

**PRINCIPAL COMPONENT ANALYSIS IN *Paspalum* GERMPLASM
CHARACTERIZATION**

A. Regitano Neto and L. A. R. Batista

EMBRAPA-CPPSE, São Carlos, SP, C.P. 339, Brasil, 13560-970,

amadeu@cppse.embrapa.br

Abstract

In a survey, through principal component analyses, for the most representative descriptors to be used in *Paspalum* germplasm characterization, 16 agronomic traits, 16 reproductive traits and 21 vegetative traits were evaluated in 32 accessions comprising *Paspalum plicatulum* (19) and *Paspalum compressifolium* (13) from the institutional germplasm collection. ANOVA for agronomic descriptors showed significant differences among accessions, except for percentage of dry matter in winter (DMW) which was not significant. Principal component analyses allowed reduction of 69%, 69% and 62%, respectively, in the initial number of descriptors, and the election of the most important for agronomic traits: DMYS, DMYW, RGS, CPPW and NDFW (dry matter yield in summer, dry matter yield in winter, regrowth score in summer, crude protein percentage in winter and neutral detergent fiber in winter, respectively); reproductive descriptors: NSG, PE, PFA, PS and LE (number of seeds per gram, pubescence of the escape, pubescence of the floral axis, pubescence of the spikelet and length of the escape, respectively); and vegetative descriptors: PLDI, PLL, PSDE, PINDE, LL, LWH, SWB, SWT (pubescence of leaf – distribution, pubescence of leaf – length, pubescence of sheath – density, pubescence of internode –

density, leaf length, leaf width – half, sheath width – base and sheath width – top, respectively).

Supported by Embrapa and FAPESP (99/02421-4)

Keywords: *Paspalum plicatulum*, *Paspalum compressifolium*, multivariate analysis, native pastures, tropical grasses, plant genetic resources

Introduction

Feeding of animals in Brazilian cattle production is based mainly on pastures. Pastures consist in the most economic form of animal production in Brazil and are formed essentially by varieties from another continents, mainly from Africa. The use of extensive covered areas with only one grass variety presents the risk of genetic vulnerability, besides the aspect of the introduction of plagues and diseases with the importation of exotic germplasm.

The genus *Paspalum* comprises native forage grasses from the south of South America, highly adapted to the environment of this region, with good agronomic potential which is still unexplored (Batista & Godoy, 2000). The germplasm of the genus is quite diverse and comprises several groups, within them the Plicatula group with the species *Paspalum plicatulum* and *Paspalum compressifolium*, objects of this study.

Genetic improvement of a genus requires estimation of genetic parameters, to assure the exploration of its variability and direction of the breeding program. Employment of multivariate techniques has been recommended to estimate genetic divergence among accessions and for optimization of germplasm collections. The study of principal components has been indicated in discarding descriptors of less importance for explanation of the variation, saving unnecessary work.

The present work aimed to characterize the accessions and to select the most representative botanical-agronomic descriptors of the presented variation, through principal components analysis of the accessions of *Paspalum plicatulum* and *Paspalum compressifolium* maintained by the Embrapa's Southeast – Cattle Research Center, in São Carlos, SP.

Material and Methods

In a survey for the most representative descriptors to be used in *Paspalum* germplasm characterization through principal component analysis, 16 reproductive traits, 16 agronomic traits and 21 vegetative traits were evaluated in 32 accessions comprising *Paspalum plicatulum* (19) and *Paspalum compressifolium* (13) from the institutional germplasm collection. The studied accessions were collected in central-south Brazil, in the following states: 47% in Rio Grande do Sul, 25% in Paraná, 19% in Santa Catarina, 6% in Mato Grosso do Sul and 3% in São Paulo. Agronomic data were taken from a completely randomized block experiment with two replications and yield was evaluated in successive cuts, while reproductive and vegetative traits were measured in a third replication specifically used for these evaluations. Experimental plots measured 4 m² without borders.

Principal component methodology can be found in details in Morrison (1976) and Mardia et al. (1979). In order to reduce the effects of different scales of measurement in different characters, all variables in matrix data were transformed by the formula: $z_{ij} = (x_{ij}/s_j)$, where s_j is the standard deviation of the variable j .

Relative contribution of a component was evaluated by its proportional contribution to the total variance. Descriptors were selected based on their importance and redundancy, the latter defined by the correlation coefficients estimated among the traits. According to Jolliffe

(1972, 1973) and Mardia et al.(1979), variables with the highest coefficients in their components with eigenvalues smaller than 0.70 should be discarded.

