

Caracterização agronômica das pimentas biquinho vermelha e bode amarela por análise de componentes principais

Beatriz B. de Carvalho¹, Beatriz R. Morales¹, Christiane de F. M. França², Fernando C. Sala³, Gilberto R. Liska², Josiane Rodrigues^{2†}

¹ Discente do curso de Bacharelado em Engenharia Agrônômica; Centro de Ciências Agrárias (CCA); Universidade Federal de São Carlos (UFSCar); Araras – São Paulo, Brasil.

² Docente do Departamento de Tecnologia Agroindustrial e Socioeconomia Rural (DTAiSeR); Centro de Ciências Agrárias (CCA); Universidade Federal de São Carlos (UFSCar); Araras – São Paulo, Brasil.

³ Docente do Departamento de Departamento de Biotecnologia e Produção Vegetal e Animal (DBPVA), Centro de Ciências Agrárias (CCA), Universidade Federal de São Carlos (UFSCar), Araras – São Paulo, Brasil.

Resumo: As pimentas são parte da riqueza cultural brasileira e um valioso patrimônio da biodiversidade do Brasil. Foram domesticadas ao longo dos anos e, atualmente, existem diversas variedades de pimenta, cada uma delas com suas características próprias. Devido aos benefícios trazidos pelas pimentas, a sua caracterização agronômica é importante pois pode servir como base para o desenvolvimento de novas cultivares. Assim, o presente trabalho teve por objetivo caracterizar e comparar as pimentas biquinho vermelha e bode amarela quanto a algumas de suas características agronômicas, utilizando a análise de componentes principais. As variáveis analisadas na maturidade comercial das plantas foram: altura, altura da primeira bifurcação, diâmetro transversal e longitudinal da copa, compacidade da copa, número de frutos, número de folhas, índice SPAD, diâmetro de fruto, comprimento de fruto, peso fresco de parte aérea, cobertura de vaso, comprimento de folha, largura de folha, número de ramificações laterais, comprimento de ramificações laterais e dias após o transplântio para florescimento (DAT). Pelo critério de Kaiser, os quatro primeiros componentes principais foram selecionados na análise, os quais captaram aproximadamente 92% da variabilidade dos dados. A pimenta biquinho vermelha se caracterizou por valores maiores de número de frutos, diâmetro de fruto, comprimento de fruto, peso fresco de parte aérea, cobertura de vaso, comprimento de folha e largura de folha. Já a pimenta bode amarela foi caracterizada por valores maiores de altura da planta, altura da primeira bifurcação, compacidade de copa, número de folhas e DAT. Não houve diferença entre as pimentas quanto às outras variáveis agronômicas analisadas.

Palavras-chave: Análise multivariada; *Capsicum chinense*; melhoramento genético.

Agronomic characterization of biquinho vermelha and bode amarela peppers by principal component analysis

Abstract: Peppers are part of Brazilian cultural richness and a valuable patrimony of Brazil's biodiversity. They have been domesticated over the years and, currently, there are several varieties of pepper, each with its own characteristics. Due to the benefits brought by peppers, their agronomic characterization is important as it can serve as a basis for the development of new cultivars. Thus, the present work aimed to characterize and compare the biquinho vermelha and bode amarela peppers in terms of some of their agronomic characteristics, using principal component analysis. The variables analyzed in the commercial maturity of the plants were: height, height of the first bifurcation, transversal and longitudinal diameter of the plant top, crown compactness, number of fruits, number of leaves, SPAD index, fruit diameter, fruit length, fresh weight of part aerial, pot coverage, leaf length, leaf width, number of lateral branches, length of lateral branches and days after transplanting for flowering (DAT). Using Kaiser's criterion, four principal components were selected in the analysis, which captured approximately 92% of the data variability. The biquinho vermelha pepper was characterized by higher values for number of fruits, fruit diameter, fruit length, fresh weight of aerial part, pot coverage, leaf length and leaf width. The bode amarela pepper was characterized by higher values of plant height, height of the first fork, crown compactness, number of leaves and DAT. There was no difference between the peppers regarding the other agronomic variables analyzed.

Keywords: Multivariate analysis; *Capsicum chinense*; genetical enhancement.

