RONALDO SILVA GOMES

CARACTERIZAÇÃO DE GERMOPLASMA DE *Cucurbita moschata* D. VISANDO O MELHORAMENTO GENÉTICO DE ASPECTOS AGROMORFOLÓGICOS E DO ÓLEO DE SEMENTES

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Fitotecnia, para obtenção do título de *Doctor Scientiae*.

Orientador: Derly José Henriques da Silva

VIÇOSA - MINAS GERAIS
2021

Ficha catalográfica elaborada pela Biblioteca Central da Universidade Federal de Viçosa - Campus Viçosa

T
G633c
2021

Gomes, Ronaldo Silva, 1991-
Caracterização de germoplasma de C*ucurbita moschata* D. visando o melhoramento genético de aspectos agromorfológicos e do óleo de sementes / Ronaldo Silva Gomes. – Viçosa, MG, 2021.
1 tese eletrônica (110 f.): il. (algumas color.).

Texto em inglês.
Inclui apêndice.
Orientador: Derly José Henriques da Silva.
Tese (doutorado) - Universidade Federal de Viçosa.
Inclui bibliografia.
DOI: https://doi.org/10.47328/ufvbbt.2021.058
Modo de acesso: World Wide Web.

1. *Cucurbita moschata* - Melhoramento genético. 2. Ácidos graxos essenciais. 3. Carotenoides. 4. Correlação (Estatística). I. Universidade Federal de Viçosa. Departamento de Agronomia. Programa de Pós-Graduação em Fitotecnia. II. Título.

CDD 22. ed. 635.622

Bibliotecário(a) responsável: Alice Regina Pinto CRB6 2523

RONALDO SILVA GOMES

CARACTERIZAÇÃO DE GERMOPLASMA DE *Cucurbita moschata* D. VISANDO O MELHORAMENTO GENÉTICO DE ASPECTOS AGROMORFOLÓGICOS E DO ÓLEO DE SEMENTES

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Fitotecnia, para obtenção do título de *Doctor Scientiae*.

APROVADA: 06 de agosto de 2021.

Assentimento:

Ronaldo Silva Gomes

Ronaldo Silva Gomes
Autor

Derly José Henriques da Silva
Orientador

À minha querida esposa, Raimunda Alves, com muito amor, dedico.

AGRADECIMENTOS

A Deus pelo dom da vida e pelo cuidado infalível.

À minha família; especialmente à minha esposa pelo companheirismo, amor e pela sabedoria compartilhada; aos meus pais Riba e Antônia e à minha irmã Railda, pelo incentivo e apoio durante todo esse percurso.

À minha primeira orientadora, professora Maria da Cruz, pelo incentivo, excelente orientação e altruísmo.

À Universidade Federal de Viçosa e ao Programa de Pós- Graduação em Fitotecnia pela oportunidade de realização do curso de doutorado.

Ao meu estimado orientador, professor Derly, pelo acolhimento nesta Universidade, pela sábia orientação, exemplo e ensinamentos transmitidos.

Aos meus colegas Cleverson, Ronaldo e Rebeca pelo grande companheirismo e amizade.

Aos funcionários da Horta Velha, Wilson, Felicinho, Zin, Nilsinho, Marcos, Isaias, Zé Nilson e Cocó, pela ajuda na condução do experimento de campo.

À primeira igreja batista de viçosa, na pessoa do Pr. Sergio, pelo acolhimento tão aconchegante nesta cidade.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico- CNPq, pela conceção da bolsa de doutorado.

BIOGRAFIA

RONALDO SILVA GOMES, filho de Antônio Silva Gomes e Maria Antônia Silva Gomes, nasceu em Pio XII, Maranhão, Brasil, aos 16 de setembro de 1991.

Iniciou seus estudos básicos em Pio XII, no ano de 2000, na Unidade Escolar Governador Luiz Rocha e concluiu no ano de 2009, no Centro de Ensino Jansén Veloso, sempre na rede pública.

Em março de 2010 iniciou a graduação no curso de Agronomia na Universidade Federal do Maranhão, Campus de Chapadinha. Trabalhou durante toda a sua graduação como bolsista, sob a orientação da professora Maria da Cruz. Participou de mobilidade acadêmica internacional, permanecendo por mais de um ano em graduação sanduiche na University of Wisconsin-River Falls, WI, Estados Unidos. Nesse período foi estagiário no Programa de Melhoramento de Milho para a Região Oeste dos Estados Unidos, conduzido pela Iowa State University

Em agosto de 2015 iniciou o mestrado no Programa de Pós-Graduação em Genética e Melhoramento na Universidade Federal de Viçosa (UFV), em Viçosa-MG, Brasil, onde obteve o título de mestre em Julho de 2017, com a dissertação intitulada "Variabilidade morfoagronômica e nutricional de acessos de *Cucurbita moschata* Duch. do Banco de Germoplasma de Hortaliças - UFV". Em agosto de 2017, iniciou o doutorado no Programa de Pós-Graduação em Fitotecnia, também na UFV, submetendo-se à defesa de tese em 06 de agosto de 2021 com a tese intitulada "Caracterização de germoplasma de *Cucurbita moschata* D. visando o melhoramento genético de aspectos agromorfológicos e do óleo de sementes". Nesse mesmo período já havia sido selecionado para pós-doutorado júnior (PDJ/CNPq), para desenvolver o projeto de pesquisa intitulado "Aplicação de redes neurais artificiais visando à otimização do manejo e uso de germoplasma vegetal conservado em bancos".

"*Imagine sua vida como se fosse um banquete onde você deve se comportar com cortesia. Quando os pratos lhe forem passados, estenda a mão e sirva-se de uma porção moderada. Se algum prato não lhe for apresentado, aproveite o que já está no seu. Ou se o prato ainda não chegou a você, espere pacientemente a sua vez*".

(Epicteto)

RESUMO

GOMES, Ronaldo Silva, D.Sc., Universidade Federal de Viçosa, agosto de 2021. **Caracterização de germoplasma de *Cucurbita moschata* D. visando o melhoramento genético de aspectos agromorfológicos e do óleo de sementes**. Orientador: Derly José Henriques da Silva.

A *Cucurbita moschata* D. é uma hortaliça de elevada importância socioeconômica em decorrência da qualidade nutricional de seus frutos, sementes e óleo de sementes. Diante disto, o objetivo geral deste trabalho foi avaliar 91 acessos de *C. moschata* mantidos BGH-UFV, visando à identificação de acessos promissores para aspectos agromorfológicos, produtividade do óleo de sementes e para o perfil de ácidos graxos do óleo. Avaliações agromorfológicas foram realizadas em experimento em campo, conduzido entre janeiro e julho de 2016, enquanto o perfil de ácidos graxos foi analisado a partir de cromatografia gasosa. Os acessos expressaram elevada variabilidade para as características número de graus-dias acumulados para o florescimento, teor total de carotenoides na polpa de frutos, produtividade de frutos, massa de sementes por fruto e produtividade de sementes. A análise *per se* identificou os acessos BGH-6749, BGH-5639 e BGH-219 como aqueles com floração mais precoce. Os acessos BGH-5455A e BGH-5598A apresentaram os maiores teores de carotenóides, com médias superiores a 170,00 µg g^{-1} de massa fresca. O BGH-4610, BGH-5485A, BGH-6590, BGH-5556A, BGH-5472A e BGH-5544A foram os acessos mais promissores em termos de produtividade do óleo de sementes, com média (µ+g) de aproximadamente 0,20 t ha^{-1}. Os acessos mais promissores para maior teor de ácido oleico foram o BGH-5456A, BGH-3333A, BGH-5361A, BGH-5472A, BGH-5544A, BGH-5453A e BGH-1749, com média (µ+g) de aproximadamente 30%. Conclui-se que os acessos de *C. moschata* avaliados no presente estudo possuem elevado potencial para uso no melhoramento genético dessa hortaliça visando obtenção de florescimento precoce e incremento do teor total de carotenoides da polpa de frutos. Esses acessos são promissores para uso no melhoramento visando maior produtividade do óleo de sementes e maior teor de ácido oleico.

Palavras-chave: Ácidos graxos. Análise de correlação. Carotenoides. *Cucurbita moschata*. Melhoramento.

ABSTRACT

GOMES, Ronaldo Silva, D.Sc., Universidade Federal de Viçosa, August, 2021. **Characterization of *Cucurbita moschata* D. germplasm aiming the genetic breeding of agromorphological aspects and seed oil**. Adviser: Derly José Henriques da Silva.

Cucurbita moschata D. is a vegetable of high socioeconomic importance due to the nutritional quality of its fruits, seeds, and seed oil. Therefore, the general objective of this work was to assess 91 accessions of *C. moschata* maintained in BGH-UFV, aiming to identify promising accessions for agromorphological aspects, and for productivity of seed oil and fatty acid profile. Agromorphological assessments were carried out from a field experiment, carried out between January and July 2016, while the fatty acid profile was analyzed from gas chromatography. The accessions expressed high variability for the characteristics number of accumulated degree-days for flowering, total content of fruit pulp carotenoids, fruit yield, mass of seeds per fruit and seed yield. *Per se* analysis identified the accessions BGH-6749, BGH-5639, and BGH-219 as those with earlier flowering. The accessions BGH-5455A and BGH-5598A had the highest levels of carotenoids, with averages above 170.00 $\mu g\ g^{-1}$ of fresh mass. BGH-4610, BGH-5485A, BGH-6590, BGH-5556A, BGH-5472A, and BGH-5544A were the most promising accessions in terms of seed oil productivity, with an average (μ+g) of approximately 0.20 $t\ ha^{-1}$. The most promising accessions for higher oleic acid content were BGH-5456A, BGH-3333A, BGH-5361A, BGH-5472A, BGH-5544A, BGH-5453A, and BGH-1749, with an average (μ+g) of approximately 30 %. It is concluded that the accessions of *C. moschata* assessed in this study have a high potential for use in the genetic breeding of this vegetable aiming at obtaining early flowering and increase of total content of fruit pulp carotenoids. These accessions are promising for use in the breeding of *C. moschata* aiming at higher seed oil productivity and higher oleic acid content.

Keywords: Carotenoids. Correlation analysis. *Cucurbita moschata. B*reeding. Fatty acids.

SUMÁRIO

INTRODUÇÃO GERAL

A *Cucurbita moschata* D. é cultivada mundialmente, consistindo em uma hortaliça de elevada importância socioeconômica em decorrência da qualidade nutricional de seus frutos, sementes e óleo de sementes. A princípio, seu cultivo era voltado principalmente para a produção de frutos, os quais constituem fontes de uma série componentes essenciais à nutrição e saúde humana. Corroborando isto, estudos demonstram que a polpa dos frutos da *C. moschata* possui elevados teores de carotenoides como α e β-caroteno (CARVALHO *et al.* 2012; NAKKANONG *et al.*, 2012; PRIORI *et al.*, 2017), os principais precursores da vitamina A, além de expressarem pronunciada atividade antioxidante. Adicionalmente, seus frutos são excelentes fontes de minerais como K, Ca, P, Mg e Cu (NAGAR *et al.*, 2018; PRIORI *et al.*, 2017).

A produção do óleo de sementes da *C. moschata* para alimentação humana configura excelente oportunidade na produção dessa hortaliça. O óleo de suas sementes é constituído por aproximadamente 75% de ácidos graxos insaturados, com elevado teor de ácido monoinsaturado como o oleico (JARRET *et al.*, 2013; SOBREIRA, 2013; VERONEZI e JORGE 2015). Com isto, o óleo das sementes da *C. moschata* configura opção de consumo mais saudável, constituindo, por exemplo, excelente substituto de fontes lipídicas prejudiciais à saúde humana como aquelas com teores elevados de ácidos graxos saturados.

Estudos têm destacado que as sementes e o óleo de sementes da *C. moschata* contêm elevados teores de componentes antioxidantes como a vitamina E e carotenoides (VERONEZI e JORGE 2012; DASH *et al.*, 2017), os quais, além de serem benéficos à saúde humana, atuam na produção do óleo contra processos oxidativos. Além dos aspectos nutricionais, a *C. moschata* expressa produtividade de semente de até 0.58 t. ha^{-1} e teor do óleo de sementes de até 49% (JARRET *et al.*, 2013; PATEL, 2013; GOMES *et al.*, 2020), corroborando seu elevado potencial para produção de óleo de sementes. Soma-se a isto o fato de que a *C. moschata* é cultivada mundialmente, perfazendo, juntamente com outras cucurbitas como a *C. pepo* e *C. máxima*, área cultivada e produção próximas de 2 milhões de hectares e 27,6 milhões de toneladas, respectivamente (FAO, 2020). No Brasil, a área cultivada e a produção da *C. moschata*, juntamente com a *C. maxima* e *C. pepo* é de aproximadamente 78,67 mil hectares e 417,83 toneladas, respectivamente, confirmando a importância socioeconômica dessa hortaliça.

Visando à conservação e ao uso da diversidade genética de hortaliças como a *C. moschata*, o Banco de Germoplasma da Universidade Federal de Viçosa (BGH-UFV), mantém uma coleção com aproximadamente 350 acessos de *C. moschata*, consistindo em uma das maiores coleções dessa espécie no Brasil (FONSECA *et al.*, 2015). A avaliação desse germoplasma tem levado à identificação e uso de acessos com resistência a importantes fito patógenos da cultura e de acessos com elevado potencial para uso visando o melhoramento da qualidade nutricional da polpa de frutos e do óleo de sementes (MOURA, *et al.*, 2005; LIMA NETO, 2013; SOBREIRA, 2013; OLIVEIRA *et al.*, 2021). Isto corrobora a importância da continuação dos estudos envolvendo a avaliação e uso do germoplasma de *C. moschata* mantido no BGH-UFV.

Assim, o objetivo geral deste trabalho foi avaliar o germoplasma de *C. moschata* mantido no BGH-UFV, visando à identificação de acessos promissores para aspectos agromorfológicos, produtividade do óleo de sementes e para o perfil de ácidos graxos do óleo. Os objetivos específicos foram: a) avaliar agromorfologicamente 91 acessos de *C. moschata* mantido no BGH-UFV e as correlações genéticas dessas características; b) analisar a variabilidade agromorfológica dos acessos, visando identificar genótipos de floração precoce, com altos teores totais de carotenóides na polpa dos frutos, e genótipos com alto potencial de produtividade de sementes e óleo de sementes; c) avaliar a produtividade do óleo de sementes e os perfis de ácidos graxos, assim como as correlações dessas características de 91 acessos de

C. moschata; e d) examinar a variabilidade dos acessos quanto aos aspectos de sementes e do perfil do óleo, visando identificar os acessos com alta produtividade de óleo de sementes e alto teor de ácido oleico.

REFERÊNCIAS

CARVALHO, L.M.J.; GOMES, P.B.; GODOY, R.L.O.; PACHECO, S.; MONTE, P.H.F.; CARVALHO, J.L.V.; NUTTI, M.R.; NEVES, A.C.L.; VIEIRA, A.C.R.A.; RAMOS, S.R.R. Total carotenoid content, α-carotene and β-carotene, of landrace pumpkins (*Cucurbita moschata* Duch): A preliminary study. **Food Research International**. v. 47, p. 337–340. 2012. Doi: 10.1016/j.foodres.2011.07.040.

DASH, P.; GHOSH, G. Proteolytic and antioxidant activity of protein fractions of seeds of *Cucurbita moschata*. **Food Bioscience**. v.18, p. 1-8, 2017. Doi: 10.1016/j.fbio.2016.12.004.

FAO- Food and Agriculture Organization of the United Nations. **FAO STATE**. Roma: FAO-Food and Agriculture Organization of the United Nations, 2020. Disponível em: http://www.fao.org/faostat/en/#data/QC, acesso em: 14/05/2021.

FONSECA, M.A.; QUEIROZ, M.A.; LOPES, J.F.; ASSIS, J.G.A.; ROMÃO, R.L.; RAMOS, S.R.R.; DO CARMO, C.A.S.; MOURA, M.C.; DIAS, R.C.; DA SILVEIRA, L.M.; BARBIERI, R.; SILVA, D.J.H.; NORONHA, S. Geographical distribution and conservation of *Cucurbita* in Brazil. **Magistra**. v. 27, p. 432-442, 2015.

GOMES, R.S.; MACHADO JÚNIOR, R.; Almeida, C.F.; CHAGAS, R.R.; OLIVEIRA, R.L.; DELAZARI, F.T.; SILVA, D.J.H. Brazilian germplasm of winter squash (*Cucurbita moschata* D.) displays vast genetic variability, allowing identification of promising genotypes for agro-morphological traits. **Plos One**. v. 15, p. 1-16, 2020. Doi: 10.1371/journal.pone.0230546.

JARRET, R.L.; LEVY, I.; POTTER, T.L.; CERMAK, S.C.; MERRICK, L.C. Seed oil content and fatty acid composition in a genebank collection of *Cucurbita moschata* Duchesne and *C. argyrosperma* C. Huber. **Genetic Resources and Crop Evolution**. v. 9, p. 701-716, 2013. Doi: 10.1017/S1479262112000512.

LIMA NETO, I.S. **Pré- melhoramento de abobra (*Cucurbita moschata*) visando a biofortificacao em carotenoides**. 2013. Tese (Doutorado em Fitotecnia) - Universidade Federal de Viçosa- UFV, Viçosa, 2013.

MOURA, M.C.C.L; ZERBINI, F.M.; SILVA, D.J.H. QUEIROZ, M.A. Reação de acessos de *Cucurbita* sp. ao *Zucchini yellow mosaic virus* (ZYMV). **Horticultura Brasileira**. v.23, p.206-210, 2005.

NAGAR, A.; SUREJA, A.K.; KAR, A.; BHARDWAJ, R.; KRISHNAN, S.G.; MUNSHI, A. Profiling of mineral nutrients and variability study in pumpkin (*Cucurbita moschata*) genotypes. **Agricultural Research**. v. 7, p. 225-231, 2018. Doi.org/10.1007/s40003-018-0329-3.

NAKKANONG, K.; YANG, J.H.; ZHANG, M.F. Carotenoid accumulation and carotenogenic gene expression during fruit development in novel interspecific inbred squash lines and their parents. **Journal of Agricultural Food Chemistry**. v. 13, p. 1-12, 2012. Doi:10.1021/jf3007135.

OLIVEIRA, R.L.; GOMES, R.S.; ALMEIDA, C.F.; MACHADO JÚNIOR, R.; ROCHA, J.R.A.S.C.; SILVA, D.J.H.; CARNEIRO, P.C.S. Multitrait selection of pumpkin genotypes aimed at reducing the growth habit and improving seed production. **Crop Science.** v. 61, p. 1-10, 2021.

PATEL, S. Pumpkin (*Cucurbita* sp.) seeds as nutraceutic: A review on status quo and scopes. **Mediterranean Journal of Nutrition and Metabolism**. v. 4, p. 51-62, 2013. Doi: 10.1007/s12349-013-0131-5.

PRIORI, D.; VALDUGA, E.; VILLELA, J.C. B. Characterization of bioactive compounds, antioxidant activity and minerals in landraces of pumpkin (*Cucurbita moschata*) cultivated in Southern Brazil**. Food Science and Technology**. v. 37, p. 33-40, 2017. Doi: 10.1590/1678-457x.05016.

SOBREIRA, F.M. **Divergência genética entre acessos de abóbora para estabelecimento de coleção nuclear e pré-melhoramento para óleo funcional**. 2013. Tese (Doutorado em Fitotecnia) - Universidade Federal de Viçosa- UFV, Viçosa, 2013.

VERONEZI, C.M.; JORGE, N. Chemical characterization of the lipid fractions of pumpkin seeds. **Nutrition and Food Science**. v. 45, p. 161-173, 2015. Doi: 10.1108/NFS-01-2014-0003.

VERONEZI, C.; JORGE, N. Bioactive compounds in lipid fractions of pumpkin (*Cucurbita* sp) seeds for use in food. **Journal of Food Science**. v. 77, p. 1-5, 2012. Doi: 10.1111/j.1750-3841.2012.02736.x.

ARTIGO I

BRAZILIAN GERMPLASM OF WINTER SQUASH (*Cucurbita moschata* D.) DISPLAYS VAST GENETIC VARIABILITY, ALLOWING IDENTIFICATION OF PROMISING GENOTYPES FOR AGRO-MORPHOLOGICAL TRAITS

ARTIGO I - BRAZILIAN GERMPLASM OF WINTER SQUASH (*Cucurbita moschata* D.) DISPLAYS VAST GENETIC VARIABILITY, ALLOWING IDENTIFICATION OF PROMISING GENOTYPES FOR AGRO-MORPHOLOGICAL TRAITS

Artigo publicado na Revista PLoS ONE 15(6): e0230546. https://doi.org/10.1371/journal.pone.0230546

Ronaldo Silva Gomes[1]*, Ronaldo Machado Júnior[1], Cleverson Freitas de Almeida[1], Rafael Ravaneli Chagas[1], Rebeca Lourenço de Oliveira[1], Fabio Teixeira Delazari[2], Derly José Henriques da Silva[1]

[1]Agronomy Department, Graduation Program on Genetics and Breeding, Federal University of Viçosa-UFV, Viçosa, MG, Brazil; [2]National Service of Rural Learning- SENAR, Campo Grande, MS, Brazil

*Corresponding author: ronaldo.s.gomes@ufv.br

ABSTRACT

Winter squash fruits (*Cucurbita moschata* D.) are among the best sources of vitamin A precursors and constitute sources of bioactive components such as phenolic compounds and flavonoids. Approximately 70% of *C. moschata* seed oil is made up of unsaturated fatty acids, with high levels of monounsaturated fatty acids and components such as vitamin E and carotenoids, which represent a promising nutritional aspect in the production of this vegetable. *C. moschata* germplasm expresses high genetic variability, especially in Brazil. We assessed 91 *C. moschata* accessions, from different regions of Brazil, and maintained at the Federal University of Viçosa (UFV) Vegetable Germplasm Bank, to identify early-flowering accessions with high levels of carotenoids in the fruit pulp and high yields of seed and seed oil. Results showed that the accessions have high variability in the number and mass of seeds per fruit, number of accumulated degree-days for flowering, total carotenoid content, and fruit productivity, which allowed selection for considerable gains in these characteristics. Analysis of the correlation between these characteristics provided information that will assist in selection to improve this crop. Cluster analysis resulted in the formation of 16 groups, confirming the variability of the accessions. *Per se* analysis identified accessions BGH-6749, BGH-5639, and BGH-219 as those with the earliest flowering. Accessions BGH-5455A and BGH-5598A had the highest carotenoid content, with averages greater than 170.00 $\mu g\ g^{-1}$ of fresh mass. With a productivity of 0.13 t ha^{-1}, accessions BGH-5485A, BGH-4610A, and BGH-5472A were the most promising for seed oil production. These last two accessions corresponded to those with higher seed productivity, averaging 0.58 and 0.54 t ha^{-1}, respectively. This study confirms the high potential of this germplasm for use in breeding for promotion of earlier flowering and increase in total content of fruit pulp carotenoids and in seed and seed oil productivity.

Key words: agro-morphological, bioactive compounds, carotenoids, correlation, clustering, Cucurbita moschata, genotypic variance, seed oil.

1 INTRODUCTION

Winter squash (*Cucurbita moschata* D.) is one of the vegetables of greater socio-economic importance in the Cucurbita genus, largely due to the high nutritional value of its fruits and seeds. The pulp of its fruits constitutes an important source of carotenoids such as β-carotene, the precursor of greater pro-vitamin A activity [1, 2, 3]. The pulp is also an excellent source of minerals such as K, Ca, P, Mg, and Cu [4, 5]. The socio-economic importance of *C. moschata* is also linked to the high volume and value of its production. It is estimated that, together with other cucurbits such as *C. pepo* and *C. maxima*, the cultivated area and the world production of this vegetable in 2017 were approximately 2 million hectares and 25 million tons, respectively [6], most of it concentrated in China and India. In Brazil, this crop is of high socio-economic importance, with a cultivated area of approximately 90 thousand hectares, an estimated production of more than 40 thousand tons / year, and an annual production value of around R$ 1.5 million [7].

C. moschata brings together characteristics that are fundamental to biofortification programmes, such as high productivity and profitability potentials, high efficiency in reducing micronutrient deficiencies in humans, and good acceptability by producers and consumers in regions where it is grown [8]. This has caused this vegetable to be chosen as a strategic crop for breeding programmes promoting biofortification, such as the Brazilian Biofortification Program (BioFORT), led by the Brazilian Agricultural Research Corporation (Embrapa), which aims for biofortification in vitamin A precursors [9].

The crop also has potential for the production of edible seed oil. Its seed oil comprises about 70% unsaturated fatty acids, and it has a high content of monounsaturated fatty acids [10, 11, 12], so it is a good substitute for other lipid sources that have higher contents of saturated fatty acids. The oil is also rich in bioactive components such as vitamin E and carotenoids [13], which are important antioxidants in the human diet, in addition to protecting the oil itself against oxidative processes. In addition, this species is commonly cultivated in low-technology systems [14], making it fundamental to ensuring healthier diets and promoting food security in the regions where it is grown, particularly in less-developed regions and in the context of family-based farming.

Associated with its socio-economic importance, *C. moschata* germplasm commonly expresses high genetic variability in all regions where it occurs [15, 16, 17, 18], especially in Brazil [19, 20, 21]. Archaeological evidence indicates that this species was present in Latin America prior to colonization, and appears to have already been an important component in the diet of the native peoples living there [22, 23, 24]. Currently, the variability of this vegetable in

Brazil is closely tied to the human populations involved in its cultivation, who are predominantly family-based farmers.

The selection practised over time by these populations, associated with the exchange of seeds between them, and the natural occurrence of hybridisation in the germplasm of this species has contributed to its increased variability. The high variability in agronomic, nutritional and bioactive characteristics displayed by *C. moschata* and the intercrossability of *Cucurbita* species have enabled these characteristics to be transferred from *C. moschata* to other species of this genus [25, 26, 27, 28]. This is of strategic importance and may aid the worldwide cultivation of species of the *Cucurbita* genus.

The usefulness of plant germplasm conserved in banks depends on the amount and quality of information associated with it, such as genetic and phenotypic data, which highlights the importance of its proper evaluation. On the other hand, the high volume of germplasm and limitations in resources and area available for the establishment of field trials commonly make its assessment difficult. In view of this, the FAO's Second Global Action Plan for Plant Genetic Resources for Food and Agriculture sets out guidelines that provide greater efficiency in the conservation and use of plant germplasm [29].

This is essential information for the management and use of germplasm [30, 31, 32, 33]. Evaluation of the germplasm maintained in banks makes it possible to estimate the magnitude of the genetic and statistical parameters of characteristics of interest, which can provide information on the nature of variability observed for these traits, in addition to elucidating which characteristics or groups of characteristics most contribute to germplasm variability. From this assessment, it is also possible to assess the association between the characteristics evaluated. Together, the information obtained from these assessments is essential for optimising the use and management of plant germplasm.

