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

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

92 part of the *C. moschata* germplasm maintained in the BGH-UFV revealed high variability
93 among 54 accessions in terms of oleic acid content (Sobreira, 2013). In that study, BGH-7765,
94 with oil oleic acid content of 28.39%, was identified as a promising accession for use as parent
95 germplasm in breeding programs aimed at improving *C. moschata* seed oil.

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

101 **2. MATERIAL AND METHODS**

102 **2.1 Germplasm origin**

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

109

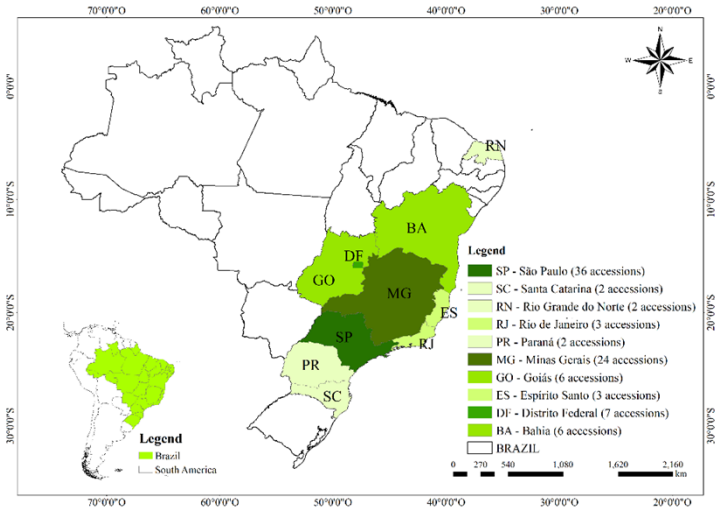

Figure 1. Brazilian map displaying the regions of origin of the *Cucurbita moschata* accessions tested in the present study.

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):

$$PS = NFP \times MSF \times 1{,}111^{*} \quad (1)$$

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

$$SOP = \frac{PS \times SOC}{100} \quad (3)$$

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·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·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)

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

$$DS = h^{2*}$$

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.

201

202 **3. RESULTS**

203 **3.1 Variance components and genetic-statistical parameters of characteristics related to**
204 **SOP and fatty acid profiles**

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

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

213 The predicted selection gains for PS and SOP were 0.266 and 0.10 t ha^{-1}, respectively
214 (Table 1). Among the characteristics related to oil fatty acid profile, oleic acid content achieved
215 the greatest selection gain (6.99%), followed by linoleic acid content (-5.12%) and PUFA
216 content (-5.11%), which also achieved considerable selection gains (Table 1). The phenotypic
217 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}
218 (phenotypic mean, 0.050 t ha^{-1}) for SOP (Table 1). Oleic acid content ranged from 16.01 to
219 40.18% (phenotypic mean, 24.55%), and linoleic acid content from 36.58 to 58.33%
220 (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_p	σ_g	σ_b	σ	*A*	h^2	*SG*	*Phenotypic range*	μ	CV_g%	CV_P %	CVr %
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_p	σ_g	σ_b	σ	*A*	h^2	*SG*	*Range*	μ	CV_g%	CV_P %	CVr %
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 (3	0.050	126.49	167.332	89.442
Traits of seed oil profile												
Traits	σ_p	σ_g	σ_b	σ	*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 (2	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 (2	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 (CVr, %) 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).