† Autor correspondente: josirodrigues@ufscar.br

Manuscrito recebido em: 11/07/2024

Manuscrito revisado em: 24/09/2024

Manuscrito aceito em: 30/09/2024

Introdução

As pimentas são parte da riqueza cultural brasileira e um valioso patrimônio da biodiversidade do Brasil. Foram domesticadas ao longo dos anos e, atualmente, existem diversas variedades de pimenta, cada uma delas com suas características próprias, tais como forma, cor, sabor e pungência (COSTA *et al.*, 2015). O cultivo de pimentas ocorre em todas as regiões do Brasil, e seu consumo é feito de diversas formas, seja *in natura* ou processada em geleias e compotas (FINGER; PEREIRA, 2016). Além do seu uso na alimentação e culinária, as pimentas também são utilizadas na forma medicinal, e há também um mercado promissor de venda de pimentas ornamentais (COSTA *et al.*, 2019).

O Brasil é o centro da diversidade do gênero *Capsicum*, sendo que as espécies mais cultivadas no país são a *C. annum*, *C. baccatum*, *C. chinense* e *C. frutescens*. Dentre elas, *Capsicum chinense* é a mais brasileira das espécies de *Capsicum* domesticadas, e se adapta facilmente ao clima quente e úmido, e dentre elas estão as pimentas biquinho vermelha e bode amarela. A pimenta biquinho vermelha possui o formato de uma gota, sendo reconhecida pela ausência de pungência, bem como pelo seu sabor doce. Além disso, ela é rica em vitaminas, é considerada um antioxidante, ajuda o sistema imunológico, ajuda a controlar o açúcar no sangue, a prevenir inflamações e a diminuir dores musculares. Por sua vez, a pimenta bode amarela possui frutos redondos ou achatados, aromáticos e picantes. Também é fonte de vitaminas, tem efeito antioxidante, e atua no fortalecimento do sistema imunológico, além de estimular a produção de endorfina (REIFSCHNEIDER, 2000).

Devido aos benefícios trazidos pelas pimentas, bem como à demanda por variedades com características superiores, a caracterização agrônômica das pimentas já disponíveis é importante pois pode servir como base para o desenvolvimento de novas cultivares. Na caracterização agrônômica das pimentas, entretanto, uma série de variáveis devem ser consideradas de forma simultânea. Nesse contexto, a técnica multivariada da análise de componentes principais (ACP) (PEARSON, 1901; HOTELLING, 1933; 1936) pode ser utilizada. Costa *et al.* (2020), por exemplo, empregaram a técnica para caracterizar e analisar a divergência genética entre diferentes genótipos do gênero *Capsicum* com relação a 11 caracteres quantitativos e 15 qualitativos. Da mesma forma, Nascimento *et al.* (2021) utilizaram a ACP para estimar a variabilidade genética com base em vários descritores morfoagronômicos de diferentes genótipos de pimenta.

De forma geral, a ACP possui como principal objetivo transformar as variáveis originais, correlacionadas entre si, em outro agrupamento de variáveis, de mesma dimensão, mas sem correlação, denominadas componentes principais, os quais são combinações lineares das variáveis originais e buscam reter, por ordem de estima, o máximo de informações contidas nas variáveis usadas. Sua análise é associada à redução massiva de dados retendo grande parte das informações originais. Assim, essa técnica reduz as variáveis e agrupa indivíduos conforme suas similaridades, facilitando assim o seu estudo (HONGYU *et al.*, 2016).

Diante o exposto, o presente trabalho teve por objetivo caracterizar e comparar as pimentas biquinho vermelha e bode amarela quanto a algumas de suas características agrônômicas, utilizando para isso a técnica multivariada da ACP, a qual ainda não foi utilizada na caracterização e comparação das duas variedades de pimenta. Com isso, pretende-se também verificar a eficiência da técnica na comparação agrônômica das duas variedades, bem como auxiliar os programas de melhoramento de pimentas, no intuito de desenvolver novas cultivares com características superiores àquelas já disponíveis no mercado.

