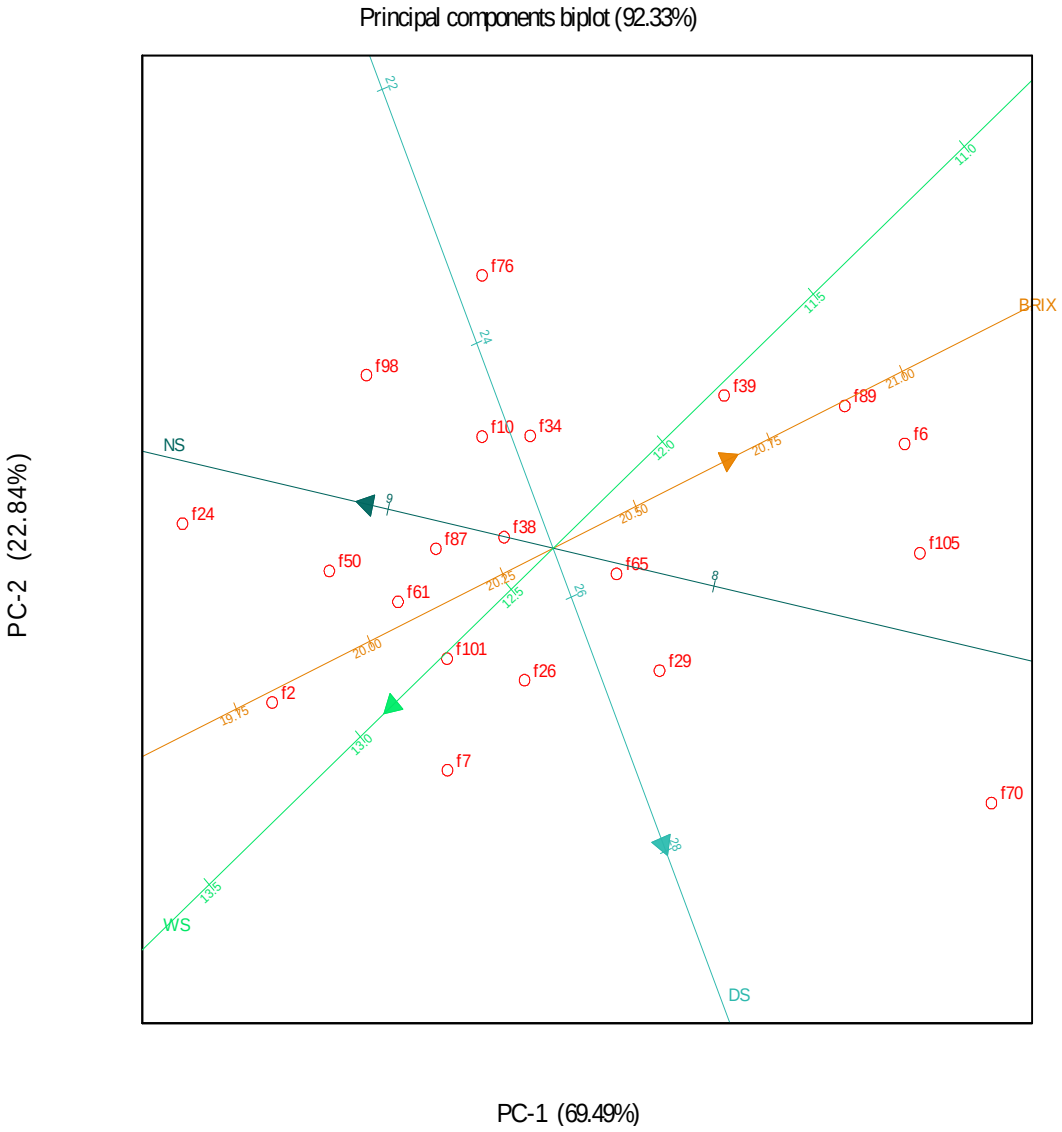

Figure 1: Biplot from all families and traits following number of stalks (NS), diameter of stalks (DS), weight of stalks (WS) and juice percentage of solids soluble (BRIX). The color vectors correspond to each variable and its comportment with this analysis. Red circles correspond to each family.

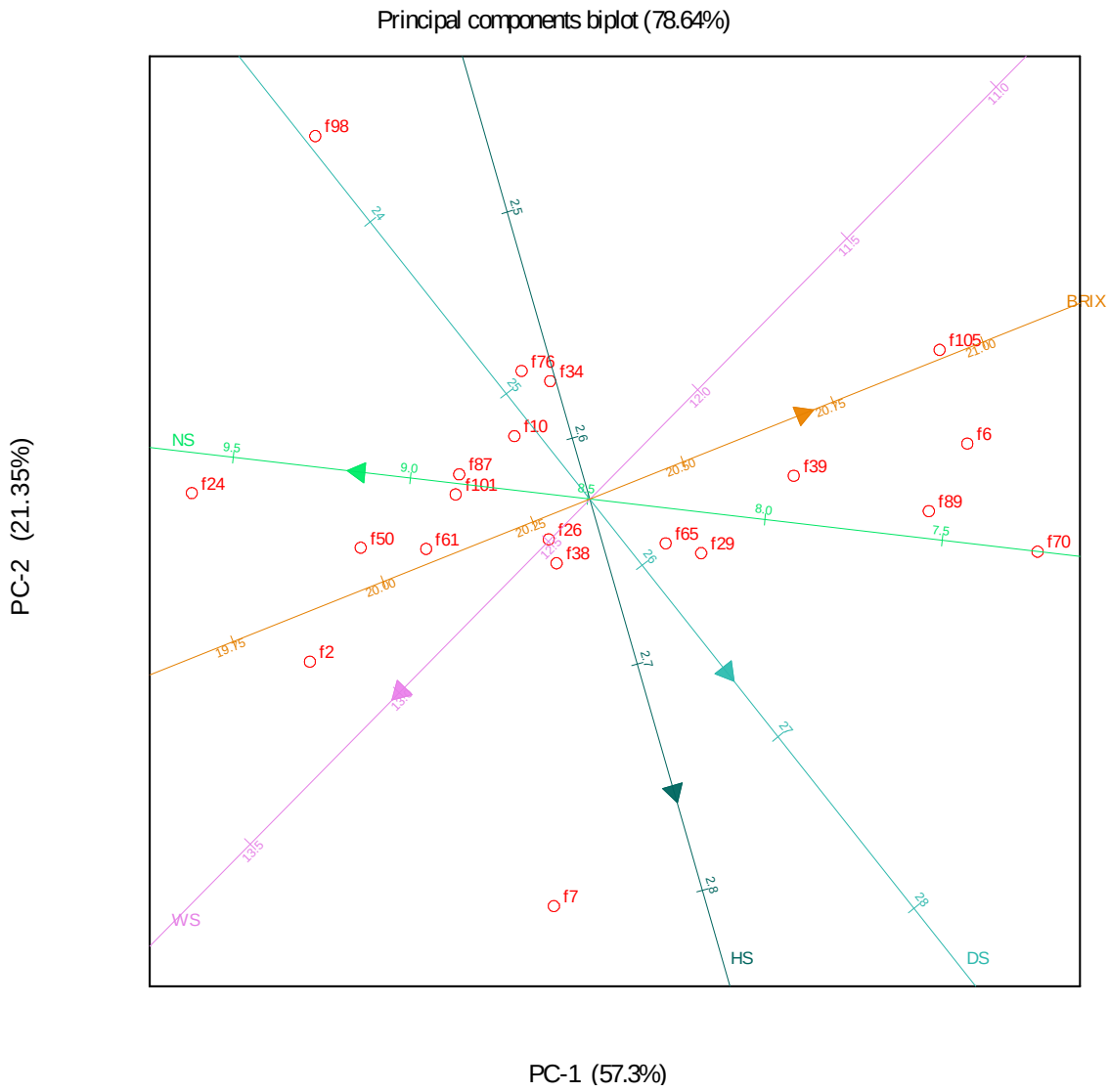

Figure 2: Biplot from all families and number of stalk (NS), height of stalks (HS), diameter of stalks (DS), weight of stalks (WS) and juice percentage of solids soluble (BRIX). The color vectors correspond to each variable and its comportment with this analysis. Red circles correspond to each family.

