Brazilian Journal of Biology

ISSN 1519-6984 (Print)
ISSN 1678-4375 (Online)

Original Article

Genetic diversity based on multivariate techniques in elephant grass genotypes for bioenergy

Diversidade genética com base em técnicas multivariadas em genótipos de capim-elefante para bioenergia

A. K. F. Vidal[a]*, R. F. Daher[a], R. S. Freitas[a], J. G. S. Santana[a], W. F. Stida[a], M. Ambrósio[a], V. B. Silva[b], A. G. Souza[a], M. R. Nascimento[a], R. M. Santos[a], C. L. Leite[a] and J. E. C. Farias[a]

[a]Universidade Estadual do Norte Fluminense Darcy Ribeiro – UENF, Centro de Ciências e Tecnologias Agropecuárias – CCTA, Campos dos Goytacazes, RJ, Brasil

[b]Universidade Federal do Piauí – UFPI, Centro de Ciências Agrárias – CCA, Programa de Pós-graduação em Agronomia – PPGA, Campus Universitário Ministro Petrônio Portella, Teresina, PI, Brasil

Abstract

The use of biomass for energy production constitutes a promising strategy that warrants the search for new sources of biomass. Elephant grass has been gaining notoriety due to its high dry matter yield and rapid growth. The present study was carried out to quantify the genetic divergence of nine elephant grass half-sib families in order to identify genotypes with greater genetic divergence and productive potential for hybridization, using the hierarchical clustering methodology based on principal components. Half-sib families were generated using genotypes from the Active Germplasm Bank of Elephant Grass. The experiment was laid out in a randomized-block design with nine half-sib families, three replicates, and eight plants per plot. A total of 216 genotypes of elephant grass were evaluated. Principal component (PC), biplot, and hierarchical clustering analyses for diversity estimation were conducted using R software. The first two PCs of biplot analysis accounted for 64% of the cumulative variation. Dry matter yield was the most important trait for genotype discrimination (0.89), followed by plant height (0.67) and stem diameter (0.61) in PC1. In this analysis, the distances between accessions were considered and there were no family links, which indicates the existence of wide variability within the evaluated families, since genotypes belonging to the same family were not grouped together, but rather distributed into different groups. Crosses between genotypes of group three and genotypes of groups one and two are recommended for the development of high-yielding genotypes when aiming at energy production.

Keywords: biomass, *Cenchrus purpureus* (Schumach.) *Morrone*, principal component, hierarchical clustering.

Resumo

A utilização de biomassa para a produção de energia representa uma estratégia promissora, justificando a busca por novas fontes de biomassa. O capim-elefante tem ganhado destaque devido ao seu elevado rendimento de matéria seca e rápido crescimento. Este estudo teve como objetivo quantificar a divergência genética entre novas famílias de meio-irmãos de capim-elefante, com a finalidade de identificar genótipos com maior divergência genética e potencial produtivo para hibridização. Utilizou-se uma metodologia de agrupamento hierárquico baseada em componentes principais. As famílias de meio-irmãos foram obtidas a partir de genótipos do Banco Ativo de Germoplasma de Capim-Elefante. O experimento foi conduzido em delineamento de blocos casualizados, com nove famílias de meio-irmãos, três repetições e oito plantas por parcela, totalizando a avaliação de 216 genótipos de capim-elefante. Foram realizadas análises de componentes principais (PC), biplot e agrupamento hierárquico para estimar a diversidade, utilizando o software R. Os dois primeiros PCs da análise biplot explicaram 64% da variação total. A produtividade de matéria seca foi a característica mais relevante para a discriminação genotípica (0,89), seguida pela altura da planta (0,67) e pelo diâmetro do caule (0,61) no PC1. Nesta análise, foram consideradas as distâncias entre os acessos, sem considerar parentesco, o que indica uma variabilidade significativa dentro das famílias avaliadas, uma vez que os genótipos pertencentes à mesma família não foram agrupados, mas sim distribuídos em grupos distintos.

Palavras-chave: biomassa, *Cenchrus purpureus* (Schumach.) *Morrone*, componente principal, cluster hierárquico.

---

*e-mail: anakesia.vidal@hotmail.com

Received: September 19, 2023 – Accepted: January 2, 2024

1. Introduction

With the growing population and technological innovations, energy consumption is expected to increase at a rapid pace in the coming years (Coyle and Simmons, 2014). This increase in consumption will cause the depletion of non-renewable fossil fuel resources and global warming. Fossil fuels account for 81% of the world's total primary energy supply. Renewables contributed only 13.8% of the primary energy supply in 2018 (WBA, 2019).

As a renewable energy source, biomass plays a vital role in meeting global demand in various sectors, providing 55.6% of renewable energy and accounting for a significant share on all continents (WBA, 2019). The use of biomass for energy production constitutes a promising strategy, thus warranting the search for new sources of biomass.

Among the crops used for biomass energy production, elephant grass [*Cenchrus purpureus* (Schumach.) *Morrone*] has been gaining notoriety due to its high dry matter yield (59.26 t ha^{-1} year^{-1}) (Silva et al., 2020) and rapid growth, which allows for harvesting at 10 months (Vidal et al., 2022). However, despite its relevance, there is still a need to advance research focusing on the development of new cultivars.

