

**Caracterização de genótipos de mandioca por técnicas multivariadas**  
**Characterization of cassava genotypes by multivariate techniques**  
**Caracterización de los genotipos de la mandioca mediante técnicas multivariadas**

Recebido: 02/05/2020 | Revisado: 02/05/2020 | Aceito: 04/05/2020 | Publicado: 12/05/2020

**Diego Alves Gomes**

ORCID: <https://orcid.org/0000-0003-2108-4672>

Universidade Federal Rural de Pernambuco, Brasil

E-mail: [diegoalvesestatistica@gmail.com](mailto:diegoalvesestatistica@gmail.com)

**Lucas Silva do Amaral**

ORCID: <https://orcid.org/0000-0003-1096-3241>

Universidade Federal Rural de Pernambuco, Brasil

E-mail: [lucasihua@gmail.com](mailto:lucasihua@gmail.com)

**Denise Stéphanie de Almeida Ferreira**

ORCID: <https://orcid.org/0000-0001-5150-9331>

Universidade Federal Rural de Pernambuco, Brasil

E-mail: [deniialmeida@gmail.com](mailto:deniialmeida@gmail.com)

**Guilherme Rocha Moreira**

ORCID: <https://orcid.org/0000-0001-6344-1151>

Universidade Federal Rural de Pernambuco, Brasil

E-mail: [guirocham@gmail.com](mailto:guirocham@gmail.com)

**Frank Gomes - Silva**

ORCID: <https://orcid.org/0000-0002-3481-3099>

Universidade Federal Rural de Pernambuco, Brasil

E-mail: [franksinatrags@gmail.com](mailto:franksinatrags@gmail.com)

**Maria Lindomárcia Leonardo da Costa**

ORCID: <https://orcid.org/0000-0003-1345-0169>

Universidade Federal da Paraíba, Brasil

E-mail: [lindomarciacosta@gmail.com](mailto:lindomarciacosta@gmail.com)

**Moacyr Cunha Filho**

ORCID: <https://orcid.org/0000-0002-3466-8143>

Universidade Federal Rural de Pernambuco, Brasil

E-mail: [moacyr2006@gmail.com](mailto:moacyr2006@gmail.com)

**André Luiz Pinto dos Santos**

ORCID: <https://orcid.org/0000-0002-7703-2102>

Universidade Federal Rural de Pernambuco, Brasil

E-mail: [andrefensor@hotmail.com](mailto:andrefensor@hotmail.com)

**Mércia Regina Pereira de Figueiredo**

ORCID: <https://orcid.org/0000-0002-6397-1547>

Instituto Capixaba de Pesquisa, Assistência Técnica e Extensão Rural, CPDI Norte, Brasil

E-mail: [merreg5@gmail.com](mailto:merreg5@gmail.com)

**Patrícia Guimarães Pimentel**

ORCID: <https://orcid.org/0000-0001-6037-5232>

Universidade Federal do Ceará, Brasil

E-mail: [pgpimentel@hotmail.com](mailto:pgpimentel@hotmail.com)

**Resumo**

Objetivou-se a seleção genotípica da mandioca (*Manihot esculenta crantz*) para formação de grupos divergentes de genótipos. Inicialmente foram utilizadas 23 variáveis quantitativas de 28 genótipos; após a análise de componentes principais esse número foi reduzido a 7 variáveis (diâmetro do caule, número manivas-sementes por planta, comprimento da raiz, diâmetro da raiz, potencial hídrico foliar antemanhã, potencial hídrico foliar ao meio-dia, produtividade da parte aérea). A análise de agrupamento por meio da distância de Chebyshev e do método de ligação Mcquitty foi utilizada para gerar um dendrograma cuja validação se deu a partir de um coeficiente cofenético de 0,8. Foram formados dois grupos compostos de 24 e 4 genótipos de mandioca. Esses foram indicados no dendrograma por meio dos índices de Ratkowsky, McClain e KL. O segundo grupo que é formado pelos genótipos tussuma, Caititi, Poti Branca e mulatinha apresentou maior frequência das variáveis que descrevem a parte superior da mandioca. Essas informações são importantes para a criação de banco de informações de espécies rústicas e o melhoramento das mesmas.

