BRAZILIAN JOURNAL OF FOOD TECHNOLOGY

ISSN 1981-6723
Online version

ORIGINAL ARTICLE

Soybean germplasm characterization for human consumption aptitude in Uruguay

Caracterização do germoplasma de soja com aptidão ao consumo humano no Uruguai

Giuliana Vera[1], Federico Condón Priano[1], Daniel Vázquez[2]*

[1]Instituto Nacional de Investigación Agropecuaria (INIA), Sistema Agrícola Ganadero, La Estanzuela/Colonia - Uruguay
[2]Instituto Nacional de Investigación Agropecuaria (INIA), Agroalimentos, La Estanzuela/Colonia - Uruguay

***Corresponding Author:** Daniel Vázquez, Instituto Nacional de Investigación Agropecuaria (INIA), Agroalimentos, Ruta 50, km 11.500, 70000, La Estanzuela/Colonia - Uruguay, e-mail: dvazquez@inia.org.uy

Cite as: Vera G., Condón Priano F., & Vázquez, D. (2024). Soybean germplasm characterization for human consumption aptitude in Uruguay. *Brazilian Journal of Food Technology*, 27, e2023048. https://doi.org/10.1590/1981-6723.04823

Abstract

Soybean (*Glycine max* (L.) Merrill) is a crop of world economic importance; and its grain constitutes a significant source of protein and carbohydrates for human consumption. This work aimed to characterize soybean genotypes available in Uruguay for human consumption regarding protein quantity and quality, carbohydrate composition and oil content in relation to yield, both in genetically modified and conventional genotypes. In this study, 13 genotypes grown in three different environments (two locations, two years), a conventional set (22 genotypes) and a genetically modified set (36 genotypes), grown in a single environment were subject to study. The traits measured were yield, total protein, oil content, soluble protein, glycinin (11S), conglycinin (7S), the 11S/7S ratio, soluble carbohydrates, sucrose and total raffinose family oligosaccharides (RFOs). A significant environmental and genotype effect was detected for most of the parameters. The interaction between genotype and environment was significant for total protein, oil content and sucrose. Soluble protein and the 11S fraction were only affected by environment; soluble carbohydrates and the 7S fraction were only affected by genotype. It was possible to identify genotypes with good characteristics for human consumption (high total protein, sucrose, ratio 11S/7S and low oligosaccharides) across environments, both genetically modified and conventional genotypes. Variability was found among the genotypes available in Uruguay in the parameters studied.

Keywords: Soybean; Soy foods; Genotypes of Uruguay; Soybean protein; Soybean carbohydrates; Soybean composition; Genotype variability; Environmental variability.

Resumo

A soja (*Glycine max* L. Merrill) é uma cultura de importância econômica mundial, sendo que seu grão constitui uma importante fonte de proteínas e carboidratos para a alimentação humana. O objetivo deste trabalho é caracterizar genótipos de soja disponíveis no Uruguai para o consumo humano em termos de quantidade e qualidade proteicas, composição de carboidratos e teor de óleo, em relação ao rendimento, tanto em genótipos geneticamente modificados quanto em convencionais. Neste trabalho, foram estudados 13 genótipos de soja cultivados em três ambientes diferentes (duas localidades, dois anos), além de um grupo convencional (22 genótipos) e outro geneticamente modificado (36 genótipos), cultivados em um único ambiente. Os parâmetros avaliados foram

rendimento, proteína total, teor de óleo, proteína solúvel, glicinina (11S), conglicinina (7S), relação 11S/7S, carboidratos solúveis, sacarose e oligossacarídeos totais da família da rafinose. Efeito ambiental e genotípico significativo foi detectado para a maioria dos parâmetros analisados. A interação entre genótipo e ambiente foi significativa para proteína total, teor de óleo e sacarose. A proteína solúvel e a fração 11S foram afetadas apenas pelo ambiente, e os carboidratos solúveis e a fração 7S foram afetados apenas pelo genótipo. Foi possível identificar genótipos com boas características para consumo humano (alto teor de proteína total, sacarose, relação 11S/7S e baixos oligossacarídeos) em todos os ambientes, tanto transgênicos quanto genótipos convencionais. Os genótipos disponíveis no Uruguai apresentaram variabilidade a partir dos parâmetros estudados.

Palavras-chave: Soja; Alimentos à base de soja; Genótipos uruguaios; Proteína da soja; Carboidratos da soja; Composição da soja; Variabilidade genotípica; Variabilidade ambiental.

Highlights

- Soybean genotypes in Uruguay showed variability in human consumption quality parameters
- Genotype, environment, and their interaction influenced key quality parameters
- Genotypes with a desirable combination of quality parameters were identified

1 Introduction

There is a growing awareness in the population about consuming healthy and sustainable foods that has led to a worldwide interest rise in plant protein sources as alternatives to meat. Ethical and religious issues are also leading to concerns surrounding animal-based proteins (Hartmann & Siegrist, 2017). Population growth, expected to reach 9 billion by 2050 (Gu et al., 2021) is another factor driving the growing interest in plant-based protein sources to meet the increasing demand for proteins (Seto & Ramankutty, 2016). Soybean (*Glycine max* (L.) Merrill) seed is the most used and characterized plant protein source (Zheng et al., 2022; Qin et al., 2022; Messina et al., 2022). Soybean derived ingredients have a significant presence in the plant based-protein industry due to their nutritional properties, bioavailability, and techno-functionalities that enhance the textural characteristics of end products (Samard & Ryu, 2019; Balestra & Petracci, 2019). In addition, they usually have low cost compared with other food ingredients. Soy proteins have long been used in different foods such as natto, tempeh, and tofu (Fukushima, 1981). Currently, these proteins are also used to make meat analogs, that is, restructured products that mimic processed meats, such as hamburgers, patties, and nuggets (Sha & Xiong, 2020).

Soybean production has grown significantly over the past 20 years, becoming one of Uruguay's leading export products; at present it is the main summer crop and the main agricultural crop of the country (Garance & Arbeletche, 2020). Additionally, the global demand of soybean for food production is increasing (Zheng et al., 2022), however, no studies on the suitability of soybean for human consumption were found.

Soybean grain has an average protein content of 38% to 40%, but it can range from 35% to 50% (Hwang et al., 2014). Genotype variability in both protein content and composition has an impact on the yield and quality of soy foods (Khatib et al., 2002; Min et al., 2005). Furthermore, Murphy & Resurreccion (1984) found that protein composition is also influenced by the environment; the most important fraction of the protein is the soluble one (15% 30%) because it can be processed and utilized in traditional soy foods. Glycinin (11S) and β-conglicinin (7S) are the major storage proteins (globulins); they both are the main groups of soluble protein and have a significant effect on food texture. The 11S fraction is richer in disulfide bonds and sulfhydryl groups which leads to a gel with higher hardness and elasticity more desirable in food production. On the other hand, 7S fraction has more hydrophilic amino acids, creating gels with lower

hardness and elasticity. Therefore, it is desirable to increase the 11S content and reduce the 7S content for both nutritive properties and processing soy food (Zhou et al., 2019). Furthermore, the ratio of 11S to 7S globulins (11S/7S) was positively and significantly correlated with tofu yield (Mujoo et al., 2003). When using SDS-PAGE (sodium dodecyl sulfate-polyacrylamide gel electrophoresis) technique, the soluble proteins 11S and 7S can be identified. The fraction 11S is a trimeric glycoprotein consisting of three major types of subunits (α′, α, β) with different combinations and physicochemical properties. The fraction 7S is a hexamer consisting of acidic (A) and basic (B) polypeptides that are linked by disulfide bridges and composed of glycinin subunits (Iibuchi & Imahori, 1978; Kitamura et al., 1976). The grain oil content is on average 20% with a variation range of 15.4-22.0% (Kumar et al., 2006). This variation can be related to both genotype and environmental factors (Assefa et al., 2018), but does not influence soybean food products (Schaefer & Love, 1992). Soluble carbohydrates represent a range of 7.0-10.0% of the grain weight and are composed mostly of sucrose (4.7-5.7%) and the raffinose family oligosaccharides (RFOs) (4.2% to 5.9%): stachyose, raffinose, and verbascose (Kumar et al., 2010; Choung, 2005). Raffinose, stachyose, and verbascose are sucrose linked with one, two, and third galactose molecules, respectively, via α1→6 glycosidic (Obendorf et al., 1998). Due to the absence of α1→6 galactosidase enzyme required to break down this linkage in the human gastrointestinal tract, the RFOs are not digested producing gases and abdominal discomfort (Karr-Lilienthal et al., 2005). Therefore, sugars affect soyfood quality and nutritional values. The benefits of soybeans for human consumption have been demonstrated, but many consumers avoid their use mainly due to the presence of off-odours and off-flavours (Esteves et al., 2010). Sucrose content contributes to the sweetness to soy-based food products like soymilk, tofu, and natto; moreover, it is the major energy source for fermentation (Taira, 1990).

The main objective in this study was to describe the diversity observed among soybean genotypes and environments for human consumption in Uruguay; specifically, we wanted to measure the variability in yield, total protein content, soluble protein, carbohydrates, and oil in adapted germplasm, including both Genetically Modified Organism (GMO) and conventional, and both commercial and advanced breeding genotypes. In addition, this study also sought to understand the effect of the genotype and the environment on the different parameters and identify genotypes with a good combination of them for human consumption.

