agriculture

MDPI

Article

Genetic Diversity of Korean Black Soybean (*Glycine max* L.) Germplasms with Green Cotyledons Based on Seed Composition Traits

Ji Yun Lee [1,†], Hyun Jo [2,3,†], Chang Ki Son [1], Jeong Suk Bae [1] and Jeong-Dong Lee [2,4,*]

[1] Gyeongsangbuk-do Provincial Agricultural Research & Extension Service, Daegu 41404, Republic of Korea
[2] Department of Applied Biosciences, Kyungpook National University, Daegu 41566, Republic of Korea
[3] Upland-Field Machinery Research Center, Kyungpook National University, Daegu 41566, Republic of Korea
[4] Department of Integrative Biology, Kyungpook National University, Daegu 41566, Republic of Korea
* Correspondence: jdlee@knu.ac.kr; Tel.: +82-53-950-5709; Fax: +82-53-958-6880
† These authors contributed equally to this work.

Abstract: The demand for black soybeans (*Glycine max* (L.) Merr.) with green cotyledons is increasing because of their health benefits. Therefore, it is important to understand the genetic diversity of black soybean germplasms and to develop a new soybean cultivar. This study aimed to evaluate genetic diversity among 469 black soybean germplasms with green cotyledons based on seed composition traits. Twenty seed composition traits were analyzed to conduct correlation analysis, principal component analysis (PCA), and cluster analysis, which indicated that black soybean germplasms were divided into four clusters. Black soybean germplasms in cluster 1 had higher crude fat, lutein, chlorophyll *a*, chlorophyll *b*, and total chlorophyll contents, but lower cyanidin-3-glucoside content than those in clusters 2 and 3. However, germplasms in clusters 2 and 3 had the highest cyanidin-3-glucoside content. Moreover, germplasms in cluster 1 had significantly higher palmitic acid content than those in clusters 2 and 3. Germplasms in clusters 2 and 3 had relatively high α-linolenic acid content. Germplasms in cluster 4 had the highest oleic acid content. This study highlights the genetic diversity of black soybean germplasms with different seed composition traits, and the results of this study can be beneficial for soybean breeding programs, enabling them to develop new black soybean cultivars with green cotyledons and improved seed composition traits.

Keywords: genetic diversity; black soybean; green cotyledon; seed composition; breeding

check for updates

Citation: Lee, J.Y.; Jo, H.; Son, C.K.; Bae, J.S.; Lee, J.-D. Genetic Diversity of Korean Black Soybean (*Glycine max* L.) Germplasms with Green Cotyledons Based on Seed Composition Traits. *Agriculture* **2023**, *13*, 406. https://doi.org/10.3390/agriculture13020406

Academic Editors: Edyta Paczos-Grzeda, Volker Mohler and Sylwia Sowa

Received: 10 January 2023
Revised: 6 February 2023
Accepted: 7 February 2023
Published: 9 February 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Black soybeans (*Glycine max* (L.) Merr.) are one of the most significant food crops in Asian countries, including Korea, due to the high amount of protein, oil, and chemicals with health benefits in the soybean seeds [1]. Soybean paste and soy sauce are fermented foods that have been proven to have antiaging, anticancer, antibacterial, and antidiabetic properties [2]. In addition, soybeans are a valuable source of nutrition and one of the most important raw materials for vegetable oil as they contain 20% vegetable fat [3]. Soybean oil consists of five major fatty acids: 12% palmitic acid (16:0, PA), 4% stearic acid (18:0, SA), 23% oleic acid (18:1, OA), 53% linoleic acid (18:2, LA), and 8% α-linolenic acid (18:3, LNA) [4]. Furthermore, the alternation of fatty acid profiles affects the physicochemical properties of soybean oil, indicating that fatty acids are a major factor in determining the commercial use of soybean oil [5,6]. In addition to serving as a source of essential fatty acids, soybean oil exhibits pharmacological effects, such as lowering cholesterol levels and preventing cardiovascular diseases. Moreover, soybeans can be used as cosmetic raw materials, biodiesel, and industrial oil [5,7–17]. Soybeans also contain phytochemicals, such as isoflavones, tocopherols, and saponins, as well as antioxidants, including minerals

and vitamins. Consequently, it is widely known that soybeans provide excellent health benefits [18,19].

In Korea, black soybeans with green cotyledons have received attention as a functional food. Black soybeans contain a greater number of biologically functional compounds than yellow commodity soybeans, but they contain very similar nutrients. The seed coats of black soybeans contain pigments known as anthocyanins, which affect the color of the seed coat [20–30]. The seed coats of black soybeans consist of 11 anthocyanins: cyanidin-3-O-glucoside (C3G), delphinidin-3-O-glucoside (D3G), petunidin-3-O-glucoside (Pt3G), peonidin-3-O-glucoside (Pn3G), malvidin-3-O-glucoside (M3G), pelargonidin-3-O-lucoside (Pl3G), cyanidin-3-O-galactoside, cyanidin, catechin-cyanidin-3-O-glucoside, delphinidin-3-O-galactoside, and cyanidin-3-O-galactoside. Among them, the three main anthocyanins in the seed coats of black soybeans are C3G, D3G, and Pt3G [21,22,25,31].

Cotyledons in soybean seeds are green or yellow. Pigments, such as chlorophyll and lutein, play a major role in determining the color of the cotyledons in soybean seeds [32–34]. In general, the cotyledons, leaves, and pods of soybeans are green during the development stages. The plants turn yellow as they reach maturity because the chlorophyll and chloroplast components are decomposed. This is a natural aging process that occurs in various plants, including soybeans, at the maturity stage [35–38]. However, green senescent leaves and green cotyledons of matured seeds have been reported in various crops, and this phenomenon is known as "stay-green" [39–42]. Furthermore, nuclear inheritance and cytoplasmic inheritance have been identified as impairing chlorophyll degradations in soybeans, resulting in the green cotyledons seed trait. *D1* (*Glyma.01g214600* in W82.a2.v1 assembly) and *D2* (*Glyma.11g027400*) are involved in chlorophyll degradation and are nuclear genes inherited by green cotyledons in soybean seeds. Soybeans having double recessive mutations of *D1* and *D2* produce green colored cotyledons [43]. In addition, *cytG* in the chloroplast genome is associated with the decomposition of chlorophyll during senescence [40,43–47]. Lutein, a carotenoid pigment component that protects the photosynthetic system and absorbs light energy, was more abundant in green soybeans than in yellow ones [48,49]. Carotenoids and chlorophyll are biosynthesized in the chloroplast and the metabolisms of the two components are closely linked [50].

Black soybeans with green cotyledons contain anthocyanin, chlorophyll, and lutein, which reduce the risks of cardiovascular diseases and increase anticancer and antioxidant effects [21,51–55]. The seed compositions in soybean seeds are the major factors that determine the availability and use of soybeans. Therefore, it is essential to secure black soybean germplasms by evaluating morphological traits, such as seed size, color, maturity, and length, as well as various nutritional and bioactive substances (protein, fat, anthocyanin, and pigment components). This approach will aid in the development of new cultivars with a high content of functional substances. In the present study, the objective was to analyze the genetic diversity of 469 black soybean germplasms with green cotyledons and with 20 seed composition traits, and this information will be helpful in developing a new black soybean cultivar with green cotyledons.

2. Materials and Methods

2.1. Growth Conditions of Black Soybean Germplasms with Green Cotyledons

To understand the genetic diversity based on the seed composition traits, a collection of 470 soybean accessions, including three check cultivars, was used in this study. The collection consisted of a yellow soybean check cultivar, 'Uram' [56], and 469 black soybean germplasms with green cotyledons, including check cultivars 'Cheongja 3' [57] and 'Cheongja' [58]. Among the 469 black soybean germplasms with green cotyledons, 405 accessions were obtained from the National Agrobiodiversity Center in Jeonju, Republic of Korea, and 47 accessions were derived from the pure line selection of germplasms with different seed characteristics, plant appearance, flowering, and maturity. Fifteen black soybean accessions with green cotyledons were collected from Gyeongsanbuk-do, Republic of Korea. A total of 470 soybean accessions were grown at Gyeongsanbuk-do Agricultural

Research and Extension Service, Daegu, Republic of Korea, over 3 years from 2013 to 2015, and the planting dates were 14 June 2013, 29 May 2014, and 15 June 2015. The planting density was 80 × 15 cm, and each plot was 1.5 m in length (a single row). Each plot was planted with two replications, and two seedlings per hill were grown during the growing season. Finally, each plot was harvested in bulk and combined into one [33,59].

2.2. Seed Composition Analysis

The anthocyanin, chlorophyll, and lutein contents in the black soybean accessions with green cotyledons were determined using the methods described by Jo et al. [59]. The values for chlorophyll *a*, chlorophyll *b*, lutein, and six anthocyanins (D3G, C3G, Pt3G, Pl3G, Pn3G, and M3G) were obtained from our previous studies [33,59].

The fatty acid composition analysis was conducted using the method described by Bilyeu et al. [60], with some modifications. Briefly, 1.5 mL of chloroform:hexane:methanol (8:5:2, $v/v/v$) solution was added to 0.5 g of pulverized sample, followed by extraction at room temperature for 12 h for the fatty acid analysis. Later, 75 µL of a methylating reagent (0.25 M sodium methoxide:petroleum ether:ethyl ether (1:5:2, $v/v/v$)) was added to 100 µL of the extracted solvent, and hexane was added to adjust the total volume to 1 mL. The methylating agent was used to obtain methyl derivatives of fatty acids (fatty acid methyl esters, FAME), decrease polarity, increase volatility, and separate fatty acids more efficiently. A gas chromatography system (7890A, Agilent Technologies Inc., Santa Clara, CA, USA) attached to a flame ionization detector and an Agilent DB-FFAP capillary GC column (30 m × 0.25 mm, 0.25 µm) was used for the qualitative analysis of the fatty acids. The temperature of the sample injector was set to 230 °C and the temperature of the detector was set to 250 °C. The peak areas for the fatty acids were calculated based on standard curves of standard fatty acids (Fame #16, Restek) and the retention time of standard fatty acids.