**Palavras-chave:** Componentes principais; Distância de Chebyshev; Diâmetro da raiz.

### **Abstract**

The objective was the genetic selection of cassava (*Manihot esculenta crantz*) to formation of groups. Initially, 23 quantitative variables from 28 genotypes were used. After principal-components analysis, this number was reduced to 7 variables (stalk diameter, number of maniva-seed per plant, root length, root diameter, leaf water potential in the morning, leaf water potential in the midday and productivity of aerial part. A cluster analysis using Chebyshev distance and Mcquitty connection methods were used to generate a dendrogram whose validation was based from cophenetic coefficient of 0.8. Two groups composed of 24 and 4 cassava genotypes were formed. These were indicated in the dendrogram using Ratkowsky, McClain and KL indexes. The second group formed by the tussuma, Caititi, Poti Branca and mulatinha genotypes showed higher frequency of the variables that describe an upper part of cassava. This information is important for the creation of database for rustic species and their improvement.

**Keywords:** Chebyshev distance; Plant varieties; Principal-components.

### **Resumen**

El objetivo es seleccionar genotipos de manihot (*Manihot esculenta crantz*) para formar grupos divergentes de genotipos. Inicialmente se utilizaron 23 variables cuantitativas de 28 genotipos, luego del análisis de componentes principales se redujo de 23 a 7 variables (diámetro del tallo, número de semillas de maní por planta, longitud de la raíz, diámetro de la raíz, potencial hídrico de las hojas en la mañana, potencial hídrico de las hojas al mediodía, productividad aérea). Se utilizó el análisis de agrupamiento, con distancia de Chebyshev y el método de unión Mcquitty, generando un dendrograma, que fue validado con un coeficiente confenético de 0,8. Se formaron dos grupos, uno con 24 genotipos y el segundo grupo compuesto por 4 genotipos de mandioca respectivamente. Los índices de Ratkowsky, McClain y KL indicaron ambos grupos en el dendrograma. El segundo grupo, formado por los genotipos tussuma (22), Caititi (6), Poti Branca (15) y mulatinha (12), presenta una mayor frecuencia de las variables que describen la parte superior de la yuca. Esta información es importante para la creación de un banco de información sobre las especies rusas y su mejora.

**Palabras clave:** Componentes principales; Distancia de Chebyshev; Diámetro de la raíz.

## 1. Introdução

A mandioca (*Manihot esculenta* Crantz), popularmente conhecida por macaxeira ou aipim, é uma das principais fontes de carboidratos para milhões de pessoas ao redor do mundo e em especial nos países em desenvolvimento (Iyer et al., 2010). O Brasil é o país que possivelmente originou essa espécie, que é cultivada praticamente em todo território brasileiro, desempenhando papel importante na indústria e alimentação (Zago et al., 2017).

O estado da Bahia é o terceiro maior produtor da raiz e obteve uma produção de 1.956.103 t/ha de mandioca no ano de 2016, e em especial, o município de Cândido Sales - BA que responde por uma área plantada de 6.000 hectares (IBGE, 2016), constitui-se o maior produtor dessa cultura neste estado. Com o passar dos anos a produção diminuiu na cidade e no estado da Bahia.

Uma das causas da queda de produtividade da mandioca no Brasil, é a falta tanto de cultivares mais resistentes aos vários tipos de solos e clima existentes quanto de informação a respeito do manuseio desse plantio por parte dos pequenos agricultores (Lopes et al., 2010). A mandioca é uma espécie rústica e de plantio simples, geradora de renda no meio rural, ofertando aos agricultores oportunidades de trabalho. Por este motivo, torna-se importante o seu aprimoramento por parte dos programas de melhoramento genético auxiliando os agricultores na promoção de cultivos mais resistentes ao processo do plantio até a colheita.

Uma etapa fundamental nos programas de melhoramento é a compreensão da divergência genética com o intuito de realizar cruzamentos entre as espécies. A avaliação da divergência genética de espécies populares como a mandioca, auxilia na conservação dessas culturas no âmbito regional mantendo vivas espécies selvagens. Vale ressaltar que a variação do clima tem se intensificado ao passar dos anos e é preciso apostar em cultivares mais resistentes a essas mudanças (FAO, 2019).

Alguns procedimentos vêm sendo muito utilizados para caracterizar a divergência genética, como aqueles biométricos que podem ser realizados a partir de análise dialélica. Nesses procedimentos é preciso realizar combinações entre as variedades, e observar sua viabilidade quanto ao número de genótipos. Por outro lado, nos meios preditivos existe a vantagem de conhecer os fatores que indicam divergência entre os indivíduos antes mesmo da realização dos cruzamentos e, portanto, facilitar a escolha: de genótipos mais propícios a geração de espécies superiores.