2 Material and methods

2.1 Materials

A total of 48 soybean genotypes, including 38 from the Instituto Nacional de Investigación Agropecuaria (INIA) soybean breeding program and 10 commercial cultivars, were selected to represent the soybean genotype variability available in Uruguay (Table 1). These genotypes were sown in different environments, defined by planting date, location, and year. The evaluated genotypes included a range of maturity groups from 4.9 to 6.8, representing the maturity groups used in Uruguay (Garner & Allard, 1930). The duration of the vegetative period and the beginning of flowering represented a difference of 20 to 25 days, depending on each year's conditions and planting date, since flower induction is controlled by temperature and photoperiod. A total of 217 samples were analyzed. Each trial had two replications of 1.28 m x 4.00 m plots. A total of 13 genotypes (four conventional breeding lines, from INIAs breeding program, and 9 commercial GMO cultivars) were grown in three environments to study the influence of genotype, environment, and their interaction on different factors (this set is referred as "GxE"). The first environment (LED20) was characterized by a late planting date was December 19th 2019 at INIA La Estanzuela (34°20′16.89′′ S; 57°41′25.90′′ W); the total rainfall in this period was 375 mm (low rainfall). The second environment was November 06th 2018 sowing date, also at INIA La Estanzuela (LEN19), where the rainfall was 604 mm (medium rainfall). The third environment was obtained sowing on November 08th 2018 in Young (32°42′06.97′′ S; 57°38′17.82′′ W) (YON19), with a rainfall of 1138 mm (high rainfall). A second set was configured by 22 genotypes, composed of 12 non-GMO experimental lines and 10 GMO check commercial

cultivars (CONV). The location and sowing date were the same that LED20. The third set included 36 GMO genotypes, including 10 checks and 26 from the breeding program (GM), sown on December 18th 2020 at INIA La Estanzuela. All GMO genotypes had the Roundup Ready (RR) event transgenic.

2.2 Determination of protein and oil contents

The nitrogen content was determined in two reps by Kjeldahl Foss 2100 (Foss, Denmark), and the total protein content was estimated using factor 6.25 (ISO 20483:2013). The samples were ground in a Perten Laboratory mill 3303 model (Perten Instruments, Sweden), using position 5 and then reground in position 1. The oil content was obtained with no replicates by nuclear magnetic resonance (NMR) spectroscopy using a Spinlock SLK SG 100 model (Spinlock Magnetic Resonance Solutions, Argentina), calibrated with hexane extraction data obtained with a Twisselmann system.

Table 1. List of soybean genotypes included in the three sets characterized.

Genotypes	GxE	CONV	GM	MG	Genotypes	GxE	CONV	GM	MG
DM 50i17 IPRO*	x	x	x	5.0	SJ13619*			x	5.8
NA 5009 RG*	x	x	x	5.0	SJ12395*			x	5.2
NA 5909 RG*	x	x	x	5.9	SJ12394*			x	5.2
GE 590 CI*	x	x	x	5.9	SJ14507*			x	6.6
NA 5509 RG*	x	x	x	5.5	SJ13616*			x	6.5
5958 RSF IPRO*	x	x	x	5.9	SJ13425*			x	6.8
62R63 RSF*	x	x	x	6.2	SJ13327*			x	6.7
DM 6.8i*	x	x	x	6.8	SJ12210*			x	6.6
DM 6.2i*	x	x	x	6.2	SJ13618*			x	6.5
5351 RSF*		x	x	5.3	SJ14502*			x	6.6
SJ13621	x	x		5.7	SJ13626*			x	5.5
SJ13623	x	x		5.9	SJ13371*			x	6.0
SJ13624	x	x		6.2	SJ14504*			x	5.9
SJ13625	x	x		6.2	SJ14490*			x	5.0
SJ14494		x		6.0	SJ13064*			x	5.0
SJ14497		x		4.9	SJ13615*			x	5.2
SJ14498		x		5.0	GENESIS 6201*			x	6.2
SJ14508		x		6.2	GENESIS 5601*			x	5.6
SJ14509		x		6.6	GENESIS 6301*			x	6.3
SJ14511		x		5.7	GENESIS 5602*			x	5.8
SJ14513		x		6.1	GENESIS 5501*			x	5.5
SJ14514		x		5.5	GENESIS 6602*			x	6.6
SJ13614*			x	5.9	GENESIS 5901*			x	5.9
SJ14505*			x	6.0	FS 59*			x	5.9

GxE: set of 13 genotypes grown in three environments; CONV: set of 22 genotypes grown in one environment; GM: set of 36 genotypes grown in one environment. MG: Maturity groups. * indicates GMO genotypes. Genotypes 1 through 10 are commercial checks used as standard references in yield and adaptation trials. Genotypes with the prefix "SJ", "Genesis" and "FS" belong to INIA`s breeding program.

2.3 Preparation of sample

The grown samples obtained for Kjeldahl were reground in a SYSPRO Lab Instruments (Agro Uruguay, Uruguay) mill. Samples were defatted with a Twisselmann system (Matthäus & Brühl, 2001); basically, 10 g were left for 3 h in the system using hexane (petroleum ether 62 ºC to 68 °C, Cicarelli, Argentina). The defatted samples were used to measure soluble protein and carbohydrates.

2.4 Protein fraction

2.4.1 Soluble protein content

Protein extraction was done as described by Stanojevic et al. (2011); basically, 50 mg of sample were extracted in an Eppendorf tube with 1 mL of extraction buffer (0.03 M Tris-HCl, pH 8.0, 0.01 M β-mercaptoethanol) obtaining a sample to buffer ratio of 1:20. The tube was vortexed every 30 min for 2 h and it was centrifuged using a centrifuge Hermle Z 300 K model (Labnet, USA) at 8000 rpm (5018g) for 20 min, obtaining the protein extract. The protein content in the supernatant was determined in duplicates using Bradford reagent (Sigma-Aldrich, USA), with bovine serum albumin (Amresco, USA) as standard.

2.4.2 Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE)

The procedure of SDS-PAGE was according to Laemmli (1970), with minor modifications. The separating gel was 10% acrylamide, 0.375 M Tris-HCl pH 8.0, 0.1% SDS, 0.1% tetramethylethylenediamine (TEMED), 0.04% ammonium persulphate (APS); and stacking gel was 5% acrylamide, 0.125 M Tris-HCl pH 6.8, 0.1% SDS, 0.1% TEMED and 0.04% APS.

The protein extract was diluted to a concentration of 2.5 µg/µL of soy protein in a ratio 1:8 sample/extraction buffer; 10 µL sample was loaded into each well (i.e., about 25 µg of protein). Then, loading buffer (0.01 M Tris-HCl, 1% SDS, 0.1% dithiothreitol, 0.05% bromophenol blue, 10% glycerin), was added in a 1:1 ratio, heated at 95 °C for 10 min, and cooled to room temperature. Each sample was seeded in duplicate. The gels were run in a buffer solution of pH 8.3 [0.025 M Tris, 0.192 M glycine, and 0.1% SDS] at 80 V until the line passed the partition between the gels and then changed to 120 V. Gels were stained by shaking in the staining solution (0.1% Comassie Brillant Blue R250 dissolved in 50% trichloroacetic acid) for 1h and they were destained with 7% trichloroacetic acid overnight. The molecular weight of the bands was estimated using molecular weight markers of Thermo Scientific (Lithuania). These included β-galactosidase (116.0 kDa), bovine serum albumin (66.2 kDa), ovalbumin (45.0 kDa), lactate dehydrogenase (35.0 kDa), REase Bsp98I (25.0 kDa), β-lactoglobulin (18.4 kDa) and Lysozyme (14.4 kDa).

SDS-PAGE separated the subunits of the proteins 7S and 11S as it is shown in Figure 1. The separation between the proteins 11S and 7S is at 44 kDa (Liu et al., 2007). The identification of the bands was based on previous reports (Pesic et al., 2005; Fontes et al., 1984). The subunits identified of the 7S protein were α′, α, and β with molecular weights of 80, 70, and 50, respectively. The subunits of the 11S protein are grouped into acidic (An) and basic (Bn) ones. The band with a molecular weight of 40 kDa is the acidic A3 polypeptide and the group of polypeptides close to 35 kDa corresponded to the major group of acidic polypeptides (A1, A2, A4, A6, A7; Figure 1). The other acidic polypeptide A5 located at the end of the gel had a molecular weight of about 15 kDa. The group of protein bands with molecular weight values of approximately 20 kDa were basic components (B1, B2, B4). The other band above the basic components is the B3 polypeptide of a basic subunit of the 11S fraction.

SDS-PAGE was performed in electrophoresis unit EIDO NA-1114 (Nihon Eido Co., Japan). The gels were analyzed using Gel Doc EZ System (BIO-RAD, USA) and Image LabTM software (BIO-RAD, USA) for densitometry.

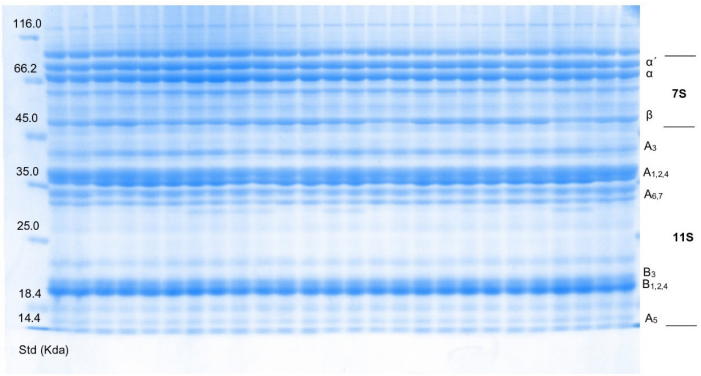

Figure 1. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) analysis of protein composition. Std, molecular weight standards (kDa). A_3, $A_{1,2,4}$, $A_{6,7}$, A_5, B_3, $B_{1,2,4}$ are polypeptides of glycine (11S; A, acidic; B, basic), and α', α, β are subunits of β-conglycinin (7S).