In breeding programs, new cultivars are produced from crosses. Because the obtained genotypes have genetic variability, it is possible to identify superior genotypes with traits relevant for energy production (Rodrigues et al., 2017). In this respect, research should be undertaken to elucidate the genetic diversity of the crop, thus aiding the crossing process. This will contribute to the development of new cultivars of elephant grass that can be obtained through generation advancement.

In the selection of parents, separate studies are usually carried out to assess genetic variability and agronomic performance (Rocha et al., 2015). For this study, we employed a methodology in which hierarchical clustering is done on principal components, combining factor analysis and a subsequent clustering process, based on Ward's hierarchical classification. This technique allows the association of agronomic and diversity performance in a single analysis (Maciel et al., 2018), providing the simultaneous evaluation of all genotypes under study with a reduction in the original size of the dataset and minimal loss of information.

In view of the importance of the crop and the advantages of using this technique, the present study was carried out to quantify the genetic divergence of nine elephant grass half-sib families in order to identify genotypes with greater genetic divergence and productive potential for hybridization, using the hierarchical clustering methodology based on principal components.

2. Material and Methods

2.1. Characterization of the experimental area

The experiment was implemented in the experimental area of the State Center for Research in Agroenergy and Waste Utilization (CEPAAR-PESAGRO-RIO), located in the municipality of Campos dos Goytacazes - RJ, Brazil (21°44'47" S, 41°18'24" W, 11 m above sea level).

Meteorological data were obtained from the automatic agrometeorological station, located near the experimental area. Figure 1 shows the monthly precipitation and temperature values during the experimental period (November 2019 to August 2021) and the climatological normal of the municipality (1981-2010). During the 22 months of the experiment, an accumulated precipitation of 2,197 mm was recorded. Temperature values ranged from 16 to 33 °C.

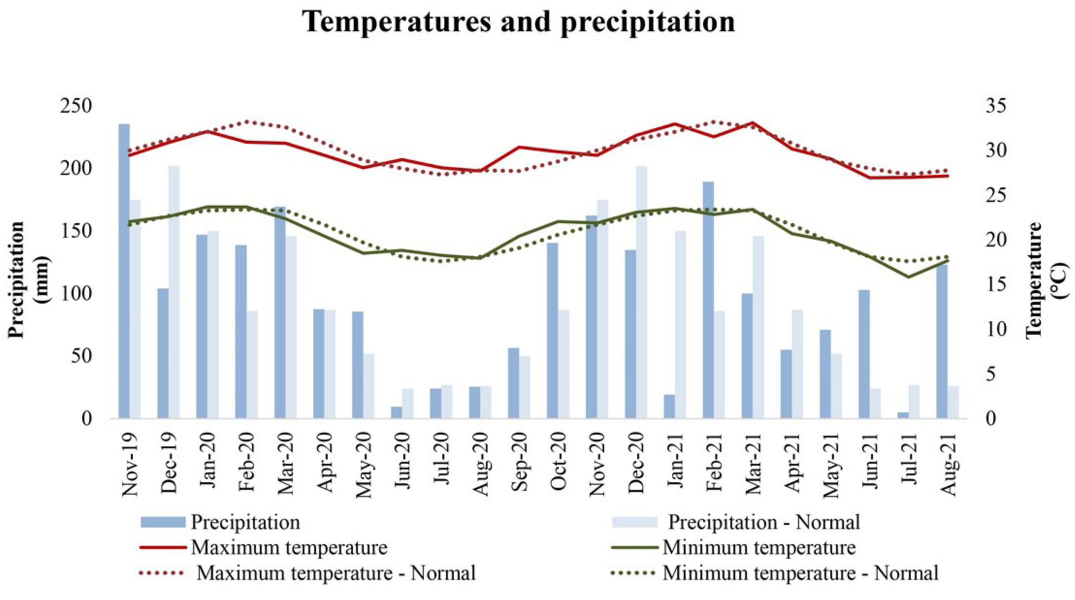

Figure 1. Temperatures and precipitation during the experiment with elephant grass and Climatological Normal (1981 to 2010). Campos dos Goytacazes, 2019-2021.

Elephant grass grows well in environments with temperatures between 18 and 35 °C, with the optimal temperature being 24 °C. According to the Köppen classification, the climate in the municipality of Campos dos Goytacazes, is the Aw (tropical) type. Rainfall occurs much more frequently in summer than in winter. According to the climatological normal (1981-2010), the average annual maximum temperature is 29.90 °C and the minimum is 20.80 °C. Average annual precipitation is 1,055.3 mm. Therefore, the municipality of Campos dos Goytacazes has climatic characteristics that favor the optimal development of the crop.

The soil in the experimental area where the genotypes were cultivated is characterized as a terrace soil and classified as dystric yellow Oxisol (*argissolo*). The soil properties are as follows: pH - 5.3; phosphorus - 36 mg dm^{-3}; potassium - 166 mg dm^{-3}; Ca - 1.5 cmolc dm^{-3}; Mg - 0.6 cmolc dm^{-3}; Al - 0.1 cmolc dm^{-3}; H + Al - 2.4 cmolc dm^{-3}; Na - 0.05 cmolc dm^{-3}; C - 1.18%; and organic matter - 2.03%.

