RELATIONSHIPS BETWEEN YIELD AND NUTRITIONAL COMPONENTS OF VEGETABLE-TYPE SOYBEAN GENOTYPES USING CORRELATIONS AND PRINCIPAL COMPONENT ANALYSIS

Rouxléne VAN DER MERWE*; Angeline VAN BILJON; Arno HUGO; Jacques VAN DER MERWE

Department of Plant Sciences. University of the Free State. Bloemfontein, South Africa

* VanDerMerweR@ufs.ac.za

Received: September-06, 2021
Accepted: October-02, 2021
Published on-line: December-05, 2021

Citation:
Van Der Merwe R, Van Biljon A, Hugo A, Van der Merwe J. 2021. Relationships between yield and nutritional components of vegetable-type soybean genotypes using correlations and principal component analysis. Mol 21: 2

Abstract

Vegetable-type soybean is a special type soybean that is desired due to its rich nutritional composition and health benefits. A desirable cultivar should be high yielding, but the fresh beans should also have a high protein content, low oil content and a sweet taste. Thus, during selection for simultaneous improvement of yield and nutritional quality, trait associations need to be taken into account. The aim of this study was to determine correlation coefficients among pod yield, yield components and nutritional quality components in nine vegetable-type soybean cultivars over two seasons (2016/2017 and 2018/2019) at one location. The trials were planted in a randomised complete block design with three replications on a farm in the Free State province of South Africa. Results showed no significant correlations between pod yield and any of the nutritional components (except a positive correlation between total pod yield and ash content).

However, some significant correlations were observed between yield components and nutritional components. Significant negative correlations were found between crude fat respectively with number of pods on branches, number of branches per plant, number of pods at R6 growth stage and total number of pods per plant. The only significant negative correlation for sucrose was with number of nodes on the main stem. Monounsaturated fatty acids showed significant negative correlations with number of nodes on the main stem and number pods at the R6 growth stage, while polyunsaturated fatty acids showed significant positive correlations with six yield components respectively. Principal component analysis confirmed most of the correlations observed and further showed that most of the genetic variation could be attributed to yield components than nutritional components.

Generally, no strong associations were observed between total pod yield and marketable pod yield respectively with the nutritional quality components. However, significant correlations between yield components and nutritional components that need to be considered in the breeding programme include the positive correlations between the number of pods per plant respectively with polyunsaturated fatty acids and ash content and the negative correlation with crude protein content. Sucrose content will not be influenced when selection is applied for improved yield. Components that contributed most to the variability and that should be considered as selection criteria for improved yield and nutritional quality include crude protein, crude fat, sucrose and total polyunsaturated fatty acids.

Resumen

La soja de tipo hortícola es de un tipo especial con una composición nutricional que aporta beneficios para la salud. Una variedad de este tipo debe tener elevado rendimiento, y los granos frescos también deben tener un alto contenido de proteínas, bajo contenido de aceite y sabor dulce. Por lo tanto, durante la selección para la mejora simultánea del rendimiento y la calidad nutricional, deben tenerse en cuenta las asociaciones de diferentes caracteres.

El objetivo de este estudio fue determinar los coeficientes de correlación entre el rendimiento de vaina, los componentes del rendimiento y los componentes de la calidad nutricional en nueve variedades de soja de tipo hortícola, durante dos temporadas (2016/2017 y 2018/2019). Los ensayos de campo se llevaron a cabo según un diseño de bloques completos aleatorizados con tres repeticiones en una finca experimental en la provincia de Free State de Sudáfrica. Los resultados no mostraron correlaciones significativas entre el rendimiento de las vainas y ninguno de los componentes nutricionales (excepto una correlación positiva entre el rendimiento total de vainas y el contenido de cenizas). Sin embargo, se observaron algunas correlaciones significativas entre los componentes del rendimiento y los componentes nutricionales. Se encontraron correlaciones negativas significativas entre la grasa bruta con el número de vainas, el número de ramificaciones por planta, el número de vainas en la etapa de crecimiento R6 y el número total de vainas por planta. La única correlación negativa significativa en el contenido en sacarosa fue con el número de nudos en el tallo principal. Los ácidos grasos monoinsaturados mostraron correlaciones negativas significativas con el número de nudos en el tallo principal y el número de vainas en la etapa de crecimiento R6, mientras que los ácidos grasos poliinsaturados mostraron correlaciones positivas significativas con seis componentes de rendimiento. El análisis de componentes principales confirmó la mayoría de las correlaciones observadas y mostró, además, que la mayor parte de la variación genética podría atribuirse a componentes de rendimiento más que a componentes nutricionales. En general, no se observaron asociaciones relevantes entre el rendimiento total de las vainas y el rendimiento de las vainas comercializables con los componentes de la calidad nutricional. Sin embargo, las correlaciones significativas entre los componentes del rendimiento y los componentes nutricionales, que deben tenerse en cuenta en un programa de mejora, incluyen las correlaciones positivas entre el número de vainas por planta con los ácidos grasos poliinsaturados y el contenido de cenizas y la correlación negativa con el contenido de proteína cruda. El contenido de sacarosa no se verá afectado cuando por la selección para mejorar el rendimiento. Los componentes que más contribuyeron a la variabilidad, que deberían considerarse como criterios de selección para mejorar el rendimiento y la calidad nutricional, incluyen la proteína bruta, la grasa bruta, la sacarosa y los ácidos grasos poliinsaturados totales.

