Plot size in sugarcane family selection experiments

Márcio Henrique Pereira Barbosa*[1]; Luiz Alexandre Peternelli[2] and Luís Cláudio Inácio da Silveira[1]

[1] Department of Plant Science - Universidade Federal de Viçosa – UFV, CEP 36571-000, Viçosa-MG, Brasil; [2]Department of Computer Science - Universidade Federal de Viçosa – UFV, CEP 36571-000, Viçosa-MG, Brasil. E-mail: ptrnelli@dpi.ufv.br (* Corresponding Author. E-mail: mbarbosa@mail.ufv.br)

ABSTRACT

This study aimed at verifying the effect of the number of plants per plot on the evaluation of sugarcane families. Five full-sib families were evaluated in randomized complete blocks with 250 genotypes per plot in 3 replications. Each plot consisted of 5 rows with 50 plants each. Based on stalk weight means per row and different sample sizes, it was concluded that 50 plants per plot are adequate for family evaluation in replicated trials. In order to estimate refractometer solids (brix), a sample of at least 10 plants per plot is suggested, using a stalk in each plant for juice extraction. It was also concluded that family stalk yield estimates can be obtained by the product between the number of stalks of the plot and the mean weight of 30 stalks taken from different plants of the family.

KEY WORDS: Sample, *Sacharum* spp., sugarcane, family selection.

INTRODUCTION

Visual selection of individuals (mass selection) for low heritability traits in the early stages of plant development has been inefficient in various crops, including sugarcane (Skinner, 1971). Based on that, sugarcane breeding programs have applied low selection pressure (10 to 30%) in early stage trials (seedling trials) to reduce the possibility of discarding superior genotypes (Skinner et al., 1987).

Family selection is an alternative to improve sugarcane mass selection efficiency (Cox et al., 1996). In Australia, family selection has been used in Bureau of Sugar Experiment Stations (BSES) since 1986. Nowadays, all the regional programs are using family selection routinely. It has been reported that the number of superior clones detected in the later phases of the clonal test has significantly increased when family selection is combined with mass selection (Cox et al., 1996).

Applying mass selection to superior families increases the probability of identifying elite clones, thus improving the efficiency of the available resources in a breeding program. However, previous selection of superior families is necessary. Such inferences may be obtained by evaluating a group of clones which would represent each family, arranged in experimental designs (Jackson et al. 1998). Another alternative is to evaluate families using replicated trials, with plots consisting of non-cloned individuals, which would together provide information on the genetic value of the evaluated families, as applied by Stringer et al. (1996)

However, an assessment of the number of individuals necessary to represent the mean and variance of the families is required. Some studies have suggested sample sizes between 20 to 150 plants for evaluating characteristics such as Brix, stalk yield and yield components (Skinner, 1971; Wu et al., 1977; Wu et al, 1978; Mariotti et al. 1981; Peixoto and Ricci, 1984).

This study aimed at verifying the effect of the number of plants per plot for selection of sugarcane families in replicated trials.

MATERIAL AND METHODS

Genetic material

Five full–sib families obtained in 1996 at the Serra do Ouro Cross Station at the Universidade Federal de Alagoas derived from the following crosses: RB765418 x SP79-2233, SP79-1011 x RB855036, SP70-1143 x RB72454, SP70-1143 x IAC68-12 and RB72454 x SP80-1520, were used. A total of 750 genotypes were obtained for each family. Sowing, seedling transplant procedures and plant cutting without selection were those described by Barbosa and Silveira (2000).

Locality and experimental details

The experiment was set up in January 1997, at the Centro de Melhoramento da Cana-de-açúcar – CECA of the Universidade Federal de Viçosa, in Oratórios-MG. CECA is located at the Zona da Mata in Minas Gerais at 400 m altitude, $20°24'17''S$ latitude and $42°40'13''W$ longitude.

The experiment was arranged in randomized complete blocks with three replications. For each family, the experimental plot consisted of 250 genotypes arranged in subplots of 5 rows with 50 genotypes each. Plants were spaced at 50 cm with 1.40m between rows. Fertilization and other cropping procedures were those normally used for conducting trials at CECA.