The UFV Vegetable Germplasm Bank (BGH-UFV) maintains more than 350 accessions of *C. moschata*, constituting one of the largest collections of this species in Brazil [34]. This bank continually carries out work on the characterisation and evaluation of this germplasm [35], which has allowed the sources of resistance to important phyto-pathogenic agents to be identified [36], and its production [21] and nutritional aspects of fruits and seed oil to be improved [10, 37]. The potential of this germplasm as a source of genes for the improvement of this crop, along with the possibility of elucidating the genetic mechanisms linked to important production parameters, justifies the continuation of studies on its assessment and use.

This study therefore aimed to: a) agro-morphologically assess some of the *C. moschata* accessions maintained by BGH-UFV, b) analyse the genetic relationships of these agro-morphological characteristics, and c) analyse their agro-morphological variability, with a view to identifying earlier-flowering genotypes, genotypes with high total levels of carotenoids in the fruit pulp, and those with high potential for seed and seed oil productivity.

2 MATERIALS AND METHODS

2.1 Origin of germplasm and preparation of seedlings

In this study, we assessed 95 genotypes, comprising 91 accessions of *C. moschata* maintained in the BGH-UFV, and four control genotypes (Figure 1). The controls comprised the commercial hybrids Tetsukabuto and Jabras, and the cultivars Jacarezinho and Maranhão, all widely cultivated and commercialised in Brazil. The accessions came from different regions of Brazil [35], and consisted, for the most part, of landraces collected from family-based farmers, who commonly select the genotypes and conserve their seeds.

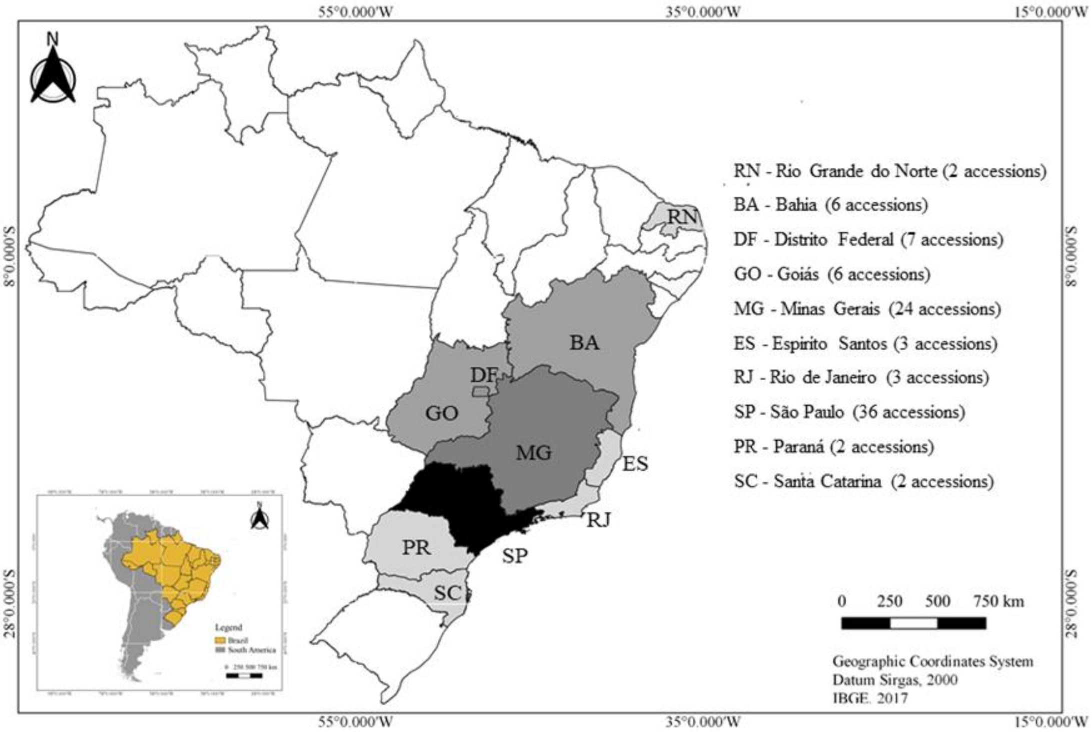


Figure 1. Brazilian map showing the states of origin of the *C. moschata* accessions assessed in this study.

Seedlings were produced in a 72-cell expanded-polystyrene tray containing commercial substrate. Seedling transplantation and cultural treatments were carried out according to local recommendations for the cultivation of pumpkins [38].

2.2 Experiment location and experimental design

The experiment was carried out from January to July 2016, at “Horta Velha” (200° 45'14" S, 420° 52'53" W and 648.74 m alt.), an experimental unit of the Agronomy Department of the Federal University of Viçosa, Viçosa-MG, Brazil.

The experiment was arranged in a Federer's augmented block design [39], with five replications for each control. The four controls, also called common treatments, were randomly distributed in each of the five blocks, and the 91 accessions, called regular treatments, were randomly assigned to all blocks. A spacing of 3x3 m between plants and rows was adopted, which resulted in a stand of 1,111 plants ha^{-1}. Each plot consisted of five plants, and all assessments were carried out from three central plants. The evaluations of fruit and seed characteristics were carried out on three fruits per plant.

2.3 Assessments of agro-morphological aspects, total carotenoid content of fruit pulp, and seed and seed oil yields

For the assessment involving multi-categorical characteristics, we adopted the morphological descriptors suggested by Bioversity International and the European Cooperative Programme for Plant Genetic Resources (ECPPGR), plus some additional descriptors. These descriptors comprised agro-morphological characteristics of plants, fruits, and seeds (Supplementary Table 1). Assessment was also based on agronomic characteristics, the total content of fruit pulp carotenoids, productivity of seeds, and seed oil productivity (Table 1).

Table 1. Descriptors involving agronomic aspects of plants, fruits and seeds, used in the assessment of the *C. moschata* germplasm maintained by BGH-UFV

Phase/organ	**Descriptors**
Reproductive phase	Accumulated degree-days for flowering (DDF).
Fruit	Number of fruits per plant (NFP), average mass of fruits (MF), productivity of fruits (PF), height of fruit (HF), diameter of fruit (DF), thickness of fruit peel (TFP), resistance of fruit peel to penetration (RFP), resistance of fruit pulp to penetration (RP), thickness of fruit pulp (PT), diameter of internal cavity of fruit (DIC), total content of fruit pulp carotenoids (TC), and the lutein content of fruit pulp (L).
Seed	Number of seeds per plant (NSF), mass of seeds per fruit (MSF), ratio of seed to fruit mass (MS/F), mass of one hundred seeds (MOH), productivity of seeds (PS), seed thickness (ST), seed length (SL), and seed width (SW).
Seed oil	Seed oil content (SOC) and seed oil productivity (SOP).

The estimates of the total carotenoid (TC) and lutein contents (L) of fruit pulp were based on colorimetric parameters. For this, the fruit pulp colour was characterised with the aid of a manual tri-stimulus colorimeter, Colour Reader CR-10 Konica Minolta, by parameters related to luminosity, and the contribution of red (a) and yellow (b). The fruit pulp was characterised from a fruit from each of the three central plants of the plot. This was carried out on pulp from four different parts of the fruit (part facing the sun, part facing the soil, part by the peduncle, and floral insertion part). The values of each parameter consisted of averages obtained from the pulp of fruits harvested from each of the plots' central plants. The estimates of TC were obtained using the equations proposed by [40], described below:

$$C = \sqrt{a^2 + b^2}$$

$$TC = 6.1226 + 1.7106 * a$$

$$L = -6.3743 + 0.2818 * C$$

Where:

C corresponds to the saturation or chroma of the fruit pulp;

a corresponds to the contribution of red to the colour of fruit pulp;

b corresponds to the contribution of yellow to the colour of fruit pulp;

TC corresponds to the total content of fruit pulp carotenoids ($\mu g\ g^{-1}$ of fresh pulp mass); and

L corresponds to the lutein content of fruit pulp ($\mu g\ g^{-1}$ of fresh pulp mass).

The seed oil was extracted by cold pressing, with the aid of a 30- ton-capacity press, with the necessary adaptations for pressing. For this, the seeds were previously dried in a forced-air-circulation oven for 72 hours, at 23°C. To standardise the process, 50 g seed samples were weighed from each accession and all samples were equally pressed for approximately 10 minutes.

2.4 Estimation of genotypic values, components of variance and genetic-statistical parameters

Phenotypic data were analysed using restricted maximum likelihood (REML) procedures and the best linear unbiased prediction (BLUP). These procedures were carried out with the aid of the R program, using the "lme4" package [41]. The estimates of variance components were obtained from the REML procedure, while the genotypic values of accessions (BLUPS) and controls (BLUES) were obtained from the BLUP procedure. All estimates were based on the following model:

$y=Wb+Xa+Zt+e$

in which:

y corresponds to the phenotypic data vector;

b corresponds to the vector comprising the effect of blocks, assumed to be random;

a corresponds to the vector comprising the effect of accessions, assumed to be random:

t corresponds to the vector comprising the effect of controls, assumed to be fixed: and

e corresponds to the error vector.

The letters W, X and Z correspond to the incidence matrices of parameters b, a, and t, respectively, with the data vector y.

The estimates of variance components comprised the phenotypic (σ^2_p), genotypic (σ^2_g), and residual (σ^2) variances, and the variance associated with the block effect (σ^2_b). The genetic-statistical parameters comprised the broad sense heritability (h^2), the selection accuracy (A), selection gain (SG), the phenotypic mean of the characteristics (μ), and the genotypic (CV_g %), phenotypic (CV_P %), and residual (CV_r %) coefficients of variance. These were obtained from the following estimators: $h^2 = 1- (Pev/\sigma^2_g)$, where Pev corresponds to the prediction of error variance [42];]; $A = \sqrt{1 - (Pev/\ \sigma_g^2\ )}$; $SG = h^2*DS$, where DS corresponds to the selection differential, estimated from the average of the top 15% most promising accessions: CV_g % = (σ^2_g/μ) x 100; CV_p % = (σ^2_p/μ) x 100; e CV_r % = (σ^2/μ) x 100.

2.5 Correlation analysis

This analysis was based on the matrix of genetic correlations, obtained from the following estimator:

$$rg = Cov\ (x,y)/\sqrt{\sigma_g^2(x)\ \sigma_g^2(y)}$$

in which;

Cov (x, y), corresponds to the genetic covariance between two variables X and Y, and σ^2_g (x) and σ^2_g (y) correspond to the genetic variances of variables X and Y, respectively.

The correlations were analysed using a procedure known as a correlation network, which allows all relationships between the variables under study to be analysed in relation to a specific function. This procedure also allows the direction and magnitude of the correlations to be distinguished. The direction is denoted by colours: dark green is used for the lines that connect positively-correlated variables, and red for the lines that connect negatively-correlated variables. The magnitude of the correlations is denoted by the thickness of the lines connecting the variables: the thicker the line, the greater the correlation. The significance of the correlations was analysed using Mantel's Z test at 1 and 5% probability. The correlation analysis was performed with the aid of the Genes program [43].

2.6 Analysis of variability and clustering

The analysis of variability was carried out using both quantitative and multi-categorical information. For quantitative data, the distance matrix between the genotypes was obtained from the BLUPS estimates in the case of accessions, and from the BLUES in the case of the controls; the genetic distances were obtained based on the negative average Euclidean distance, with data standardisation.

The matrix was obtained from negDistMat, a function of the APCluster package [44] implemented in the R program, version 3.5.1 [45]. The distances d (x; y) between the accession pairs, exemplified here as any two accessions x (x1, ..., xn) and y (y1, ..., yn), were estimated from the following equation:

$$d(x,y) = -(1/v)\sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$$

in which v corresponds to the number of quantitative descriptors evaluated.

The distance matrix for the qualitative data was obtained using the arithmetic complement of the simple coincidence index. The variability analysis was performed from a single distance matrix, obtained from the sum of the distance matrices of the quantitative and qualitative data. For the sum of matrices, they were standardised and each received an equal weight in the sum procedure. The variability analysis was performed using the procedure known as the Affinity propagation method [46]. The grouping was carried out from 100 independent rounds, aiming to assess the consistency of grouping.

The operation of Affinity initially involves the identification, in a set of components, of samples that will function as centres of this set. This method simultaneously considers all the set components as potential centres, i.e. as nodes in an interconnected network. Following the identification of potential centres, messages are transmitted between the set components along the network until a good set of centres and their corresponding groups emerge. The messages exchanged between the components in *Affinity* can be "responsiveness" *r* (*i, k)* and "availability *a* (*i, k*). This first case reflects the accumulated evidence of how appropriate point k is to serve as an example for point i, considering all other potential examples for this point. The "availability", in turn, reflects the accumulated evidence of how appropriate it would be for point i to choose point k as an exemplar, considering the other points for which point k can be an exemplar [46]. In the analysis of the present study, availability was initially established as zero.

A principal component analysis was implemented in order to identify the contribution of traits in the clustering of the genotypes. This analysis considered the data of quantitative and multi-categorical traits, according to the methodology of [47]; and was implemented using the FactoMineR package [48].

2.7 Identification of promising accession groups and per se identification of accessions

In order to facilitate the identification of promising groups of accessions for each characteristic, we carried out a grouping of means of the genotypic values corresponding to the groups obtained from the analysis of variability. This was based on Tocher's method of grouping means. The identification per se of the most promising accessions for each trait was carried out by ranking the respective genotypic effects, genetic gain and the new predicted average of the accessions, and the top 15% were considered the most promising accessions.

3. RESULTS

3.1 Variance components and genetic-statistical parameters of the agronomic aspects, total content of fruit pulp carotenoids, and the characteristics of seeds and seed oil

Estimates of the variance components and the genetic-statistical parameters are presented in Table 2. The estimates of genotypic variance were highest for number of seeds per fruit (NSF) and mass of seeds per fruit (MSF), decreasing to accumulated degree-days for flowering (DDF), and total content of fruit pulp carotenoids (TC). Among these variance estimates, only the genotypic variance of DDF was not significant. The estimates of variance associated with the block effect were low for all characteristics (Table 2).

For mass of seeds per fruit (MSF), number of seeds per fruit (NSF), total content of fruit pulp carotenoids (TC), and accumulated degree-days for flowering (DDF), most of the phenotypic variance was attributable to genotypic variance, with residual variance contributing less for most of the characteristics (Table 2).

As can also be seen in Table 2, most of the characteristics had high values for selection accuracy (*A*). Heritability estimates were 0.525, 0.495, and 0.774 for accumulated degree-days for accumulated-days for flowering (DDF), productivity of fruits (PF), and total content of fruit pulp carotenoids (TC), respectively, While productivity of seeds (PS) had a heritability of 0.481 and seed oil productivity (SOP) 0.291. Heritability was high (>0.50) for most of the characteristics, and very high for seed characteristics, such as mass of seeds per fruit (MSF), ratio of seed to fruit mass (MS/F), and number of seeds per fruit (NSF), and fruit characteristics, such as total content of fruit pulp carotenoids (TC) and average mass of fruit (MF), as shown in Table 2.

The high estimates of genotypic variance and heritability showed that considerable selection gain could be obtained for most of the characteristics (Table 2). For the number of accumulated degree-days for flowering (DDF), the gain was -92.947. It was also possible to obtain gains of 7.817 t ha^{-1} for productivity of fruits (PF) and 20.426 µg g^{-1} of fresh pulp mass for total content of fruit pulp carotenoids (TC), while the potential gains for productivity of seeds (PS) and seed oil productivity (SOP) were 0.187 and 0.072 t ha-1, respectively (Table 2).

The phenotypic range between accessions for accumulated degree-days for flowering (DDF) was 120.0 to 820.4 (average 606. 642) (Table 2). The range for productivity of fruits (PF) was 0.7 to 44.6 t ha^{-1} (average 12.946 t ha^{-1}), and that for total content of fruit pulp carotenoids (TC) was 43.4 to 187.2 µg g^{-1} of fresh pulp mass (average 65.763 µg g^{-1}), while that for productivity of seeds (PS) was 0.01 to 0.9 t ha^{-1} (average 0.269 t ha-1). The phenotypic range between accessions for seed oil productivity (SOP) was 0.004 to 0.40 t ha^{-1} (average 0.050 t ha^{-1}) (Table 2).

The greatest ranges between accessions for the coefficients of genotypic variation (CV_g%) were for mass of fruit (MF) and seed oil content (SOC), while for the coefficient of phenotypic variation (CV_P%), the greatest ranges between accessions were for seed oil productivity (SOP) and accumulated degree-days for flowering (DDF). The estimates of residual variation coefficient ranged from 7.502 to 71.582 for total content of fruit pulp carotenoids (TC) and SOP, respectively (Table 2).

Table 2. Estimates of variance components and genetic-statistical parameters of agronomic aspects, total content of fruit pulp carotenoids, and yields of seeds and seed oil

Vegetative trait

Traits	σ^2_p	σ^2_g	σ^2_r	σ^2_b	*A*	h^2	*SG*	*Range*	µ	CV$_g$%	CV$_P$ %	CV*r* %
DDF	10781.493	6385.892^{ns}	3909.203	486.397800	0.725	0.525	-92.947	120.0- 820.4	606.642	13.172	17.116	10.306

Fruit traits

Traits	σ^2_p	σ^2_g	σ^2_r	σ^2_b	*A*	h^2	*SG*	*Range*	µ	CV$_g$%	CV$_P$ %	CV*r* %
NFP	8.724	3.583^{ns}	4.614	0.527	0.655	0.429	2.303	1- 15	4.783	39.575	61.752	44.909
MF	2.738	2.373**	0.364	0.000	0.841	0.707	2.189	0.45-10.0	2.735	56.323	60.500	22.059
PF	73.954	38.598*	29.279	6.076	0.704	0.495	7.817	0.7- 44.6	12.946	47.989	66.427	41.796
TC	387.206	362.902**	24.303	0.000	0.880	0.774	20.426	43.4- 187.2	65.763	28.967	29.921	7.496

Seed and oil traits

Traits	σ^2_p	σ^2_g	σ^2_r	σ^2_b	*A*	h^2	*SG*	*Range*	µ	CV$_g$%	CV$_P$ %	CV*r* %
NSF	25274.617	20784.317**	2817.703	1672.597	0.844	0.712	167.873	78.6- 805.7	454.188	31.741	35.003	11.685
MSF	490.881	465.357**	16.141	9.382	0.899	0.808	27.428	4.4- 119.3	51.929	41.541	42.665	7.736
MS/F	0.000142523	0.000125343**	0.000015091	0.000002089	0.854	0.729	0.015	0.00- 0.05	0.023	48.676	51.905	16.890
MOHS	7.395	4.391*	3.003	0.000	0.721	0.519	2.210	6.3- 23.6	11.701	17.908	23.240	14.809
PS	0.042	0.019^{ns}	0.016	0.006	0.694	0.481	0.187	0.01- 0.9	0.269	51.241	76.185	47.022
SOC	13.010	0.462^{ns}	11.822	0.725	0.512	0.262	1.254	28.50- 54.4	18.516	3.670	19.480	18.569
SOP	0.001743	0.000172^{ns}	0.001300	0.000	0.540	0.291	0.072	0.004- 0.40	0.050	26.037	83.498	72.111

Accumulated degree-days for flowering (DDF), number of fruits per plant (NFP), average mass of fruit (MF), productivity of fruit (PF), total content of fruit pulp carotenoids (TC), number of seeds per fruit (NSP), mass of seeds per fruit (MSF), ratio of seed to fruit mass (MS/F), mass of one hundred seeds (MOHS), productivity of seeds (PS), seed oil content (SOC), and seed oil productivity (SOP). Components of variance involving phenotypic (σp), genotypic (σg), and residual (σ) variances, and the variance associated with the block effect (σb). Genetic-statistical parameters involving accuracy (*A*), broad-sense heritability (h^2), selection gain (*SG*), average (μ), coefficients of genotypic (CV$_g$ %), phenotypic (CV$_P$ %), and residual variation (CV$_r$ %). ns not significant; **, * significant at $p < 0.01$and 0.05, respectively by the likelihood ration test.

3.2 Genotypic correlations

A genotypic correlation network analysis and visualization of agronomic aspects, including the total content of fruit pulp carotenoids, and characteristics of seeds and seed oil is given in Figure 2, which shows cohesion of groups involving some of the fruit characteristics and those involving some of the characteristics of seeds. Cohesion is also shown between fruit productivity (PF) and other characteristics of this group, such as average mass of fruits (MF), diameter of internal cavity of fruit (DIC), height of fruit (HF), diameter of fruit (DF,), and thickness of fruit peel (TFP). As can be inferred from the colour and thickness of the lines, this set of variables showed high positive correlations. The highest correlations in this group were for PF with MF, and PF with DIC, with values equivalent to 0.61 and 0.54, respectively, both of which were significant ($p<0.01$). The productivity of fruits (PF) and number of fruits per plant (NFP) showed a correlation of 0.39, and each of these showed high correlations with the productivity of seeds (PS), 0.74 and 0.51, respectively all of which were significant ($p<0.001$), (Figure 2).

Accumulated degree-days for flowering (DDF) had low correlation with others characteristics. Seed oil content (SOC) had negative and low-magnitude correlations with soluble solids of fruit pulp (SS) and resistance of fruit pulp to penetration (RP) (Figure 2).

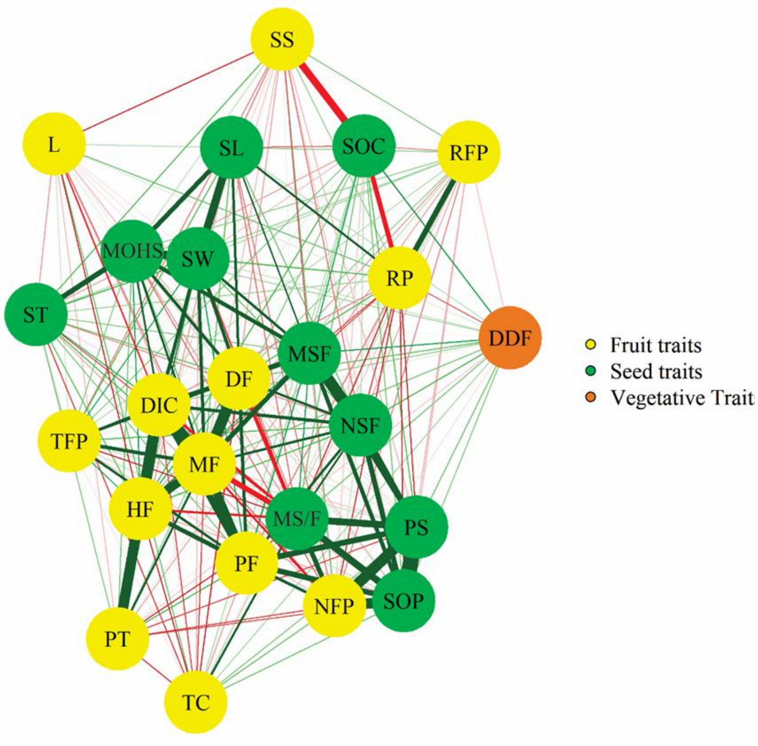


Figure 2. Network of genotypic correlations of agronomic aspects, the total content of fruit pulp carotenoids, and the characteristics of seeds and seed oil of the *C. moschata* germplasm assessed in this study and maintained by the BGH-UFV. The green and red lines denote positive and negative correlations, respectively. Thicker lines indicate greater magnitudes of correlation while the thinner lines indicate lesser magnitudes. Accumulated degree-days accumulated for flowering (DDF), number of fruits per plant (NFP), average mass of fruits (MF), productivity of fruits (PF), height of fruit (HF), diameter of fruit (DF), thickness of fruit peel (TFP), resistance of fruit peel to penetration(RFP), resistance of fruit pulp to penetration (RP), pulp thickness (PT), diameter of internal cavity of fruit (DIC), soluble solids of fruit pulp (SS), total content of fruit pulp carotenoids (TC), lutein content of fruit pulp (L), mass of seeds per fruit (MSF), productivity of seeds (PS), ratio of seed to fruit mass (MS/F), mass of one hundred seeds (MOHS), seed oil content (SOC), and seed oil productivity (SOP).

There was cohesion between the group of variables involved in seed productivity and variables such as the ratio of seed to fruit mass (MS/F), number of seeds per fruit (NSF), and mass of seeds per fruit (MSF). This set of variables had positive and high-magnitude correlations, of which the correlation of seed productivity (SP) with MS/F, equivalent to 0.56 and significant ($p<0.01$), was the highest. The group involving the mass of one hundred seeds (MOHS) and characteristics such as seed width (SW), seed thickness (ST), and seed length (SL) was also a cohesive group. This group had positive correlations, of which the correlation of

MOHS with seed width SW, equivalent to 0.62 and significant ($p<0.01$), was the highest (Figure 2).

3.3 Genetic variability and clustering

Cluster analysis, based on the agro-morphological aspects, the total content of fruit pulp carotenoids, and the characteristics related to the yields of seed and seed oil of the germplasm, placed the accessions into 16 groups (Table 3).

Based on the clustering pattern, high variability was observed between the accessions. About 17% of the genotypes were in group 11, together with the control, Jabras. Group 1, the second largest, contained 13.18% of the accessions and two controls, Jacarezinho and Maranhão. Groups 5 and 14 contained 10 and 11 accessions, respectively, making them the next largest groups formed. The grouping of genotypes in the other groups did not occur equitably and some of them contained only one genotype (Table 3).

The visual pattern of the clustering in heatmap format showed low similarity between the groups formed, as denoted by the predominance of yellow and orange colouring (Figure 3). Visual analysis of this clustering also shows homogeneity of the distances between groups, denoted by the uniformity of the colouring. The morphological pattern of fruits representative of part of the groups obtained with genotypes clustering is shown in the figure 4.