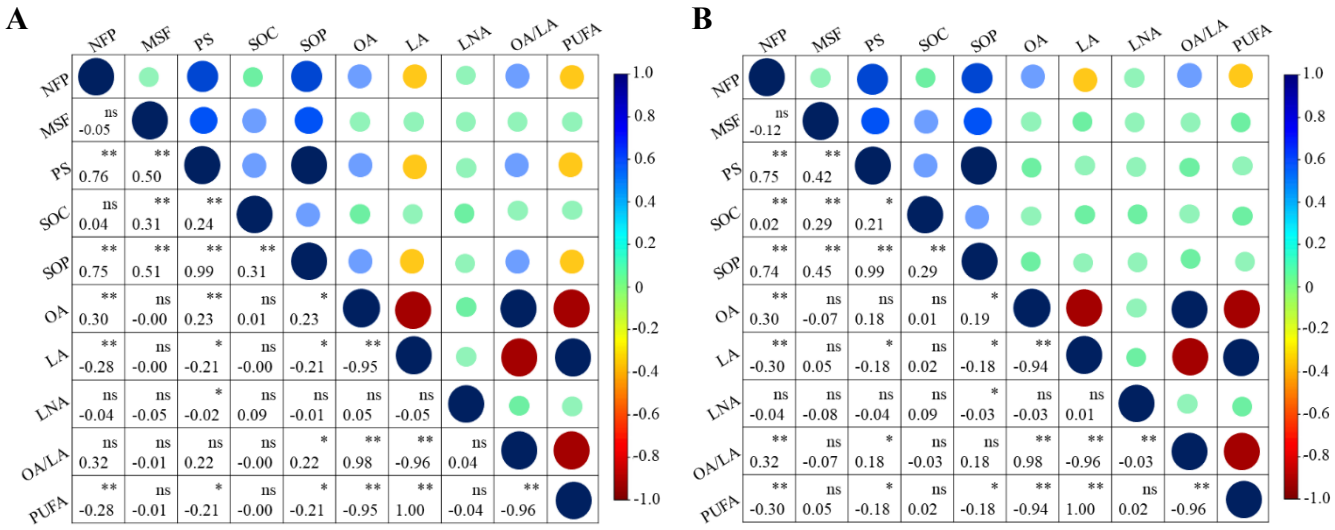

Figure 2. Phenotypic (A) and genotypic (B) correlations of characteristics related to seed oil productivity and fatty acids. Number of fruits per plant (NFP), mass of seeds per fruit (MSF), productivity of seeds (PS), seed oil content (SOC), seed oil productivity (SOP), oleic acid (OA) content, linoleic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

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

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).

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

287

Clusters	SOP	Oleico	Linoleico	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

288
289
290
291
292 The genotypic averages followed by the same letters in the column do not differ from one another based
293 on the Tocher's method of average grouping.
294

295 **3.4 Principal components analysis**

296 The first three principal components (PCs) explained 82.13% of total variation among
297 accession in terms of characteristics related to SOP and fatty acid profiles. PC1 explained
298 47.16% of total variation. The oleic acid:linoleic acid ratio and oleic acid content were the
299 characteristics with the highest positive loading, while linoleic acid content and PUFA content
300 were the characteristics with the highest negative loading on PC1 (Figure 3).

