Journal of Integrative Agriculture 2023, 22(9): 2632–2647

Available online at www.sciencedirect.com

ScienceDirect

RESEARCH ARTICLE

Natural variations and geographical distributions of seed carotenoids and chlorophylls in 1 167 Chinese soybean accessions

Berhane S. GEBREGZIABHER[1, 2*], ZHANG Sheng-rui[1*], Muhammad AZAM[1], QI Jie[1], Kwadwo G. AGYENIM-BOATENG[1], FENG Yue[1], LIU Yi-tian[1], LI Jing[1], LI Bin[3#], SUN Jun-ming[1#]

[1] *National Engineering Laboratory for Crop Molecular Breeding, Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing 100081, P.R.China*

[2] *Crop Sciences Research Department, Mehoni Agricultural Research Center, Maichew 7020, Ethiopia*

[3] *MARA Key Laboratory of Soybean Biology (Beijing), Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing 100081, P.R.China*

Abstract

Understanding the composition and contents of carotenoids in various soybean seed accessions is important for their nutritional assessment. This study investigated the variability in the concentrations of carotenoids and chlorophylls and revealed their associations with other nutritional quality traits in a genetically diverse set of Chinese soybean accessions comprised of cultivars and landraces. Genotype, planting year, accession type, seed cotyledon color, and ecoregion of origin significantly influenced the accumulation of carotenoids and chlorophylls. The mean total carotenoid content was in the range of 8.15–14.72 μg g^{-1} across the ecoregions. The total carotenoid content was 1.2-fold higher in the landraces than in the cultivars. Soybeans with green cotyledons had higher contents of carotenoids and chlorophylls than those with yellow cotyledons. Remarkably, lutein was the most abundant carotenoid in all the germplasms, ranging from 1.35–37.44 μg g^{-1}. Carotenoids and chlorophylls showed significant correlations with other quality traits, which will help to set breeding strategies for enhancing soybean carotenoids without affecting the other components. Collectively, our results demonstrate that carotenoids are adequately accumulated in soybean seeds, however, they are strongly influenced by genetic factors, accession type, and germplasm origin. We identified novel germplasms with the highest total carotenoid contents across the various ecoregions of China that could serve as the genetic materials for soybean carotenoid breeding programs, and thereby as the raw materials for food sectors, pharmaceuticals, and the cosmetic industry.

Received 11 May, 2022 Accepted 9 September, 2022
Berhane S. GEBREGZIABHER, E-mail: berhane76@gmail.com; ZHANG Sheng-rui, E-mail: zhangshengrui@caas.cn; [#]Correspondence SUN Jun-ming, Tel: +86-10-82105805, E-mail: sunjunming@caas.cn; LI Bin, Tel/Fax: +86-10-82105805, E-mail: libin02@caas.cn
[*] These authors contributed equally to this study.

doi: 10.1016/j.jia.2022.10.011

Keywords: accession type, carotenoid, chlorophyll, ecoregion, geographical distribution, Pearson's correlation, soybean (*Glycine max* L. Merrill)

1. Introduction

Soybean (*Glycine max* L. Merrill), which belongs to the family Leguminosae, is the most important economic legume crop globally, and the fourth crop after rice, wheat and maize in terms of global production. It has been used as a major source of nutritious feed for humans and livestock and has long been an essential, year-round component of traditional Asian and vegetarian diets (Kaga *et al.* 2012). Soybean is considered as a rich source of protein and oil, as well as various bioactive compounds (Wang *et al.* 2007; Singh *et al.* 2019). The bioactive compounds are associated with human health and nutrition due to their provitamin A, antioxidant and anti-inflammatory activities (Rezaei *et al.* 2016). Carotenoids are among the bioactive phytochemical compounds of soybean seeds that play significant roles in reducing age-related macular degeneration, cardiovascular diseases, some types of cancer, diabetes, and UV-induced skin damage, as well as promoting bone health, and improving metabolism, among others (Chatterjee *et al.* 2018). They are also involved in light harvesting, photoprotection through singlet oxygen quenching and non-photochemical quenching, vision, communication between or within species through coloration, protection against oxidants, modulation of the properties of membranes, fertility and reproduction (Dias *et al.* 2018).

More than 750 carotenoids have been identified in nature, and they comprise two classes: carotenes and xanthophylls, which are hydrocarbon and oxygenated derivatives, respectively. They are widely distributed in plants, algae, bacteria, and some fungi, while animals get them from their diet. The most dominant carotenoids in the human diet include β-carotene, α-carotene, β-cryptoxanthin, lutein, zeaxanthin, and lycopene, and they are the most abundant carotenoids in human blood plasma (Rao and Rao 2007). Chlorophylls, primarily chlorophyll *a* (Chl *a*) and chlorophyll *b* (Chl *b*), are another group of plant pigments found in soybean seeds. Chlorophyll and its derivatives are among the phytochemical compounds that are believed to be potentially associated with the prevention of chronic diseases, including cancer (Ferruzzi and Blakeslee 2007), and they are also used as food product additives because of their color and physicochemical properties (Schoefs 2002). Therefore, it is important to quantify the concentrations of chlorophylls in food crops, including soybeans.

The soybean was domesticated in China about 4 500 years ago, and immense germplasm resources have been collected and dispersed throughout the world (Qiu *et al.* 2011). Previous reports have shown the geographical distribution and genetic diversity of over 20 000 Chinese soybean accessions (Dong *et al.* 2004), which were screened based on qualitative and quantitative traits. Several studies have evaluated Chinese accessions with respect to disease resistance (Huang *et al.* 2016), abiotic stress tolerance (Guan *et al.* 2014), and nutritional quality attributes such as isoflavones (Azam *et al.* 2020), fatty acids (Abdelghany *et al.* 2020), tocopherols (Ghosh *et al.* 2021) and folates (Agyenim-Boateng *et al.* 2022). However, there is limited information pertaining to the carotenoid contents of Chinese soybean accessions, implying that carotenoids in soybean seeds have not been given adequate attention in breeding research programs, despite of their role in health and nutrition. Furthermore, soybeans are rarely considered as a primary carotenoid source compared to other food crops, such as corn and peas. However, soybean derived food is a major part of the traditional diets of many Asian countries, including China, Japan, and Korea (Teng *et al.* 2017). Apparently, increasing the carotenoid content in soybean seeds is believed to be an effective way to improve the nutritional value of soybean derived foods. Hence, soybean cultivars with greater carotenoid contents in the seeds would be the desirable breeding materials for soybean breeders. The large-scale evaluation of unexploited soybean germplasms with diversified phytochemical properties is needed for the improvement of soybean seed carotenoid accumulation by conventional breeding. The high genetic diversity of available soybean germplasms provides an opportunity for plant breeders to develop cultivars with traits of interest (Wang *et al.* 2016). Therefore, in the present study, a collection of 1 167 Chinese soybean accessions (cultivars and landraces) conserved in the National Crop Genebank of China (NCGC) were analyzed to obtain a diversified and detailed account of their carotenoid and chlorophyll contents.

The genetic background of a cultivar, seed cotyledon color (Lee *et al.* 2017), and environmental factors, including temperature, salinity, field mold stress and others (Deng *et al.* 2022), influence the distribution and variation of targeted nutritional components in soybean seeds. Among the nutritional components, carotenoid accumulation can be affected by genotype (G), environment (E) and climatic factors. Previous studies (Whent *et al.* 2009) showed that genetic factors contributed 78% of the variation in carotenoid contents, while environment and G×E interactions represented 18 and 4% of the variation, respectively. Soybeans can grow in the various ecoregions of China, with diverse environmental conditions. Evidence suggests that differences in climate and soil have made Chinese soybeans topographically distributed in three main growing areas, the northern region (>40°N), Huang–Huai–Hai Valley region (~30 to 40°N) and southern region (19 to 31°N), owing to different planting seasons (Li *et al.* 2008, 2009; Song *et al.* 2016). The germplasm accessions included in this study were collected from these regions.

Notably, screening a considerable number of germplasms is important to facilitate the selection of genotypes with higher levels of carotenoid accumulation for breeding programs. However, detailed information concerning to carotenoid and chlorophyll profiles and concentrations in soybean seeds for a set of diverse soybean germplasms from different ecoregions of origin, and aligned with their geographical distribution, is extremely limited. To date, we are aware of only such two studies, both in Japan, that profiled and analyzed variations of carotenoid concentrations (particularly lutein and β-carotene) in 50 soybean varieties (Monma *et al.* 1994) as well as in 490 soybean cultivars and 610 wild soybean accessions (Kanamaru *et al.* 2006). As mentioned above, some other small-scale studies have also reported the effects of genotype, environmental conditions and management practices on carotenoid concentrations (mainly lutein) of 8 (Whent *et al.* 2009), 15 (Lee *et al.* 2009) and 20 (Seguin *et al.* 2011) soybean genotypes.

To remedy this lack of important information, this study aimed to (i) analyze the natural variations of carotenoids and chlorophylls in a diverse panel of 1 167 Chinese soybean germplasm accessions; (ii) investigate the correlations of carotenoid and chlorophyll components with other important soybean nutritional characteristics (such as protein, tocopherol, isoflavone, oil and fatty acids); and (iii) visualize the geographical distribution of carotenoids to identify hotspot regions and select elite accessions with the desired carotenoid compositions. The information obtained in this study will help breeders and producers to develop novel breeding strategies for enhancing the carotenoid content in soybean crops, thereby meeting consumer demands, while enhancing the market value of the seeds for producers.