Table 3. Clustering of the *C. moschata* germplasm assessed in this study and maintained by BGH-UFV, based on agro-morphological aspects, the total content of fruit pulp carotenoids, and the yields of seeds and seed oil

Clusters	**Accessions**
1	BGH-117(BA); BGH-5616A(DF); BGH-5630A(DF); BGH-6590(GO); BGH-4281(MG), BGH-4454A(MG), BGH-6116(MG); BGH-5472A(SP), BGH-5541(SP), BGH-5556A(SP); Jacarezinho (BR); Maranhão (BR).
2	BGH-4459A(MG); BGH-5548A(SP).
3	BGH-4590A(MG).
4	BGH-5653(BA); BGH-1927(MG), BGH-4681A(MG).
5	BGH-1749(BA); BGH-7219A(PR), BGH-7668(PR); BGH-5051(RJ); BGH-5453A(SP), BGH-5473A(SP), BGH-5544A(SP), BGH-5591A(SP), BGH-5593(SP), BGH-5596A(SP).
6	BGH-4610A(MG), BGH-5361A(MG); BGH-3333A(RJ); BGH-5440A(SP), BGH-5485A(SP).
7	BGH-5455A(SP), BGH-5598A(SP).
8	BGH-5624A(DF); BGH-6587A(GO), BGH-6595(GO); BGH-5247A(MG), BGH-6115(MG); BGH-5493A(SP) BGH-5494A(SP), BGH-5559A(SP).
9	BGH-315(DF); BGH-6593(GO); BGH-1004(MG); BGH-5499A(SP), BGH-5530A(SP), BGH-5606A(SP).
10	BGH-1961(ES); BGH-4516(MG), BGH-5248(MG), BGH-5648(MG), BGH-5659A(MG); BGH-5442(SP), BGH-5538(SP), BGH-5554A(SP).
11	BGH-95(BA); BGH-5638(DF); BGH-1945A(ES); BGH-6794(GO); BGH-4453(MG), BGH-4607A(MG), BGH-6155(MG); BGH-5301(SP), BGH-5451(SP), BGH-5528(SP), BGH-5551(SP), BGH-5552(SP), BGH-5553(SP), BGH-5560A(SP), BGH-5597(SP); Jabras (BR).
12	BGH-5649A(BA).
13	GBH-5694(DF); BGH-6099(RN); BGH-900(SP).
14	BGH-5240(BA); BGH-5639(DF); BGH-4287A(MG), BGH-4598A(MG), BGH-5224A(MG), BGH-6117A(MG); BGH-1461A(SC), BGH-6749(SC); BGH-5466(SP), BGH-5497(SP), BGH-5603(SP).
15	BGH-1992(ES); BGH-6594(GO); BGH-305A(MG); BGH-6096(RN); BGH-291(RJ); BGH-5456A(SP).
16	Tetsukabuto (BR).

The letters next to the names refer to the initials of the genotypes' states of origin. Bahia (BA), Distrito Federal (DF), Espírito Santo (ES), Goiás (GO), Minas Gerais (MG), Paraná (PR), Rio Grande do Norte (RN), Rio de Janeiro (RJ), São Paulo (SP), and Santa Catarina (SC), Brazil (BR).

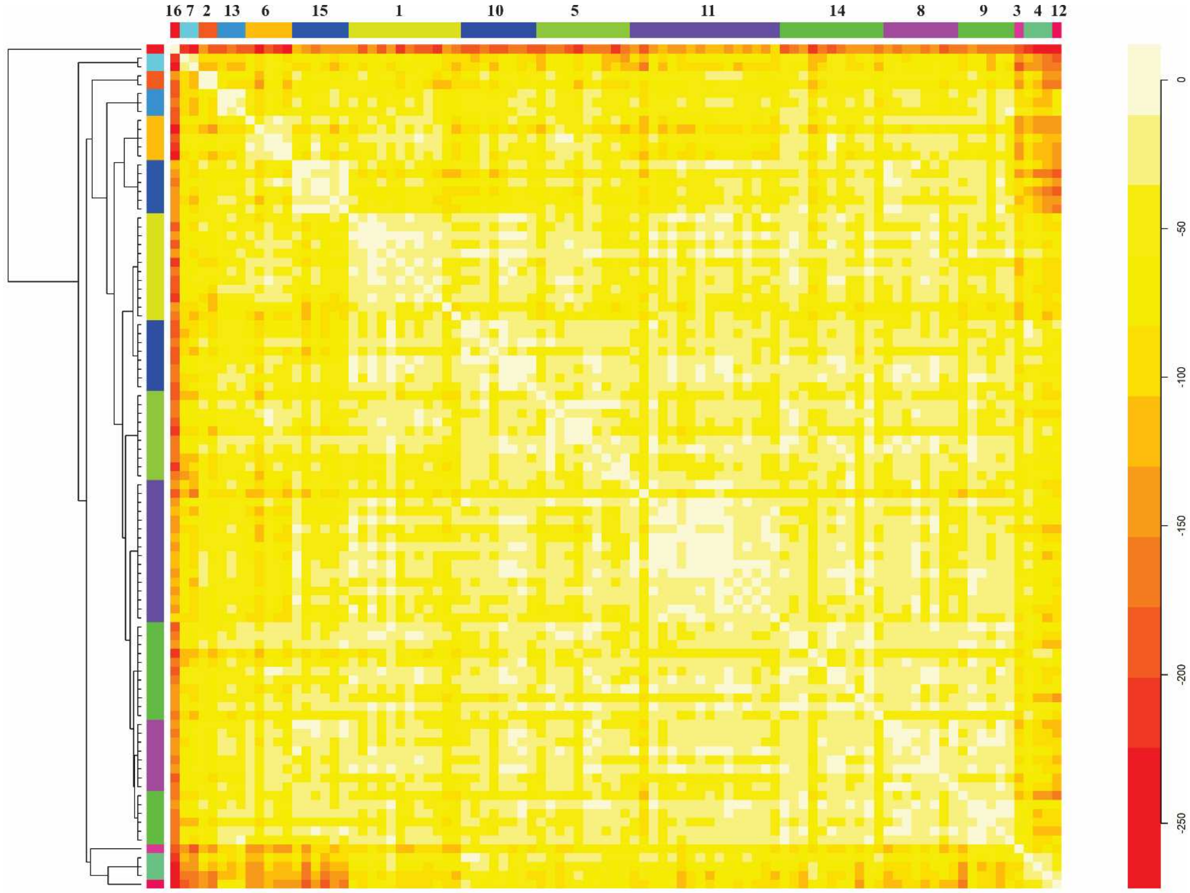


Figure 3. Heatmap and hierarchical clustering of the genetic distances of the *C. moschata* accessions, based on agro-morphological traits, the total content of fruit pulp carotenoids, and the yields of seeds and seed oil. The coloured bars on the upper and lower axis correspond to the groups obtained in the clustering. The dissimilarity between each pair of accessions and between groups is indicated by the colour, which varies from white to red. Red indicates the pairs of genotypes with the highest dissimilarity and white indicates the pairs of genotypes with lowest dissimilarities.

Figure 4. Figure showing the morphological pattern of fruits representative of part of the groups obtained with genotypes clustering. BGH-4590A (group 3), BGH-5560A and BGH-6155 (group 11), BGH-5649A (group 12), Jacarezinho and BGH-117 (group 1), BGH-4598A and BGH-5639 (group 14), BGH-5548A (group 2), BGH-5453A and BGH-5544A (group 5), and BGH-900 (group 13).

The result of the principal components analysis (PCA) refers to the first 15 independent components, which explained 55.56% of the total variation observed between the genotypes

(Figure 5). Component 1, which explained 7.35% of the total variation, had a greater contribution from quantitative variables, mainly from the average mass of fruits (MF), diameter of fruits (DF), and diameter of internal cavity of fruit (DIC). Component 2 had a greater contribution from the multi-categorical traits and explained 5.93% of the total variation (Figure 5). The result of PCA regarding the fifteen principal components and the relative contribution of traits in each component is provided in the Supplementary Table 2.

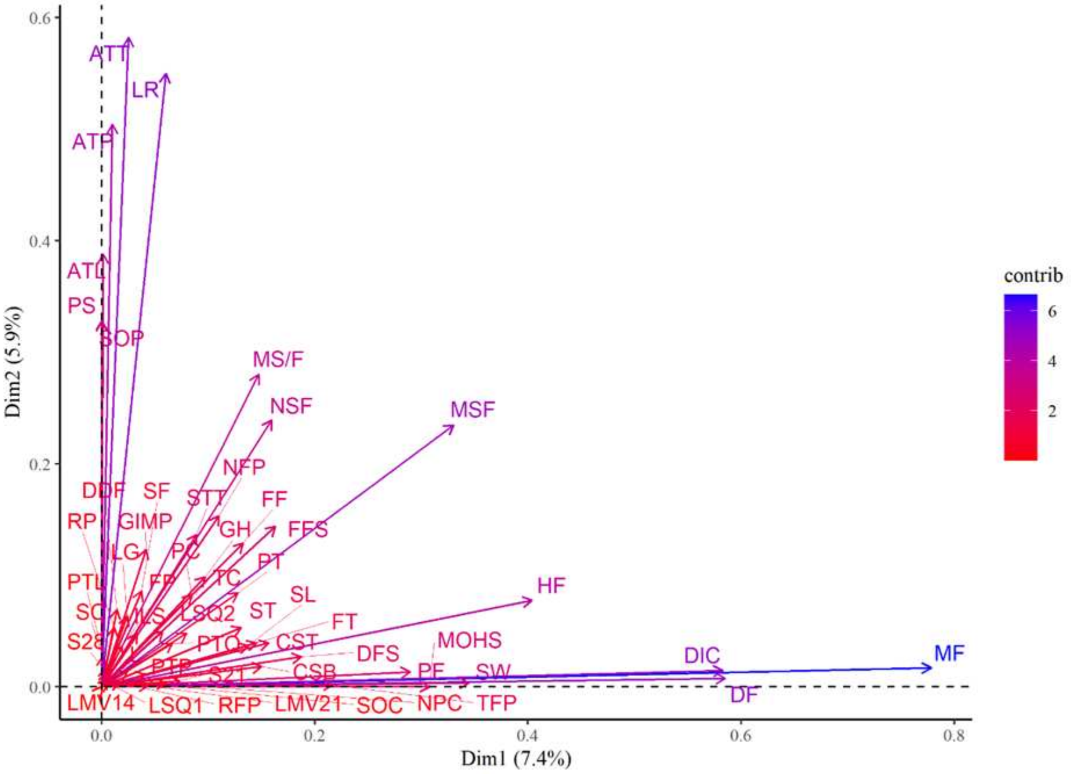


Figure 5. Dispersion of quantitative and multi-categorical characteristics in relation to the first two components: leaf SPAD index at 21 days after transplanting (S21) and at 28 days (S28); length of main vine at 14 days after transplanting (LMV14) and at 21 (LMV21); accumulated degree-days for flowering (DDF); number of fruits per plant (NFP); average mass of fruits (MF); productivity of fruits (PF); height of fruit (HF); diameter of fruit (DF); thickness of fruit peel (TFP); resistance of fruit peel to penetration (RFP); resistance of fruit pulp to penetration (RP); thickness of fruit pulp (PT); diameter of internal cavity of fruit (DIC); total content of fruit pulp carotenoids (TC), and lutein (L); number of seeds per plant (NSF); mass of seeds per fruit (MSF); ratio of seed to fruit mass (MS/F); mass of one hundred seeds (MOH); productivity of seeds (PS); seed thickness (ST); seed length (SL); seed width (SW); seed oil content (SOC); seed oil productivity (SOP). Growth habit (GH); stem colour (SC); intensity of leaf green (ILG), leaf silvering (LS), intensity of leaf silvering (ILS); leaf serration (LS); presence of trichomes in the leaves (PAL); amount of trichomes in the adaxial surface of leaves (ATT); amount of trichomes in the abaxial surface of

leaves (ATL); leaf recess (LR); presence of trichomes in the petiole (PTP); amount of trichomes in the petiole (ATP); green intensity of male pedicel (GIMP); format of fruits (FF); format of peduncle (FP); number of colours of fruit peel (NPC); topography of fruit surface (FT); format of floral scar (FFS); peel texture (PT); predominant colour of fruit peel (PC); depth of fruits slices (DFS); seed format (SF); aspect of seed tegument (AST); seed tegument texture (STT); colour of seed tegument (CST); colour of seed border (CSB).

3.4 Identification of promising clusters and per se identification of promising genotypes

In order to facilitate the visualisation of clusters with the most desirable characteristics, a grouping of means of clusters was performed by the Tocher method (Table 4).

The lowest mean for accumulated degree-days for flowering (DDF) occurred in Group 16, which contained only the control Tetsukabuto, although most groups expressed intermediate averages for this characteristic (Table 4). The group with the highest mean for productivity of fruits (PF) was Group 4, formed by the accessions BGH-1927, BGH-4681A, and BGH-5653. This group also expressed one of the highest averages for mass of fruits (MF) and an intermediate average for number of fruits per plant (NFP). As for the total content of fruit pulp carotenoids (TC), the highest average occurred in Group 7, formed by the accessions BGH-5455A and BGH-5598A. Groups 1 and 6 expressed the highest averages for seed (PS) and seed oil productivity (SOP). Group 1 contained the largest number of accessions (Table 4).

The identification per se of the most promising accessions for each trait, based on their respective genotypic effects, is shown in Tables 5 and 6. Also in these tables are the estimates, for each accession, of their genetic gains and the new predicted average for each trait.

The selected accessions had averages for accumulated degree-days for flowering (DDF) that were much lower than the general average of the accessions (606.64) and the average of the controls (526.41), with their new predicted averages ranging from 474.39 to 251.09, and genetic gains from -132.25 to -355.55. Notably, the accessions BGH-6749, BGH-5639, and BGH-2191 were the most promising for DDF (Table 5).

Table 4. Grouping of means of the genotypic values of the groups obtained in the analysis of variability for agro-morphological aspects, the total content of fruit pulp carotenoids, and productivities of seed and seed oil

Groups	DDF	NFP	MF	PF	TC	NSF	MSF	MS/F	MOHS	PS	SOC	SOP
1	19.15^{b}	0.48^{b}	-0.22^{d}	1.46^{c}	2.09^{b}	94.22^{b}	0.52^{b}	0.52^{b}	0.63^{c}	0.12^{a}	-0.07^{a}	0.04^{a}
2	32.20^{b}	-0.69^{b}	-0.58^{e}	-2.80^{c}	-8.77^{b}	-95.99^{b}	-0.80^{b}	-0.80^{c}	-1.54^{d}	-0.08^{b}	-0.16^{a}	-0.03^{b}
3	-35.38^{b}	0.15^{b}	1.17^{b}	5.61^{b}	-7.46^{b}	-207.22^{c}	-1.22^{b}	-1.22^{d}	3.62^{a}	-0.05^{b}	0.12^{a}	-0.02^{b}
4	-19.47^{b}	-0.67^{b}	4.53^{a}	10.41^{a}	-4.95^{b}	129.02^{b}	-0.53^{b}	-0.53^{c}	1.66^{b}	0.02^{b}	0.07^{a}	0.01^{b}
5	2.76^{b}	0.33^{b}	0.40^{c}	2.89^{c}	-8.54^{b}	44.57^{b}	-0.40^{b}	-0.40^{c}	-0.85^{c}	0.02^{b}	0.01^{a}	0.01^{b}
6	11.73^{b}	2.69^{a}	-1.23^{g}	1.50^{c}	-0.84^{b}	79.53^{b}	1.76^{b}	1.76^{a}	-1.16^{d}	0.21^{a}	-0.01^{a}	0.07^{a}
7	36.35^{b}	1.24^{a}	-1.36^{g}	-2.48^{c}	111.02^{a}	13.62^{b}	1.90^{b}	1.90^{a}	-0.08^{c}	0.06^{b}	0.06^{a}	0.03^{b}
8	-9.88^{b}	0.14^{b}	-0.47^{d}	-0.09^{c}	2.01^{b}	43.54^{b}	0.56^{b}	0.56^{b}	-1.79^{d}	-0.01^{b}	-0.14^{a}	-0.01^{b}
9	-8.08^{b}	-0.66^{b}	-0.93^{f}	-3.83^{c}	-6.11^{b}	-36.84^{b}	0.42^{b}	0.42^{b}	-0.41^{c}	-0.06^{b}	0.02^{a}	-0.02^{b}
10	40.41^{b}	-0.65^{b}	0.84^{b}	-0.10^{c}	-5.74^{b}	151.09^{b}	-0.3^{b}	-0.31^{c}	0.97^{b}	0.00^{b}	-0.03^{a}	0.00^{b}
11	-10.76^{b}	-0.80^{b}	0.29^{c}	-1.32^{c}	-1.03^{b}	-96.72^{b}	-0.79^{b}	-0.79^{c}	0.91^{b}	-0.09^{b}	-0.05^{a}	-0.03^{b}
12	127.82^{a}	-1.17^{b}	4.33^{a}	5.60^{b}	-11.10^{b}	228.21^{a}	-0.65^{a}	-0.65^{c}	2.62^{a}	0.03^{b}	0.20^{a}	0.02^{b}
13	82.43^{a}	-0.30^{b}	-0.24^{d}	-1.24^{c}	-5.32^{b}	-46.40^{b}	-0.02^{b}	-0.02^{c}	-0.19^{c}	-0.03^{b}	0.11^{a}	-0.01^{b}
14	-66.85^{b}	-0.17^{b}	-0.02^{d}	-0.47^{c}	1.13^{b}	22.01^{b}	0.48^{b}	0.48^{b}	-0.08^{c}	0.01^{b}	0.04^{a}	0.00^{b}
15	-43.65^{b}	0.72^{b}	-1.39^{g}	-2.61^{c}	-0.99^{b}	-243.74^{c}	-0.86^{b}	-0.86^{c}	-1.34^{d}	-0.08^{b}	-0.03^{a}	-0.03^{b}
16	-138.72^{c}	-0.06^{b}	-1.32^{g}	-4.92^{c}	-10.69^{b}	-393.92^{d}	-1.82^{c}	-1.82^{d}	0.07^{c}	-0.22^{c}	-5.50^{b}	-0.05^{c}

The genotypic values of the accessions (BLUPS) and controls (BLUES) vary from negative to positive; therefore the signal of group means is a reflection of the genotypic values in each group. Accumulated degree-days accumulated for flowering (DDF), number of fruits per plant (NFP), average mass of fruits (MF), productivity of fruits (PF), total content of fruit pulp carotenoids (TC), number of seeds per fruit (NSF), mass of seeds per fruit (MSF), ratio of seed to fruit mass (MS/F), mass of one hundred seeds (MOHS), productivity of seeds (PS), seed oil content (SOC), and seed oil productivity (SOP). The letters a, b, c, d, f, and g refer to the groups formed in the clustering of means obtained by the Tocher method.

For productivity of fruits (PF), the selected accessions had higher averages than the general average of the accessions (12.95 t ha^{-1}) and the average of the controls (11.85 t ha^{-1}), with their new predicted averages ranging from 15.49 to 29.27 t ha^{-1}. As for total content of fruit pulp carotenoids (TC), the selected accessions also had much higher averages than the general average of the accessions (65.76 µg g^{-1} of fresh weight) and that of the controls (65.58 µg g^{-1} of fresh weight). The new averages predicted for this characteristic among those selected ranged from 72.34 to 179.46 µg g^{-1} of fresh pulp mass, and the most promising accessions for this characteristic were BGH-5455A and BGH-5598A (Table 5).

The identification per se of the most promising accessions for productivity of seeds (PS), seed oil content (SOC) and seed oil productivity (SOP), together with their respective genetic gains and new predicted averages for these characteristics is shown in Table 6.

As for productivity of seeds (PS), the new predicted averages among the selected accessions ranged from 0.33 to 0.58 t ha^{-1} and the genetic gains from 0.06 to 0.31 t ha^{-1}. Notably, the accessions BGH-4610A, BGH-5485A, and BGH-6590 were the most promising for this characteristic (Table 6). The selected accessions displayed small differences in seed oil content (SOC); however, the average of these was higher than that of the controls (16.73%). Finally, for seed oil productivity (SOP), the new predicted averages ranged from 0.12 to 0.13 t ha^{-1} and the genetic gains from -0.07 to -0.08 t ha^{-1}. The accessions BGH-5485A, BGH-4610A, and BGH-5472A were the most promising for this characteristic (Table 6).

Table 5. Estimates of the genotypic effects, genetic gain and new predicted averages for the accumulated degree-days for flowering (DDF), fruit productivity (PF) and total content of fruit pulp carotenoids (TC), for the top 15% most promising accessions and the controls

	DDF				**PF**				**TC**		
Accessions	g	Gain	New Average	Accessions	g	Gain	New Average	Accessions	g	Gain	New Average
BGH-6749	-291.35	-355.55	251.09	BGH-4453	17.32	16.32	29.27	BGH-5455A	113.85	113.70	179.46
BGH-5639	-152.11	-216.29	390.35	BGH-5653	5.60	15.44	28.38	BGH-5598A	108.19	108.03	173.80
BGH-291	-119.80	-183.97	422.67	BGH-5544A	13.82	12.82	25.76	BGH-1461A	14.18	14.03	79.80
BGH-5638	-102.40	-166.57	440.07	BGH-4681A	11.74	10.74	23.69	BGH-5616A	11.50	11.35	77.12
BGH-6587A	-91.41	-155.57	451.07	BGH-5224A	10.21	9.21	22.16	BGH-6794	11.08	10.93	76.70
BGH-5624A	-83.56	-147.73	458.91	BGH-6587A	-4.26	8.35	21.29	BGH-5556A	9.85	9.70	75.47
BGH-1004	-83.32	-147.48	459.16	BGH-4590A	5.61	4.61	17.55	BGH-5606A	8.78	8.63	74.40
BGH-1749	-76.12	-140.28	466.36	BGH-5649	5.60	4.59	17.54	BGH-5497	8.73	8.58	74.34
BGH-5301	-76.12	-140.28	466.36	BGH-5051	5.30	4.30	17.25	BGH-5451	8.68	8.53	74.29
BGH-5456A	-75.88	-140.04	466.60	BGH-5596A	4.52	3.52	16.46	BGH-5493A	8.62	8.47	74.24
BGH-5485A	-75.88	-140.04	466.60	BGH-5248	4.09	3.09	16.03	BGH-5247A	7.34	7.19	72.95
BGH-5530A	-75.88	-140.04	466.60	BGH-5472A	4.06	3.06	16.00	BGH-6749	6.91	6.76	72.53
BGH-6794	-68.94	-133.11	473.53	BGH-5473A	3.62	2.62	15.57	BGH-6587A	13.49	6.71	72.47
BGH-4598A	-68.09	-132.25	474.39	BGH-5556A	3.54	2.54	15.49	BGH-95	6.73	6.58	72.34
Average			606.64	Average			12.95	Average			65.76
Controls				Controls				Controls			
	BLUES	Gain	New Average		BLUES	Gain	New Average		BLUES	Gain	New Average
Jabras	-192.84	-54.12	413.99	Jabras	8.53	8.87	20.88	Jabras	14.93	16.19	80.68
Tetsukabuto	-138.72	0.00	468.11	Tetsukabuto	-1.03	-0.69	11.31	Tetsukabuto	0.23	1.49	65.98
Maranhão	3.99	142.71	610.82	Maranhão	-4.60	-4.26	7.75	Maranhão	-5.13	-3.88	60.61
Jacarezinho	5.89	144.61	612.72	Jacarezinho	-4.92	-4.57	7.43	Jacarezinho	-10.69	-9.43	55.06
Average			526.41	Average			11.85	Average			65.58

Table 6. Estimates of the genotypic effects, genetic gain and new predicted averages for the productivity of seeds (PS), seed oil content (SOC), and seed oil productivity (SOP), for the top 15% most promising accessions and the controls

Accessions	PS			Accessions	SOC			Accessions	SOP		
	g	Gain	New Average		g	Gain	New Average		g	Gain	New Average
BGH-4610A	0.34	0.31	0.58	BGH-7219A	0.43	-0.98	17.53	BGH-5485A	0.01	-0.07	0.13
BGH-5485A	0.30	0.28	0.54	BGH-5649	0.20	-1.21	17.30	BGH-4610A	0.01	-0.07	0.13
BGH-6590	0.29	0.26	0.53	BGH-5653	0.16	-1.25	17.27	BGH-5472A	0.01	-0.07	0.13
BGH-5556A	0.22	0.19	0.46	BGH-5466	0.16	-1.25	17.27	BGH-5556A	0.01	-0.07	0.12
BGH-5472A	0.22	0.19	0.46	BGH-900	0.16	-1.26	17.26	BGH-6590	0.01	-0.07	0.12
BGH-5544A	0.19	0.17	0.44	BGH-6155	0.16	-1.26	17.26	BGH-5544A	0.01	-0.07	0.12
BGH-5440A	0.18	0.15	0.42	BGH-5544A	0.15	-1.26	17.25	BGH-4281	0.01	-0.08	0.12
BGH-4281	0.16	0.13	0.40	BGH-6794	0.15	-1.26	17.25	BGH-5440A	0.01	-0.08	0.12
BGH-5361A	0.16	0.13	0.40	BGH-5472A	0.15	-1.27	17.25	BGH-5630A	0.01	-0.08	0.12
BGH-5630A	0.15	0.12	0.39	BGH-305A	0.14	-1.27	17.24	BGH-5473A	0.01	-0.08	0.12
BGH-5473A	0.15	0.12	0.39	BGH-5455A	0.13	-1.28	17.23	BGH-5361A	0.01	-0.08	0.12
BGH-5453A	0.13	0.10	0.37	BGH-5240	0.12	-1.29	17.23	BGH-5453A	0.00	-0.08	0.12
BGH-4287A	0.10	0.08	0.34	BGH-4681A	0.12	-1.29	17.22	BGH-5455A	0.00	-0.08	0.12
BGH-4454A	0.09	0.06	0.33	BGH-4590A	0.12	-1.29	17.22	BGH-5466	0.00	-0.08	0.12
Average			0.27	Average			18.52	Average			0.11
Controls				Controls				Controls			
	BLUES	Gain	New Average		BLUES	Gain	New Average		BLUES	Gain	New Average
Jacarezinho	0.28	0.30	0.53	Jacarezinho	0.39	2.38	18.98	Jacarezinho	0.05	0.01	0.10
Maranhão	0.03	0.06	0.29	Maranhão	-0.91	1.07	17.67	Maranhão	0.01	-0.03	0.06
Jabras	-0.14	-0.12	0.11	Jabras	-1.40	0.59	17.19	Jabras	-0.03	-0.08	0.01
Tetsukabuto	-0.22	-0.19	0.04	Tetsukabuto	-5.50	-3.52	13.08	Tetsukabuto	-0.05	-0.09	0.00
Average			0.24	Average			16.73	Average			0.04

4. DISCUSSION

4.1 Variance components and genetic-statistical parameters of the agronomic aspects, total content of fruit pulp carotenoids, and the characteristics of seeds and seed oil

As with other species, the usefulness of *C. moschata* germplasm conserved in banks depends on the level and quality of information associated with it [30, 31, 32, 33, 49,]. The samples of *C. moschata* maintained by BGH-UFV constitute one of the largest collections of this species in Brazil [34]. Studies involving the assessment of this germplasm have allowed the identification of accessions with crucial characteristics for this crop, such as phytopathogenic resistance, and for its genetic improvement in terms of production and nutritional aspects of its fruits and seed oil [10, 21, 36, 37]. Although BGH-UFV maintains more than 350 accessions of *Cucurbita* ssp. [35], part of this germplasm has not yet been assessed, demonstrating the importance of continuing these studies.