2.5 Determination of soluble carbohydrates

Soluble carbohydrate content was determined by the enzymatic kit procured from Megazyme International Ltd., Ireland. Since the samples were defatted, and assuming that the oil content to be around 20%, 0.40 g of sample was weighed out instead of the 0.50 g suggested by the supplier. Then, 5 mL of ethanol (95% v/v) were added and they were incubated in the water bath at 84 º C to 88 °C for 5 min to inactivate enzymes. The soy-flour reference sample included in the kit was defatted using chloroform, following the standard procedure. Volume was adjusted to 50 mL with sodium acetate buffer (50 mM, pH 4.5). Then, it was incubated for 15 min and mixed thoroughly to obtain a uniform slurry. 1.50 mL of this solution was transferred to an Eppendorf tube and centrifuged at 1000 rpm (627 g) for 10 min. 0.20 mL of the upper aqueous solution was transferred into three tubes (A, B and C). A volume of 0.20 mL of sodium acetate buffer, invertase, and a mixture of α-galactosidase/invertase was added to tubes A, B and C, respectively. All three tubes were incubated at 50 °C for 20 min. Reagent blank (0.40 mL of sodium acetate buffer) and glucose control for quadruplicated (0.10 mL of standard glucose solution and 0.30 mL of sodium acetate buffer) were included. Then, 3.0 mL of GOPOD (glucose oxidase-peroxidase) reagent was added in all tubes and incubated at 50 °C for 20 min. Absorbance was measured at 510 nm; glucose content was determined for each solution through a calibration curve; the value of sucrose and RFOs was obtained by difference. Two replications were made for each sample.

2.6 Statistical analysis

The data were analyzed with SAS software version 9.4 (SAS Institute Inc., 1993) and with InfoStat version 2020 (Di Rienzo et al., 2020). The grain composition data was analyzed using a mixed model using PROC GLIMMIX; genotype, environment, and genotype x environment interaction were considered fixed effects, while the technical repetition of each measurement for each sample at the laboratory was considered random. Adjusted means were calculated for each of the effects and their interaction. Yield and oil content were analyzed by SAS PROC GLM using a complete randomized block design. In both analyses, the means were separated using the least significant difference at the 5% significance level. The principal components analysis and the correlations among traits were calculated using InfoStat. The correlations between different parameters were also done and were significant at $p < 0.05$ level.

3 Results and discussion

3.1 Study of genotypes and environments

3.1.1 Descriptive statistics

The yield and seed quality mean, minimum, and maximum observed for the GxE study, 13 genotypes grown in three environments representative of local variability, are presented in Table 2. There were yield differences

between environments, with LED20 expressing the highest mean yield, YON19 the lowest yield and LEN19 with the widest range. Although YON19 had the best hydric conditions (1138 mm), the yield was the lowest due to lodging caused by excessive plant growth (1.40 m on average). Also, the soil in YON19 was medium textured, degraded, and generally, crop rotation was improperly (data not shown).

The results of oil, total protein, soluble carbohydrates, and soluble protein content were observed within the range of what was reported in the literature. Qin et al. (2014) measured oil values in China commercial varieties ranging from 14.2% to 22.7%; Kumar et al. (2006) reported values from 15.4% to 22.0%, also in soybean commercial from India. For total protein content, Qin et al. (2014) found values from 31.7% to 49.8% and Kumar et al. (2006) from 32.2% to 42.1%. Hwang et al. (2014) analyzed the germplasm of USDA finding a wider range in total protein (35% to 50%). Also, these protein and oil values were within the range observed by Cuitiño et al. (2019, 2020) for Uruguay. For soluble protein, Yu et al. (2016) observed a range of 26.5% to 36.0% working with 35 different commercial varieties in China and Stanojevic et al. (2011) detected values from 23.3% to 31.0% in six commercial varieties. Studying soluble carbohydrate content, Choung (2005) reported a range of 7.1% to 10.6% for 32 soybean genotypes and Yu et al. (2016) obtained broader variability (8.5% to 14.1%).

The 7S and 11S fractions content were comparable to that reported by Cai & Chang (1999), studying 13 soybean varieties (7.3-9.9% and 14.1% to 22.9%, respectively). The 11S/7S protein ratio was similar to some literature reports but also it was higher compared to others. Cai & Chang (1999) observed values from 1.6 to 2.5 and Stanojevic et al. (2011) from 1.7 to 1.9; whereas Murphy & Resurreccion, (1984) and Zilić et al. (2011) observed similar ranges (2.1 to 3.4 and 2.4 to 3.3, respectively).

The sucrose and RFOs contents were comparable with the values of Kumar et al. (2010) (1.2% to 5.7% and 2.3% to 6.1%, respectively), as well as Choung (2005) (2.6% to 6.8% and 2.2% to 5.1%, respectively) and the ones found by Yu et al. (2016) (sucrose: 1.5% to 7.3%; RFOs: 3.5% to 6.9%). However, in a wide genetic diversity study, Hou et al. (2009) investigated 241 genotypes from 28 origins (within Africa, Asia, Europe, and North and South America) detecting a broader range in both sucrose (0.2% to 9.5%) and stachyose plus raffinose (RFOs) content (0.03% to 9.0%) than that observed in this work and the cited references.

The results were within the range reported previously by other authors who studied soybean quality. In general, the observed values were similar to those found in commercial cultivars, though the study of extensive germplasm collections showed a more extended range. This indicates that the intra-specific diversity was winder than included in the studied germplasm set; this is expected since the number of genotypes studied is restricted and corresponds to commercial cultivars and breeding lines from a single breeding program, all selected for yield, not for grain quality.

Table 2. Mean, minimum and maximum of studied parameters of 13 soybean genotypes in three environments (GxE set).

Parameters	LED20			LEN19			YON19		
	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max
Yield (kg/ha)	4561	3581	5374	3898	2966	5846	3820	3000	5237
Oil (%)	20.5	20.1	21.2	20.1	19.2	20.8	21.2	20.6	22.3
Total protein (%)	41.5	40.2	43.8	41.1	39.6	43.6	39.1	37.8	40.9
Soluble protein (%)	22.9	21.7	23.9	24.1	22.7	26.1	22.4	20.3	23.6
7S (%)	6.5	5.2	7.2	6.8	6.1	7.5	6.7	5.7	7.4
11S (%)	16.4	15.3	17.7	17.3	15.9	19.0	15.7	14.4	16.9
11S/7S ratio	2.4	2.1	3.2	2.5	2.3	2.8	2.2	2.0	2.5
Sol. carb. (%)	9.5	8.8	10.0	9.6	9.1	9.9	9.4	9.1	10.0
Sucrose (%)	5.4	4.6	6.3	5.4	4.5	6.2	5.0	4.3	5.8
RFOs (%)	3.9	3.2	4.3	4.0	3.5	4.7	4.4	3.6	4.9

Min: Minimum; Max: Maximum; RFOs: Raffinose family oligosaccharides; Sol. carb.: Soluble carbohydrates.

3.1.2 Genotype and environmental effects

Table 3 exhibits the significance of the genotype, environment and their interaction effects for the studied parameters; both genotype and environment had significant effects on most of them. The environmental effect was significant for all traits except soluble carbohydrates and 7S content. For soluble protein and 11S fraction, only the environmental effect was significant. Yield, soluble carbohydrates, RFOs, 7S, and 11S/7S ratio were parameters not affected by the interaction between genotype and environment and could be differentiated using the genotype's mean. The genotype x environment interaction (GxE) was significant for oil, total protein and sucrose content; for these traits, although a high proportion of the variation was due to environment, with proportions of 87%, 90%, and 50%, respectively (Table 4), the genotype effect in sucrose accounted for 38% of the total variation.

Table 3. Genotype, environment, and their interaction effects on yield and grain quality parameters; and observed means for the GxE genotype set (sorted by yield).

	Yield (kg/ha)	Oil (%)	Total protein (%)	Soluble protein (%)	7S (%)	11S (%)	11S/7S ratio	Sol. carb. (%)	Sucrose (%)	RFOs (%)
G	S	S	S	NS	S	NS	S	S	S	S
E	S	S	S	S	NS	S	S	NS	S	S
GxE	NS	S	S	NS	NS	NS	NS	NS	S	NS
5958 RSF IPRO	5429 ^{a}	21.1 ^{bc}	40.8 ^{abcd}	22.2	6.1 ^{cd}	16.0	2.5 ^{abc}	9.1 ^{d}	5.1 ^{fg}	3.9 ^{cd}
DM 6.8i	4738 ^{b}	20.4 ^{fg}	41.7 ^{a}	23.2	7.0 ^{ab}	16.2	2.2 ^{cd}	9.7 ^{ab}	5.6 ^{bcd}	4.0 ^{cd}
62R63 RSF	4355 ^{bc}	20.5 ^{ef}	39.8 ^{de}	22.9	6.7 ^{abc}	16.2	2.3 ^{bcd}	9.8 ^{ab}	5.7 ^{ab}	4.0 ^{cd}
SJ13625	4179 ^{bcd}	20.3 ^{fg}	39.8 ^{de}	23.1	6.6 ^{abc}	16.5	2.4 ^{abcd}	9.2 ^{d}	4.6 ^{i}	4.5 ^{a}
NA 5509 RG	4125 ^{cd}	20.7 ^{de}	40.7 ^{bcd}	24.4	6.9 ^{ab}	17.4	2.2 ^{cd}	9.7 ^{ab}	5.7 ^{abc}	3.9 ^{cd}
DM 6.2i	4050 ^{cde}	20.8 ^{bcd}	40.7 ^{bcd}	23.4	6.9 ^{ab}	16.5	2.3 ^{bcd}	9.9 ^{a}	5.9 ^{a}	3.9 ^{d}
GE 590 CI	4031 ^{cde}	21.1 ^{b}	40.6 ^{bcd}	22.5	6.3 ^{cd}	16.2	2.5 ^{abc}	9.7 ^{ab}	5.4 ^{cde}	4.2 ^{cd}
NA 5909 RG	4010 ^{cde}	21.4 ^{a}	41.5 ^{ab}	23.3	6.9 ^{ab}	16.3	2.3 ^{bcd}	9.6 ^{bc}	5.3 ^{def}	4.2 ^{bc}
SJ13624	3976 ^{cde}	20.3 ^{fg}	39.6 ^{e}	23.0	6.8 ^{abc}	16.3	2.2 ^{d}	9.3 ^{cd}	4.8 ^{hi}	4.5 ^{ab}
SJ13623	3889 ^{cde}	20.2 ^{g}	41.3 ^{abc}	22.9	6.7 ^{abc}	16.2	2.3 ^{bcd}	9.2 ^{d}	4.7 ^{hi}	4.5 ^{ab}
DM 50i17 IPRO	3650 ^{def}	20.8 ^{cde}	40.4 ^{cde}	22.4	5.9 ^{d}	16.5	2.6 ^{a}	9.1 ^{d}	4.9 ^{gh}	4.1 ^{cd}
SJ13621	3510 ^{ef}	20.4 ^{fg}	39.4 ^{e}	23.1	6.4 ^{bcd}	16.7	2.5 ^{ab}	9.3 ^{cd}	5.8 ^{ab}	3.4 ^{e}
NA 5009 RG	3273 ^{f}	20.2 ^{g}	41.1 ^{abc}	23.9	7.0 ^{a}	16.8	2.4 ^{abc}	9.4 ^{cd}	5.2 ^{ef}	4.1 ^{cd}

S: Significant ($p < 0.05$); NS: No significant ($p > 0.05$); G: Genotype; E: Environment; GxE: Genotype x Environment interaction; RFOs: Raffinose family oligosaccharides; Sol. carb.: Soluble carbohydrates. Means in the same column with different letters are significantly different ($p < 0.05$) for genotype effect.