2. Materials and methods

2.1. Chemicals and reagents

The standards of carotenoid components, such as zeaxanthin (CAS: 144-68-3, purity≥85%) and β-carotene (CAS: 7235-40-7, purity≥98%), and chlorophyll standards including Chl *a* (CAS: 479-61-8, purity≥85%) and Chl *b* (CAS: 519-62-0, purity≥90%), were purchased from Shanghai Yuanya Biotechnology Co., Ltd. (Shanghai, China), while the carotenoid standard of lutein (CAS: 127-40-2, purity≥96%) was purchased from Sigma-Aldrich (St. Louis, MO, USA). Methanol, acetone, ethanol, and ammonium acetate were all HPLC grade and purchased from Thermo Fisher Scientific Co., Ltd. (New Jersey, USA). Methyl *tert*-butyl ether, and butylated hydroxytoluene, both HPLC grade, were purchased from Mreda Technology Inc., USA and Shanghai Macklin Biochemical Co., Ltd., Shanghai, China, respectively. The stock solution of each standard was prepared and manipulated under dim light to prevent light-induced isomerization or degradation. Individual working standards and stock solutions were stored in amber glass vials at –20 and –80°C, respectively. The calibration curve equation for each standard, with R^2 of 0.999 for accurate linearity, was prepared and used to determine the carotenoid and chlorophyll contents of soybean seeds (Appendix A). More details on this process were described in a previous study (Gebregziabher *et al.* 2021).

Three standards of tocopherol components: α-tocopherol (CAS: 10191-41-0, purity≥96%), γ-tocopherol (CAS: 54-28-4, purity≥96%) and δ-tocopherol (CAS: 119-13-1, purity≥90%), were purchased from Sigma-Aldrich Chemical Company (St. Louis, MO, USA). The standards of isoflavone components, including daidzin (552–66–9, purity ≥98), glycitin (GL) (CAS: 40246-10-4, purity≥98), genistin (CAS: 529-59-9, purity≥98%), malonyldaidzin (MD) (CAS: 124590-31-4, purity≥90%), malonylglycitin (MGL) (CAS: 137705-39-6, purity≥90%), malonylgenistin (MG) (CAS: 51011-05-3, purity≥90%), daidzein (CAS: 486-66-8, purity≥98%), and genistein (CAS: 446-72-0, purity≥98%), were provided by Dr. Akio Kikuchi, National Agricultural Research Center for Tohoku Region, Japan. HPLC grade ethanol and acetonitrile were purchased from Thermo Fisher Scientific Co., Ltd. (New Jersey, USA). Acetic acid (Chengdu Xiya Chemical Technology Co., Ltd., Chengdu, China) was also used in the extraction of isoflavones.

Concerning the fatty acids, for a mixed standard of five fatty acid methyl esters, methyl palmitate (C16:0, palmitic acid (PA), purity≥98.5%), methyl stearate (C18:0, stearic acid (SA), purity≥99%), methyl oleate (C18:1, oleic acid (OA), purity≥99%), methyl linoleate (C18:2, linoleic acid

(LA), purity≥98.5%), and methyl linolenate (C18:3, Linolenic acid (LNA), purity≥99%), were purchased from Sigma-Aldrich (St. Louis, MO, USA). Other reagents, including sodium methoxide from Sigma-Aldrich (St. Louis, MO, USA), as well as HPLC grade *n*-hexane and methanol from Thermo Fisher Scientific Co., Ltd. (Shanghai, China), were deployed in the determination of fatty acids.

2.2. Plant materials and field experiments

The experiments included 1 167 Chinese soybean accessions collected from three major soybean growing ecoregions in China: the northern region (NR, 241 accessions), the Huang–Huai–Hai Valley region (HR, 434 accessions), and the southern region (SR, 492 accessions). Among these accessions, 926 were landraces and 241 were modern cultivars selected from the Chinese primary core collection. This core collection was developed to capture (as much as possible) the phenotypic diversity and geographic distribution of the full collection of 23 587 soybean germplasm accessions conserved in the Chinese National Soybean GeneBank (CNSGB), Institute of Crop Sciences, Chinese Academy of Agricultural Sciences (Li *et al.* 2008; Huang *et al.* 2016). The set used in this study included accessions with both green and yellow seed cotyledon colors. The detailed information on the numbers and types of accessions is provided in Appendix B.

Field experiments were conducted at Changping, Beijing (40°13´N, 116°12´E) and Sanya, Hainan Province (18°24´N, 109°5´E) in the 2017 and 2018 cropping seasons. The soil pH, total nitrogen, phosphorus and potassium levels of the experimental site were 8.22, 80.5 mg kg^{-1}, 68.7 mg kg^{-1} and 12.31 g kg^{-1}, respectively, at Changping; and 5.27, 98.59 mg kg^{-1}, 39.68 mg kg^{-1}, and 80.78 g kg^{-1}, respectively, at Sanya. The experiments were laid out in a randomized incomplete block design across locations due to the large number of accessions and limited land resources. Planting materials were planted in 3.0-m long rows with 0.5 and 0.1 m inter- and intra-row spacings, respectively. Each line in each plot contained 20 plants as a source of seeds that were used for subsequent carotenoid and chlorophyll content measurements. Other recommended agronomic practices were carried out as previously reported (Abdelghany *et al.* 2020). The mean monthly temperature, rainfall, and sunshine of the experimental locations are shown in Appendix C.

2.3. Carotenoid and chlorophyll extraction and determination

The carotenoids and chlorophylls of mature soybean seeds were extracted and analyzed according to the method described by Gebregziabher *et al.* (2021). Briefly, 20 g of seeds from each accession were ground to a fine powder with a sample preparation Mill (Retsch ZM100, Φ=1.0 mm, Rheinische, Germany). An analytical balance (Sartorius BS124S, Gottingen, Germany) was used to measure out 100 mg of the powder from each sample, which was then transferred into a 2-mL micro-centrifuge tube preloaded with 1.5 mL of a mixture of ethanol and acetone solvents in a 1:1 ratio. These solvents were premixed with 0.1% butylated hydroxytoluene (w/v) to protect the carotenoids and chlorophylls from degradation and oxidation. The sample was homogenized by vortexing (IKA® MS 3 basic, Germany) at 300 r min^{-1} for 1 min and then placed in an ultrasonic water bath (Ningbo Scientz Biotechnology Co. Ltd., Ningbo, China) for 20 min. The supernatant was collected after centrifugation at 13 000 r min^{-1} for 10 min at 4°C and transferred to a new centrifuge tube for another centrifugation at 13 000 r min^{-1} for 5 min at 4°C. The supernatant was then filtered using a YMC duo-filter (YMC Co., Kyoto, Japan) through a 0.22-µm pore dimension with the help of a sterile syringe (Jiangsu Zhiyu Medical Equipment Co., Ltd., Jiangsu, China) and placed in a 1.5-mL amber glass HPLC vial (AS ONE, Ningbo, China) for subsequent analysis. The analysis was carried out in an Agilent 1100 Model HPLC instrument (Agilent Technologies, Santa Clara, CA, USA) fitted with a Hewlett-Packard Model 1050 solvent delivery system, using the mobile phases of methanol-10 mmol L^{-1} ammonium acetate, methyl *tert*-butyl ether (100%), and ultrapure water (100%) delivered at a 0.9-mL min^{-1} flow rate through a C30-YMC Carotenoid (250 mm×4.6 mm I.D., S-5 µm, Kyoto, Japan) column coupled with a UV-Vis detector (Santa Clara, CA, USA) set at 450 nm. Finally, the concentrations of each component were calculated using the formula: Carotenoid (mg g^{-1})=[Cx·V·D]/Wt; where, Cx is the concentration of each carotenoid component calculated from the standard calibration curve equation, V is the volume of the extraction solvent, D is any dilution factor, and Wt is the sample weight on a dry basis.

2.4. Analysis of other soybean seed quality components

Soybean seed quality traits, including protein, oil, fatty acids, tocopherols and isoflavones, were also determined for the soybean accessions evaluated in this study. The analysis of individual and total tocopherol components was carried out in an Agilent 1200 Series HPLC instrument (Agilent Technologies, Santa Clara, CA, USA) with a C18 reversed-phase column (YMC ODS AM-303, 250 mm×4.6 mm, S-5 μm, YMC Co., Kyoto, Japan) coupled with a UV detector (Santa Clara, CA, USA). The detailed procedure for the extraction and determination of tocopherols has been reported recently (Ghosh *et al.* 2021). The protein and oil contents were measured by Fourier transform near-infrared spectroscopy (MPA, Bruker Fourier Rheinstetten, Germany). The five predominant fatty acids, PA, SA, OA, LA, and LNA acids, were derivatized into their methyl esters and determined by using gas chromatography (GC-2010, Shimadzu Inc., Kyoto, Japan). The extraction and determination procedure for the fatty acids was similar to that in a previous report (Abdelghany *et al.* 2020). The determinations of individual and total isoflavone concentrations were carried out using the method of Sun *et al.* (2011), and its detailed extraction and determination procedure has also been reported recently (Azam *et al.* 2020).

2.5. Statistical analysis

Analysis of variance (ANOVA) was performed on the combined data using the procedure of a general linear model (PROC GLM) (SAS version 9.1, SAS Institute Inc., Cary, NC, USA). Differences were deemed significant when the *P*-value was less than 0.05. Multiple comparisons of means were performed using Tukey's honestly significant difference (HSD) test. ANOVAs were calculated to determine the effects of accession, accession type, seed cotyledon color, and ecoregion on the carotenoid and chlorophyll concentrations. Accession, ecoregion, accession type and cotyledon color were considered as fixed effects, while location nested within years was set as a random effect. Boxplots were drawn to show the distributions and variations of seed carotenoid and chlorophyll compositions among the three ecoregions, seed cotyledon colors and accession types (i.e., landraces and cultivars). Pearson's correlation analysis (*r*) was calculated to identify associations between the various chemical compositions of soybean seeds. Principal component analysis (PCA) was performed to identify the components with high discriminatory properties, which were in turn used to group accessions, and also to show the contribution of each component to the total variation among different ecoregions and accession types. PCA, *r*, and boxplots were determined using R statistical software version 3.6.3 (R Foundation for Statistical Computing, Vienna, Austria). In the R project, the 'ggplot2' package was used for data visualization and drawing boxplots, while the corrplot package was used to graphically display the correlation matrix. Geographical distribution maps of soybean seed carotenoid and chlorophyll composition means were constructed with ArcGIS Pro 2.7 (Esri, Redlands, CA, USA).