Evaluated Traits

All plants were cut in July 1997, approximately 7 months after planting. Data were collected during the first 15 days of April 1998, at the first rattoon crop.

Five bundles with 10 stalks each obtained from different stools were sampled from each one of the five rows of the plots. Rows presenting less than 50 plants due to non-sprouting genotypes had some stools contributing with more than one stalk for the formation of the five bundles with 10 stalks. Thus, a total of 25 bundles with 10 stalks each was obtained per plot. All bundles were weighed separately in order to obtain the row weights in each plot as necessary.

A hydraulic press was used to extract the juice by crushing the fifth internode starting from the base of each of the 10 stalks per bundle. Brix readings of the extracted juice was carried out with a manual refractometer. All the stalks in each row were weighed on a scale.

Data analysis

Variance analyses were conducted considering the following models:

Model 1: $Y_{ijl} = m + b_j + F_i + e_{(ij)} + S_{l(i)} + d_{l(ij)}$,

used to analyse the stalk yield trait, where: Y_{ijl} = observed value at the experimental unit corresponding to the ith family, jth block and lth observation; m = overall mean; b_j = random effect of the jth block (j = 1, 2, 3); F_i = fixed effect of the ith family (i = 1, 2, 3, 4, 5); $e_{(ij)}$ = between plots experimental error; $S_{l(i)}$ = fixed effect of the lth row within the ith family (l = 1, 2, 3, 4, 5); and $d_{l(ij)}$ = within plots experimental error. And Model 2: $Y_{ijz} = m + b_j + F_i + e_{(ij)} + A_{z(i)} + d_{z(ij)}$, used to analyse the sampling effect within families for the average brix of 10 stalks and the mean weight of 10, 20 and 30 stalks traits, where: Y_{ijz} = observed value at the experimental unit corresponding to the ith family, jth block and zth observation; $A_{z(i)}$ = fixed effect of the zth sample within the ith family (z = 1, 2,..., s, where s corresponds to the maximum number of bundles or groups of bundles necessary to give the samples required for the analysis); and $d_{z(ij)}$ = within plots experimental error. The other model components are as previonsly defined.

Different within plot stalk samplings were evaluated for the stalk yield estimation. Stalk yield (kg), the total number of stalks, and the weight (kg) of 5 samples with 10 stalks each were obtained in each row. Based on these data, five stalk yield estimates per row were thus obtained: 1) estimated yield with 10 stalks = (number of stalks in a row-NC) x (mean weight of a 10 stalks sample on the same row); 2) estimated yield with 20 stalks = (NC x mean weight of a 20 stalks sample on the same row); 3) estimated yield with 30 stalks = (NC x mean weight of a 30 stalks sample on the same row); 4) estimated yield with 40 stalks = (NC x mean weight of a 40 stalks sample on the same row); and 5) estimated yield with 50 stalks = (NC x mean weight of a 50 stalks sample on the same row). Simple and partial correlations, as well as their respective tests, were performed according to Steel and Torrie (1980).

RESULTS AND DISCUSSION

A breeder always tries to apply methodologies which are suitable for the experimental evaluation of the populations submitted to selection. Plot size and shape are important aspects to be considered when evaluating sugarcane families. The experimental methodology for genotype evaluation is directly related with the efficacy of the breeding program to obtain genetic progress in the selection of quantitative inheritance traits.

The analysis of variance procedures for inferring sample size and shape used in this study to evaluate sugarcane families were: A) model 1: used for stalk yield trait. Based on this model, the F test between rows within families was calculated. B) model 2: In this case, the F

test was also calculated for sample within families for the average brix of 10 stalks and mean weight of 10, 20 and 30 stalks traits.

According to the analysis of variance (model 1) for stalk yield, no significant difference was detected between rows within families by the F test ($P>0.05$), at a reasonable experimental precision (CV% = 13.70; Table 1). This showed that the rows with 50 plants had the same mean. Thus, the basic unit of 50 plants provided the same stalk weight information given by the whole-plot with 250 plants. It is important to mention that this study assumed that a 250 genotype per plot sample would be sufficient to estimate the population mean by plot, since the genetic effect of the ith family was expected to be the same in the three replications.