Most of the *C. moschata* germplasm express vigorous growth and indeterminate growth habit [50], and *C. moschata* plants commonly occupy a large area of cultivated land, making it difficult to phenotypically assess its germplasm in experimental designs such as in randomised blocks. The main limitation in the evaluation of *C. moschata* germplasm in randomised blocks is the difficulty of ensuring satisfactory homogeneity throughout the experimental area. In addition, the germplasm seed samples kept in banks in most cases are small, making it impossible to repeat accessions throughout the experimental area and assess quantitative characteristics. In view of this, we proposed in this study to evaluate part of the *C. moschata* germplasm maintained at BGH-UFV using the design known as Federer's augmented blocks [39]. The details of all aspects inherent to this design are very well described by Federer and, according to him, the design circumvents the limitations mentioned above and can be adopted even when the propagating material is insufficient for the establishment of more than one plot and where the quantity of samples to be evaluated is too great.

The present study describes the evaluation of one of the largest germplasm volumes of C. moschata. The high estimates of genotypic variance for characteristics related to seed production observed in this study corroborate those reported by [51], who also observed higher estimates of genotypic variance for the number of seeds per fruit and flowering characteristics, and also a greater contribution of genotypic variance to the phenotypic variance in these characteristics. Additionally, most of the characteristics assessed in this study gave high estimates of heritability (>0.50), considering the classification of [52], especially the characteristics of seeds such as mass of seeds per fruit (MSF), ratio of seed to fruit mass (MS/F), and number of seeds per fruit (NSF), as well the aspects related to fruits, such as total content of fruit pulp carotenoids (TC) and mass of fruit (MF). High estimates of heritability point to a greater correlation between the phenotype and the genotype [53], indicating that most of the variability observed for these characteristics resulted from genotypic effects.

The high estimates of genotypic variances may be associated with the quantitative nature of these characteristics, which may be the result of the influence of a high number of genes [54]. Most of the germplasm evaluated in this study came from the land of family-based farmers, who do not carry out selection either for seed characteristics or to obtain earlier-flowering genotypes. As already mentioned, the exchange of seeds between farmers and the natural occurrence of hybridisation between populations of *C. moschata* has increased the variability of this species, even for characteristics for which selection is commonly carried out, such as fruit productivity.

Considerable predicted gains were obtained for most of the characteristics, considering the overall average of accessions. This result was associated with the high estimates of genotypic variance and heritability observed for most of the characteristics (Table 2).

The average relationship between the coefficient of genetic variation and the residual coefficient was close to one unit for most of the characteristics. Although the estimates of the residual coefficients of variation for most characteristics were high, in general they tended to be lower in relation to their corresponding coefficients of genotypic variability, which demonstrates that most of the variability expressed by germplasm was due to genetic factors (Table 2).

4.2 Genetic correlation network

Analysis of correlations between characteristics has been widely used in plant breeding, where often a high number of characteristics must be considered simultaneously [55, 56]. This analysis is often used to assist in indirect selection for certain characteristics [55, 57]. However, as highlighted by [58], in cases where one intends to practise indirect selection for a primary characteristic by means of a secondary one, the heritability of the latter characteristic must be greater than that of the former for efficient selection. In view of this, the selection of genotypes with higher average mass of fruits (MF) seems to be a promising alternative for obtaining higher fruit productivity in C. moschata.

It should, however, be highlighted that when selecting genotypes for increasing fruit productivity in C. moschata, crucial aspects for their acceptability in the consumer market, such as the shape and size of fruits, must be considered. Currently, important pumpkin consumption centres like the state of Minas Gerais and most of the southeast region of Brazil demand smaller fruits, and most of the consumption in these regions is represented by fruits from hybrid cultivars, such as Jabras and Tetsukabuto, which have a globular shape and weigh from 2 to 3 kg [14]. On the other hand, in the north and northeast regions of Brazil, larger fruits, which are commonly sold in slices, are more acceptable. The prevention of waste and the ease of transport are determining aspects for the acceptability of fruit shapes, and the search for greater productivity in the cultivation of *C. moschata* must therefore also consider these characteristics, equating them with aspects such as the number of fruits per plant (NFP), height of fruit (HF) and diameter of fruit (DF).

Based on the correlations obtained in this study, the simultaneous consideration of aspects such as higher number of fruits per plant (NFP), higher productivity of fruits (PF) and higher ratio of seed to fruit mass (MS/F) seems to be a promising alternative for obtaining higher seed productivity (PS) in *C. moschata*. The heritability estimates obtained for these characteristics (>0.42), suggest that reasonable gains are feasible with selection for each one of them (Table 2). With this, besides greater PF and NFP, the selection of genotypes with higher PS should also prioritise greater translocation of photoassimilates for seed production, something indicated by a higher ratio of seed to fruit mass (MS/F).

Despite its applicability, correlation analysis has some limitations, and, as warned by [59], the quantification and interpretation of the correlation coefficients between two or more characteristics can result in errors during the selection process. According to them, this occurs because high estimates of correlations between these characteristics may be the effect of one or more secondary characteristics. It is therefore recommended that analysis of the association between a primary and secondary characteristic be accompanied by information on the direct and indirect effects of secondary variables on the primary [60], an approach currently known as path analysis [59].

Despite some limitations, correlation analysis has proven to be quite useful in plant breeding, mainly in the indirect selection for one or more main characteristics that have low heritability or are difficult to assess. This indirect selection is based on secondary characteristics with greater heritability or ease of assessment, providing faster genetic gains than with direct selection. In fact, correlation analysis has assisted in the indirect selection for characteristics of roots [61], for productivity in different crops [62, 63, 64], and for nutritional aspects and quality of fruits [65, 66]. Correlation analysis can also be very useful in the characterisation and management of plant germplasm, as it may optimise the choice and number of descriptors to be used in this process.

4.3 Genetic variability and clustering

The analysis of variability provides important assistance in the initial phase of plant breeding programmes and in the management of plant germplasm. In this first case, it provides allocation of accessions in groups, guiding crossbreeding. *C. moschata* is allogamous, and analysing the variability of its germplasm can assist in the orientation of crossings between more diverse genotypes, thereby aiding the exploration of hybrid vigour [67, 68]. Variability analysis also allows duplicates in the germplasm collections [69, 70, 71], which correspond to pairs or groups of accessions with high similarity, to be identified. In fact, it is estimated that less than 30% of the accessions maintained in the collections worldwide are distinct, which hinders their maintenance [29]. Therefore, in addition to optimising the use of germplasm, variability analysis reduces the cost of its maintenance by reducing its volume [72].

The accessions of *C. moschata* assessed in this study displayed high genetic variability in their agro-morphological characteristics, the total content of fruit pulp carotenoids (TC), and the productivity of seeds (PS) and seed oil (SOP), resulting in the formation of 16 clusters (Table 3). The clustering of Jacarezinho and Maranhão in the same group (Group 1) reflects its consistency since these two cultivars have similar characteristics.

Clustering did not reflect a smaller genetic distance between those accessions from the same state or geographic region of Brazil. Group 11, for example, grouped accessions from different states and regions; and the preponderance of accessions from Minas Gerais (MG) and São Paulo (SP) in this group was probably only a result of the greater number of accessions from these states. This trend was repeated for other groups with higher numbers of accessions such as 1, 5 and 14. A study involving the assessment of *C. moschata* accessions from different regions of Brazil and maintained at BGH-UFV [73] also did not report smaller genetic distance between the accessions from the same state or region.

It is notable that the two hybrids used as controls, Jabras and Tetsukabuto, clustered in different groups. Although they have similar fruit shape and size, the groups to which they were allocated differed in most characteristics (Table 4), and their different genotypic values for most characteristics (Tables 5 and 6) justified their clustering in different groups. Tetsukabuto, which is an interspecific hybrid between *C. moschata* and C. maxima [74], corresponded to the group with lowest genotypic average for accumulated degree-days for flowering (DDF), in addition to expressing genotypic averages quite different from the other groups in relation to the characteristics of seeds and seed oil (Table 4), justifying its clustering separately from the other genotypes.

The predominance of yellow colour in the hierarchical clustering in heatmap format denoted low similarity between the clusters formed (Figure 3). As can also be seen in Figure 3, the uniformity in the yellow coloration for the genetic distances between groups confirms the homogeneity of distances between them.

The variability denoted by the clustering of the accessions corroborates the high estimates of genetic variances and heritabilities displayed by most of the agronomic characteristics; the total content of fruit pulp carotenoids (TC); and seed characteristics such as mass of seeds per fruit (MSF), ratio of seed to fruit mass (MS/F), and number of seeds per fruit (NSF) (Table 2). This is also analogous to other studies involving the analysis of variability in this crop in Brazil [19, 21].

The greater contribution of the average mass of fruits (MF), diameter of fruit (DF), diameter of internal cavity of fruit (DIC), as well as the mass of seeds per fruit (MSF), and number of seeds per fruit (NSF) for component 1, suggests that there was greater variability for these characteristics, and that they contributed more to genotype discrimination (Figure 5). This result seems to be related to the estimates of genotypic variance, since MSF and NSF also corresponded to characteristics with the greatest genotypic variances (Table 2). The greatest contribution, in component 2, of variables such as the amount of trichomes (AT), leaf recess (LR) and amount of trichomes in the petiole (ATP) shows the importance of multi-categorical characteristics in the discrimination of the studied germplasm.

4.4 Identification of promising groups of genotypes

In *C. moschata*, the identification of promising groups of genotypes can assist in the orientation of crossings targeting hybrid vigour exploitation and the segregation of populations for their characteristics of interest [75, 76].

As shown in Table 4, Group 1 expressed a high genotypic average for total content of fruit pulp carotenoids (TC) and the highest averages for productivity of seeds (PS) and seed oil content (SOC), confirming the high number of promising accessions for these characteristics. The negative correlations between SOC and characteristics related to the quality of fruit pulp in *C. moschata*, such as content of soluble solids (SS) and resistance of fruit pulp to penetration, might hinder simultaneous gains for these characteristics. This can be managed by conducting individualised breeding subprogrammes, aiming in one case to improve seed oil production, and in another, to improve fruit production and quality.

The highest average for total content of fruit pulp carotenoids (TC) occurred in Group 7, formed by the accessions BGH-5455A and BGH-5598A (Table 4). These accessions were also identified as the most promising for TC in the identification per se, with new predicted averages greater than 170 μg g^{-1} of fresh pulp mass (Table 5). This result is much higher than those reported in previous studies [4, 37, 77]. Among these, the study involving the characterisation of 55 accessions of *C. moschata*, also maintained by the BGH-UFV, reported a total content of fruit pulp carotenoid averages not greater than 118.70 μg g^{-1} of fresh pulp mass [37]. On the other hand, averages of up to 404.98 μg g^{-1} of fresh pulp mass have been reported [1, 72], when evaluating *C. moschata* germplasm from northeast Brazil. The differences observed for the total content of fruit pulp carotenoids between the present study and previous studies might be mainly associated with the genetic aspects of the germplasm evaluated in each study. According to [72], in northeast Brazil there is a preference for winter squash fruits with more orange pulp, a characteristic associated with higher levels of carotenoids, which corroborates the results obtained for this characteristic in studies involving the evaluation of *C. moschata* germplasm from this region.

Studies with *C. moschata* commonly involve the analysis of fruit pulp carotenoids and generally report high levels of these components [1, 4, 78, 79]. Among these studies, about 19 different carotenoids in the carotenogenic profile of the fruit pulp were identified [1], and *β*- and *α*-carotene constituted the largest proportion of the total carotenoid content in this species. In fact, this vegetable has been considered one of the best sources of carotenoids such as *β*-carotene, with levels above those found in other important carotenogenic vegetables, such as carrots [80].

The main biological functions of components such as *α*- and *β*-carotene are their pronounced pro-vitamin A activity [81, 82], and a series of bioactive functions, especially antioxidant activity [83, 84]. Along with its bioactive functions, *C. moschata* brings together fundamental characteristics for biofortification programmes, such as high production potentials and profitability, high efficiency in reducing deficiencies in micronutrients in humans, and good acceptance by producers and consumers in the regions where this crop is grown [8]. *C. moschata* has therefore been strategically used in programmes targeting biofortification in vitamin A precursors, among them the Brazilian Biofortification Programme (BioFORT), led by the Brazilian Agricultural Research Corporation (Embrapa) [9].

The main interest in the assessment of productivity of seeds (PS) and seed oil productivity (SOP) in *C. moschata* corresponds to the high potential for using its seed oil for food purposes. Governments and health experts are interested in encouraging the consumption of unsaturated fatty acids rather than saturated ones, based on the consensus that this reduces the risk of cardiovascular diseases [85, 86, 87], and this vegetable not only has a high oil content, with the lipid fraction of its seeds reaching up to 49% of its composition [88], but the lipid profile of this oil consists of more than 70% unsaturated fatty acids, with a preponderance of fatty acids such as linoleic C18: 2 ($\Delta^{9,12}$) and oleic C18: 1 (Δ^{9}).

C. moschata seed oil is also rich in bioactive components such as vitamin E and carotenoids [13], which have important antioxidant activity, in addition to protecting the oil against oxidative processes. Despite this, most of the seeds from the production of *C. moschata* in Brazil are still discarded during consumption. Their use, therefore, represents an alternative way of supplementing diets as well as increasing the income of farmers involved in the production of this vegetable.

Group 16, consisting solely of the control Tetsukabuto, displayed the lowest average for accumulated degree-days for flowering (DDF), indicating that this genotype has the earliest flowering period (Table 4). As can also be seen in the Table 4, most groups had intermediate averages for DDF. Normally, *C. moschata* plants have very long internodes, and this, coupled with the vigorous growth of this species, limits its cultivation, since plants with a greater internode length require much larger areas for cultivation. The interest in assessing precocity in *C. moschata* is based on the possible relationship of this characteristic with aspect such determinate growth habit. According to [89], the *Bu* gene, identified as being responsible for the formation of shorter internodes in pumpkins, is also linked to earlier flowering in this species. In a study evaluating hybrids and segregating winter squash populations for oil production and plant size reduction [50], the cultivars Piramoita and Tronco Verde, which have determinate growth habits, displayed the smallest number of days for female flowering. Greater precocity is an important characteristic for most crops, especially in the cultivation of vegetables, as it optimises the use of cultivation areas, reduces the risks of exposure of the crop to adverse abiotic and biotic factors, and reduces management costs.

In view of the low correlation observed between accumulated degree-days for flowering (DDF) and the other characteristics, it is unlikely that accessions that simultaneously express earlier-flowering and other important characteristics in *C. moschata* will be identified. Therefore, the initial identification of earlier-flowering accessions, followed by incorporation of this trait in germplasm that is promising for other characteristics seems appropriate in *C. moschata* breeding.

Group 4, formed by BGH-1927, BGH-4681A and BGH-5653, had the highest average for productivity of fruits (PF) (Table 4). It also had one of the highest averages for mass of fruits (MF) and an intermediate average for number of fruits per plant (NFP), corroborating the estimates of the correlations between these characteristics and productivity of fruits (Figure 2). The accessions BGH-4681A and BGH-5653 were also identified as the most promising for PF in the per se identification, with averages above 20 t ha^{-1} (Table 5). These averages were much higher than the world average, estimated at 13.4 t ha^{-1} [6].

Although the cultivation of *C. moschata* is primarily intended for fruit production, as already mentioned, the selection of genotypes for greater fruit productivity in this crop must also consider crucial aspects for the acceptability of fruits such as shape and size. In general, winter squash production must currently prioritise the adoption of cultivars with smaller fruits. In addition to obtaining fruits of greater mass, greater productivity in *C. moschata* can also be achieved by obtaining cultivars with higher number of fruits per plant (NFP), based on the estimated correlation observed between productivity of fruits (PF) and NFP (Figure 2).

4.5 *Per se* identification of promising accessions

Per se identification of promising accessions can guide selection for a specific trait, allowing the identification of promising accessions for the development of superior inbred lines and/or open-pollinated cultivars. In fact, from a brief survey of the Brazilian National Cultivar Register (RNC), it appears that, of the 182 cultivars of *C. moschata* registered at the moment, most consist of open-pollinated cultivars [74]. This survey also found a considerable number of intra- and interspecific hybrids, confirming the feasibility of applying inbreeding in certain stages of *C. moschata* breeding.

The selected accessions displayed averages for accumulated degree-days for flowering (DDF) much lower than the general averages of the accessions and the controls. Notably, the accessions BGH-6749, BGH-5639, and BGH-219 expressed the lowest new predicted averages for DDF, making them the earliest-flowering accessions (Table 5). Regarding productivity of fruits (PF), the notably more promising accessions were BGH-4453, BGH-5653, BGH-5544A, BGH-4681A, BGH-5224A, and BGH-6587A, which expressed gains above 8 t ha-1 and new predicted averages for PF above 20 t ha^{-1} (Table 5). It should be highlighted that the BGH-5544A accession also expressed high averages for productivity of seeds (PS) and seed oil (SOP), corroborating the correlations of these characteristics with productivity of fruits (Figure 2).

The most promising accessions for total content of fruit pulp carotenoids (TC) were BGH-5455A and BGH-5598A (Table 5). These accessions expressed gains and new predicted averages for TC higher than 108.03 and 173.80 μg g^{-1} of fresh pulp mass, respectively, which were much higher than those of the controls. For the characteristics of seed and seed oil, it was found that the accessions BGH-4610A, BGH-5485A, and BGH-6590 were the most promising for productivity of seeds (PS) (Table 6). These accessions expressed gains and new predicted averages for PS of up to 0.31 and 0.58 t ha^{-1}, respectively. The most promising accessions for seed oil productivity (SOP) were BGH-5485A, BGH-4610A, and BGH-5472A, which had new predicted averages for SOP of 0.13 t ha^{-1}. It is worth highlighting that these accessions corresponded to those with higher PS, corroborating the strong correlation between productivity of seeds and seed oil productivity (Figure 2).

5. CONCLUSIONS

The accessions of *C. moschata* assessed in this study expressed high genetic variability for agro-morphological characteristics and for agronomic aspects related to the production of seeds such as number and mass of seeds per fruit, for accumulated degree-days for flowering, for total content of fruit pulp carotenoids, and for productivity of fruits, which allowed considerable gains to be obtained from selection for each of these characteristics.

The network of genetic correlations showed that higher fruit productivity in *C. moschata* might be achieved from the selection of aspects considered crucial in the production of this crop such as higher number of fruits per plant, and height and diameter of fruit. It also showed that greater seed productivity might be achieved with selection for a higher ratio of seed to fruit mass, number and mass of seeds per fruit; this information will assist in selection for higher productivity of fruit, seed and seed oil.

The clustering analysis resulted in 16 groups, with low similarity between the groups, which corroborates the variability of these accessions.

Grouping the averages of the clusters and identification per se allowed the most promising groups and accessions to be recognised for each characteristic, an approach that will guide the use of these accessions in breeding programmes.

Per se analysis identified the accessions BGH-6749, BGH-5639, and BGH-219 as those with the lowest averages for accumulated degree-days for flowering, highlighting them as the earliest flowering accessions. The most promising accessions for productivity of fruits were BGH-4453, BGH-5653, BGH-5544A, BGH-4681A, BGH-5224A, and BGH-6587A, with new predicted averages greater than 20 t ha^{-1}. The accessions with the highest averages for total content of fruit pulp carotenoids were BGH-5455A and BGH-5598A, with averages greater than 170.00 µg g^{-1} of fresh pulp mass. The accessions BGH-5485A, BGH-4610A, and BGH-5472A were the most promising for seed oil productivity, which, in the case of the former two, also corresponded to the highest averages for productivity of seeds. The accessions of *C. moschata* assessed in this study are a promising source for the genetic improvement of characteristics such as early flowering, total content of fruit pulp carotenoids, and productivity of seeds and seed oil.

ACKNOWLEDGMENTS

We are thankful to the Coordenação de Aperfeiçoamento de Pessoal e Nível Superior - Brasil (CAPES) - Finance Code 001. The authors are thankful to the National Council for Scientific and Technological Development–CNPq, for the doctorate scholarship (doctorate-GD grant), granted to the first author.

6. REFERENCES

1. Carvalho LMJ De, Gomes PB, Godoy RL De O, Pacheco S, Do monte PHF, De Carvalho JLV, Nutti MR, Neves ACL, Vieira ACRA, Ramos SRR. Total carotenoid content, α-carotene and β-carotene, of landrace pumpkins (*Cucurbita moschata* Duch): A preliminary study. Food Res Int. 2012; https://doi.org/10.1016/j.foodres.2011.07.040.
2. Ribeiro EMG, Chitchumroonchokchaib C, De Carvalho LMJ, De Moura FF, De Carvalho JLV, Faillab ML. Effect of style of home cooking on retention and bioaccessibility of pro-vitamin A carotenoids in biofortified pumpkin (*Cucurbita moschata* Duch.). Food Res Int. 2015; http://dx.doi.org/10.1016/j.foodres.2015.08.038.
3. Bergantin C, Maietti A, Tedeschi P, Font G, Manyes L, Marchetti N. HPLC-UV/Vis-APCI-MS/MS determination of major carotenoids and their bioaccessibility from "Delica" (*Cucurbita maxima*) and "Violina"(*Cucurbita moschata*) pumpkin as food traceability markers. Molecules. 2018; doi:10.3390/molecules23112791.
4. Priori D, Valduga E, Villela JCB, Mistura CC, Vizzotto M, Valgas RA, Barbieri RL. Characterization of bioactive compounds, antioxidant activity and minerals in landraces of pumpkin (*Cucurbita moschata*) cultivated in Southern Brazil. Food Sci Tech-Brazil. 2017; http://dx.doi.org/10.1590/1678-457x.05016.
5. Nagar A, Sureja A K, Kar A, Bhardwaj R, Krishnan SG, Munshi A Das. Profiling of mineral nutrients and variability study in pumpkin (*Cucurbita moschata*) genotypes. Agr Res. 2018; https://doi.org/10.1007/s40003-018-0329-3.
6. Food and Agriculture Organization of the United Nations- FAO STATE. 2019 [cited 15 October 2019]. In: FAO STATE [Internet]. Rome: FAO STATE 2019. Available from: http://www.fao.org.faostat/em/#home (2019).
7. Instituto Brasileiro de Geografia e Estatística- IBGE: Produção Agrícola Municipal. 2017 [cited 15 October 2019]. In: IBGE: Produção Agrícola Municipal [Internet]. Rio de Janeiro: IBGE 2017. Available from: https://sidra.ibge.gov.br/pesquisa/censo-agropecuario/censo-agropecuario-2017#horticultura.
8. Bouis H E, Welch RM. Biofortification- a sustainable agricultural strategy for reducing micronutrient malnutrition in the global south. Crop Sci. 2010; doi: 10.2135/cropsci2009.09.0531.
9. Saltzman A, Birol E, Bouis HE, Boy E, De Moura FF, Islam Y, Pfeiffer WH. Biofortification: progress toward a more nourishing future. Glob Food Secur. 2013; https://doi.org/10.1016/j.gfs.2012.12.003.

10. Sobreira FM. Divergência genética entre acessos de abóbora para estabelecimento de coleção nuclear e pré-melhoramento para óleo funcional. D. Sc. Thesis, Universidade Federal de Viçosa. 2013. Available from: https://www.locus.ufv.br/handle/123456789/1367.
11. Jarret RL, Levy I, Potter TL, Cermak SC, Merrick LC. Seed oil content and fatty acid composition in a genebank collection of *Cucurbita moschata* Duchesne and *C. argyrosperma* C. Huber. Genet Resour Crop Ev. 2013; doi: 10.1017/S1479262112000512.
12. Veronezi C M, Jorge N. Chemical characterization of the lipid fractions of pumpkin seeds. Nutr Food Sci. 2015; doi 10.1108/NFS-01-2014-0003.
13. Veronezi C, Jorge N. Bioactive compounds in lipid fractions of pumpkin (*Cucurbita* sp) seeds for use in food. J Food Sci. 2012; doi: 10.1111/j.1750-3841.2012.02736.x.
14. Gomes RS. Variabilidade morfoagronômica e nutricional de acessos de *Cucurbita moschata* Duch. do Banco de Germoplasma de Hortaliças – UFV. M. Sc. Thesis, Universidade Federal de Viçosa. 2017. Available from: https://www.locus.ufv.br/bitstream/handle/123456789/16359/texto%20completo.pdf?sequence=1&isAllowed=y.
15. Barboza N, Albertazzi FJ, Sibaja-Cordero JA, Mora-Umana F, Astorga C, Ramírez P. Analysis of genetic diversity of *Cucurbita moschata* (D.) germplasm accessions from Mesoamerica revealed by PCR SSCP and chloroplast sequence data. Sci Hortic-Amsterdam. 2012; https://doi.org/10.1016/j.scienta.2011.10.028.
16. Ferriol M, Picó B, Córdova PFDE, Nuez F. Molecular diversity of a germplasm collection of squash (*Cucurbita moschata*) determined by SRAP and AFLP markers. Crop Sci. 2004; doi:10.2135/cropsci2004.6530.
17. Wu J, Chang Z, Wu Q, Zhan H, Xie S. Molecular diversity of Chinese *Cucurbita moschata* germplasm collections detected by AFLP markers. Sci Hortic-Amsterdam. 2011; doi https://doi.org/10.1016/j.scienta.2010.12.006.
18. Hernández-Rosales HS, Castellanos-Morales G, Sánchez-de la Veja G, Aguirre-Planter E, Montes-Hernández S, Lira-Saade R, Eguiarte LE. Phylogeographic and population genetic analyses of *Cucurbita moschata* reveal divergence of two mitochondrial lineages linked to an elevational gradient. Am J Bot. 2020; doi:10.1002/ajb2.1424.