3. Results

3.1. Variations of the carotenoid and chlorophyll contents in various soybean accessions

The contributions of various soybean accessions to seed composition were statistically analyzed (Appendix D). Consequently, large variations in the contents of carotenoids and chlorophylls were observed ($P<0.01$). Overall means and variation of the traits across the two planting years are summarized in Table 1. Different carotenoid components, such as lutein, zeaxanthin, β-carotene and total carotenoids, were found to have means of 11.37, 0.59, 0.69, and 11.69 μg g^{-1}, respectively (Table 1). Additionally, Chl *a* and Chl *b* were determined to have concentrations ranging from 1.02–111.46 and 0.34–36.87 μg g^{-1}, with means of 11.83 and 3.01 μg g^{-1}, respectively (Table 1). The Chl *a* and Chl *b* values were only obtained from 59 and 91% of the total accessions, respectively, implying that the seeds of many soybean accessions do not contain chlorophylls. The highest (111.46 μg g^{-1}) and lowest (1.02 μg g^{-1}) concentrations of Chl *a* were obtained from landrace ZDD13590 and *cv*. ZDD24801, respectively, both from SR. Similarly, landrace ZDD10132 gave the highest (36.87 μg g^{-1}) content of Chl *b*, while landrace ZDD12327 and *cv*. ZDD14190 had the lowest (0.34 μg g^{-1}) levels.

Table 1 Variations in the carotenoid and chlorophyll contents of Chinese soybean accessions[1)]

Component	Minimum (μg g^{-1})	Maximum (μg g^{-1})	Range (μg g^{-1})	Mean (μg g^{-1})	Standard deviation	Coefficient of variation (%)
Lutein	1.35	37.44	36.09	11.37	6.62	58.30
Zeaxanthin	0.02	3.38	3.37	0.59	0.43	73.53
β-Carotene	0.04	3.62	3.57	0.69	0.46	67.35
Chlorophyll *a*	1.02	111.46	110.44	11.83	9.42	79.68
Chlorophyll *b*	0.34	36.87	36.54	3.01	3.79	126.20
Total carotenoid	1.35	40.76	39.41	11.69	7.08	60.59

[1)] Data are expressed as the means of two replicates.

Lutein was detected in all accessions, while only 37 and 65% of the accessions contained β-carotene and zeaxanthin components, respectively. The highest total carotenoid content was observed in landrace ZDD13590, while the lowest was in *cv*. ZDD25115 and landrace ZDD21672, all from SR, showing a 30-fold difference between these accessions. As shown in Appendix E, landrace ZDD00294 from NR contained the highest lutein content (37.44 μg g^{-1}), and *cv*. ZDD25115 and landrace ZDD21672 from SR had the lowest lutein content (1.35 μg g^{-1}). The content of zeaxanthin was in the range of 0.02–3.38 μg g^{-1} with the highest value observed in landrace ZDD06562, originally collected from SR. In the case of β-carotene, ZDD10132 (landrace) and ZDD14267 (landrace) showed the highest (3.62 μg g^{-1}) and lowest (0.04 μg g^{-1}) contents, respectively. This study deployed a wide range of diverse soybean accessions in order to help breeders to demonstrate the extreme variability in carotenoid concentrations among soybean seeds. Interestingly, some accessions in the present study (Appendices E and F) showed a remarkable accumulation of carotenoid components, which will benefit soybean breeding for carotenoid enhancement.

3.2. Seed carotenoid and chlorophyll contents vary among soybean accession types

The concentrations of carotenoids and chlorophylls were significantly influenced (P<0.01) by accession type (Fig. 1; Appendix D). Moreover, cultivation year and the accession type by year interaction significantly influenced the chlorophyll and carotenoid components, except for Chl *b* (Appendix D), indicating that seasonal variation had a pronounced impact on the soybean seed carotenoid and chlorophyll profiles. As shown in Fig. 1, the total carotenoid content of landraces ranged from 1.35–40.76 μg g^{-1} with a mean of 12.09 μg g^{-1}, which was 1.20-fold higher than that of cultivars, indicating that landraces are rich sources of carotenoids. Similarly, the lutein (1.35–37.44 μg g^{-1}), zeaxanthin (0.02–3.38 μg g^{-1}) and β-carotene (0.09–3.62 μg g^{-1}) contents of landraces were 1.30–2.10-fold higher than the respective components in cultivars. By comparison, landraces also contained considerably higher chlorophyll content than cultivars, suggesting the genetic potential and environmental adaptability of landraces to withstand harsh conditions.

3.3. Variations of carotenoid and chlorophyll contents across ecoregions

Comparisons of the carotenoid profiles and concentrations of soybean accessions that originated in the three regions of China (NR, HR and SR) are shown in Fig. 2. The ANOVA results showed significant differences (P<0.05) among ecoregions for all carotenoid and chlorophyll components, which might be attributed to both genetic and environmental factors. The ecoregion by year interaction revealed a very highly significant effect (P<0.001) on the carotenoid components. Cultivation year also significantly (P<0.05) affected the carotenoid components. Total carotenoid showed the greatest difference in HR (14.72 μg g^{-1}), followed by NR (13.46 μg g^{-1}) and then the SR (8.15 μg g^{-1}) accessions. Similarly, accessions originating from HR contained the highest levels of β-carotene and zeaxanthin (Fig. 2). The concentration of lutein, as shown in Fig. 2, revealed a decreasing trend from high (north) to low (south) latitudes, with no significant difference between the NR (13.18 μg g^{-1}) and HR (14.23 μg g^{-1}) accessions, while the SR accessions contained 39.60 and 44.06% lower levels than the NR and HR accessions, respectively. It is noteworthy that the HR accessions contained a higher concentration of chlorophylls compared with the other two regions, which were significantly the same (Fig. 2). These results provide new insights into soybean seed carotenoid variability among germplasms with diverse geographical origins. Taken together, our results suggest that HR accessions could be preferable for producing high quality soybean nutrition.

3.4. Seed cotyledon color has a pronounced effect on the carotenoid and chlorophyll contents

The seed cotyledon color of accessions highly significantly (P<0.001) influenced the concentrations of carotenoids and chlorophylls (Fig. 3; Appendix D). The interaction of planting year and cotyledon color significantly affected (P<0.05) the accumulations of Chl *b* and β-carotene, while it had no impact on the other components (Appendix D). The variations in the carotenoid and chlorophyll compositions between the seed cotyledon colors are shown in Fig. 3. The mean concentrations of total carotenoid, lutein, zeaxanthin and β-carotene in green cotyledon accessions were 2.39-, 2.28-, 1.8-, and 2.83-fold higher than the respective components in yellow cotyledon-colored accessions, respectively. Similarly, accessions with green cotyledons contained higher concentrations of Chl *a* (9.34 μg g^{-1}) and Chl *b* (2.57 μg g^{-1}) as compared to the yellow ones. These results suggest that green cotyledon-colored accessions have higher carotenoid and chlorophyll concentrations, as reflected by the contents of lutein, zeaxanthin, β-carotene, total carotenoids, Chl *a* and *b* in the various accessions,

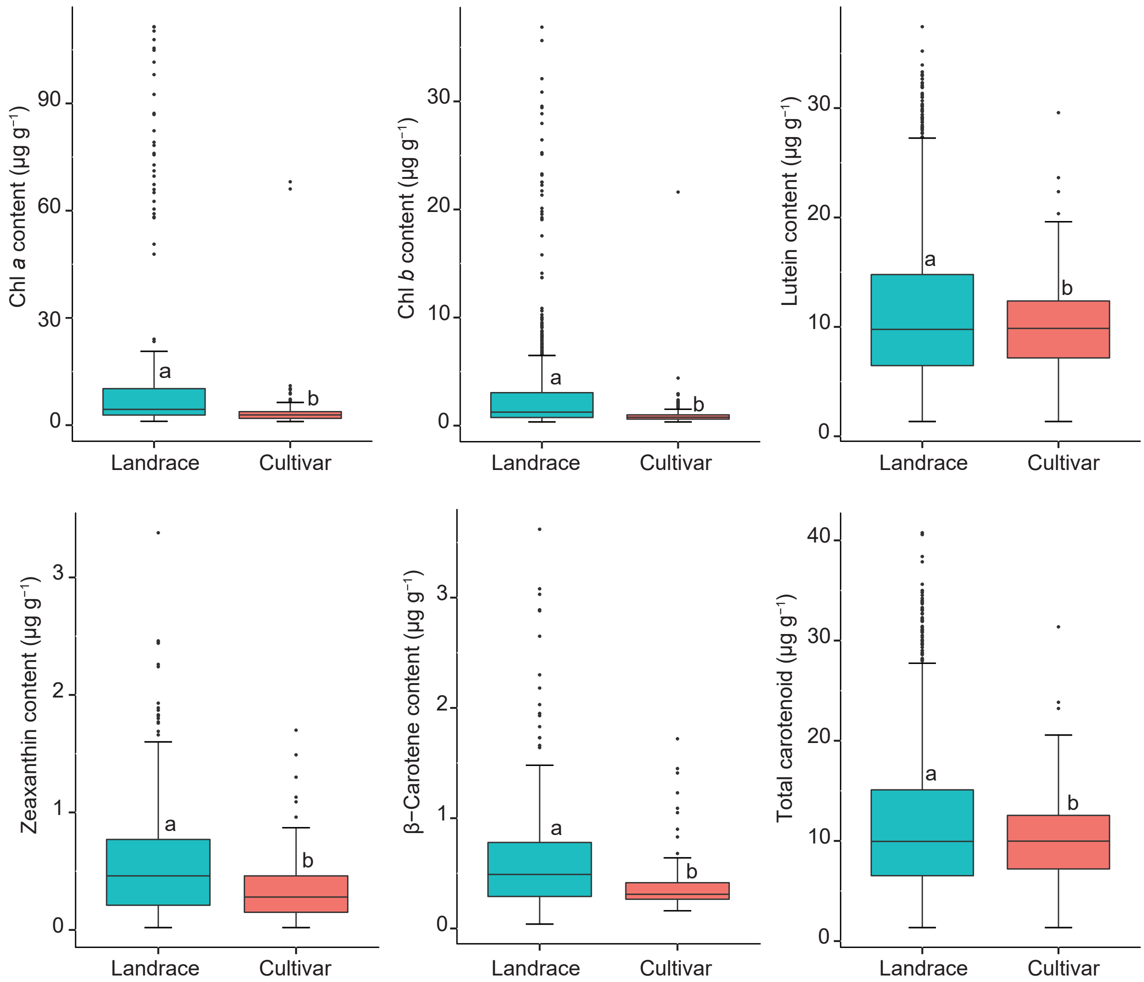


Fig. 1 Variations in carotenoid and chlorophyll (Chl) concentrations of soybean cultivars and landraces. Different lower-case letters (a and b) indicate significant differences at the P<0.05 level.

so they could be ultimately preferred for soybean carotenoid breeding.