No método preditivo destaca-se a análise multivariada, onde não é necessário ter cruzamentos prévios entre espécies (Fonseca et al., 2006). A utilização desse conjunto de

técnicas combinado com a caracterização morfológica de cultivares tem proporcionado bons resultados em estudos de divergência genética com a romã (Mahajan et al., 2018), azeite (Abuzayed et al., 2018), e batata doce (Andrade et al., 2017).

A fim de contribuir com a agricultura da mandioca na cidade de Candido Sales – BA, esse trabalho foi realizado para compreender as diferenças genéticas de 28 genótipos de mandioca utilizando a estatística multivariada por meio da análise de componentes principais e análise de agrupamentos. Demonstrar as principais características em cada grupo de genótipos é de suma importância para os programas de melhoramento genético, uma vez que por meio de seus cruzamentos com espécies superiores informações úteis sobre os genótipos podem ser obtidas.

## 2. Metodologia

Segundo Pereira et al., (2018) o presente artigo pode ser classificado como pesquisa exploratória de natureza quantitativa. O experimento foi instalado na região de Bomba situado a 15°18'13" de latitude Sul e 41°17'32" de longitude, no município de Cândido Sales – Ba (Guimarães, 2013). Os dados são oriundos de um experimento organizado em blocos casualizados com 28 tratamentos e três repetições, com parcelas de 8,4 m de comprimento e 3,0 m de largura, com espaço de plantio de 1,0 m entre linhas e 0,60 m entre plantas, 26 plantas úteis totalizando 15,6 m<sup>2</sup>.

Com a finalidade de reduzir as variáveis, a análise de componentes principais foi aplicada aos dados de genótipos de mandioca. As componentes principais são obtidas a partir de um conjunto de variáveis aleatórias observadas  $X_1, X_2, X_3, \dots, X_i$  e surgem de uma combinação linear  $Y_j = b_{j1}X_1 + b_{j2}X_2 + \dots + b_{ji}X_i$ , com base na matriz de correlação  $R$  ou variância-covariância  $S$ . A solução é calculada através da equação  $|R - \lambda I| = 0$ , em que  $I$  é a matriz identidade,  $\lambda$  são os autovalores (os autovalores representam a variância explicada em cada componente), e  $b_j$  são os autovetores associados aos autovalores.

A análise de componentes principais foi aplicada nas variáveis, totalizando 28 genótipos de mandioca e 23 variáveis. A análise de componentes principais é utilizada em estudos sobre dimensão de dados (Lu et al., 2019).

Para organizar um dendrograma é comum utilizar uma distância multivariada, combinada com um método de ligação que demonstre quanto um elemento ou indivíduo está distante ou é semelhante a outro. A distância de Chebyshev dada entre dois objetos  $a_i$  e  $a_j$  é o

valor máximo para todas as variáveis, a saber  $D(a_i, a_j) = \max_v |a_{iv} - a_{jv}|$ .

O método de ligação Mcquitty é definido segundo Charrad et al. (2014) como a distância entre os grupos  $C_i$  e  $C_j$ ; é a média ponderada entre as diferenças dos grupos  $D_{ij} = \frac{(D_{ik}+D_{il})}{2}$ , em que  $C_j$  é formado a partir da soma do grupos  $C_k$  e  $C_l$ .

A distância de Chebyshev combinada com método de ligação Mcquitty gera uma estrutura gráfica denominada dendrograma. Na literatura existem várias distâncias que podem ser utilizadas em análise de agrupamento (Johnson & Wichern, 2014).

A validação do dendrograma foi realizado com o coeficiente cofenético que é definido por:

$$r = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2 \sum_i (y_i - \bar{y})^2}}.$$

Foram utilizados 3 índices (Ratkowsky, McClain e KL) para definir os grupos no dendrograma havendo indicação majoritária de 2 grupos divergentes.