The crude protein content (% crude protein = % N × 6.25) was calculated with an elemental analyzer (Vario Max CNS, Elementar Analysensysteme GmbH, Langenselbold, Main-Kinzig, Germany) using the Dumas method based on the AOAC method [61]. The pulverized sample (200 mg) was heated to 1200 °C to measure the nitrogen gas and obtain the nitrogen content. The crude fat content was extracted by utilizing a Soxtherm Automated Soxhlet Extraction System (Gerhardt, Königswinter, North Rhine-Westphalia, Germany) according to the AOAC method [61], and petroleum ether was used as the extraction solvent.

The crude fat was extracted with boiling, rinsing, solvent recovery, and drying steps after placing 2 g of the pulverized sample in a cylindrical filter paper, adding 140 mL petroleum ether, and connecting it to the machine. The crude fat content value was calculated as crude fat (%) = (weight of the glass and extracted fat—weight of the empty glass) × (100/sample weight).

2.3. Statistical Analysis

The results from the experiments were statistically analyzed using the SAS package (version 9.4 and Enterprise Guide 7.1, SAS Institute, Inc., Cary, NC, USA) and R analysis (version 3.5.3 and R studio). The principal component analysis (PCA) and the cluster analysis were conducted using the sample correlation matrix for the 468 germplasms, in which the 20 seed composition traits, including crude protein, crude fat, fatty acids, lutein, chlorophyll, and anthocyanin, were studied. Lutein, chlorophyll, and anthocyanin data from our previous studies were used [33,59]. However, one accession (IT186154) was excluded from the PCA and cluster analysis due to missing data for several measured seed compositions. In the principal component analyses, the optimal number of principal components was determined by referring to Kaiser's rule [62] and the Scree test [63]. The PCA was conducted using the SAS PRINCOMP procedure in SAS, and the results of the PCA were analyzed using the ggplot2 package in R software.

K-means cluster analysis was utilized for the cluster analysis in this study. R-square (RSQ) and RSQ/(1-RSQ) values were used as indicators to determine the influence of variables on cluster formation [64], and whether the number of clusters was optimal was determined by using the Cubic Clustering Criterion (CCC) and Pseudo F (PSF) values. The CCC is a statistic indicating the bias degree of clusters under the assumption that observation values follow a uniform distribution. The optimal number of clusters can be confirmed through the CCC distribution based on the change in the number of clusters, and the study shows that the larger CCC value ($\geq$3) provides a better clustering result. PSF is a statistic that measures the degree of clusters, and the higher the PSF value indicates the degree of separation between clusters [65–67]. The SAS' FASTCLUS procedure was utilized for the cluster analysis, and the results were shown on the scatterplots through the SAS' CANDISC and GPLOT procedures.

3. Results and Discussion

3.1. Seed Composition Traits of Black Soybean Germplasms with Green Cotyledons

The results of a statistical analysis of the measured seed composition traits in 470 soybean germplasms, including Cheongja 3, Cheongja, and a yellow soybean cultivar, Uram, are shown in Table 1. The range of crude protein content in the 470 accessions was 35.6–45.8%, with an average of 39.7%. A total of 55 germplasms had <38.0% protein content, whereas 387 germplasms had 38.1–42.0% protein content. A total of 27 germplasms had >42.1% protein content. The mean protein content values of the Cheongja 3, Cheongja, and Uram cultivars were 42.5%, 42.7%, and 40.2%, respectively. Furthermore, the crude fat content in the 470 accessions ranged from 15.9% to 22.4%, with an average of 18.6%. Only 2 black soybean germplasms had <16.0% fat content, whereas 124 germplasms had 16.1–18.0% fat content. A total of 315 germplasms had 18.1–20.0% fat content. A total of 28 germplasms had >20.1% fat content. The fat content values of the Cheongja3, Cheongja, and Uram cultivars were 19.5%, 21.8%, and 18.8%, respectively.

Yellow soybeans have been used as a main ingredient in foods such as tofu, soy sauce, and oil due to their seed color. Although the seed compositions of black soybeans are similar to those of yellow soybeans, the availability of black soybeans is limited. The protein content of 179 black soybeans with green cotyledon accessions was higher than that of Uram. In addition, 191 black soybean accessions had higher fat content than Uram. These black soybeans with green cotyledons containing high amounts of protein or fat can serve as useful genetic materials for expanding the availability of black soybeans.

Regarding the fatty acid compositions, the following results were obtained: PA ranged from 9.4% to 15.3%, with an average of 11.3%; SA ranged from 2.4% to 4.9%, with an average of 3.7%; OA ranged from 13.7% to 40.6%, with an average of 23.1%; LA ranged from 40.2% to 60.2%, with an average of 53.7%; and LNA ranged from 5.2% to 12.3%, with an average of 8.2%. In Cheongja 3, Cheongja, and Uram, the respective PA contents were 13.1%, 12.1%, and 11.1%; the respective SA contents were 4.4%, 3.7%, and 4.8%; the respective OA contents were 21.2%, 22.9%, and 24.4%; the respective LA contents were 54.0%, 53.7%, and 52.0%; and the respective LNA contents were 7.4%, 7.6%, and 7.8%. These results were similar to those of previous studies [5,6,16]. Of the 469 black soybean germplasms with green cotyledons, 10 germplasms contained <10.0% PA, 6 contained >30.0% OA, 16 contained <50.0% LA, 60 contained <7.0% LNA, and 9 contained >10.0% LNA. Altered fatty acid profiles in soybean seeds affect the quality of soybean oil and play a significant role in determining its use. In saturated fatty acids (SFAs), PA acts as a health-threatening compound as it increases the cholesterol content, whereas SA has the least effect on increased cholesterol levels [6,16]. However, PA and SA have the advantage of reducing the generation of trans fats because there is no need for hydrogenation during food processing [4,68]. Since elevated OA content increases the oxidative stability of soybean oil, soybean oils with high OA content can be used in various industries. In addition, OA exhibits excellent health benefits, such as lowering cholesterol levels and preventing cardiovascular diseases and cancer; hence, it is highly useful as a functional food [6,7,9,10,16]. Moreover, higher

amounts of LA (ω-6) and LNA (ω-3) lower the oxidative stability of soybean oil, thereby decreasing its quality. Since ω-6 and ω-3 are essential fatty acids, they have high utility as health foods and can reduce cardiovascular diseases and increase anti-inflammatory and antioxidant functions [6,7,10,13,15–17]. Thus, soybean seeds with altered fatty acid profiles can be used in various food-based industries [6,69].

Table 1. Statistical data for seed composition traits of 469 black soybean germplasms with green cotyledons (including check cultivars 'Cheongja 3' and 'Cheongja'), and a yellow soybean check cultivar, 'Uram', over three years (2013–2015).

Trait	Mean	SD	CV	Maximum	Minimum	Median	PCS	Kurtosis	Uram
Crude protein (%)	39.7	1.5	3.8	45.8	35.6	39.5	0.7	1.3	40.2
Crude fat (%)	18.6	1.0	5.3	22.4	15.9	18.5	0.4	0.8	18.8
Lutein (μg/g)	47.6	2.7	5.7	59.7	42.7	47.3	0.9	1.2	41.6
Chlorophyll a (μg/g)	33.9	15.6	46.0	88.4	11.5	28.6	1.2	0.3	0.0
Chlorophyll b (μg/g)	21.9	5.9	27.0	40.8	7.7	23.5	−0.3	−0.6	0.0
Total chlorophyll (μg/g)	55.8	15.4	27.6	120.0	22.6	54.1	0.7	1.3	0.0
C3G (mg/g)	9.7	4.1	41.8	19.8	1.7	11.2	−0.7	−0.9	0.0
D3G (mg/g)	2.5	0.9	36.8	4.8	0.0	2.7	−1.0	0.5	0.0
Pt3G (mg/g)	0.6	0.3	56.3	1.6	0.0	0.6	0.6	1.1	0.0
Pl3G (mg/g)	0.2	0.1	86.5	0.7	0.0	0.2	0.7	0.1	0.0
Pn3G (mg/g)	0.1	0.0	45.6	0.2	0.0	0.1	−0.6	0.3	0.0
M3G (mg/g)	0.0	0.1	116.6	0.3	0.0	0.0	2.4	5.4	0.0
Total anthocyanin (mg/g)	13.0	5.0	38.6	24.3	2.3	15.0	−0.9	−0.6	0.0
PA (%)	11.3	0.8	6.9	15.3	9.4	11.2	1.5	4.4	11.1
SA (%)	3.7	0.4	9.7	4.9	2.4	3.7	0.2	0.8	4.8
OA (%)	23.1	2.7	11.6	40.6	13.7	22.7	1.4	7.3	24.4
LA (%)	53.7	2.0	3.7	60.2	40.2	54.0	−1.3	7.6	52.0
LNA (%)	8.2	1.0	11.8	12.3	5.2	8.3	−0.1	0.4	7.8
SFA (%)	15.0	0.9	6.1	19.5	12.9	14.8	1.5	4.0	15.8
UFA (%)	85.0	0.9	1.1	87.1	80.5	85.2	−1.5	4.0	84.2

SD: standard deviation; CV: coefficient of variation; PCS: Pearson's coefficient of skewness (PCS > 0, skewed to the right = positively skewed distribution; PCS < 0, skewed to the left = negatively skewed distribution; PCS = 0, normal distribution); K: kurtosis (K > 3, narrow curved; K < 3, broad curved; K = 3, normal distribution); Uram: check cultivar with yellow seed coat and cotyledon; crude protein~crude fat: in the whole seed; lutein~total chlorophyll: in cotyledons removed seed coats; the number of black soybean resources is 469 for lutein and chlorophyll, including check cultivars Cheongja and Cheongja 3; C3G~total anthocyanin: in seed coats (C3G: cyanidin-3-glucoside; D3G: delphinidin-3-glucoside; Pt3G: petunidin-3-glucoside; Pl3G: pelargonidin-3-glucoside; Pn3G: peonidin-3-glucoside; M3G: malvidin-3-galactoside); palmitic acid~unsaturated fatty acid: in the whole seed (PA: palmitic acid; SA: stearic acid; OA: oleic acid; LA: linoleic acid; LNA: α-linolenic acid; SFA: saturated fatty acid; UFA: unsaturated fatty acid).

