scientific reports

Check for updates

OPEN

Flavonoids and their relationship with the physiological quality of seeds from different soybean genotypes

Izabela Cristina de Oliveira[1], Dthenifer Cordeiro Santana[1], João Lucas Gouveia de Oliveira[1], Elber Vinícius Martins Silva[2], Ana Carina da Silva Candido Seron[2], Matildes Blanco[2], Larissa Pereira Ribeiro Teodoro[2], Carlos Antônio da Silva Júnior[3], Fabio Henrique Rojo Baio[2], Charline Zaratin Alves[2] & Paulo Eduardo Teodoro[2]✉

Flavonoids are compounds that result from the secondary metabolism of plants and play a crucial role in plant development and mitigating biotic and abiotic stresses. The highest levels of flavonoids are found in legumes such as soybean. Breeding programs aim to increase desirable traits, such as higher flavonoid contents and vigorous seeds. Soybeans are one of the richest sources of protein in the plant kingdom and the main source of flavonoid derivatives for human health. In view of this, the hypothesis of this study is based on the possibility that the concentration of isoflavones in soybean seeds contributes to the physiological quality of the seeds. The aim of this study was to analyze the content of flavonoids in soybean genotypes and their influence on the physiological quality of the seeds. Seeds from thirty-two soybean genotypes were obtained by carrying out a field experiment during the 2021/22 crop season. The experimental design was randomized blocks with four replications and thirty-two F_3 soybean populations. The seeds obtained were subjected to germination, first germination counting, electrical conductivity and tetrazolium vigor and viability tests. After drying and milling the material from each genotype, liquid chromatography analysis was carried out to obtain flavonoids, performed at UPLC level. Data were submitted to analysis of variance and, when significant, the means were compared using the Scott-Knott test at 5% probability. The results found here show the occurrence of genotypes with higher amounts of flavonoids when compared to their peers. The flavonoid FLVD_G2 had the highest concentration and differed from the others. Thus, we can assume that the type and concentration of flavonoids does not influence the physiological quality of seeds from different soybean genotypes, but it does indirectly contribute to viability and vigor, since the genotypes with the highest FLVD_G2 levels had better FGC values. The findings indicate that there is a difference between the content of flavonoids in soybean genotypes, with a higher content of genistein. The content of flavonoids does not influence the physiological quality of seeds, but contributes to increasing viability and vigor.

Keywords Chromatography, Genistin, Germination, Isoflavones, Seed viability

Flavonoids are plant polyphenols biosynthesized in secondary metabolism, with an antioxidant function by protecting the plant from biotic and abiotic stresses[1]. These molecules play crucial roles in plant metabolism, including development, growth and ripening, preventing damage caused by pathogens and pests, and acting as chemical messengers in associations with mycorrhizae and bacteria[2].

Flavonoids generally accumulate in the vacuole of plant cells in the form of glycosides, constituting the class of isoflavones, the main molecules that accumulate in legumes, especially soybeans and their derivatives[3]. The antioxidant capacity of flavonoids is related to their stabilized structure, allowing them to act on reactive free radicals, protecting plant cells against events that damage their metabolism[4]. The benefits of flavonoids are not

[1]Department of Agronomy, State University of São Paulo (UNESP), Ilha Solteira, SP 15385-000, Brazil. [2]Federal University of Mato Grosso do Sul (UFMS), Chapadão do Sul, MS 79560-000, Brazil. [3]Department of Geography, State University of Mato Grosso (UNEMAT), Sinop, MT 78550-000, Brazil. ✉email: eduteodoro@hotmail.com

only for vegetables, but also for human health, as they can contribute to the treatment of diseases and reduce the risk of several types of cancer[5].

Soybean (*Glycine max* (L.) Merr.) is an important crop worldwide, mainly due to its nutritional properties, which include proteins, amino acids, fibers, minerals, vitamins and isoflavones that are beneficial to humans[6]. The expansion of soybean plantations and the improvement of its characteristics and grain composition are the outcome of plant breeding based on the selection of genotypes with gene expression for the desirable contents and that are expressed during seed development[7].

Soybean seeds, in turn, are responsible for perpetuating the species and propagating characteristics inherited from their parents like as the content of isoflavones, even in smaller quantities, such as genistein and daidzein, and their glycosides daidzin and genistin[8]. Variation in the composition and yield of seeds is the result of their physiological quality, which is influenced by genetic, physical, physiological and health factors[9]. Seed attributes are measured through germination and vigor tests, which seek to determine the ability of seeds to germinate and generate normal seedlings even under unfavorable growing conditions[10]. Among the vigor tests, the first germination counting (FGG), tetrazolium vigor (TZVG) and viability (TZVB) tests are the most widely used, providing fast and accurate information on the quality of the seed lot.