2.2. Progeny formation

The families were formed using genotypes from the Active Germplasm Bank of Elephant Grass (BAGCE) at LEAG/CCTA/UENF, located in Campos dos Goytacazes - RJ, Brazil. For this purpose, the nine most productive genotypes in terms of dry matter yield were selected, following Rocha et al. (2015) (Table 1).

The crosses were carried out from June to August 2019, during the flowering stage of the crop. The crosses were allowed to occur naturally, and the panicles of the respective genotypes were harvested later. Panicles were harvested at the beginning, middle, and end of the flowering season. This procedure was adopted to enable the collection of seeds pollinated by early and late parents, thus ensuring greater variability for the generated families. After collection, the seeds of each genotype were mixed, packed in aluminum foil, and stored in a refrigerator. On September 18, 2019, sowing was carried out in 128-cell Styrofoam trays, which were filled with forest substrate and kept for 60 days in a greenhouse equipped with an irrigation system to provide ideal conditions for germination and seedling maintenance.

2.3. Establishment and conduct of the experiment

Soil preparation consisted of two disc harrowing operations. The seedlings were transplanted to the field on November 18, 2019. Supplementary irrigation was carried out by a conventional sprinkler system only in the phase of implementation and establishment of the plants (November and December), to ensure their establishment and development in this early stage.

During the experimental period, fertilizer application was carried out according to the results of soil analysis and recommendations given in the liming and fertilization manual harvest (Freire et al., 2017). Fertilizer treatment was split into three applications: at planting, after plot-uniformity and after the first cut.

The families were evaluated in a randomized-block design with three replicates. The plot consisted of a 15-m row with a spacing of 1.50 × 1.50 m, totaling 10 plants per plot. The usable area was represented by the eight central plants, and the plants at the ends of the line were considered borders. Table 1 displays the identification of the 216 genotypes that composed the usable area according to the families to which they belong.

The plot-uniformity cut was made on April 14, 2020, marking the beginning of the evaluation period. After eight months of plant growth, the grass was harvested for evaluations on two occasions: December 1, 2020, and July 30, 2021.

2.4. Evaluated traits

Eight individual plants of each plot were evaluated for the following traits:

- Dry matter yield (DMY, $t.ha^{-1}$) - the harvested plant was weighed in the field, and afterward, a random sample was taken from each plant. This sample was chopped, placed in a labeled paper bag, weighed, and then taken to an oven at 65 °C for 72 hours. Subsequently, the samples were weighed again to obtain the air-dried sample. The dried material was later ground in a Wiley mill with a 5 mm sieve and placed in plastic bags for determining the oven-dried sample. This sample was determined by drying 2 g of each ground material at 105 °C for 18 hours and then weighing it again

Table 1. Identification of the nine parents from the Active Germplasm Bank of Elephant Grass (BAGCE) at LEAG/CCTA/UENF, used as female parents to generate the half-sib families, and the 216 genotypes obtained from the crosses. Campos dos Goytacazes, RJ.

Identification		Female parent	BAGCE	Origin
Family	Genotype			
F1	1-24	Gigante de Pinda	7	Brazil
F2	25-48	Pasto Panamá	74	Panama
F3	49-72	903-77 ou Australiano	68	Brazil
F4	73-96	Gramafante	45	Colombia
F5	97-120	Três Rios	3	Brazil
F6	121-144	Elefante de Pinda	26	Colombia
F7	145-168	Duro de Volta Grande	34	Brazil
F8	169-192	Taiwan A-46	33	Brazil
F9	193-216	King Grass	51	Brazil

- Dry matter percentage (DM, %) - calculated by dividing the plant's dry weight by its fresh weight and multiplying the result by 100;
- Number of tillers (NT) - determined by counting the number of tillers of each of the evaluated plants;
- Plant height (PH, m) – measured from the ground to the inflection of the last fully expanded leaf of each of the evaluated plants;
- Stem diameter (SD, mm) – defined as the average of three tillers of each evaluated plant, measured with a digital caliper at 1 m above the ground.

2.5. Statistical analysis

Descriptive statistics involved the means (each value being the average of the two evaluation harvests) of five morpho-agronomic traits related to energy production, which were evaluated in 216 genotypes belonging to nine elephant grass half-sib families. The minimum value, maximum value, standard deviation, and coefficient of variation (CV%) were obtained using the *pastecs* package in R software.

Firstly, principal component analysis (PCA) was performed using the data collected from the 216 genotypes. The data were subjected to Bartlett's test of sphericity (95% significance level), which tested the hypothesis that the correlation matrix is an identity matrix, and the Kaiser-Meyer-Olkin test, which tested the suitability of our dataset for factor analysis. The eigenvalues of all principal components were presented. The eigenvectors were used to differentiate the genotype effects in terms of strength and direction for the traits evaluated in the principal components. The first two PCs (PC1, PC2) and PC3 showed the greatest variability in the tested parameters.

The biplot was generated based on the genotype values, using PC1 on the horizontal axis and PC2 on the vertical axis. The effective contribution of genotypes was used in the first three components. Pearson's correlation between distance and cophenetic matrices was used to determine cluster cohesion.