3.2. Correlation Analysis of Germplasms

The results of a correlation analysis of the seed component traits of black soybean germplasms with green cotyledons is presented in Table 2 and is intended to explain the linear associations between two traits. The anthocyanin, chlorophyll, and lutein contents of the black soybean accessions with green cotyledons were obtained from our previous studies [33,59]. The crude protein content showed significantly negative correlations with chlorophyll *b* ($r = -0.56$, $p < 0.001$), C3G ($r = -0.42$, $p < 0.001$), D3G ($r = -0.43$, $p < 0.001$), and total anthocyanin contents ($r = -0.44$, $p < 0.001$). Kim et al. [12] reported a significant negative correlation between the protein and fat contents of black soybeans ($r = -0.22$ $p < 0.05$). Similarly, Tajuddin et al. [70] reported a negative correlation between protein and fat contents. In the present study, there was a weak negative correlation between crude protein and crude fat contents ($r = -0.22$, $p < 0.001$). Crude fat content showed a significant positive correlation with lutein ($r = 0.37$, $p < 0.001$), chlorophyll *a* ($r = 0.35$, $p < 0.001$), total chlorophyll ($r = 0.34$, $p < 0.001$), and OA ($r = 0.32$, $p < 0.001$) contents. However, crude fat content showed a significant negative correlation with LNA content ($r = -0.38$, $p < 0.001$).

Table 2. Simple linear correlation coefficients for 20 seed composition traits of 469 black soybean germplasms with green cotyledons, including check cultivars 'Cheongja 3' and 'Cheongja', over three years (2013–2015).

Trait	CP	CF	Lut	Chl *a*	Chl *b*	Total Chl	C3G	D3G	Pt3G	Pl3G	Pn3G	M3G	Total Ant	PA	SA	OA	LA	LNA	SFA	UFA
CP	1																			
CF	−0.21 ***	1																		
Lut	−0.11 *	0.37 ***	1																	
Chl *a*	0.14 **	0.35 ***	0.82 ***	1																
Chl *b*	−0.56 ***	−0.04 ns	0.10 *	−0.22 ***	1															
Total Chl	−0.07 ns	0.34 ***	0.86 ***	0.93 ***	0.16 ***															
C3G	−0.42 ***	−0.21 ***	−0.54 ***	−0.74 ***	0.51 ***	−0.56 ***	1													
D3G	−0.43 ***	−0.13 **	−0.31 ***	−0.47 ***	0.61 ***	−0.24 ***	0.63 ***	1												
Pt3G	−0.11 *	−0.17 ***	−0.52 ***	−0.69 ***	0.25 ***	−0.60 ***	0.61 ***	0.65 ***	1											
Pl3G	−0.42 ***	−0.17 ***	−0.34 ***	−0.48 ***	0.46 ***	−0.31 ***	0.65 ***	0.43 ***	0.27 ***	1										
Pn3G	−0.22 ***	−0.08 ns	−0.46 ***	−0.62 ***	0.43 ***	−0.46 ***	0.71 ***	0.56 ***	0.71 ***	0.38 ***	1									
M3G	0.11 *	−0.10 *	−0.44 ***	−0.55 ***	−0.03 ns	−0.56 ***	0.37 ***	0.28 ***	0.89 ***	0.03 ns	0.58 ***	1								
Total Ant	−0.44 ***	−0.29 ***	−0.54 ***	−0.75 ***	0.56 ***	−0.55 ***	0.99 ***	0.75 ***	0.69 ***	0.65 ***	0.74 ***	0.42 ***	1							
PA	0.12 *	0.07 ns	0.07 ns	0.13 **	−0.08 ns	0.10 *	−0.17 ***	−0.10 *	−0.10 *	−0.07 ns	−0.12 *	−0.08 ns	−0.16 ***	1						
SA	0.22 ***	−0.08 ns	−0.06 ns	0.01 ns	−0.05 ns	−0.01 ns	−0.10 *	0.02 ns	0.07 ns	−0.20 ***	−0.06 ns	0.11 *	−0.08 ns	0.16 ***	1					
OA	0.15 **	0.32 ***	0.38 ***	0.47 ***	−0.43 ***	0.31 ***	−0.42 ***	−0.43 ***	−0.26 ***	−0.42 ***	−0.31 ***	−0.05 ns	−0.45 ***	−0.21 ***	−0.09 ns	1				
LA	−0.14 **	−0.26 ***	−0.33 ***	−0.41 ***	0.39 ***	−0.27 ***	0.36 ***	0.41 ***	0.25 ***	0.37 ***	0.29 ***	0.07 ns	0.39 ***	−0.06 ns	−0.02 ns	−0.92 ***	1			
LNA	−0.29 ***	−0.38 ***	−0.40 ***	−0.56 ***	0.48 ***	−0.38 ***	0.58 ***	0.43 ***	0.26 ***	0.55 ***	0.39 ***	0.02 ns	0.58 ***	−0.15 **	−0.22 ***	−0.69 ***	0.54 ***	1		
SFA	0.18 ***	0.03 ns	0.04 ns	0.12 *	−0.09 ns	0.09 ns	−0.18 ***	−0.08 ns	−0.06 ns	−0.14 **	−0.13 **	−0.02 ns	−0.17 ***	0.92 ***	0.53 ***	−0.21 ***	−0.07 ns	−0.21 ***	1	
UFA	−0.18 ***	−0.03 ns	−0.04 ns	−0.12 *	0.09 ns	−0.09 ns	0.18 ***	0.08 ns	0.06 ns	0.15 **	0.13 **	0.02 ns	0.17 ***	−0.92 ***	−0.53 ***	0.21 ***	0.07 ns	0.21 ***	−1.00 ***	1

CP: crude protein; CF: crude fat; Lut: lutein; Chl a: chlorophyll a; Chl b: chlorophyll b; Total Chl: total chlorophyll; C3G: cyanidin-3-glucoside; D3G: delphinidin-3-glucoside; Pt3G: petunidin-3-glucoside; Pl3G: pelargonidin-3-glucoside; Pn3G: peonidin-3-glucoside; M3G: malvidin-3-galactoside; Total Ant: total anthocyanin; PA: palmitic acid; SA: stearic acid; OA: oleic acid; LA: linoleic acid; LNA: α-linolenic acid; SFA: saturated fatty acid; UFA: unsaturated fatty acid. *** Significant at the 0.001 probability level. ** Significant at the 0.01 probability level. * Significant at the 0.05 probability level. NS: non-significance.

The correlation between lutein and chlorophyll demonstrates that lutein showed a highly significant positive correlation with chlorophyll *a* (r = 0.82, $p < 0.001$) and total chlorophyll content (r = 0.86, $p < 0.001$). Monma et al. [49] also reported that there was a correlation between lutein content and total chlorophyll content. The significant correlation between lutein and chlorophyll is a result of their biosyntheses and metabolisms being closely linked [50]. However, despite this close correlation, chlorophyll *b* seems to have a weak positive correlation with lutein and a negative correlation with chlorophyll *a*. However, the scatterplots below show that there were three divided groups (arrows) showing positive correlations (Figure 1). These results could be due to the involvement of genetic traits in green cotyledons, which are divided into germplasms with nuclear inheritance (*d1d2*) and cytoplasmic inheritance (*cytG*) [59]. Studies have reported that black soybeans with nuclear inheritance (*d1d2*) show higher total chlorophyll content than soybeans with cytoplasmic inheritance (*cytG*), and that the values of chlorophyll *a*/*b* are higher in the composition ratio [32,45,59].

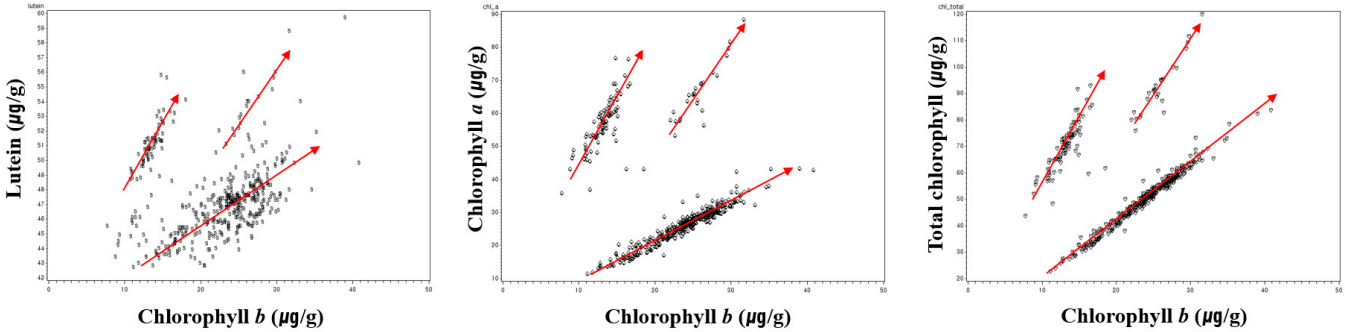


Figure 1. Scatterplots showing positive correlations between chlorophyll *b* and lutein, chlorophyll *b* and chlorophyll *a*, and chlorophyll *b* and total chlorophyll in 469 black soybean germplasms with green cotyledons, including check cultivars 'Cheongja 3' and 'Cheongja', over three years (2013–2015).

The total anthocyanin content, including C3G, showed significantly negative correlations with lutein (r = –0.54, $p < 0.001$), chlorophyll *a* ($r = -0.75$, $p < 0.001$), and total chlorophyll ($r = -0.55$, $p < 0.001$) contents and significantly positive correlations with chlorophyll *b* content ($r = 0.56$, $p < 0.001$). Furthermore, according to previous studies, anthocyanin and chlorophyll are closely linked in the biosynthesis stage in poinsettia leaves and apple peels [71–73].