4. CONCLUSIONS (AND RECOMMENDATIONS)

For single sites analysis:

-The variable height of stalks, BRIX and diameter of stalks can be used for family selection.

-Height, diameter of stalks and BRIX are good traits to be considered on breeding program in sugarcane.

- Height of stalks shows the higher heritability of crosses.

For multiple-site:

- Interesting non-additive effects (i.e. dominance and dominance-by-site) were found for WS, DS and HS.
- It is recommended the use of HS for family and parental selection.
- BRIX and DS should be used for parental selection rather than family selection.

For PCA analyses

- The top 20 families was separated in two groups according with the average families and the PCA analyses proven them structural difference.

CHAPTER 2

COMPARISON OF AMONG AND WITHIN SUGARCANE FAMILY SELECTION METHODS UNDER DIFFERENT GENETIC SCENARIOS

RESUMO

O principal objetivo desse estudo é avaliar através de simulação de dados a eficiência do BLUPIS como um método de seleção de clones de cana-de-açúcar nos programas de melhoramento. Para o desenvolvimento desse trabalho foram considerados alguns cenários genéticos diferindo nas proporções das variâncias aditivas, dominância e epistasia, junto com a herdabilidade no sentido amplo e restrito. Os dados experimentais foram os mesmo do capitulo um, baseados nos resultados das analises do capitulo um através do modelo parental foram criados seis cenários genéticos divididos em dois grupos. O primeiro grupo, tem maiores valores de dominância (1D e 3D) e o segundo grupo possui maiores valores de epistasia (1I e 3I), vale ressaltar que o grupo 2 (2D e 2I) os valores foram extrapolados e não concordam com essa divisão entre D e I. De acordo com modelos estatísticos foram simulados 500 bancos de dados para cada cenário. Os dados foram analisados através de modelos individuais e em nível de parcela, com os resultados foram realizadas as comparações dos dois métodos (BLUPI and BLUPIS) com o valor genotípico real. Para as comparações foram usados os valores absolutos das diferenças entre os valores de correlação entre o real valor genotípico e o BLUPI, entre o real valor genotípico e o BLUPIS e entre BLUPI e BLUPIS. O uso do BLUPIS é recomendado no melhoramento genético e provou ser efetivo para diferentes cenários baseados em diversos valores de variância aditiva, dominância e epistasia, junto com os diferentes valores de herdabilidade com a mínima perda de informações.

ABSTRACT

The main objective of this study is to evaluate through computer simulation, the efficiency of BLUPIS as a method to select sugarcane clones in a breeding program.

To develop this study were considered several scenarios differing in proportions of additive, dominance and epistasis variance, together with different narrow and broad sense heritabilities. The dataset and experimental design were the same of chapter 1, based in the results from the parental model of chapter one, were created six scenarios divided in two groups. First group, there are majority of dominance (1Dand 3D) and, second one with majority of epistasis (1Iand 3I), highlight for group 2 (2D and 2I), the values were extrapolated and are out of this division of majority dominance and majority of epistasis. In according with statistical model were simulated 500 datasets for each scenario. This dataset were analyzed through of individual and plot model, after were made the comparison of two methods (BLUPI and BLUPIS) with real genetic values. For the comparison were used the absolute value of differences and correlation between real genetic value and BLUPI, real genetic value and BLUPIS and between BLUPI and BLUPIS. The used of BLUPIS is recommended for genetic improvement and proved to be effective to different scenarios based in diverse values of additive, dominance and epistasis variance component with minimal loss information.

1. INTRODUCTION

Sugarcane (*Saccharum officinarum* L.) is an alogamous species belonging to the Gramineae family (Poaceae), which is mainly asexually propagated. It is a complex polyploidy and provides raw material for the production of white sugar and alcohol (Zhou et al., 2005).

Sugarcane is an important industrial crop, totaling a production of more than 560 million tons in 2011, mainly from Brazil, Hawaii, Australia and USA. Brazil has a total of 851 million hectares or arable land, of which 6 million hectares are occupied by sugarcane crops, with a total production of about 600 million tons for the 2010/2011 harvest year (Ministério da Agricultura, 2011).

Sugarcane breeding programs aim to generate new cultivars that increase productivity and reduce economic cost (Oliveira, 2007). Sugarcane genotype selection is executed in all phases of a breeding program. It begins with parental selection followed by selection of crosses and then individuals belonging to the population, and finishes with clonal selection (Souza, 1989; Calija et al., 2001). The main objectives of these programs are the selection of potential interesting genotypes (Resende and Barbosa, 2005). Usually the measurement and selection traits are the number, height, diameter and weight of stalks, juice percentage of soluble solids (i.e. BRIX), together with production of biomass, sugar and percentage of sugar (Oliveira, 2007).

In the southern hemisphere, the typical sugarcane breeding program is composed of the following phases: T1, T2, T3, and FE. In phase T1, seedlings are obtained from crosses and transplanted into the field. In general, in this phase, large numbers of genotypes are evaluated, which are not replicated and are planted in a single environment. Mass (i.e. individual) or family selection are used, but frequently the

former is preferred. In phase T2, the genotypes selected from the previous phase are vegetative propagated (or 'cloned') and planted in a single plot using an augmented block design with commercial controls. For phase T3, the latest selected clones from phase T2 are now typically tested in several environments and followed through many years. These are planted as augmented blocks or randomized complete block designs. In phases T2 and T3, selection is based in choosing the best families (i.e. family selection) followed by mass selection of individuals within these families. Programs use different methods to determine the number of individuals within a family to select (e.g. simulated individual BLUP; Resende and Barbosa, 2006). Finally, during the experimental phase (FE), clones selected from phase T3 are intensively propagated and operationally evaluated in various industries and distilleries. Here, the genotypes are evaluated in several environments on large plots with three replications or more depend of number of clones and with sufficient number of harvests (Ferreira et al., 2005).

Depending on the phase, selection of individuals, families or a combination of these is used (Resende and Barbosa, 2006). Several sugarcane breeding programs have preferred family selection, as this has been shown to be superior to individual selection by providing with better genetic gains, cost of operation and resource efficiency (Stringer et al., 2011). Australian and west Indian sugarcane breeding programs use family selection (Cox and Hogarth, 1993; Kennedy and Bellamy, 1997) and in Brazil, Colombia and Argentina, a modified version of this selection is used (Bressiani et al., 2005).

The field trials commonly used for family selection can also assist on the estimation of parental breeding value that produced the families (Stringer et al., 2011). In this situation, parents need to be crossed in some pre-planned arrangement, such as factorial or diallel. However, this is often difficult due to differential sexual maturity timing. The above situation, together with different amounts of seedling per family,

generates an unbalanced experiment that requires to be analyzed using linear mixed models with the recommended statistical procedure REML (restricted maximum likelihood) allowing the simultaneous estimation of variance components and BLUP (Best Linear unbiased prediction) values for individual genotypes, families and/or parent for balance or unbalanced situations (Resende, 2004).

Sugarcane field trials are typically established with multiple plants per plot belonging to the same family. These individuals correspond to full-sibs and could be originated from sexual or asexual propagation. Yield measurements, and other traits, are obtained at the plot level, as all the individuals in a plot are harvested together. The best family is selected considering this plot level information; however, selection of the best genotypes within a family is not possible as individual measurements within a plot are not available. This could be done using within-plot mass selection, but a recommended alternative is the Simulated Individual BLUP (BLUPIS; Resende and Barbosa, 2005 and Resende, 2007). BLUPIS determines the families from which individuals (e.g. clones) that will be advanced in the program, together with the number of families and numbers individuals within a family to be selected. Resende and Barbosa (2005) report that this methodology increases the efficiency of the selection process and, compared to mass selection, it reduces costs, because in BLUPIS a smaller number of clones are advanced in the selection.

Most breeding program strategies can be cheaply and quickly evaluated with the use of stochastic simulation, which allows evaluation of several hypothetical scenarios with great flexibility. Stochastic simulation describes a real phenomenon through a model that should be a realistic replica of the situation of interest. It has basically two objectives: to understand the existing system and, to prescribe recommendations (Ross, 2006).

A model used for simulation must represent properly an object, system or idea, i.e. representative of the real phenomenon, and it should be flexible and easy to interpret, and the output should be comparable with the real object (Cruz, 2001).