3.5. Principal component analysis based on eco-region and accession type

Two PCA constructed based on accession type and ecoregion, respectively, are shown in Figs. 4 and 5. In both biplots, the first component of PCA (PC1) and the second component (PC2) accounted for 80% of the total variability, explaining 61.6 and 18.4% of the total variance, respectively. All carotenoid and chlorophyll components (lutein, zeaxanthin, β-carotene, Chl *a* and *b*) positively contributed to the total variance. Chl *b* (28.5%) contributed the most to the variance in PC1, followed by Chl *a* (26.1%) and β-carotene (21.1%), while zeaxanthin (66.9%) followed by lutein (12.2%) were the two best contributors to the total variance in PC2. The accession type PCA (Fig. 4) showed that landraces are more diverse than cultivars, resulting in very high contents of carotenoids and chlorophylls, so they can be preferred as reliable sources for those components. On the other hand, although the number of HR accessions evaluated (434) was lower than the number of SR accessions (492), the HR accessions are more diverse in the PCA biplot (Fig. 5), signifying that HR accessions tend to contain the highest levels of carotenoid and chlorophyll components.

3.6. Correlation analysis between carotenoids and other seed nutritional characteristics

The correlation coefficients between each chemical component of the soybean seeds are summarized in Fig. 6. Interestingly, a perfectly strong significant positive correlation (r=1.00***) was observed between total carotenoid and lutein, indicating that lutein contributes the highest amount to the total carotenoid accumulation. Lutein was followed by β-carotene, which showed a

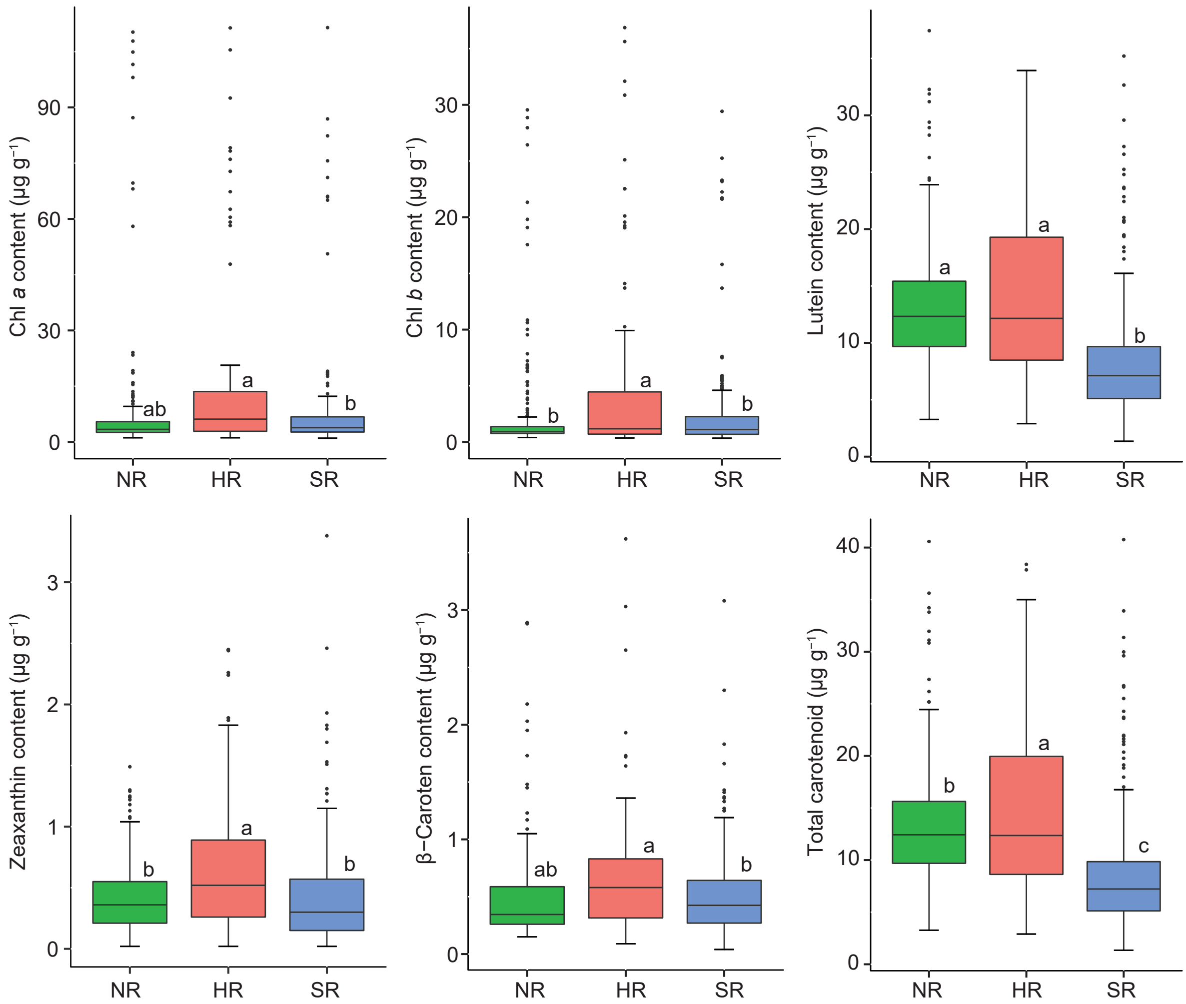


Fig. 2 Variations in carotenoid and chlorophyll (Chl) concentrations among accessions with diverse ecoregions of origin. NR, northern region; HR, Huang–Huai–Hai Valley region; SR, southern region. Different lower-case letters (a, b, and c) indicate significant differences at the *P*<0.05 level.

strong correlation (r=0.81***) with total carotenoid. Chl *a* was positively correlated with Chl *b* (r=0.94***), lutein (r=0.56***), zeaxanthin (r=0.28***), β-carotene (r=0.78***) and total carotenoids (r=0.59***); while Chl *b* was positively associated with lutein (r=0.68***), zeaxanthin (r=0.35***), β-carotene (r=0.75***) and total carotenoids (r=0.70***).

The carotenoid components were also correlated with other seed nutritional quality characteristics. As shown in Fig. 6, total carotenoid and lutein were positively correlated with LA (r=0.22***), GL (r=0.34***), MD (r=0.20***), and MGL (r=0.42***), implying that a consistent increase in lutein and total carotenoid is accompanied by corresponding increases of these components. On the other hand, these components were negatively correlated with OA (r=–0.22***), δ-tocopherol (r=–0.41*** and r=–0.42***, respectively), and total tocopherol (r=–0.20*** and r=–0.20***, respectively). Notably, zeaxanthin had more significant and positive correlations with the derivatives of the malonylglycosides components (MD, MGL and MG), which are the major components of isoflavones in soybean seeds. The β-carotene also had good positive correlations with these components, as well as glycitin and LNA; however, it was negatively correlated with oil and SA, suggesting that accessions with higher β-carotene contents tend to have lower oil and fatty acids (specifically SA). Overall, the carotenoid components were significantly correlated with other quality traits, except for protein and saturated fatty acids (PA). The carotenoid components showed a weak negative correlation with oil, and comparatively strong correlations with LA, glycitin, and MGL. The existence of negative or positive correlations between traits is quite helpful for allowing breeders to improve the soybean seed carotenoid profile without sacrificing other beneficial traits under the appropriate environmental conditions.