The results of PCA were subjected to hierarchical clustering (HCPC) using Ward's method to group similar genotypes into clusters. The HCPC approach is one of the most widely used descriptive data methods to group genotypes with similar patterns of responses from quantitative data. It is used to group genotypes with similar patterns of responses from quantitative data. The objective is to classify the genotypes into groups as homogeneous as possible. The ideal number of groups was determined by a dendrogram and inertia criterion. In the dendrogram, significant changes between two cut-off levels suggest an optimal number of groups.

To visualize the groups, a plot was constructed in which the genotypes were projected as a function of the distances between the groups, using the first two principal components.

The analyses applied to the principal component, the biplot, and the hierarchical clustering for diversity estimation were conducted using R software. The *psych*, *FactoMineR*, and *factoextra* packages were used for PCA, whereas the *hclust*, *ggplot2*, *ggpbubr*, and *factoextra* packages were used for cluster analysis.

3. Results and Discussion

Descriptive analysis showed that the highest coefficients of variation were found for the traits of dry matter yield (DMY) and number of tillers (NT), followed by stem diameter (SD). The other CV% obtained for the variables were considered low. These results indicate reliability in the choice of the experimental designs used as well as in the strategies for conducting and evaluating the experiments for the improvement of elephant grass (Table 2).

The high coefficients of variation (CV) obtained for DMY (37.53%), NT (30.51%), and SD (21.40%) are acceptable, considering that the traits evaluated in this study are quantitative and, consequently, highly influenced by the environment. Gravina et al. (2020) and Souza et al. (2017) also worked with elephant grass and found similar results regarding CV in experiments evaluating the same quantitative traits.

Principal component analysis (PCA) reduces the number of primary variables to describe the total variation of a population and explain the contribution of traits to the total variation (El-Hashash, 2016). In this way, it improves the interpretation of data, allowing the breeder to obtain the maximum information possible in a smaller number of PCs.

In this study, there were three significant components (Table 3). The significance of the components was considered based on the latent root criterion (Barbosa et al., 2021), according to which eigenvalues greater than one are considered significant in PCA. Factor loadings show the correlation between the variable and the component, indicating the quality of the representation. Factor loadings are considered most relevant when above 0.58.

Table 2. Descriptive statistics for morpho-agronomic traits of 216 genotypes belonging to nine elephant grass half-sib families. Campos dos Goytacazes, 2019 - 2021.

Trait	Unit	Minimum	Maximum	Mean	Standard deviation	CV (%)
DMY	t ha^{-1}	4.40	36.35	14.92	5.60	37.53
% DM	%	21.90	45.30	32.82	3.62	11.04
PH	m	2.10	3.85	3.11	0.28	9.01
SD	mm	3.15	9.15	5.28	1.13	21.40
NT	Tillers	12.00	70.00	34.82	10.62	30.51

CV: coefficient of variation; DMY: dry matter yield; %DM: dry matter percentage; PH: plant height; SD: stem diameter; NT: number of tillers.

The first two PCs of biplot analysis accounted for 64% of the cumulative variation (Table 3). According to Yang et al. (2009), at least 60% of the data variation must be explained by the first two PCs. Similar values for total variation detected by PCs were reported by Gravina et al. (2020), who evaluated morpho-agronomic traits in elephant grass genotypes and described that 70.07% of the total variation was explained in the first two PCs.

Researchers working with quantitative traits in other species found similar results. In string bean, agronomic traits in the first two PCs explained 63.50% of the total variation (Oliveira et al., 2018). Yildirim et al. (2018), evaluated grain quality in wheat genotypes and had 71.15% of the variation explained in the first two PCs. In pepper, Baba et al. (2019), examined the concentrations of secondary metabolites and reported an explained variation of 63.14 and 59.61% in the first two PCs in unripe and ripe fruits, respectively.

Table 3 presents in detail the descriptors that most contributed to the discrimination of accessions, which showed the highest eigenvalues for the respective components. Dry matter yield was the most important trait for genotype discrimination (0.89), followed by PH (0.67) and SD (0.61) in PC1, which is the most important component, as it explained 37% of the total variation. According Al-Naggar et al. (2020), the importance of a PC is defined according to the total variance it explains.

For PC2, the traits that stood out were NT and SD. In PC3, only %DM contributed to the variance of the component. Principal component analysis allowed reducing the size of the interrelated variables, proving to be efficient in the multivariate correlation of genotypes and studies of selection from graphic representation (Maia et al., 2016).

Figure 2 shows the 25 genotypes with the greatest contribution, according to their scores in the principal components. Genotypes of the nine families evaluated were selected, demonstrating variability within the families, given that all of them have promising genotypes. Family 2 contributed 20% of the selected genotypes, followed by family 8, with 16%. Families 3, 5, and 9 contributed 12% each, whereas families 2, 4, and 7 contributed 8%. Only one genotype was selected from family 6, corresponding to 4% of the total selected. In addition to the available variability, superior genotypes must possess traits of interest for elephant grass genetic breeding programs and favorable to bioenergy production.