Several simulation studies for breeding strategies have been reported. Gurgel (2004), working with autogamous plants, determined the optimal number of selected families, which can be easily extrapolated to other situations. Fernandes (2006) evaluated the efficiency of circular diallels compared to complete diallels, with respect to the estimation of general and specific combining ability. Souza et al. (2006) compared the relative efficiency for clone selection and estimation of genetic parameter based on augmented designs against randomized complete block designs. Pinto Junior (2004) used simulation for indentify the best individual and families for compose a population and them respectively clone of *Eucalyptus grandis* W HILL EX MIDEN. Peternelli et al. (2009) used simulation to compare the efficiency in the selection of genotypes and the quality of the estimates of the variance components and heritability for augmented block designs, duplicated augmented block designs and randomized complete block designs experiments.

The main objective of this study is to evaluate, the efficiency of the BLUPIS as a method to select sugarcane potential clones in a breeding program, through computer simulation. This will be done considering several scenarios differing in proportions of additive, dominance and epistasis variance, together with different narrow and broad sense heritabilities that mimic a range of traits of interest.

2. MATERIAL AND METHODS

2.1. Analysis of sugarcane field trials

In order to consider realistic situations (scenarios) for the simulations, a series of datasets belonging to current field trials was used as the basis to define genetic

parameters. These datasets originated from five experiments carried by the sugarcane breeding program, at Federal University of Viçosa, Minas Gerais, Brazil (see details in Chapter 1). Each experiment was carried out on May, 2007, in a randomized complete block design (RCBD) with five replication and 24 treatments. The treatments consisted of 22 full-sib families and two controls checks per trial. Therefore, the total number of treatments (or families) was 110. Each full-sib family was only represented in one experiment. For analyses, the two controls were dropped.

The phenotypic traits evaluated were: number (NS) and weight (WS, kg ha^{-1}) of stalks, juice percentage of soluble solids (BRIX, %), and diameter (DS, cm), and height (HS, m) of stalks. For each plant, NS was measured by counting the number of stalks longer than 1 m, and WS was obtained by weighting the complete plant and dividing by its NS only on those plants with more than five leaves. For a single randomly selected stalk for each of the 20 plants in the plot, BRIX was determined by refractometry, DS was measured using a digital caliper at ground level, and HS correspond to the height from ground level until dewlap.

A multi-site analysis was performed using a linear mixed model fitted by REML/BLUP using ASReml v. 3 (Gilmour et al. 2009). The linear mixed model considered was:

$$\boldsymbol{y} = \boldsymbol{1\mu} + \boldsymbol{Xs} + \boldsymbol{Z_1 b} + \boldsymbol{Z_2 p} + \boldsymbol{Z_3 m} + \boldsymbol{Z_4 f} + \boldsymbol{Z_5 m{\times}s} + \boldsymbol{Z_6 f{\times}s} + \boldsymbol{Z_7 mf} + \boldsymbol{e} \qquad (1)$$

where, $\boldsymbol{y}$ is the vector of observation for all trials; $\boldsymbol{\mu}$ is a constant representing the overall mean, $\boldsymbol{s}$ is the vector of trials fixed effects; $\boldsymbol{b}$ is the vector of random block effects within trial, with $\boldsymbol{b} \sim$ MVN $(\boldsymbol{0}, \sigma_b^2 \boldsymbol{I}_b)$, $\boldsymbol{p}$ is the vector of random plot effects within trial, with $\boldsymbol{p} \sim$ MVN $(\boldsymbol{0}, \sigma_p^2 \boldsymbol{I}_p)$; $\boldsymbol{m}$ and $\boldsymbol{f}$ are the vector of random parental effects (corresponding to overlaid mother and father effects), with $\boldsymbol{m} \sim$ MVN $(\boldsymbol{0}, \sigma_m^2 \boldsymbol{A})$ and $\boldsymbol{f} \sim$ MVN $(\boldsymbol{0}, \sigma_m^2 \boldsymbol{A})$; $\boldsymbol{m{\times}s}$ and $\boldsymbol{f{\times}s}$ are the vector of random interaction effects between parents and trials (also overlaid), with $\boldsymbol{m{\times}s} \sim$MVN $(0, \sigma_{ms}^2 \boldsymbol{A} \otimes \boldsymbol{I}_s)$ and $\boldsymbol{f{\times}s} \sim$ MVN

$(0,\sigma^2_{ms}\boldsymbol{A}\otimes\boldsymbol{I}_s)$; ***mf*** is the vector of random family effects, with ***mf*** ~ MVN($\mathbf{0}$,$\sigma^2_{mf}\boldsymbol{I}_{mf}$)and ***e*** is the vector of random residual effects, with ***e*** ~ MVN ($\mathbf{0}$,***R***). The matrices ***X***, ***Z*****1**, ***Z*****2**, ***Z*****3**, ***Z*****4**,***Z*****5**, ***Z*****6**and ***Z*****7**are incidence matrices that relate the effects to the data, and $\boldsymbol{I}_x$ are identity matrices of the proper size. The ***A*** and ***R*** matrices correspond to the numerator relationship matrix and a block diagonal matrix with independent and identically distributed errors, respectively

The results of these analyses are presented in Table 1 (more details are provided in Chapter 1). In Table 1, the summary of the genetic and environmental variance components, and other genetic parameters, are presented across all trials.

Table 1: The average values and estimates of genetic parameters, of real data via multiple-sites analysis, considering the five trials for the variables number (NS) and weight (WS) of stalks, juice percentage of soluble solids (BRIX%), diameter (DS) and height (HS) of stalks of 110 families via ASReml v.3 *parental model*, Oratórios, MG, 2009.

Component	NS	WS	BRIX	DS	HS
V_a *	1.364	0.864	0.796	2.000	0.032
V_d	0.372	7.340	0.288	2.230	0.088
Block	0.167	0.363	0.173	0.069	0.003
Plot	0.069	0.036	0.192	0.144	0.003
Error	36.56	95.97	6.570	17.97	0.219
V_T	37.23	98.42	7.25	19.24	0.255
h^2_a	0.036	0.009	0.110	0.104	0.125
d^2	0.010	0.074	0.039	0.115	0.345

*V_a is additive variance; V_d is dominance variance component; Block is variance component of block; Plot is variance component of plot of each trial; Error is variance component of error; V_T is total variance; h^2 is narrow-heritability and d^2 is a dominance ratio.

The above analyses provided with an estimate of the additive, dominance variance, together with narrow-sense heritability, and magnitude of design and environmental components for a variety of traits. The proportions of total variance for additive and dominance effect ranged from 0.88% to 12.55% and 0.99% to 34.51%, respectively. The higher values were found for HS on both effects. Note that the dominance variance, V_d, contain dominance-by-site interaction, as a family was only

planted in a single site. No estimate of epistasis is available from these analyses, due to the confounding of this component with the residual term.

For sugarcane, literature has reported a wide range of proportions of epistasis that depends on the trait. Hong-kai et al. (2009) found significant epistasis for WS and biomass, where the former was considerably larger than the latter. In contrast, these authors reported not significant presence of epistasis for the trait NS.

2.2. Field design and simulation

Six simulation scenarios were considered in this study, which varied according to their proportions of additive, dominance and epistasis components (Table 2). These scenarios were divided into three groups according to the magnitude of additive and non-additive variance, these were: 1) small additive and small non-additive, 2) small additive and large non-additive and, 3) large additive and small non-additive variation. In addition, within each group, different partitions of the non-additive into dominant and epistasis variance were considered: 1) with majority of dominance (D, composed by scenarios 1D and 3D) and, 2) those with majority of epistasis (I composed by scenario 1I and 3I), the scenario 2 (2D and 2I) are one exception, the values of epistasis were extrapolated and they are higher than dominance. The causal genetic components of variance were determined in proportion, for simplicity but without loss of generality, of the total variance, V_T, which was set to 1. These proportions were selected, partially based on the results presented in Table 1, together with results commonly reported in the literature, particularly for epistasis. Some extreme situations were considered; for example, a large epistasis was defined for Scenario 2.

Table 2: Causal genetic components of additive (V_a), dominance (V_d), epistasis (V_i), genetic variance (V_g), narrow-sense heritability (h^2_a) and broad sense heritability (H^2) in each scenario.