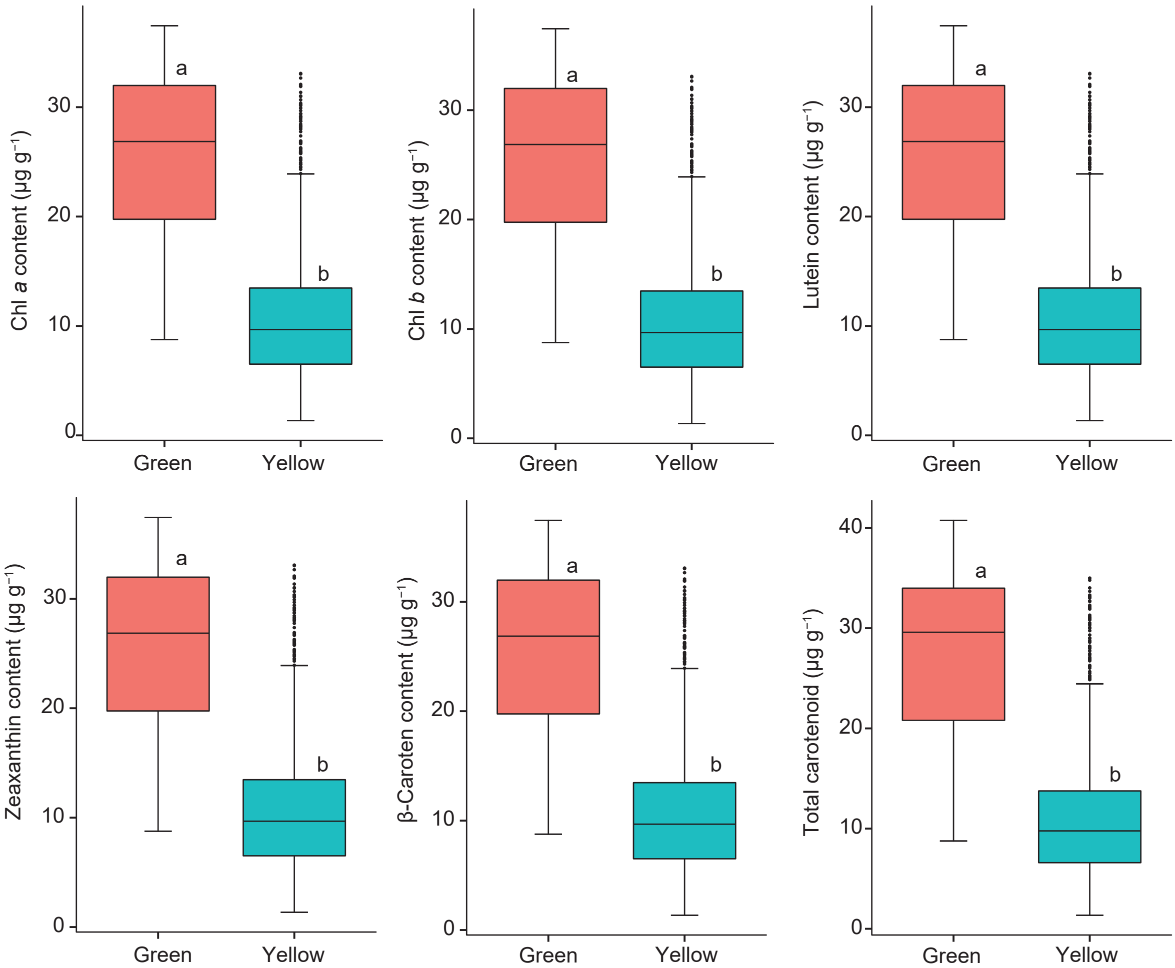


Fig. 3 Variations in carotenoid and chlorophyll (Chl) concentrations among accessions with different seed cotyledon colors. Different lower-case letters (a and b) indicate significant differences at the *P*<0.05 level.

3.7. Geographic distribution of carotenoid contents in soybean seeds

The inter-relationships of geographical factors (latitude, longitude, and altitude) with carotenoid and chlorophyll components are shown in Appendix G. Lutein and total carotenoid were significantly and positively correlated with the latitude, longitude, and altitude of the ecoregion origins of the corresponding soybean accessions, with the highest correlations occurring in latitude (r=0.36*** and r=0.35***, respectively). The geographical distribution map of these components (Fig. 7) shows that the locations located in the plain of HR, including Hebei, Shanxi, Shaanxi, Ningxia and some parts of Gansu, represent accessions with higher accumulations of carotenoids, attributed to their appropriate weather factors such as temperature and precipitation. Similarly, the Inner Mongolia, Heilongjiang, Jilin and Liaoning of the northern and northeastern regions are hotspots for accessions with adequate carotenoid concentrations, whereas accessions originating in the SR parts of China were adapted to high temperatures and produced less carotenoid contents. As shown in Appendix G, zeaxanthin showed negative (r=–0.12***) and positive (r=0.14***) correlations with longitude and altitude, respectively, but no relationship with latitude. The β-carotene showed extremely weak nonsignificant correlations. Similarly, chlorophylls had an extremely weak significant positive correlation with latitude and nonsignificant associations with longitude and altitude, indicating a comparative increment of chlorophyll contents in accessions from high latitude areas.

3.8. Focusing on the soybean accessions with higher contents of lutein and total carotenoid

From this study, 44 soybean accessions (one cultivar

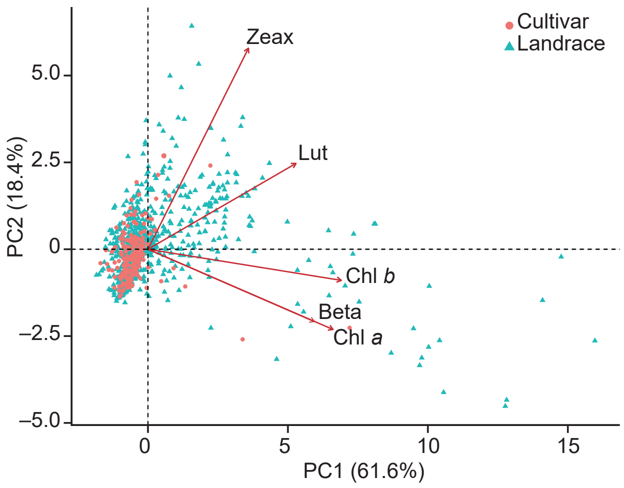


Fig. 4 Principal component analysis (PCA) for soybean seed carotenoid and chlorophyll (Chl) compositions based on accession types. Each point on the PCA-biplot represents a single accession and the color-coded accessions with different symbols represent the accession types (cultivars and landrace). Zeax, zeaxanthin; Lut, lutein; Beta, β-carotene.

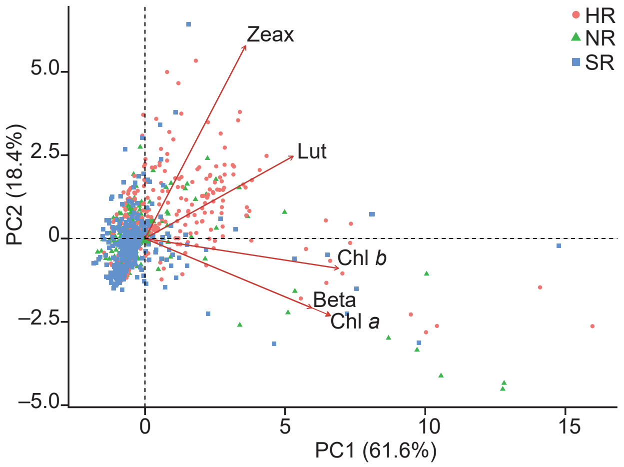


Fig. 5 Principal component analysis (PCA) for soybean seed carotenoid and chlorophyll (Chl) compositions based on ecoregions. Each point on the PCA-biplot represents a single accession and the color-coded accessions with different symbols represent the region of origin. HR, Huang–Huai–Hai Valley region; NR, northern region, SR, southern region; Zeax, zeaxanthin; Lut, lutein; Beta, β-carotene.

and 43 landraces) containing more than 27 μg g^{-1} lutein were screened as the high lutein soybeans (Appendix E), and they are suggested to be used in breeding programs to enhance soybean lutein. In addition, we identified 18 elite soybean accessions with total carotenoid contents of greater than 33.00 μg g^{-1}; among which, 12 were from HR, four were from NR and the remaining two were of SR origin (Appendix F). Surprisingly, all 18 of them were landraces.

This study also identified 64 accessions with total carotenoid concentrations from 25.00–33.00 μg g^{-1}, 118 accessions with total carotenoid concentrations from 17.00–25.00 μg g^{-1}, 472 accessions ranging from 9.00–17.00 μg g^{-1}, and 495 accessions with less than 9.00 μg g^{-1}. Interestingly, the first 18 top accessions contained substantial concentrations of lutein, zeaxanthin and β-carotene components. As shown in Fig. 6, these components had strong associations with each other and with total carotenoids, which will help soybean breeders to enhance total carotenoid without affecting the other components. In summary, the accessions with higher or lower individual and total carotenoids will be used as sources of genetic material in soybean breeding programs, and thereby natural ingredients for food products of the modern food industries, with a focus on supplying a healthy and safer daily diet to consumers worldwide.

4. Discussion

4.1. Seed carotenoids and chlorophylls exhibit a wide range of variations in soybean accessions

The carotenoid and chlorophyll concentrations of soybean germplasm accessions evaluated in this study showed a wide range of variability. The total carotenoids showed wider ranges in the present study than in previously reported studies (Monma *et al.* 1994). The variability of carotenoid contents between this study and previous works is most likely due to the different sets of accessions and the presence of large-scale germplasm collections in our study, which helps to facilitate the selection of genotypes with higher accumulations of carotenoids for breeding programs.

Lutein is a key component of total carotenoids, and it contributes the greatest amount. In the present study, lutein accounted for 95.37% of the total carotenoid concentration, while β-carotene accounted for 2.81%. This is in agreement with previous study (Kim *et al.* 2012) which showed that lutein and β-carotene represented 96.6 and 2.5% of the total carotenoids, respectively. Some studies have analyzed the variability of lutein and found it to be the most abundant component of carotenoids in soybean seeds (Monma *et al.* 1994; Kanamaru *et al.* 2006; Lee *et al.* 2009), as well as other legumes, such as pea and chickpeas (Marles *et al.* 2012; Ashokkumar *et al.* 2015) and lentils (Zhang *et al.* 2014), suggesting that its higher concentration could contribute to the antioxidant potential of soybean, which ultimately maintains eye health and alleviates other health problems. Lutein was