Figure 3 is a Biplot with the projection of the evaluated variables. The plot shows a multivariate relationship between the variables and demonstrates the possibility of selecting genotypes according to the variable with which they correlate. The values of the genotypes and variables are plotted on different scales. Variables with angles less than 90º are positively correlated, whereas those with angles greater than 90º are negatively correlated. Variables with an angle of 90º do not have any correlation (Oliveira et al., 2018).

Studies of correlations between traits help the breeder in their decision making, considering that the selection process is hampered by the existence of traits of low heritability or difficult to measure. Therefore, it is extremely important to know how and to which extent the variation of a trait can be reflected in the others (Santana et al., 2021).

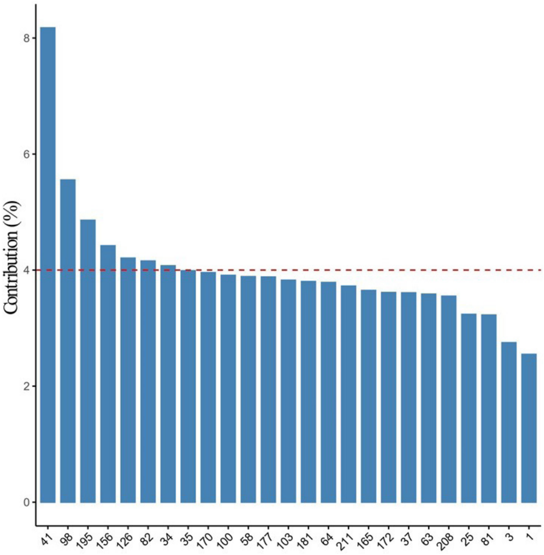

Figure 2. Contribution (%) of the top 25 elephant grass genotypes to variability based on three principal components.

Table 3. Loadings, eigenvalues, variance, and cumulative variance in the principal components of 216 genotypes belonging to nine elephant grass half-sib families evaluated at two harvests. Campos dos Goytacazes, 2019-2021.

Trait	Component				
	PC1	PC2	PC3	PC4	PC5
Dry matter yield	0.89[a]	0.31	0.01	0.05	-0.33
Dry matter percentage	0.08	0.41	0.90[a]	0.09	0.1
Plant height	0.67[a]	-0.45	0.16	-0.55	0.11
Stem diameter	0.61[a]	-0.58[a]	0	0.52	0.14
Number of tillers	0.48	0.74[a]	-0.4	-0.01	0.26
Eigenvalues	1.85[b]	1.35[b]	1.00[b]	0.59	0.21
% Variance	37.03	27.00	19.89	11.80	4.28
% Cumulative variance	37.03	64.03	83.92	95.72	100

[a]Loadings > 0.58 are significant. [b]Eigenvalues > 1 are significant.

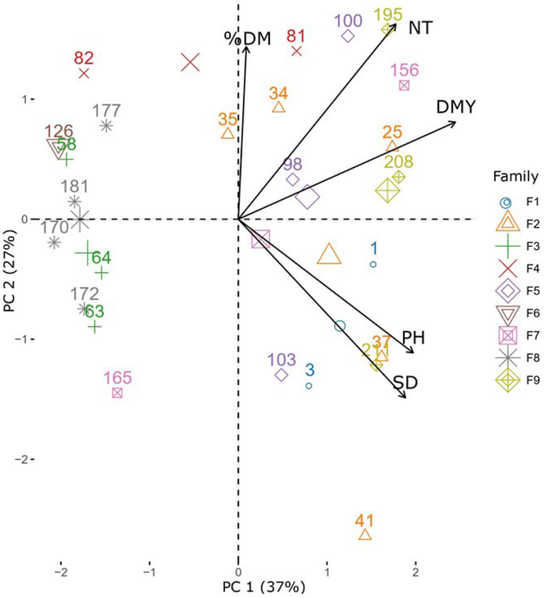

Figure 3. Biplot with the projection of dry matter yield (DMY), dry matter percentage (%DM), plant height (PH), stem diameter (SD), and number of tillers (NT) of 216 elephant grass genotypes, with the identification of only the top 25 genotypes.

Dry matter yield correlated with all other evaluated traits, with the highest correlation being with NT. When aiming at energy production, DMY is the most important trait for the crop. According to Rossi et al. (2011), plants with a high number of tillers have smaller height and diameter, which leads to low fiber and high nitrogen levels and, consequently, low calorific value. On the other hand, when a plant has few tillers, its stems are more robust and the plant is taller, with high fiber and low nitrogen contents and, therefore, high calorific value.

Dry matter percentage correlated negatively with SD and had no correlation with PH or NT. Dry matter percentage is of great importance for the crop for energy purposes, considering that this trait directly influences the calorific value of biomass. During burning, in case of low %DM, heat loss occurs in the flue gases in the form of water vapor. Thus, the higher the moisture content, the lower the amount of energy provided by the fuel. In this way, selecting genotypes with high scores in PC1 (DMY, PH, and SD) and PC3 (%DM), will increase the chances of finding promising genotypes for energy production (Rossi et al., 2011; Daher et al., 2018).