301

302

303

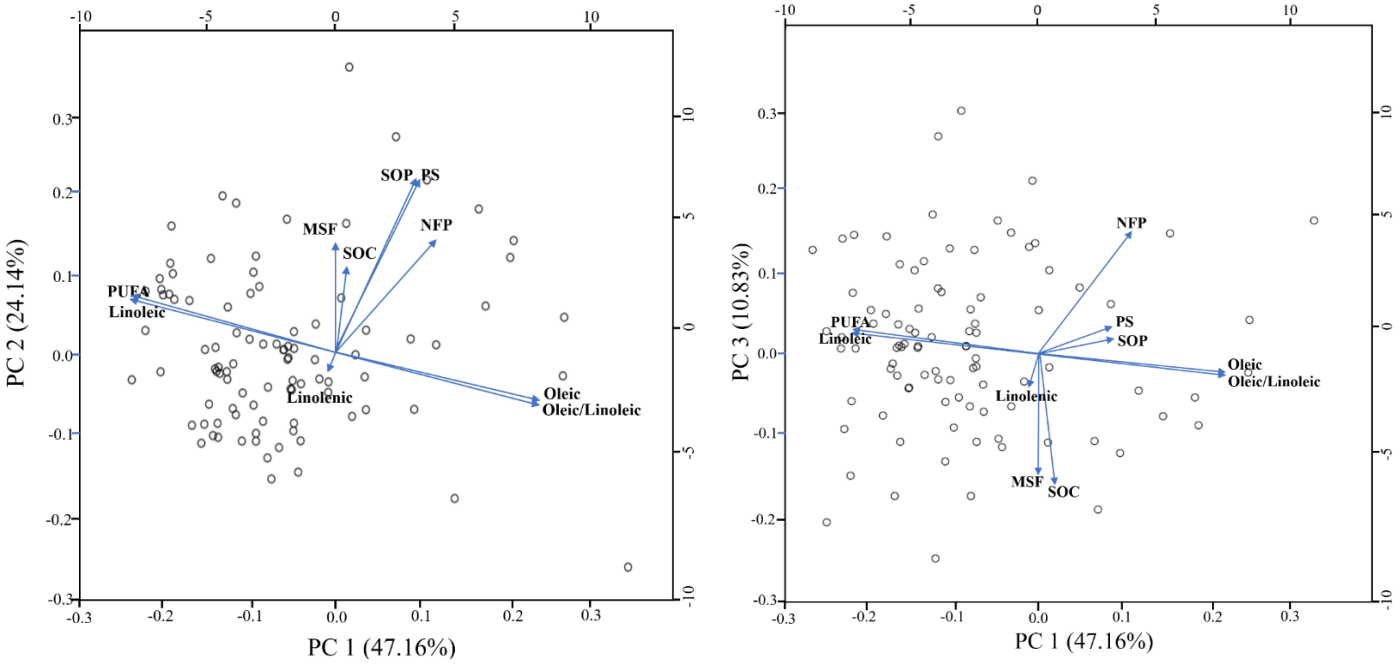

Figure 3. Dispersion of characteristics related to seed oil productivity and fatty acid profile in relation to the first three principal components. Number of fruits per plant (NFP), mass of seeds per fruit (MSF), productivity of seeds (PS), seed oil content (SOC), seed oil productivity (SOP), and polyunsaturated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

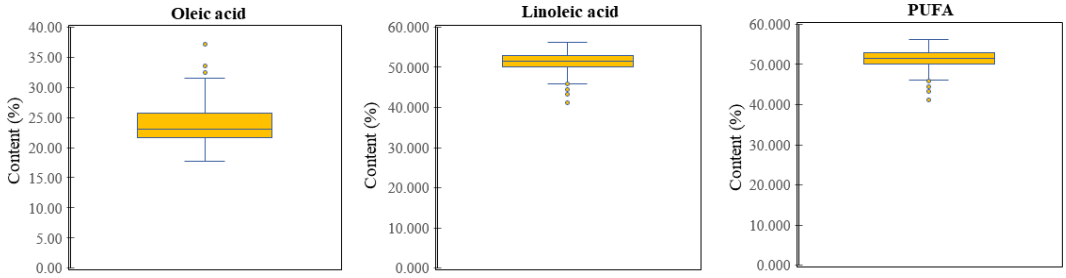

Figure 4. Box plots showing the variability in the oleic, linoleic, and polyunsaturated fatty acid content in the seed oil of the tested accessions.

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).

339 **Table 4.** Estimates of genotypic values (g) and environmental interaction-free genotypic values

340 (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).

341

342

343

344

345

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 $t \cdot ha^{-1}$ 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.*

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

375 Studies on other oilseeds, such as soybean, rapeseed, and sunflower, have often reported
376 variations in the fatty acid profiles among germplasm samples and attributed such variations to
377 genetic factors (Hemingway, et al., 2015; Yol et al., 2017). Given the quantitative nature of the
378 fatty acid profile, it is reasonable to assume that this characteristic is also influenced by
379 environmental factors, and some studies have shown the strong influence of environmental
380 factors, such as temperature, on the fatty acid profiles of various oilseeds (Werteker et al.,
381 2010). Corroborating the findings of Gomes et al. (2020), most of the accessions evaluated in
382 the present study were acquired from family farmers, who do not typically select accessions
383 for the characteristics of seed or seed oil, which possibly contributed to the maintenance of
384 high variability in these traits.

385 **4.2 Correlations of characteristics related to SOP and fatty acid profiles**

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

393 In the present study, we expected the variables NFP, MSF, and PS to be strongly
394 correlated with SOP, consistent with previous reports of strong correlation between seed yield
395 and oil productivity (Assefa et al., 2018). Thus, in addition to PS and MSF, NFP appears to be

396 a determinant of greater SOP in *C. moschata*. These correlations of SOP with its directly
397 associated characteristics can be strategic to guide indirect selection aimed at increasing *C.*
398 *moschata* oil productivity. Thus, the selection of genotypes with higher PS and MSF may be a
399 promising alternative to obtain greater SOP in *C. moschata*.

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

411 Regarding the correlations among the components of oil fatty acid profile, the strong,
412 negative but significant correlations of oleic acid content with linoleic acid and PUFA content
413 indicate an antagonistic relationship between these components of *C. moschata* seed oil (Figure
414 2). The phenotypic correlations among these characteristics may be attributed to environmental
415 and genetic factors. For instance, environmental factors, such as temperature, may strongly
416 influence the correlations among fatty acids such as oleic and linoleic acid. In line with this,
417 studies on other oilseeds, such as soybean, have shown that the oleic acid content of oil tends
418 to decrease with increasing temperature, contrary to that observed for linoleic acid content
419 (Bachlava et al., 2008), corroborating the negative correlations among these two fatty acids
420 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.

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.