This study was conducted on five sugarcane families only, and since its objective was to maximize plot size up to 250 genotypes, a larger number of individuals were used within each plot rather than a large number of families. Also, it was impossible to increase the number of families considerably due to the limited financial resources available for carrying out the experiment. Table 1 shows no significant differences by the F test among families, as expected.

It should be emphasized that the objective of this study was to obtain inferences on the family means rather than on within family variance. In fact, there is a higher probability of selecting superior individuals in families with a superior mean and genetic variance. However, the estimates of variance within families per stalk yield and its components are strongly influenced by the environment since the estimates are obtained at individual levels. Collecting such information is a rather slow and laborious process to be implemented a breeding program routine task.

Early family selection is an alternative procedure to improve mass selection efficiency, since the selection will be applied to families with higher

Table 1 – Summary of variance analysis for the stalk yield per row

Variation source	Gl	QM
Replication	2	4911.32
Family	4	3312.07 [n.s.]
Error a	8	6041.90
Row (family)	20	481.93 [n.s.]
Error b	40	1049.85
Total	74	

[n.s.] - non-significant at 5% probability by the F test.
$CV_a\%=32.8$ $CV_b\%=13.7$ mean=236.37Kg

means. Family assays can also be applied for top–cross evaluation. Thus, clones or varieties presenting good combining ability could be later used in a large scale to obtain populations which will be submitted to mass selection.

Stalk yield estimation can also be obtained by the product between the number of stalks and stalk mean weight traits. This is especially important when the goal is to obtain seedlings for immediate multiplication of the individuals selected or when it is not possible to weigh all the plots, which, in the final analysis, would be the best option. Thus, within plot stalk samplings were used to obtain yield estimates. The interpretation of the results was done through simple and partial correlations, presented in Table 2.

The yield component number of stalks was found to be more important ($r=0.71**$) than the stalk mean weight for yield estimation under different sampling sizes (r varying from 0.13 up to 0.49; data not shown). When the effect of number of stalks was excluded, smaller correlations were also detected between the estimated and observed stalk yield (Table 2). This was also highlighted when the effect of the mean weight of 10 stalks was excluded from the correlations between the estimated and observed stalk yield. For the other sample sizes (20, 30, 40 and 50 stalks), the partial correlations, when the effect of the mean weight of stalks is excluded were practically similar to the simple correlations obtained between the stalk yield and the estimated stalk yield.

In order to verify the weight variability of the 10-stalk sample weights, analyses of variance were carried out according to model 2. Since there were 25 bundles of 10-stalks in each plot, 25 was the number of 10-stalk samples used in this first analysis. In Table 3, significant differences ($P < 0.05$) were observed among 10-stalk samples within families. This shows that the sample size of, approximately, 5%, i.e., a bundle with 10 stalks, was not sufficient to give reliable estimates of family stalks mean weight. Thus, a similar analysis was carried out with a 20-stalk sample, i.e., approximately 10% of the total number of stalks in the plot. A total of 12 different 20-stalk samples were obtained in each plot. Each 20-stalk sample mean was obtained as the mean weight of two 10-stalk bundle in a particular plot. According to Table 3, significant differences ($P < 0.05$) among samples within the family were observed, providing the same conclusion as that for the 10-stalk sample analysis. A similar analysis

was also carried out with a 30-stalk sample, i.e., approximately 15% of the total number of stalks in the plot. In this case, a total of 8 different 30-stalk samples were obtained in each plot. Each 30-stalk sample mean was obtained as the mean weight of three 10-stalk bundle in a particular plot. For this analysis no significant differences (P > 0,05) among samples within families were observed (Table 3). This indicates that the mean weight of 30-stalk samples did not differ significantly by the F test. Through correlations (Table 2) and analyses of variance (Table 3), it can be concluded that samples sizes equal to 30 stalks in the plot are necessary to estimate stalk mean weight, and, consequently, yield estimated by the number of stalks in the plot and the stalk mean weight. This result is in agreement with that obtained by Mariotti and Lascano (1969), that suggest a sample of 30 stalks (r = 0.80**) to obtain yield estimates in clone evaluation.