This methodology was applied using the software GENES (Cruz, 1998).

Results and Discussion

Analysis of variance (ANOVA) for agronomic descriptors showed significant differences among accessions, except for DMW (percentage of dry matter in winter) which was not significant. Coefficients of variation ($CV_e\%$) ranged from 1.30 to 24.25 considering the original 16 traits in the experiment.

The relative importance of the principal components was estimated from the percentage of total variance contained on its first corresponding eigenvalues. Table 1 shows the eigenvalues and the absolute and relative contribution of the descriptors in the explanation of total observed variation for the first components and three types of studied characteristics.

For agronomic characters it was evidenced that the first three principal components explained 79,86% of the available variation, while for reproductive and vegetative characters (Table 1) the first five principal components were necessary, so that 72,37 and 75,83% of the variation could be represented, respectively. Dispersion of the variance in several components was also observed by Strapasson (1997) in *Paspalum*, Cury (1993) in cassava and Dias (1994) in cocoa. Dilution of the variation among the components can be attributed to the diversity of the appraised descriptors.

Principal component analysis allowed reduction of 69%, 69% and 62% in the initial number of descriptors for agronomic, reproductive and vegetative traits, respectively. It also allowed the election of the following most important descriptors: 1) agronomic descriptors: DMYS, DMYW, RGS, CPPW and NDFW (dry matter yield in summer, dry matter yield in winter, regrowth score in summer, crude protein percentage in winter and neutral detergent

fiber in winter, respectively); 2) reproductive descriptors: NSG, PE, PFA, PS and LE (number of seeds per gram, pubescence of the escape, pubescence of the floral axis, pubescence of the spikelet and length of the escape, respectively); and vegetative descriptors: PLDI, PLL, PSDE, PINDE, LL, LWH, SWB, SWT (pubescence of leaf – distribution, pubescence of leaf – length, pubescence of sheath – density, pubescence of internode – density, leaf length, leaf width – half, sheath width – base and sheath width – top, respectively).

References

Batista, L.A.R. and Godoy R. (2000). Avaliação preliminar e seleção de germoplasma do gênero *Paspalum* para produção de forragem. Revista da Sociedade Brasileira de Zootecnia, **29**:23-32.

Cruz, C.D. (1998). Programa Genes – Aplicativo computacional em estatística aplicada à genética. Genetics and Molecular Biology, **21**:135-138.

Cury, R. (1993). Dinâmica evolutiva e caracterização de germoplasma de mandioca (*Manihot esculenta*, Crantz) na agricultura autóctone do sul do Estado de São Paulo. Piracicaba, 103p. Dissertação (M.S.) - Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo.

Dias, L.A.S. (1994). Divergência genética e fenética multivariada na predição de híbridos e preservação de germoplasma de cacau (*Theobroma cacao* L.). Piracicaba, 94p. Tese (Doutoramento) - Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo.

Jolliffe, I.T. (1972). Discarding variables in a principal component analysis. I. Artificial data. Applied Statistics, **21**: 160-173.

Jolliffe, I.T. (1973). Discarding variables in a principal component analysis. II. Real data Applied Statistics, **22**: 21-31.

Mardia, K.V.; Kent J.T. and Bibby J.M. (1979) *Multivariate analysis*. London, Academic Press. 521p.

Morrison, D.F. (1976) *Multivariate statistical methods*. New York, McGraw-Hill, 415p.

Strapasson, E. (1997) Seleção de descritores na caracterização de germoplasma de *Paspalum* através de componentes principais. Piracicaba, 95p. Dissertação (M.S.) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo.

Table 1 - Estimates of the eigenvalues associated to the principal components and its relative and accumulated contribution, obtained in the study of 16 agronomic, 16 reproductive and 21 vegetative descriptors evaluated in 32 accessions of *Paspalum*.

Components	Eigenvalues	%	% accumulated
Agronomic descriptors			
1	6,8576	42,85	42,85
2	3,4338	21,46	64,31
3	2,4873	15,54	79,86
Reproductive descriptors			
1	4,3224	27,02	27,02
2	2,4424	15,26	42,28
3	1,8604	11,63	53,91
4	1,5056	9,41	63,32
5	1,4473	9,05	72,37
Vegetative descriptors			
1	5,7736	27,49	27,49
2	4,2541	20,26	47,76
3	2,2822	10,87	58,62
4	2,1304	10,15	68,76
5	1,4827	7,06	75,83