Soybean seeds are one of the richest sources of protein in the plant kingdom and the main source of flavonoid derivatives for human health. In light of this, the hypothesis of this study is that the concentration of isoflavones in soybean may contribute to the physiological quality of the seeds. The aim of this study is to analyze the content of flavonoids present in soybean genotypes and their influence on the physiological quality of the seeds.

Material and methods

Obtaining the seeds

Seeds from thirty-two soybean populations were obtained from the UFMS/CPCS breeding program by conducting a field experiment at the Federal University of Mato Grosso do Sul, Chapadão do Sul campus (Latitude 18°41′33″S, Longitude 52°40′45″W and Altitude 810 m), State of Mato Grosso do Sul, Brazil. The soil is classified as a clayey red oxisol with the following chemical properties in the 0–0.2 m layer: pH (H_2O) = 6.2; Exchangeable Al (cmolc dm^{-3}) = 0.0; Ca + Mg (cmolc dm^{-3}) = 4.31; P (mg dm^{-3}) = 41.3; K (cmolc dm^{-3}) = 0.2; organic matter (g dm^{-3}) = 19.74; V (%) = 45.0; m (%) = 0.0; sum of bases (cmolc dm^{-3}) = 2.3; CEC (cmolc dm^{-3}) = 5.1. The climate is classified as Tropical Savannah (Aw), with dry winter and rainy summer. Figure 1 shows the average rainfall and temperature during the experiment.

The field experiment was carried out in the 2022/2023 crop season in a randomized block design with four replications and thirty-two genotypes. The experimental plots had five rows 1.5 m long, spaced 0.45 m apart and a population of 15 plants m^{-1}.

Manual sowing took place in October 2021 with conventional soil preparation using plowing and harrowing. The seeds were treated with fungicide (Pyraclotrobin + Methyl Thiophanate), insecticide (Fipronil) and inoculant (*Bradyrhizobium* spp.) at a rate of 200 mL of the products per 100 kg of seeds. Other crop treatments were carried out according to the needs of the soybean crop to control pathogens, insects and weeds.

After the genotypes had matured, the seeds were harvested and taken to the laboratory for seed testing. The seed physiological quality and liquid chromatography tests were carried out using a composite sample of the four blocks harvested in the field, and a completely randomized experimental design was then adopted.

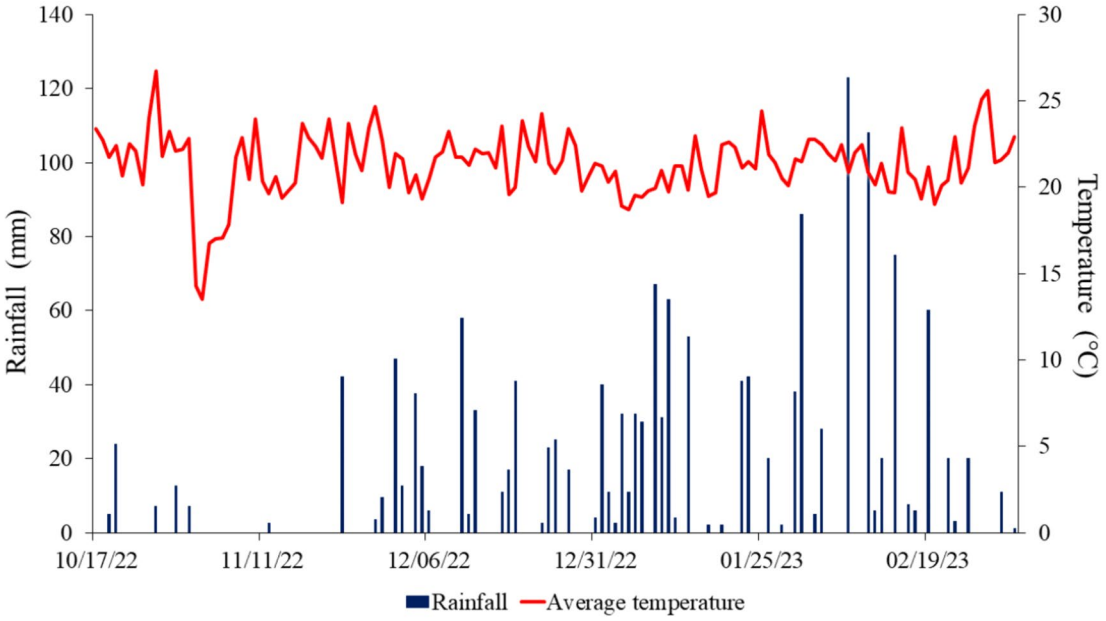

Figure 1. Rainfall and average temperature graph during the 2022/23 crop season at the experimental area.

Analysis of the physiological quality of seeds

The seeds were first subjected to water content determination using the oven method, where two sub-samples of each genotype with approximately 4.0 g of seeds were conditioned in a forced-air circulation oven at 105 ± 3 °C for 24 h (Brasil, 2009). The results were expressed as a percentage on a wet basis.