	Scenarios					
	1D	1I	2D	2I	3D	3I
V_a	0.100	0.100	0.100	0.100	0.400	0.400
V_d	0.075	0.025	0.100	0.020	0.075	0.025
V_i	0.025	0.075	0.300	0.380	0.025	0.075
V_g	0.200	0.200	0.500	0.500	0.500	0.500
h^2_a	0.100	0.100	0.100	0.100	0.400	0.400
H^2	0.200	0.200	0.500	0.500	0.500	0.500

The experimental design generated for each of the simulated field trials consisted on typical sugarcane trial based on a RCBD with five replications and 20 rectangular plots (2 x 10 plants). The treatments consisted of 20 full-sib families without controls.

The genetic components from the scenarios defined above (Table 2), explain between 20 to 50% of the total variability. The remainder, i.e. non-genetic portion, corresponds to the variability due to design effects together with environmental noise. Based on these elements, the linear model used to simulate the data was:

$$y_{iklm} = \mu + p_{ikl} + g_{klm} + e_{iklm} \quad (2)$$

where, g_{klm} was partitioned into:

$$g_{klm} = m_k + f_l + mf_{kl} + c(mf)_{klm} \quad (3)$$

where, y_{iklm} corresponds to the observation belonging to the i^{th} block, kl^{th} plot, μ is the fixed overall mean; p_{ikl} is the random effect of the kl^{th} plot within the i^{th} block, ikl =1,…,20, $p_{ikl} \sim N(0, \sigma^2_p)$; g_{klm} is the random effect of genotype, k = 1,…,20; l = 1,..,20 and m = 1,…,100, $g_{klm} \sim N(0, \sigma^2_g)$ and e_{jkl} is the random error effect, $e_{jkl} \sim N(0, \sigma^2)$.The genotype effect, is formed by a male, female, family and effect within family (individual effect), where m_k is the random effect of k^{th} male, $m_k \sim N(0, \sigma^2_m)$; f_l is the random effect of l^{th} female, $f_l \sim N(0, \sigma^2_m)$; mf_{kl} is the random effect of kl^{th} family within

plot, $mf_{kl} \sim N(0, \sigma^2_{mf})$ and $c(mf)_{klm}$ is the random effect of klm^{th} individual within family effect, $c(mf)_{klm} \sim N(0, \sigma^2_{cmf})$. Therefore, the total genetic variance is: $\sigma^2_g = 2\sigma^2_m + \sigma^2_{mf} + \sigma^2_{cmf}$. All random effects, including the error term, were assumed independent and identically normally distributed. Therefore, parents are assumed to be unrelated.

Based in the above linear model and the values specified in Table 2, the observed variance components for the genetic population terms (i.e. male, female, family and clonal), together with the design terms (block and plot) and environmental noise used to generate the simulated datasets are specified in Table 3.

Table 3: Observed components of variance of female (σ^2_f), male (σ^2_m), male and family (σ^2_{mf}), individual within family (σ^2_{cmf}), block (σ^2_b), plot (σ^2_p), error (σ^2) and total (σ^2_T) using for the six simulation scenarios.

	Scenarios					
	1D	1I	2D	2I	3D	3I
σ^2_f	0.025	0.025	0.025	0.025	0.100	0.100
σ^2_m	0.025	0.025	0.025	0.025	0.100	0.100
σ^2_{mf}	0.01875	0.00625	0.025	0.050	0.01875	0.00625
σ^2_{cmf}	0.13125	0.14375	0.425	0.445	0.28125	0.29375
σ^2_b	0.010	0.010	0.010	0.010	0.010	0.010
σ^2_p	0.010	0.010	0.010	0.010	0.010	0.010
σ^2	0.780	0.780	0.480	0.480	0.480	0.480
σ^2_T	1.000	1.000	1.000	1.000	1.000	1.000

The response variable for each of these scenarios was simulated based on equations 2 and 3, considering an overall mean equal to zero and the magnitude of the

variance component as specified in the Table 3. In total, 500 simulation files were generated for each of the six scenarios, in a total of 3,000 datasets available for analyses. Note that each simulation was based in an independent genetic population (i.e. 500 sets). The simulation code was implemented using the statistical software R (R Development Core Team, 2010), and the analyses were done using ASREml v.3 (Gilmour et al.,2009).

2.3. Statistical analysis of simulated data

Two statistical models were fitted to each of the datasets. First, a model based on individual plant data, called *individual model*, which used the following expression:

$$y_{ijkl} = \mu + b_i + p_{ijk} + g_{ijkl} + e_{ijkl} \quad (4)$$

where, y_{ijkl} corresponds to the observation belonging to the i^{th} block, jk^{th} plot, l^{th} genotype, μ is the fixed overall mean of each experiment; b_i is random effect of the i^{th} block, $b_i \sim N(0,\sigma_b^2)$; p_{ijk} is random effect of the jk^{th} plot within the i^{th} block, $p_{ijk} \sim N(0,\sigma_p^2)$; g_{ijkl} is random effect of the l^{th} individual genotype within the jk^{th} plot and i^{th} block, $g_{ijkl} \sim N(0,\sigma_g^2)$; and e_{ijkl} is the residual term, $e_{ijkl} \sim N(0,\sigma^2)$, which were assumed identically and independently distributed. This model provides with a breeding value prediction for each of the individuals to rank and select specific genotypes. The ASReml code for extract BLUPI values was $g = a + SCA + (d + i)_d$. Where, g is the individual genetic effect, a is the additive effect from male and female, SCA is the specific combining ability, d is the dominance effect of families and i is the epistasis effect. This equation was used together with the pedigree file. All these are important when extract the BLUPI values because of comparison with BLUPIS values

Second, plot averages obtained from the individual plant data were calculated and used to fit a linear mixed model, called *plot model*, defined as:

$$y_{ijk}=\mu + b_i + mf_{jk} + e_{ijk} \quad (5)$$

where, y_{ijk} corresponds to the observation belonging to the i^{th} block, jk^{th} family, μ is fixed the overall mean, b_i is random effect of the i^{th} block, $b_i \sim N(0, \sigma_b^2)$;$mf_{jk}$ is random effect of the jk^{th} family, $mf_{jk} \sim N(0, \sigma_{mf}^2)$; and e_{ijk} is the residual term, $e_{ijk} \sim N(0, \sigma^2)$, which were assumed identically and independently distributed. All random effects were considered independent, including families. This model was used to extract the BLUPIS values through of $g = 0.5(a_m + a_f) + SCA$ where g is the genetic effect of each family composed of a_m and a_f which are the dominance effect from families and SCA is the specific combining ability which give the additive effect from male and female (parents). The need of this type of model is common where the only information available is plot means (e.g. BRIX from a sample of plants). This model provides with a BLUP value for each of the families, allowing for ranking and selection of the best families. However, there are no individual predictions.

2.4. Selection methodologies

Under ideal situations, to perform forward selections an individual BLUP values (genetic or breeding value) should be available for each plant. These predictions can be used to make selections of individual plants within families or across families. Several criteria of family or individual selections can be implemented. In this study, an overall ranking of all individuals across families using the results from the individual model was obtained and the top *n* genotypes were selected. This methodology is called BLUPI as is based in individual BLUP values.

Alternatively, using partial information, i.e. plot means, it is possible to obtain a BLUP values for each of the families from which to make selections. Using BLUPIS, a proportion of families is determined together with the number of individuals n_k to select within those families. Later, *n* individual plant phenotypes of those identified families

are chosen by mass-selection of the best genotypes within a plot based on the trait of interest or a highly correlated trait. This methodology, described by Resende and Barbosa (2006) is called BLUPIS. Further details are provided in Section 2.5. For the present study, the number of individuals to select from the top family, i.e. n_0 was set to 50, which corresponds to an effective population size number of 96% of the maximum Ne of a family.

2.5. Simulated individual BLUP methodology

This methodology was development by Resende and Barbosa (2006) aiming to determine the number of individual to select within the best families, without considering individual plant phenotypes. BLUPIS does not consider those families with estimated negative genotypic effects, i.e., those with value below the average (Resende and Barbosa, 2006). The total number of individuals to be selected, n_k, is determined by the expression: $n_k = (g_k/g_0)\times n_0$, where g_0 refers to the predicted genotype value of the best family, g_k refers to the genotypic value of the k^{th} family, and, n_0 equals the number of individuals selected from the best family (Resende, 2004; and Resende and Barbosa 2005, 2006). The determination of $n_0$is based on the concept of effective population size (*Ne*), which according to Vencovsky (1978) is expressed as:

$$Ne = (2n_0)/(n_0+1)$$

where, n_0 here is the number of individuals required to achieve the maximum percentage of the population representativeness, where *Ne* is the effective population size. For example, to achieve a 0.96 of the effective population size (maximum family representative), n_0 should be equal 50 individuals.