In addition, lutein and chlorophyll had no significant correlation with PA and SA, which are SFAs. However, OA, which is a monounsaturated fatty acid, showed a significant positive correlation with lutein ($r = 0.38$, $p < 0.001$), chlorophyll *a* ($r = 0.47$, $p < 0.001$), and total chlorophyll content ($r = 0.31$, $p < 0.001$), but it showed a significant negative correlation with chlorophyll *b* ($r = -0.43$, $p < 0.001$). Conversely, LA and LNA showed a significant negative correlation with lutein ($r = -0.33$, $r = -0.40$, $p < 0.001$), chlorophyll *a* ($r = -0.41$, $r = -0.56$, $p < 0.001$), and total chlorophyll content ($r = -0.27$, $r = -0.38$, $p < 0.001$). However, LA and LNA showed a significant positive correlation with chlorophyll *b* ($r = 0.39$, $r = 0.48$, $p < 0.001$). These results are supported by other studies on the correlations between lutein and fatty acid compositions [74]. During the embryogenesis of "green seeds" such as soybeans, Arabidopsis, and rapeseed seeds, the genes involved in photosynthesis and fatty acid synthesis pathways are very tightly linked, resulting in controlled chlorophyll and fatty acid synthesis [75,76].

In the correlation analysis of anthocyanin components, C3G showed a highly significant positive correlation with D3G ($r = 0.63$, $p < 0.001$), Pt3G ($r = 0.61$, $p < 0.001$), and Pn3G ($r = 0.71$, $p < 0.001$). Furthermore, D3G showed a highly significant positive correlation with C3G ($r = 0.63$, $p < 0.001$), Pt3G ($r = 0.65$, $p < 0.001$), and Pn3G ($r = 0.56$, $p < 0.001$). In addition, Pt3G showed a highly significant positive correlation with Pn3G ($r = 0.71$, $p < 0.001$) and M3G ($r = 0.89$, $p < 0.001$). These results support the hypothesis that the anthocyanin pigments should show a highly significant positive correlation with each other [77]. However, Pl3G, Pn3G, and M3G showed statistically weak correlations with the components due to the small amounts of anthocyanin components they contain. In addition, since C3G has the highest content, the correlation analysis between total anthocyanin content and other component traits showed a trend similar to that of C3G.

The correlations between fatty acid compositions indicated that SFA and unsaturated fatty acid (UFA) showed a highly significant negative correlation ($r = -1.00$, $p < 0.001$). However, SFA (PA and SA) showed a weak negative correlation with OA and LA. OA showed a highly significant negative correlation with LA ($r = -0.92$, $p < 0.001$) and LNA ($r = -0.69$, $p < 0.001$). In addition, a highly significant positive correlation was shown between LA and LNA ($r = 0.54$, $p < 0.001$). These results are consistent with those of previous studies [5,78,79].

3.3. Principal Component Analysis

A PCA was conducted on 469 germplasms with 20 seed composition traits using a sample correlation matrix (Table 3). The anthocyanin, chlorophyll, and lutein contents in black soybean accessions with green cotyledons were obtained from our previous PCA study (Jo et al. 2021; Lee et al. 2021). Furthermore, the four principal components were the most important components in determining variations in the 469 soybean germplasms based on the criterion of Kaiser [62] and the Scree plot. The four components accounted for 76.2% of the total variations. The first principal component had an eigenvalue of 7.5477 and a 37.7% rate of contribution to the total variation. The second principal component had an eigenvalue of 3.2648 and a contribution rate of 16.3%. The third principal component had an eigenvalue of 2.6188 and a contribution rate of 13.0%. The fourth principal component had an eigenvalue of 1.8189 and a contribution rate of 9.0% (Table 3).

Table 3. Eigenvalues and proportions of principal components to 20 quantitative traits of 468 black soybean germplasms with green cotyledons, including check cultivars 'Cheongja 3' and 'Cheongja'.

Principal Component	Eigenvalue	Difference	Proportion	Cumulative
1	7.5477	4.2829	0.3774	0.3774
2	3.2648	0.6460	0.1632	0.5406
3	2.6188	0.7999	0.1309	0.6716
4	1.8189	0.7363	0.0909	0.7625
5	1.0826	0.2485	0.0541	0.8166
6	0.8341	0.1371	0.0417	0.8583
7	0.6970	0.2301	0.0349	0.8932
8	0.4669	0.0507	0.0233	0.9165
9	0.4163	0.0381	0.0208	0.9374
10	0.3782	0.0856	0.0189	0.9563
11	0.2925	0.0670	0.0146	0.9709
12	0.2255	0.0152	0.0113	0.9822
13	0.2103	0.0815	0.0105	0.9927
14	0.1288	0.1136	0.0064	0.9991
15	0.0152	0.0130	0.0008	0.9999
16	0.0021	0.0019	0.0001	1.0000
17	0.0002	0.0002	0.0000	1.0000
18	0.0001	0.0000	0.0000	1.0000
19	0.0000	0.0000	0.0000	1.0000
20	0.0000	0.0000	0.0000	1.0000

Table 4 shows the correlations between these 4 principal components and the 20 component traits. The first principal component showed a positive correlation with chlorophyll *b*, total anthocyanin, C3G, LNA, LA, and UFAs, whereas it showed a negative correlation with crude protein, crude fat, lutein, chlorophyll *a*, total chlorophyll, PA, SA, OA, and SFA. In particular, total anthocyanin (0.337) and C3G (0.325) were the phenotypic measurements with the biggest positive contributions to the first principal component, whereas chlorophyll *a* (−0.319) contributed negatively to the first principal component. The second principal component showed a highly positive correlation with the total SFAs (0.516), PA (0.452), and SA (0.325) (Table 4). The second principal component showed a highly negative correlation with the total unsaturated fatty acids (−0.516). Chlorophyll *b*, total chlorophyll content, and crude protein were correlated with the third principal component. The fourth principal component showed a highly positive correlation with crude fat and OA, and a negative correlation with LA and LNA.

In this study, the first and second principal components explained 54.0% of the black soybean germplasms with green cotyledons. Figure 2 presents a scatter plot showing the first principal component against the second principal component. On the upper side of the horizontal axis, black soybean germplasms with high total UFA are present, whereas on the lower side, black soybean germplasms with high total SFA are present. The results indicate that C3G and total anthocyanin positively contributed to the PCA plot consisting

of principal components 1 and 2, whereas lutein, chlorophyll *a*, total chlorophyll, and OA contributed negatively.

Table 4. The 4 principal components among 20 seed composition traits of 468 black soybean germplasms with green cotyledons, including check cultivars 'Cheongja 3' and 'Cheongja'.

Trait	Principal Component (Eigenvectors)			
	1	2	3	4
Crude protein	−0.127	0.176	−0.335	−0.267
Crude fat	−0.122	−0.081	0.086	0.413
Lutein	−0.248	−0.141	0.273	0.202
Chlorophyll *a*	−0.319	−0.085	0.181	0.083
Chlorophyll *b*	0.185	−0.086	0.413	0.179
Total chlorophyll	−0.252	−0.119	0.341	0.152
C3G	0.325	−0.028	0.047	0.119
D3G	0.263	0.003	0.166	0.219
Pt3G	0.274	0.096	−0.211	0.287
Pl3G	0.238	−0.062	0.206	−0.052
Pn3G	0.279	0.024	−0.057	0.257
M3G	0.183	0.122	−0.362	0.285
Total anthocyanin	0.337	−0.016	0.057	0.156
PA (16:0)	−0.077	0.452	0.185	0.067
SA (18:0)	−0.034	0.325	−0.032	0.075
OA (18:1)	−0.220	−0.192	−0.275	0.330
LA (18:2)	0.208	0.052	0.212	−0.338
LNA (18:3)	0.255	−0.059	0.189	−0.298
SFA	−0.080	0.516	0.145	0.089
UFA	0.080	−0.516	−0.146	−0.089

C3G: cyanidin-3-glucoside; D3G: delphinidin-3-glucoside; Pt3G: petunidin-3-glucoside; Pl3G: pelargonidin-3-glucoside; Pn3G: peonidin-3-glucoside; M3G: malvidin-3-galactoside; PA: palmitic acid; SA: stearic acid; OA: oleic acid; LA: linoleic acid; LNA: α-linolenic acid; SFA: saturated fatty acid; UFA: unsaturated fatty acid. The underlined values are the traits that showed a higher correlation with the principal component.

3.4. Cluster Analysis with Seed Composition Traits

The anthocyanin, chlorophyll, and lutein contents in the black soybean accessions with green cotyledons were obtained from our previous cluster analysis study [33,59]. A K-means cluster analysis was performed on 469 black soybean germplasms with green cotyledons to classify them into 4 clusters. In addition, this study used CCC and PSF values to determine the appropriate number of clusters. Furthermore, based on the CCC and PSF values for the population (*K*) numbers from 2 to 6, the highest CCC value was 26 and the highest PSF value was 655 when *K* was 4, indicating that the 469 black soybean germplasms in this study were divided into 4 clusters (Table 5). In addition, a canonical discriminant analysis with *K* = 4 revealed that cluster 1 is the smallest group, consisting of 10 germplasms, which is 2.1% of the total number of germplasms. Cluster 2 consisted of 108 germplasms, which is 23.1% of the total number of germplasms. Cluster 3 is the largest group, containing 254 germplasms, which is 54.3% of the total number of germplasms. Cluster 4 consisted of 96 germplasms, which is 20.5% of the total number of germplasms (Figure 3, Table S1). Our previous study revealed that there were four clusters of black soybean accessions with green cotyledons based on their agronomic traits [33]. However, genotypic data from 470 soybean accessions with 6 K single nucleotide polymorphic loci determined 3 clusters based on PCA and phylogenetic tree analyses [59].

Chlorophyll *a*, total chlorophyll, and total anthocyanin had relatively higher R-square (RSQ) and RSQ/(1-RSQ) values than other traits (Table 5). The results of the present study revealed that chlorophyll *a*, total chlorophyll, and total anthocyanin greatly contributed to the classification of the clusters. Furthermore, chlorophyll *a* showed the highest RSQ/(1-RSQ) value (9.7482), suggesting that it had the greatest influence on cluster classification in these black soybean germplasms. The black soybean germplasms in cluster 1 had higher

crude fat, lutein, chlorophyll *a*, chlorophyll *b*, and total chlorophyll contents than the other clusters, whereas the C3G content in cluster 1 was significantly lower than that in clusters 2 and 3 (Figure 4). Clusters 2 and 3 were located closest to each other (Figure 4), indicating that the black soybean germplasms included in these two clusters showed a similar pattern of measured seed compositions in this study. The black soybean germplasms in clusters 2 and 3 had the highest anthocyanin (e.g., C3G) content, whereas they had lower lutein, chlorophyll *b*, and total chlorophyll contents than the other clusters (Figure 4). Figure 5 shows fatty acid profiles in four clusters of distribution. Cluster 1 had significantly higher PA content than clusters 2 and 3. However, there was no significant difference in SA content among the clusters. Clusters 2 and 3 had relatively high LNA content. Cluster 4 had the highest OA content and the lowest LA and LNA contents.