19. Borges RME, De Resende GM, De Lima MAC. Phenotypic variability among pumpkin accessions in the Brazilian Semiarid. Hortic Bras. 2011; http://dx.doi.org/10.1590/S0102-05362011000400003.
20. De Lima GKL, De Queiroz MA, Silveira L M da. Rescue of *Cucurbita* spp. germplasm in Rio Grande do Norte. Rev Caatinga. 2016; http://dx.doi.org/10.1590/1983-21252016v29n130rc.
21. Ferreira MG, Salvador FV, Lima MNR. Parâmetros genéticos, dissimilaridade e desempenho per se em acessos de abóbora. Hortic Bras. 2016; http://dx.doi.org/10.1590/s0102-053620160413.
22. Dillehay T, Rossen J, Andres TC, Williams DE. Preceramic Adoption of peanut, squash, and cotton in Northern Peru. Science. 2007; doi: 10.1126/science.1141395.
23. Piperno DR, Stothert K E. Phytolith evidence for early holocene cucurbita domestication in Southwest Ecuador. Science. 2003; doi: 10.1126/science.1080365.
24. Whitaker TW, Bemi WP. Cucurbits – cucumis, citrullus, cucurbita, lagenaria (Cucurbitaceae). In: Simmonds NW, editors. Evolution of crop plants. London: Longman Press; 1976. pp. 64-69.
25. Ara N, Nakkanong K, Lv W, Yang J, Hu Z, Zhang M. Antioxidant enzymatic activities and gene expression associated with heat tolerance in the stems and roots of two Cucurbit species ("*Cucurbita maxima*" and "*Cucurbita moschata*") and their interspecific inbred line "maxchata". Int J Mol Sci. 2013; doi 10.3390/ijms141224008.
26. Krajnc AU, Janzekovic I, Sober A, Ivanic A. The impact of interspecific hybridization on the chemical compositions of oil pumpkin seeds. Phyton-Ann Rei Bot A. 2016; doi 10.12905/0380.phyton56(1)2016-0061.
27. Karaagac O, Balkaya A. Interspecific hybridization and hybrid seed yield of winter squash (Cucurbita maxima Duch.) and pumpkin (*Cucurbita moschata* Duch.) lines for rootstock breeding. Sci Hortic- Amsterdam. 2013; doi http://dx.doi.org/10.1016/j.scienta.2012.10.021.
28. Uretsky J, Loy JB. Evaluation of morphological traits associated with productivity in F1 Interspecific (*Cucurbita maxima* Duch. x *C. moschata* Duch.) hybrid processing squash. Hortscience. 2017; doi 10.21273/HORTSCI12018-17.
29. The Second Report on The State of the World's Plant Genetic Resources for Food and Agriculture– FAO. 2010 [cited 21 October 2019]. In: FAO [Internet]. Rome: FAO 2010. Available from: http://www.fao.org/3/i1500e/i1500e00.htm.

30. Gomes RS, Almeida CF, Costa JRS, Machado Junior R, Delazzari F, Silva FCS, Silva DJH. Genetic diversity in sweet cassava from the Brazilian Middle North Region and selection of genotypes based on morpho-agronomical descriptors. Afr J Agr Res. 2016; doi: 10.5897/AJAR2016.11267.
31. Oliveira EJ D, Santana FA, de Oliveira LA, Santos V Da S. Genotypic variation of traits related to quality of cassava roots using affinity. Sci Agr. 2014; http://dx.doi.org/10.1590/0103-9016-2014-0043.
32. Rana JC, Sharma TR, Tyagi RK, Chahota RK, Gautam N, Singh M, Sharma PN, Ojha SN. Characterisation of 4274 accessions of common bean (*Phaseolus vulgaris* L.) germplasm conserved in the Indian gene bank for phenological, morphological and agricultural traits. Euphytica. 2015; doi 10.1007/s10681-015-1406-3.
33. Torres LG, Caixeta DG, Rezende WM, Schuster A, Azevedo CF, Silva FF, De Lima RO. Genotypic variation and relationships among traits for root morphology in a panel of tropical maize inbred lines under contrasting nitrogen levels. Euphytica. 2019; https://doi.org/10.1007/s10681-019-2373-x.
34. Fonseca MA, Queiroz MA de, Lopes JF, Assis JG de A, Romão RL, Ramos SRR, Do Carmo CAS do, Moura M da C, Dias R de C, Da Silveira LM, Barbieri R, Da Silva DJH, Noronha S. Geographical distribution and conservation of *Cucurbita* in Brazil. Magistra. 2015; 27: 432-442.
35. Silva DJH, Moura MCC, Casali VWD. Recursos genéticos do banco de germoplasma de hortaliças da UFV: histórico e expedições de coleta. Hortic Bras. 2001; http://dx.doi.org/10.1590/S0102-05362001000200002.
36. Moura MCCL, Zerbini FM, Silva DJH da, De Queiroz MA. Reação de acessos de Cucurbita sp. ao *Zucchini yellow mosaic virus* (ZYMV). Hortic Bras. 2005; 23: 206-210.
37. Lima Neto IS. Pré- melhoramento de abobra (*Cucurbita moschata*) visando a biofortificacao em carotenoides. D. Sc. Thesis, Universidade Federal de Viçosa. 2013. Available from: https://www.locus.ufv.br/handle/123456789/1367.
38. Filgueira FAR. Novo manual de olericultura: agrotecnologia moderna na produção e comercialização de hortaliças. 3rd ed. Viçosa: Universidade Federal de Viçosa Press; 2013.
39. Federer WT. Augmented or (Hoonuiaku) designs. Hawaiian Planter's Record. 1956; 55: 191-208.

40. Itle RA, Kabelka EA. Correlation between L* a* b* color space values and carotenoid content in pumpkins and squash (*Cucurbita* spp.). Hortscience. 2009; 44: 633-637.
41. Bates D, Maechler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. J. Stat. Softw. 2015; https://doi.org/10.18637/jss.v067.i01.
42. Cullis BR, Smith AB, Coombes NE. On the design of early generation variety trials with correlated data. J Agric Biol Envir S. 2006; doi: 10.1198/108571106X154443.
43. Cruz CD. GENES - a software package for analysis in experimental statistics and quantitative genetics. Acta Sci-Agron. 2013; doi: 10.4025/actasciagron.v35i3.21251.
44. Bodenhofer U, Kothmeier A, Hochreiter S. APCluster: an R package for affinity propagation clustering. Bioinformatics. 2011; doi: 10.1093/bioinformatics/btr406.
45. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2018; URL http://www.R-project.org/.
46. Frey BJ, Dueck D. Clustering by passing messages between data points. Science. 2007; doi: 10.1126/science.1136800.
47. Pagès J. Analyse factorielle de données mixtes. Rev Stat Appl. 2004; 52: 93-11.
48. Lê S, Josse J, Husson F. FactoMineR: an R package for multivariate analysis. J Stat Softw. 2008; 25: 1-18.
49. Darrudi VN, Soltania F, Shokrpoura M, Ercolanob MR. Genetic diversity *of Cucurbita pepo* L. and *Cucurbita moschata* Duchesne accessions using fruit and seed quantitative traits. J Appl Res Med Arom Plant. 2018; doi.org/10.1016/j.jarmap.2017.11.003.
50. Laurindo RDF, Laurindo BS, Delazari FT, Carneiro PC De S, Da Silva DJH. Potencial de híbridos e populações segregantes de abóbora para teor de óleo nas sementes e plantas com crescimento do tipo moita. Ver Ceres. 2017; http://dx.doi.org/10.1590/0034-737x201764060004.
51. Mohsin GM, Islam MS, Rahman MS, Ali L, Hasanuzzaman M. Genetic variability, correlation and path coefficients of yield and its components analysis in pumpkin (*Cucurbita moschata* Duch ex Poir). Int J Agr Res Innov Technol. 2017; doi:http://www.ijarit.webs.com.
52. Resende MDV. Delineamento de experimentos de seleção para maximização da acurácia seletiva e do progresso genético. Rev Arvore. 1995; 19: 4, 479- 500.
53. Schmidt P, Hartung J, Rath J, Piepho HP. Estimating broad-sense heritability with unbalanced data from agricultural cultivar trials. Crop Sci. 2019; doi: 10.2135/cropsci2018.06.0376.

54. Heino M. Quantitative traits. In: Cadrin SX, Kerr RA, Mariani S, editors. Stock identification methods, 2nd edition. Cambridge: Academic Press; 2014.
55. Hashemi S, Khadivi A. Morphological and pomological characteristics of white mulberry (*Morus alba* L.) accessions. Sci Hortic-Amsterdam. 2020; https://doi.org/10.1016/j.scienta.2019.108827.
56. Dias F dos S, Dos Santos FC, Gomes RS, da Silva DC, Soares D de A, Rosa DP, da Silva AF. Path analysis in fruits of the buriti (*Mauritia flexuosa* L. f) in populations from the Brazilian Middle North Region. Tree Genet Genomes. 2017; https://doi.org/10.1007/s11295-017-1205-1.
57. Maurya KN, Pal PK, Shukla S. Relationship of opium yield with yield contributing traits in segregating populations derived through biparental mating in opium poppy (Papaver somniferum L.). Ind Crop Prod. 2019; 139, 1-9. https://doi.org/10.1016/j.indcrop.2019.111557.
58. Hallauer AR, Carena MJ, Miranda Filho JB. Quantitative Genetics in Maize Breeding. 1 st edition. New York, Dordrecht, Heidelberg and London: Springer Press; 2010.
59. Cruz CD, Regazzi AJ, Carneiro PCS. Modelos biométricos aplicados ao melhoramento genético, volume 1. 4th edition. Viçosa: Universidade Federal de Viçosa Pres; 2012.
60. Wright S. Correlation and causation. J Agric Res. 1921; 20: 557-585.
61. Velho LPS, Gemeli MS, Trevisane N, Pereira TCV, Cerutti PH, De Mel RC, Guidolin AF, Coimbra JLM, Corrêa SC. Phenotypic correlation and direct and indirect effects of aerial part components with root distribution of common bean. Pesqui Agropecu Bras. 2017; doi: 10.1590/S0100-204X2017000500006.
62. Silva TN, Moro GV, Moro FV, dos Santos DMM, Buzinaro R. Correlation and path analysis of agronomic and morphological traits in maize. Ver Cienc Agron. 2016; doi: 10.5935/1806-6690.20160041.
63. Nascimento-Júnior I, Môro GV, Môro FV. Indirect selection of maize genotypes based on associations between root agronomic and anatomical characters. Chil J Agr Res. 2018; doi: 10.4067/S0718-58392018000100039.
64. Gurmu F, Shimelis H, Laing MD. Correlation and path-coefficient analyses of root yield and related traits among selected sweet potato genotypes. S Afr J Plant S. 2017; http://dx.doi.org/10.1080/02571862.2017.1354405.

65. MratiNić E, PoPoVSki B, MilošEVić T, PoPoVSka M. Evaluation of apricot fruit quality and correlations between physical and chemical attributes. Czech J Food Sci. 2011; 29: 161-170.

66. Do Rêgo ER, Do Rêgo M M, Cruz CD, Finger FL, Casali VW D. Phenotypic diversity, correlation and importance of variables for fruit quality and yield traits in Brazilian peppers (*Capsicum baccatum*). Genet Resour Crop Ev. 2011; doi 10.1007/s10722-010-9628-7.

67. Singh N, Sidhu MK, Dhatt A. Heterosis and combining ability for bushy and butternut traits in pumpkin (*Cucurbita moschata*). Indian J Agr Sci. 2018; 88: 877-883.

68. Kumar V, Mishra DP, Yadav GC, Yadav S. Exploitation of heterobeltiosis and economic heterosis for horticultural yield, and its attributes and biochemical traits in pumpkin (*Cucurbita moschata* Duch. ex. Poir) under salt affected soil. Curr Sci India. 2018; 115: 1550-156.

69. Albuquerque HYG, Oliveira EJ, Brito AC, de Andrade LRB, Carmo CD do, Morgante CV, Vieira EA, Moura E, Faleiro FG. Identification of duplicates in cassava germplasm banks based on single-nucleotide polymorphisms (SNPs). Sci Agr. 2019; doi: http://dx.doi.org/10.1590/1678-992X-2017-0389.

70. Marconi G, Ferradini N, Russi L, Concezzi L, Veronesi F, Albertini W. Genetic characterization of the apple germplasm collection in central Italy: the value of local varieties. Front Plant Sci. 2018; doi: 10.3389/fpls.2018.01460.

71. Raatz B, Mukankusi C, Lobaton JD, Male A, Chisale V, Amsalu B, Fourie D, Mukamuhirwa F, Muimu K, Mutari B, Nchimbi-Msolla S, Nkalubo S, Tumsa K, Chirwa R, Maredia MK, He C. Analyses of African common bean *(Phaseolus vulgaris* L.) germplasm using a SNP fingerprinting platform: diversity, quality control and molecular breeding. Genet Resour Crop Ev. 2019; doi: https://doi.org/10.1007/s10722-019-00746-0(0123456789().,-volV)(01234567.

72. Valois ACC, Salomão AN, Aliem AC. Glossário de recursos genéticos vegetais. Brasília: Embrapa Cenargen Press; 1996.

73. Moura MCCL. Identificação de fontes de resistência ao potyvirus ZYMV e diversidade genética e ecogeográfica em acessos de abóbora. D. Sc. Thesis. Universidade Federal de Viçosa. 2003. Available from: https://www.locus.ufv.br/handle/123456789/10265.

74. Brasil. 2019 [cited 20 October 2019]. In: Registro Nacional de Cultivares– RNC [Internet]. Brasília: Registro Nacional de Cultivares– RNC 2019. Available from: http://sistemas.agricultura.gov.br/snpc/cultivarweb/cultivares_registradas.ph.
75. Restrepo-Salazar JA, Vallejo-Cabrera FA, Restrepo-Salazar EF. Combining ability as a function of inbreeding for fruit traits in *Cucurbita moschata* Duch. ex Poir. Rev Fac Nac. Agron. 2019; doi: 10.15446/rfnam.v72n3.77320.
76. Singh N, Sidhu MK, Dhatt A. Heterosis and combining ability for bushy and butternut traits in pumpkin (*Cucurbita moschata*). Indian J Agr Sci. 2018; 88: 877-883.
77. Azevedo-Meleiro CH, Rodriguez-Amaya DB. Qualitative and quantitative differences in carotenoid composition among *Cucurbita moschata*, *Cucurbita maxima*, and *Cucurbita pepo*. J Agr Food Chem. 2007; 55: 4027-4033.
78. De Carvalho LMJ, Smiderle L de ASM, De Carvalho JLV, Cardoso F de SN, Koblitz MGB. Assessment of carotenoids in pumpkins after different home cooking conditions. Food Sci Tech-Brazil. 2014; doi http://dx.doi.org/10.1590/fst.2014.0058.
79. Nakkanong K, Yang JH, Zhang MF. Carotenoid accumulation and carotenogenic gene expression during fruit development in novel interspecific inbred squash lines and their parents. J Agr Food Chem. 2012; doi: https://doi.org/10.1021/jf3007135.
80. Rodriguez-Amaya DB, Kimura M, Godoy HT, Amaya-Farfan J. Updated Brazilian database on food carotenoids: Factors affecting carotenoids composition. J Food Compos Anal. 2008; doi: https://doi.org/10.1016/j.jfca.2008.04.001.
81. Bohn T, Desmarchelier C, El SN, Keijer J, Schothorst EV, Rühl R, Borel P. β-carotene in the human body: metabolic bioactivation pathways – from digestion to tissue distribution and excretion. P Nutr Soc. 2019; doi:10.1017/S0029665118002641.
82. Green AS, Fascetti AJ. Meeting the vitamin a requirement: the efficacy and importance of β-carotene in animal species. The Scientific World J. 2016; doi: http://dx.doi.org/10.1155/2016/7393620.
83. Gammone MA, Riccioni G, D’Orazio N. Carotenoids: potential allies of cardiovascular health? Food Nutr Res. 2015; doi:http://dx.doi.org/10.3402/fnr.v59.26762.
84. Jayedi A, Rashidy-Pour A, Parohan M, Zargar MS, Shab-Bidar S. Dietary antioxidants, circulating antioxidant concentrations, total antioxidant capacity, and risk of all-cause mortality: a systematic review and dose-response meta-analysis of prospective observational studies. Adv Nutr. 2018; doi:https://doi.org/10.1093/advances/nmy040.

85. Skeaff CM, Miller J. Dietary fat and coronary heart disease: summary of evidence from prospective cohort and randomised controlled trials. Ann Nutr Metab. 2009; doi: 10.1159/000229002.
86. Hooper L, Summerbell CD, Thompson R, Sills D, Roberts FG, Moore HJ, Smith GD. Reduced or modified dietary fat for preventing cardiovascular disease. Cochrane Db Syst Rev. 2012; doi: 10.1002/14651858.CD002137.pub3.
87. Mozaffarian D, Micha R, Wallace S. Effects on coronary heart disease of increasing polyunsaturated fat in place of saturated fat: a systematic review and meta-analysis of randomized controlled trials. Plos One. 2010; doi: https://doi.org/10.1371/journal.pmed.1000252.
88. Patel S. Pumpkin (*Cucurbita* sp.) seeds as nutraceutic: A review on status quo and scopes. Mediterr J Nutr Metab. 2013; doi: https://doi.org/10.1007/s12349-013-0131-5.
89. Wu T, Cao J. Differential gene expression of tropical pumpkin (*Cucurbita moschata* Duchesne) bush mutant during internode development. Sci Hortic- Amsterdam. 2008; doi https://doi.org/10.1016/j.scienta.2008.04.002.

ARTIGO II

IDENTIFICATION OF HIGH SEED OIL YIELD AND HIGH OLEIC ACID CONTENT IN BRAZILIAN GERMPLASM OF WINTER SQUASH (*Cucurbita moschata* D.)

ARTIGO II - IDENTIFICATION OF HIGH SEED OIL YIELD AND HIGH OLEIC ACID CONTENT IN BRAZILIAN GERMPLASM OF WINTER SQUASH (*Cucurbita moschata* D.)

Em processo de revisão na Saudi Journal of Biological Sciences- https://www.journals.elsevier.com/saudi-journal-of-biological-sciences.

Ronaldo Silva Gomes[a*], Ronaldo Machado Júnior[a], Cleverson Freitas de Almeida[a], Rebeca Lourenço de Oliveira[a], Rafael Ravaneli Chagas[a], Ednângelo Duarte Pereira[a], Fabio Teixeira Delazari[b], Derly José Henriques da Silva[a]

[a] Agronomy Department, Federal University of Viçosa-UFV, PH Rolfs Avenue, Zip Code 36570-000, Viçosa, MG, Brazil

[b] National Service of Rural Learning- SENAR, Marcino dos Santos Street, number 401, Zip Code 79042-140, Campo Grande, MS, Brazil

***Corresponding author**: PH Rolfs Avenue, Genetic Resources Management Laboratory, Zip Code 36570-900, Viçosa, MG, Brazil, ronaldo.s.gomes@ufv.br, +55 31 99751-6432

NFP: number of fruits per plant; MSF: total mass of seeds per fruit: PS: productivity of seeds; SOC: total seed oil content; SOP: seed oil productivity; UFV: Federal University of Viçosa; BLUPs: best non-biased linear predictions; BLUEs: best non-biased linear estimates; PUFAs: polyunsaturated fatty acids; RML: restricted maximum likelihood; MUFA: monounsaturated fatty acids

ABSTRACT

Cucurbita moschata D. seed oil contains approximately 75% unsaturated fatty acids, with high levels of monounsaturated fatty acids and antioxidant compounds such as vitamin E and carotenoid, constituting a promising food in nutritional terms. Associated to this, the Brazilian germplasm of *C. moschata* exhibits remarkable variability, representing an important source for the genetic breeding of this vegetable and other cucurbits. In this context, the present study evaluated the productivity and profile of the seed oil of 91 *C. moschata* accessions from different regions of Brazil and maintained in the Vegetable Germplasm Bank of the Federal University of Viçosa (BGH-UFV). A field experiment was conducted between January and July 2016. The tested *C. moschata* accessions showed high genetic variability in terms of characteristics related to seed oil productivity (SOP), such as the mass of seeds per fruit and productivity of seeds, providing predicted selection gains of 29.39 g and 0.26 t ha^{-1}, respectively. Based on the phenotypic and genotypic correlations, greater SOP can be achieved while maintaining high oleic acid content and low linoleic acid content, providing oil of better nutritional and chemical quality. In variability analysis, the accessions were clustered into five groups, which presented different averages for SOP and fatty acid content of seed oil; approach that will guide the use of appropriate germplasm in programs aimed at genetic breeding for SOP and seed oil profile. *Per se* analysis identified BGH-4610, BGH-5485A, BGH-6590, BGH-5556A, BGH-5472A, and BGH-5544A as the most promising accessions in terms of SOP, with average (μ+g) of approximately 0.20 t ha^{-1}. The most promising accessions for higher oleic acid content of seed oil were BGH-5456A, BGH-3333A, BGH-5361A, BGH-5472A, BGH-5544A, BGH-5453A, and BGH-1749, with average (μ+g) of approximately 30%, and almost all of these accessions were also the most promising in terms of lower linoleic acid content of seed oil, with average (μ+g) of approximately 45%. Overall, part of the *C. moschata* accessions evaluated in the present study can serve as a promising resource in genetic breeding programs for SOP and fatty acid profile, aiming at the production of oil with better nutritional and physicochemical quality.

KEYWORDS: bioactive compounds, genetic correlation, clustering, *Cucurbita moschata*, genetic parameters, seed oil.

1 INTRODUCTION

Winter squash (*Cucurbita moschata* D.) is one of the cucurbit vegetables of great socioeconomic importance owing to the high nutritional value of its fruits and seeds. Cultivated mainly for fruit production, *C. moschata* has been strategically used in biofortification programs for vitamin A because of its high content of carotenoids in fruits such as *β*- and *α*-carotene—the major precursors of vitamin A (Carvalho *et al.*, 2012; Saltzman *et al.*, 2013). In addition, winter squash fruits are an excellent source of minerals such as K, Ca, P, Mg, and Cu (Nagar *et al.*, 2018; Priori *et al.*, 2018). *C. moschata* is cultivated across a wide geographical range worldwide, and together with other cucurbits such as *C. pepo* and *C. maxima*, the area under the cultivation and worldwide production of *C. moschata* was estimated to be nearly 2 million hectares and 27.6 million tons, respectively, in 2018 (FAO, 2020), highlighting the socioeconomic importance of this vegetable.

Furthermore, *C. moschata* seed oil can serve as an excellent product due to its nutritional and physicochemical properties, associated with the high seed production potential of this vegetable. Lipids account for up to 49% of *C. moschata* seed components (Jarret *et al.*, 2013; Patel, 2013), and studies on the germplasm of this cucurbit have already identified accessions that can produce up to 0.58 $t \cdot ha^{-1}$ of seeds (Gomes *et al.*, 2020). *C. moschata* seed oil contains approximately 75% unsaturated fatty acids, with high content of monounsaturated fatty acids (MUFAs), such as oleic acid (Jarret *et al.*, 2013; Sobreira, 2013; Veronezi and Jorge 2015). Thus, it is an excellent substitute for vegetable lipid sources that contain high levels of saturated fatty acids, which are harmful to human health. Associated with this, some studies have reported that *C. moschata* seeds and seed oil contain high levels of antioxidant compounds, such as vitamin E and carotenoids, components beneficial to human health (Veronezi and Jorge 2012; Dash *et al*., 2017), which also protect the seed oil from oxidative processes that may lead to rancidity. In this line, *C. moschata* seed oil may serve as a health food in the cultivation regions of this vegetable, particularly in less developed regions, and in the family farming context (Gomes *et al*., 2020).

Studies on the Brazilian germplasm of *C. moschata* have emphasized the evaluation of agromorphological characteristics reporting remarkable variability in these traits (De Lima *et al*., 2016; Ferreira *et al*., 2016; Oliveira *et al*., 2016; Gomes *et al*., 2020). As an allogamous species, variability in the *C. moschata* germplasm is associated with the occurrence of natural hybridization across different populations. Already present in the diet of native Latin American people (Dillehay *et al*., 2007; Piperno *et al*., 2003), and with a widespread cultivation in the American continent, the variability of *C. moschata* may also be related to anthropogenic actions, such as frequent exchange of seeds among family farmers (Gomes *et al*., 2020). Additionally, the variability of the Brazilian germplasm of *C. moschata* reflects its adaptation to a broad ecological range, constituting different edaphoclimatic conditions; thus, these accessions represent an important source for the genetic breeding of this vegetable and other cucurbits.

The Vegetable Germplasm Bank of the Federal University of Viçosa (BGH-UFV) maintains approximately 350 accessions of *C. moschata*, mostly landraces, with a collection period of over five decades, from different geographic regions of Brazil (Silva *et al*., 2001). The *C. moschata* collection maintained in the BGH-UFV constitutes a substantial sample of the Brazilian germplasm, being one of the largest collections of this species in the country (Fonseca *et al*., 2015). A preliminary assessment of the seed oil fatty acid profile of a small part of the *C. moschata* germplasm maintained in the BGH-UFV revealed high variability among 54 accessions in terms of oleic acid content (Sobreira, 2013). In that study, BGH-7765, with oil oleic acid content of 28.39%, was identified as a promising accession for use as parent germplasm in breeding programs aimed at improving *C. moschata* seed oil.

To this end, the objectives of the present study were (a) to evaluate the seed oil productivity (SOP) and oil fatty acid profiles of 91 *C. moschata* accessions from different regions of Brazil maintained in the BGH-UFV; (b) to analyze the correlations between these characteristics; and (c) to examine the variability of this germplasm for identifying accessions with high SOP and high oleic fatty acid content but low linoleic acid content in seed oil.

2 MATERIAL AND METHODS

2.1 Germplasm origin

The present study evaluated 91 *C. moschata* accessions maintained in the BGH-UFV. These accessions, mostly landraces, have been collected by the BGH-UFV over a period of more than five decades (Silva *et al.*, 2001) from different geographical regions of Brazil (Figure 1). The accessions were evaluated with four genotypes used as controls, the cultivars 'Jacarezinho' and 'Maranhão' and the hybrids Jabras and Tetsukabuto, which are widely cultivated and commercialized in Brazil.

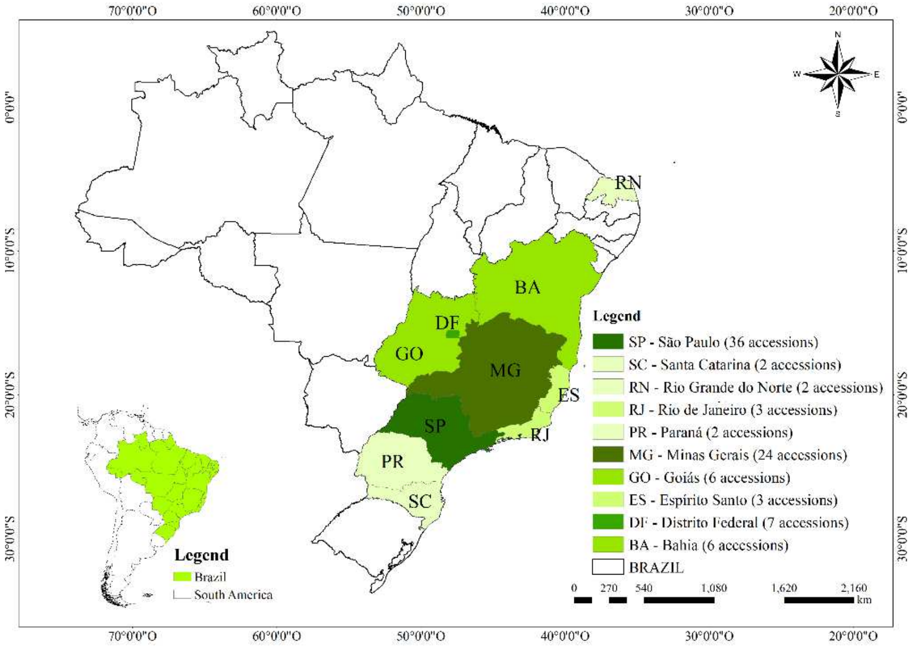


Figure 1. Brazilian map displaying the regions of origin of the *Cucurbita moschata* accessions tested in the present study.