The germination test (GERM) was carried out according to the methodology proposed by Brasil (2009), where four sub-samples of 50 seeds from each genotype were distributed on germitest paper previously moistened by 2.5 times their weight and then kept in a germinator at 25 °C. The results were expressed as a percentage of normal seedlings at eight days. The first germination counting test (FGC) was carried out together with the germination test by counting normal seedlings five days after the test, and the result was expressed as a percentage of normal seedlings.

The electrical conductivity (EC) test was carried out by weighing four subsamples of 25 seeds of each genotype on a precision scale (0.0001 g) and placing them in plastic cups containing 75 mL of distilled water, which remained in the germinator for 24 h at 25 °C[12]. The results were obtained by taking readings using a conductivity meter (DIGMED DM-31), with values expressed in $\mu S\ cm^{-1}\ g^{-1}$ seeds.

The tetrazolium test was carried out according to the methodology described by França-Neto and Krzyzanowski (2018), in which four subsamples of 25 seeds from each genotype were initially pre-soaked in Germitest paper moistened 2.5 times their weight and kept in the germinator at 25 °C for 16 h. After this period, the seeds were soaked in a 2,3,5-trigenyltetrazolium chloride solution and placed in a B.O.D. in the dark at 40 °C for three hours. The seeds were then individually assessed and classified in terms of vigor (TZVG) and viability (TZVB), with the results expressed as a percentage.

After carrying out the physiological quality tests, 40 g of seeds from each genotype were dried and ground and then 0.005 g were weighed to quantify the flavonoids present in the soybean seeds.

Liquid chromatography analysis to obtain flavonoids

For the analysis, the seed samples were dried and then ground. To extract the isoflavones, 50 mg of the samples were added to a 2 mL eppendorf, in which 1.5 mL of 70% methanol containing acetic acid (0.1%) was added. The solution was shaken briefly and then incubated for 2 h in ultrasound. Subsequently, the samples were centrifuged at 5.000 rpm for 20 min and the supernatant obtained was filtered using a syringe with a 0.2 µm filter and transferred to 1.5 mL vials before injection into an ultra-performance liquid chromatography (UPLC) system. Aliquots of 10 µL were used for direct injection into the equipment. Each sample was analyzed three times[14].

The isoflavones were separated and quantified using a Waters Acquity 1100 series UPLC liquid chromatograph with automatic sample injector. An HSS C18 reverse phase column, 1.8 µm (internal diameter 2.1 mm (i.d.) £ 100 mm) with an Acquity HSS C18 pre-column, 1.8 µm (2.1 mm i.d. £ 5 mm) was used. A binary linear gradient system was used to separate the isoflavones, with the following mobile phases: Milli-Q water and 0.1% acetic acid as solvent A and acetonitrile and 0.1% acetic acid as solvent B. The initial gradient was 99% for solvent A and 1.0% for solvent B from 0 to 9 min, 41.2% A and 58.8% B from 9 to 9.1 min, 100% B from 9.1 to 11 min and returning to 99% A and 1% B at 11 min and remaining that way until 15 min, which was the running time for each sample[14]. The mobile phase flow rate was 0.289 mL min⁻1 and the column temperature during the run was 30 °C. Isoflavones were detected using a Waters photo diode array detector set to a wavelength of 254 nm. For detecting isoflavones, we used commercially acquired standards of daidzein (FLVD_D1), genistein (FLVD_G1) and genistin (FLVD_G2) solubilized in 70% methanol and acetic acid (0.1%) at the following concentrations: 0.000125, 0.0002, 0.0005, 0.001, 0.01, and 0.02 mg mL^{-1}. The qualitative and quantitative identity of the peak was confirmed by comparing the retention times and UV spectra of individual compounds and by the standard addition method.

All the solvents used in the chromatographic analyses were HPLC grade, and before use were vacuum filtered through a 0.2 µm pore membrane and then degassed in a vacuum system using ultrasound. The water used was distilled and then ultra-purified in a Milli-Q system and then degassed.

Statistical analysis

Data was initially subjected to analysis of variance. For the statistical analyses, data were analyzed in a completely randomized design using the F test at 5% probability and, when significant, the means were compared using the Scott-Knott test at 5% probability[15].

Subsequently, a canonical variables graph was constructed to evaluate the relationship between the variables analyzed and the thirty-two soybean populations, and a Pearson correlation analysis was carried out to observe the correlation between the variables studied. All the analyses were carried out using the R software[16].

Results

The physiological variables and flavonoid content of the thirty-two soybean genotypes differed in relation to the FGC and genistin (G2) variables (Table 1). Thus, the FGC and G2 variables were compared in relation to their means, highlighting the genotypes that differed from the others.

In the mean comparison test, it can be seen that genotypes G14, G15, G17, G18, G19, G23 and G24 showed inferior behavior for FGC (Table 2). The other genotypes showed superior behavior for this variable, with results above 88% PCG. Genotypes G1, G2 and G3 stood out from the others in terms of G2 content.