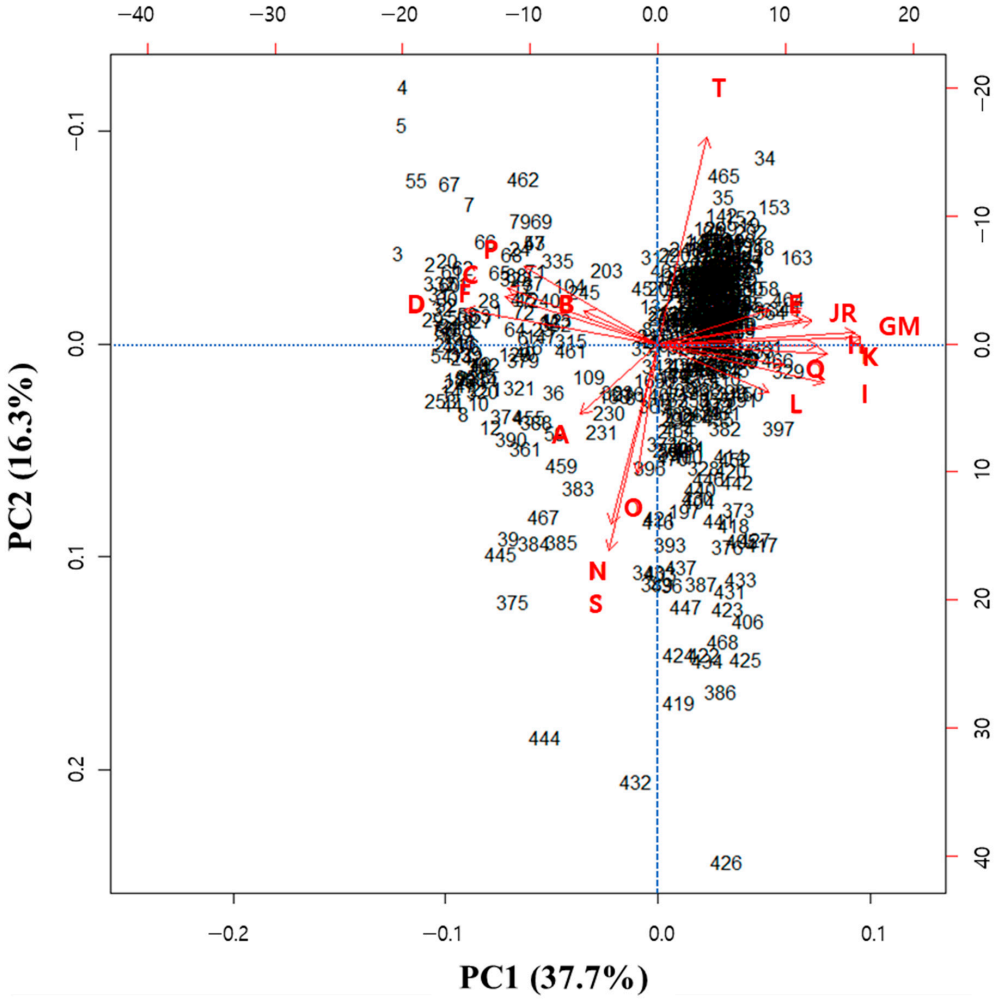


Figure 2. Scatterplot of principal component 1 (PC1: 37.7%) and principal component 2 (PC2: 16.3%) based on 20 seed composition traits of 468 black soybean germplasms with green cotyledons, including check cultivars 'Cheongja 3' and 'Cheongja'. A: crude protein; B: crude fat; C: lutein; D: chlorophyll a; E: chlorophyll b; F: total chlorophyll; G: C3G (cyanidin-3-glucoside); H: D3G (delphinidin-3-glucoside); I: Pt3G (petunidin-3-glucoside); J: Pl3G (pelargonidin-3-glucoside); K: Pn3G (peonidin-3-glucoside); L: M3G (malvidin-3-galactoside); M: total anthocyanin; N: palmitic acid; O: stearic acid; P: oleic acid; Q: linoleic acid; R: α-linolenic acid; S: saturated acid; T: unsaturated acid.

Table 5. Statistical variables for clusters of 20 seed composition traits of 468 black soybean germplasms with green cotyledons, including check cultivars 'Cheongja 3' and 'Cheongja'.

Variable	Total SD	within SD	R-Square	RSQ/(1-RSQ)
Crude protein	1.5157	1.3176	0.2492	0.3319
Crude fat	0.9748	0.9065	0.1408	0.1638
Lutein	2.7305	1.6903	0.6193	1.6265
Chlorophyll *a*	15.6020	4.7743	0.9070n	9.7482
Chlorophyll *b*	5.9057	4.1003	0.5210	1.0879
Total chlorophyll	15.4317	6.4765	0.8250	4.7141
C3G	4.0595	2.4189	0.6472	1.8348
D3G	0.9037	0.6641	0.4635	0.8638
Pt3G	0.3145	0.2241	0.4957	0.9831
Pl3G	0.1413	0.1112	0.3847	0.6251
Pn3G	0.0292	0.0213	0.4718	0.8933
M3G	0.0459	0.0363	0.3780	0.6077
Total anthocyanin	5.0218	2.8425	0.6817	2.1413
PA (16:0)	0.7834	0.7757	0.0259	0.0266
SA (18:0)	0.3627	0.3632	0.0037	0.0037
OA (18:1)	2.6856	2.2171	0.3228	0.4767
LA (18:2)	1.9999	1.7277	0.2585	0.3486
LNA (18:3)	0.9684	0.7336	0.4299	0.7540
SFA	0.9111	0.9037	0.0224	0.0229
UFA	0.9116	0.9043	0.0222	0.0227
Overall	5.4053	2.3687	0.8092	4.2412
Pseudo F Statistic = 655.97				
Cubic Clustering Criterion (CCC) = 26.025				

Total SD: standard deviation for the total sample; Within SD: pooled within cluster standard deviation; R-Square: for predicting the variable from the cluster; RSQ/(1-RSQ): the ratio of between-cluster variance to within-cluster variance ($R^2/(1-R^2)$); Pseudo F Statistic: estimating the number of clusters; Cubic Clustering Criterion (CCC): computed under the assumption that the variables are uncorrelated; C3G: cyanidin-3-glucoside; D3G: delphinidin-3-glucoside; Pt3G: petunidin-3-glucoside; Pl3G: pelargonidin-3-glucoside; Pn3G: peonidin-3-glucoside; M3G: malvidin-3-galactoside; PA: palmitic acid; SA: stearic acid; OA: oleic acid; LA: linoleic acid; LNA: α-linolenic acid; SFA: saturated fatty acid; UFA: unsaturated fatty acid.

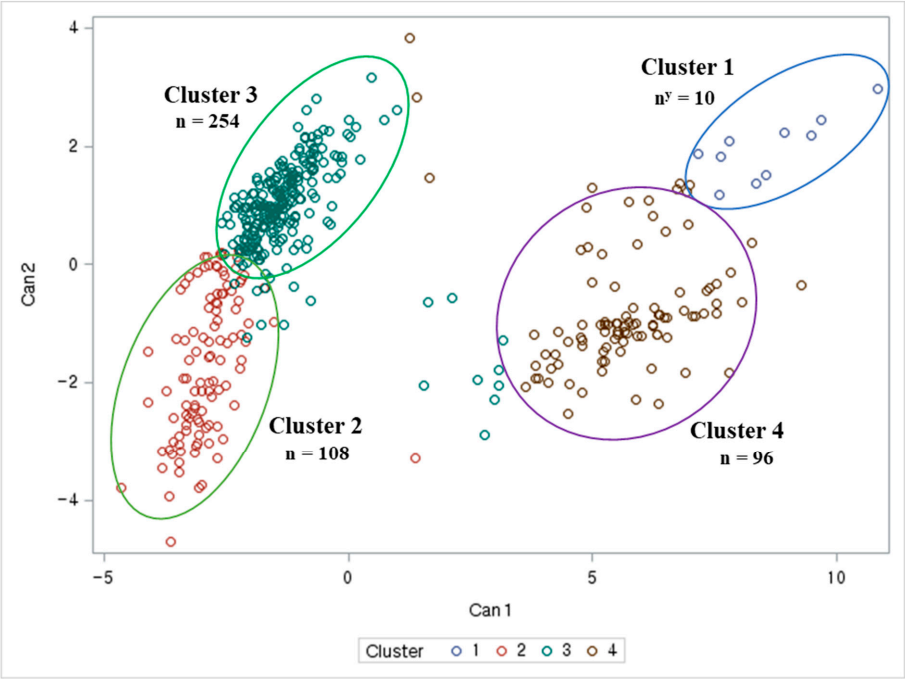


Figure 3. The 4 clusters based on 5 principal components of 468 black soybean germplasms with green cotyledons, including check cultivars 'Cheongja 3' and 'Cheongja'. Can: canonical variables; n: number of resources in each cluster.