2.2 Location and conduct of the experiment

A field experiment was conducted between January and July 2016 at "Horta Velha"—an experimental unit of the Department of Agronomy of UFV (20°45′24″S, 42°50′45″W; altitude, 648.74 m).

The genotype seedlings were cultivated in expanded polystyrene trays with 72 cells containing a commercial substrate. Subsequently, the seedlings were transplanted in the experimental area following to the augmented block design proposed by Federer (1956), with five replicates for each control. The plants were distributed with 3 × 3 m spacing between plants and rows, resulting in a stand of 1,111 plants ha^{-1}. Each experimental plot contained five plants, and all evaluations of fruits, seeds, and seed oil were performed on three fruits each from three central plants in a plot.

2.3 Assessment of seed oil productivity

The genotypes were evaluated for the number of fruits per plant (NFP), total mass of seeds per fruit (MSF), productivity of seeds (PS), and total seed oil content (SOC). The PS, SOC, and SOP estimates were obtained using the following equations (equations 1, 2, and 3, respectively):

$$PS = NFP \times MSF \times 1{,}111^{*} \tag{1}$$

$$SOC = \frac{\text{Weight of oil extracted from the seed sample (mg)}}{\text{Total weight of the seed sample (mg)}} \times 100 \tag{2}$$

$$SOP = \frac{\text{PS} \times \text{SOC}}{100} \tag{3}$$

where PS is the productivity of seeds (t ha^{-1}); NFP is the number of fruits per plant; MSF is the mass of seeds per fruit (g); *1,111 is the number of plants per hectare; SOC is the oil content expressed as the percentage of the seed mass on a dry basis (%); and SOP is the seed oil productivity (t ha^{-1}).

Initially, the seeds were dried in a forced air circulation oven for 72 h at 30°C. Next, 20 g seeds from each genotype were ground in a Willey knife mill with a 1 mm sieve. SOC was determined in an extractor (ANKOM XT15) following extraction with petroleum ether using a standard AOAC method (Thiex *et al.*, 2003). Before loading in the extractor, the ground seed samples were dried in an unventilated oven at 105°C for 2 h. Then, approximately 2 g samples were transferred to filter envelopes (XT4; ANKOM technology), which were sealed and placed in the extractor. Oil was extracted from the samples with ether circulation for 30 min at 90°C, and the percentage of oil was calculated as the difference between the sample weight before and after extraction. SOC was expressed in grams per 100 grams of seeds on a dry basis.

2.4 Analysis of the fatty acid profile of seed oil

Seed oil extracted by cold pressing was subjected to gas chromatography. Oil was extracted from approximately 50 g of seeds using a hydraulic press aid according to the methodology described by Gomes *et al*. (2020).

The composition of the methyl esters of oil fatty acids was determined according to the methodology described by Bubeck *et al*. (1989), with some modifications. The Shimadzu GC-17A gas chromatograph equipped with an automatic insertion platform, a flame ionization detector, and a Carbowax capillary column (30 m × 0.25 nm) was used. Chromatography was performed at an injector temperature of 230°C and detector temperature of 250°C. The column operation was programmed to start at 200°C, with an increase of $3°C{\cdot}min^{-1}$ until reaching the final temperature of 225°C. Nitrogen was used as the carrier gas at a flow rate of $1.3\ L{\cdot}min^{-1}$. The content of each methyl ester of fatty acids was expressed as a percentage of relative peak area.

2.5 Obtaining of best non-biased linear predictions (BLUPs), best non-biased linear estimates (BLUEs), variance components, and genetic-statistical parameters

Data were analyzed using a mixed model based on the predictions of restricted maximum likelihood (RML) and BLUPs. The “lme4” package in R version 3.6.1 was used (Bates *et al*., 2015). The variance components were obtained based on RML, and the genotypic values were obtained based on BLUPs and BLUEs using the following model:

$$y = Wb + Xa + Zt + e$$

where y represents the vector comprising the phenotypic values of variables; b represents the vector comprising the effect of blocks (random effect); a represents the vector comprising the effect of accessions (random effect); t represents the vector comprising the control effect (fixed effect); and e represents the error vector. W, X, and Z represent the incidence matrices of parameters b, a, and t, respectively, with the data vector y. All statistical analyses were performed based on the genotypic values.

The genetic parameters were obtained based on the following estimators:

$h^2 = 1 - (Pev/2\sigma^2_g)$

where Pev represents the prediction of error variance (Cullis *et al.*, 2006)

$$A = \sqrt{1 - (Pev/\ \sigma_g^2)}$$

$DS = h^{2*}$

where DS represents the selection differential, estimated from an average of 15% of the most promising accessions for each characteristic.

$CV_g\,\% = (\sigma_g/\mu) \times 100$

$CV_p\,\% = (\sigma_p/\mu) \times 100$

$CVr\,\% = (\sigma/\mu) \times 100.$

2.6 Analysis of correlations between characteristics

The correlations between characteristics were analyzed based on the following model:

$$rg = Cov\ (x, y)/\sqrt{\sigma_g^2(x)\ \sigma_2^2(y)}$$

where $Cov\ (x, y)$ represents the genetic covariance between two variables, X and Y, and $\sigma^2_g\ (x)$ and e $\sigma^2_g\ (y)$ represent the genetic variances corresponding to the variables X and Y, respectively.

The significance of correlations was analyzed in GENES (Cruz, 2013) using the Mantel test (Z statistic) at 1% and 5% probability.

2.7 Analysis of the germplasm variability

The matrix of distances between the genotypes was obtained based on the genotypic values of characteristics related to SOP and seed oil profile. The distances between genotypes were estimated as the average Euclidean distance with data standardization. Based on this, variability was assessed using Tocher's clustering in GENES (Cruz, 2013).

Principal component analysis was used to identify the contribution of the characteristics to genotype variability using GENES (Cruz, 2013).

2.8 Identification of promising accessions

Promising accession clusters were identified, and *per se* analysis of the most promising accessions for each characteristic was performed. *Per se* identification was performed based on the ranking of the respective genotypic values and the environmental interaction-free genotypic value (μ+g), considering 15% of the most promising accessions.

3. RESULTS

3.1 Variance components and genetic-statistical parameters of characteristics related to SOP and fatty acid profiles

MSF was the characteristic associated with SOP with the highest genotypic variance (Table 1). Regarding oil fatty acid profile, oleic acid content exhibited the highest genotypic variance. Linoleic acid and polyunsaturated fatty acid (PUFA) content exhibited genotypic variances of 10.66% and 10.65%, respectively. All these variances were significant (p <0.01), as shown in Table 1.

Most of the tested characteristics showed very high heritability (>0.70) (Table 1), according to the classification of Resende (1995). The heritability estimates for PS, SOP, oleic acid content, and linoleic acid content were 0.705, 0.667, 0.857 and 0.729, respectively.

The predicted selection gains for PS and SOP were 0.266 and 0.10 t ha^{-1}, respectively (Table 1). Among the characteristics related to oil fatty acid profile, oleic acid content achieved the greatest selection gain (6.99%), followed by linoleic acid content (-5.12%) and PUFA content (-5.11%), which also achieved considerable selection gains (Table 1). The phenotypic amplitude was up to 0.90 t ha^{-1} (phenotypic mean, 0.26 t ha^{-1}) for PS and up to 0.36 t ha^{-1} (phenotypic mean, 0.050 t ha^{-1}) for SOP (Table 1). Oleic acid content ranged from 16.01 to 40.18% (phenotypic mean, 24.55%), and linoleic acid content from 36.58 to 58.33% (phenotypic mean, 50.66%).

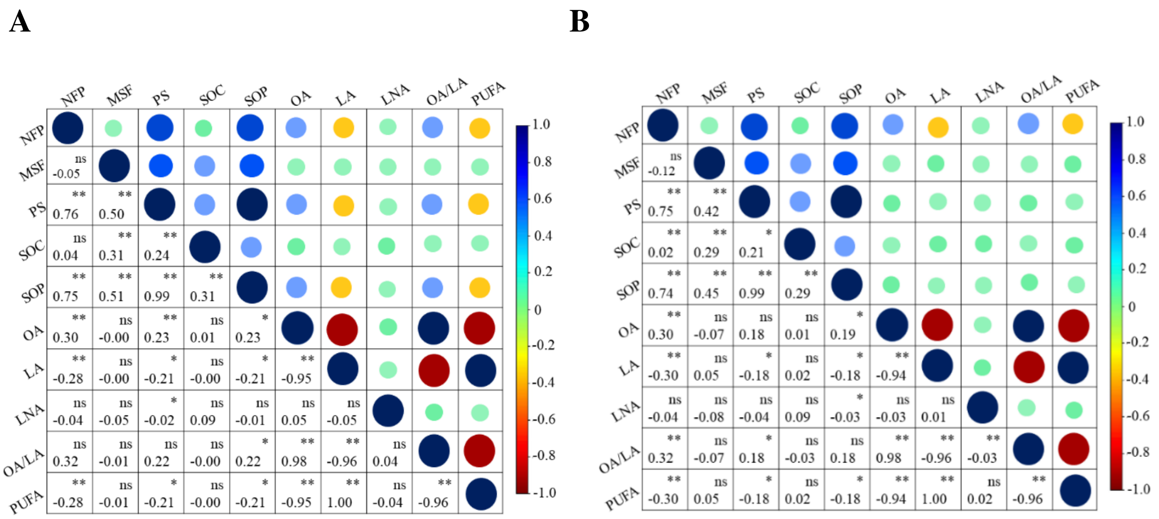
Table 1. Variance components and genetic-statistical parameters of characteristics related to seed oil productivity and fatty acid profiles

Fruit trait											
Trait	σ^2_p	σ^2_g	σ^2_b	σ^2_r	*A*	h^2	*SG*	*Phenotypic range*	μ	CV$_g$%	CV$_P$ %	CV*r* %
NFP	8.683	2.876^{ns}	0.585	5.222	0.600	0.361	1.961	1- 15	4.783	35.456	61.607	47.776
Seed traits											
Traits	σ^2_p	σ^2_g	σ^2_b	σ^2_r	*A*	h^2	*SG*	*Phenotypic Range*	μ	CV$_g$%	CV$_P$ %	CV*r* %
MSF	491.465	427.158**	16.394	47.913	0.943	0.891	29.399	4.4- 119.3	51.929	39.800	42.691	13.329
PS	0.042	0.025^{ns}	0.006	0.010	0.839	0.705	0.266	0.01- 0.9	0.269	58.778	76.185	37.174
SOC	11.840	1.988^{ns}	0.000	9.851	0.425	0.181	0.797	25.57- 48.89	18.516	7.614	18.583	16.950
SOP	0.007	0.004^{ns}	0.001	0.002	0.816	0.667	0.102	0.003- 0.36	0.050	126.49	167.332	89.442
Traits of seed oil profile											
Traits	σ^2_p	σ^2_g	σ^2_b	σ^2_r	*A*	h^2	*SG*	*Phenotypic range*	μ	CV$_g$%	CV$_P$%	CV$_r$ %
Palmitic	1.064	0.000^{ns}	0.077	0.987	0.000	0.000	0.000	12.04- 18.14	15.16	0.000	6.804	6.553
Estearic	1.159	0.000^{ns}	0.066	1.093	0.000	0.000	0.000	6.31- 12.3	9.60	0.000	11.214	10.890
Oleic	24.863	17.182**	5.268	2.412	0.925	0.857	6.999	16.01- 40.18	24.55	16.884	20.310	6.326
Linoleic	17.914	10.664**	3.629	3.620	0.853	0.729	-5.121	36.58- 58.33	50.66	6.446	8.354	3.755
Linolenic	0.001	0.001^{ns}	0.000	0.000	0.992	0.986	-0.013	0.01- 0.33	0.18	17.568	17.568	0.000
Oleic/linoleic	0.023	0.016^{ns}	0.004	0.001	0.937	0.879	0.232	0.27- 1.09	0.049	258.14	309.505	0.000
saturated	2.074	0.000^{ns}	0.151	1.923	0.000	0.000	0.000	21.03- 28.14	24.76	0.000	5.816	5.600
PUFA	17.906	10.653**	3.617	3.634	0.853	0.728	-5.116	36.60- 58.34	50.68	6.440	8.349	3.761

Number of fruits per plant (NFP), mass of seeds per fruit (MSF), productivity of seeds (PS), seed oil content (SOC), seed oil productivity (SOP), and polyunsaturated fatty acid (PUFA) content. Components of phenotypic (σ_p), genotypic (σ_g), and residual (σ) variances as well as variance associated with the block effect (σ_b). Accuracy (*A*), broad-sense heritability (h^2), selection gain (*SG*), phenotypic range, phenotypic average (μ), coefficient of genotypic (CV$_g$, %), phenotypic (CV$_P$, %), and residual (CV*r*, %) variation. Not significant (ns); $^{**}p < 0.01$ and $^{*}p < 0.05$, likelihood-ratio test.

3.2 Correlations of characteristics related to SOP and fatty acid profiles

For convenience, NFP, MSF, and PS were assumed to be directly related to SOP. The phenotypic correlations of these first variables with SOP ranged from 0.51 (SOP × MSF) to 0.99 (SOP × PS), with all correlations being significant ($p < 0.01$). The phenotypic correlations between SOP × PS (0.99) and SOP × NFP (0.75) were classified as very strong and strong, respectively, according to Shimakura and Ribeiro Júnior (2012). The genotypic correlations between characteristics directly related to SOP were similar to the phenotypic correlations in terms of direction and significance, although their magnitude was slightly smaller (Figure 2).



A

	NFP	MSF	PS	SOC	SOP	OA	LA	LNA	OA/LA	PUFA
NFP										
MSF	-0.05 ns									
PS	0.76 **	0.50 **								
SOC	0.04 ns	0.31 **	0.24 **							
SOP	0.75 **	0.51 **	0.99 **	0.31 **						
OA	0.30 **	-0.00 ns	0.23 **	0.01 ns	0.23 *					
LA	-0.28 **	-0.00 ns	-0.21 *	-0.00 ns	-0.21 *	-0.95 **				
LNA	-0.04 ns	-0.05 ns	-0.02 *	0.09 ns	-0.01 ns	0.05 ns	-0.05 ns			
OA/LA	0.32 ns	-0.01 ns	0.22 ns	-0.00 ns	0.22 *	0.98 **	-0.96 **	0.04 ns		
PUFA	-0.28 **	-0.01 ns	-0.21 *	-0.00 ns	-0.21 *	-0.95 **	1.00 **	-0.04 ns	-0.96 **	

B

	NFP	MSF	PS	SOC	SOP	OA	LA	LNA	OA/LA	PUFA
NFP										
MSF	-0.12 ns									
PS	0.75 **	0.42 **								
SOC	0.02 **	0.29 **	0.21 *							
SOP	0.74 **	0.45 **	0.99 **	0.29 **						
OA	0.30 **	-0.07 ns	0.18 ns	0.01 ns	0.19 *					
LA	-0.30 **	0.05 ns	-0.18 *	0.02 ns	-0.18 *	-0.94 **				
LNA	-0.04 ns	-0.08 ns	-0.04 ns	0.09 ns	-0.03 *	-0.03 ns	0.01 ns			
OA/LA	0.32 **	-0.07 ns	0.18 *	-0.03 ns	0.18 ns	0.98 **	-0.96 **	-0.03 **		
PUFA	-0.30 **	0.05 ns	-0.18 *	0.02 ns	-0.18 *	-0.94 **	1.00 **	0.02 ns	-0.96 **	

Figure 2. Phenotypic (A) and genotypic (B) correlations of characteristics related to seed oil productivity and fatty acids. Number of fruits per plant (NFP), mass of seeds per fruit (MSF), productivity of seeds (PS), seed oil content (SOC), seed oil productivity (SOP), oleic acid (OA) content, linoleic acid (LA) content, linolenic acid (LNA) content, OA:LA ratio, and polyunsaturated fatty acid (PUFA) content were measured.

SOP showed the strongest positive correlation with oleic acid content (0.23) and oleic acid:linoleic ratio (0.22) and strongest negative correlations with linoleic (-0.21) and PUFA (-0.21) content. SOP showed weak correlations with fatty acids, according to Shimakura and Ribeiro Júnior (2012), ranging from 0.20 to 0.39. Most of the genotypic correlations between SOP and oil fatty acids were similar to the phenotypic correlations in terms of direction and significance, although their magnitude was smaller (Figure 2).

Among fatty acids, linoleic and PUFA content showed the strongest positive correlation (1; p <0.01) (Figure 2). The correlation between the oleic content × oleic acid:linoleic ratio (0.98; $p < 0.01$) was classified as very strong (Figure 2), according to Shimakura and Ribeiro Júnior (2012). Oleic acid:linoleic ratio × linoleic acid content and oleic acid:linoleic ratio × PUFA acid content showed very strong negative correlations (both -0.96; $p < 0.01$). Similarly, oleic acid content × linoleic acid content showed very strong negative correlation (-0.95; p <0.01). Most of the genotypic correlations between fatty acids were similar to the phenotypic correlations in terms of direction, significance, and magnitude (Figure 2).

3.3 Clustering and variability of accessions based on characteristics related to SOP and fatty acid profiles

In clustering analysis, the tested accessions and controls formed five groups. Group 1 included 83 accessions (91.20%) and the controls Jabras, Jacarezinho, Maranhão, and Tetsukabuto. Group 2 comprised eight accessions (8.79%), corresponding to the second largest group. Group 3 comprised two accessions, and the remaining groups (4 and 5) comprised a single accession each (Table 2).

Group 5, formed by the accession BGH-4610A, presented the highest average (μ+g) value for SOP, estimated at 0.277 t ha^{-1}. Group 2, formed by the accessions BGH-5472A, BGH-5544A, BGH-5556A, BGH-5361A, BGH-5453A, BGH-3333A, BGH-6590, and BGH-5485A, also presented a high average (μ+g) value for SOP, estimated at 0.19 t ha^{-1} (Tables 2 and 3).

Regarding the fatty acid content of seed oil, group 3, formed by the accessions BGH-5456A and BGH-1749, presented the highest average (μ+g) value for oleic acid content (34.18%), followed by group 2 (30.62%). Group 3, formed by the accessions BGH-5456A and BGH-1749, presented the lowest average (μ+g) value for linoleic acid content (43.16%) (Tables 2 and 3).

Table 2. Tocher's clustering of accessions based on the genotypic values of characteristics related to seed oil productivity and oil fatty acid profiles

Clusters	**Accessions**
1	BGH-5224 BGH-5240 BGH-5560A BGH-5301 BGH-6099 BGH-315 BGH-5603 BGH-1004 BGH-5499A BGH-5548A BGH-95 BGH-5554A BGH-5593 BGH-5638 BGH-5494A BGH-5530A BGH-5551 BGH-5597 BGH-5606A BGH-5552 BGH-900 BGH-6115 BGH-4598A BGH-5596A BGH-4590A BGH-5247A BGH-6096 BGH-5616A BGH-5553 BGH-5493A BGH-7668 BGH-5451 BGH-5659A BGH-4607A BGH-1461A BGH-6594 BGH-6794 BGH-5559A Maranhão BGH-6749 BGH-4516 BGH-5497 BGH-1927 BGH-4681A GBH-5694 BGH-117 BGH-6587A BGH-6117A BGH-6595 BGH-291 BGH-5648 BGH-5455A BGH-6116 BGH-5528 BGH-5639 BGH-5248 BGH-4287A BGH-4454A BGH-5624A BGH-6593 BGH-5541 Tetsukabuto BGH-5442 BGH-5538 BGH-5591A BGH-5051 BGH-4459A Jabras BGH-1992 BGH-305A BGH-5630A BGH-7219A BGH-5598A BGH-4453 BGH-5466 Jacarezinho BGH-4281 BGH-5473A BGH-1961 BGH-5649A BGH-5653 BGH-6155 BGH-5440A
2	BGH-5472A BGH-5544A BGH-5556A BGH-5361A BGH-3333A BGH-5453A BGH-6590 BGH-5485A
3	BGH-1749 BGH-5456A
4	BGH-1945A
5	BGH-4610

Table 3. Grouping of genotypic averages of the accession groups based on Tocher's method of average grouping

Clusters	SOP	Oleic	Linoleic	PUFA
1	0.087 b	22.996 b	51.850 b	51.864 b
2	0.198 a	30.626 a	46.482 a	46.506 a
3	0.041 b	34.184 a	43.168 a	43.186 a
4	0.046 b	22.012 b	51.873 b	52.121 b
5	0.277 a	22.759 b	51.820 b	51.835 b

The genotypic averages followed by the same letters in the column do not differ from one another based on the Tocher's method of average grouping.

3.4 Principal components analysis

The first three principal components (PCs) explained 82.13% of total variation among accession in terms of characteristics related to SOP and fatty acid profiles. PC1 explained 47.16% of total variation. The oleic acid:linoleic acid ratio and oleic acid content were the characteristics with the highest positive loading, while linoleic acid content and PUFA content were the characteristics with the highest negative loading on PC1 (Figure 3).

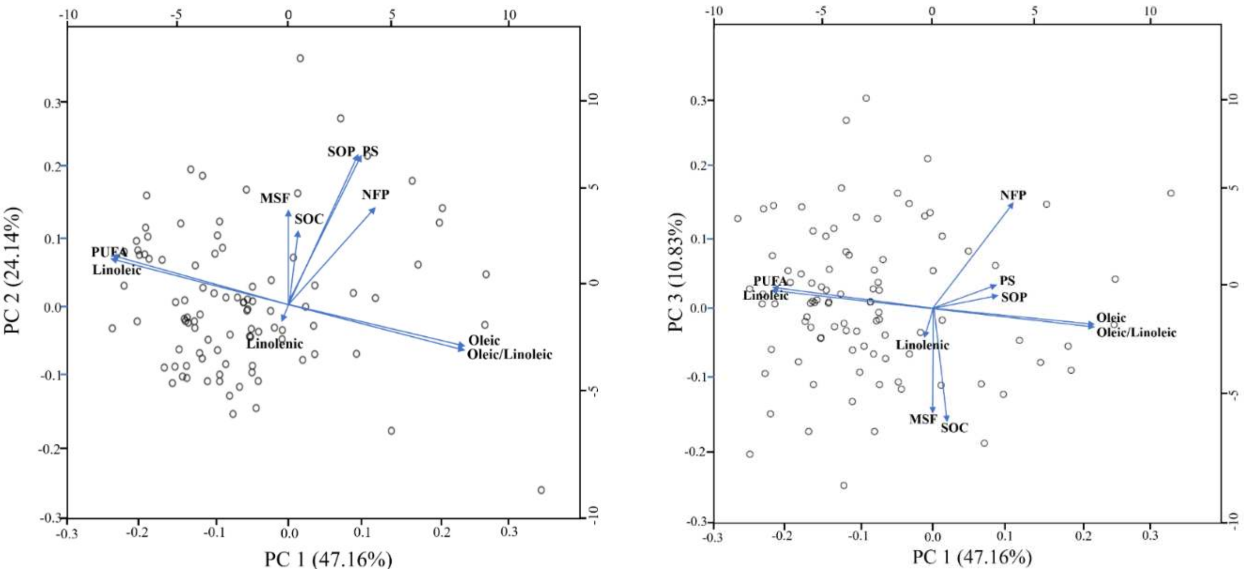


Figure 3. Dispersion of characteristics related to seed oil productivity and fatty acid profile in relation to the first three principal components. Number of fruits per plant (NFP), mass of seeds per fruit (MSF), productivity of seeds (PS), seed oil content (SOC), seed oil productivity (SOP), and polyunsaturated fatty acid (PUFA).

PC2 explained 24.14% of total variation, and PS, SOP, and MSF were the characteristics with the highest positive loading on this PC. PC3 explained 10.83% of total variation, and SOC and MSF were the characteristics with the highest loading on this PC.

Furthermore, PCA revealed the relationships between the characteristics. Along PC1, SOP was strongly and positively correlated with PS and NFP; moreover, along the same PC, there was a strong positive correlation between oleic acid content and oleic acid:linoleic acid ratio and a strong negative correlation between linoleic acid content and PUFA content (Figure 3).

3.5 Identification of promising accessions in terms of SOP and fatty acid profiles

Among the tested accessions, the (μ+g) estimate for SOP ranged from 0.14 to 0.27 t ha^{-1}, which was much higher than the general average for accessions (0.05 t ha^{-1}). Notably, the accessions BGH-4610A, BGH-5485A, BGH-6590, BGH-5556A, BGH-5472A, and BGH-5544A were the most promising in terms of SOP, with values close to 0.20 t ha^{-1} (Table 4).

Among fatty acids of seed oil, oleic acid content showed the highest (μ+g) amplitude (17.71 to 37.20%) (Figure 4). Associated with this, the (u+g) estimates for oleic acid content among the selected accessions ranged from 27.28 to 37.20% (Table 4). The accessions BGH-5456A, BGH-3333A, BGH-5361A, BGH-5472A, BGH-5544A, BGH-1749, BGH-5653, and BGH-5453A were the most promising in terms of oleic acid content, with values close to 30.00%.

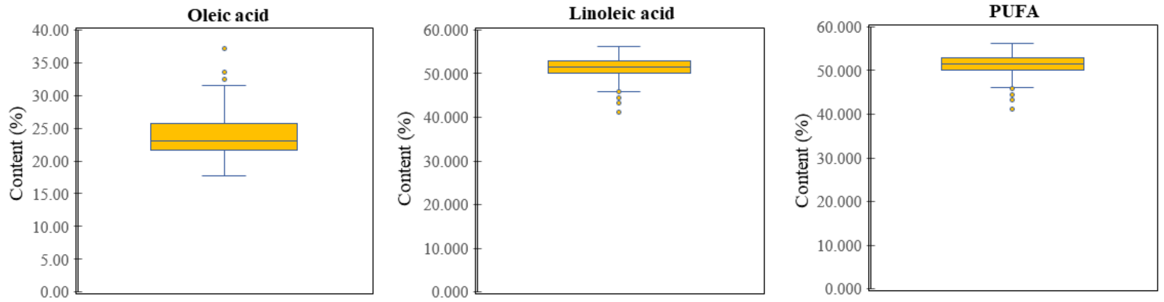


Figure 4. Box plots showing the variability in the oleic, linoleic, and polyunsaturated fatty acid content in the seed oil of the tested accessions.

Among the selected accessions, the (μ+g) estimate for linoleic content ranged from 41.09 to 48.77%. The accessions BGH-5456A, BGH-3333A, BGH-5361A, BGH-1749, BGH-5544A, and BGH-5556A expressed the lowest estimates for linoleic acid content, with values close to 40.00% (Table 4).