Introduction

Vegetable-type soybean (*Glycine max* L.) is a special type of soybean that is harvested at the R6 growth stage and consumed as a vegetable (Shanmugasundaram et al. 1991). The fresh beans are mainly desired by consumers due to their rich nutritional composition and health benefits (Xu et al. 2016; Zeipina et al. 2017). In South Africa, vegetable-type soybean has been introduced as alternative high protein food source that can be cultivated by resource poor and small-scale farmers on a contract basis in order to alleviate poverty and malnutrition in the sub-urban communities. However, due to poor adaptation of introduced cultivars and low yield realised by farmers, there is a need to develop better adapted cultivars for local production conditions. In addition, no information regarding the nutritional composition as well as the relationship between nutritional components and yield is available for these cultivars.

Many traits that are focussed on within a breeding programme are complex and the result of an interaction between a number of components. Not only is it important to understand the relationship between yield and its components, but also the relationship between yield and nutritional components.

An understanding of these relationships is of utmost importance in order to adopt the most appropriate selection criteria in breeding (Sarawgi et al. 1997; Govindaraj et al. 2009; Djanta et al. 2020). A desired vegetable-type soybean cultivar should not only be high yielding, but the fresh beans should also have a high protein and low oil content with a sweet taste (Rao et al. 2002; Zeipina et al. 2017). To obtain this desired cultivar, trait associations need to be taken into account during selection in the breeding programme.

In order to determine the degree of relationships among characteristics, correlation coefficients are determined (Sarutayophat 2012). Correlation studies enable breeders to evaluate the strength of relationships between various characteristics as well as to determine the magnitude and direction of changes expected during selection (Aondover et al. 2013). A number of studies have determined the relationships among yield and nutritional quality traits in commodity soybean (Akram et al. 2011; Assefa et al. 2018) and in vegetable-type soybean (Sarutayophat 2012; Jiang et al. 2018). These studies mainly showed that yield was significantly and positively correlated with number of pods per plant, number of branches per plant, plant height and maturity type. Previous studies on correlations among nutritional components showed a significant negative correlation between protein content and fat content while Assefa et al. (2018) reported positive correlations between yield respectively with protein and fat content.

Since both yield and nutritional quality traits are polygenic and determined by several components, a technique is required to identify and prioritise the important traits by minimising the number of traits for effective selection and genetic gain in the breeding programme. Principal component analysis (PCA) is a data reduction technique, which identifies the minimum number of traits that contribute to maximum variability and it can rank genotypes based on PC scores (Dubey et al. 2018). Previous studies have used PCA to assess genetic variability among commodity and vegetable-type soybean genotypes using yield and its related traits (Vianna et al. 2013) as well as to select superior soybean progeny from segregating populations and to identify more efficient crosses and parents in the breeding programme (Dallastra et al. 2014).

Since no information is available on relationships between yield and nutritional components of vegetable-type soybean cultivars that are currently tested under South African production conditions, the aims of this study were 1) to determine correlation coefficients among pod yield, yield components and nutritional quality components in nine vegetable-type soybean cultivars over two seasons and 2) to identify yield and quality components that will account for most of the variance in the observed variables and that can be used for selection in the breeding programme.

Material and Methods

Nine vegetable-type soybean genotypes, obtained from the Edamame Development Programme (EDP) in South Africa were used in this study (Table 1). These are introductions that were initially imported from the Word Vegetable Centre (AVRDC) based in Taiwan.

Field trials were planted on farm Sorgvliet (29°29'30.5"S 25°31'12.1"E; 1293 masl) in the Fauresmith district of the Free State province of South Africa. The warm and drier climate for this province, compared to KwaZulu-Natal province is not ideal for soybean production; however, it was shown to be a stable location for vegetable-type soybean cultivar evaluation (Smit 2019). Trials were repeated over two summer seasons of 2016/2017 and 2018/2019. Trials were established during November after the first spring rain and when the chances for late frost were slim. The first trial was planted on 16 December 2016 while the second field trial was planted on 20 November 2018. Both trials were planted in a randomised complete block design with three replications (Figure 1). The plots consisted of four rows, each 3 m in length with an inter-row spacing of 0.7 m. Before planting, seed were inoculated with *Bradyrhizobium japonicum* (strain WB74), which was obtained from Stimuplant. For both trials, seed were planted using an Earthway Precision Garden Seeder 1001-B. Seed were planted to a final population of 350 000 plants/ha^{-1}.

Table 1. List of vegetable-type soybean genotypes used in the study.

Name	AVDRC code	Seed coat colour	Flower colour	Maturity group*
AGS292	VI060637	Yellow	Purple	Quick
AGS354	-	Green	Purple	Medium
AGS457	AVSB0803	Brown	White	Quick
UVE3	AVSB0308	Light-green	White	Quick
UVE9	AVSB0808	Yellow	White	Quick
UVE10	AVSB0809	Light-green	White	Quick
UVE16	AVSB1003	Green	White	Medium
UVE21	TAI 75	Green	White	Medium-quick
UVE23	ZHONGKE-1117	Green	White	Medium-quick

*Maturity grouping for South Africa was based on the number of days to grain harvest (quick = 100 days, medium-quick = 115 days, medium = 130 days, medium-late = 145 days, late = 160 days); Sources: Van der Merwe et al. (2018); AVRDC (2019)

Trial management practices included irrigation and nutrients applied to the soil. Supplemental irrigation using a linear irrigation system was applied twice a week to ensure optional plant population and to prevent plants experiencing drought stress. Nutrients that were applied to the soil consisted of Omnia NPK 2:3:4(30) applied at 300 kg/ha before planting and 300 kg/ha as top-dressing. Trials were hand-weeded as necessary.