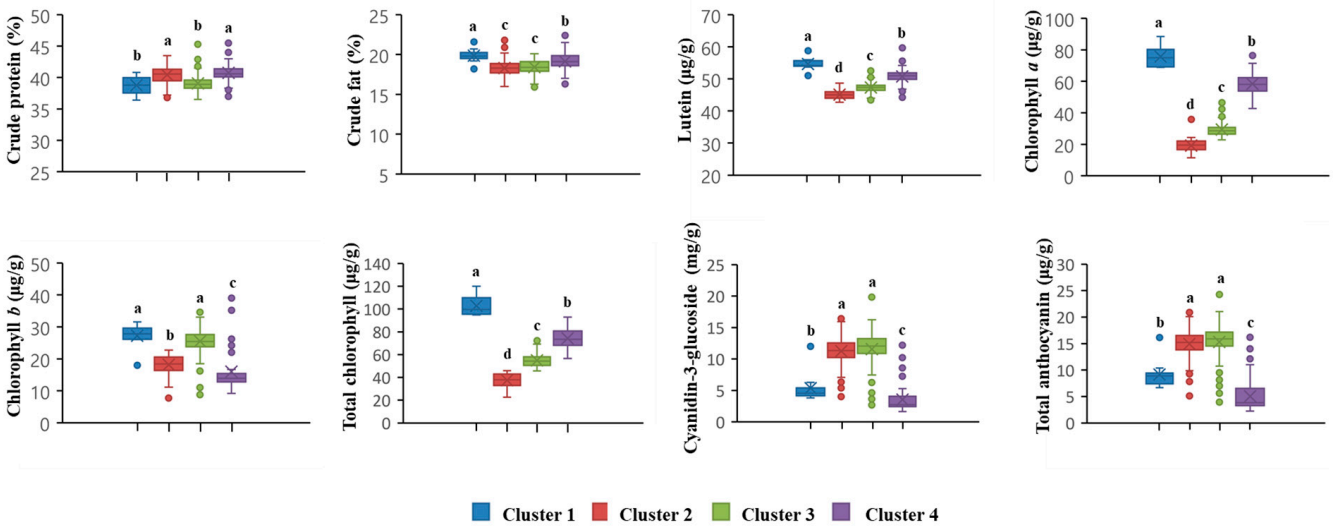


Figure 4. Box plots of crude protein, crude fat, lutein, chlorophylls, and anthocyanins in 4 clusters of 468 black soybean germplasms with green cotyledons, including check cultivars 'Cheongja' and 'Cheongja 3'. The letters above the bars are statistically different based on the least square difference.

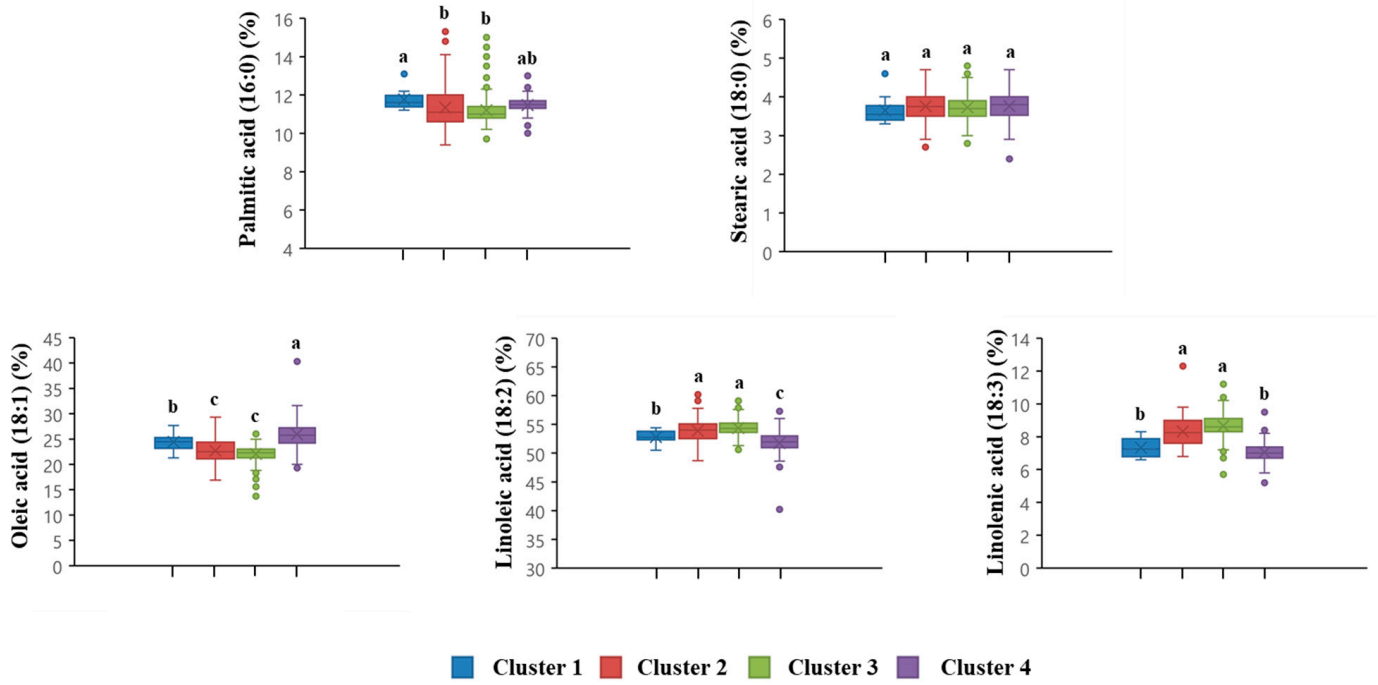


Figure 5. Box plots of fatty acids in 4 clusters of 468 black soybean germplasms with green cotyledons, including check cultivars 'Cheongja 3' and 'Cheongja'. The letters above the bars are statistically different based on the least square difference.

Anthocyanin, lutein, and chlorophyll are major antioxidants in black soybean germplasms with green cotyledons [20–30]. Furthermore, OA prevents cardiovascular diseases, and LNA is an ω-3 fatty acid directly related to brain health [5,6]. Black soybeans with high levels of anthocyanin, lutein, chlorophyll, OA, and LNA are highly utilized as raw materials for preparing healthy foods [6,7,9,10,13,15,21,30,51–55,80]. Protein in soybean seeds is one of the most important sources of vegetable protein. From a nutritional perspective, a combination of higher protein and anthocyanins can provide novel functional foods with desirable health benefits. However, we did not discover a black soybean germplasm with high protein and anthocyanin content in this study, so this could be a breeding goal. There has been a change in the consumer perception of black soybeans with green cotyledons.

Therefore, the development of new black soybean cultivars with green cotyledons and with altered seed compositions is warranted. This study provides useful information for improving seed composition in soybean breeding programs.

4. Conclusions

Black soybeans with green cotyledons contain anthocyanin, chlorophyll, and lutein, which reduce the risks of cardiovascular diseases and increase anticancer and antioxidant effects [21,51–55]. Therefore, it is essential to secure black soybean germplasms with green cotyledons by evaluating various nutritional and bioactive substances. This study analyzed the genetic diversity of 469 black soybean germplasms with green cotyledons and with 20 seed composition traits. Twenty seed composition traits were analyzed to conduct correlation analysis, principal component analysis (PCA), and cluster analysis, which indicated that black soybean germplasms were divided into four clusters. Black soybean germplasms in cluster 1 had higher crude fat, lutein, chlorophyll *a*, chlorophyll *b*, and total chlorophyll contents, but lower C3G content than those in clusters 2 and 3. However, germplasms in clusters 2 and 3 had the highest C3G content. Moreover, germplasms in cluster 1 had significantly higher PA content than those in clusters 2 and 3. Germplasms in clusters 2 and 3 had relatively high LNA content. Germplasms in cluster 4 had the highest OA content. This study highlights the genetic diversity of black soybean germplasms with different seed composition traits, and the results of this study can be beneficial for soybean breeding programs, enabling them to develop new black soybean cultivars with green cotyledons and improved seed composition traits.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/agriculture13020406/s1, Figure S1: Calibration curves of anthocyain (D3G, C3G, Pt3G, Pl3G, Pn3G, M3G) for the anaylsis of black soybean germplasms with green cotyledon. D3G: delphinidin-3-glucoside, C3G: cyanidin-3-glucoside, Pt3G: petunidin-3-glucoside, Pl3G: pelargonidin-3-glucoside, Pn3G: peonidin-3-glucoside, M3G: malvidin-3-galactoside. Table S1: Four clusters based on 20 seed composition traits of 468 black soybean accessions including two checks (Cheongja 3, and Cheongja) by K-means cluster analysis.

Author Contributions: Conceptualization, J.-D.L.; methodology, J.Y.L., C.K.S. and J.S.B.; formal analysis, J.Y.L. and H.J.; writing—original draft preparation, J.Y.L. and H.J.; writing—review and editing, H.J., C.K.S., J.S.B. and J.-D.L.; supervision, J.-D.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was carried out with the support of the "Cooperative Research Program for Agriculture Science and Technology Development (Project No. PJ01416803)", Rural Development Administration, Republic of Korea.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: The datasets generated during this study are available from the corresponding author on reasonable request.

Acknowledgments: The authors would like to acknowledge the personnel from the Plant Genetics and Breeding Lab at the Kyungpook National University and the Gyeongsangbuk-do Provincial Agricultural Research and Extension Service for their time and work on the field experiments.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lee, C.Y.; Choi, M.S.; Kim, H.T.; Yun, H.T.; Lee, B.; Chung, Y.S.; Kim, R.W.; Choi, H.K. Soybean [*Glycine max* (L.) Merrill]: Importance as a crop and pedigree reconstruction of Korean varieties. *Plant Breed. Biotech.* **2015**, *3*, 179–196. [CrossRef]
2. Sanjukta, S.; Rai, A.K. Production of bioactive peptides during soybean fermentation and their potential health benefits. *Trends Food Sci. Tech.* **2016**, *50*, 1–10. [CrossRef]
3. Goldsmith, P.D. Economics of soybean production, marketing, and utilization. In *Soybeans: Chemistry, Production, and Utilization*; AOCS Press: Urbana, IL, USA, 2015; pp. 117–150. [CrossRef]
4. Fehr, W.R. Breeding for modified fatty acid composition in soybean. *Crop Sci.* **2007**, *47*, 72–87. [CrossRef]