Similarly, for the brix variable, the sample size needed to represent the family mean was investigated. Thus, brix was obtained in bundle samples with 10 stalks per each row, with a total of 25 samples per plot.

Significant differences among samples within families where not detected by the F test (P > 0,05), according to model 2 (table 4).Therefore, it was concluded that only 10 stalks taken from different plants in the same plot provided reliable estimates of average Brix per family. These results agree with those obtained by Barbosa et al. (1999), who used Lessman and Atkins (1963) and Meier and Lessman (1971) methodology to conclude that a sample of 7% of the genotypes, approximately 16 plants per plot, represent the family brix average.

Few studies have related sample sizes in family evaluation. Some inferences on the family performance have been obtained through combined data on the clones which were evaluated in one-row plots, normally 5 to 8 m long. Mariotti et al. (1981) concluded that a random sample of 20 clones, evaluated in the stage 2 trial, is sufficient to estimate the family mean for sucrose production and yield with a reasonable precision.

Another alternative procedure to obtain inferences on the families can be employed

Table 2 – Simple and partial correlations between stalk yield and estimated stalk yield, adjusted for stalk mean weight and number of stalks, for different sample sizes.

Variable Estimated production (column X)	Partial correlation of the column X variable with stalk yield without the effect of number of stalks	Variable Mean weight (column Y)	Partial correlation of column X variable with stalk yield without the effect of column Y variable	Simple correlation column X variable with stalk yield
with 10 stalks	0.11	of 10 stalks	0.63**	0.67**
With 20 stalks	0.49**	of 20 stalks	0.70**	0.78**
With 30 stalks	0.51**	of 30 stalks	0.78**	0.79**
With 40 stalks	0.59**	of 40 stalks	0.82**	0.82**
With 50 stalks	0.59**	of 50 stalks	0.82**	0.81**
Number of stalks	-	-	-	0.71**

** significant at 1% probability by the t test (Ho: $r_{xy} = 0$).

Table 3 – Summary of the variance analysis for the mean weight of stalks using 10, 20 and 30-stalk samples

Variation Sources	10-stalk samples		20-stalk samples		30-stalk samples	
	Df	MS	Df	MS	Df	MS
Replication	2	65.62	2	32.54	2	21.66
Family	4	65.37*	4	30.23 [n.s.]	4	20.15 [n.s.]
Error a	8	15.77	8	7.97	8	5.32
Sample (family)	120	4.05*	55	2.70*	35	1.41 [n.s.]
Error b	240	3.00	110	1.60	70	1.08
Total	374		179		119	
$CV_a\%$		38.06		27.02		22.07
$CV_b\%$		16.61		12.11		9.97
Mean		10.44		10.44		10.44

* - significant at 5% probability by the F test;
[ns] - non-significant at 5% probability by the F test.

Tabela 4 - Summary of variance analysis for brix[1].

Variation source	Df	MS
Replication	2	21.36
Family	4	5.03 [n.s.]
Error a	8	3.15
Sample (family)	120	1.18 [n.s.]
Error b	240	1.06
Total	374	

[1] Brix: reading on the juice extracted from 10 stalks of different genotypes;
[n.s.] - non-significant at 5% probalility by the F test.

during stage 1 trial, at the stage of individual plant-cane plant seedlings. In this case, plots would contain genotypes which, if combined, could provide the family mean value for a particular character. Within this same context, Wu et al (1977 and 1978), applied the Monte Carlo method to determine that the plot size of 40 genotypes per family was adequate to estimate Brix mean and variance, stalk number, stalk diameter and stalk height. On the other hand, Skinner (1971) recommended a larger plot containing 71 plants. These results are similar to those obtained in this study, where plots with 50 genotypes would be sufficient to obtain inferences in yield of stalks and their components.

CONCLUSIONS

a) Plots with 50 plants are sufficient to estimate a family stalk yield.
b) The family brix could be precisely estimated by a sample of 10 plants per plot, using one stalk from each stool for juice extraction.
c) Family stalk yield estimates can be obtained by the product between the number of stalks of the plot and the mean weight of 30 stalks taken from different plants of the family.