Figure 1. Nine vegetable-type soybean cultivars planted in a randomised complete block design with three replications over two seasons (2016/2017 and 2018/2019) on farm Sorgvliet in the Free State province of South Africa.

Whole plants were harvested by hand at the R6 growth stage; however, harvest date differed across genotypes since they differed in terms of maturity type (Table 1). Yield component data were collected on 10 randomly selected plants per plot and the averages used for analysis. Traits measured on the 10 plants included plant height (PH), which was measured from the plant base to the highest node on the plant in centimeters using a tape measure; number of branches per plant (BpP), which was determined by counting all branches that contained pods; number of pods on branches (BPo), which was determined by counting all pods on all branches per plant; number of pods on the main stem (MstPo), which was determined by counting all pods on the main stem; total number pods per plant (TPo), which was determined by counting all pods on all branches and the main stem per plant;

number of fully filled pods (marketable pods) that are at the R6 growth stage (R6Po), which was determined by counting all R6 pods per plant; number of nodes on the main stem (MstN), which was determined by counting all nodes on the main stem per plant; and number of nodes containing pods (NPo), which was determined by counting all nodes that contained pods per plant.

On a per plot basis, traits that were measured in grams (g) on 50 plants harvested from the middle rows of the plot included total pod mass (all green pods harvested from the plant) and pod mass of R6 pods only. Total pod yield and R6 pod yield (or marketable pod yield) were each determined in kg/ha using:

Yield (kg/ha) = [Mass of pods (kg) x 10 000 m^2] / plot area (m^2), where, the plot area used in the equation was adjusted to the plot area (m^2) of 50 plants.

From the 50 plants harvested at the R6 growth stage, all pods were removed from the stems. The R6 pods were separated from the immature and aborted pods and these were stored at -18ºC in the freezer until needed for laboratory analyses. For nutritional quality analysis, the R6 pods were defrosted overnight, shelled and the beans dried in a drying oven. The oven temperature was gradually increased from 70ºC for day one, to 80ºC for day two and 100ºC for days three to five. The dried beans were taken from the oven and milled using an IKA Yellow Line A10 grinder and the flour used for analyses.

For crude protein content determination, a procedure similar to the method described by Zeeman (2005) was followed for protein extraction. Approximately 0.2 g dry flour of each sample was weighed into a tiny foil cup. Each foil cup and sample were then inserted into a Leco Nitrogen Analyser (Wright and Baily 2001). Nitrogen (N) content was determined automatically by combustion in oxygen. The calculation was done by using a factor of 6.25, to convert the N content of the samples to crude protein content (g/kg).

Crude fat content was determined using the Soxhlet extraction method (AOAC 1995). Samples of 4 g dry flour were weighed into each filter paper thimble. Thimbles were placed in a vacuum oven at 50°C and left overnight. Phosphorus pentoxide was used as a moisture absorbent. The thimbles were placed in the Soxhlet apparatus and the fat extracted from the flour using hexane as solvent. Following the extraction, the hexane was evaporated on a water bath and the fat weighed. The flasks were left overnight in the refrigerator and the fat was filtered the next day. The filtered fat was placed in a vacuum oven for two hours at 50°C. The total extractable fat content was determined gravimetrically and expressed as g/kg.

For fatty acid determination, total fat from the seed sample was quantitatively extracted, according to the method of Folch et al. (1957). A rotary evaporator was used to dry the fat extracts under vacuum and the extracts were dried overnight in a vacuum oven at 50°C, using phosphorus pentoxide as moisture adsorbent. The extracted fat was stored in a polytop (glass vial, with push-in top) under a blanket of nitrogen and frozen at -20°C pending fatty acid analyses. A lipid aliquot (20 mg) of lipid was transferred into a Teflon-lined screw-top test tube by means of a disposable glass Pasteur pipette. Fatty acids were trans-esterified to form methyl esters using 0.5 N sodium hydroxide (NaOH) in methanol and 14% boron trifluoride in methanol (Slover and Lanza 1979; Hur et al. 2004; Diaz et al. 2005). Non-adecanoic acid (C19:0) was used as internal standard. Fatty acid methyl esters (FAME's) from fat were quantified using a Varian 430 flame ionization gas chromatograph. Galaxy Chromatography Software recorded the chromatograms.