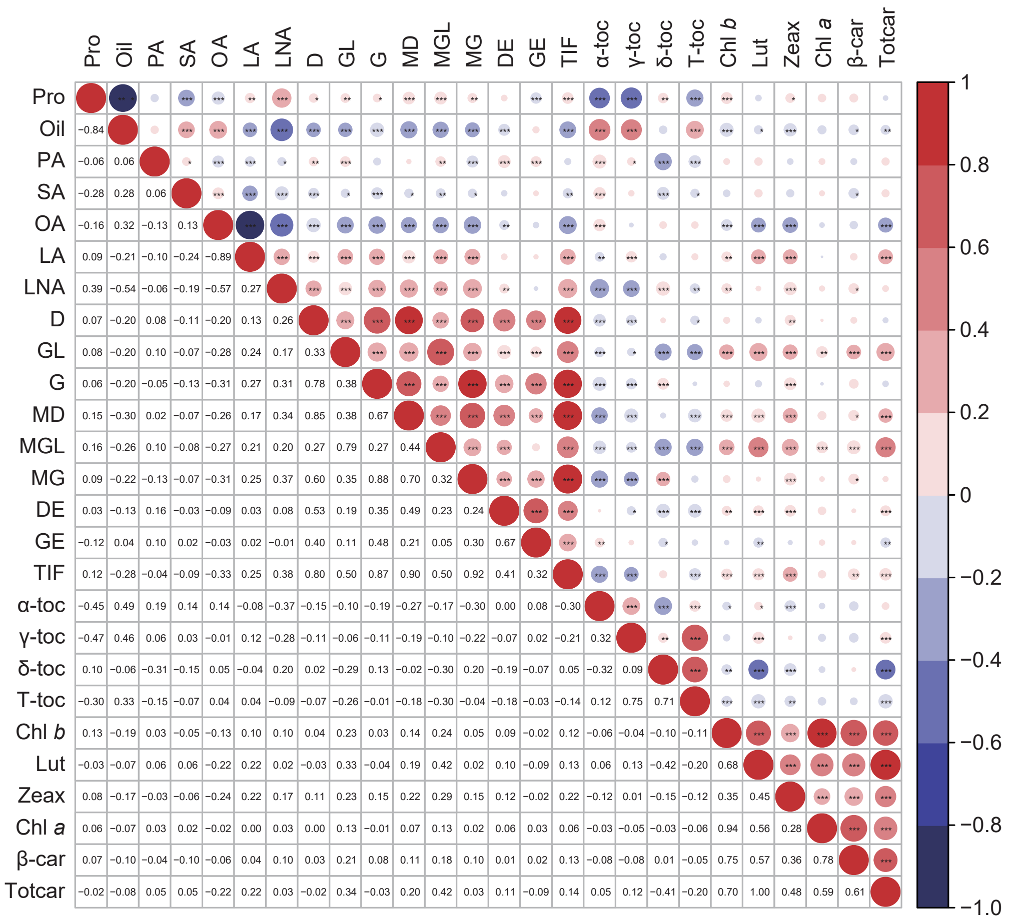
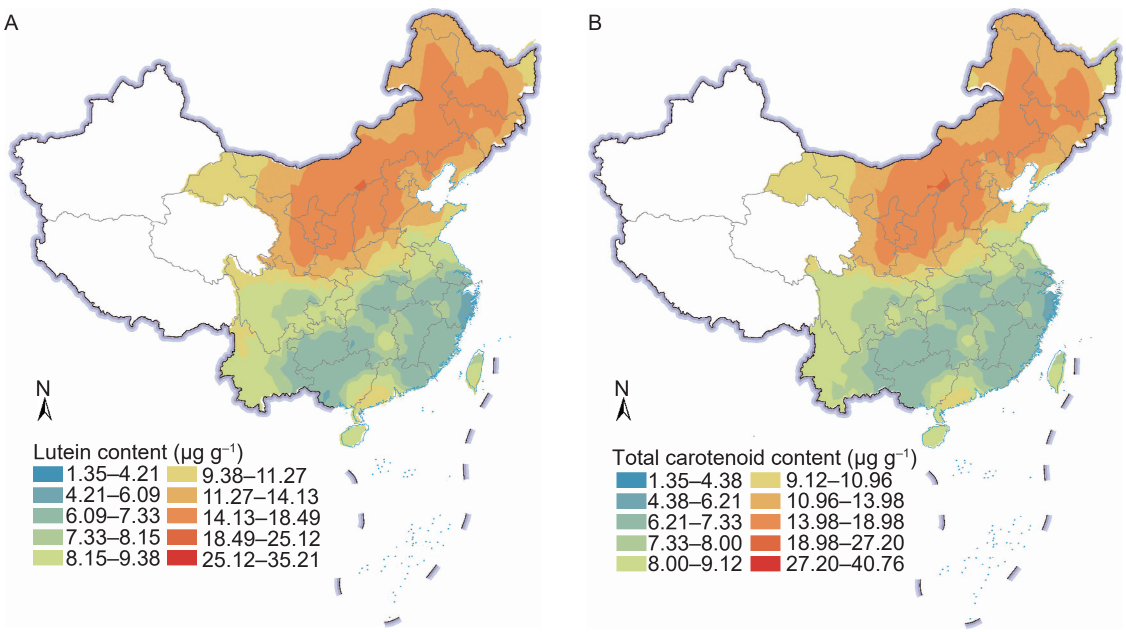


Fig. 6 Heatmap response of the Pearson's correlation coefficient (*r*) for soybean seed quality traits of soybean accessions. The lower diagonal represents the correlation coefficients; the upper diagonal plots reflect for significant differences indicated by * ($P<0.05$), ** ($P<0.01$), and *** ($P<0.001$); while empty plots represent no correlation between the traits. Pro, protein; PA, palmitic acid; SA, stearic acid; OA, oleic acid; LA, linoleic acid; LNA, linolenic acid; D, daidzin; GL, glycitin; G, genistin; MD, malonyldaidzin; MGL, malonylglycitin; MG, malonylgenistin; DE, daidzein; GE, genistein; TIF, total isoflavone; α-toc, α-tocopherol; γ-toc, γ-tocopherol; δ-toc, δ-tocopherol; T-toc, total tocopherol; Chl *b*, chlorophyll *b*; Lut, lutein; Zeax, zeaxanthin; Chl *a*, chlorophyll *a*; β-car, β-carotene; Totcar, total carotenoids.

found to dominate the carotenoids in all the accessions evaluated in our study. Previous studies also confirmed that lutein is found in the seeds of all soybean varieties tested, whereas the other components such as β-carotene, α-carotene, lycopene, zeaxanthin, β-cryptoxanthin were not found in all soybean seeds (Monma *et al.* 1994; Kim *et al.* 2012) or other legume seeds (Fernández-marín *et al.* 2017; Kan *et al.* 2018; Gebregziabher *et al.* 2022). Lutein was also found to be the most relatively stable component compared with the coefficient of variation (CV) values of the others, suggesting it has strong genetic control, which in turn shows that high lutein genotypes are likely to maintain their performance even when grown under diverse field conditions. Overall, the wide natural variation in carotenoids in this soybean panel can provide a good opportunity to identify genetic resources for soybean breeding.

4.2. The wide variations in seed carotenoids and chlorophyll concentrations may be attributed to both genetic and environmental factors

The variations in carotenoid and chlorophyll accumulations in soybean seeds could be due to the effects of plant genetic characteristics, as well as environmental and agronomic factors, as previously reported (Lee *et al.* 2009; Whent *et al.* 2009; Seguin *et al.* 2011). In this study, accession type, ecoregion of origin, seed cotyledon color and cultivation year significantly influenced the soybean seed carotenoid and chlorophyll contents. Higher contents of nutritionally valuable carotenoids were found in the landrace soybean accessions compared to the cultivars (Fig. 1) and the accession type PCA confirmed that landraces are more diverse than cultivars, resulting in excellent contents of carotenoids and



审图号 : GS 京 (2023) 0932 号

Fig. 7 The geographical distribution maps of lutein contents (A) and total carotenoid contents (B) in soybean seeds, with locations based on the accession's region of origin.

chlorophylls. Wide variations in the lutein and β-carotene contents among soybean cultivars and wild accessions were previously reported (Kanamaru *et al.* 2006), and that study found the concentrations of lutein and β-carotene in wild soybeans to be within the ranges of 5.80–32.8 and 0.9–2.8 μg g^{-1}, respectively, while the lutein contents of cultivated soybeans ranged from 1.60–4.80 μg g^{-1}, with no substantial amounts of β-carotene detected. In addition, Ebert *et al.* (2017) found a significantly higher content of lutein in mung bean (*Vigna radiata* L.) landraces (9.80 μg g^{-1}) in comparison with improved varieties (8.30 μg g^{-1}). Therefore, landraces are undisputedly genetically diverse and well adapted to climatic fluctuations and stressful environments (such as water stress, salinity, and high temperatures), and they show tolerance to biotic stresses and harbour genes and gene complexes for quality traits (Li *et al.* 2008; Dwivedi *et al.* 2016; Agyenim-Boateng *et al.* 2022). On the other hand, cultivars have low genetic diversity, are more vulnerable to extreme environmental conditions, and have been deliberately selected for desired traits in the course of domestication, resulting in low carotenoid accumulation (Fernández-Marín *et al.* 2014). Landraces are ideal genetic resources for exploring novel genetic variations that can overcome the challenges to crop production. They have been recognized as good sources of traits for local adaptation, stress tolerance, and seed nutrition (Dwivedi *et al.* 2016). Thus, the results of this study suggested that soybean landraces are reliable genetic resources for high carotenoid production and can contribute carotenoid producing alleles to the development of modern cultivars in soybean breeding programs.

Carotenoids and chlorophylls showed significant variations among the ecoregions of origin, so the geographical origin can be a major factor in the diversification of soybeans in China. Our accessions were examined both within and outside of their original ecoregion of origin. Taking this into account, HR accessions produced substantial concentrations of carotenoids and chlorophylls, followed by NR accessions (Fig. 2), and this was supported by the PCA-biplot (Fig. 5). These results signify that HR accessions tend to contain the highest levels of carotenoid and chlorophyll components, owing to the higher genetic diversity in HR accessions (Dong *et al.* 2004; Li *et al.* 2008), which in turn implies the presence of alleles conferring the adaptability of soybean accessions in diverse environments irrespective of geographical origin. Importantly, this finding could help breeders to broaden the range of germplasms utilized and encourage the exploitation of these valuable genetic resources in carotenoid-focused breeding programs. Consistent with our results, soybean seed quality traits, including soluble sugars (Qi *et al.* 2022), fatty acids (Abdelghany *et al.* 2020), isoflavone (Azam *et al.* 2020), tocopherols (Ghosh *et al.* 2021) and folates (Agyenim-Boateng *et al.* 2022), are affected by the ecoregion of origin. Generally, the major differences in carotenoid and chlorophyll compositions across the ecoregions are apparently related to environmental and climatic factors (temperature, rainfall, and light), geographical locations (latitude, and longitude), and growing seasons (Britz *et al.* 2008; Li *et al.* 2009). Therefore, it is important to note that nutritional composition (including carotenoids) varies greatly based on the geographical origin of soybean accessions, suggesting that soybean breeders should focus on the origin of the accessions when developing modern cultivars for a particular trait of interest, and this notion is supported by earlier studies (Zhang *et al.* 2018).