5. Choung, M.G. Variation of oil contents and fatty acid compositions in Korean soybean germplasms. *Korean J. Crop Sci.* **2006**, *51*, 139–145.
6. Lee, J.D.; Bilyeu, K.D.; Shannon, J.G. Genetics and breeding for modified fatty acid profile in soybean seed oil. *J. Crop Sci. Biotech.* **2007**, *10*, 201–210.
7. Asif, M. Health effects of omega-3,6,9 fatty acids: *Perilla frutescens* is a good example of plant oils. *Orient Pharm. Exp. Med.* **2011**, *11*, 51–59. [CrossRef]
8. Cahoon, E.B. Genetic enhancement of soybean oil for industrial uses: Prospects and challenges. *AgBioForum* **2003**, *6*, 11–13.
9. Chang, N.W.; Huang, P.C. Effects of the ratio of polyunsaturated and monounsaturated fatty acid to saturated fatty acid on rat plasma and liver lipid concentrations. *Lipids* **1998**, *33*, 481–487. [CrossRef]
10. De Jong, A.J.; Kloppenburg, M.; Toes, R.E.; Ioan-Facsinay, A. Fatty acids, lipid mediators, and T-cell function. *Front. Immunol.* **2014**, *5*, 483. [CrossRef]
11. Gerde, J.A.; White, P.J. Lipid. In *Soybeans: Chemistry, Production, and Utilization*; Johnson, L.A., White, P.J., Galloway, R., Eds.; AOCS Press: Urbana, IL, USA, 2015; pp. 193–227. [CrossRef]
12. Kim, S.L.; Lee, Y.H.; Chi, H.Y.; Lee, S.J.; Kim, S.J. Diversity in lipid contents and fatty acid composition of soybean seeds cultivated in Korea. *Korean J. Crop Sci.* **2007**, *52*, 348–357.
13. Kris-Etherton, P.M.; Harris, W.S.; Appel, L.J. Fish consumption, fish oil, omega-3 fatty acids, and cardiovascular disease. *Arterioscler. Thromb. Vas.* **2003**, *23*, 20–30. [CrossRef]
14. Neff, W.E. Oxidative stability of natural and randomized high palmitic-and high-stearic-acid oils from genetically modified soybean varieties. *J. Am. Oil Chem. Soc.* **1999**, *76*, 825–831. [CrossRef]
15. Wall, R.; Ross, R.P.; Fitzgerald, G.F.; Stanton, C. Fatty acids from fish: The anti-inflammatory potential of long-chain omega-3 fatty acids. *Nutr. Rev.* **2010**, *68*, 280–289. [CrossRef]
16. Wilson, R.F. Seed composition. In *Soybeans: Improvement, Production and Uses*, 3rd ed.; Boerma, H.R., Specht, J.E., Eds.; ASA, CSSA, SSSA: Madison, WI, USA, 2004; pp. 621–629. [CrossRef]
17. Zhu, X.; Wang, B.; Zhang, X.; Chen, X.; Zhu, J.; Zou, Y.; Li, J. Alpha-linolenic acid protects against lipopolysaccharide-induced acute lung injury through anti-inflammatory and anti-oxidative pathways. *Microb. Pathogenesis* **2020**, *142*, 104077. [CrossRef]
18. Barnes, S. Evolution of the health benefits of soy isoflavone. *Proc. Soc. Exp. Biol. Med.* **1998**, *217*, 386–392. [CrossRef]
19. Tomomatsu, H. Health effects of oligosaccharides. *Food Technol.* **1994**, *48*, 61–65.
20. Astadi, I.R.; Astuti, M.; Santoso, U.; Nugraheni, P.S. In vitro antioxidant activity of anthocyanins of black soybean seed coat in human low density lipoprotein (LDL). *Food Chem.* **2009**, *112*, 659–663. [CrossRef]
21. Cho, K.M.; Ha, T.J.; Lee, Y.B.; Seo, W.D.; Kim, J.Y.; Ryu, H.W.; Jeong, S.H.; Kang, Y.M.; Lee, J.H. Soluble phenolics and antioxidant properties of soybean (*Glycine max* L.) cultivars with varying seed coat colors. *J. Funct. Foods* **2013**, *5*, 1065–1076. [CrossRef]
22. Choung, M.G.; Baek, I.Y.; Kang, S.T.; Han, W.Y.; Shin, D.C.; Moon, H.P.; Kang, K.H. Isolation and determination of anthocyanins in seed coats of black soybean (*Glycine max* (L.) Merr.). *J. Agric. Food Chem.* **2001**, *49*, 5848–5851. [CrossRef]
23. Correa, C.R.; Li, L.; Aldini, G.; Carini, M.; Chen, C.Y.O.; Chun, H.K.; Cho, S.M.; Park, K.M.; Russell, R.M.; Blumberg, J.B.; et al. Composition and stability of phytochemicals in five varieties of black soybeans (*Glycine max*). *Food Chem.* **2010**, *123*, 1176–1184. [CrossRef]
24. Dajanta, K.; Janpum, P.; Leksing, W. Antioxidant capacities, total phenolics and flavonoids in black and yellow soybeans fermented by *Bacillus subtilis*. A comparative study of Thai fermented soybeans (*thuanao*). *Int. Food Res. J.* **2013**, *20*, 3125–3132.
25. Lee, J.H.; Kang, N.S.; Shin, S.O.; Shin, S.H.; Lim, S.G.; Suh, D.Y.; Baek, I.Y.; Park, K.Y.; Ha, T.J. Characterisation of anthocyanins in the black soybean (*Glycine max* L.) by HPLC-DAD-ESI/MS analysis. *Food Chem.* **2009**, *112*, 226–231. [CrossRef]
26. Lee, K.J.; Lee, J.R.; Ma, K.H.; Cho, Y.H.; Lee, G.A.; Chung, J.W. Anthocyanin and isoflavone contents in Korean black soybean landraces and their antioxidant activities. *Plant Breed. Biotech.* **2016**, *4*, 441–452. [CrossRef]
27. Lee, K.J.; Baek, D.Y.; Lee, G.A.; Cho, G.T.; So, Y.S.; Lee, J.R.; Ma, K.H.; Chung, J.W.; Hyun, D.Y. Phytochemicals and antioxidant activity of Korean black soybean (*Glycine max* L.) landraces. *Antioxidants* **2020**, *9*, 213. [CrossRef]
28. Todd, J.J.; Vodkin, L.O. Pigmented soybean (*Glycine max*) seed coats accumulate proanthocyanins during development. *Plant Physiol.* **1993**, *102*, 663–670. [CrossRef]
29. Xu, B.; Chang, S.K.C. Antioxidant capacity of seed coat, dehulled bean, and whole black soybeans in relation to their distributions of total phenolics, phenolic acids, anthocyanins, and isoflavones. *J. Agric. Food Chem.* **2008**, *56*, 8365–8373. [CrossRef]
30. Yamashita, Y.; Wang, L.; Nakamura, A.; Nanba, F.; Saito, S.; Toda, T.; Nakagawa, J.; Ashida, H. Black soybean improves the vascular function through an increase in nitric oxide and a decrease in oxidative stress in healthy women. *Arch. Biochem. Biophys.* **2020**, *688*, 108408. [CrossRef]
31. Choi, Y.M.; Yoon, H.; Lee, S.; Ko, H.C.; Shin, M.J.; Lee, M.C.; Hur, O.S.; Ro, N.Y.; Desta, K.T. Isoflavones, anthocyanins, phenolic content, and antioxidant activities of black soybeans (*Glycine max* (L.) Merrill) as affected by seed weight. *Sci. Rep.* **2020**, *10*, 1–13. [CrossRef]
32. Kang, S.T.; Seo, M.J.; Moon, J.K.; Yun, H.T.; Lee, Y.H.; Kim, S.J.; Hwang, Y.S.; Lee, S.K.; Choung, M.G. Introduction of stay green mutant for the development of black seed coat and green cotyledon soybean variety. *Korean J. Crop Sci.* **2010**, *55*, 187–194.
33. Lee, J.Y.; Choi, H.J.; Son, C.K.; Bae, J.S.; Jo, H.; Lee, J.D. Genetic diversity of black soybean germplasms with green cotyledons based on agronomic traits and cotyledon pigments. *Korean J. Breed. Sci.* **2021**, *53*, 127–139. [CrossRef]