Table 4. Components of variance of the parameters with significant GxE interaction.

	Oil (%)	Total protein (%)	Sucrose (%)
Environment	87	90	50
Genotype	10	7	38
GxE	2	3	4
Rep	1	0	8

GxE: Genotype x Environment interaction; Rep: replicate.

The genotypes with the highest yield were 5958 RSF IPRO (5429 kg/ha) and DM 6.8i (4738 kg/ha), while NA 5009 RG was the genotype with the lowest yield (3273 kg/ha) (Table 3). Although the GxE interaction effect was detected on oil content, the same genotypes maintained in the top ranking across environments (NA 5909 RG, 5958 RSF IPRO, and GE 590; Figure 2A). For total protein, the GxE interaction was also statistically significant, but several genotypes were located at the top in the different environments, presenting values over 40%: NA 5009 RG, DM 50i17 IPRO, and SJ13621 (Figure 2B). The genotype effect was significant

for the 7S protein fraction. The concentration of 7S ranged from 5.9% to 7.0% (Table 3). It has been reported that soy protein is deficient in sulfur amino acids (methionine and cysteine); the 11S fraction contains more S-amino acids than 7S fraction (Krishnan, 2005). The 11S/7S ratio had significant differences between the genotypes. Particularly, those that had highest 11S/7S ratio were DM 50i17 IPRO (2.6), SJ13621 (2.5), 5958 RSF IPRO (2.5) and GE 590 CI (2.5). Total soluble carbohydrates and their components are very relevant factors to analyze. The genotype DM 6.2i had a high content of soluble carbohydrates (9.9%; Table 3) with a low percentage of RFOs (3.9%). The genotype with the lowest content of oligosaccharides was SJ13621 (3.4%), while SJ13623, SJ13624 and SJ13625 had highest RFOs content (average value 4.5%). Although the GxE interaction effect was detected on sucrose content, two genotypes (DM 6.2i and SJ13621) consistently showed the highest levels of sucrose; they presented values over 5.5% in all environments (Figure 2C). Thus, these genotypes had the most desirable combination (high sucrose and low RFOs) for soy food (Mozzoni et al., 2013).

Figure 3 shows principal component analysis among the different parameters for the 13 genotypes; parameters with significant GxE interaction or without significant genotype effect have a vector to each environment. The first and second principal components (PC 1 and PC 2) represented 52.0% of the total variation among genotypes for the 10 traits considered. The yield and oil content vectors were correlated positively, and negatively correlated to total protein, for each of the environments. Although the GxE interaction effect was significant on oil, sucrose, and total protein content, the vectors were correlated. Contrary, the vectors for the 11S fraction and soluble protein had different magnitudes and these were in different sectors. Even PC 3 contribution is relatively low (15.7%), when it is considered both 11S fraction and soluble protein groups of vectors are closer (figure not shown). The sucrose and RFOs vectors showed a negative correlation. Confirming the high oil mean of NA 5909 RG, 5958 RSF IPRO, and GE 590 CI in the biplot these genotypes were near to oil vectors. Also, 5958 RSF IPRO had the highest yield and high values of 11S/7S ratio. The genotypes SJ13623, SJ13624, and SJ13625 had high RFOs content and they were close to the RFOs vector. The genotype SJ13621 had high values of total protein and soluble protein (mostly in the environment LED20), and it was represented on the same side of these vectors. The genotype NA 5009 RG, was at the bottom of the biplot because it had high total protein content, mostly in LEN19. On the left, was the point of DM 50i17 IPRO indicating a high 11S/7S ratio, total protein, and RFOs content. Conversely, the points of DM 6.2i, DM 6.8i and 62R63 RSF were on the right of the biplot confirming the high soluble carbohydrate content with high values of sucrose.

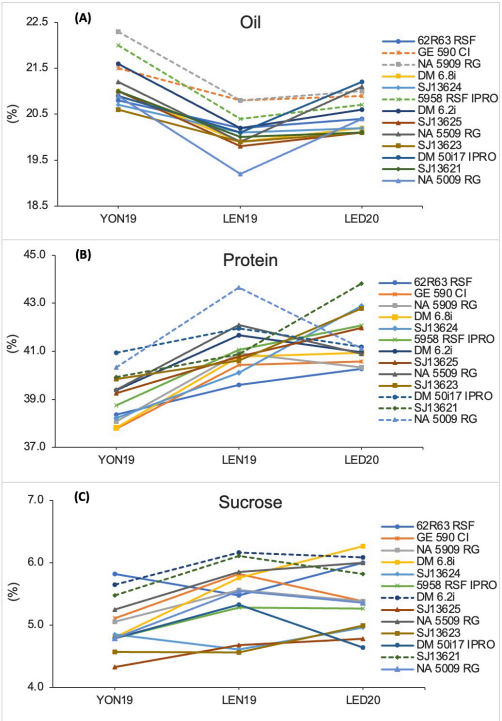

Figure 2. Mean of the 13 genotypes to each environment of the parameters with interaction GxE, including (A) oil, (B) total protein and (C) sucrose. Dotted lines correspond to genotypes with high parameter values in all three environments.

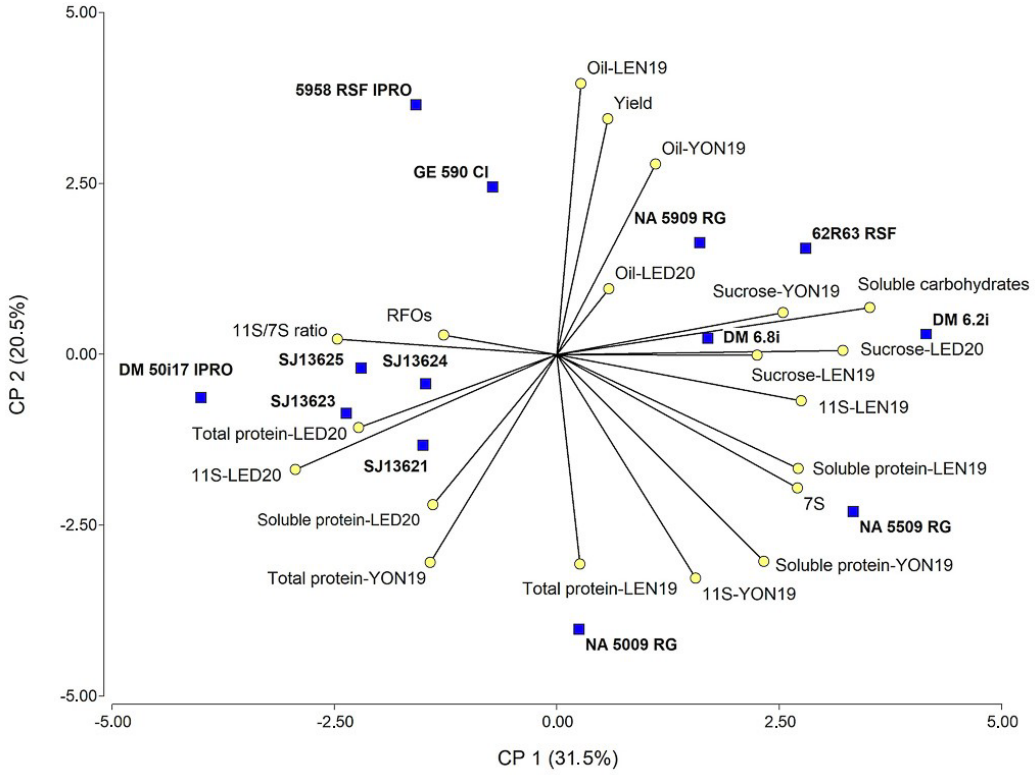

Figure 3. Principal component analysis of the 13 genotypes.

3.2 Genotype variability

3.2.1 Descriptive statistics

In order to search genotype diversity, two sets of samples were also studied. The first one included 10 GMO as checks and 12 non-GMO lines, totalizing 22 genotypes (CONV), and other with 36 GMO lines genotypes (GM). In this case, they had been grown in a single environment each.

Table 5 shows the averages and ranges of each set (CONV and GM). Both sets yield means were similar (around 4500 kg/ha). The range was wider in CONV than in GM. The oil content mean was the same in both sets (20.8%) and the range was higher in GM. Contrary, the total protein content was higher in CONV than GM but had less variability. Means of oil and total protein content were similar to the GxE set. The content of soluble protein, the fractions 7S and 11S and 11S/7S ratio content were higher in GM than in CONV. Values of soluble protein and their fractions were similar to the range observed in the GxE study previously described. The results of carbohydrates, were similar between sets and the range was wider in GM than in CONV. Additionally, these results were comparable with the detected in the GxE study.