2.6. Comparisons of the Methodologies

Both of the methodologies described in Section 2.4, BLUPI and BLUPIS, were compared using different statistics, such as bias and correlations.

In order to compare the selection methods, first, for each of the simulations, the top *n* individuals per family were determined using their real genetic values. Later, the results from the individual analyses (BLUP values) were used to perform the selection of the top *n* individuals to conform BLUPI method. For BLUPIS, the results from the family model identified the top families, and later, using the phenotypic values within each plot.

Initially, the top individuals per family were determined using the real genetic values, BLUPI and BLUPIS. The number of concurrencies between all possible pairs of methods was obtained, and used to obtain summary statistics.

2.6.1. Differences of absolute values between the real and predicted value by both methods BLUPI and BLUPIS

The dataset estimated by true value, BLUPI and BLUPIS were used to obtain the three absolute difference values: First, the absolute difference between the real genetic value and BLUPI, second are between the real genetic value and BLUPIS and the third are between BLUPI and BLUPIS.

2.6.2. Correlations values

After obtaining the simulated dataset, a comparison was made between the coincidence of the number of superior individuals selected in each family by two methodologies BLUPI and BLUPIS with the real genetic value. For each of the 500 files, based on the number of individuals selected in each of the 20 families were estimated three correlations values:

1) Between the real genetic value and BLUPI: $r_{realgenetic;BLUPI} = \frac{COV_{(realgenetic;BLUPI)}}{\sigma_{realgentic} * \sigma_{BLUPI}}$

2) Between the real genetic value and BLUPIS: $r_{realgenetic;BLUPIS} = \frac{COV_{(realgenetic;BLUPIS)}}{\sigma_{realgenetic} * \sigma_{BLUPIS}}$

and,

3) Between BLUPI and BLUPIS: $r_{BLUPI;BLUPIS} = \frac{COV_{(BLUPI;BLUPIS)}}{\sigma_{BLUPI} * \sigma_{BLUPIS}}$

Where, *COV* is the covariance between respective selection method and σ is the standard deviation between the respective methods. The totals of correlation were 500 for each comparison.

3. RESULTS AND DISCUSSION

Table 4 has the absolute values of the difference between the selection methodologies; the smallest differences in this case are the best results. The difference d1 showed highest average for the scenario 2I. It can be explained due to the high values of epistasis (Table 2) and low values of dominance and additives presented in this scenario. The variance component of epistasis (σ_i^2) is high within family than among families.

The scenario 3D showed the best average for d1, and it is not very different from scenario 3I, which can be explained by present high values of additive variance and low values of epistasis. Oliveira (2007) reports, that the additive variance component is important influence on the response to selection in sugarcane.

The average differences (d1) found in scenario 1 (D and I) almost not differ from those in scenario 2 (D and I), the scenario 2 showed a bit higher values than scenario 1, it could be explained because both scenarios were generated with the same values of additive variance, different non-additive and different values of environmental. The high environmental variance for scenario 1 (Table 3) can be

confounded with the high values of epistasis (Table 1) from scenario 2.For scenario 3 (D and I) all results indicated smaller average values and standard deviation when compared to the other scenarios.

The results for d2, i.e. the comparison between the real genetic values with BLUPIS, was better for all scenarios than the ones obtained for d1; however, standard deviations tended to be larger than d1. This indicates that better agreements are obtained using BLUPIS but at the cost of a small increase on variability.

As expected, the average values of the differences d3 (BLUPI and BLUPIS) were lower than the other comparison of differences (d1 and d2). The average of difference ranged from 2.79 to 4.97, and the standard deviation ranged from 0.75 to 1.13. The scenario 3 (D and I) showed the lowest values of differences, it can be explained due the lower component of non-additive variance because the BLUPIS is really on $\hat{g}$ family. This agreement between BLUPI and BLUPIS is due to the same derivation of estimators

In summary, it can be concluded that the BLUPIS methodology showed good results for all scenarios, where the best results are obtained for the situation with high additive variance and a low non-additive contribution.

A further comparison, used to evaluate BLUPI (see Table 5) was obtained by calculating the relative coincidence (in percentage) on the selected individuals between real genetic values and BLUPI.

Table 4: Average, standard deviation (SD) and coefficient of variation (CV%) of absolute value of difference between the number of selection through of the real genetic value and BLUPI (d1), real genetic value and BLUPIS (d2) and between BLUPI and BLUPIS (d3).

	Scenarios					
	1D	1I	2D	2I	3D	3I
	d1	d1	d1	d1	d1	d1
Average	12.45	12.96	12.89	13.45	9.01	9.36
SD	2.04	2.08	2.01	2.01	1.94	1.84
CV(%)	16.38	16.05	15.6	14.94	21.53	19.66
	d2	d2	d2	d2	d2	d2
Average	9.78	10.12	10.5	10.7	8.7	9.1
SD	2.08	2.08	2.15	2.04	2.05	2.00
CV(%)	21.27	20.55	20.48	19.07	23.64	21.97
	d3	d3	d3	d3	d3	d3
Average	4.75	4.97	4.43	4.91	2.79	2.85
SD	1.12	1.09	1.12	1.08	0.78	0.75
CV(%)	23.58	21.93	25.51	21.99	27.95	26.31

To prove the efficiency of BLUPI (Table 5) the comparison of relative coincidence of number of selected between real genetic values and BLUPI were realized.

Table 5: Comparison the coincidence of number of selected between the real value and Individual BLUP.

	Scenarios					
	1D	1I	2D	2I	3D	3I
Average	24.6	23.9	31.6	29.6	34.6	35.1
SD	5.2	5.4	6.5	6.1	5.5	4.9
Range	[9.4-38]	[10.4-55]	[11-48.6]	[10.2-45.1]	[4.3-63.1]	[21.9-51.3]

In scenario 1 there was the lowest percentage of efficiency (coincidence), followed by the scenario 2I and scenario 3 (D and I) showed the high percentage of coincidence, both are higher than 30%. It can be concluded that BLUPI was more efficient with high additive variance and heritability in narrow and broad-sense.

The Table 6 gives the information gain between the real genetic value and BLUPI. It expresses the ratio of realized genetic gain (based on BLUPI selection), in comparison to the maximum gain that could be achieved based on selection using the real genetic values. As before, the best realized gains were found on scenario 3 (D and I). These results reinforce the idea that for high values of the additive variance and low non-additive variability, it is possible to achieve interesting gains.

Table 6: Percentage of gain of selection in comparison true value and individual BLUP selection method.

	Scenarios					
	1D	1I	2D	2I	3D	3I
Average	33	32	46	43	52	52
SD	0.09	0.08	0.08	0.08	0.08	0.07
Range	[0.072-0.58]	[-0.02-0.51]	[0.2-0.65]	[0.18-0.74]	[0.05-0.79]	[0.25-0.66]

Correlation values (Table 7) were obtained the correlation values based on the coincidence of the number of individual selected by family in each method within each scenario.

The correlation for the scenario 1D showed average of0.533, 0.593 and 0.936 respectively. Correlations between BLUPI and BLUPIS were high (average of 0.936), with little variation. The methodologies BLUPI and BLUPIS reported similar results. It can be explained because both methods use the predicted genotypic value to make the selection, where the method BLUPIS automatically eliminates the families with negative genotypic values and BLUPI method has the highest concentration of selected individuals in families with positive genotypic effect.

The scenario 1D showed higher correlation values than scenario1I. These results can be explained due to differences between the values of the components of variance

non-additive (dominance and epistasis effects) that are larger in the scenario1I, since the values of additive variance and heritability are the same for both scenarios. According to Cruz (2005), dominance causes difficulties in improving when you want to make a distinction between a homozygous dominant to heterozygous. There was not marked difference between the magnitude of the correlation values when comparing the methodologies BLUPIS with BLUPI in both scenarios.