ACKNOWLEDGEMENTS

RESUMO

Tamanho da parcela em experimentos para seleção entre famílias em cana-de-açucar

O presente trabalho teve como objetivo verificar o efeito do número de plantas por parcela na avaliação de famílias de cana-de-açúcar. Foram avaliadas cinco famílias de irmãos completos em blocos casualisados com 250 genótipos por parcela em três repetições. Cada parcela foi constituída por cinco sulcos com 50 plantas cada um. Considerando as médias de peso de colmos por sulco, conclui-se que 50 plantas por parcela são suficientes para avaliação de famílias em ensaios com repetição. Para estimativa do Brix em refratômetro de campo sugere-se uma amostra mínima de 10 plantas por parcela, tomando-se para extração do caldo um colmo em cada touceira. Também foi concluído que estimativas da produção de colmos por familia podem ser obtidos pelo produto entre o número de colmos na parcela e o peso médio de 30 colmos obtidos de plantas diferentes na familia.

REFERENCES

Barbosa, M.H.P and Silveira, L.C.I. 2000. 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. 18(3):30-32.

Cox, M.C.; McRae, T.A .; Bull, J.K. and Hogarth, D.M. 1996. Family selection improves the efficiency and effectiveness of sugarcane improvement program. p.42-43.

In: Wilson, J.R.; Hogarth, D.M.; Campbell, J.A.; Garside, A.L. (Ed.). Sugarcane: Research towards efficient and sustainable production. CSIRO Divison of Tropical Crops and Pasture, Brisbane.

Jackson, P.; McRae, T.; Morgan, T. and Elliott, S. 1998. F1 optimal plot size, replication and selection criteria in early selection stages. In: ISSCT breeding and germplasm workshop. 4th , May, 1998. (Abstract:http://www.sugarnet.com/issct/f1.htm).

Lessman, K.J. and Atkins, R.E. 1963. Optimum plot size and relative efficiency of lattice designs for grain sorghum yield tests. Crop Science. 3:477-481.

Mariotti, J.A. and Lascano, O.G. 1969. Estudios de muestro para la evaluacion del rendimiento de la caña de azucar. Revista Industrial y Agrícola de Tucumán. 46(2):37-44.

Mariotti, J.A.; Scandaliaris, J.; Ahmed, M.A. and Cuenya, M.I. 1981. Analisis del comportamiento familiar en progenies hibridas de cana de azucar. Revista Industrial y Agrícola de Tucumán. 58(2):1-14.

Meier, V.D. and Lessman, K.J. 1971. Estimation of optimum field plot shape and size for testing yield in *Cramble abyssinica* Hochst. Crop Science. 11:648-650.

Peixoto, T.C. and Ricci Jr. A. 1984. Amostra mínima representativa para caracterização da variabilidade expressa na descendência de progenitores de cana-de-açúcar. Boletim Técnico Copersucar. 28:27-30.

Skinner, J. 1971. Selection in sugarcane: a review. Proceedings of the International Society of Sugarcane Technologists. 14:149-162.

Skinner, J.C.; Hogarth, D.M. and Wu, K.K. 1987. Selection methods, criteria and indices. p.409-453. In: Heinz, D.J. (Ed.). Sugarcane improvement through breeding. Elsevier, Amsterdam.

Steel, R.G.D. and Torrie, J.H. 1980. Principles and procedures of statistics. A biometrics approach. McGraw Hill, New York. 633p.

Stringer, J.K.; McRae, T.A. and Cox, M.C. 1996. Best linear unbiased prediction as a method of estimating breeding value in sugarcane. p.39-41. In: Wilson, J.R.; Hogarth, D.M.; Campbell, J.A. and Garside, A.L. (Ed.). Sugarcane:Research towards efficient and sustainable production. CSIRO Divison of Tropical Crops and Pastur, Brisbane.

Wu, K.K.; Heinz, D.J.; Meyer, H.K. and Ladd, S.L. 1978. Minimum sample size for estimating progeny mean and variance. Crop Science. 18(1):57-62.

Wu, K.K.; Heinz, D.J.; Meyer, H.K. and Ladd, S.L. 1977. Selection studies in sugarcane (*Saccharum* spp. hybrids). III. A method to determine sample size for the estimation of population variance. Theoretical and Applied Genetics. 51:29-33.