3.2.2 Variability of genotypes

Tables 6 and 7 exhibit the genotype effect on the parameters in both sets. Yield, oil and total protein had significant effect both in CONV and GM, as already observed in GxE study. Conversely, the genotype effect was nonsignificant on soluble protein and 11S in both sets and in GxE study. The results of soluble carbohydrates, sucrose and RFOs showed that they had significant genotype effect in CONV and GM. The protein 7S and the 11S/7S ratio had significant genotype effect only in CONV.

Table 5. Mean and range for different parameters of the groups CONV and GM.

Parameters	CONV			GM		
	Mean	Min	Max	Mean	Min	Max
Yield (kg/ha)	4524	3581	5374	4575	3744	5316
Oil (%)	20.8	20.1	21.9	20.8	19.7	22.1
Total protein (%)	41.4	39.5	43.8	40.3	37.2	42.9
Soluble protein (%)	22.9	21.7	23.9	23.6	21.0	25.4
7S (%)	6.5	5.2	7.2	6.8	5.8	8.4
11S (%)	16.3	15.0	17.7	16.8	14.4	18.1
11S/7S ratio	2.4	1.9	3.2	2.3	1.6	3.2
Soluble carbohydrates (%)	9.4	8.4	10.0	9.9	9.2	10.4
Sucrose (%)	5.4	4.6	6.3	5.9	4.8	6.9
RFOs (%)	3.9	3.2	4.3	4.0	3.3	4.7

Min: Minimum; Max: Maximum; RFOs: Raffinose family oligosaccharides; CONV: conventional; GM: genetically modified.

Table 6. Genotype effect on the different parameters and differentiation of means in 22 genotypes (CONV), including conventional genotypes and GMO checks.

	Yield (kg/ha)	Oil (%)	Total protein (%)	Soluble protein (%)	7S (%)	11S (%)	11S/7S ratio	Sol. carb. (%)	Sucrose (%)	RFOs (%)
G	**S**	**S**	**S**	**NS**	**S**	**NS**	**S**	**S**	**S**	**S**
DM 6.8i*	5374 ^{a}	20.2 ^{ghi}	41.0 ^{fgh}	23.0	7.1 ^{abc}	16.0	2.2 ^{cde}	9.9 ^{a}	6.3 ^{a}	3.5 ^{def}
SJ14494	5338 ^{ab}	20.4 ^{fghi}	41.8 ^{cdefg}	23.0	7.2 ^{a}	15.8	2.0 ^{de}	9.3 ^{bcde}	5.4 ^{efg}	3.7 ^{abcdef}
5958 RSF IPRO*	5204 ^{abc}	20.7 ^{def}	42.1 ^{bcde}	22.3	6.2 ^{de}	16.1	2.5 ^{bcd}	9.1 ^{def}	5.3 ^{ghi}	3.6 ^{cdef}
SJ14514	5104 ^{abcd}	21.2 ^{bc}	39.5 ^{j}	22.1	7.1 ^{ab}	14.9	1.9 ^{e}	9.1 ^{de}	5.4 ^{fgh}	3.7 ^{cdef}
SJ13623	5065 ^{abcd}	20.1 ^{i}	42.8 ^{abc}	22.9	6.3 ^{cd}	16.6	2.5 ^{bcd}	9.2 ^{de}	4.9 ^{jk}	4.1 ^{abcd}
62R63 RSF*	4785 ^{abcde}	20.4 ^{fghi}	40.3 ^{hij}	21.9	6.6 ^{abcd}	15.3	2.2 ^{cde}	9.8 ^{abc}	6.0 ^{bc}	3.8 ^{abcde}
SJ13625	4755 ^{bcde}	20.1 ^{i}	41.9 ^{bcdef}	22.8	6.1 ^{de}	16.7	2.6 ^{bc}	9.2 ^{bcde}	4.8 ^{kl}	4.3 ^{a}
NA 5509 RG*	4725 ^{cde}	21.1 ^{bcd}	40.9 ^{ghi}	23.6	7.2 ^{a}	16.4	2.1 ^{de}	9.9 ^{ab}	6.0 ^{bc}	3.8 ^{abcde}
SJ14511	4690 ^{cdef}	20.7 ^{def}	41.8 ^{cdefg}	23.9	6.8 ^{abcd}	17.1	2.4 ^{bcde}	9.5 ^{abcd}	5.6 ^{de}	3.9 ^{abcde}
SJ13624	4680 ^{cdef}	20.2 ^{hi}	42.9 ^{ab}	23.4	7.1 ^{ab}	16.3	2.0 ^{de}	9.0 ^{def}	4.9 ^{jk}	4.0 ^{abcd}
DM 6.2i*	4561 ^{defg}	20.6 ^{efg}	41.0 ^{fgh}	21.7	6.4 ^{abcd}	15.3	2.3 ^{cde}	10.0 ^{a}	6.1 ^{ab}	3.8 ^{abcde}
GE 590 CI*	4525 ^{defgh}	20.9 ^{cde}	40.6 ^{hi}	23.4	6.6 ^{abcd}	16.8	2.4 ^{bcde}	9.6 ^{abcd}	5.4 ^{fgh}	4.1 ^{abc}
SJ14498	4523 ^{defgh}	21.3 ^{bc}	41.7 ^{defg}	22.8	5.5 ^{ef}	17.3	2.8 ^{ab}	9.1 ^{def}	5.1 ^{hij}	3.9 ^{abcde}
SJ14497	4441 ^{efghi}	21.9 ^{a}	41.9 ^{bcdef}	23.3	6.4 ^{abcd}	16.8	2.3 ^{bcde}	8.4 ^{f}	5.1 ^{ij}	3.3 ^{ef}
DM 50i17 IRPO*	4112 ^{fghij}	21.3 ^{bc}	41.2 ^{defgh}	22.9	5.2 ^{f}	17.7	3.2 ^{a}	8.8 ^{ef}	4.6 ^{l}	4.1 ^{abcd}
5351 RSF*	4091 ^{fghij}	21.4 ^{b}	39.9 ^{ij}	22.8	6.3 ^{cde}	16.5	2.3 ^{bcde}	8.7 ^{ef}	4.6 ^{l}	4.0 ^{abcd}
NA 5909 RG*	4076 ^{ghij}	21.0 ^{bcd}	40.3 ^{hij}	22.8	6.6 ^{abcd}	16.1	2.2 ^{cde}	9.6 ^{abcd}	5.4 ^{fgh}	4.2 ^{abc}
SJ14513	4051 ^{ghij}	20.5 ^{fgh}	42.2 ^{bcd}	23.1	6.6 ^{abcd}	16.4	2.2 ^{cde}	9.2 ^{cde}	5.1 ^{hij}	4.0 ^{abcd}
SJ14509	4046 ^{ghij}	21.2 ^{bc}	41.2 ^{defgh}	21.8	6.4 ^{bcd}	15.4	2.3 ^{cde}	9.9 ^{a}	5.6 ^{def}	4.3 ^{a}
SJ14508	3955 ^{hij}	21.2 ^{bc}	41.0 ^{fgh}	22.8	7.1 ^{ab}	15.6	2.0 ^{de}	9.9 ^{ab}	5.7 ^{de}	4.1 ^{abc}
NA 5009 RG*	3854 ^{ij}	20.4 ^{fghi}	41.1 ^{efgh}	23.2	6.5 ^{abcd}	16.7	2.4 ^{bcde}	9.6 ^{abcd}	5.3 ^{ghi}	4.2 ^{ab}
SJ13621	3581 ^{j}	20.1 ^{i}	43.8 ^{a}	23.9	6.3 ^{cd}	17.6	2.7 ^{bc}	9.1 ^{def}	5.8 ^{cd}	3.2 ^{f}

S: Significant ($p < 0.05$); NS: No significant ($p > 0.05$); G: Genotype; E: Environment; RFOs: Raffinose family oligosaccharides; Sol. carb.: Soluble carbohydrates. * indicates GMO genotypes. Means in the same column with different letters are significantly different ($p < 0.05$).

Table 6 also presents the differentiation of means in the parameters that had significant effect, with genotypes sorted by yield. The genotypes with the highest yield were DM 6.8i and SJ14494 (5374 kg/ha and 5338 kg/ha, respectively); while SJ13621 had the lowest yield (3581 kg/ha). The oil content was higher in the genotype SJ14497 with a mean of 21.9%. In total protein, high variability was observed among genotypes, being INIA lines the ones with highest values: SJ13621 (43.8%), SJ13624 (42.9%) and SJ13623 (42.8%). The 11S/7S ratio was significantly higher in the genotypes DM 50i17 IPRO (3.2), SJ14498 (2.8) and SJ13621 (2.7). The highest value of soluble carbohydrates was 10% in the genotype DM 6.2i, mostly explained by high content of sucrose (6.1%). The genotype DM 6.8i had the highest sucrose with a mean of 6.3%; it also had high content of this carbohydrate: 62R63 RSF (6.0%), NA 5509 RG (6.0%) and SJ13621 (5.8%). Of those genotypes with high sucrose, DM 6.8i and SJ13621 had low content of RFOs with means of 3.5% and 3.2%, respectively. Also, SJ14497 showed low RFOs (3.3%).

Table 7 presents the means difference of the genotypes in GM, which are sorted according to yield. In yield, although it had a wide range, it was difficult to differentiate the genotypes from each other, but a significant variability was observed for quality parameters (oil, total protein and soluble carbohydrates). The mentioned negative correlation between oil and protein content was evidenced when the genotype with highest oil content had the lowest content of total protein (SJ12394) and genotypes high total protein had low percentage of oil, like SJ14502 and SJ13618. SJ13619 had the highest value of sucrose content (6.9%) and the lowest of RFOs (3.3%); furthermore, this genotype had high yield (4833 kg/ha) and total protein (41.3%). Similar situation was observed in the genotype DM 6.8i with 6.6% of sucrose, 3.3% of RFOs and 4991 kg/ha of yield.