A canonical variables graph was constructed to assess the relationship between the variables analyzed and the soybean genotypes (Fig. 2). There is a certain proximity between the FLVD_G2 vector and the G1, G2 and G3 soybean genotypes, as also indicated by the mean comparison test (Table 2). Genotypes G14, G17, G18, G19 and G26 were also close to each other and genotypes G15 and G24 were related to the vectors of the flavonoid

SV	DF	EC	FGC	GERM	TZVG	TZVB
Genotypes	31	0.0027^{ns}	65.6*	31.2^{ns}	$75,84^{ns}$	$27,08^{ns}$
Residual	62	0.0028	31	20.84	77,78	19,11
CV (%)		78.02	6.33	4.95	11,61	4,7
SV	DF	FLVD_D1	FLVD_G1		FLVD_G2	
Genotypes	31	5.80^{-08ns}	2.20^{-05ns}		5.78^{-07*}	
Residual	62	4.31^{-08}	2.23^{-05}		3.50^{-07}	
CV (%)		100.85	343.84		107.65	

Table 1. Summary of the analysis of variance for the seed physiological variables electrical conductivity (EC), first germination count (FGC), germination (GERM), tetrazolium vigor (TZVG) and tetrazolium viability (TZVB) and flavonoid contents daidzein (FLVD_D1), genistein (FLVD_G1) and genistin (FLVD_G2) for thirty-two soybean genotypes. Source: Elaborated by the authors. *FV* source of variation, *CV (%)* coefficient of variation, *DF* degrees of freedom. *Significant and ns not significant at 5% probability by the F test.

Genotypes	FGC (%)	FLVD_G2 (mg mL^{-1})
G1	89.33 a	0.00214 a
G2	90.33 a	0.00168 a
G3	86.67 a	0.00182 a
G4	93.00 a	0.00045 b
G5	89.67 a	0.00039 b
G6	90.33 a	0.00043 b
G7	92.00 a	0.00038 b
G8	89.00 a	0.00043 b
G9	90.67 a	0.00043 b
G10	89.67 a	0.00037 b
G11	92.00 a	0.00047 b
G12	88.00 a	0.00037 b
G13	88.00 a	0.00039 b
G14	80.67 b	0.00044 b
G15	74.00 b	0.00043 b
G16	90.67 a	0.00040 b
G17	84.67 b	0.00042 b
G18	82.00 b	0.00038 b
G19	83.67 b	0.00041 b
G20	90.00 a	0.00044 b
G21	88.67 a	0.00037 b
G22	90.00 a	0.00043 b
G23	83.33 b	0.00041 b
G24	75.00 b	0.00041 b
G25	92.00 a	0.00042 b
G26	87.67 a	0.00039 b
G27	88.00 a	0.00041 b
G28	92.00 a	0.00043 b
G29	91.00 a	0.00043 b
G30	90.67 a	0.00040 b
G31	91.00 a	0.00041 b
G32	91.67 a	0.00042 b

Table 2. Comparison of means for the variables first germination counting (FGC) and flavonoid genistein concentration (FLVD_G2) for thirty-two soybean genotypes. Different letters in the same column differ by the Scott-Knott test at 5% probability.