FAME samples were identified by comparing the retention times of FAME peaks from samples with those of standards obtained from Supelco (Supelco 37 Component Fame Mix 47885-U, Sigma-Aldrich Aston Manor, Pretoria, South Africa). Fatty acids were expressed as the proportion of each individual fatty acid to the total of all fatty acids present in the sample. For this study, total monounsaturated fatty acid (MUFA), total polyunsaturated fatty acid (PUFA) and total saturated fatty acid (SFA) contents were determined (in percentage) using the equations:

MUFA (%) = Palmitoleic (%) + Heptadecenoic (%) + Oleic (%) + Vaccenic (%)
PUFA (%) = Linolelaidic (%) + Linoleic (%) + α-Linolenic (%) + Eicosadienoic (%) + Eicosatrienoic (%)

SFA (%) = Myristic (%) + Palmitic (%) + Margaric (%) + Stearic (%) + Arachidic (%) + Tricosanoic (%) + Lignoceric (%)

For sucrose content determination, samples of 0.25 g dry flour was used for extraction with 4 ml of 80% ethanol. The tubes were placed in an 80°C water bath for 30 min and homogenised for 10 min. The mixture was centrifuged at 5000 revolutions per minute (rpm) for 10 min and the supernatant was removed to another tube. The pellet was washed twice with 2 ml 80% ethanol and filled to a 10 ml volume. The supernatant was used to quantify the content of sucrose. The content of sucrose was analysed by using the resorcinol hydrochloric acid method (Huber et al. 1982) where 0.2 ml of 0.1% resorcinol and 3 ml of 30% hydrochloric acid (HCl) were added to the tube containing the supernatant to be incubated at 80ºC for 10 min. Absorbance of the developed colour was measured at 480 nm. From the absorbance readings a standard curve was drawn to calculate the sucrose content (g/kg).

For ash content determination, samples of 1 g dry flour were heated to 400ºC where the weight loss was measured. The weight lost by the heating process were essentially the organic matter (OM) burnt off. The ash content is determined by subtracting the OM from 1000 and is expressed in g/kg.

All statistical analyses were performed using Genstat release 18 software (VSN International 2017). In order to determine the relationships between pod yield, yield and nutritional components, Pearson's correlation coefficients were calculated from the mean values of all measured traits. Principal component analysis (PCA) was used to determine the genetic divergence between genotypes, which permits the identification and selection of the most promising genotypes for cultivation and improvement, and also to evaluate the relative importance of all traits in the total variation available among the genotypes (Vianna et al. 2013).

Results and Discussion

According to De Mello Filho et al. (2004) a soybean breeding programme should not be focussed on the genetic improvement of isolated traits but rather on the genetic improvement of a set of traits. This will give a breeder an idea on how the intervention in one trait, such as yield can cause alternation in other traits (e.g. nutritional quality components). Results on the correlation analysis of pod yield, yield components and nutritional components combined across nine vegetable-type soybean cultivars and two seasons are presented in Table 2.

Generally, crude protein content showed no strong or significant correlations with total pod yield, marketable pod yield or any of the eight yield components. This was in contrast with previous studies where a negative correlation between grain yield and protein content was observed for commodity soybean (De Mello Filho et al. 2004; Popović et al. 2011). The non-significant correlations between the yield traits and protein content was also in contrast with Assefa et al. (2018) who indicated a positive association between yield and protein content. Therefore, selection for improved yield in the vegetable-type soybean breeding programme using these genotypes would not negatively influence protein content.

On the other hand, crude fat content showed significant negative correlations with four yield components, e.g. number of pods on branches, number of branches per plant, number of pods at R6 growth stage and total number of pods per plant. These negative correlations indicate that an increase in these yield components could result in a reduction in crude fat content.

The significant negative correlation between total number of pods and crude fat content was not in accordance with the non-significant positive correlation found by Adeshina et al. (2019) in commodity soybeans. Even though negative correlations were observed for some yield components, crude fat content showed no significant correlations with total pod yield or marketable pod yield, which are the most important traits in vegetable-type soybean.

Vegetable-type soybean desirably should have a high sugar content while the sucrose content should be greater than 10% of dry weight during mid-pod development until maturity (Konovsky et al. 1994; Masuda and Harada 2000). Therefore, knowledge on the association between yield and sucrose content is essential for the breeding programme. Results from this study indicated that sucrose content was not significantly correlated with pod yield, marketable pod yield and the yield components apart from the number of nodes on the main stem. Therefore, selection for improved yield in the vegetable-type soybean breeding programme would not negatively influence sucrose content. In addition, this finding was in accordance with results from Mebrahtu and Mohamed (2006) who also reported a non-significant correlation between sucrose content and green pod yield in vegetable-type soybean. However, these authors reported a significant and positive correlation between sucrose content and plant height but this observation was not found in the current study.

The fatty acid profile showed variable magnitudes and directions of correlations between pod yield and the yield components. Results indicated that total monounsaturated fatty acids would be less affected when selection is based on yield components as compared to total polyunsaturated fatty acids. While total saturated fatty acids showed no significant correlations with yield or yield components, indicating that these traits are independent from each other. Generally, soybean fat contains saturated fatty acids such as palmitic (16:0) and stearic (18:0) acids; however, the levels of saturated fatty acids are low (Abdelghany et al. 2020). On the other hand, soybean fat contains a high level of unsaturated fatty acids including oleic (18:1), linoleic (18:2) and linolenic (18:3) acids (Abdelghany et al. 2020).

The polyunsaturated fatty acids and especially the omega-3 fatty acids are perceived as a beneficial dietary intervention for the prevention and treatment of cardiovascular disease (Ander et al. 2003). Therefore, a diet that contains higher levels of polyunsaturated fatty acids as compared to saturated fatty acids is regarded healthier.