Materiais e Métodos

Dados

O experimento foi conduzido em casa de vegetação no Centro de Ciências Agrárias (CCA), da Universidade Federal de São Carlos (UFSCar), município de Araras, São Paulo, Brasil (latitude -22.3114976, longitude 47.3847851, elevação 685 m). Os tratamentos utilizados no ensaio foram duas variedades de pimenta da espécie *Capsicum chinense*: pimenta biquinho vermelha e pimenta bode amarela, e o delineamento considerado foi o inteiramente casualizado, com cinco repetições para cada um dos tratamentos testados.

As variáveis analisadas no momento de maturidade comercial das plantas foram: altura (X1 - cm), altura da primeira bifurcação (X2 - cm), diâmetro longitudinal da copa (X3 - cm), diâmetro transversal da copa (X4 - cm), compacidade da copa (X5), número de frutos (X6), número de folhas (X7), índice SPAD (X8), diâmetro de fruto (X9 - cm), comprimento de fruto (X10 - cm), peso fresco de parte aérea (X11 - g), cobertura de vaso (X12), comprimento de folha (X13 - cm), largura de folha (X14 - cm), número de ramificações laterais (X15), dias após o transplântio (X16) e comprimento de ramificações laterais (X17 - cm).

Após a coleta, os dados foram analisados estatisticamente pela técnica multivariada de análise de componentes principais (ACP) (PEARSON, 1901; HOTELLING, 1933; 1936), a qual permite comparar as variedades de pimenta biquinho vermelha e bode amarela com relação a cada uma das 17 variáveis agrônômicas analisadas de forma simultânea. Em suma, o objetivo da técnica é a obtenção de um pequeno número de combinações lineares (componentes principais) do conjunto de variáveis originais, de forma que esses componentes retenham o máximo possível da informação contida nas variáveis originais.

Análise de Componentes Principais

Sejam X_1, X_2, \dots, X_p as variáveis agrônômicas avaliadas em cada um dos n indivíduos do ensaio. Este conjunto de medidas originais forma uma matriz de dados \mathbf{X} de dimensão $n \times p$:

$$\mathbf{X} = \begin{bmatrix} x_{11} & \cdots & x_{1p} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{np} \end{bmatrix}.$$

Sendo as variáveis X_1, X_2, \dots, X_p correlacionadas entre si, e apresentando médias $\mu_1, \mu_2, \dots, \mu_p$ e variâncias $\sigma_1^2, \sigma_2^2, \dots, \sigma_p^2$, respectivamente, é possível afirmar que existe uma covariância entre a i -ésima e a k -ésima variável, a qual é definida por σ_{ik} . Dessa forma, as p variáveis podem ser expressas na forma vetorial por

$$\mathbf{X} = \begin{bmatrix} X_1 \\ \dots \\ X_p \end{bmatrix},$$

com vetor de médias

$$\mu = \begin{bmatrix} \mu_1 \\ \dots \\ \mu_p \end{bmatrix}$$

e matriz de variâncias e covariâncias

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \dots & \sigma_{1p} \\ \vdots & \ddots & \vdots \\ \sigma_{n1} & \dots & \sigma_p^2 \end{bmatrix}.$$

A partir da matriz Σ é possível determinar os seus autovalores $\lambda_1, \lambda_2, \dots, \lambda_p$, bem como os seus respectivos autovetores e_1, e_2, \dots, e_p , em que $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$. O i -ésimo componente principal da ACP, $i=1, \dots, p$, é definido da forma (JOHNSON; WICHERN, 1998):

$$Z_i = e_i' X = e_{i1} X_1 + e_{i2} X_2 + \dots + e_{ip} X_p.$$

Por sua vez, a variância de cada Z_i é dada por (JOHNSON; WICHERN, 1998):

$$\text{Var}(Z_i) = \text{Var}(e_i' X) = e_i' \text{Var}(X) e_i = e_i' \Sigma e_i.$$

Utilizando a decomposição espectral, é possível escrever a matriz de variâncias e as covariâncias como $\Sigma = P \Lambda P'$, sendo P a matriz composta pelos autovetores de Σ em suas colunas e Λ a matriz diagonal contendo os autovalores de Σ , donde segue que:

$$\text{tr}(\Sigma) = \text{tr}(P \Lambda P') = \text{tr}(\Lambda P P') = \text{tr}(\Lambda I) = \text{tr}(\Lambda) = \sum_{i=1}^p \lambda_i.$$