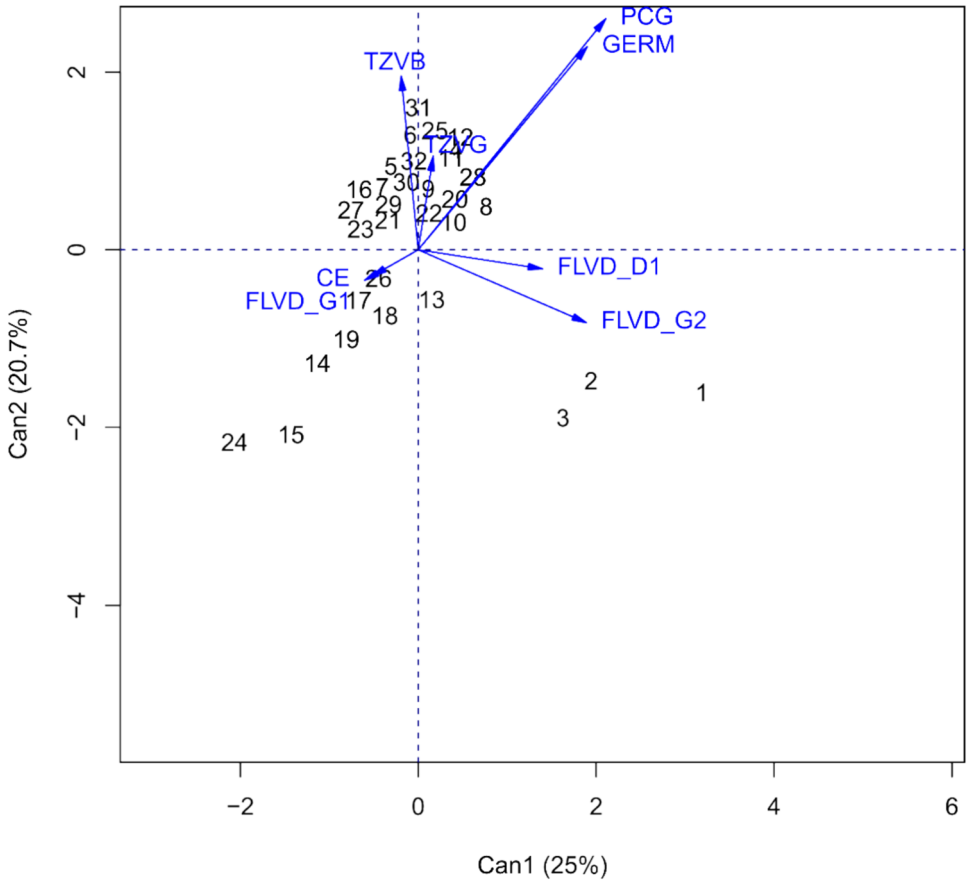

Figure 2. Canonical variable analysis for the seed physiological variables electrical conductivity (EC), first germination count (PCG), germination (GERM), tetrazolium vigor (TZVG) and tetrazolium viability (TZVB) tests, flavonoid contents daidzein (FLVD_D1), genistein (FLVD_G1) and genistein (FLVD_G2) for thirty-two soybean genotypes.

genistein (FLVD_G1) and electrical conductivity (EC) variables. The other genotypes were grouped together and close to the other variables.

Pearson's correlation graph shows high correlations between the variables FGC and GERM, and negative correlations between tetrazolium vigor (TZVB) and FLVD_G2 (Fig. 3). A moderate correlation can be observed between the flavonoid daidzein (FLVD_D1) and FLVD_G2. The other correlations between the variables were very weak or non-existent.

Discussion

The genotypes studied were significant for the variables FGC and genistein (FLVD_G2) (Table 1). FGC is a vigor test carried out on seeds with the purpose of identifying the genotypes with the greatest capacity to germinate and generate normal seedlings quickly under unfavorable environmental conditions[17]. Meanwhile, flavonoids are compounds that result from the secondary metabolism of plants and play an important role in the antioxidant and defense metabolism of plants[18]. Thus, the two variables assess which seeds have the best physiological characteristics, directly related to greater control of free radicals and oxidative metabolism, factors that interfere with the ability to germinate and generate vigorous plants[1].

The highest concentrations of FLVD_G2 were observed in genotypes G1, G2 and G3 (Table 2). With the exception of genotypes G14, G15, G17, G18, G19, G23 and G24, the other genotypes showed values above 85% FGC. Flavonoids are compounds secreted by plant roots as a mechanism for establishing an association with rhizobia, helping to form nodules on the roots and improving nutrient uptake[19]. However, there are classes of flavonoids that are also found in leaves, flowers and seeds, playing a protective role against ultraviolet light, defending against abiotic stresses and regulating the movement of auxins in plants[20]. Genistein is an isoflavone, a class of flavonoids, which accounts for approximately 50% of all the isoflavones present in soybeans. However, the bioactive properties present in the leaves and seeds are dependent on the genomic composition and environmental conditions where the plants are grown[21]. Among legumes, soybeans are richest in isoflavones, including genistein and daidzein, which play an important role in defense metabolism[22]. Thus, we can infer that genotypes G1, G2 and G3 have a higher concentration of FLVD_G2 due to their genetic constitution, contributing to the seeds showing better FGC values.