O índice de Ratkowsky segundo Charrad et al., (2014) é determinado como o número ótimo de grupos com base na seguinte razão:

$$Ratkowsky = \frac{\bar{S}}{q^{1/2}}, \quad (1)$$

em que:

$$\bar{S}^2 = \frac{1}{p} \sum_{j=1}^p \frac{BGSS_j}{TSS_j}, \quad (2)$$

com

$$BGSS_j = \sum_{k=1}^q n_k (c_{kj} - \bar{x}_j) \quad (3)$$

$$TSS_j = \sum_{i=1}^n (x_{ij} - \bar{x}_j)^2. \quad (4)$$

Vale salientar que  $q$  é o número ótimo de grupos para o qual a razão dada em (1) é máxima e  $p$  é o número de variáveis. O índice de McClain (ou índice de Rao) segundo Charrad et al., (2014) é a relação entre dois termos da equação a seguir  $McClain = \frac{\bar{S}_w}{\bar{S}_b}$ . O numerador é a média das distâncias intra-grupos (dentro dos grupos) e o denominador é a média das distâncias entre-grupos (entre os grupos). O valor mínimo do índice é usado para indicar o número ideal de grupos. O índice de KL segundo Charrad et al. (2014) e dado por:

$$KL(q) = \left| \frac{DIFF_q}{DIFF_{q+1}} \right|$$

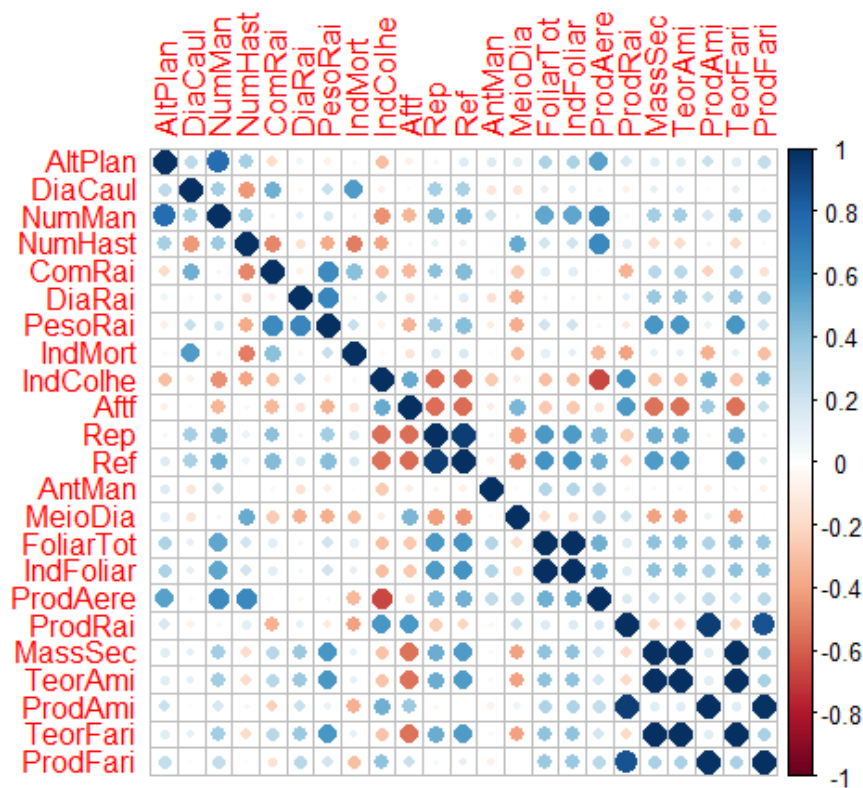
em que  $DIFF_q = (q - 1)^{2/p}$  traço  $(w_{q-1}) - q^{2/p}$  traço  $(w_q)$ . O valor de  $q$ , o qual maximiza  $KL(q)$ , é considerado como o número ideal de grupos e  $w_q$  é a soma de quadrados dos erros intra-grupos. Aplicações desses índices podem ser vistas em (Melendez-Melendez et al., 2019)

A variedade Sergipe aparece em dois tratamentos, um com manivas de outras regiões (Sergipe) e outra com manivas criadas do método de multiplicação rápida (Sergipe MR) (Guimarães, 2013). A análise foi realizada no software R versão 3.4.0.

### 3. Resultados e Discussão.

Na Figura 1 encontra-se as correlações de todas as variáveis. Pode-se observar correlação positiva entre número de manivas e sementes por planta com as variáveis altura por planta e produtividade da parte aérea.