The (μ+g) estimates for oleic acid:linoleic acid ratio ranged from 1.00 to 0.57 among the selected accessions, and the accessions BGH-5456A, BGH-5361A, BGH-3333A, BGH-1749, BGH-5472A, and BGH-5544A expressed the lowest estimates, with values close to 0.75.

The (μ+g) estimates for PUFA content ranged from 41.12 to 48.78%, and the accessions with the lowest estimates were BGH-5456A, BGH-3333A, BGH-5361A, BGH-1749, and BGH-5544A, with values close to 40% (Table 4).

Table 4. Estimates of genotypic values (g) and environmental interaction-free genotypic values (μ+g) for seed oil productivity and fatty acid profiles

Accessions	**SOP**		Accessions	**Oleic**		Accessions	**Linoleic**	
	g	μ + g		g	μ + g		g	μ + g
BGH-4610A	0.227	0.277	BGH-5456A	12.651	37.201	BGH-5456A	-9.563	41.097
BGH-5485A	0.203	0.253	BGH-3333A	9.458	34.008	BGH-3333A	-7.354	43.306
BGH-6590	0.176	0.226	BGH-5361A	8.987	33.537	BGH-5361A	-6.129	44.531
BGH-5556A	0.174	0.224	BGH-5472A	7.965	32.515	BGH-1749	-5.421	45.239
BGH-5472A	0.166	0.216	BGH-5544A	6.913	31.463	BGH-5544A	-4.821	45.839
BGH-5544A	0.152	0.202	BGH-1749	6.617	31.167	BGH-5556A	-4.663	45.997
BGH-5440A	0.133	0.183	BGH-5653	5.904	30.454	BGH-5453A	-4.262	46.398
BGH-5630A	0.126	0.176	BGH-5453A	5.702	30.252	BGH-5472A	-4.021	46.639
BGH-4281	0.126	0.176	BGH-6155	5.128	29.678	BGH-6155	-3.976	46.684
BGH-5473A	0.126	0.176	BGH-5466	5.042	29.592	BGH-5653	-3.408	47.252
BGH-5361A	0.123	0.173	BGH-5556A	4.515	29.065	BGH-5648	-2.367	48.293
BGH-5453A	0.103	0.153	BGH-6590	3.384	27.934	BGH-5466	-2.315	48.345
BGH-4287A	0.096	0.146	BGH-5639	2.752	27.302	BGH-5639	-1.938	48.722
BGH-5598A	0.093	0.143	BGH-5648	2.733	27.283	BGH-5624A	-1.889	48.771
AO		0.050	OA		24.55	OA		50.66
AS		0.195	SA		30.818	SA		46.222

Accessions	**Linolenic**		Accessions	**Oleic/Linoleic**		Accessions	**PUFAs**	
	g	μ + g		g	μ + g		g	μ + g
BGH-5455A	-0.180	0.000	BGH-5456A	0.959	1.008	BGH-5456A	-9.555	41.125
BGH-4598A	-0.175	0.005	BGH-5361A	0.789	0.838	BGH-3333A	-7.348	43.332
BGH-1749	-0.175	0.005	BGH-3333A	0.768	0.817	BGH-5361A	-6.131	44.549
BGH-5301	-0.174	0.006	BGH-1749	0.707	0.756	BGH-1749	-5.433	45.247
BGH-1927	-0.173	0.007	BGH-5472A	0.684	0.733	BGH-5544A	-4.821	45.859
BGH-5240	-0.173	0.007	BGH-5544A	0.676	0.725	BGH-5556A	-4.661	46.019
BGH-4287A	-0.172	0.008	BGH-5453A	0.636	0.685	BGH-5453A	-4.261	46.419
BGH-5630A	-0.171	0.009	BGH-5653	0.612	0.661	BGH-5472A	-3.989	46.691
BGH-5051	-0.171	0.009	BGH-6155	0.597	0.646	BGH-6155	-3.974	46.706
BGH-5248	-0.170	0.010	BGH-5556A	0.595	0.644	BGH-5653	-3.407	47.273
BGH-5591A	-0.170	0.010	BGH-5466	0.584	0.633	BGH-5648	-2.368	48.312
BGH-4454A	-0.169	0.011	BGH-6590	0.527	0.576	BGH-5466	-2.311	48.369
BGH-5224	-0.169	0.011	BGH-5648	0.527	0.576	BGH-5639	-1.940	48.740
BGH-5485A	-0.169	0.011	BGH-5639	0.521	0.570	BGH-5624A	-1.895	48.785
AO		0.180	OA		0.049	OA		50.68
AS		0.008	SA		0.705	SA		46.245

Seed oil productivity (SOP), polyunsaturated fatty acids (PUFAs), original average (OA), and average of selected genotypes (AS).

4 DISCUSSION

4.1 Variance components and genetic-statistical parameters of characteristics related to SOP and fatty acid profiles

The greatest genotypic variances for MSF and PS, which were associated with their high heritability (0.89% and 0.70%, respectively), confirmed the marked genetic variability in these two characteristics, particularly for MSF, in the germplasm (Table 1). From these results, the predicted selection gains were significant and up to 29.39 g for MSF and 0.26 t ha^{-1} for PS (Table 1). Consistent with our results, remarkable variability in characteristics related to seed production, such as MSF and PS, in the *C. moschata* germplasm has been reported previously (Lima Neto, 2013; Darrudi *et al.*, 2018; Oliveira *et al.*, 2020).

Regarding the fatty acid profile of seed oil, the highest genotypic variance and very high heritability for the content of oleic acid, demonstrate the high genetic variability and feasibility of identifying *C. moschata* accessions that can produce oil with higher oleic acid content. As a result, the predicted selection gain for oleic acid was 6.99%, which corresponded to the greatest selection gain among the components of seed oil (Table 1). Furthermore, the linoleic acid and PUFA content exhibited high genetic variability, also demonstrating the feasibility of identifying *C. moschata* accessions that can produce oil with lower linoleic acid content, providing predicted selection gains of -5.12% for this first fatty acid (Table 1).

The results for the fatty acid profiles obtained in the present study are consistent with previously reported profiles (Sobreira, 2013; Jarret *et al.*, 2013). For instance, Jarret *et al.* (2013) evaluated the fatty acid profile of 38 seed samples corresponding to *C. moschata* accessions in the Plant Germplasm Collection of the United States Department of Agriculture, Griffin, and reported that oleic acid presented the greatest amplitude (10 to 53.80%), followed by linoleic acid (24.70 to 61.70%). Previous studies also corroborate high linoleic and oleic acid content, low palmitic acid content, and trace stearic and linolenic acid content in *C.*

moschata seed oil (Applequist *et al*., 2006; Kim *et al*., 2012; Veronezi and Jorge, 2015). To the best of our knowledge, however, no previous study on *C. moschata* has analyzed the genetic parameters associated with the components of the fatty acid profile of seed oil, such as genetic variance, heritability, and selection gain.

Studies on other oilseeds, such as soybean, rapeseed, and sunflower, have often reported variations in the fatty acid profiles among germplasm samples and attributed such variations to genetic factors (Hemingway, *et al*., 2015; Yol *et al*., 2017). Given the quantitative nature of the fatty acid profile, it is reasonable to assume that this characteristic is also influenced by environmental factors, and some studies have shown the strong influence of environmental factors, such as temperature, on the fatty acid profiles of various oilseeds (Werteker *et al*., 2010). Corroborating the findings of Gomes *et al*. (2020), most of the accessions evaluated in the present study were acquired from family farmers, who do not typically select accessions for the characteristics of seed or seed oil, which possibly contributed to the maintenance of high variability in these traits.

4.2 Correlations of characteristics related to SOP and fatty acid profiles

The phenotypic correlations of NFP, MSF, and PS with SOP ranged from moderate to very strong ($p < 0.001$), corroborating the notion that these characteristics are directly related to SOP (Figure 2). The genotypic correlations of these first three variables with SOP were very similar to the phenotypic correlations, indicating that the relationship of the variables NFP, MSF, and PS with SOP may be attributed to genetic factors (Figure 2). These trends were repeated for correlations of other variables, which tended to express genotypic correlations similar to the phenotypic.

In the present study, we expected the variables NFP, MSF, and PS to be strongly correlated with SOP, consistent with previous reports of strong correlation between seed yield and oil productivity (Assefa *et al*., 2018). Thus, in addition to PS and MSF, NFP appears to be a determinant of greater SOP in *C. moschata*. These correlations of SOP with its directly associated characteristics can be strategic to guide indirect selection aimed at increasing *C. moschata* oil productivity. Thus, the selection of genotypes with higher PS and MSF may be a promising alternative to obtain greater SOP in *C. moschata*.

The positive correlations of SOP with oleic acid content and oleic acid:linoleic acid ratio observed in the present study suggest the feasibility of obtaining greater oil productivity while maintaining a desirable oil profile, with high oleic acid content and oleic acid:linoleic acid ratio (Figure 2). Additionally, the negative correlations of SOP with linoleic acid and PUFA content indicate the feasibility of obtaining greater SOP while maintaining lower PUFA such as the linoleic acid. To the best of our knowledge, no study on *C. moschata* has addressed the relationship between SOP and components of seed oil fatty acid profile. Meanwhile, studies on other oilseeds, such as peanut, have reported weak phenotypic correlations between SOP and oleic acid content (-0.034) (Yol *et al.*, 2017), suggesting a small change in the acid content with increase in oil productivity in this crop. In contrast to the present study, Yol *et al.* (2017) reported a weak but positive correlation between SOP and oleic acid content.

Regarding the correlations among the components of oil fatty acid profile, the strong, negative but significant correlations of oleic acid content with linoleic acid and PUFA content indicate an antagonistic relationship between these components of *C. moschata* seed oil (Figure 2). The phenotypic correlations among these characteristics may be attributed to environmental and genetic factors. For instance, environmental factors, such as temperature, may strongly influence the correlations among fatty acids such as oleic and linoleic acid. In line with this, studies on other oilseeds, such as soybean, have shown that the oleic acid content of oil tends to decrease with increasing temperature, contrary to that observed for linoleic acid content (Bachlava *et al.*, 2008), corroborating the negative correlations among these two fatty acids observed in the present study.

The metabolic pathways of fatty acids may be the key genetic factor responsible for the correlations among the components of oil fatty acid profile. In this context, desaturases in the plastids and endoplasmic reticulum play a central role in fatty acid synthesis and catalyze their conversion to MUFAs or PUFAs (Long *et al.*, 2018). Among these enzymes, delta-12 fatty acid desaturase 2 (Δ12-FAD2) converts oleic acid precursors into linoleic acid precursors (Ohlrogge and Browse, 1995), and according to Dehghan and Yarizade (2014), the *FAD2* gene family is rather ubiquitous and diverse in plants. Thus, the strong negative correlation between oleic and linoleic acid content observed in the present study may also be related to the action of FAD2 during the biosynthesis of these two fatty acids.

The analysis of correlations among characteristics is an important subsidy for plant breeding, which must contemplate several variables simultaneously (Dias *et al.*, 2017), proving very useful when determining selection strategies. As shown in the present study, the plant germplasms maintained in banks commonly constitute a representative sample of the species gene pool, thus providing comprehensive information on the relationships among germplasm characteristics, which makes the correlation analysis very useful during the initial evaluation of plant germplasms conserved in banks. The information on correlations among the components of seed oil observed in the present study may be particularly meaningful in breeding programs aimed at improving the fatty acid profile of *C. moschata* seed oil, considering the feasibility of increasing oleic acid content while decreasing PUFA content, given the strong negative correlation between these two components.

4.3 Clustering and variability of accessions based on characteristics related to SOP and fatty acid profiles

The clustering of accessions confirmed the variability in characteristics related to SOP and fatty acid profiles among the studied accessions (Table 2). This clustering is consistent with the high estimates of genotypic variance and heritability for the evaluated characteristics related to SOP, including MSF and PS, as well as those related to the fatty acid profile of seed oil, including oleic and linoleic acid content (Table 1). The variability observed among the accessions tested in the present study also corroborates previous reports emphasizing remarkable variability in both agromorphological and molecular characteristics in the *C. moschata* germplasm (Ferriol *et al.*, 2004; Barboza *et al.*, 2012; Ferreira *et al.*, 2016).

The clustering of accessions in the present study did not reflect a greater similarity between accessions from the same state or geographic region, consistent with the reports of agromorphological characteristics of the *C. moschata* germplasm from different geographic regions (Moura, 2003; Gomes *et al.*, 2020). Unlike the present study, previous studies involving the analysis of fatty acid profile of the seeds of other oilseed crops, such as soybean, have reported that germplasms from regions at higher latitudes tended to express higher palmitic, stearic, and oleic acid content than germplasms from lower latitudes (Wu *et al.*, 2017; Abdelghany *et al.*, 2020), suggesting a greater similarity between the germplasms from the same region. In this line, Bachlava *et al.* (2008) observed that the oleic acid content of soybean tended to increase from lower to higher latitudes, indicating increase in the content of this fatty acid with decrease in temperature, contrary to the trends for linoleic and linolenic acid. Similarly, Song *et al.* (2016) observed that the oleic acid content of soybean seeds was negatively correlated with the duration of sunlight incidence. Thus, the greater similarity between the soybean germplasms from the same region may reflect regional ecogeographic characteristics, different from the results of the present study.

Already present in the diet of native peoples (Piperno *et al.*, 2003; Dillehay *et al.*, 2007), *C. moschata* is widely cultivated in Latin America, which is an important center of diversity for this vegetable. In addition, previous studies have highlighted the variability in the Brazilian germplasm of *C. moschata* (De Lima *et al.*, 2016; Ferreira *et al.*, 2016; Gomes *et al.*, 2020), possibly as a result of the adaptation of this germplasm to a wide ecological range, constituted by diverse edaphoclimatic conditions. Additionally, the intrinsic characteristics of *C. moschata*, such as the occurrence of natural hybridization across populations, associated with the processes of selection and seed exchange practiced by populations involved in its cultivation, also contribute to the variability in the Brazilian germplasm of this vegetable (Gomes *et al.*, 2020).

The similarity between accessions of different geographic regions observed in the present study suggests the adaptability of the germplasm, indicating the feasibility of its cultivation under edaphoclimatic conditions different from those in its regions of origin. Compared with other crops, such as soybean (Abdelghany *et al.*, 2020), the possible adaptability of *C. moschata* germplasm observed in the present study represents an opportunity to meet the diverse demands of various cultivation regions and systems of this vegetable, specifically in Brazil, which is characterized by its continental dimension.

Additionally, crossbreeding between divergent genotypes has been conveniently explored in *C. moschata*, aiming at the exploitation of hybrid vigor for aspects related to growth habit, fruit and seed production, as well as for fruit chemical–nutritional aspects (El-Tahawey *et al.*, 2015; Kumar *et al.*, 2018). Based on genotypic data and environmental interaction-free genotypic values, our results of variability analysis will be particularly useful to assist the crossing of promising genotypes aimed at the exploration of hybrid vigor for characteristics related to SOP and fatty acid profiles.

4.4 Principal components analysis

In consonance with their respective variance components and genetic parameters (Table 1), the oleic acid, linoleic acid, and PUFA content made the greatest contributions to the discrimination of the genotypes in PCA, confirming the high variability of these characteristics (Figure 3). The results of PCA regarding the variability of accessions were consistent with the results of clustering analysis (Table 2), also corroborating the strong positive correlations of PS with SOP, as well as the positive correlation of SOP with the oleic acid content (Figure 2).

4.5 Identification of promising accessions in terms of SOP and fatty acid profiles

Based on their highest genotypic averages for SOP, the groups 5 and 2 were identified as the most promising for this characteristic (Table 3). Consistent with this result, *per se* analysis identified accession BGH-4610 from group 5 as the most promising in terms of SOP, with the (μ+g) estimate of 0.27 t ha^{-1} (Table 4). In addition, accessions BGH-5485A, BGH-6590, BGH-5556A, BGH-5472A, and BGH-5544A from group 2 were also identified as promising in terms of SOP in *per se* analysis, with the (μ+g) estimate of ~0.20 t ha^{-1} (Table 4).

The groups 3 and 2 were identified as the most promising in terms of SOP with higher oleic acid content (Table 4). Associated with this, the accession BGH-5456A from group 3 was identified as most promising in terms of high oleic acid content in *per se* analysis, with the (μ+g) estimate of 37.20% (Table 4). The accessions BGH-3333A, BGH-5361A, BGH-5472A, BGH-5544A, and BGH-5453A from group 2 and the accession BGH-1749 from group 3 were also identified as promising in terms of high oleic acid content, with the (μ+g) estimate of ~30.00% (Table 4). The identification of promising groups and *per se* identification of accessions with high oleic acid content in the seed oil are associated with the high amplitude of (μ+g) estimates for this characteristic (Figure 4).

The groups 3 and 2 also presented the lowest averages for linoleic acid and PUFA content, confirming them as the most promising in terms SOP with lower PUFA content (Table 4). Consistent with this result, the accession BGH-5456A, BGH-3333A, BGH-5361A, BGH-5544A, BGH-5556A, and BGH-1749 were identified as promising in terms of low linoleic acid content, with the (μ+g) estimate of ~45% (Table 4). In *per se* analysis, the accessions identified as the most promising in terms of low linoleic acid also were the most promising in terms of low PUFA content, demonstrating the predominance of linoleic acid among PUFAs.

Recently, the development of cultivars with an oil profile that is better suited to human nutrition and health has been emphasized in the genetic breeding of oilseed crops. This is in agreement with a series of studies demonstrating the association between the consumption of lipid sources predominantly comprising saturated fatty acids and the high risk of cardiometabolic pathologies, particularly cardiovascular diseases and type II diabetes mellitus (Harris *et al.*, 2009; Keys *et al.*, 2017; Wu *et al.*, 2019). This has encouraged the replacement of saturated lipids in human food by unsaturated fatty acids, with a particular focus on vegetable oils—the main source of unsaturated fatty acids in the human diet.

Associated with high levels of unsaturated fatty acids, vegetable oils should ideally have high stability against environmental stressors, such as humidity, light, heat, and oxygen. These oils must also be resistant to oxidative actions, which are related to the production of secondary components responsible for triggering allergic responses and cardiovascular diseases, such as atherosclerosis (Yanishlieva *et al.*, 2001; Garbin *et al.*, 2013), also responsible deteriorating the sensory quality (Choe and Min, 2006) and reduce the shelf life of oils (Xie *et al.*, 2019). Given these demands, breeding programs for oilseed crops such as soybean, rapeseed, and corn have emphasized the development of cultivars that produce more stable oils, prioritizing the increase in the content of oleic acid, a MUFA, which has greater oxidative stability (Burton *et al.*, 2006; Bachlava *et al.*, 2008; Wang *et al.*, 2009; Long *et al.*, 2018). In this line, some studies have confirmed the higher oxidative stability of oleic acid [C18: 1 (Δ^{9})] compared to PUFAs such as linoleic acid [C18: 2 ($\Delta^{9, 12}$] and linolenic acid [C18: 3 ($\Delta^{9, 12, 15}$], indicating that oleic acid is 10 times more stable than linoleic acid and 20 times more stable than linolenic acid (Liu *et al.*, 1992). Thus, genetic breeding for improving the fatty acid profile of *C. moschata* seed oil should simultaneously target an increase in oleic acid content and decrease in polyunsaturated fatty acids, particularly linoleic acid content, aiming at greater nutritional and physicochemical quality of oil.

The identification of *C. moschata* accessions with high oleic acid content in the seed oil indicates the feasibility of identifying accessions that produce seed oil with higher nutritional quality and stability. Of note, *C. moschata* seed oil contains high levels of bioactive components, such as vitamin E and carotenoids (Veronezi and Jorge, 2012), which are important antioxidants in the human diet and protect the oil against oxidative processes. Despite of their high nutritional value, a large portion of seeds produced during *C. moschata* cultivation is still discarded (Li *et al*., 2019), particularly in Brazil. As highlighted by Gomes *et al*. (2020), the use of *C. moschata* seeds for oil production represents an alternative strategy to complement the diet, in addition to increasing the income of farmers involved in the production of this vegetable.

To the best of our knowledge, the present study is the first to analyze the SOP and fatty acid profile of *C. moschata* based on a relatively large number of accessions representative of different geographic regions of Brazil. Our results demonstrate the marked potential of some *C. moschata* accessions for oil production, as confirmed by the *per se* identification of accessions with high SOP and high oleic acid content in the oil. Our results corroborate the productive potential as well as the nutritional and physicochemical quality of seed oil from the Brazilian germplasm of *C. moschata*.

5 CONCLUSIOSNS

The tested accessions expressed high genetic variability in terms of MSF and PS, providing the predicted selection gains of 29.39 g and 0.26 ha^{-1}, respectively.

Phenotypic and genotypic correlations indicated that a greater *C. moschata* SOP can be achieved by selecting for higher PS and MSF. Correlations also indicated that a greater SOP can be obtained while maintaining high oleic fatty acid content and low linoleic acid content, providing oil with better nutritional and chemical quality.

In the analysis of variability, the 91 accessions tested in this study were clustered into five groups, allowing the identification of the most promising groups in terms of greater SOP and higher oleic acid content in the oil, an approach that will guide the use of this germplasm in breeding programs aimed at improving the SOP and fatty acid profile.

Per se analysis identified the accessions BGH-4610, BGH-5485A, BGH-6590, BGH-5556A, BGH-5472A, and BGH-5544A as the most promising in terms of SOP, with the (μ+g) estimate of ~0.20 t ha^{-1}. Accessions BGH-5456A, BGH-3333A, BGH-5361A, BGH-5472A, BGH-5544A, BGH-5453A, and BGH-1749 were identified as the most promising in terms of higher oleic content in oil were, with the (μ+g) estimate of ~30%, and most of these accessions were also the most promising in terms of lower linoleic acid content in oil, with the (μ+g) estimate ~40%. Therefore, part of the *C. moschata* germplasm evaluated in the present study is a promising source for the genetic improvement of SOP and fatty acid profile, aiming at the production of oil with better nutritional and physicochemical quality.

AUTHOR CONTRIBUTIONS

Conceptualization: Derly José Henriques da Silva.

Data curation: Ronaldo Silva Gomes, Ronaldo Machado Júnior, Cleverson Freitas de Almeida, Rebeca Lourenço de Oliveira.

Investigation: Ronaldo Silva Gomes, Ronaldo Machado Júnior, Cleverson Freitas de Almeida, Rebeca Lourenço de Oliveira, Ednângelo Duarte Pereira, Fabio Teixeira Delazari.

Software: Rafael Ravaneli Chagas and Ronaldo Silva Gomes.

Supervision: Derly José Henriques da Silva.

Writing – original draft: Ronaldo Silva Gomes.

Writing – review & editing: Ronaldo Silva Gomes, Derly José Henriques da Silva, Rebeca Lourenço de Oliveira, Cleverson Freitas de Almeida.

FUNDING

This work had support from the Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG). This work was supported by the Coordenação de Aperfeiçoamento de Pessoal e Nível Superior (CAPES) [grant number 001]. This work had the support from the National Council of Technological and Scientific Development (CNPq), which corresponded to a doctorate scholarship (doctorate-GD grant), granted to the first author, Ronaldo Silva Gomes. There was no additional external funding received for this study.

6 REFERENCES

Abdelghanya, A. M., Zhanga, S., Azama, M., Shaibua, A. S., Fenga, Y., Lia, Y., Tiana, Y., Honga, H., Lia, B., Suna, J., 2020. Profiling of seed fatty acid composition in 1025 Chinese soybean accessions from diverse ecoregions. The Crop J. 8, 635–644. https://doi.org/10.1016/j.cj.2019.11.002.

Applequist, W. L., Avula, B., Schaneberg, B.T., Wang, Y. H., Khan, I.A., 2006. Comparative fatty acid content of four *Cucurbita* species grown in a common (shared) garden. J. Food Compos. Anal. 19, 606–611. https://doi.org/10.1016/j.jfca.2006.01.001.

Assefa, Y., Bajjalieh, N., Archontoulis, S., Casteel, S., Davidson, D., Kovács, P., Naeve, P., Ciampitti, I. A., 2018. Spatial characterization of soybean yield and quality (amino acids, oil, and protein) for United States. Sci Rep-UK. 8, 1-11. https://doi:10.1038/s41598-018-32895-0.

Barboza, N., Albertazzi, F. J., Sibaja-Cordero, J. A., Mora-Umana, F., Astorga, C., Ramírez, P., 2012. Analysis of genetic diversity of *Cucurbita moschata* (D.) germplasm accessions from Mesoamerica revealed by PCR SSCP and chloroplast sequence data. Sci. Hortic-Amsterdam. 134, 60-71. https://doi.org/10.1016/j.scienta.2011.10.028.

Bates, D., Maechler, M., Bolker, B., Walker, S., 2015. Fitting linear mixed-effects models using lme4. J. Stat. Softw. 67, 1-48. https://doi.org/10.18637/jss.v067.i01.

Bachlava, E., Burton, J. W., Brownie, C., Wang, S., Auclair, J., Cardinal, A. J., 2008. Heritability of oleic acid content in soybean seed oil and its genetic correlation with fatty acid and agronomic traits. Crop Sci. 48, 1764–1772. https:// doi:10.2135/cropsci2004-0769.

Burton, J.W., Wilson, R.F., Rebetzke, G. J., Pantalone, V. R., 2006. Registration of N98-4445A mid-oleic soybean germplasm line. Crop Sci. 46, 1010–1012. https://doi:10.2135/cropsci2004-0769.

Carvalho, L. M. J. De., Gomes, P. B., Godoy, R. L. De O., Pacheco, S., Do monte, P. H. F., De Carvalho, J. L. V., Nutti, M. R., Neves, A. C. L., Vieira, A. C. R. A., Ramos, S. R. R., 2012. Total carotenoid content, *α*-carotene and *β*-carotene, of landrace pumpkins (*Cucurbita moschata* Duch): A preliminary study. Food Res. Int. 47, 237- 340. https://doi.org/10.1016/j.foodres.2011.07.040.

Choe, E., Min, D. B., 2006. Mechanisms and factors for edible oil oxidation. Compr. Rev. Food Sci. F. 5, 169–186. https://doi.org/10.1111/j.1541-4337.2006.00009.x.

Cruz, C. D., 2013. GENES—a software package for analysis in experimental statistics and quantitative genetics. Acta Sci-Agron. 35, 271-276. https://doi: 10.4025/actasciagron.v35i3.21251.

Cullis, B. R., Smith, A. B., Coombes, N. E., 2006. On the design of early generation variety trials with correlated data. J Agric Biol. Envir. S. 11, 381-393. https://doi: 10.1198/108571106X154443.

Dash, P., Ghosh, G., 2017. Proteolytic and antioxidant activity of protein fractions of seeds of *Cucurbita moschata.* Food Biosc. 18, 1-8.http://dx.doi.org/10.1016/j.fbio.2016.12.004.