The present results showed that soybeans with a green cotyledon color contained considerably higher contents of carotenoids than accessions with a yellow cotyledon color, which is consistent with previous studies (Monma *et al.* 1994). Furthermore, green cotyledon peas (Ashokkumar *et al.* 2015) and green cotyledon chickpea cultivars (Rezaei *et al.* 2016) were reportedly 13-fold richer in β-carotene concentration, and had twice as many total carotenoids compared with yellow cotyledon cultivars, which was attributed to the greater expression of lycopene cycle genes in the green cotyledon accessions (Rezaei *et al.* 2016). A similar trend was also observed in the chlorophyll components. The mean concentrations of Chl *a* and Chl *b* in green cotyledon accessions were 9.80- and 9.08-fold higher than the yellow cotyledon-colored accessions, respectively, which is in agreement with previous studies (Monma *et al.* 1994). Moreover, similar findings were reported in field peas by Marles *et al.* (2012).

4.3. The correlations of seed carotenoids with other seed quality traits will facilitate soybean breeding for high quality

Understanding the correlations of lipophilic phytochemicals in soybeans may help to reveal the synergistic effects among each component, contribute to the development of soybean cultivars with multiple desirable traits, and ultimately produce functional foods related to the health benefits of soybeans. The carotenoid components had significantly positive correlations among them and were positively associated with chlorophyll components, which can be attributed to their common functions including light harvesting, energy transfer, photochemical redox reactions, and photoprotection (Brotosudarmo *et al.* 2018), as well as antioxidant activity (Gálvez *et al.* 2020). These positive correlations among the carotenoid components could also be due to the presence of the common intermediate precursor geranylgeranyl pyrophosphate (GGPP) in their biosynthetic pathways (Hirschberg 1999). Carotenoids also showed significantly positive or negative correlations with other seed quality traits (Fig. 6). Lutein and total carotenoid were negatively correlated with OA content but positively correlated with LA content, which is consistent with the results of previous studies (Whent *et al.* 2009). As shown in Fig. 6, lutein and total carotenoid were significantly and negatively associated with δ-tocopherol and total tocopherol, which agrees with the report of Whent *et al.* (2009) on soybean lutein and δ-tocopherol. On the other hand, lutein was significantly and positively correlated with α-tocopherol, indicating that it is possible to either simultaneously increase their contents or increase one of the two without lowering the content of the other in the seeds of progenies during breeding, confirming a previous report (Wang *et al.* 2007). It is also important to note that the correlation analysis between soybean seed quality traits will help breeders to identify the relationships that exist within the traits, and thereby serve to set a breeding program.

The geographical distribution map of the lutein and total carotenoid contents revealed a decreasing trend towards the south (low latitude). This is consistent with the study of Song *et al.* (2018), which found that the majority of bioactive components showed declining trends from the north (high latitude) to the south (low latitude). The correlation of lutein and total carotenoid with latitude indicated that the concentrations of these components increase with higher latitudes and low temperatures, which characterize the conditions in the ecoregions of HR and NR. China has abundant and diverse germplasm resources that are incorporated into a complex cropping system. This provides a good opportunity to evaluate the quality trait variations across various ecoregions with different climatic conditions. Obviously, geographical differentiation and photo-thermal conditions can play a significant role in the genetic differentiation of soybeans, which in turn leads to seed quality trait variations (Zhang *et al.* 2020). This variation allows researchers to have the possibility of selecting highly adapted and elite candidate accessions with higher carotenoids across various locations, and is beneficial for soybean quality breeding improvement.

5. Conclusion

Carotenoids are considered to be very versatile secondary metabolites that are beneficial to human health. This is the first study on a comprehensive variability analysis of the carotenoid and chlorophyll components across a wide range of diverse soybean accessions. The results showed significantly wide variations among the carotenoid contents, with total carotenoid content variation of over 30-fold and some elite accessions with total carotenoid concentrations >33 μg g^{-1}. Moreover, the results showed that genotype, planting year, accession type, cotyledon color and ecoregion of origin significantly influenced the accumulation of carotenoids and chlorophylls. The landrace accessions contained total carotenoid contents 1.20-fold higher than the cultivars. The geographical origin of soybean accessions influenced the carotenoid and chlorophyll contents because of their various environmental and climatic conditions. Soybean seed carotenoid contents were geographically distributed across the main ecoregions of China, with a decreasing trend towards south (low latitude). The components of carotenoids and chlorophylls were significantly and positively correlated with each other. Moreover, they showed significant correlations with other quality traits, which can help to set breeding strategies for enhancing soybean carotenoids without affecting other components. In general, by analyzing a wide range of soybean accessions, we demonstrated that soybeans can be reliable sources and candidates for carotenoid biofortification, and the novel germplasms identified should be preferably selected in breeding specialty soybeans for soy foods, animal feed and other industrial products to satisfy consumer demands.

Acknowledgements

We would like to thank Dr. Qiu Lijuan from Institute of Crop Sciences, Chinese Academy of Agricultural Sciences (CAAS), for providing us with all of the soybean accessions. This work was financially supported by the National Natural Science Foundation of China (32161143033 and 32001574) and the Agricultural Science and Technology Innovation Program of CAAS (2060203-2).

Declaration of competing interest

The authors declare that they have no conflict of interest.

Appendices associated with this paper are available on https://doi.org/10.1016/j.jia.2022.10.011

References

Abdelghany A M, Zhang S, Azam M, Shaibu A S, Feng Y, Li Y, Tian Y, Hong H, Li B, Sun J. 2020. Profiling of seed fatty acid composition in 1025 Chinese soybean accessions from diverse ecoregions. *The Crop Journal*, **8**, 635–644.

Agyenim-Boateng K G, Zhang S, Islam S, Gu Y, Li B, Azam M, Abdelghany A M, Qi J, Ghosh S, Shaibu A S, Gebregziabher B S, Feng Y, Li J, Li Y, Zhang C, Qiu L, Liu Z, Liang Q, Sun J. 2022. Profiling of naturally occurring folates in a diverse soybean germplasm by HPLC-MS/MS. *Food Chemistry*, **384**, 132520.

Ashokkumar K, Diapari M, Jha A B, Tar'an B, Arganosa G, Warkentin T D. 2015. Genetic diversity of nutritionally important carotenoids in 94 pea and 121 chickpea accessions. *Journal of Food Composition and Analysis*, **43**, 49–60.

Azam M, Zhang S, Abdelghany A M, Shaibu A S, Feng Y, Li Y, Tian Y, Hong H, Li B, Sun J. 2020. Seed isoflavone profiling of 1168 soybean accessions from major growing ecoregions in China. *Food Research International*, **130**, 108957.

Britz S J, Kremer D F, Kenworthy W J. 2008. Tocopherols in soybean seeds: Genetic variation and environmental effects in field-grown crops. *Journal of the American Oil Chemists' Society*, **85**, 931–936.

Brotosudarmo T H P, Limantara L, Chandra R D, Heriyanto. 2018. Chloroplast pigments: Structure, function, assembly and characterization. In: Ratnadewi D, Hamim H, eds., *Plant Growth and Regulation – Alterations to Sustain Unfavorable Conditions*. IntechOpen, London, UK. pp. 44–59.

Chatterjee C, Gleddie S, Xiao C W. 2018. Soybean bioactive peptides and their functional properties. *Nutrients*, **10**, 1211.

Deng J C, Li X M, Xiao X L, Wu H J, Yang C Q, Long X Y, Zhang Q H, Iqbal N, Wang X C, Yong T W, Du J B, Yang F, Liu W G, Zhang J, Wu X L, Wu Y S, Yang W Y, Liu J. 2022. Field mold stress induced catabolism of storage reserves in soybean seed and the resulting deterioration of seed quality in the field. *Journal of Integrative Agriculture*, **21**, 336–350.

Dias M G, Olmedilla-alonso B, Hornero-mendez D, Osorio C, Vargas-murga L, Meléndez-martínez A J. 2018. A comprehensive database of carotenoid contents in Ibero-American foods. A valuable tool in the context of functional foods and the establishment of recommended intakes of bioactive. *Journal of Agricultural and Food Chemistry*, **66**, 5055–5107.

Dong Y S, Zhao L M, Liu B, Wang Z W, Jin Z Q, Sun H. 2004. The genetic diversity of cultivated soybean grown in China. *Theoretical and Applied Genetics*, **108**, 931–936.

Dwivedi S L, Ceccarelli S, Blair M W, Upadhyaya H D, Are A K, Ortiz R. 2016. Landrace germplasm for improving yield and abiotic stress adaptation. *Trends in Plant Science*, **21**, 31–42.

Ebert A W, Chang C H, Yan M R, Yang R Y. 2017. Nutritional composition of mungbean and soybean sprouts compared to their adult growth stage. *Food Chemistry*, **237**, 15–22.

Fernández-marín B, Míguez F, Méndez-fernández L, Agut A, Becerril J M, García-Plazaola J I, Kranner I, Colville L. 2017. Seed carotenoid and tocochromanol composition of wild fabaceae species is shaped by phylogeny and ecological factors. *Frontiers in Plant Science*, **8**, 1428.

Fernández-marín B, Milla R, Martín-Robles N, Arc E, Kranner I, Becerril J M, García-Plazaola J I. 2014. Side-effects of domestication: Cultivated legume seeds contain similar tocopherols and fatty acids but less carotenoids than their wild counterparts. *BMC Plant Biology*, **14**, 1599.