Table 7. Genotype effect on the different parameters and means of 36 GMO genotypes (GM).

	Yield (kg/ha)		Oil (%)		Total protein (%)		Soluble protein (%)	7S (%)	11S (%)	11S/7S ratio	Sol. carb. (%)		Sucrose (%)		RFOs (%)	
G	**S**		**S**		**S**		**NS**	**NS**	**NS**	**NS**	**S**		**S**		**S**	
5958 RSF IPRO	5316	a	20.6	hijk	40.9	cdef	21.0	7.6	16.3	1.6	9.7	abcdef	5.7	fghij	4.0	defghijk
FS 59	5236	ab	20.4	ijk	40.4	cdefghij	24.6	7.3	17.3	2.4	10.0	abc	5.5	ghijk	4.4	abcdef
62R63 RSF	5125	abc	20.3	ijk	40.5	cdefghi	23.1	6.0	17.0	3.0	10.1	abc	6.2	bcdef	3.8	hijkl
NA 5909 RG	5067	abcd	20.8	fghi	40.1	efghij	23.7	6.7	17.0	2.5	10.0	abc	6.0	bcdefghi	3.9	efghijkl
GENESIS 5501	4994	abcde	21.4	bcde	39.5	ijk	24.1	7.5	16.6	1.9	10.4	a	6.4	abc	3.9	ghijkl
SJ14504	4993	abcde	20.4	ijk	42.3	ab	23.4	7.9	15.6	1.8	9.2	ef	5.1	klm	4.1	bcdefghij
DM 6.8i	4991	abcde	20.3	ijk	40.5	cdefghij	21.0	6.6	14.4	2.3	9.9	abcde	6.6	ab	3.3	m
SJ14505	4916	abcdef	20.7	ghijk	40.8	cdef	23.3	7.0	15.6	2.1	10.0	abc	6.0	bcdefghi	3.9	fghijkl
SJ13619	4833	abcdefg	20.6	hijk	41.3	cd	23.9	6.3	17.6	2.7	10.2	ab	6.9	a	3.3	m
SJ14502	4803	abcdfgh	20.3	ijk	42.9	a	25.4	8.4	17.1	1.9	9.2	f	5.4	jkl	3.8	ijkl
SJ13614	4792	abcdefgh	20.2	jk	41.3	bc	23.8	6.6	17.2	2.6	9.8	abcdef	5.6	ghij	4.1	cdefghij
GENESIS 5901	4777	abcdefgh	21.4	cde	38.4	l	24.2	6.3	17.9	2.8	10.4	a	6.4	abc	3.9	fghijkl
GENESIS 6201	4758	abcdefghi	20.4	ijk	39.9	fghijk	23.0	7.3	15.9	1.9	9.9	abcdef	5.5	hijk	4.3	abcdefg
SJ13064	4720	abcdefghij	20.9	efghi	39.2	jkl	23.2	6.4	16.8	2.4	10.2	abc	5.9	bcdefghij	4.1	abcdefghij
SJ13615	4698	bcdfghij	20.5	ijk	41.4	bc	23.3	6.5	16.7	2.4	10.3	ab	5.8	efghij	4.4	abcd
SJ14490	4638	bcdefghij	21.7	abc	39.6	ghijk	23.7	6.8	16.9	2.4	10.2	abc	6.0	bcdefgh	4.0	cdefghij
NA 5509 RG	4623	bcdefghij	20.8	fghi	39.6	hijk	23.8	5.8	17.9	3.2	10.2	abc	6.1	bcdefg	4.0	defghijk
GENESIS 5602	4595	cdefghij	21.3	cde	39.5	ijk	24.0	7.0	17.0	2.2	9.8	abcdef	5.4	jkl	4.4	abcde
SJ13327	4529	cdefghijk	21.0	efgh	40.0	efghij	22.5	6.8	15.7	1.9	10.3	ab	6.4	abc	3.8	hijkl
GENESIS 6602	4524	cdefghijk	20.4	ijk	40.2	defghij	23.4	7.4	15.9	1.8	9.9	abcd	5.9	cdefghij	4.0	defghijk
SJ12210	4519	cdfghijk	19.7	l	41.3	bc	24.3	6.2	18.0	2.7	9.9	abcd	6.3	bcde	3.6	jklm
GE 590 CI	4517	cdefghijkl	21.2	def	39.2	jkl	24.3	6.4	17.9	2.8	9.7	bcdef	5.4	jkl	4.2	abcdefghi

Table 7. Continued...

	Yield (kg/ha)		Oil (%)		Total protein (%)		Soluble protein (%)	7S (%)	11S (%)	11S/7S ratio	Sol. carb. (%)		Sucrose (%)		RFOs (%)	
SJ13618	4479	defghijkl	20.2	kl	42.5	a	25.2	7.3	17. 9	2.5	9.8	abcdef	5.8	efghij	4.0	defghijkl
SJ14507	4445	efghijkl	20.8	fghi	40.6	cdefgh	23.2	6.4	16.7	2.7	10.1	abc	6.0	bcdefghi	4.0	defghijk
DM 50i17 IPRO	4412	efghijkl	21.7	abc	40.1	efghij	23.7	6.9	16.8	2.2	9.3	def	4.9	lm	4.4	abcdef
SJ13626	4383	efghijkl	21.0	efgh	39.7	ghijk	22.8	6.2	16.6	1.9	9.9	abcdef	6.2	bcdef	3.6	klm
GENESIS 5601	4368	fghijkl	21.2	defg	40.5	cdefghij	22.4	6.0	16.4	2.7	9.5	cdef	5.7	efghij	3.7	jklm
DM 6.2i	4300	ghijklm	20.7	fghij	40.1	efghij	23.4	6.7	16.7	2.5	10.3	ab	6.4	abcd	3.8	hijkl
GENESIS 6301	4252	ghijklm	21.5	bcde	38.9	kl	23.0	6.5	16.6	2.3	10.1	abc	5.8	defghij	4.2	abcdefghi
SJ12394	4211	hijklm	22.1	a	37.2	m	24.7	7.6	17.0	1.9	9.9	abcd	5.6	ghij	4.2	abcdefgh
SJ13616	4155	ijklm	19.7	1	41.1	cde	23.3	6.1	17.2	2.9	10.0	abc	6.1	bcdefg	3.9	ghijkl
SJ12395	4115	jklm	21.7	abcd	39.2	jkl	23.7	6.6	17.1	2.4	10.4	a	5.8	defghij	4.5	abc
SJ13371	4025	jklm	20.5	hijk	40.7	cdefg	23.9	7.3	16.6	1.9	9.5	cdef	5.6	fghij	3.8	hijkl
SJ13425	3957	klm	20.7	fghij	40.7	cdefg	22.3	6.3	16.1	2.4	9.4	cdef	5.8	defghij	3.5	lm
5351 RSF	3903	lm	21.9	ab	38.2	lm	22.5	6.9	15.6	1.9	9.5	cdef	4.8	m	4.7	a
NA 5009 RG	3744	m	20.8	fghi	40.4	cdefghij	24.2	6.4	17.8	2.6	10.1	abc	5.5	ijk	4.5	ab

S: Significant ($p < 0.05$); NS: No significant ($p > 0.05$); G: Genotype; E: Environment; RFOs: Raffinose family oligosaccharides; Sol. carb.: Soluble carbohydrates. Means in the same column with different letters are significantly different ($p < 0.05$).

3.3 Correlations

Table 8 shows the correlations between traits that have been reported as significant in the literature or found to be significant in our analyses. In the three sets, yield was not significantly correlated with quality parameters (data not shown). Which indicates that within the germplasm set considered and Uruguay's cropping environment it is possible to select genotypes with relatively good quality composition independently from their yield values. Eventually, it may be possible to select good quality breeding lines without affecting their yield performance. However, other authors had found significant correlation of yield with oil and total protein content (Wilcox & Shibles, 2001; Chung et al., 2003).

Protein and oil content were negatively correlated in two of the three environments in GxE set, LED20 r = -0.60 and LEN19 r = -0.56, respectively, and also in CONV and GM (r = -0.48 and r = -0.78, respectively). Qin et al. (2014) also reported that content of protein was negatively correlated with content of oil. Additionally, this correlation is frequent in Uruguay (Marina Castro, personal communication, 2022), which is undesirable because the goal is to obtain high values on both parameters.

In GxE set, total protein content showed a negative and statistically significant correlation with soluble carbohydrates, with coefficient of correlations of -0.67 in LED20 and -0.56 in YON19, in agreement with Wilcox & Shibles (2001). This correlation also was significant in the set GM (r = -0.34).

Harada et al. (1983) and Fehr et al. (2003) did not found significant correlations between total protein and the fractions. It was the same in our study, with the exception of 7S fraction in LEN19 and 11S in CONV. In our study, total and soluble protein were not correlated in the three environments of GxE nor in GM set; although this relationship was found significant in the studies of Zhang et al. (2017) and Pesic et al. (2005). The fractions 11S and 7S were positively and significantly correlated with soluble protein both in GxE study and the sets CONV and GM.

The correlation between sucrose and RFOs was negative and significant across the three environments: coefficient of correlation was -0.65, -0.91 and -0.71 for LED20, LEN19 and YON19, respectively. In GM this correlation was also significant (r = -0.70). This negative relationship was also observed by Mozzoni et al. (2013) and Qin et al. (2014). This is valuable for developing new soybean cultivars with improvements in flavor (high in sucrose) and digestibility (low in RFOs).

Table 8. Correlation coefficients between investigated parameters for the three sets.