Results from this study indicated that the total polyunsaturated fatty acids content was generally positively correlated with the yield components and showed significant correlations specifically with number of nodes on the main stem, number of pods on main stem, number of pods per node, plant height, number of pods at the R6 growth stage and total number of pods per plant. On the other hand, the monounsaturated fatty acids content showed significant negative correlations specifically with number of pods on main stem and number of pods at the R6 growth stage. Thus, enhancement of total pod yield and marketable pod yield would not necessary lead to a reduction in the level of monounsaturated fatty acids. In general, correlations among crude fat, fatty acid profiles and the yield traits showed that breeding for enhanced yield, indirectly through selection for more pods per plant, could lead to a reduction in the fat content of the fresh beans but then an increase in the level of polyunsaturated fatty acids.

Ash content is a measure of the total amount of minerals present within a food and its determination is part of proximate analysis for nutritional evaluation. Thus it is an important quality attribute for some food ingredients (Ismail 2017). Results from this study indicated that ash content was not significantly correlated with pod yield and marketable pod yield; however, it was significantly and positively correlated with some of the yield components, including number of nodes on the main stem, number of pods on the main stem, number of nodes with pods, plant height and total number of pods per plant. Thus, selection for improved yield, through direct selection for these yield components, would have a positive influence on ash content in the vegetable-type soybean breeding programme.

Table 2. Correlation coefficients and their significance between pod yield, yield components and nutritional components.

	BPo	BpP	MstN	MstPo	NPo	PH	R6Po	TPo	TPY	MPY	CP	CF	Sucr	MUFA	PUFA	SFA
BpP	0.736*	-														
MstN	0.164	-0.017	-													
MstPo	0.141	0.235*	0.358*	-												
NPo	0.544*	0.434*	0.558*	0.536*	-											
PH	0.101	-0.096	0.666*	0.375*	0.480*	-										
R6Po	0.795*	0.509*	0.582*	0.378*	0.803*	0.378*	-									
TPo	0.867*	0.567*	0.448*	0.472*	0.756*	0.331*	0.934*	-								
TPY	0.369*	0.064	-0.015	-0.207	0.233	0.084	0.260	0.246	-							
MPY	0.503*	0.247	0.082	-0.267	0.286*	0.107	0.382*	0.329*	0.903*	-						
CP	0.190	0.122	0.034	-0.117	0.122	-0.146	0.205	0.165	-0.031	0.045	-					
CF	-0.306*	-0.290*	-0.209	-0.053	-0.233	-0.057	-0.383*	-0.323*	0.213	0.072	-0.432*	-				
Sucr	-0.086	-0.169	-0.410*	-0.242	-0.258	-0.234	-0.271	-0.176	0.182	0.093	-0.018	0.418*	-			
MUFA	-0.223	-0.191	-0.308*	-0.115	-0.217	-0.116	-0.280*	-0.258	-0.121	-0.222	-0.239	-0.197	0.282*	-		
PUFA	0.225	0.084	0.505*	0.314*	0.380*	0.336*	0.385*	0.396*	0.252	0.280*	0.175	0.148	-0.108	-0.581*	-	
SFA	-0.038	-0.056	-0.210	-0.211	-0.241	-0.261	-0.150	-0.160	-0.099	-0.126	0.323*	-0.486*	0.016	-0.059	-0.425*	-
Ash	0.155	-0.016	0.367*	0.435*	0.371*	0.430*	0.323*	0.384*	-0.150	-0.141	0.179	-0.127	-0.413*	-0.223	0.225	-0.091

* Significant at $P < 0.05$, BPo = number of pods on branches, BpP = number of branches per plant, MstN = number of nodes on the main stem, MstPo = number of pods on the main stem, NPo = number of nodes with pods, PH = plant height, R6Po = number of pods at R6 growth stage, TPo: total number of pods, TPY = total pod yield MPY = marketable pod yield, CP = crude protein, CF = crude fat, Sucr = sucrose, MUFA = monounsaturated fatty acids, PUFA = polyunsaturated fatty acids.

Results from the principle component analysis (Table 3 and Figure 2) showed that a larger amount of genotype divergence was due to yield traits than nutritional quality components. Principal component 1 (PC1), which accounted for the highest variability (53.18%), was mostly associated with total pod yield, marketable pod yield (R6 pod yield) and the yield components (number of pods on branches, number of nodes on main stem, number of nodes with pods, plant height, number of pods at R6 growth stage, total number of pods). On the other hand, PC2 varied positively with crude protein, saturated fatty acids and ash content and varied negatively with crude fat, monounsaturated fatty acids and number of branches per plant. These findings were in accordance with previous reports by Vianna et al. (2013) and Khan et al. (2014).

Results from the PCA biplot confirmed most of the correlations observed between the yield and nutritional components found in the correlation matrix. Polyunsaturated fatty acids showed strong positive associations with number of nodes on main stem, plant height, marketable pod yield (R6 pod yield), total pod yield, total number of pods, number of pods at R6 growth stage and number of nodes with pods. Similarly, ash content showed strong positive associations with number of nodes on the main stem, plant height and total number of pods per plant. Sucrose content showed a strong negative association with number of nodes on the main stem.