**Figura 1.** Correlação das 23 variáveis observadas nos 28 genótipos.

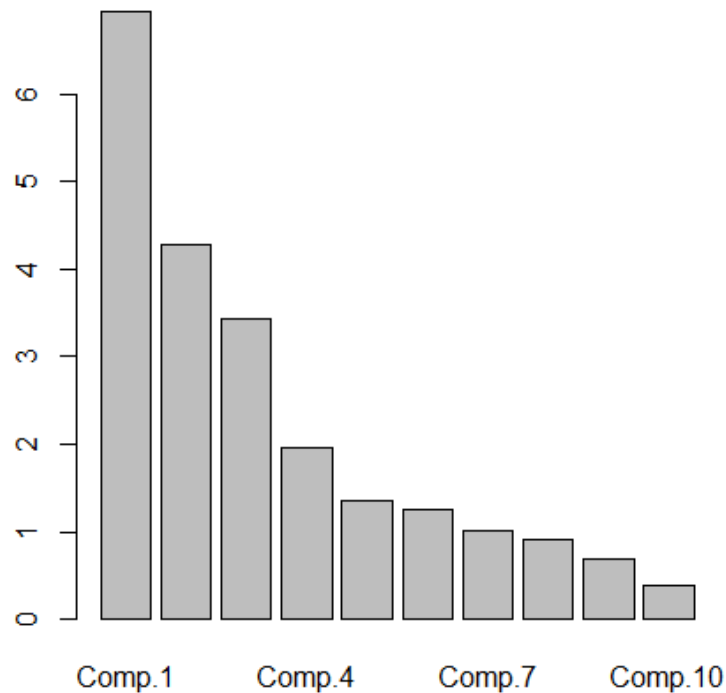


Diversos são os estudos de correlações de variáveis de espécies rústicas (Neto et al., 2017). O potencial hídrico foliar ao meio dia possui correlação negativa com o comprimento e

diâmetro da raiz. As variáveis teor de amido, produtividade de amido, teor de farinha e produtividade de farinha possuem correlações positivas.

Na Figura 2 podemos observar que a partir da componente 7 a variação da variância explicada pelas demais componentes se aproxima de zero, sugerindo que 7 componentes principais são o suficiente para explicar a variação dos dados.

**Figura 2.** Número de componentes necessários para explicar a variação dos dados.



As variáveis de maior correlação com as componentes 8, 9 e 10 foram excluídas, totalizando 16 variáveis descartadas e restando 7 variáveis na base de dados a saber: *DiaCaule* (Diâmetro do caule), *NumMan* (Número manivas-sementes por planta), *ComRai* (Comprimento da Raiz), *DiaRai* (Diâmetro da Raiz), *AntMan* (Potencial hídrico foliar antemanhã), *MeioDia* (Potencial hídrico foliar ao meio-dia) e *ProdAere* (Produtividade da parte aérea).

As variáveis que admitem forte correlação com as componentes principais e que retém pouca variância devem ser descartadas (Barbosa et al., 2005).

Na Tabela 1, estão descritas as 7 variáveis finais nos 28 genótipos de mandioca que foram aplicados os métodos de agrupamento multivariados para criação de grupos divergentes.



**Tabela 1.** Valores das 7 variáveis selecionadas das 23, por meio da análise de componentes principais dos 28 genótipos de mandioca.