Ferruzzi M G, Blakeslee J. 2007. Digestion, absorption, and cancer preventative activity of dietary chlorophyll derivatives. *Nutrition Research*, **27**, 1–12.

Gálvez A P, Viera I, Roca M. 2020. Carotenoids and chlorophylls as antioxidants. *Antioxidants*, **9**, 505.

Gebregziabher B S, Zhang S, Ghosh S, Shaibu A S, Azam M, Abdelghany A M, Qi J, Agyenim-Boateng K G, Htway H T P, Feng Y, Ma C, Li Y, Li J, Li B, Qiu L, Sun J. 2022. Origin, maturity group and seed coat color influence carotenoid and chlorophyll concentrations in soybean seeds. *Plants*, **11**, 848.

Gebregziabher B S, Zhang S, Qi J, Azam M, Ghosh S, Feng Y, Huai Y, Li J, Li B, Sun J. 2021. Simultaneous determination of carotenoids and chlorophylls by the HPLC-UV-VIS method in soybean seeds. *Agronomy*, **11**, 758.

Ghosh S, Zhang S, Azam M, Qi J, Abdelghany A M, Shaibu A S, Gebregziabher B S, Feng Y, Huai Y, Htway H T P, Agyenim-Boateng K G, Liu Y, Feng H, Li J, Song W, Li B, Sun J. 2021. Seed tocopherol assessment and geographical distribution of 1 151 Chinese soybean accessions from diverse ecoregions. *Journal of Food Composition and Analysis*, **100**, 103932.

Guan R, Qu Y, Guo Y, Yu L, Liu Y, Jiang J, Chen J, Ren Y, Liu G, Tian L, Jin L, Liu Z, Hong H, Chang R, Gilliham M, Qiu L. 2014. Salinity tolerance in soybean is modulated by natural variation in *GmSALT3*. *The Plant Journal*, **80**, 937–950.

Hirschberg J. 1999. Production of high-value compounds: Carotenoids and vitamin E. *Current Opinion in Biotechnology*, **10**, 186–191.

Huang J, Guo N, Li Y, Sun J, Hu G, Zhang H, Li Y, Zhang X. 2016. Phenotypic evaluation and genetic dissection of resistance to *Phytophthora sojae* in the Chinese soybean mini core collection. *BMC Genetics*, **17**, 85.

Kaga A, Shimizu T, Watanabe S, Tsubokura Y, Katayose Y, Harada K, Vaughan D A, Tomooka N. 2012. Evaluation of soybean germplasm conserved in NIAS genebank and development of mini core collections. *Breeding Science*, **61**, 566–592.

Kan L, Nie S, Hu J, Wang S, Bai Z, Wang J, Zhou Y, Jiang J, Zeng Q, Song K. 2018. Comparative study on the chemical composition, anthocyanins, tocopherols and carotenoids of selected legumes. *Food Chemistry*, **260**, 317–326.

Kanamaru K, Wang S, Abe J, Yamada T, Kitamura K. 2006. Identification and characterization of wild soybean (*Glycine soja* Sieb. Et Zecc.) strains with high lutein content. *Breeding Science*, **56**, 231–234.

Kim M, Kim J K, Kim H J, Pak J H, Lee J, Kim D, Choi K, Jung H W, Lee J, Chung Y, Ha S. 2012. Genetic modification of the soybean to enhance the β-carotene content through seed-specific expression. *PLoS ONE*, **7**, e48287.

Lee J, Hwang Y S, Kim S T, Yoon W B, Han W Y, Kang I K, Choung M G. 2017. Seed coat color and seed weight contribute differential responses of targeted metabolites in soybean seeds. *Food Chemistry*, **214**, 248–258.

Lee J D, Shannon J G, So Y S, Sleper D A, Nelson R L, Lee J H, Choung M G. 2009. Environmental effects on lutein content and relationship of lutein and other seed components in soybean. *Plant Breeding*, **128**, 97–100.

Li X, Wang K, Jia J. 2009. Genetic diversity and differentiation of Chinese wild soybean germplasm (*G. soja* Sieb. & Zuce.) in geographical scale revealed by SSR markers. *Plant Breeding*, **128**, 658–664.

Li Y, Guan R, Liu Z, Ma Y, Wang L, Li L, Lin F, Luan W, Chen P, Yan Z, Guan Y, Zhu L, Ning X, Smulders M J M, Li W, Piao R, Cui Y, Yu Z, Guan M, Chang R, *et al.* 2008. Genetic structure and diversity of cultivated soybean (*Glycine max* (L.) Merr.) landraces in China. *Theoretical and Applied Genetics*, **117**, 857–871.

Marles M A S, Warkentin D, Bett K E. 2012. Genotypic abundance of carotenoids and polyphenolics in the hull of field pea (*Pisum sativum* L.). *Journal of the Science of Food and Agriculture*, **93**, 463–470.

Monma M, Ito M, Saito M, Chikuni K. 1994. Carotenoid components in soybean seeds varying with seed color and maturation stage. *Bioscience*, *Biotechnology and Biochemistry*, **58**, 926–930.

Qi J, Zhang S, Azam M, Shaibu A S, Abdelghany A M, Feng Y, Huai Y, Feng H, Liu Y, Ma C, Gebregziabher B S, Ghosh S, Li J, Li B, Qiu L, Sun J. 2022. Profiling seed soluble sugar compositions in 1164 Chinese soybean accessions from major growing ecoregions. *The Crop Journal*, **10**, 1825–1831.

Qiu L J, Chen P Y, Liu Z X, Li Y H, Guan R X, Wang L H, Chang R Z. 2011. The worldwide utilization of the Chinese soybean germplasm collection. *Plant Genetic Resources-Characterization and Utilization*, **9**, 109–122.

Rao A V, Rao L G. 2007. Carotenoids and human health. *Pharmacological Research*, **55**, 207–216.

Rezaei M K, Deokar A, Tar B. 2016. Identification and expression analysis of candidate genes involved in carotenoid biosynthesis in chickpea seeds. *Fronteirs in Plant Science*, **7**, 1867.

Schoefs B. 2002. Chlorophyll and carotenoid analysis in food products. Properties of the pigments and methods of analysis. *Trends in Food Science and Technology*, **13**, 361–371.

Seguin P, Tremblay G, Pageau D, Liu W, Turcotte P. 2011. Soybean lutein concentration: Impact of crop management and genotypes. *Crop Science*, **51**, 1151–1160.

Singh B P, Yadav D, Vij S. 2019. Soybean bioactive molecules: Current trend and future prospective. In: Mérillon J M, Ramawat K G, eds., *Bioactive Molecules in Food*. Springer International Publishing, Cham. pp. 267–294.

Song W, Yang R, Wu T, Wu C, Sun S, Zhang S, Jiang B, Tian S, Liu X, Han T. 2016. Analyzing the effects of climate factors on soybean protein, oil contents, and composition by extensive and high-density sampling in China. *Journal of Agricultural and Food Chemistry*, **64**, 4121–4130.

Song W, Yang R, Yang X, Sun S, Mentreddy S R, Jiang B, Wu T, Tian S, Sapey E, Wu C, Hou W, Ren G, Han T. 2018. Spatial differences in soybean bioactive components across

China and their influence by weather factors. *The Crop Journal*, **6**, 659–668.
Sun J, Sun B, Han F, Yan S, Yang H, Akio K. 2011. Rapid HPLC method for etermination of 12 isoflavone components in soybean seeds. *Agricultural Sciences in China*, **10**, 70–77.
Teng W L, Feng W J, Zhang J Y, Xia N, Guo J, Li W, Wu D P, Zhao X, Han Y P. 2017. Identification of quantitative trait loci underlying lutein content in soybean seeds across multiple environments. *Journal of Agricultural Science*, **155**, 1263–1271.
Wang L X, Lin F Y, Li L H, LI W, Yan Z, Luan W J, Piao R H, Guan Y, Ning X C, Zhu L, Ma Y S, Dong Z M, Zhang H Y, Zhang Y Q, Guan R X, Li Y H, Liu Z X, Chang R Z, Qiu L J. 2016. Genetic diversity center of cultivated soybean (*Glycine max*) in China — New insight and evidence for the diversity center of Chinese cultivated soybean. *Journal of Integrative Agriculture*, **15**, 2481–2487.
Wang S, Kanamaru K, Li W, Abe J, Yamada T, Kitamura K. 2007. Simultaneous accumulation of high contents of α-tocopherol and lutein is possible in seeds of soybean (*Glycine max* (L.) Merr.). *Breeding Science*, **57**, 297–304.
Whent M, Hao J, Slavin M, Zhou M, Song J, Kenworthy W I, Yu L L. 2009. Effect of genotype, environment, and their interaction on chemical composition and antioxidant properties of low-linolenic soybeans grown in Maryland. *Journal of Agricultural and Food Chemistry*, **57**, 10163–10174.
Zhang B, Deng Z, Tang Y, Chen P, Liu R, Ramdath D D, Liu Q, Hernandez M, Tsao R. 2014. Fatty acid, carotenoid and tocopherol compositions of 20 Canadian lentil cultivars and synergistic contribution to antioxidant activities. *Food Chemistry*, **161**, 296–304.
Zhang L X, Liu W, Tsegaw M, Xu X, Qi Y P, Sapey E, Liu L P, Wu T T, Sun S, Han T F. 2020. Principles and practices of the photo-thermal adaptability improvement in soybean. *Journal of Integrative Agriculture*, **19**, 295–310.
Zhang J, Wang X, Lu Y, Bhusal S J, Song Q, Cregan P B, Yen Y, Brown M, Jiang G L. 2018. Genome-wide scan for seed composition provides insights into soybean quality improvement and the impacts of domestication and breeding. *Molecular Plant*, **11**, 460–472.

Executive Editor-in-Chief ZHANG Xue-yong
Managing Editor WANG Ning