Darrudi, V. N., Soltania, F., Shokrpoura, M., Ercolanob, M. R., 2018. Genetic diversity of *Cucurbita pepo* L. and *Cucurbita moschata* Duchesne accessions using fruit and seed quantitative traits. J. Appl. Res. Med. Arom. Plant. 8, 60-66. https://doi.org/10.1016/j.jarmap.2017.11.003.

De Lima, G. K. L., De Queiroz, M. A., Silveira, L. M. da. 2016. Rescue of *Cucurbita* spp. germplasm in Rio Grande do Norte. Ver. Caatinga. 29, 257-262. https://doi.org/10.1590/1983-21252016v29n130rc.

Dehghan, N. F., Yarizade, K., 2014. Bioinformatics study of delta-12 fatty acid desaturase 2 (FAD2) gene in oilseeds. Mol. Biol. Rep. 41, 5077–5087. https://doi: 10.1007/s11033-014-3373-5.

Dias, F. dos S., Dos Santos, F. C., Gomes, R. S., Da Silva, D. C., Soares, D. de A., Rosa, D. P., Da Silva, A. F., 2017. Path analysis in fruits of the buriti (*Mauritia flexuosa* L. f) in populations from the Brazilian Middle North Region. Tree Genet Genomes. 13, 1-6. https://doi.org/10.1007/s11295-017-1205-1.

Dillehay, T., Rossen, J., Andres, T. C., Williams, D. E., 2007. Preceramic Adoption of peanut, squash, and cotton in Northern Peru. Science. 316, 1890-1893. https://doi: 10.1126/science.1141395.

El-Tahawey, M.A.F.A., Kandeel, A. M., Youssef, S. M. S., Abd El-Salam, M. M. M., 2015. Heterosis, potence ratio, combining ability and correlation of some economic traits in diallel crosses of pumpkins. Egypt J. Plant Breed. 19, 419 – 439.

Federer, W. T., 1956. Augmented or (Hoonuiaku) designs. Hawaiian Planter's Record. 55, 191–208.

Ferreira, M. G., Salvador, F. V., Lima. M. N. R., 2016. Parâmetros genéticos, dissimilaridade e desempenho per se em acessos de abóbora. Hortic. Bras. 34, 537-546. https://doi.org/10.1590/s0102-053620160413.

Ferriol, M., Picó, B., De Córdova, P. F., Nuez, F., 2004. Molecular diversity of a germplasm collection of squash (*Cucurbita moschata*) determined by SRAP and AFLP markers. Crop Sci. 44, 653–664.

Food and Agriculture Organization of the United Nations- FAO STATE, 2020. http://www.fao.org.faostat/em/#home/ (accessed 15 February 2021).

Fonseca, M. A., Queiroz, M. A. de., Lopes, J. F., Assis, J. G. de A., Romão, R. L., Ramos, S. R.R., Do Carmo, C. A. S., Moura, M. C. L., Dias, R. C. S., Silveira, L. M. Da., Barbieri, R. L., Da Silva, D. J. H., Noronha, S., 2015. Geographical distribution and conservation of *Cucurbita* in Brazil. Magistra. 27, 432–442.

Garbin, U., Baggio, E., Stranieri, C., Pasini, A., Manfro, S., Mozzini, C., Vallerio, P., Lipari, G., Merigo, F., Guidi, G., Cominacini, L., Pasini, A.F., 2013. Expansion of necrotic core and shedding of Mertk receptor in human carotid plaques: a role for oxidized polyunsaturated fatty acids? Cardiovasc. Res. 97, 125–33. https://doi: 10.1093/cvr/cvs301.

Gomes, R. S., Machado Júnior, R., Cleverson, F. De A., Chagas, R. R., Oliveira, R. L. De., Delazari, F. T., Da Silva, D. J. H. 2020. Brazilian germplasm of winter squash (*Cucurbita moschata* D.) displays vast genetic variability, allowing identification of promising genotypes for agro-morphological traits. Plos One. 15, 1-16. https://doi.org/10.1371/journal.pone.0230546.

Harris, W. S., Mozaffarian, D., Rimm, E., Kris-Etherton, P., Rudel, L. L., Appel, L. J., Engler, M. M., Engler, M. B., Sacks, F., 2009. Omega-6 fatty acids and risk for cardiovascular disease. A science advisory from the American Heart Association Nutrition Subcommittee of The Council on Nutrition, Physical Activity, and Metabolism; Council on Cardiovascular Nursing; and Council on Epidemiology and Prevention. Circulation. 119, 1-8. https://doi: 10.1161/circulationaha.108.191627.

Hemingway, J., Eskandari, M., Rajcan, I., 2015. Genetic and environmental effects on fatty acid composition in soybeans with potential use in the automotive industry. Crop Sci. 55, 658-668. https://doi: 10.2135/cropsci2014.06.0425.

Jarret, R. L., Levy, I. J., Potter, T. L., Cermak, S. C., Merrick, L. C., 2013. Seed oil content and fatty acid composition in a genebank collection of *Cucurbita moschata* Duchesne and *C. argyrosperma* C. Huber. Plant Genet. Resour-C. 11, 149-157. http://doi:10.1017/S1479262112000512.

Kumar, V., Mishra, D. P., Yadav, G. C., Yadav, S., 2018. Exploitation of heterobeltiosis and economic heterosis for horticultural yield, and its attributes and biochemical traits in pumpkin (*Cucurbita moschata* Duch. ex. Poir) under salt affected soil. Curr. Sci. India. 115, 1550–156. https://doi: 10.18520/cs/v115/i8/1550-1556.

Keys, A., Menotti, A., Karvonen, M. J., Aravanis, C., Blackburn, H., Buzina, R., Djordjevic, B. S., Dontas, A. S., Fidaza, F., Key, M. H., Kromhout, D., Nedeljkovic, S., Punsar, S., Seccareccia, F., Toshima, H., 2017. The diet and 15-year death rate in the seven countries study. Am. J. Epidemiol.185, 1130-1142. https://doi: 10.1093/aje/kwx101.

Kim, M. Y., Kim, E. J., Kim, Y-N., Choi, C., Lee, B-H., 2012. Comparison of the chemical compositions and nutritive values of various pumpkin (*Cucurbitaceae*) species and parts. Nutr. Res. Pract. 6, 21–27. http://dx.doi.org/10.4162/nrp.2012.6.1.21.

Li, H., 2019. Evaluation of bioactivity of butternut squash (*Cucurbita moschata* D.) seeds and skin. Food Sci. Nutr. 8, 3252–3261. https://doi: 10.1002/fsn3.1602.

Lima Neto, I. S., 2013. Pré- melhoramento de abobra (*Cucurbita moschata*) visando a biofortificacao em carotenoides. D. Sc. Thesis, Universidade Federal de Viçosa. Available from: https://www.locus.ufv.br/handle/123456789/1367.

Long, W., Hu, M., Gao, J., Chen, S., Zhang, J., Cheng, L., Pu, H., 2018. Identification and functional analysis of two new mutant BnFAD2 alleles that confer elevated oleic acid content in rapeseed. Front. Genet. 9, 1-12. https://doi: 10.3389/fgene.2018.00399.

Liu, H.-R., White, P. J., 1992. Oxidative stability of soybean oils with altered fatty acid. J. Am. Oil Chem. Soc. 69, 528–532.

Moura, M. C. C. L., 2003. Identificação de fontes de resistência ao *potyvirus ZYMV* e diversidade genética e ecogeográfica em acessos de abóbora. D. Sc. Thesis. Universidade Federal de Viçosa. Available from: https://www.locus.ufv.br/handle/123456789/10265.

Nagar, A., Sureja, A. K., Kar, A., Bhardwaj, R., Krishnan, S. G., Munshi, A. Das. 2018. Profiling of mineral nutrients and variability study in pumpkin (*Cucurbita moschata*) genotypes. Agr. Res. 7, 225-231. https://doi.org/10.1007/s40003-018-0329-3.

Oliveira, R. L. O., Gonçalves, L. S. A., Rodrigues, R., Baba, V. Y., Sudré, C. P., Santos, M. H., Aranha, F. M. 2016. Genetic divergence among pumpkin landraces. Semin-Cienc Agrar. 37, 1-10. https:// doi10.5433/1679-0359.2016v37n2p547.

Oliveira, R. L., Gomes, R. S., Almeida, C. F., Machado Júnior, R., Rocha, J. R. A. S. C., Silva, D. J. H., Carneiro, P. C. S. 2020. Multitrait selection of pumpkin genotypes aimed at reducing the growth habit and improving seed production. Crop Sci. 61, 1-10. https:// doi 10.1002/csc2.20386.

Ohlrogge, J., Browse, J., 1995. Lipid biosynthesis. Plant Cell. 7, 957–970. https://doi.org/10.2307/3870050.

Patel, S., 2013. Pumpkin (*Cucurbita* sp.) seeds as nutraceutic: A review on status quo and scopes. Mediterr. J. Nutr. Metab. 6, 183-189. https://doi.org/10.1007/s12349-013-0131-5.

Piperno, D. R., Stothert, K. E., 2003. Phytolith evidence for early Holocene *Cucurbita* domestication in Southwest Ecuador. Science. 299, 1054-1057. https://doi: 10.1126/science.1080365.

Priori, D., Valduga, E., Villela, J. C. B., Mistura, C. C., Vizzotto, M., Valgas, R., Barbieri, R. L., 2047. Characterization of bioactive compounds, antioxidant activity and minerals in landraces of pumpkin (*Cucurbita moschata*) cultivated in Southern Brazil. Food Sci. Tech-Brazil. 37, 33-40. https://doi.org/10.1590/1678-457x.05016.

Resende, M. D.V., 1995. Delineamento de experimentos de seleção para maximização da acurácia seletiva e do progresso genético. Rev. Arvore. 19, 479–500.

Saltzman, A., Birol, E., Bouis, H. E., Boy, E., De Moura, F. F., Islam, Y., Pfeiffer, W. H., 2013. Biofortification: progress toward a more nourishing future. Glob. Food Secur. 2, 9-17. https://doi.org/10.1016/j.gfs.2012.12.003.

Shimakura, S.E., Ribeiro Júnior, P. J. 2012. Estatística descritiva: interpretação do coeficiente de correlação. Universidade Federal do Paraná, Curitiba.

Silva, D. J. H, Moura, M. C. C., Casali, V. W. D., 2001. Recursos genéticos do banco de germoplasma de hortaliças da UFV: histórico e expedições de coleta. Hortic. Bras. 19, 108-114. https://doi.org/10.1590/S0102-05362001000200002.

Sobreira, F. M., 2013. Divergência genética entre acessos de abóbora para estabelecimento de coleção nuclear e pré-melhoramento para óleo funcional. D. Sc. Thesis, Universidade Federal de Viçosa. Available from: https://www.locus.ufv.br/handle/123456789/1367.

Song, W., Yang, R., Wu, T., Wu, C., Sun, S., Zhang, S., Jiang, B., Tian, S., Liu, X., Han, T., 2016. Analyzing the effects of climate factors on soybean protein, oil contents, and composition by extensive and high-density sampling in China. J. Agric. Food Chem. 64, 4121–4130. https://doi: 10.1021/acs.jafc.6b00008.

Thiex, N. J., Anderson, S., Gildemeister, B., 2003. Crude fat, hexanes extraction, in feed, cereal grain, and forage (Randall/Soxtec/Submersion Method): Collaborative study. Journal of AOAC International, 86, 899-908.

Veronezi, C. M., Jorge, N., 2015. Chemical characterization of the lipid fractions of pumpkin seeds. Nutr. Food Sci. 45, 164-173. https://doi 10.1108/NFS-01-2014-0003.

Veronezi, C., Jorge, N., 2012. Bioactive compounds in lipid fractions of pumpkin (*Cucurbita* sp) seeds for use in food. J. Food Sci. 77, 53-57. http://doi: 10.1111/j.1750-3841.2012.02736.x.

Wang, H., Wu, B., Song, T., Chen, S., 2009. Effects of long-term selection for kernel oil concentration in KYHO, a high-oil maize population. Crop Sci. 49, 459–466. https://doi: 10.2135/cropsci2008.06.0336.

Werteker, M., Lorenz, A., Johannes, H., Berghofer, E., Findlay, C. S., 2010. Environmental and varietal influences on the fatty acid composition of rapeseed, soybeans and sunflowers. J. Agron. Crop Sci. 196, 20-27. https://doi:10.1111/j.1439-037X.2009.00393.x.

Wu, J. H. Y., Micha, R., Mozaffarian, D., 2019. Dietary fats and cardiometabolic disease: mechanisms and effects on risk factors and outcomes. 16, 1-21. https://doi.org/10.1038/s41569-019-0206-1.

Wu, T., Yang, X., Sun, S., Wang, C., Wang, Y., Jia, H., Man, W., Fu, L., Song, W., Wu, C., Yan, H., Jiang, B., Hou, W., Ren, G., Han, T., 2017. Temporal–spatial characterization of seed proteins and oil in widely grown soybean cultivars across a century of breeding in China. Crop. Sci. 57, 748–759. https://doi: 10.2135/cropsci2016.05.0336.

Xie, Y., Jiang, S., Li, M., Guo, Y., Cheng, Y., Qian, H., Yao, W., 2019. Evaluation on the formation of lipid free radicals in the oxidation process of peanut oil. LWT-Food Sci. Technol. 104, 24–29. https://doi: 10.1016/j.lwt.2019.01.016.

Yol, E., Ustun, R., Golukcu, M., Uzun, B., 2017. Oil content, oil yield and fatty acid profile of groundnut germplasm in Mediterranean climates. J. Am. Oil Chem. Soc.94, 787–804. https://doi 10.1007/s11746-017-2981-3.

Yanishlieva, N. V., Marinova, E. M., 2001. Stabilisation of edible oils with natural antioxidants. Eur. J. Lipid Sci. Tech. 103, 752–767. https://doi: 10.1002/(ISSN)1438-9312.

7 CONSIDERAÇÕES FINAIS

A avaliação agromorfológica realizada no presente trabalho identificou acessos com florescimento precoce, promissores para a produtividade de frutos e acessos com elevado teor total de carotenoides na polpa de frutos.

A avaliação dos aspectos relacionados às sementes e ao óleo identificou acessos com elevada produtividade de óleo de sementes e elevado teor de ácido oleico.

Foram obtidas as correlações entre as características agromorfológicas e entre os aspectos relacionados às sementes e ao óleo, assim como a estimativa da variabilidade genética dos acessos, informações que poderá auxiliar nas estratégias de melhoramento da *C. moschata*.

Parte dos acessos avaliada nesse trabalho é promissora para uso visando o melhoramento genético da *C. moschata* para aspectos agromorfológicos, produtividade do óleo de sementes o perfil do óleo.

8 APÊNDICE

Brazilian germplasm of winter squash (C*ucurbita moschata* D.) displays vast genetic variability, allowing identification of promising genotypes for agro-morphological traits

Ronaldo Silva Gomes[1*], Ronaldo Machado Júnior[1], Cleverson Freitas de Almeida[1], Rafael Ravaneli Chagas[1], Rebeca Lourenço de Oliveira[1], Fabio Teixeira Delazari[2], Derly José Henriques da Silva[1]

[1]Agronomy Department, Graduation Program on Genetics and Breeding, Federal University of Viçosa-UFV, Viçosa, MG, Brazil; [2]National Service of Rural Learning- SENAR, Campo Grande, MS, Brazil

***Corresponding author**: ronaldo.s.gomes@ufv.br

SUPPLEMENTARY MATERIAL

Supplementary table 1. Multi-categorical *descriptors used in the assessment of the C. moschata germplasm maintained by BGH-UFV*

Phase/organ	Descriptors
Vegetative	Growth habit, stem colour (SC), intensity of leaf green (ILG), leaf silvering (LS), intensity of leaf silvering (ILS), leaf serration (LS), presence of trichomes in the leaves (PAL), amount of trichomes in the adaxial surface of leaves (ATT), amount of trichomes in the abaxial surface of leaves (ATL), leaf recess (LR), presence of trichomes in the petiole (PTP), amount of trichomes in the petiole (ATP), and green intensity of male pedicel (GIMP).
Fruit	Format of fruits (FF), format of peduncle (FP), number of colours of fruit peel (NPC), topography of fruit surface (FT), format of floral scar (FFS), peel texture (PT), predominant colour of fruit peel (PC), and depth of fruits slices (DFS).
Seed	Seed format (SF), aspect of seed tegument (AST), seed tegument texture (STT), colour of seed tegument (CST), and colour of seed border (CSB).

Supplementary table 2. Result of principal component analysis showing the fifteen principal components and the relative contribution of traits in each component

Traits	PC 1	PC 2	PC 3	PC 4	PC 5	PC 6	PC 7	PC 8	PC 9	PC 10	PC 11	PC 12	PC 13	PC 14	PC 15
S21	1.10	0.10	0.03	0.08	0.17	0.03	1.52	0.25	0.02	7.48	1.03	2.04	2.19	0.01	0.59
S28	0.07	0.24	0.01	1.30	0.12	0.01	3.22	1.36	0.19	2.83	0.20	0.88	1.69	3.02	6.53
LMV14	0.00	0.01	0.35	2.26	0.21	3.99	4.82	0.00	0.06	2.12	4.21	0.00	1.84	0.04	0.56
LMV21	0.66	0.02	0.09	2.98	0.15	3.21	2.11	0.82	0.18	1.22	4.83	0.28	1.89	1.10	0.48
DDF	0.21	1.30	0.94	0.30	0.13	0.47	4.29	0.01	0.00	0.06	2.40	0.01	1.15	0.14	5.39
NFP	1.66	2.88	2.44	0.90	3.02	0.43	1.33	3.77	0.18	0.08	0.85	0.00	3.20	0.05	0.41
MF	11.76	0.32	0.17	0.39	0.03	0.01	0.20	0.06	0.04	0.91	0.00	0.09	0.03	0.28	0.54
PF	4.38	0.25	2.70	0.07	0.85	0.81	1.42	5.81	0.00	0.90	0.82	0.03	1.97	0.23	0.25
HF	6.10	1.46	1.64	1.08	2.96	0.11	0.18	1.26	1.32	1.32	1.20	0.13	0.00	0.07	0.03
DF	8.84	0.14	0.48	0.09	0.45	0.00	0.20	0.05	0.32	0.82	1.73	0.40	0.35	0.21	0.50
TFP	3.27	0.04	0.25	0.21	0.00	0.82	0.05	0.74	0.25	1.07	0.00	7.42	1.00	2.11	0.01
RFP	0.26	0.07	3.39	3.21	0.78	0.08	0.00	0.13	0.83	0.50	1.19	0.22	0.05	0.04	0.07
RP	0.18	1.00	8.51	1.07	1.49	0.01	0.13	0.56	1.76	0.05	0.26	0.20	0.00	0.03	0.70
PT	1.93	1.58	0.34	2.20	1.12	0.05	0.50	0.68	0.70	3.37	3.67	0.39	0.59	0.04	0.07
DIC	8.81	0.27	0.37	0.00	1.21	0.01	0.25	2.06	1.11	0.00	0.02	0.29	0.08	0.45	0.06
TC	0.86	0.93	0.11	0.19	0.17	0.02	0.13	4.45	0.00	0.60	3.50	0.33	0.51	0.40	0.38
L	0.36	0.26	0.53	0.01	0.62	0.00	0.13	0.91	1.15	0.32	0.55	0.03	0.79	4.83	7.31
NSF	2.41	4.48	3.53	0.47	0.00	0.09	0.01	0.17	0.04	0.33	0.00	0.11	3.55	3.61	0.00
MSF	4.99	4.40	1.49	0.05	0.33	0.63	0.62	1.19	0.47	0.60	0.13	0.05	2.60	0.15	0.18
MS/F	2.22	5.25	3.21	0.14	0.50	0.33	0.03	0.67	0.15	0.36	0.47	0.36	0.11	0.00	1.20
MOHS	4.65	0.01	0.25	0.00	1.13	0.30	2.79	0.20	3.50	0.87	1.31	0.49	0.16	4.01	0.28
PS	0.00	6.15	5.79	1.44	4.07	0.00	0.27	0.37	0.57	0.09	0.76	0.07	0.24	0.00	0.12
SL	2.08	0.69	3.80	0.55	0.88	0.01	0.03	3.24	3.82	2.00	0.72	0.91	0.34	0.38	0.01
SW	5.20	0.07	0.97	0.02	0.24	0.37	1.01	0.09	3.96	2.95	0.41	0.41	0.79	0.15	0.17
ST	1.98	1.00	0.01	0.07	1.39	0.06	0.84	1.48	0.49	0.01	0.71	0.47	0.02	7.03	0.26
SOC	0.10	0.13	5.93	0.26	1.38	0.29	1.71	0.39	0.07	1.65	0.01	1.70	0.00	0.01	1.14
SOP	0.00	6.13	5.57	1.30	3.90	0.06	0.20	0.25	0.52	0.18	1.31	0.08	0.29	0.00	0.28
GH	1.46	1.85	0.04	0.20	0.00	1.85	1.33	0.06	0.61	0.50	3.96	0.57	0.08	0.47	0.10
SC	0.01	0.52	2.23	0.55	3.98	4.94	2.20	1.40	0.02	3.27	5.17	3.01	16.29	1.03	0.22
ILG	0.36	1.19	0.37	4.40	1.03	6.84	1.24	2.38	5.28	2.47	4.82	1.76	1.11	2.68	2.83
LSQ1	0.10	0.07	2.03	0.07	0.07	1.81	0.06	2.11	0.97	0.00	4.52	0.00	0.53	0.00	4.59
ILS	0.50	0.89	1.00	0.04	7.02	1.20	0.13	0.17	0.82	0.32	3.65	0.54	0.43	0.06	0.10
LSQ2	1.21	0.91	0.60	0.14	1.12	5.57	1.68	0.25	2.49	0.61	0.31	0.40	1.80	11.88	0.60
PTL	0.06	0.32	0.77	0.21	0.07	1.30	3.01	0.99	0.02	5.98	2.43	0.01	0.74	0.55	3.52
ATT	0.38	10.91	2.47	17.20	5.62	0.81	0.93	0.66	0.18	0.40	0.17	0.52	0.16	1.62	0.06
ATL	0.02	7.27	2.78	13.03	0.85	1.99	0.12	0.62	0.87	0.21	1.34	1.06	0.38	1.04	2.86
LR	0.91	10.30	6.96	9.00	8.98	1.10	6.11	3.71	2.71	3.01	2.33	2.82	1.54	3.17	2.11
PTP	0.61	0.21	1.09	0.08	0.00	0.02	1.04	1.99	2.10	6.15	0.01	7.99	0.18	0.52	0.08
ATP	0.15	9.45	2.37	17.16	5.11	0.67	1.81	1.40	0.39	0.28	0.27	0.46	0.24	1.73	0.16
GIMP	0.63	2.31	0.27	0.61	0.22	0.94	0.29	4.61	7.92	1.02	4.27	6.92	0.17	0.91	1.42
FF	2.01	2.42	8.11	4.47	8.54	4.60	4.99	2.95	7.15	8.89	10.03	9.58	2.58	7.61	4.31
FP	0.28	0.68	0.93	1.13	0.53	4.96	0.80	3.40	5.13	8.92	3.05	3.21	0.69	2.09	2.75

NPC	0.91	0.06	0.01	1.03	0.20	13.76	0.03	2.85	0.91	0.48	0.57	0.16	0.02	0.17	2.45
FT	2.19	0.74	1.57	1.46	2.66	4.22	10.07	1.27	9.31	0.00	0.24	2.22	1.39	0.98	5.83
FFS	2.46	2.70	0.35	1.10	7.61	4.41	1.48	3.76	9.12	5.45	3.87	9.62	9.75	6.96	4.33
PTQ	1.00	0.73	5.40	0.44	3.60	4.62	3.56	1.32	4.31	4.46	6.01	1.89	18.54	4.00	7.47
PC	1.28	1.55	1.51	2.57	5.81	14.23	6.85	5.05	3.45	6.42	2.55	12.53	2.77	7.45	12.29
DFS	2.83	0.50	2.64	1.52	3.03	2.44	8.56	1.12	4.66	3.76	1.17	7.52	2.35	12.99	5.07
SF	0.56	1.61	0.36	0.06	1.65	1.12	0.06	6.51	1.50	2.72	4.19	3.50	0.45	1.28	0.79
STT	1.34	2.56	1.58	0.21	2.00	1.08	7.18	3.99	2.21	0.10	0.35	5.49	8.64	1.19	0.54
CST	2.37	0.74	0.65	0.98	0.93	2.46	3.97	7.25	2.42	1.77	0.31	0.68	3.56	0.30	2.96
CSB	2.27	0.34	1.03	1.67	2.05	0.89	4.49	9.20	3.77	0.13	2.13	0.16	0.21	0.95	5.01
Eigen value	6.62	5.34	5.13	4.18	3.49	3.25	2.95	2.91	2.55	2.52	2.40	2.31	2.23	2.12	2.03
% variance	7.35	5.93	5.69	4.64	3.88	3.61	3.28	3.24	2.83	2.79	2.67	2.57	2.48	2.35	2.26
Cumulative	7.35	13.28	18.97	23.62	27.50	31.10	34.38	37.62	40.45	43.24	45.91	48.48	50.95	53.30	55.56

PC refers to the principal components. Leaf SPAD index at 21 days after transplanting (S21) and at 28 days (S28); length of main vine at 14 days after transplanting (LMV14) and at 21 (LMV21); accumulated degree-days for flowering (DDF); number of fruits per plant (NFP); average mass of fruits (MF); productivity of fruits (PF); height of fruit (HF); diameter of fruit (DF); thickness of fruit peel (TFP); resistance of fruit peel to penetration (RFP); resistance of fruit pulp to penetration (RP); thickness of fruit pulp (PT); diameter of internal cavity of fruit (DIC); total content of fruit pulp carotenoids (TC), and lutein (L); number of seeds per plant (NSF); mass of seeds per fruit (MSF); ratio of seed to fruit mass (MS/F); mass of one hundred seeds (MOH); productivity of seeds (PS); seed thickness (ST); seed length (SL); seed width (SW); seed oil content (SOC); seed oil productivity (SOP). Growth habit (GH); stem colour (SC); intensity of leaf green (ILG), leaf silvering (LS), intensity of leaf silvering (ILS); leaf serration (LS); presence of trichomes in the leaves (PAL); amount of trichomes in the adaxial surface of leaves (ATT); amount of trichomes in the abaxial surface of leaves (ATL); leaf recess (LR); presence of trichomes in the petiole (PTP); amount of trichomes in the petiole (ATP); green intensity of male pedicel (GIMP); format of fruits (FF); format of peduncle (FP); number of colours of fruit peel (NPC); topography of fruit surface (FT); format of floral scar (FFS); peel texture (PT); predominant colour of fruit peel (PC); depth of fruits slices (DFS); seed format (SF); aspect of seed tegument (AST); seed tegument texture (STT); colour of seed tegument (CST); colour of seed border (CSB).