From the PCA biplot, good performing genotypes that are related to desired yield and nutritional components were identified, as done by Aondover et al. (2013) and Vianna et al. (2013). These included AGS354 and UVE10 that were related to yield components, polyunsaturated fatty acids and ash content and AGS292 that was related to number of pods on the main stem, number of pods on branches, crude fat and monounsaturated fatty acids content. Although UVE23 was negatively associated with the yield components, it was positively associated with sucrose content. Genotype UVE16, which showed average performance for yield traits, was strongly and positively associated crude protein, saturated fatty acids and ash content.

Table 3. Principal component analysis and factor loadings of pod yield, yield components and nutritional components across two seasons.

	PC1	PC2	PC3	PC4	PC5
Eigen values	9.04	2.76	1.64	1.16	1.03
Proportion of the total variance (%)	53.18	16.22	9.65	6.80	6.03
Cumulative variance (%)	53.18	69.40	79.05	85.85	91.88
Variables					
Number of pods on branches (BPo)	0.302	0.056	-0.215	0.083	-0.066
Number of branches per plant (BpP)	0.240	-0.326	-0.273	-0.187	-0.105
Number of nodes on main stem (MstN)	0.279	0.135	0.314	-0.252	0.030
Number of pods on main stem (MstPo)	0.221	-0.245	0.232	-0.173	0.211
Number of nodes with pods (NPo)	0.296	-0.108	0.261	0.092	0.0567
Plant height (PH)	0.284	0.134	0.238	0.049	-0.192
Number of pods at R6 growth stage (R6P	0.318	-0.024	0.149	0.139	0.062
Total number of pods (TPo)	0.321	-0.008	0.002	0.091	0.083
Total pod yield	0.274	0.033	0.025	0.363	-0.296
Marketable pod yield (R6 pod yield)	0.299	0.103	0.001	0.175	-0.302
Crude protein (CP)	-0.012	0.523	0.029	0.161	0.206
Crude fat (CF)	0.124	-0.348	-0.101	0.145	0.664
Sucrose	-0.132	-0.081	0.641	0.145	0.174
Monounsaturated fatty acids (MUFA)	-0.145	-0.362	0.063	0.543	-0.163
Polyunsaturated fatty acids (PUFA)	0.261	0.189	-0.037	-0.375	0.123
Saturated fatty acids (SFA)	-0.217	0.307	0.230	0.061	0.063
Ash	0.138	0.329	-0.314	0.395	0.389

Conclusions

The correlation analysis generally showed no strong correlations between total pod yield and marketable pod yield respectively with the nutritional quality components. However, significant correlations between some yield components and nutritional quality components were observed that need to be considered in the breeding programme. Positive selection for total number of pods per plant and number of marketable pods per plant might have a positive influence on both polyunsaturated fatty acids and ash content but a negative influence on crude protein content. Sucrose content will not be influenced when selection is applied for improved yield. Components that contributed most to the variability and that should be considered as selection criteria for improved yield and nutritional quality include crude protein, crude fat, sucrose and total polyunsaturated fatty acids content. Results not presented in this paper indicated significant genotype differences across most yield and nutritional components. The top performing genotypes in terms of marketable pod yield were AGS354 (4436 k/hag) and UVE10 (3103 kg/ha) and these ranked in the top three positions across all yield components. Genotypes that performed well in terms of nutritional components were UVE16 for high protein and low fat contents, and UVE21 and UVE23 for sucrose content. Genotypes UVE10, UVE9 and UVE3 showed the best fatty acid profile in terms of high unsaturated and low saturated fatty acid contents.