Genótipos	DiaCaul	NumMan	ComRai	DiaRai	AntMan	MeioDia	ProdAere
Amansa_Burro	1,83	8,59	5,08	5,11	0,48	1,01	10,8
Aramaris	1,97	7,34	5,60	5,59	0,48	0,85	6,04
Bom_Jardim	2,02	9,22	5,13	5,53	0,43	1,15	7,78
Bromadeira	1,94	6,32	5,55	5,14	0,58	1,12	4,66
Caipira	2,03	10,7	5,38	5,92	0,37	0,89	8,66
Caitite	1,75	12,8	5,06	5,23	0,78	1,15	14,1
Caravela	1,78	9,30	5,49	5,58	0,43	1,02	6,71
Kiriris	2,08	10,5	5,22	5,37	0,30	1,09	6,40
Lagoao	1,94	6,37	5,80	5,84	0,42	1,02	5,92
Lavra_Velha	1,80	4,59	5,96	4,51	0,43	1,03	7,49
Malacacheta_MR	1,84	7,25	5,54	5,33	0,38	1,17	10,6
Mulatinha	1,81	10,3	5,60	5,74	0,52	1,13	20,3
Parazinha	1,95	10,9	5,36	5,01	0,57	1,09	9,85
Peru	1,87	6,94	4,89	5,17	0,34	1,09	12,3
Poti_Branca	2,08	16,2	5,31	5,26	0,47	1,14	16,5
Salangor	2,22	9,63	6,35	4,17	0,56	1,05	12,0
Sergipana	2,06	9,00	5,58	4,74	0,45	1,25	9,08
Sergipe	1,97	10,5	5,55	5,41	0,46	1,13	12,5
Sergipe_MR	1,98	10,1	5,54	5,06	0,24	1,11	10,0
Simbe	1,86	9,74	5,03	5,02	0,48	0,96	11,4
Tapioqueira	1,99	6,63	5,23	5,59	0,51	0,77	6,72
Tussuma	2,32	13,5	6,12	5,49	0,53	1,01	12,6
Verdinha	1,99	7,40	5,50	5,13	0,58	1,01	9,12
G2006_4	1,94	5,35	5,31	5,70	0,42	1,21	6,63
G2006_5	1,88	7,29	5,14	5,21	0,50	1,17	10,1
G2006_8	1,79	8,82	4,78	4,64	0,50	1,32	10,9
G2006_10	2,09	6,21	5,05	4,72	0,40	1,22	7,63
G2006_12	1,81	6,16	5,07	5,02	0,517	1,04	7,64

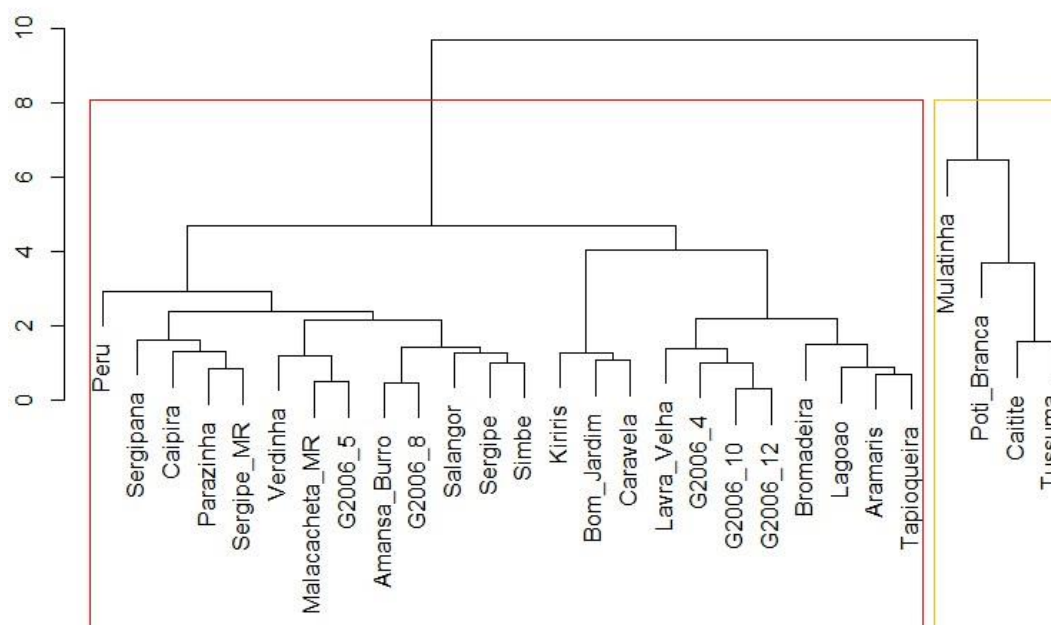
Na Tabela 2 encontra-se o resultado da análise descritiva dos dados (para cada variável) após a redução da dimensão, via componentes principais. A variável *MeioDia* possui o menor coeficiente de variação ( $CV = 0,296$ ) e um desvio padrão  $S = 0,119$  indicando uma baixa variação nos valores da mesma. Por outro lado, para a variável *DiaRai*, o coeficiente de variação ( $CV = 88,160$ ) demonstra alta variedade de dimensões das raízes. Ainda na Tabela 2, pode-se observar os valores médios e medianos além dos máximos e mínimos para cada variável.