Assim,

$$\text{tr}(\Sigma) = \sum_{i=1}^p \sigma_i^2 = \sum_{i=1}^p \lambda_i.$$

Portanto, a variabilidade total das variáveis originais é igual a variabilidade contida nos componentes principais. A contribuição de cada componente principal, por sua vez, é expressa em porcentagem e, considerando o k -ésimo componente, ela é dada por (JOHNSON; WICHERN, 1998; HONGYU *et al.*, 2016):

$$C_k = \frac{\text{Var}(Z_i)}{\sum_{i=1}^p \text{Var}(Z_i)} \times 100 = \frac{\lambda_i}{\sum_{i=1}^p \lambda_i} \times 100 = \frac{\lambda_i}{\text{tr}(\Sigma)} \times 100.$$

Em prática, escolhe-se primeiramente o componente principal de maior variância ($\max_i \lambda_i$), responsável por captar o máximo de variabilidade dos dados. Em segundo lugar, escolhe-se o componente que apresenta a segunda maior variância, e assim sucessivamente, até o componente de menor variância (MANLY, 1986). Em muitos casos, adota-se modelos que expliquem ao menos 80% da variabilidade total do conjunto de dados (HONGYU *et al.*, 2016). Outro método bastante utilizado por pesquisadores, conhecido como critério de Kaiser (KAISER, 1958), consiste em selecionar os componentes principais com autovalores maiores que a unidade ($\lambda_i > 1$).

Uma vez que os componentes principais são influenciados pela escala das variáveis originais, justamente porque a matriz de variâncias e covariâncias Σ é sensível à escala de um par de variáveis, no presente trabalho a ACP foi desenvolvida sobre a matriz de correlações das variáveis originais, denominada R , ao invés da matriz Σ . Todas as análises estatísticas foram desenvolvidas no *software* R (R CORE TEAM, 2024), utilizando-se dos pacotes *FactoMineR*, *factoextra*, *ggcorrplot* e *ggplot2*.

Resultados e Discussão

Pelo critério de Kaiser, os quatro primeiros componentes principais foram selecionados na análise ($\lambda_i > 1$), os quais foram responsáveis por captar cerca de 92% da variabilidade total contida nos dados (Tabela 1). A ACP, portanto, possibilitou uma redução no número de variáveis, facilitando assim a análise e a interpretação do conjunto de dados. Com o objetivo de verificar quais variáveis agronômicas se destacaram em cada um desses quatro componentes, foram então calculados os autovetores associados a eles, permitindo assim enxergar cada um dos componentes como combinações lineares das variáveis originais mensuradas no experimento (Tabela 2).

Table 1: Principal components, eigenvalues, proportion of total variability explained by each component, and cumulative proportion.

Componente principal	Autovalor	Proporção	Proporção acumulada
Z_1	10,50	61,77	61,77
Z_2	2,69	15,83	77,60
Z_3	1,36	8,01	85,61
Z_4	1,07	6,27	91,89
Z_5	0,63	3,71	95,59
Z_6	0,29	1,68	97,27
Z_7	0,27	1,60	98,87
Z_8	0,14	0,82	99,69
Z_9	0,05	0,31	100,00

Source: from the authors (2024).

Nos componentes, foram observadas as variáveis com cargas superiores a 0,60. Dessa forma, no primeiro componente principal destacaram-se de forma positiva as variáveis número de frutos (X6), diâmetro de fruto (X9), comprimento de fruto (X10), peso fresco de parte aérea (X11), cobertura de vaso (X12), comprimento de folha (X13) e largura de folha (X14), enquanto as variáveis altura (X1), altura da primeira bifurcação (X2), compacidade de copa (X5), número de folhas (X7) e dias após o transplântio para florescimento (X16) se destacaram de forma negativa

nesse componente. Por sua vez, no segundo componente principal destacaram-se as variáveis diâmetro transversal da copa (X4), número de ramificações laterais (X15) e comprimento de ramificações laterais (X17), enquanto no terceiro e no quarto componente se destacaram as variáveis índice SPAD (X8) e diâmetro longitudinal da copa (X3), respectivamente (Tabela 2).