	GxE			CONV	GM
	LED20	LEN19	YON19		
Total protein/oil	-0.60*	-0.56*	-0.36	-0.48*	-0.78**
Total protein/sol. carb.	-0.67*	-0.03	-0.56*	-0.32	-0.34*
Total protein/11S	0.47	0.30	0.16	0.50*	0.01
Total protein/7S	-0.12	0.60*	-0.23	-0.16	0.22
Total protein/ 11S/7S ratio	0.26	0.44	0.22	0.31	0.05
Total protein/soluble protein	0.44	0.43	0.02	0.48*	0.16
Soluble protein/11S	0.72**	0.96**	0.82**	0.72**	0.75**
Soluble protein/7S	0.23	0.71**	0.75**	0.15	0.42*
Sucrose/RFOs	-0.65*	-0.91**	-0.71**	-0.36	-0.70**

* $p < 0.05$, ** $p < 0.01$, Sol. carb.: Soluble carbohydrates. GxE: Genotype x Environment interaction; CONV: conventional; GM: genetically modified.

3.4 General discussion

This study verified that genotype, environment and genotype x environment interaction are all important as determinants of the grain quality parameters for food use in soybean. The traits in which it was possible to detect differences among genotypes were yield, soluble carbohydrates, RFOs, 7S and 11S/7S ratio, that allow to select the genotypes with the best values for these characteristics in all environments. In GxE set, the parameters that only had environment influence were soluble protein and 11S, indicating it may be necessary to elucidate which environmental factors influence them, and if it is possible to change them through agronomic management practices to ensure the achievement of the quality values desired. Those parameters in the CONV and GM sets did not have a significant genotype effect either. The parameters that showed genotype x environment interaction were oil, total protein and sucrose. For them, it is necessary to select the best combination of environment/genotype to obtain the best values.

Regarding the diversity, the set studied presented similar ranges to the studies of Qin et al. (2014), Yu et al. (2016), Stanojevic et al. (2011) and Kumar et al. (2006) in commercial cultivars, and lower than the ranges observed in the characterization of collections of diverse germplasm (Hou et al., 2009; Hwang et al., 2014). This indicates that even local variability is similar to the observed in most studies, it is possible to expand diversity at the soybean breeding program to achieve better quality values.

In both GxE study and CONV (Table 3 and Table 6), conventional genotypes alternated their place in the ranking the GMO ranks ones. Therefore, since consumers prefer non-GMO cultivars, it suggests that breeders do not need to include GMOs in their program in order to find more variability.

Some genotypes were identified in the top or the bottom in all set, indicating that a differentiation is possible. In yield, the genotypes that had high values were 5958 RSF IPRO, DM 6.8i and 62R63 RSF in the GxE study as well as in CONV and GM (Table 3, Table 6 and Table 7). Conversely, NA 5009 RG, DM 50i17 and 5351 RSF were at the bottom of the tables. The genotype 5351 RSF had oil content over 21% in the analyzed groups. In the case of total protein, the genotype SJ13621 was in the top both in the GxE study and in CONV; and DM 50i17 IPRO and NA 5009 RG had high total protein in the three sets. In opposition, GE 590 CI and 5351 RSF had low values of total protein in all set. Observing the profile of carbohydrates, 5 genotypes (DM 6.2i, SJ13621, DM 6.8i, NA 5509 RG and 62R63 RSF) maintained high content of sucrose in all sets. It is important to emphasize that the same genotypes had low content of RFOs. The conventional genotypes SJ13623 and SJ13625 were in the top in GxE study as well as in CONV with high values of RFOs; NA 5009 RG and DM 50i17 IRPO had high oligosaccharides content both in CONV and GM. In the case of the 11S/7S ratio, the genotypes that maintained high values in both sets were DM 50i17 IPRO, 5958 RSF IPRO and SJ13621; while NA 5509 RG, SJ13624 and DM 6.8i had low values (Table 3 and Table 6).

Integrating the results, it was observed that some genotypes had interesting combinations of different parameters in all the groups. For example, DM 6.8i had high yield in all sets and had good profile of carbohydrates (high sucrose-low RFOs) in CONV and GM; DM 6.2i comparatively high sucrose and low RFOs content, with the total protein content never below of 40%; SJ13621 had high content of total protein, high sucrose, low values of RFOs and high of the ratio 11S/7S.

4 Conclusions

Although the observed grain quality values were strongly influenced by environment both in their absolute value and through interaction with genotype, it was possible to identify genotypes with a better combination of quality parameters for human consumption and processing. Among these parameters, the sucrose content can be considered as the most relevant to select genotypes for human consumption. It should be selected in breeding programs for soybean improvements, but also protein related parameters should be selected. Furthermore, since non-GMO materials are preferred by the consumers, it is very relevant to state that no apparent difference was found between studied GMO and conventional genotypes; moreover, a non-GMO line had one of the best combinations of desired parameters.

Further than understanding the effect of genotype, environment and their interaction on studied parameters, more variability was explored with two specific set of samples. Comparing the three analyzed sets, the variability among genotypes was higher in the sets with more genotypes (CONV and GM), suggesting that a higher variability may be available. In addition, some genotypes were stable at different environments, suggesting that it is possible to select by stability in the studied parameters. However, further study including more environments would be recommendable to confirm these observations.

References

Assefa, Y., Bajjalieh, N., Archontoulis, S., Casteel, S., Davidson, D., Kovács, P., Naeve, S., & Ciampitti, I. A. (2018). Spatial characterization of soybean yield and quality (amino acids, oil, and protein) for United States. *Scientific Reports*, *8*(1), 14653. PMid:30279447. http://dx.doi.org/10.1038/s41598-018-32895-0

Balestra, F., & Petracci, M. (2019). Technofunctional ingredients for meat products: Current challenges. In C. M. Galanakis (Ed.), *Sustainable meat production and processing* (pp. 45-68). Cesena: Academic Press. http://dx.doi.org/10.1016/B978-0-12-814874-7.00003-1.

Cai, T., & Chang, K. C. (1999). Processing effect on soybean storage proteins and their relationship with tofu quality. *Journal of Agricultural and Food Chemistry*, *47*(2), 720-727. PMid:10563959. http://dx.doi.org/10.1021/jf980571z

Choung, M. G. (2005). Determination of soluble carbohydrates in soybean seeds. *Hangug Jagmul Haghoeji*, *50*(5), 319-324. Retrieved in 2023, August 19, from http://www.koreascience.or.kr/journal/JMHHBK.page

Chung, J., Babka, H. L., Graef, G. L., Staswick, P. E., Lee, D. J., Cregan, P. B., Shoemaker, R. C., & Specht, J. E. (2003). The seed protein, oil, and yield QTL on soybean linkage group I. *Crop Science*, *43*(3), 1053-1067. http://dx.doi.org/10.2135/cropsci2003.1053

Cuitiño, M. J., Manasliski, S., Morales, X., & Cardozo, V. (2020). Resultados. In: *Resultados experimentales de la evaluación nacional de cultivares de soja* (pp. 15-62). La Estanzuela, Uruguay: INIA. Retrieved in 2023, August 19, from http://www.inia.org.uy/convenio_inase_inia/Evaluacion_CV/Ano2020/PubSojaPeriodo2020.pdf

Cuitiño, M. J., Manasliski, S., Stewart, S., Morales, X., & Cardozo, V. (2019). Resultados. In: *Resultados experimentales de la evaluación nacional de cultivares de soja* (pp. 15-53). La Estanzuela, Uruguay: INIA. Retrieved in 2023, August 19, from http://www.inia.org.uy/convenio_inase_inia/Evaluacion_CV/Ano2019/PubSojaPeriodo2019.pdf

Di Rienzo, J. A., Casanoves, F., Balzarini, M. G., Gonzalez, L., Tablada, M., & Robledo, C. W. (2020). *InfoStat versión 2018. Centro de Transferencia InfoStat*. Argentina: FCA, Universidad Nacional de Córdoba.

Esteves, E. A., Martino, H. S. D., Oliveira, F. C. E., Bressan, J., & Costa, N. M. B. (2010). Chemical composition of a soybean cultivar lacking lipoxygenases (LOX2 and LOX3). *Food Chemistry*, *122*(1), 238-242. http://dx.doi.org/10.1016/j.foodchem.2010.02.069

Fehr, W. R., Hoeck, J. A., Johnson, S. L., Murphy, P. A., Nott, J. D., Padilla, G. I., & Welke, G. A. (2003). Genotype and environment influence on protein components of soybean. *Crop Science*, *43*(2), 511-514. http://dx.doi.org/10.2135/cropsci2003.0511

Fontes, E. P. B., Moreira, M. A., Davies, C. S., & Nielsen, N. C. (1984). Urea-elicited changes in relative electrophoretic mobility of certain glycinin and β-conglycinin subunits. *Plant Physiology*, *76*(3), 840-842. PMid:16663936. http://dx.doi.org/10.1104/pp.76.3.840

Fukushima, D. (1981). Soy proteins for foods centering around soy sauce and tofu. *Journal of the American Oil Chemists' Society*, *58*(3), 346-354. http://dx.doi.org/10.1007/BF02582376

Garance, M. P., & Arbeletche, P. (2020). Descripción y análisis de la cadena de valor de la soja en Uruguay. *Ciencias Agronómicas*, *007*(35), e007. http://dx.doi.org/10.35305/agro35.294

Garner, W. W., & Allard, H. A. (1930). Photoperiodic response of soybeans in relation to temperature and other environmental factors. *Journal of Agricultural Research*, *41*, 719-735.