For scenarios 2D and 2I, the range of the correlations for two comparisons between the real genetic value with BLUPI, and it with BLUPIS presented minimum values negative. The same occurred in the scenarios 1D and 1I. Correlations between BLUPI and BLUPIS, for both scenarios (2D and 2I) showed high correlation values with equal range, with average 0.942 (2D) and 0.933 (2I). Due these values were very close, can be considered that the method BLUPIS may be used with good efficiency relative than BLUPI.

The 2D scenario (Table 7) showed values slightly higher than scenario 1I. These values can be explained by high values of broad sense heritability for the 2D scenario. It must also consider the scenario 2D shows high values of non additive variance, which could result in lower values of correlation. However this did not occur due to high values of broad sense heritability and genetic variance which is half of the total variance. According to Falconer (1989) high values of heritability in the broad sense are important for the selection, as in the case of plan with vegetative propagation; the genotype will be inherited by offspring entirely.

For the 3D scenario (Table 7) the average values of each correlation were, respectively, 0.69, 0.726 and 0.967. For the scenario 3D and 3I, no negative values were found. The average correlation values presented in the scenario 3I do not differ from those found in the 3D scenario. The high values found in this scenario are due to the high values of additive and genotype variance; heritability (narrow and broad-sense)

and low non additive variance. According to Cruz (2005), high values of heritability result in high correlation between genotypic and phenotypic values that the differences measured between families translate into true genetic differences and ensure the success of the selection strategy that is being adopted.

According to Resende (2002), the optimal strategy for selection of individuals in the early stages of the improvement of sugarcane would be through the genotypic BLUPI that would include information simultaneously from family and individual within family. The Individual BLUP uses all the effects of statistical model, consider unbalance, uses the genetic relationship between individuals in evaluation, and consider the coincidence between the unit of selection and recombination. However, this method has not been used in the improvement of sugarcane because of practical difficulties in obtaining data from individual plants. The BLUPIS can be an alternative to selection method in this case (Resende and Barbosa, 2006), because the high correlation between BLUPI and BLUPIS show that BLUPIS is a good option for selection methods. For all scenarios the correlations were above 0.93, which indicates the efficiency of BLUPIS.

Resende and Barbosa (2006) found optimum rate agreement between the BLUPI and BLUPIS in terms of coincidence of number of individuals selected by the two procedures. The correlation was 0.955. The authors concluded that BLUPIS is recommended for the genetic improvement of species for which data recorded at the family level (total harvest of the plot) is operationally easier than at individual plant level; such is the case of sugarcane. This present work found correlation between de number of individual selected within families by BLUPI and BLUPIS methodologies ranged from0.42 to 0.99 for scenario 1D, from 0.87 to 0.99 for scenario 1I and scenario 2(D and I) and from0.9 to 0.99 for scenario 3D and 3I.

Studies comparing BLUPI and BLUPIS in *Stylosanthes* spp., found correlation ranging from 0.6 to 0.83. This study validate the efficiency of the BLUPIS method and

based on this results they suggest the use of BLUPIS for other species such as *Arachis* spp., *Cajamus cajan* and other forage legumes of economics importance, where the evaluations are of plot totals. This method is also adequate in the evaluation of progeny test with *Panicum maximum* and *Brachiaria* spp. (Resende et al., 2006).

Oliveira et al., (2011) comparing BLUPIS with individual selection in full sib families of sugarcane found that for individual selection, the number of total individual selection only in families with positive genetic effect was inferior to BLUPIS, might be due the individual selection based in the indirect characters of production. It could be to lower efficiency of selection for this character. Also, the individual selection uses the phenotypic values to make the selection and BLUPIS only use the positive genetic values.

In addition, the BLUPIS methodology provides the possibility to evaluate a larger number of families in less time, increasing the breeding program and allowing that more strategies can be used after the initial phase (Resende and Barbosa, 2006 and Silva, 2009).

Table 7: Means of correlation values (Means), Standard deviation (Sd) and Range between real genotypic value and individual BLUP (1), real genotypic value and simulated individual BLUP (2) and individual BLUP and simulated individual BLUP (3).

Correlation values						
	1D			1I		
	1	2	3	1	2	3
Mean	0.533	0.593	0.936	0.486	0.552	0.932
Sd	0.19	0.17	0.03	0.20	0.19	0.02
Range	[-0.17-0.93]	[-0.08-0.96]	[0.42-0.99]	[-0.34-0.93]	[-0.18-0.93]	[0.87-0.99]
	2D			2I		
	1	2	3	1	2	3
Mean	0.561	0.620	0.942	0.508	0.573	0.933
Sd	0.17	0.16	0.023	0.18	0.17	0.023
Range	[-0.27-0.89]	[-0.1-0.93]	[0.7-0.91]	[-0.23-0.87]	[-0.05-0.92]	[0.87-0.99]
	3D			3I		
	1	2	3	1	2	3
Mean	0.698	0.726	0.967	0.688	0.716	0.967
Sd	0.14	0.13	0.017	0.13	0.126	0.018
Range	[0.16-0.96]	[0.22-0.96]	[0.91-0.99]	[0.17-0.96]	[0.20-0.94]	[0.9-0.99]

Table 8, the total of number of individual selection through the BLUPI and real genotypic value were very similar in the scenario 1 (D and I). The intersection values also were very similar between the number of selection by real genotypic value and BLUPI for this scenario. The standard deviation also were very closed, the scenario 1I being the lower values.

The scenario 2 already, show the differences between the parameters estimated. There are higher values of number of individual selection than 2I and with higher

number of concordance between the individual selected, but with high standard deviation. There were difference between the number of individual selected for this scenario being the scenario 2D more representative.

The average of number of individual selected in scenario 3 (D and I) were higher for 3I than 3D, also the scenario 3I show higher coincidence values of number of individual selected than 3D. The values of standard deviation were closed for scenarios (D and I).

Due the number of individual selected did not show higher differences, can be concluded that the variation of genetics scenarios were not affect the selection methods. It is means that can use the BLUPI for make selection.

Table 8: Number of individual selection between BLUPI and real genotypic value.

	Scenarios					
	1D	1I	2D	2I	3D	3I
Average	219	218	226	221	216	223
Intersection*	55	53	74	67	76	80
Minimum	94	109	109	112	111	116
Maximum	369	366	388	343	391	394
SD	47.06	45.97	49.88	46.39	47.60	48.03

* Intersection is the concordance between numbers of individual selection in BLUPI and real genotypic value.

For the table 9, there are not differences between the number of individual selected through the BLUPI and BLUPIS methods. The higher values of number of total individual selected were found for scenario 2D, follow by scenario 3I. The maximum and minimum values found were the same for each method inside of each scenario. Also its means that can use the BLUPI for making selection than BLUPI due the facility.

Table 9: Number of individual selection between BLUPI and BLUPIS.

	Scenarios			
	1D		1I	
	BLUPIS	BLUPI	BLUPIS	BLUPI
Average	218	219	218	218
Minimum	94	94	109	109
Maximum	369	369	366	366
SD*	47.05	47.06	45.98	47.97
	2D		2I	
	BLUPIS	BLUPI	BLUPIS	BLUPI
Average	226	226	221	221
Minimum	109	109	112	112
Maximum	388	388	343	343
SD	49.90	49.88	46.37	46.39
	3D		3I	
	BLUPIS	BLUPI	BLUPIS	BLUPI
Average	216	216	223	223
Minimum	111	111	116	116
Maximum	391	391	394	394
SD	47.58	47.60	48.05	48.03

* SD is the standard deviation

4. CONCLUSION

The results of this study indicate that high correlation average values between the BLUPIS and BLUPI in all scenarios were obtained (0.936, 0.932, 0.942, 0.933, 0.967 e 0.967) in an efficiency studies in sugarcane referring to the number of individuals to be selected per family.

The use of BLUPIS is recommended for genetic improvement and proved to be effective for different scenarios based in diverse values of additive, dominance and epistasis variance components, with minimal loss of information.

BLUPIS provided high correlation with different additive variance and heritability values. The higher values of heritability combined with the higher values of additive, the most efficient is BLUPIS.

Also the BLUPIS method can be used with success in the selection, when the heritability and additive variance component are small.

GENERAL REFERENCES

BABU, C.N. Sugarcane. 2oEd. Allied Publish: New York, 252p. 1990.