34. Sinnecker, P.; Gomes, M.S.O.; Arêas, J.A.; Lanfer-Marquez, U.M. Relationship between color (Instrumental and visual) and chlorophyll contents in soybean seeds during ripening. *J. Agric. Food Chem.* **2002**, *50*, 3961–3966. [CrossRef]
35. Crookston, R.K.; Hill, D.S. A visual indicator of the physiological maturity of soybean seed. *Crop Sci.* **1978**, *18*, 867–870. [CrossRef]
36. Nooden, L.D. The phenomena of senescence and aging. In *Senescence and Aging in Plants*; Noodén, L.D., Leopold, A.C., Eds.; Academic Press: San Diego, CA, USA, 1988; pp. 2–50. [CrossRef]
37. Noodén, L.D.; Guiamét, J.J.; John, I. Senescence mechanisms. *Physiol. Plant.* **1997**, *101*, 746–753. [CrossRef]
38. Okatan, Y.; Kahanak, G.M.; Noodén, L.D. Characterization and kinetics of soybean maturation and monocarpic senescence. *Physiol. Plant.* **1981**, *52*, 330–338. [CrossRef]
39. Cha, K.W.; Lee, Y.J.; Koh, H.J.; Lee, B.M.; Nam, Y.W.; Paek, N.C. Isolation, characterization, and mapping of the stay green mutant in rice. *Theor. Appl. Genet.* **2002**, *104*, 526–532. [CrossRef]
40. Park, S.Y.; Yu, J.W.; Park, J.S.; Li, J.; Yoo, S.C.; Lee, N.Y.; Lee, S.K.; Jeong, S.W.; Seo, H.S.; Koh, H.J. The senescence-induced staygreen protein regulates chlorophyll degradation. *Plant Cell* **2007**, *19*, 1649–1664. [CrossRef]
41. Thomas, H.; Smart, C.M. Crops that stay green 1. *Ann. Appl. Biol.* **1993**, *123*, 193–219. [CrossRef]
42. Thomas, H.; Howarth, C.J. Five ways to stay green. *J. Exp. Bot.* **2000**, *51*, 329–337. [CrossRef]
43. Fang, C.; Li, C.; Li, W.; Wang, Z.; Zhou, Z.; Shen, Y.; Wu, M.; Wu, Y.; Li, G.; Kong, L.A. Concerted evolution of D1 and D2 to regulate chlorophyll degradation in soybean. *Plant J.* **2014**, *77*, 700–712. [CrossRef]
44. Guiamet, J.J.; Teeri, J.A.; Nooden, L.D. Effects of nuclear and cytoplasmic genes altering chlorophyll loss on gas exchange during monocarpic senescence in soybean. *Plant Cell Physiol.* **1990**, *31*, 1123–1130.
45. Guiamet, J.J.; Schwartz, E.; Pichersky, E.; Noodén, L.D. Characterization of cytoplasmic and nuclear mutations affecting chlorophyll and chlorophyll-binding proteins during senescence in soybean. *Plant Physiol.* **1991**, *96*, 227–231. [CrossRef]
46. Ott, A.; Yang, Y.; Bhattacharyya, M.; Horner, H.T.; Palmer, R.G.; Sandhu, D. Molecular mapping of D1, D2 and ms5 revealed linkage between the cotyledon color locus D2 and the male-sterile locus ms5 in soybean. *Plants* **2013**, *2*, 441–454. [CrossRef]
47. Palmer, R.; Pfeiffer, T.W.; Buss, G.R.; Kilen, T.C. Qualitative genetics. In *Soybeans: Improvement, Production, and Uses*, 3rd ed.; Boerma, H.R., Specht, J.E., Eds.; ASA, CSSA, SSSA: Madison, WI, USA, 2004; pp. 137–233. [CrossRef]
48. Gitelson, A.A.; Keydan, G.P.; Merzlyak, M.N. Three-band model for noninvasive estimation of chlorophyll, carotenoids, and anthocyanin contents in higher plant leaves. *Geophys. Res. Lett.* **2006**, *33*, L11402. [CrossRef]
49. Monma, M.; Terao, J.; Ito, M.; Saito, M.; Chikuni, K. Carotenoid components in soybean seeds varying with seed color and maturation stage. *Biosci. Biotech. Bioch.* **1994**, *58*, 926–930. [CrossRef]
50. Brotosudarmo, T.H.P.; Limantara, L.; Chandra, R.D.; Heriyanto. Chloroplast pigments: Structure, function, assembly and characterization. In *Plant Growth and Regulation-Alterations to Sustain Unfavorable Conditions*; IntechOpen: London, UK, 2018; pp. 43–80.
51. Granado, F.; Olmedilla, B.; Blanco, I. Nutritional and clinical relevance of lutein in human health. *Brit. J. Nutr.* **2003**, *90*, 487–502. [CrossRef]
52. İnanç, A.L. Chlorophyll: Structural properties, health benefits and its occurrence in virgin olive oils. *Acad. Food J. Akad. GIDA* **2011**, *9*, 26–32.
53. Kizhedath, A.; Suneetha, V. Estimation of chlorophyll content in common household medicinal leaves and their utilization to avail health benefits of chlorophyll. *J. Pharm. Res.* **2011**, *4*, 1412–1413.
54. Landrum, J.T.; Bone, R.A. Lutein, zeaxanthin, and the macular pigment. *Arch. Biochem. Biophys.* **2001**, *385*, 28–40. [CrossRef]
55. Zhang, R.F.; Zhang, F.X.; Zhang, M.W.; Wei, Z.C.; Yang, C.Y.; Zhang, Y.; Tang, X.J.; Deng, Y.Y.; Chi, J.W. Phenolic composition and antioxidant activity in seed coats of 60 Chinese black soybean (*Glycine max* L. Merr.) varieties. *J. Agric. Food Chem.* **2011**, *59*, 5935–5944. [CrossRef]
56. Ko, J.M.; Han, W.Y.; Kim, H.T.; Lee, Y.H.; Choi, M.S.; Lee, B.W.; Shin, S.U.; Seo, J.H.; Oh, K.W.; Yun, H.T.; et al. Soybean cultivar for soy-paste, 'Uram' with mechanization harvesting, large seed, disease resistance and high yield. *Korean J. Breed. Sci.* **2016**, *48*, 301–306. [CrossRef]
57. Yun, H.T.; Park, K.Y.; Moon, J.K.; Kim, Y.H.; Kim, S.L.; Ku, J.H.; Lee, Y.H.; Ryu, Y.H.; Baek, I.Y.; Han, W.Y.; et al. A new black soybean cultivar, Cheongja 3, with green cotyledon, medium-late maturity and high anthocyanin. *Korean J. Breed. Sci.* **2005**, *37*, 261–262.
58. Baek, I.Y.; Kang, S.T.; Shin, D.C.; Choung, M.G.; Han, W.Y.; Kwack, Y.H.; Moon, H.P. A new black soybean variety with green cotyledon, early maturity and large seed size "Cheongjakong". *Korean J. Breed. Sci.* **2001**, *33*, 240–241.
59. Jo, H.; Lee, J.Y.; Cho, H.T.; Choi, H.J.; Son, C.K.; Bae, J.S.; Bilyeu, K.; Song, J.T.; Lee, J.D. Genetic diversity of soybeans (*Glycine max* (L.) merr.) with black seed coats and green cotyledons in Korean germplasm. *Agronomy* **2021**, *11*, 581. [CrossRef]
60. Bilyeu, K.; Palavalli, L.; Sleper, D.; Beuselinck, P. Mutations in soybean microsomal omega-3 fatty acid desaturase genes reduce linolenic acid concentration in soybean seeds. *Crop Sci.* **2005**, *45*, 1830–1836. [CrossRef]
61. Association of Official and Analytical Chemists (AOAC). *Official Methods of Analysis*, 18th revision ed.; AOAC: Rockville, MD, USA, 2011.
62. Kaiser, H.F. The application of electronic computers to factor analysis. *Educ. Psychol. Meas.* **1960**, *20*, 141–151. [CrossRef]
63. Cattell, R.B. The screen test for the number of factors. *Multivar. Behav. Res.* **1966**, *1*, 245–276. [CrossRef]
64. Ryo, H.C. A study on the categorization of the strategy group of program provider (PP). *J. Korean Data Inf. Sci. Soc.* **2008**, *19*, 913–924.

65. Caliński, T.; Harabasz, J. A dendrite method for cluster analysis. *Commun. Stat.-Theor. M.* **1974**, *3*, 1–27. [CrossRef]
66. Cooper, M.C.; Milligan, G.W. The effect of measurement error on determining the number of clusters in cluster analysis. In *Data, Expert Knowledge and Decisions*; Gaul, W., Schader, M., Eds.; Springer: Berlin/Heidelberg, Germany, 1988; pp. 319–328. [CrossRef]
67. Milligan, G.W.; Cooper, M.C. An examination of procedures for determining the number of clusters in a data set. *Psychometrika* **1985**, *50*, 159–179. [CrossRef]
68. Hunter, J.E.; Zhang, J.; Kris-Etherton, P.M. Cardiovascular disease risk of dietary stearic acid compared with trans, other saturated, and unsaturated fatty acids: A systematic review. *Am. J. Clin. Nutr.* **2010**, *91*, 46–63. [CrossRef]
69. Cardinal, A.J.; Burton, J.W.; Camacho-Roger, A.M.; Yang, J.H.; Wilson, R.F.; Dewey, R.E. Molecular analysis of soybean lines with low palmitic acid content in the seed oil. *Crop Sci.* **2007**, *47*, 304–310. [CrossRef]
70. Tajuddin, T.; Watanabe, S.; Yamanaka, N.; Harada, K. Analysis of quantitative trait loci for protein and lipid contents in soybean seeds using recombinant inbred lines. *Breed. Sci.* **2003**, *53*, 133–140. [CrossRef]
71. Downs, R.J.; Siegelman, H.W.; Butler, W.L.; Hendricks, S.B. Photoreceptive pigments for anthocyanin synthesis in apple skin. *Nature* **1965**, *205*, 909–910. [CrossRef]
72. Kannangara, C.G.; Hansson, M. Arrest of chlorophyll accumulation prior to anthocyanin formation in *Euphorbia pulcherrima*. *Plant Physiol. Bioch.* **1998**, *36*, 843–848. [CrossRef]
73. Slatnar, A.; Mikulic-Petkovsek, M.; Veberic, R.; Stampar, F.; Schmitzer, V. Anthocyanin and chlorophyll content during poinsettia bract development. *Sci. Hortic.* **2013**, *150*, 142–145. [CrossRef]
74. Lee, J.D.; Shannon, J.G.; So, Y.S.; Sleper, D.A.; Nelson, R.L.; Lee, J.H.; Choung, M.G. Environmental effects on lutein content and relationship of lutein and other seed components in soybean. *Plant Breed.* **2009**, *128*, 97–100. [CrossRef]
75. Wu, X.L.; Liu, Z.H.; Hu, Z.H.; Huang, R.Z. BnWRI1 coordinates fatty acid biosynthesis and photosynthesis pathways during oil accumulation in rapeseed. *J. Integr. Plant Biol.* **2014**, *56*, 582–593. [CrossRef]
76. Ruuska, S.A.; Schwender, J.; Ohlrogge, J.B. The capacity of green oilseeds to utilize photosynthesis to drive biosynthetic processes. *Plant Physiol.* **2004**, *136*, 2700–2709. [CrossRef]
77. Joo, Y.H.; Park, J.H.; Choung, M.G.; Yun, S.G.; Chung, K.W. Variation of contents and color difference of anthocyanin by different cultivation year in black soybean seed. *Korean J. Crop Sci.* **2004**, *49*, 507–511.
78. Liu, K.; Orthoefer, F.; Brown, E.A. Association of seed size with genotypic variation in the chemical constituents of soybeans. *J. Am. Oil Chem. Soc.* **1995**, *72*, 189–192. [CrossRef]
79. Yoon, T.H.; Im, K.J.; Kim, D.H. Fatty acid composition of lipids obtained from Korean soybean varieties. *Korean J. Food Sci. Technol.* **1984**, *16*, 375–382.
80. Adhvaryu, A.; Erhan, S.Z.; Perez, J.M. Preparation of soybean oil-based greases: Effect of composition and structure on physical properties. *J. Agric. Food Chem.* **2004**, *52*, 6456–6459. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.