Gu, D., Andreev, K., & Dupre, M. E. (2021). Major trends in population growth around the world. *China CDC Weekly*, *3*(28), 604. http://dx.doi.org/10.46234/ccdcw2021.160

Harada, K., Toyokawa, Y., & Kitamura, K. (1983). Genetic analysis of the most acidic 11S globulin subunit and related characters in soybean seeds. *Japanese Journal of Breeding*, *33*(1), 23-30. http://dx.doi.org/10.1270/jsbbs1951.33.23

Hartmann, C., & Siegrist, M. (2017). Consumer perception and behaviour regarding sustainable protein consumption: A systematic review. *Trends in Food Science & Technology*, *61*, 11-25. http://dx.doi.org/10.1016/j.tifs.2016.12.006

Hou, A., Chen, P., Alloatti, J., Li, D., Mozzoni, L., Zhang, B., & Shi, A. (2009). Genetic variability of seed sugar content in worldwide soybean germplasm collections. *Crop Science*, *49*(3), 903-912. http://dx.doi.org/10.2135/cropsci2008.05.0256

Hwang, E. Y., Song, Q., Jia, G., Specht, J. E., Hyten, D. L., Costa, J., & Cregan, P. B. (2014). A genome-wide association study of seed protein and oil content in soybean. *BMC Genomics*, *15*(1), 1-12. PMid:24382143. http://dx.doi.org/10.1186/1471-2164-15-1

Iibuchi, C., & Imahori, K. (1978). Heterogeneity and its relation to the subunit structure of the soybean 7S globulin. *Agricultural and Biological Chemistry*, *42*(1), 31-36. http://dx.doi.org/10.1080/00021369.1978.10862918

Karr-Lilienthal, L. K., Kadzere, C. T., Grieshop, C. M., & Fahey Junior, G. C. (2005). Chemical and nutritional properties of soybean carbohydrates as related to nonruminants: A review. *Livestock Production Science*, *97*(1), 1-12. http://dx.doi.org/10.1016/j.livprodsci.2005.01.015

Khatib, K. A., Aramouni, F. M., Herald, T. J., & Boyer, J. E. (2002). Physicochemical characteristics of soft tofu formulated from selected soybean varieties. *Journal of Food Quality*, *25*(4), 289-303. http://dx.doi.org/10.1111/j.1745-4557.2002.tb01026.x

Kitamura, K., Takagi, T., & Shibasaki, K. (1976). Subunit structure of soybean 11S globulin. *Agricultural and Biological Chemistry*, *40*(9), 1837-1844. http://dx.doi.org/10.1080/00021369.1976.10862297

Krishnan, H. B. (2005). Engineering soybean for enhanced sulfur amino acid content. *Crop Science*, *45*(2), 454-461. http://dx.doi.org/10.2135/cropsci2005.0454

Kumar, V., Rani, A., Goyal, L., Dixit, A. K., Manjaya, J. G., Dev, J., & Swamy, M. (2010). Sucrose and raffinose family oligosaccharides (RFOs) in soybean seeds as influenced by genotype and growing location. *Journal of Agricultural and Food Chemistry*, *58*(8), 5081-5085. PMid:20353171. http://dx.doi.org/10.1021/jf903141s

Kumar, V., Rani, A., Solanki, S., & Hussain, S. M. (2006). Influence of growing environment on the biochemical composition and physical characteristics of soybean seed. *Journal of Food Composition and Analysis*, *19*(2-3), 188-195. http://dx.doi.org/10.1016/j.jfca.2005.06.005

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, *227*(5259), 680-685. http://dx.doi.org/10.1038/227680a0

Liu, S., Zhou, R., Tian, S., & Gai, J. (2007). A study on subunit groups of soybean protein extracts under SDS-PAGE. *Journal of the American Oil Chemists' Society*, *84*(9), 793-801. http://dx.doi.org/10.1007/s11746-007-1111-z

Matthäus, B., & Brühl, L. (2001). Comparison of different methods for the determination of the oil content in oilseeds. *Journal of the American Oil Chemists' Society*, *78*(1), 95-102. http://dx.doi.org/10.1007/s11746-001-0226-y

Messina, M., Sievenpiper, J. L., Williamson, P., Kiel, J., & Erdman Junior, J. W. (2022). Perspective: soy-based meat and dairy alternatives, despite classification as ultra-processed foods, deliver high-quality nutrition on par with unprocessed or minimally processed animal-based counterparts. *Advances in Nutrition*, *13*(3), 726-738. PMid:35325028. http://dx.doi.org/10.1093/advances/nmac026

Min, S., Yu, Y., & Martin, S. S. (2005). Effect of soybean varieties and growing locations on the physical and chemical properties of soymilk and tofu. *Journal of Food Science*, *70*(1), C8-C21. http://dx.doi.org/10.1111/j.1365-2621.2005.tb09026.x

Mozzoni, L., Shi, A., & Chen, P. (2013). Genetic analysis of high sucrose, low raffinose, and low stachyose content in V99-5089 soybean seeds. *Journal of Crop Improvement*, *27*(5), 606-616. http://dx.doi.org/10.1080/15427528.2013.812998

Mujoo, R., Trinh, D. T., & Ng, P. K. (2003). Characterization of storage proteins in different soybean varieties and their relationship to tofu yield and texture. *Food Chemistry*, *82*(2), 265-273. http://dx.doi.org/10.1016/S0308-8146(02)00547-2

Murphy, P. A., & Resurreccion, A. P. (1984). Varietal and environmental differences in soybean glycinin and beta-conglycinin content. *Journal of Agricultural and Food Chemistry*, *32*(4), 911-915. http://dx.doi.org/10.1021/jf00124a052

Obendorf, R. L., Horbowicz, M., Dickerman, A. M., Brenac, P., & Smith, M. E. (1998). Soluble oligosaccharides and galactosyl cyclitols in maturing soybean seeds in planta and in vitro. *Crop Science*, *38*(1), 78-84. http://dx.doi.org/10.2135/cropsci1998.0011183X003800010014x

Pesic, M. B., Vucelic-Radovic, B. V., Barac, M. B., & Stanojevic, S. P. (2005). The influence of genotypic variation in protein composition on emulsifying properties of soy proteins. *Journal of the American Oil Chemists' Society*, *82*(9), 667-672. http://dx.doi.org/10.1007/s11746-005-1126-x

Qin, P., Song, W., Yang, X., Sun, S., Zhou, X., Yang, R., Li, N., Hou, W., Wu, C., Han, T., & Ren, G. (2014). Regional distribution of protein and oil compositions of soybean cultivars in China. *Crop Science*, *54*(3), 1139-1146. http://dx.doi.org/10.2135/cropsci2013.05.0314

Qin, P., Wang, T., & Luo, Y. (2022). A review on plant-based proteins from soybean: Health benefits and soy product development. *Journal of Agriculture and Food Research*, *7*, 100265. http://dx.doi.org/10.1016/j.jafr.2021.100265

Samard, S., & Ryu, G. H. (2019). Physicochemical and functional characteristics of plant protein-based meat analogs. *Journal of Food Processing and Preservation*, *43*(10), e14123. http://dx.doi.org/10.1111/jfpp.14123

SAS Institute Inc (1993). *SAS 9.1.3 Help and Documentation*. Cary: SAS Institute Inc.

Schaefer, M. J., & Love, J. (1992). Relationships between soybean components and tofu texture 1. *Journal of Food Quality*, *15*(1), 53-66. http://dx.doi.org/10.1111/j.1745-4557.1992.tb00975.x

Seto, K. C., & Ramankutty, N. (2016). Hidden linkages between urbanization and food systems. *Science*, *352*(6288), 943-945. PMid:27199419. http://dx.doi.org/10.1126/science.aaf7439

Sha, L., & Xiong, Y. L. (2020). Plant protein-based alternatives of reconstructed meat: Science, technology, and challenges. *Trends in Food Science & Technology*, *102*, 51-61. http://dx.doi.org/10.1016/j.tifs.2020.05.022

Stanojevic, S. P., Barac, M. B., Pesic, M. B., & Vucelic-Radovic, B. V. (2011). Assessment of soy genotype and processing method on quality of soybean tofu. *Journal of Agricultural and Food Chemistry*, *59*(13), 7368-7376. PMid:21627319. http://dx.doi.org/10.1021/jf2006672

Taira, H. (1990). Quality of soybeans for processed foods in Japan. *Japan Agricultural Research Quarterly*, *24*(3), 224-230. Retrieved in 2023, August 19, from https://www.jircas.go.jp/en/publication/list/jarq

Wilcox, J. R., & Shibles, R. M. (2001). Interrelationships among seed quality attributes in soybean. *Crop Science*, *41*(1), 11-14. http://dx.doi.org/10.2135/cropsci2001.41111x

Yu, X., Yuan, F., Fu, X., & Zhu, D. (2016). Profiling and relationship of water-soluble sugar and protein compositions in soybean seeds. *Food Chemistry*, *196*, 776-782. PMid:26593554. http://dx.doi.org/10.1016/j.foodchem.2015.09.092

Zhang, D., Lü, H., Chu, S., Zhang, H., Zhang, H., Yang, Y., Li, H., & Yu, D. (2017). The genetic architecture of water-soluble protein content and its genetic relationship to total protein content in soybean. *Scientific Reports*, *7*(1), 5053. PMid:28698580. http://dx.doi.org/10.1038/s41598-017-04685-7

Zheng, L., Regenstein, J. M., Zhou, L., & Wang, Z. (2022). Soy protein isolates: A review of their composition, aggregation, and gelation. *Comprehensive Reviews in Food Science and Food Safety*, *21*(2), 1940-1957. PMid:35182006. http://dx.doi.org/10.1111/1541-4337.12925

Zhou, Y., Li, X., Hua, Y., Kong, X., Zhang, C., Chen, Y., & Wang, S. (2019). The absence of lipoxygenase and 7S globulin of soybeans and heating temperatures on the properties of soymilks and soy yogurts. *Lebensmittel-Wissenschaft und -Technologie-. Lebensmittel-Wissenschaft + Technologie*, *115*(4), 108431. http://dx.doi.org/10.1016/j.lwt.2019.108431

Zilić, S. M., Barać, M. B., Pešić, M. B., Mladenović Drinić, S. D., Ignjatović-Micić, D. D., & Srebrić, M. B. (2011). Characterization of proteins from kernel of different soybean varieties. *Journal of the Science of Food and Agriculture*, *91*(1), 60-67. PMid:20812371. http://dx.doi.org/10.1002/jsfa.4148

Funding: None.

Received: Apr. 19, 2023; Accepted: Nov. 21, 2023

Associate Editor: Priscila Zaczuk Bassinello.