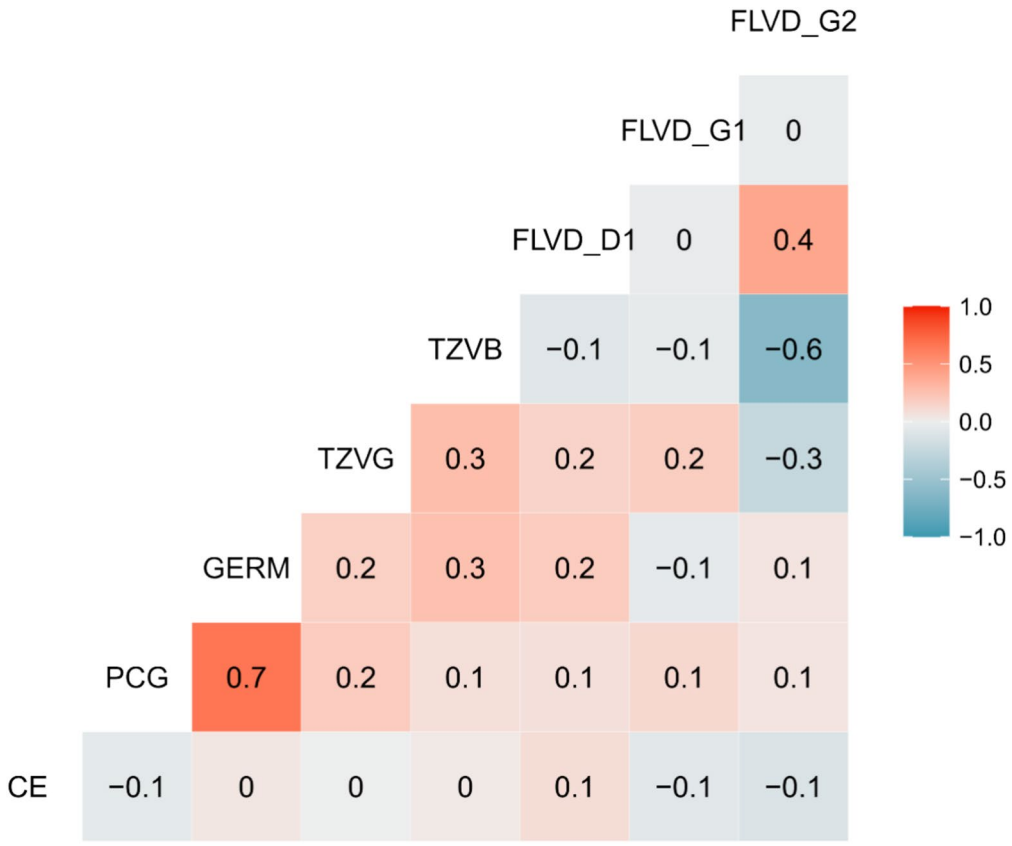

Figure 3. Pearson's correlation for the seed physiological variables electrical conductivity (CE), first germination counting (PCG), germination (GERM), tetrazolium vigor (TZVG) and tetrazolium viability (TZVB) and flavonoid contents daidzein (FLVD_D1), genistein (FLVD_G1) and genistin (FLVD_G2) for 32 soybean genotypes.

Canonical variable analysis found that the FLVD_G2 vector was close to genotypes G1, G2 and G3 (Fig. 2). Principal component analysis (PCA) are often equivalent[23], So PCA was applied to mean values of the measured traits to identify the most important factors and to explain the relationships between variables and observations[24]. This finding supports the ones shown in (Table 2), where these genotypes have higher genistein contents. The occurrence of higher genistein levels indicates that these genotypes have a high intrinsic potential to induce the elimination of reactive oxygen species (ROS) when the seeds are exposed to conditions that promote their development, such as abiotic stress that activates cellular respiration[2]. Its antioxidant action is due to the numerous hydroxyl groups in its molecules, eliminating free radicals and inhibiting enzymes that generate free radicals[25]. However, as already mentioned, the regulation of flavonoid production, but specifically isoflavones in soybeans, is a genetically mediated activity, explaining the high concentration in genotypes G1, G2 and G3 and the non-occurrence in the other genotypes.

The flavonoid genistein (FLVD_G1) and electrical conductivity (EC) vectors overlapped, with proximity to genotypes G14, G17, G18, G19 and G26. EC is a vigor test based on the cell membrane integrity, where the leachate content, mainly sugars and amino acids, determines the vigor of soybean seeds[12]. FLVD_G1, in turn, shows the presence of genistein, an isoflavone that together with daidzein, comprises the highest concentration of flavonoids in soybeans, which can also occur in the form of their glycosides, since genistein is a by-product of genistein[21]. The proximity of both vectors indicates that the seeds of the aforementioned genotypes show similar results in terms of FLVD_G1 composition, but genotypes G14, G17 and G19 do not show good results regarding vigor, as demonstrated by the FGC test. Thus, this result may be related to the higher concentrations of FLVD_G1 in genotype G26 and the high EC values for genotypes G14, G17, G18 and G19, indicating that these genotypes do not have good membrane integrity and are therefore more susceptible to pathogen attack during primary root emission, resulting in low vigor.

FGC and GERM vectors are overlapping, and the vectors for tetrazolium vigor (TZVG) and tetrazolium viability (TZVB) are close to each other and with the other genotypes close to them. Genotypes G20 and G28 are located on the FGC and GERM vectors, indicating that both show good results for these variables, which are closely related. The physiological quality of seeds, studied by GERM and vigor tests such as FGC, TZVG and TZVB, is influenced by physiological, physical, health and genetic factors, which are fundamental pillars for seeds to express their vital functions[10], because the lengths of the CVA axes and the angles between them reveal the interrelationships between the parameters[24,26]