For hierarchical clustering, we employed a methodology that combines component analysis and a subsequent clustering process based on Ward's hierarchical classification. The main objective of hierarchical clustering on PCA is to maximize the total variance of the projected points, i.e., to define the subspace that best represents the diversity of genotypes, by grouping genotypes according to their behavior within the population (Argüelles et al., 2014). The aggregation criterion used was based on the Ward index, a method based on a classic sum-of-squares criterion in which groups are formed so as to minimize within-group dispersion (Doudney and Rinaldi, 1984).

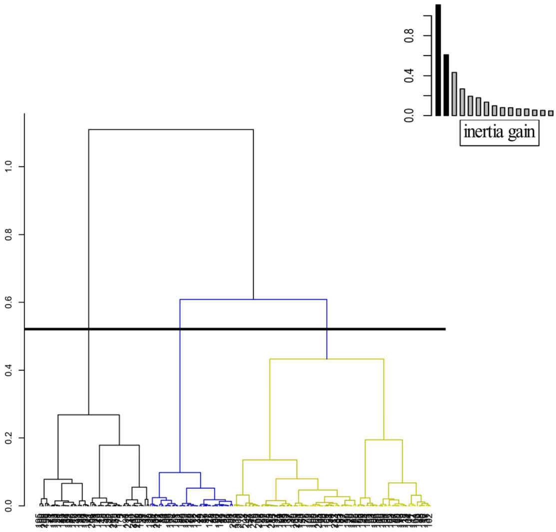

Figure 4. Hierarchical clustering and initial partition.

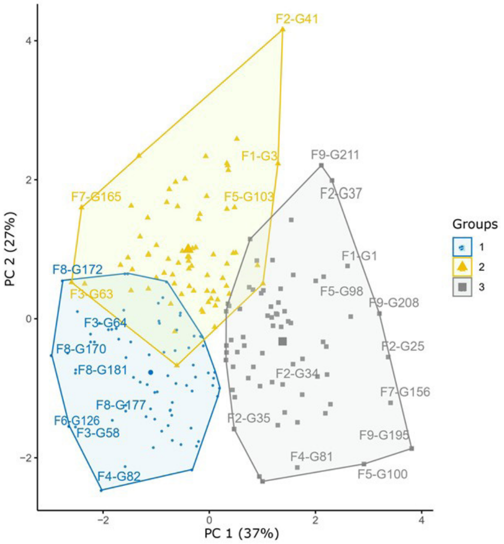

Figure 5. Hierarchical clustering based on the first two principal components and ellipses around the groups formed by 216 genotypes with identification of only the top 25 genotypes belonging to the nine elephant grass half-sib families. Campos dos Goytacazes, 2019-2021.

Figure 4 illustrates the hierarchical classification, where the terminal elements of the tree are the genotypes. The number of groups can then be chosen, taking into account that the greater the inertia, the better the clustering (Argüelles et al., 2014). A cut-off level calculated based on the inertia gains is suggested. This suggested cut-off point was used and the tree was cut automatically, resulting in the formation of three groups.

In Figure 5, the scatterplot produced by the first two principal components illustrates the partitioning into three groups. The three groups formed have ellipses around them as a method of internal validation of the clustering results.

Table 4. Description of the groups according to the five traits evaluated. Campos dos Goytacazes, 2019-2021.

Trait	Unit	Group mean		
		Group 1 (75)	Group 2 (72)	Group 3 (69)
Dry matter yield	t ha^{-1}	11.55	12.71	20.88
Dry matter percentage	%	34.09	30.37	33.98
Plant height	m	2.92	3.17	3.22
Stem diameter	mm	4.34	5.98	5.55
Number of tillers	Tillers	31.72	29.59	43.64

These ellipses create the areas of the groups formed, providing greater clarity in the separation of the groups. The 25 best genotypes have their identifications and were marked with different symbols according to the group to which they belong, whereas the other genotypes were only signaled by the symbols according to the groups to which they belong.

The first group was formed by 75 genotypes, most of which belonged to families 3, 4, 6, and 8, which contained 14, 12, 13, and 11 genotypes, respectively. This group was formed by the genotypes with the highest %DM values. Group two comprised the genotypes with the highest SD and was composed of 72 genotypes, with greater participation of family 5, with 13 genotypes, followed by families 7 and 9, with 10 genotypes each (Table 4).

Group 3 contained 69 genotypes, with the greatest contribution being from families 2 and 9, with 11 genotypes each. This group was formed by the genotypes that showed the highest values for DMY, PH, and NT (Table 4). In this analysis, the distances between accessions were considered and there were no family links, which indicates the existence of wide variability within the evaluated families, since genotypes belonging to the same family were not grouped, but were distributed into different groups.

The analyses conducted in this study allow us to confirm the variability within the nine families evaluated. These results were already expected, considering that the 216 elephant grass genotypes evaluated come from nine half-sib families and the pollen grains may have come from different male parents. With the present results, breeders will be able to manage future crosses between contrasting genotypes.

Crosses between promising genotypes of group 3, which showed high DMY and PH, and genotypes of groups 1 and 2 are recommended. These crosses will generate genotypes that in addition to having high dry matter yield potential, will also have a lower number of tillers, a larger diameter and, consequently, high fiber and low nitrogen contents and high calorific value (Rossi et al., 2011), characteristics of great importance in elephant grass for energy production.