Table 2: Principal components Z_1, Z_2, Z_3 and Z_4 as linear combinations of the original variables.

Variável	Z_1	Z_2	Z_3	Z_4
X_1	-0,88	-0,02	0,09	0,42
X_2	-0,96	-0,15	0,03	-0,06
X_3	0,57	0,19	-0,41	0,66
X_4	-0,57	0,76	0,06	0,12
X_5	-0,72	0,58	0,20	-0,13
X_6	0,83	0,38	-0,04	0,26
X_7	-0,86	0,33	-0,04	0,18
X_8	-0,25	-0,42	0,76	0,23
X_9	0,92	0,06	0,06	-0,20
X_{10}	0,99	0,00	0,01	-0,09
X_{11}	0,79	0,53	0,06	0,13
X_{12}	0,91	0,32	0,03	-0,10
X_{13}	0,85	0,24	0,41	0,01
X_{14}	0,84	0,15	0,41	-0,09
X_{15}	-0,54	0,66	0,36	0,06
X_{16}	-0,95	-0,03	0,08	-0,05
X_{17}	-0,51	0,63	-0,28	-0,42

Source: from the authors (2024).

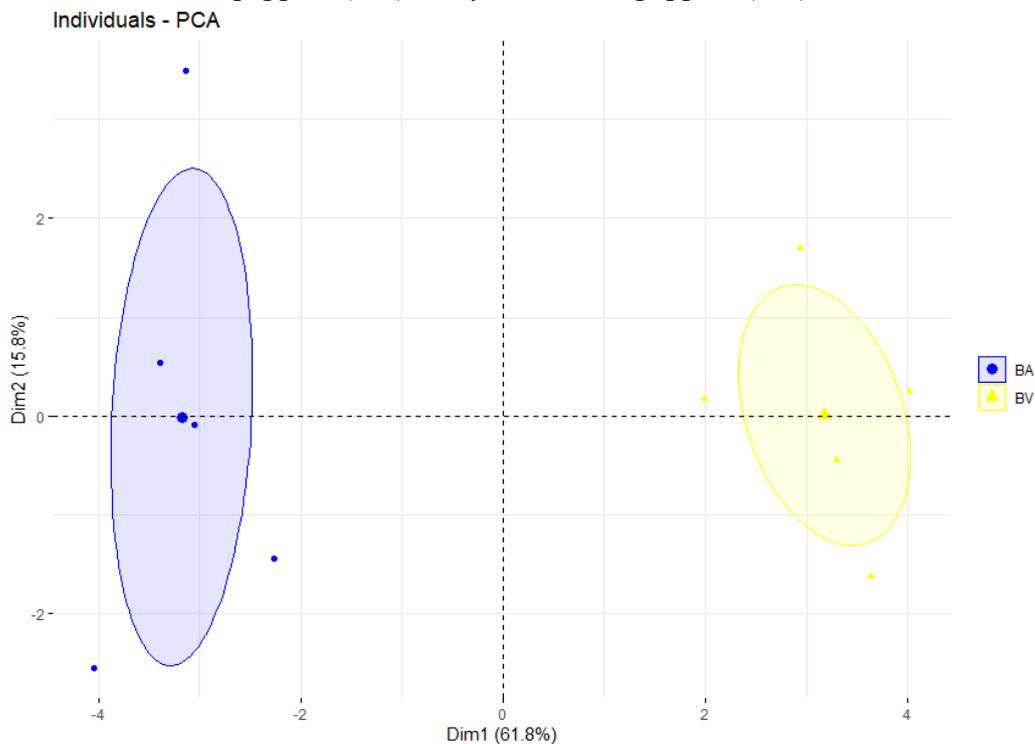
Legend: X1: altura da planta; X2: altura da primeira bifurcação; X3: diâmetro transversal da copa; X4: diâmetro longitudinal da copa; X5: compacidade da copa; X6: número de frutos; X7: número de folhas; X8: índice SPAD; X9: diâmetro de fruto; X10: comprimento de fruto; X11: peso fresco de parte aérea; X12: cobertura de vaso; X13: comprimento de folha; X14: largura de folha; X15: número de ramificações laterais; X16: dias após o transplântio e X17: comprimento de ramificações laterais.