The correlation found between FGC and GERM (Fig. 3) is expected, since the assessments are carried out jointly in the same test. The correlation between the flavonoid daidzein (FLVD_D1) and FLVD_G2 may be related to the fact that soybean is a source of isoflavones such as daidzein and genistein, with daidzein being the most abundant product and genistein, a glycoside of genistein, the second most abundant[27]. The negative correlation between TZVB and FLVD_G2 may be the result of the low presence of daidzin in the seeds of the genotypes, especially in relation to their viability, as assessed by TZVB. Daidzin is the glycoside form of daidzein, which is more soluble in water and therefore more suitable for storage in plant cell vacuoles, as well as being more stable and resistant to enzymatic degradation[3]. The TZVB variable, in turn, theoretically assesses the ability of seeds to emit a primary root when under the action of factors that delay germination[17]. Thus, the negative correlation found between TZVB and FLVD_2 may be related to the occurrence of viable seeds, which do not require the activation of antioxidant molecules such as daidzin.

The results found here show the occurrence of genotypes with higher amounts of flavonoids when compared to their peers. The flavonoid FLVD_G2 had the highest concentration and differed from the others. Thus, we can assume that the type and concentration of flavonoids does not influence the physiological quality of seeds from different soybean genotypes, but it does indirectly contribute to viability and vigor, since the genotypes with the highest FLVD_G2 levels had better FGC values.

Conclusion

Our results evidence differences in the content of flavonoids between soybean genotypes, which is a genetically modulated characteristic. The occurrence of genotypes with higher genistein content did not influence the occurrence of seeds with higher physiological quality, but it did contribute indirectly to the viability and vigor of the seeds by minimizing free radicals that would impair the seeds from expressing their maximum potential. Future studies should investigate the quantification of isoflavone molecules in soybean seeds and their relationship with the vigor of seeds of different soybean genotypes.

Soybean genotypes differ in the content of flavonoids in the seeds. The flavonoid genistein has a higher concentration in certain soybean genotypes. The content of flavonoids does not influence the physiological quality of soybean seeds, but it does contribute to increasing viability and vigor.

Data availability

The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

Received: 19 March 2024; Accepted: 19 July 2024
Published online: 24 July 2024

References

1. Chen, Z. *et al.* Integrative analysis of metabolome and transcriptome reveals the improvements of seed quality in vegetable soybean (Glycine max (L.) Merr.). *Phytochemistry* https://doi.org/10.1016/j.phytochem.2022.113216 (2022).
2. Shen, N. *et al.* Plant flavonoids: Classification, distribution, biosynthesis, and antioxidant activity. *Food Chem.* https://doi.org/10.1016/j.foodchem.2022.132531 (2022).
3. Ku, Y. S. *et al.* Understanding the composition, biosynthesis, accumulation and transport of flavonoids in crops for the promotion of crops as healthy sources of flavonoids for human consumption. *Nutrients* https://doi.org/10.3390/nu12061717 (2020).
4. Jabbari, M. & Jabbari, A. Antioxidant potential and DPPH radical scavenging kinetics of water-insoluble flavonoid naringenin in aqueous solution of micelles. *Coll. Surf. A Physicochem. Eng. Asp.* **489**, 392–399. https://doi.org/10.1016/j.colsurfa.2015.11.022 (2016).
5. Hsiao, Y. H., Hon, C. T. & Pan, M. H. Bioavailability and health benefits of major isoflavone aglycones and their metabolites. *J. Funct. Foods* https://doi.org/10.1016/j.jff.2020.104164 (2020).
6. Jiang, G. L., Katuuramu, D. N., Xu, Y., Ren, S. & Rutto, L. K. Analysis and comparison of seed protein, oil, and sugars in edamame dried using two oven-drying methods and mature soybeans. *J. Sci. Food Agric.* **100**(10), 3987–3994. https://doi.org/10.1002/jsfa.10443 (2020).
7. Sun, S. *et al.* Analysis of spatio-temporal transcriptome profiles of soybean (glycine max) tissues during early seed development. *Int. J. Mol. Sci.* **21**(20), 1–21. https://doi.org/10.3390/ijms21207603 (2020).
8. Kuligowski, M., Sobkowiak, D., Polanowska, K. & Jasińska-Kuligowska, I. Effect of different processing methods on isoflavone content in soybeans and soy products. *J. Food Compos. Anal.* https://doi.org/10.1016/j.jfca.2022.104535 (2022).
9. Meneguzzo, M. R. R. *et al.* Seedling length and soybean seed vigor. *Cienc. R.* https://doi.org/10.1590/0103-8478cr20190495 (2021).
10. Marcos Filho, J. Seed vigor testing: An overview of the past, present and future perspective. *Sci. Agric.* **72**, 363–374 (2015).
11. Brasil. Ministério da Agricultura, Pecuária e Abastecimento. Secretaria de Defesa Agropecuária. *Regras para análise de sementes.* Brasília. Avaliable on: https://www.gov.br/agricultura/pt-br/assuntos/insumosagropecuarios/arquivos-publicacoes-insumos/2946_regras_analise__sementes.pdf (2009).
12. do Prado, J. P., Krzyzanowski, F. C., Martins, C. C. & Vieira, R. D. Physiological potential of soybean seeds and its relationshipto electrical conductivity. *J. Seed Sci.* **41**, 407–415 (2019).
13. França Neto, J. D. B., Krzyzanowski, F. C., & Henning, A. A. A importancia do Uso de Semente de Soja de Alta Qualidade. Londrina. Avaliable on: https://ainfo.cnptia.embrapa.br/digital/bitstream/item/49831/1/ID-30537.pdf (2010).
14. Carrão-Panizzi, M. C., de Favoni, S. P. & Kikuchi, A. Extraction time for soybean isoflavone determination. *Braz. Arch. Biol. Technol.* **45**, 515–518 (2002).
15. AJ Scott, Knott M. A cluster analysis method for grouping means in the analysis of variance. http://www.jstor.org (1974).
16. R Development Core Team R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Avaliable on: https://www.R-project.org/ (2013).
17. Reed, R. C., Bradford, K. J. & Khanday, I. Seed germination and vigor: Ensuring crop sustainability in a changing climate. *Heredity* **128**(6), 450–459 (2022).
18. Hernández-Rodríguez, P., Baquero, L. P. & Larrota, H. R. Flavonoids. In *Bioactive Compounds: Health Benefits and Potential Applications* (eds Hernández-Rodríguez, P. *et al.*) (Elsevier, 2019).