**Tabela 2.** Análise descritiva dos das variáveis após a redução da dimensão.

Estatísticas	DiaCaul	NumMan	ComRai	DiaRai	AntMan	MeioDia	ProdAere
Mín.	1,750	4,780	4,170	0,2430	0,772	22,34	4,660
S	0,135	2,622	0,363	0,414	0,102	0,119	3,457
$\bar{X}$	1,945	5,370	5,220	0,4735	1,091	39,74	9,485
$\bar{X}$	1,950	5,401	5,223	0,4696	1,081	40,25	9,800
CV	6,923	48,547	6,950	88,160	9,436	0,296	35,28
Máx.	2,320	6,350	5,920	0,7830	1,323	59,76	20,290

Visando uma divisão criteriosa para a formação dos grupos divergentes adotou-se a distância de Chebyshev com o método de ligação Mcquitty gerando uma estrutura denominada dendrograma Figura 3.

**Figura 3.** Dendrograma gerado através da distância de Chebyshev com o método de ligação Mcquitty, com os dois grupos indicados com os índices de Ratkowsky, McClain e KL.



Utilizou-se os índices Ratkowsky, McClain e KL para a indicação de quantos grupos poderiam ser formados no dendrograma. Diversos são os índices utilizados para validar e determinar a quantidade de grupos ideal (Charrad et al., 2014).

Os três índices indicaram a mesma quantidade de grupos; os 2 grupos podem ser vistos divididos em cor vermelha e laranja na Figura 3. O primeiro grupo foi formado por 24 genótipos e dentre esses os que já foram melhorados (G2006\_4, G2006\_5, G2006\_8, G2006\_10,

G2006\_12). Esse grupo foi formado por genótipos que descrevem as características da raiz, a saber: *ComRai*, *DiaRai* e *DiaCaul*.

Os genótipos Tussuma, Caititi, Poti Branca e Mulatinha ficaram reunidos no segundo grupo, o qual possui maiores valores nas variáveis que descrevem a parte superior da planta como *ProdAere* e *MeioDia*. A divergência genética desses genótipos de mandioca está dividida em dois grupos e o coeficiente cofenético de 0,8, utilizado para validar o dendrograma, indica um ajuste razoável na formação dos grupos.

#### **4. Considerações Finais**

A estatística multivariada demonstrou potencial para organizar informações de genótipos de mandioca no município de Cândido Sales – BA. Pesquisadores e agricultores da região podem utilizar esses resultados para melhorar as bases de informações dos genótipos.

A distância de Chebyshev com o método de ligação Mcquitty podem ser utilizados para obtenção de grupos divergentes de genótipos de mandioca. Obter grupos com características distintas é importante para o conhecimento de espécies rústicas.

As Técnicas da estatística multivariada possuem um vasto apoio didático e computacional, podendo ser expandidas para outros municípios brasileiros que se destacam nos cultivos de espécies rústicas.

#### **Agradecimentos**

A Universidade Federal Rural de Pernambuco (UFRPE). O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de nível Superior - Brasil (CAPES) - Código de Financiamento 001

#### **Referências**

Abuzayed, M., Frary, A., & Doganlar, S. (2018). Genetic diversity of some Palestinian and Turkish olive (*Olea europaea* L.) germplasm determined with SSR markers. *IUG Journal of Natural Studies*, 26 (1), 10 - 17. Recuperado de <https://journal.iugaza.edu.ps/index.php/IUGNS/article/viewFile/3362/2036>

Andrade, E. K. V. D., Andrade Júnior, V. C. D., Laia, M. L. D., Fernandes, J. S. C., Oliveira, A. J. M., & Azevedo, A. M. (2017). Genetic dissimilarity among sweet potato genotypes using morphological and molecular descriptors. *Acta Scientiarum. Agronomy*, 39 (4), 447 - 455. doi: <http://dx.doi.org/10.4025/actasciagron.v39i4.32847>.

Barbosa, L., Lopes, P. S., Regazzi, A. J., Guimarães, S. E. F., & Torres, R. D. A. (2005). Avaliação de características de carcaça de suínos utilizando-se a análise dos componentes principais. *Revista Brasileira de Zootecnia*, 34 (6), 2009 - 2217. doi: <http://dx.doi.org/10.1590/S1516-35982005000700007>.