4. Conclusions

Dry matter yield, plant height, and stem diameter are traits of greater importance for the discrimination of elephant grass genotypes.

There is great variability within the evaluated families.

Crosses between genotypes of group three and genotypes of groups one and two are recommended for the development of high-yielding genotypes when aiming at energy production.

References

AL-NAGGAR, A.M.M., SHAFIK, M.M. and MUSA, R.Y.M., 2020. Genetic diversity based on morphological traits of 19 maize genotypes using principal component analysis and GT biplot. *Annual Research & Review in Biology*, vol. 35, no. 2, pp. 68-85. http://dx.doi.org/10.9734/arrb/2020/v35i230191.

ARGÜELLES, M., BENAVIDES, C. and FERNÁNDEZ, I., 2014. A new approach to the identification of regional clusters: hierarchical clustering on principal components. *Applied Economics*, vol. 46, no. 21, pp. 2511-2519. http://dx.doi.org/10.1080/00036846.2014.904491.

BABA, V.Y., CONSTANTINO, L.V., IVAMOTO, S.T., MOREIRA, A.F.P., MADEIRA, T.B., NIXDORF, S.L., RODRIGUES, R. and GONÇALVES, L.S.A., 2019. Capsicum-Colletotrichum interaction: identification of resistance sources and quantification of secondary metabolites in unripe and ripe fruits in response to anthracnose infection. *Scientia Horticulturae*, vol. 246, no. 1, pp. 469-477. http://dx.doi.org/10.1016/j.scienta.2018.11.011.

BARBOSA, R.S., PEREIRA, G.F.M., RIBEIRO, S.S., HAGE, A.L.F., COSTA, G.F., SALOMÃO, R.P. and SCHWARTZ, G., 2021. Key species selection for forest restoration after bauxite mining in the Eastern Amazon. *Ecological Engineering*, vol. 162, pp. 106190. http://dx.doi.org/10.1016/j.ecoleng.2021.106190.

COYLE, E.D. and SIMMONS, R.A., 2014. *Understanding the global energy crisis*. West Lafayette: Purdue University Press. http://dx.doi.org/10.26530/OAPEN_469619.

DAHER, R.F., PEREIRA, A.V., MENEZES, B.S., CASSARO, S., NOVO, A.A.C., FURLANI, E., AMARAL JUNIOR, A., PEREIRA, M.G., STIDA, W.F. and VIDAL, A.K.F., 2018. Canonical correlations among morpho-agronomic and chemical traits in hybrids between elephant grass and millet. *Australian Journal of Crop Science*, vol. 12, no. 2, pp. 210-216. http://dx.doi.org/10.21475/ajcs.18.12.02.pne406.

DOUDNEY, C.O. and RINALDI, C.N., 1984. Modification of UV-induced mutation frequency and cell survival of *Escherichia coli* B/r WP2 trp E65 by treatment before irradiation. *Journal of Bacteriology*, vol. 160, no. 1, pp. 233-238. http://dx.doi.org/10.1128/jb.160.1.233-238.1984. PMid:6384193.

EL-HASHASH, E., 2016. Genetic diversity of soybean yield based on cluster and principal component analyses. *Journal of Advances in Biology & Biotechnology*, vol. 10, no. 3, pp. 1-9. http://dx.doi.org/10.9734/JABB/2016/29127.

FREIRE, L.R., BALIEIRO, F.C., ZONTA, E., DOS ANJOS, L.H.C., PEREIRA, M.G., LIMA, E., GUERRA, J.G.M., FERREIRA, M.B.C., LEAL, M.A.A., DE CAMPOS, D.A.V.B. and POLIDORO, J.C., 2017. *Manual de calagem e adubação*. Seropédica: Universidade Rural, 430 p.

GRAVINA, L.M., OLIVEIRA, T.R.A., DAHER, R.F., GRAVINA, G.A., VIDAL, A.K.F., STIDA, W.F., CRUZ, D.P., SANT'ANNA, C.Q.S.S., ROCHA, R.S., PEREIRA, A.V. and OLIVEIRA, G.H.F., 2020. Multivariate analysis in the selection of elephant grass genotypes for biomass production. *Renewable Energy*, vol. 160, no. 1, pp. 1265-1268. http://dx.doi.org/10.1016/j.renene.2020.06.094.

MACIEL, G.M., PEIXOTO, J.V.M., ROCHA, J.P.R., ALMEIDA, R.S., OLIVEIRA, C.S. and MENDONÇA, T.F.V., 2018. Multivariate techniques in the determination of genetic diversity in pest-resistant mini tomato genotypes. *Horticultura Brasileira*, vol. 36, no. 4, pp. 504-509. http://dx.doi.org/10.1590/s0102-053620180413.

MAIA, M.C.C., ARAÚJO, L.B., DIAS, C.T.S., OLIVEIRA, L.C., VASCONCELOS, L.F.L., CARVALHO JÚNIOR, J.E.V., SIMEÃO, M. and BASTOS, Y.G.M., 2016. Selection of mango rosa genotypes in a breeding population using the multivariate-biplot method. *Ciência Rural*, vol. 46, no. 10, pp. 1689-1694. http://dx.doi.org/10.1590/0103-8478cr20130722.