BARBOSA, M. H. P.; SILVEIRA, L. C. I. Metodologias de seleção, progressos e mudanças no programa de melhoramento genético da cana-de-açúcar da Universidade Federal de Viçosa. STAB, Açúcar, Álcool e Subprodutos, Piracicaba, v.18, n.3, p.30-32, 2000.

BARBOSA, M. H. P.; SILVEIRA, L. C. I. Melhoramento genético e recomendação de cultivares. 2o Ed. In: Santos, F.; Borém, A.; Caldas, C. Cana-de-açúcar bioenergia, açúcar e etanol – Tecnologias e Perspectivas. Editora UFV: Viçosa, 637p. 2012.

BASTOS I. T. et al. Análise dialélica em clones de cana-de-açúcar. Bragantia, Campinas, v.62, n.2, p.199-206, 2003.

BRESSIANI, J. A. Seleção seqüencial em cana de açúcar. 133p. Tese de Doutorado em Agronomia - Genética e Melhoramento de Plantas. Escola Superior de Agricultura Luiz de Queiroz – Universidade Federal de São Paulo, Piracicaba, 2001.

BRESSIANI, J. A., VENCOVSKY, R. and BURNQUIST, W. L. Modified family selection in sugarcane. Proc. Int. Soc. Sugarcane Technol., 25:459-467, 2005.

BRESSIANI, J.A. et al. Breeding sugarcane for leaf scald resistance: A genetic study. Journal American Society Sugarcane Technologists, v.27, p. 15-22, 2007.

CALIJA, V.; HIGGINS, A. J.; JACKSON, P. A. BIELING, L. M.; COOMANS, D. An operations research approach to the problem o the sugarcane selection. Annals of Operations Research: Netherlands, v.108, p.123-142, 2001.

CAVALCANTI, C. J. C. Estimativas de parâmetros genéticos de alguns caracteres agrícolas da cana de açúcar (Saccharum spp.). 128p. Dissertação de Mestrado em Botânica – Melhoramento Genético de Plantas. Universidade Federal Rural de Pernambuco, Recife, 1990.

CHAUDHARY, R.R. Genetic variability and heritability in sugarcane. Nepal Agric. Res. J, v. 4, n.5, p. 56-59, 2001.

COMPANHIA NACIONAL DE ABASTECIMENTO. Cana de açúcar: Safra 2009/2010 terceiro levantamento. Brasilia, 16p, 2011. (Relatório de Acompanhamento da safra Brasileira).

COX, M. C.; HOGARTH, D. M. The effectiveness of family selection in early stages of a sugarcane improvement program. In: Australian Plant Breeding Conference, v.10. Bundaberg, 1993.Procedings… Brisbane, p.53-54, 1993.

CRUZ, C. D. Programa Genes: aplicativo computacional em genética e estatística. Viçosa: Imprensa Universitária da UFV, 2001, 442p.

CRUZ, C. D. Princípios da Genética Quantitativa. Viçosa: UFV, 2005, 394p.

DUDLEY, J. W. & R. H. MOLL. Interpretation and use of estimates of heritability and genetic variances in plant breeding. Crop Sci. v.9, p. 257-262, 1969.

FALCONER, D.S. Introduction to quantitative genetics.3 ed. London: Logman, 1989.464p

FERNANDES, V.B. Eficiência de dialelos circulantes para escolha de genitores baseados na simulação de dados. 82p. (Mestrado em Estatística e Experimentação Agropecuária). Universidade Federal de Lavras, Lavras, MG, 2006.

FERREIRA, F. M. et al. Effects of inbreeding on the selection of sugarcane clones. Crop Breeding and Applied Biotechnology, Londrina, v.5, p.174-182, 2005.

FERREIRA, F.M. et al. Relações fenotípicas e genotípicas entre componentes de produção em cana de açúcar. Bragantia, v.66, n. 4, p. 527-533, 2007.

GILMOUR, A. R., THOMPSON, R. AND CULLIS, B. R. AI, an efficient algorithm for REML estimation in linear mixed models, Biometrics, 51: 1440–1450, 1995.

GILMOUR, A.R. et. al. ASReml. Userguide. VSN INTERNACIONAL Ltd, HEMEL HEMPSTEAD, 2009.

GURGEL, F. L. Simulação computacional no melhoramento genético de plantas. 187p. (Doutorado em Genética e Melhoramento). Universidade Federal de Lavras, Lavras, MG, 2004.

HENDERSON, C. R. Applications of linear models in animal breeding. Guelph: University of Guelph, 1984. 462p.

HOGARTH, D.M. Quantitative inheritance studies in sugarcane. I. Estimation of variance components. Australian Journal of Agricultural Research, v. 22, p. 103-109, 1971.

HOGARTH, D.M.; WU, K.K.; HEINZ, D.J. Estimating genetic variance in sugarcane using a factoring cross design. Crop Science, v.21, p. 21-25, 1981.

HONG-KAI, Z. et al. Genetic analysis of sugarcane biomass yield and its component traits using ADAA models. Journal of Tropical Agriculture.v.47, n. 1-2, p. 70-73, 2009.

KENNEDY, A. J. and BELLAMY, S. The experimental basis and pratical application of family selection as a tool in the early stages of selection of varieties of sugarcane in Barbados. Proc. WIST XXVI Conference, Guyana, p. 56-63, 1997.

LITTELL, R. C. et al. SAS for Mixed Models, SAS Institute Inc: North Caroline, 2006. 820p.

LIU, G. et al. Genetic analysis for brix weight per stool and its component traits in sugarcane (Saccharum officinarum).Journal of Zhejiang University Science B, v.8, n.12, p. 860-866, 2007.

MACEDO, G.A.R. et al. Recomendações técnicas para o cultivo da cana de açúcar. EPAMIG Centro Oeste, Minas Gerais, 2010.

MACHADO, E. C.; PEREIRA, A. R.; FAHL, J. I.; ARRUDA, H. V.; CIONE, J. Índices biométricos de duas variedades de cana-de-açúcar. Pesquisa Agropecuária Brasileira, Brasília, v.17, n.9, p.1323-1329, 1982.

MARTIN, A,L,A. et al. Genetic variability and pedigree analysis of Brazilian common bean elite genotypes. Scientia Agricola, v. 60, n. 2, p.283-290, 2007.

MATSUOKA, S.; GARCIA, A. A. F.; CALHEIROS, G. G. Hibridação em cana-de-açúcar. In: BORÉM, A. (ed.). Hibridação Artificial de Plantas. 1° ed. Universidade Federal de Viçosa,Viçosa : Editora da UFV, v.1, p.221-254, 1999.

MINISTÉRIO DA AGRICULTURA. Portal Desenvolvimento Agrário, Safra, 2011. Disponível em: www.mda.gov.br/portal/. Acesso em: 12/09/2011.

NATIONAL AGRICULTURAL STATISTICS SERVICE (NASS). Crop production: 2010 Summary, USDA, 2011. Disponível em: http//usda.mannilb.cornell.edu/MannUsda/viewDocumentInfo.do?documentID=1047. Acesso em: 12/09/2011.

OLIVEIRA, R. A. Seleção de famílias de maturação precoce em cana de açúcar via REML/BLUP. 142p. (Tese) Doutorado em Produção Vegetal. Universidade Federal do Paraná (UFPR), Curitiba, Paraná, 2007.

OLIVEIRA, R.A., et al. Procedimento BLUPIS e seleção massal em cana de açúcar. Bragantia, Campinas, v.70, n.4, p.796-800, 2011.

OLIVEIRA, F. J. et al. Variâncias genéticas em clones RB e variedades de cana de açúcar. In: X Jornada de Ensino, pesquisa e extensão – JEPEX 2010 – UFRPE: Recife, 18 a 22 de outubro, 2010.

PETERNELLI, L.A. et al. Delineamentos Aumentados no Melhoramento de Plantas em condições de restrições de recursos. Ciência Rural, v.39, n.9, p.2425-2430, 2009.

PINTO JUNIOR, J. E. REML/BLUP para analise de múltiplos experimentos no melhoramento genético de *Eucalyptus grandis* W. HILL EX MAIDEN. 127p. (Doutorado em Produção Vegetal). Universidade Federal do Paraná, Curitiba, Paraná, 2004.

R Development Core Team (2010) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL. Disponível em: http://www.R-project.org.