Charrad, M., Ghazzali, N., Boiteau, V., Niknafs, A., & Charrad, M. M. (2014). NbClust: An R Package for Determining the Relevant Number of Clusters in a Data Set. *Journal of statistical software*, 61 (6), 1 - 36. Recuperado de [http://cedric.cnam.fr/fichiers/art\\_2554.pdf](http://cedric.cnam.fr/fichiers/art_2554.pdf)

Fonseca, A. F., Sediya, T. S., Cruz, C. D., Sakaiyama, N. S., Ferrão, M. A., Ferrão, R. G., & Bragança, S. M. (2006). Divergência genética em café conilon. *Pesquisa agropecuária brasileira*, 41 (4), 599 - 605.

Guimarães, D. G. (2013). Avaliação de genótipos de mandioca em Cândido Sales – BA.

IBGE. (2016). Sidra. Website do Instituto Brasileiro de Geografia e Estatística [IBGE]. Recuperado de <https://sidra.ibge.gov.br/tabela/5457>

Iyer, S., Mattinson, D. S., & Fellman, J. K. (2010). Study of the early events leading to cassava root postharvest deterioration. *Tropical Plant Biology*, 3 (3), 151 - 165.

Johnson, R., & Wichern, D. (2014). *Applied Multivariate Statistical Analysis*. New Jersey, USA: Pearson.

Lopes, A. C., Viana, A. E. S., Matsomoto, S. N., Júnior, C., dos Santos, N., & São José, A. R. (2010). Complementação da irrigação e épocas de colheita de mandioca cv. coqueiro no Planalto de Conquista, BA. *Ciência e agrotecnologia*, 34 (3), 579-587. doi.org/10.1590/S1413-70542010000300008

Lu, C., Feng, J., Chen, Y., Liu, W., Lin, Z., & Yan, S. (2019). Tensor Robust Principal Component Analysis with A New Tensor Nuclear Norm. *IEEE transactions on pattern analysis and machine intelligence*, 1-22. Recuperado de <https://arxiv.org/pdf/1804.03728.pdf>

Mahajan, R., Javed, A., & Kapoor, N. (2018). Characterization of genetic diversity of wild pomegranate collected from Himachal Pradesh, India. *Annals of Plant Sciences*, 2042 - 2046. doi: <http://dx.doi.org/10.21746/aps.2018.7.2.10>

Melendez-Melendez, G., Cruz-Paz, D., Carrasco-Ochoa, J. A., & Martínez-Trinidad, J. F. (2019). An improved algorithm for partial clustering. *Expert Systems with Applications*, 282 - 291. doi: <https://doi.org/10.1016/j.eswa.2018.12.027>

Neto, S., Silva, T. O. D., Blank, A. F., Silva, J. O. D., & Araújo Filho, R. N. D. (2017). Produtividade de clones de batata doce em função de doses de nitrogênio. *Horticultura Brasileira*, 35 (3), 445 - 452. <https://doi.org/10.1590/s0102-053620170322>

Organização das Nações Unidas para Alimentação e a Agricultura. [FAO] (2019) Folha informativa. Recuperado de <http://www.fao.org/mozambique/news/detail/pt/c/1198477/>

Pereira, A.S.; Shitsuka, D.M.; Parreira, F.J. & Shitsuka, R. (2018). Metodologia da pesquisa científica. Editora UAB/NTE/UFMS, Santa Maria/RS. Disponível em:  
[http://repositorio.ufsm.br/bitstream/handle/1/15824/Lic\\_Computacao\\_Metodologia-PesquisaCientifica.pdf?sequence=1](http://repositorio.ufsm.br/bitstream/handle/1/15824/Lic_Computacao_Metodologia-PesquisaCientifica.pdf?sequence=1)

Zago, B. W., Barelli, M. A. A., Hoogerheide, E. S. S., Corrêa, C. L., Delforno, G. I. S., & da Silva, C. J. (2017). Morphological diversity of cassava accessions of the south-central mesoregion of the State of Mato Grosso, *Embrapa Agrossilvipastoril*, 16 (3), 1 - 10  
doi <http://dx.doi.org/10.4238/gmr16039725>

**Porcentagem de contribuição de cada autor no manuscrito**

Diego Alves Gomes – 50%

Lucas Silva do Amaral – 5%

Denise Stéphanie de Almeida Ferreira – 5%

Guilherme Rocha Moreira – 10%

Frank Gomes-Silva – 10%

Maria Lindomárcia Leonardo da Costa – 5%

Moacyr Cunha Filho – 5%

André Luiz Pinto dos Santos – 5%

Mércia Regina Pereira de Figueiredo – 3%

Patrícia Guimarães Pimentel – 2%