A fim de estudar o comportamento de cada uma das pimentas biquinho vermelha e bode amarela com relação aos componentes principais retidos na ACP, foram construídos os Biplots das variedades com relação aos componentes $Z_1 \times Z_2$ e $Z_3 \times Z_4$ (Figuras 1 e 2).

Analisando o Biplot apresentado na Figura 1, observa-se que a pimenta biquinho vermelha possui escores altos para o primeiro componente principal, enquanto a pimenta bode amarela mostrou baixos escores para esse componente. Em tese, esse comportamento indica que a pimenta biquinho vermelha apresentou valores maiores para as variáveis número de frutos, diâmetro de fruto, comprimento de fruto, peso fresco de parte aérea, cobertura de vaso, comprimento de folha e largura de folha, enquanto a pimenta bode amarela se caracterizou por valores maiores de altura da planta, altura da primeira bifurcação, compacidade de copa, número de folhas e dias após o transplântio para florescimento. Com relação ao segundo componente, entretanto, não foi observada diferença significativa entre as pimentas, o que indica que não houve diferença entre as variedades

com relação as variáveis agrônômicas diâmetro transversal da copa, número de ramificações laterais e comprimento de ramificações laterais.

Figure 1: Biplot $Z_1 \times Z_2$ of the individuals in the trial with 95% confidence ellipses for red biquinho peppers (BV) and yellow bode peppers (BA).

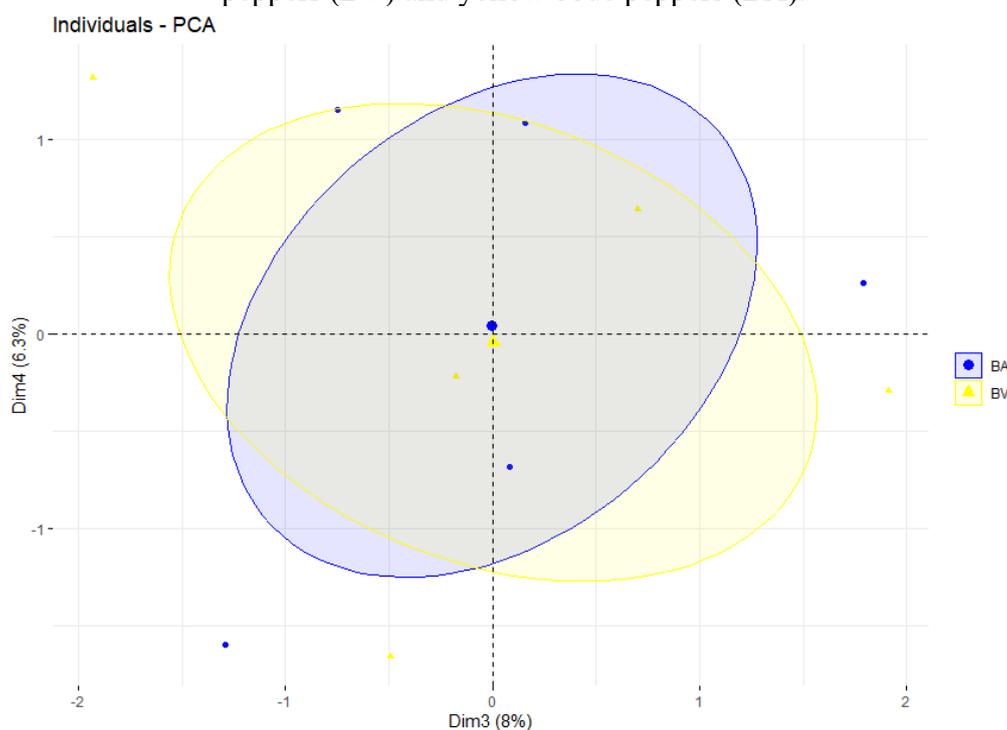


Source: from the authors (2024).

Por sua vez, ao analisar o Biplot apresentado na Figura 2, nota-se que não houve diferença entre as variedades de pimenta com relação aos componentes principais três e quatro, o que indica que as pimentas biquinho vermelha e bode amarela não se diferenciaram no que diz respeito às variáveis índice SPAD e diâmetro longitudinal da copa.