RAMALHO, M. A. P. Eficiência relativa de alguns processos de seleção intra-populacional no milho baseados em famílias não endógamas. 122 f. Dissertação (Mestrado) - Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, 1977.

RAMALHO, M. A. P.; SANTOS, J. B. dos; PINTO, C. A. B. P. Genética na agropecuária. Lavras: UFLA, 2004. 472 p.

REIS, E. Estatística multivariada aplicada. 2a Edição. Edições Silabo, LDA: Lisboa, 2001, 343p.

RESENDE, M. D. V. Genética biométrica e estatística no melhoramento de plantas perenes. Brasília: Embrapa Informação Tecnológica, 2002, 975p.

RESENDE, M. D. V. Métodos estatísticos ótimos na analise de experimentos de campo. Embrapa Florestas: Colombo, 2004. 65p. (Documento 100).

RESENDE, M. D. V.; BARBOSA, M. H. P. Melhoramento genético de plantas de propagação assexuada. Embrapa Informação Tecnológica: Colombo, 2005. 130p.

RESENDE, M.D.V.; BARBOSA, M.H.P. Selection via simulated Blup base on family genotypic effects in sugarcane. Pesquisa Agropecuária Brasileira: Brasília, v.41, n.3, p.421-429, 2006.

RESENDE, M.D.V. O software selegen REML/BLUP. Embrapa Informação Tecnológica, Campo Grande, 2007. 299p.

RESENDE, R.M.S et al. Genotypic evaluation of accessions and individual selection in *Stylosanthes spp*. by simulated BLUP method. Crop Breeding and Applied Biotechnology, v.6, p.253-260, 2006.

RICHARD, C.A. and HENDERSON, M.T. Genetic behavior of sucrose content in the parents and progeny. American Society of Sugarcane Techonologist, v. 9, p. 68-71, 1981.

ROSS, S.M. Simulation. 4oEd. Elsevier: Oxford, 2006. 298p.

SEARLE, S.R.; CASELLA, G.; McCULLOCH, C.E. Variance components. New York. J. Wiley, 1992. 528p.

SILVA, F. L. Seleção dentro de famílias de cana de açúcar via BLUP individual simulado. 65p. Tese (Doutorado em Genética e Melhoramento), Viçosa, MG: Universidade Federal de Viçosa. 2009.

SIMEÃO, R. M. et al. Avaliação genética em erva-mate pelo procedimento BLUP Individual multivariado sob interação genótipo x ambiente. Pesquisa Agropecuária Brasileira, v.37, n.11, p.1589-1596, 2002.

SKINNER, J.C., HOGARTH, D.M., and WU, K.K. Selection methods, criteria, and indices. In: Sugarcane improvement through breeding. Elsevier: Amsterdam, p. 409-453, 1987.

SOUZA JUNIOR, C. L. de. Componentes da variância genética e suas implicações no melhoramento vegetal. Piracicaba: FEALQ, 1989, 134p.

SOUZA, E.F.M., PETERNELLI, L.A., BARBOSA, M.H. Designs and model effects definitions in the initial stage of a plant breeding program. Pesquisa Agropecuária Brasileira, v.41, n.3, p.369-375, 2006.

STRINGER, J.K., et al. Family selection Improves the efficiency and effectiveness of selection original seedlings and parents. Sugar Tech, v.13, n.1, p.36-41, 2011.

TAIZ, L.; ZEIGER, E.; trad. SANTARÉM, E. R. et al. Fisiologia vegetal. 3.ed., Porto Alegre: Editora Artmed, 2004. 719p.

TYAGI, S. D. and SINGH, D. N. Studies on genetic variability for stalk characters in sugarcane. Indian Sugar XL VIII: 259-262, 1998.

VENCOVSKY, R. Effective size of monoecious populations submitted to artificial selection. Brazilian Journal of Genetics, v. 1, n.3, p. 181-191, 1978.

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

XING, Y., et al. Analysis of sugarcane mosaic virus resistance in maize in an isogenic dihybrid crossing scheme and implications for breeding potyvirus-resistant maize hybrids. Genome, v. 49, p.1274-1282, 2006.

ZHOU, H.K., et al. Genetic Analysis of genotype x environment interaction for major characters of sugarcane. J. Northwest Sci-Tech. Univ. Agric. For. (Nat. Sci. Ed.), v. 33, n.2, p. 69–73, 2005.

WU, L. et al. The ligation amplification reaction (LAR) – Amplification of Specific DNA sequences using sequential rounds of template – dependent ligation. Genomics, v.4, p.560-569, 1989.

YAN, W., HUNT, L.A. Genotype interaction and crop yield. Plant Breeding Rev., v.16, p. 135-178, 1998.

.

APPENDIX

2 sulcos de bordadura na parte superior do experimento = 8025 Plantio: 07 de maio 2007 Q7

Cada parcela representada por dois sulcos com 10 plantas cada espaçadas a 0,5m

		Rep 1				Rep 2				Rep 3				Rep 4				Rep 5			
		FAIXAS																			
	Parcela	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
PLOT	1	14	9	21	5	19	2	6	13	22	18	15	7	5046	11	18	7	6	21	2	5046
	2	16	18	1	5046	5046	5	4	21	4	5046	20	9	14	8	20	9	12	19	15	10
Exp 1	3	4	11	15	22	11	20	22	1	454	16	8	3	16	454	4	17	1	13	3	17
	4	7	17	12	19	9	15	454	12	5	19	11	2	13	22	1	5	14	454	7	4
	5	20	454	10	6	3	16	14	10	17	10	12	21	3	2	19	12	16	9	20	5
	6	8	13	3	2	7	8	17	18	14	13	6	1	21	6	15	10	22	18	8	11
	7	40	38	26	30	23	454	33	36	25	24	26	40	36	30	29	5046	42	37	44	38
	8	31	41	32	42	28	43	27	5046	29	454	38	30	39	33	38	44	25	40	26	39
Exp 2	9	35	28	44	5046	38	35	39	42	27	33	28	37	34	24	25	26	33	43	35	30
	10	33	39	24	37	26	41	44	29	23	43	32	5046	32	454	28	35	24	27	32	5046
	11	454	27	25	36	40	32	34	37	35	41	39	36	31	43	42	23	38	28	34	23
	12	43	23	34	29	24	25	31	30	42	34	44	31	37	41	27	40	29	454	41	31
	13	56	48	66	49	62	49	67	46	49	62	5046	66	45	65	51	54	55	45	5046	63
	14	45	46	65	51	54	51	63	66	51	54	50	65	48	454	66	59	61	48	64	46
Exp 3	15	50	454	68	53	50	61	5046	65	59	60	67	56	56	61	68	5046	65	454	57	68
	16	64	67	62	59	55	59	64	68	61	53	46	68	50	55	62	49	52	56	62	66
	17	61	63	57	60	53	454	52	56	52	64	55	45	67	46	57	60	49	59	50	54
	18	55	54	52	5046	60	57	48	45	454	63	48	57	64	63	52	53	60	53	67	51
	19	82	70	77	5046	73	76	118	77	76	74	70	75	89	90	73	83	78	70	47	89
	20	88	74	71	58	454	85	90	71	85	84	5046	80	88	74	77	58	74	72	85	88
Exp 4	21	89	72	69	80	70	81	89	78	82	69	89	118	454	72	71	80	90	118	81	82
	22	76	454	47	75	74	82	88	5046	81	73	83	90	82	85	69	75	76	5046	71	84
	23	118	85	78	84	69	72	75	47	454	78	58	71	76	70	47	84	69	73	75	77
	24	90	73	83	81	84	83	80	58	72	47	88	77	118	5046	78	81	454	80	58	83
	25	112	95	96	99	113	99	100	121	93	109	103	117	121	93	98	100	94	100	95	97
	26	109	107	119	117	101	117	5046	108	100	108	105	99	117	107	103	102	111	454	117	105
Exp 5	27	111	93	108	100	106	96	94	119	94	107	5046	113	454	105	113	109	103	119	99	121
	28	94	97	121	5046	102	454	109	111	97	121	111	101	106	97	5046	112	112	98	5046	96
	29	101	105	106	113	112	98	95	107	98	454	106	96	101	95	119	111	102	106	101	113
	30	454	103	98	102	105	103	93	97	95	112	102	119	99	96	108	94	107	109	108	93

FiguraI: Croqui of experimento used in this research.