OLIVEIRA, T.R.A., GRAVINA, G.A., OLIVEIRA, G.H.F., ARAÚJO, K.C., ARAÚJO, L.C., DAHER, F.R., VIVAS, M., GRAVINA, L.G. and CRUZ, D.P., 2018. The GT biplot analysis of green bean traits. *Ciência Rural*, vol. 48, no. 6, pp. 1-6. http://dx.doi.org/10.1590/0103-8478cr20170757.

ROCHA, A.S., DAHER, R.F., GRAVINA, G.A., PEREIRA, A.V., RODRIGUES, E.V., VIANA, A.P., SILVA, V.Q.R., AMARAL JUNIOR, A.T., NOVO, A.C., OLIVEIRA, M.L.F. and OLIVEIRA, E., 2015. Comparison of stability methods in elephant-grass genotypes for energy purposes. *African Journal of Agricultural Research*, vol. 10, no. 47, pp. 4283-4294. http://dx.doi.org/10.5897/AJAR2015.10218.

RODRIGUES, E.V., DAHER, R.F., SANTOS, A., VIVAS, M., MACHADO, J.C., GRAVINA, G.A., SOUZA, Y.P., VIDAL, A.K.F., ROCHA, A.D.S. and FREITAS, R.S., 2017. Selecting elephant grass families and progenies to produce bioenergy through mixed models (REML/BLUP). *Genetics and Molecular Research*, vol. 16, no. 2, pp. 1-9. http://dx.doi.org/10.4238/gmr16029301. PMid:28525651.

ROSSI, D.A., MENEZES, B.R.S., DAHER, R.F., GRAVINA, G.A., LIMA, R.S.N., LEDO, F.J.S., GOTTARDO, R.D., CAMPOSTRINI, E. and SOUZA, C., 2011. Canonical correlations in elephant grass for energy purposes. *African Journal of Agricultural Research*, vol. 6, no. 30, pp. 6348-6353. http://dx.doi.org/10.5897/AJB2014.13915.

SANTANA, J.G.S., RAMOS, H.C.C., MIRANDA, D.P., SANTA CATARINA, R., VETTORAZZI, J.C.F., POLTRONIERI, T.P.S., DUARTE, R.P., PIROVANI, A.A.V., BOHRY, D., AMARAL JÚNIOR, A.T. and PEREIRA, M.G., 2021. Genotype analysis by trait is a practical and efficient approach on discrimination of inbred lines and identification of papaya (*Carica papaya* L.) ideotypes for fruit quality. *Euphytica*, vol. 217, no. 6, pp. 1-9. http://dx.doi.org/10.1007/s10681-021-02850-8.

SILVA, V.B., DAHER, R.F., SOUZA, Y.P., MENEZES, B.R.S., SANTOS, E.A., FREITAS, R.S., OLIVEIRA, E.S., STIDA, W.F. and CASSARO, S., 2020. Assessment of energy production in full-sibling families of elephant grass by mixed models. *Renewable Energy*, vol. 146, no. 1, pp. 744-749. http://dx.doi.org/10.1016/j.renene.2019.06.152.

SOUZA, Y.P., DAHER, R.F., PEREIRA, A.V. and SILVA, R.S., 2017. Repeatability and minimum number of evaluations for morpho-agronomic characters of elephant-grass for energy purposes. *Brazilian Journal of Agricultural Sciences*, vol. 12, no. 3, pp. 391-397. http://dx.doi.org/10.5039/agraria.v12i3a5456.

VIDAL, A.K.F., DAHER, R.F., FREITAS, R.S., STIDA, W.F., LÉDO, F.J.S., SILVA, V.B., ENTRANTE, G.C., TARDIN, F.D., GRAVINA, G.A., VIVAS, M., SOUZA, A.G., AMBRÓSIO, M., SANTANA, J.G.S. and FARIAS, J.E.C., 2022. Growth curve in elephant grass genotypes based on morpho-agronomic traits for energy production. *Chilean Journal of Agricultural Research*, vol. 82, no. 1, pp. 78-87. http://dx.doi.org/10.4067/S0718-58392022000100078.

WORLD BIOENERGY ASSOCIATION – WBA, 2019 [viewed 11 February 2023]. *Global bioenergy statistics 2019* [online]. Available from: https://www.worldbioenergy.org/uploads/191129%20WBA%20GBS%202019_HQ.pdf

YANG, R.C., CROSSA, J., CORNELIUS, P.L. and BURGUEÑO, J., 2009. Biplot analysis of genotype × environment interaction: proceed with caution. *Crop Science*, vol. 49, no. 5, pp. 1564-1576. http://dx.doi.org/10.2135/cropsci2008.11.0665.

YILDIRIM, M., BARUTCULAR, C., KOC, M., DIZLEK, H., HOSSAIN, A., ISLAM, M.S., TOPTAS, I., BASDEMIR, F., ALBAYRAK, Ö., AKINCI, C. and EL SABAGH, A., 2018. Assessment of the grain quality of wheat genotypes grown assessment of the grain quality of wheat genotypes grown under multiple environments. *Fresenius Environmental Bulletin*, vol. 27, no. 7, pp. 4830-4837.