De modo geral, a ACP mostrou-se eficiente na caracterização e comparação das pimentas biquinho vermelha e bode amarela quanto às variáveis agrônômicas consideradas. De forma similar, Costa *et al.* (2020) observaram que a utilização de técnicas multivariadas, dentre elas a ACP, foi eficiente na caracterização morfoagronômica e estimativa da divergência genética entre os genótipos do gênero *Capsicum* avaliados em seu estudo. Por sua vez, Nascimento *et al.* (2021) também notaram que métodos multivariados, tal qual a ACP, foram eficientes na estimação da variabilidade genética de genótipos de pimenta, tendo por base descritores morfoagronômicos.

Figure 2: Biplot $Z_3 \times Z_4$ of the individuals in the trial with 95% confidence ellipses for red biquinho peppers (BV) and yellow bode peppers (BA).



Source: from the authors (2024).

Conclusões

A técnica multivariada da ACP mostrou-se eficiente na caracterização e na comparação das variedades de pimenta biquinho vermelha e bode amarela. Enquanto a pimenta biquinho vermelha destacou-se com relação às características de seus frutos e as dimensões de suas folhas, a pimenta bode amarela destacou-se com relação ao tamanho e compacidade da planta, entretanto, levou mais tempo para o seu florescimento. Espera-se que os resultados obtidos com a pesquisa auxiliem os programas de melhoramento de pimentas no desenvolvimento de novas cultivares.

Referências

COSTA, G. N.; SILVA, B. M. P.; LOPES, A. C. de A; CARVALHO, L. C. B; GOMES, R. L. F. Selection of pepper accessions with ornamental potential. *Ver. Caatinga*, v. 32, p. 566-574, 2019.

COSTA, L. S.; SILVA, L. S. N.; SILVA FILHA, C. M. R.; SANTOS, J. F. F.; SILVA, R. N. O. Caracterização de genótipos de *Capsicum* spp. por técnicas multivariadas no sul do Piauí. *Brazilian Journal of Development*, v. 6, n. 12, p. 97371-97385, 2020.

COSTA, L. V; BENTES, J. L. S; LOPES, M. T. G; ALVES, S. R. M; JÚNIOR, J. M. V. Caracterização de acessos de pimentas do Amazonas. *Horticultura Brasileira*, v. 33, p. 290-298, 2015.

- FINGER, F. L.; PEREIRA, G. M. Physiology and postharvest of pepper fruits. In: RÊGO, E. R.; RÊGO, M. M.; FINGER, F. L. *Production, and breeding of chilli pepper (Capsicum spp)*. New York: Springer, 2016, p. 27-40.
- HONGYU, K.; SANDANIELO, V. L. M.; OLIVEIRA JUNIOR, G. J. de. Análise de componentes principais: resumo teórico, aplicação e interpretação. *Engineering and Science*, v. 5, p. 83-90, 2016.
- HOTELLING, H. Analysis of a complex of statistical variables into principal components. *The Journal Educational Psychology*, v. 24, p. 498-520, 1933.
- HOTELLING, H. Simplified calculation of principal components. *Psychometrika*, v. 1, p. 27-35, 1936.
- JOHNSON, R. A.; WICHERN, D. W. *Applied multivariate statistical analysis*. Madison: Prentice Hall International, 1998. 816p.
- KAISER, H. F. The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, v. 23, p. 187-200, 1958.
- MANLY, B. F. J. *Multivariate statistical methods*. New York: Chapman and Hall, 1986. 159p.
- NASCIMENTO, M. F.; ARAÚJO, F. F.; SANTOS, R. M. C.; FINGER, F. L.; BRUCKNER, C. H. Genetic diversity in Capsicum and Solanum genotypes based on morphoagronomic characters. *Revista Brasileira de Ciências Agrárias*, v. 16, n. 4, p. e1169, 2021.
- PEARSON, K. On lines and planes of closest fit to systems of points in space. *Philosophical Magazine*, v. 2, p. 559-572, 1901.
- R CORE TEAM. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2023. Disponível em: <www.R-project.org/>.
- REIFSCHNEIDER, F. J. B. *Capsicum: pimentas e pimentões no Brasil*. Brasília, Embrapa Hortaliças, 103p., 2000.