19. Li, K. *et al.* Structure-activity relationship of eight high content flavonoids analyzed with a preliminary assign-score method and their contribution to antioxidant ability of flavonoids-rich extract from *Scutellaria baicalensis* shoots. *Arab. J. Chem.* **11**(2), 159–170. https://doi.org/10.1016/j.arabjc.2017.08.002 (2018).
20. Sandoval-Yañez, C., Mascayano, C. & Martínez-Araya, J. I. A theoretical assessment of antioxidant capacity of flavonoids by means of local hyper–softness. *Arab. J. Chem.* **11**(4), 554–563. https://doi.org/10.1016/j.arabjc.2017.10.011 (2018).
21. Kim, M. A. & Kim, M. J. Isoflavone profiles and antioxidant properties in different parts of soybean sprout. *J. Food Sci.* **85**(3), 689–695. https://doi.org/10.1111/1750-3841.15058 (2020).
22. Rojas, C. M., Senthil-Kumar, M., Tzin, V. & Mysore, K. S. Regulation of primary plant metabolism during plant-pathogen interactions and its contribution to plant defense. *Front. Plant Sci.* https://doi.org/10.3389/fpls.2014.00017 (2014).
23. Peltier, C., Visalli, M. & Schlich, P. Comparison of canonical variate analysis and principal component analysis on 422 descriptive sensory studies. *Food Qual. Prefer.* **40**, 326–333 (2015).
24. Ozaktan, H. & Doymaz, A. Mineral composition and technological and morphological performance of beans as influenced by organic seaweed-extracted fertilizers applied in different growth stages. *J. Food Compos. Anal.* **114**, 104741 (2022).
25. Nagarajan, S. *et al.* Antioxidant activity of synthetic polymers of phenolic compounds. *Polymers* **12**(8), 1646 (2020).
26. Çetin, N., Ozaktan, H., Uzun, S., Uzun, O. & Ciftci, C. Y. Machine learning based mass prediction and discrimination of chickpea (*Cicer arietinum* L.) cultivars. *Euphytica* **219**, 20 (2023).
27. Liu, J. Y. H. *et al.* Soy flavonoids prevent cognitive deficits induced by intra-gastrointestinal administration of beta-amyloid. *Food Chem. Toxicol.* **141**, 111396. https://doi.org/10.1016/j.fct.2020.111396 (2020).

Author contributions

Izabela Cristina de Oliveira: conducted the experiments in the field, seed quality analyzes and wrote the initial version of the manuscript; Dthenifer Cordeiro Santana: helped in conducting field experiments and seed quality analyzes; João Lucas Gouveia de Oliveira: helped in conducting field experiments Elber Vinícius Martins Silva: helped in conducting field experiments Ana Carina da Silva Candido Seron: helped in UPLC analyzes; Matildes Blanco helped in UPLC analyzes; Larissa Pereira Ribeiro Teodoro: assisted with data analysis and discussion of results; Carlos Antônio da Silva Júnior: assisted in acquiring funds to carry out the research Fabio Henrique Rojo Baio: assisted in acquiring funds to carry out the research Charline Zaratin Alves: helped in seed quality analyzes; Paulo Eduardo Teodoro: assisted in planning the research, acquiring funds to carry out the research and writing the manuscript. All authors read the manuscript and contributed to its final, improved version. All authors of the manuscript have read and agreed to its content and are accountable for all aspects of the accuracy and integrity of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to P.E.T.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.