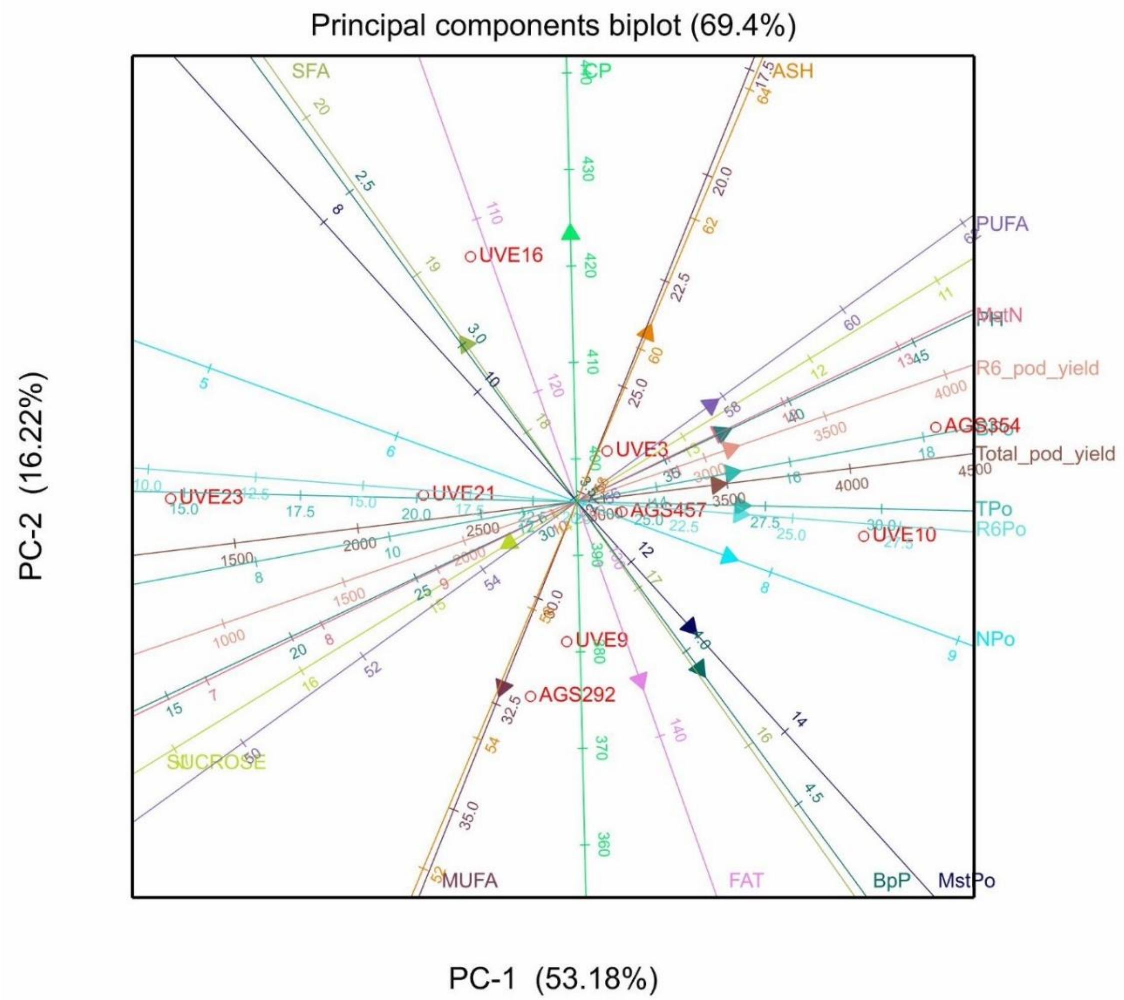

Figure 2. Principal component analysis biplot between PC1 and PC2 showing the contribution of pod yield, yield components and nutritional components to the variability of nine vegetable-type soybean genotypes across two seasons.

Acknowledgements

The authors would like to acknowledge the Edamame Development Program in South Africa for providing seed material used in this study, the National Research Foundation for providing research funding, Mrs J van der Merwe for performing nutritional quality analyses as well as Mr WJ van der Merwe for trial management.

References

Abdelghany AM, Zhanga S, Azama M Shaibua AS, Fenga Y, Lia Y, Tiana Y, Honga H, Lia B, Suna J. 2020. Profiling of seed fatty acid composition in 1025 Chinese soybean accessions from diverse ecoregions. The Crop Journal 8: 635-644

Adeshina DA, Abimiku SE, Inegbedion E, Ononokpono EG, Adeboye SE. 2019. Correlation between oil content and yield of some early maturing soybean (*Glycine max* (L.) Merill) genotypes in Keffi, Nasarawa State. International Journal of Environment, Agriculture and Biotechnology 4(3): 720-726

Akram RM, Fares WM, Fateh HAS Rizk AMA. 2011. Genetic variability, correlation and path analysis in soybean. Egyptian Journal of Plant Breeding 15(1): 89-102

Ander BP, Dupasquier CMS, Prociuk MA, Pierce GN. 2003. Polyunsaturated fatty acids and their effects on cardiovascular disease. Experimental and Clinical Cardiology 8(4): 164-172

AOAC (Association of Official Analytical Chemists International). 1995. Official methods of analyses of AOAC International, 16th edition, Method 4.5.01 (920.39), AOAC International, Maryland, USA

Aondover S, Lekan LB, Terkimbi V. 2013. Correlation, path coefficient and principal component analysis of seed yield in soybean genotypes. International Journal of Advanced Research 1(7): 1-5

Assefa Y, Bajjalieh N, Archontoulis S, Casteel S, Davidson D, Kovács P, Naeve S, Ciampitti IA. 2018. Spatial characterization of soybean yield and quality (amino acids, oil, and protein) for United States. Scientific Reports. www.nature.com/scientificreports. Accessed: 31/10/2020

AVRDC. 2019. World Vegetable Centre. Vegetable soybeans. Available at: https://avrdc.org/seed/improved-lines/vegetable-soybean. Accessed: 10/11/2019

Dallastra A, Unêda-Trevisoli SH, Ferraudo AS, Di Mauro AO. 2014. Multivariate approach in the selection of superior soybean progeny which carry the RR gene1. Revista Ciência Agronômica 45(3): 588-597

De Mello Filho OL, Sediyama CS, Moreira MA, Reis MS, Massoni GA, Piovesan ND. 2004. Grain yield and seed quality of soybean selected for high protein content. Pesquisa Agropecuária Brasileira 39(5): 445-450

Diaz MT, Alvarez I, De La Fuente J, Sanudo C, Campo MM, Oliver MA, Font i Furnols M, Montossi F, San Julian R, Nute GR, Caneque V. 2005. Fatty acid composition of meat from typical lamb production systems of Spain, United Kingdom, Germany and Uruguay. Meat Science 71: 256-263

Djanta MKA, Agoyi EE, Agbahoungba S, Quenum FJ, Chadare FJ, Assogbadja AE, Agbangla C, Sinsin B. 2020. Vegetable soybean, edamame: Research, production, utilization and analysis of its adoption in Sub-Saharan Africa. Journal of Horticulture and Forestry 12(1): 1-12

Dubey N, Avinashe HA, Shrivastava AN. 2018. Principal component analysis in advanced genotypes of soybean [*Glycine max* (L.) Merrill] over seasons. Plant Archives 18(1): 501-506

Folch J, Lees M, Sloane-Stanley GH. 1957. A simple method for the isolation and purification of total lipids from animal tissue. Journal of Biological Chemistry 226: 497-509

Govindaraj M, Selvi B, Rajarathinam S. 2009. Correlation studies for grain yield components and nutritional quality traits in pearl millet (*Pennisetum glaucum* (L.) R. Br.) germplasm, World Journal of Agricultural Sciences 5(6): 686-689

Hur SJ, Ye BW, Lee JL, Ha YL, Park GB, Joo ST. 2004. Effects of conjugated linoleic acid on color and lipid oxidation of beef patties during cold storage. Meat Science 66: 771-775

Ismail BP. 2017. Ash Content Determination. In: Nielsen SS (Ed) Food Analysis Laboratory Manual, third edition, pp 117-119. Springer International Publishing

Jiang G, Rutto L, Ren S. 2018. Evaluation of soybean lines for edamame yield traits and trait genetic correlation. American Society of Horticultural Science 53: 1732-1736

Khan MSA, Karim MA, Haque MM, Karim AJMS, Mian MAK. 2014. Variations in agronomic traits of soybean genotypes. SAARC Journal of Agriculture 12(2): 90-100

Konovsky J, Lumpkin TA, McClary D. 1994. Edamame: The vegetable soybean. In: O'Rourke AD (Ed), Understanding the Japanese Food and Agrimarket: A Multifaceted Opportunity, pp 173-181. Binghamton, NY: Haworth Press

Masuda R, Harada K. 2000. Carbohydrate accumulation in developing soybean seeds; sucrose and starch levels in 30 cultivars for soy foods. In: Proceedings of the Third International Soybean Processing and Utilization Conference. pp 67-68

Mebrahtu T, Mohamed A. 2006. Genetic variation for green pod yield and quality among vegetable soybean genotypes. Journal of Crop Improvement 16(1-2): 113-130

Popović V, Vidić M, Jocković D, Ikanović J, Jakšić S, Cvijanović G. 2012. Variability and correlations between yield components of soybean [*Glycine max* (L.) Merr.]. Genetika 44(1): 33-45

Rao MSS, Bhagsari AS, Mohamed AI. 2002. Fresh green seed yield and seed nutritional traits of vegetable soybean genotypes. Crop Science 42: 1950-1958

Sarawgi AK, Rastogi NK, Soni DK. 1997. Correlation and path analysis in rice accessions from Madhya Pradesh. Field Crops Research 52: 161-167.

Sarutayophat T. 2012. Correlation and path coefficient analysis for yield and its components in vegetable soybean. Songklanakarin Journal of Science and Technology 34(3): 273-277

Shanmugasundaram S, Cheng S-T, Huang M-T, Yan M-R. 1991. Varietal improvement of vegetable soybean in Taiwan. In: Shanmugasundaram S (Ed), Vegetable soybean: research needs for production and quality improvement, pp 30-42. Asian Vegetable Research and Development Center, Taiwan.

Singh N, Kaur N, Rana JC, Sharma SK. 2010. Diversity in seed and flour properties in field pea (*Pisum sativum*) germplasm. Food Chemistry 122: 518-525

Slover HT, Lanza E. 1979. Quantitative analysis of food fatty acids by capillary gas chromatography. Journal of the American Oil Chemists Society 56: 933-943

Smit A. 2019. Yield stability of edamame (*Glycine max* L.) introductions under South African production conditions. PhD Thesis, University of the Free State, South Africa

Taira H, Taira H. 1971. Influence of location on the chemical composition of soybean seeds: I. Protein, oil, carbohydrate, and ash contents. Japanese Journal of Crop Science 40(4): 530-544

Van der Merwe R, Tyawana S, Van Der Merwe J, Mwenye O. 2018. Evaluation of drought tolerance indices in vegetable-type soybean. Sociedad De Ciencias De Galicia – Science Society of Galicia (SCG). Mol 2018: 19-31. ISSN: 1133-3669

Vianna VF, Unêda-Trevisoli SH, Desidério JA, de Santiago S, Charnai K, Júnior JAF, Ferraudo AS, Di Mauro AO. 2013. The multivariate approach and influence of characters in selecting superior soybean genotypes. African Journal of Agricultural Research 8(30): 4162-4169

VSN International. 2017. Genstat® 18th Edition (PC/Windows XP). Copyright 2015. VSN International Ltd.

Wright AF, Bailey J. 2001. Organic carbon, total carbon, and total nitrogen determinations in soils of variable calcium carbonate contents using a Leco CN-2000 Dry Combustion Analyzer. Communications in Soil Science and Plant Analysis 32: 3243-3258

Xu Y, Cartier A, Kibet D, Jordan K, Hakala I, Davis S, Sismour E, Kering M. 2016. Physical and nutritional properties of edamame seeds as influenced by stage of development. Food Measure 10: 193-200

Zeeman DZ. 2005. Evaluation of sun-dried *Opuntia ficus-indica* var. Algerian cladodes in sheep diets. MSc Agriculture Dissertation, University of the Free State, Bloemfontein, South Africa

Zeipinia S, Alsina I, Lepse L. 2017. Insight in Edamame yield and quality parameters: A Review. Research for Rural Development 2: 40-45