

Temporospatial Characterization of Nutritional and Bioactive Components of Soybean Cultivars in China

Tingting Wu¹ · Yang Yao¹ · Shi Sun¹ · Caijie Wang¹ · Hongchang Jia² · Weiqun Man³ · Lianshun Fu⁴ · Wenwen Song¹ · Cunxiang Wu¹ · Bingjun Jiang¹ · Wensheng Hou¹ · Guixing Ren¹ · Tianfu Han¹

Received: 4 May 2016 / Revised: 19 September 2016 / Accepted: 16 October 2016 / Published online: 15 November 2016
© AOCS 2016

Abstract Nutritional and bioactive constituents, apart from protein and oil, have many beneficial functions and are of great importance in determining the quality and nutritional value of soybeans [*Glycine max* (L.) Merr.]. However, the temporospatial trends in soybean constituents have not been determined. The objective of this study was thus to identify the temporospatial trends in seed constituents among 89 widely grown soybean cultivars over a period of nine decades (1923–2007) in four soybean-producing regions of China. A combination of analytical techniques was used to quantify eight categories of 35 seed constituents. The results include the following: (1) The temporal pattern (the genetic change) after nine decades of breeding was characterized by significantly increasing trends in the

concentration of isoflavones, tocopherol, starch, sucrose, Ca, Mn, phospholipids and lectin, and decreasing trends in Zn, Cu and lutein concentrations. (2) The spatial distribution showed higher stachyose, Zn, isoflavone, saponin and lectin content in the cultivars from lower latitudes versus higher latitudes, and the content of tocopherol, carotenoid, oligosaccharides and phospholipids exhibited the reverse pattern. (3) According to principal component analysis, carotenoid, Fe and Mn, saponin components and glycitin were the most important contributors to the total variation across the 89 cultivars. These results provide us with a comprehensive profile for manipulating chemical constituents and improving various functional ingredients through breeding.

T. Wu, Y. Yao and S. Sun contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s11746-016-2908-4) contains supplementary material, which is available to authorized users.

✉ Guixing Ren
renguixing@caas.cn

✉ Tianfu Han
hantianfu@caas.cn

¹ MOA Key Laboratory of Soybean Biology (Beijing), Institute of Crop Science, the Chinese Academy of Agricultural Sciences, 12 Zhongguancun South Street, 100081 Beijing, China

² Heihe Branch of Heilongjiang Academy of Agricultural Sciences, 164300 Heihe, Heilongjiang, China

³ Soybean Research Institute, Heilongjiang Academy of Agricultural Sciences, 150086 Harbin, Heilongjiang, China

⁴ Tieling Academy of Agricultural Sciences, 112616 Tieling, Liaoning, China

Keywords Soybean · Nutritional · Bioactive · Widely grown · Breeding

Abbreviations

D	Daidzin
GL	Glycitin
G	Genistin
MD	Malonyldaidzin
MGL	Malonylglycitin
Sap I	Saponin I
Sap II	Saponin II
Sap α g	Saponin α g
Sap β g	Saponin β g
Sap β α	Saponin β α
BBI	Bowman–Birk inhibitors
PC	Phosphatidylcholine
PE	Phosphatidylethanolamine
PI	Phosphatidylinositol
α -Toc	α -Tocopherol
β + γ -Toc	β + γ -Tocopherol

δ -Toc	δ -Tocopherol
BHT	Butylated hydroxytoluene

Introduction

Soybeans are the predominant source of plant protein and vegetable oil in the human diet [1]. In addition, soybeans contain health-beneficial phytochemicals, including lipid-soluble vitamins, mineral elements, oligosaccharides, saponins, phospholipids and isoflavones [2–7]. Among these, isoflavones, saponins, carotenoid and tocopherol have important functions in protecting against cancer, cardiovascular disease and oxidative deterioration [8–11]. Moreover, soybean seeds can provide 15 minerals essential to human health [12], and thus contribute to alleviating global malnutrition. These multi-nutrient characteristics and high bioactivity have led to an increased interest in soybean cultivation and soy-based products. Understanding the nutritional and bioactive constituents of soybean seeds can help in meeting the demands of the food market and in maximizing their health and therapeutic value.

Soybeans have been cultivated in China for more than 5000 years [13]. China also has a long history of modern soybean breeding, which began in 1913 [14]. A total of 2062 soybean cultivars have been developed over the past 100 years (1913–2014) (Shuming Wang, personal communication, 2015), and more than 100 of these are widely grown throughout these areas [15]. These widely grown soybean cultivars have adapted to the environment and have excellent overall performance [15]. Changes in soybean lutein, tocopherol and isoflavone content with release date due to genetic breeding have been identified in previous research in the US [16], Canada [6, 17] and China [18]. However, changes in the content of other soybean seed constituents, such as saponins, phospholipids, oligosaccharides and mineral elements, during the nearly 100 years of breeding have rarely been documented. This study is the first to describe temporal trends in seed constituents among these widely grown Chinese soybean cultivars bred over an extended time frame.

The northeast region and the Huang-Huai-Hai Rivers Valley are the first and second largest soybean-producing regions in China, respectively, constituting more than 83% of the national production area [19]. The content of various constituents is affected by regional climatic and environmental conditions [20, 21]. Previous studies have reported that the spatial distribution of protein, oil, fatty acid and amino acid content varied among the different regions [3, 22]. However, the spatial distribution of other soybean seed constituents across a wide-ranging area of the main soybean producing-region in China has not been well documented. In the current study, we sought to determine whether soybean

cultivars that have adapted to a particular region were more likely to produce soybean seeds with higher or lower levels of various seed component constituents. Correlations among seed constituents can provide soybean breeders with insight into the best ways to increase concentrations of desirable components, while reducing concentrations of undesirable components. In addition, among seed constituents, some substances can explain a greater proportion of variation in seed composition than others, and are more important in distinguishing different genotypes.

The objectives of this study were (1) to explore the evolutionary trends in concentrations of seed constituents over the nine decades of breeding, (2) to compare concentrations of seed constituents among the four main soybean-producing regions in China, (3) to characterize the correlations among seed constituents of widely grown soybean cultivars, and (4) to identify the main soybean constituents contributing to the variation among the 89 cultivars.

Materials and Methods

Plant Materials and Field Experiments

The experiments encompassed 89 widely grown soybean (*Glycine max* (L.) Merr.) cultivars released between 1923 and 2007 [15]. These cultivars originated from four regions in China: the North Heilongjiang Province (NH, 17 cultivars), the Midsouth Heilongjiang Province (MSH, 23 cultivars), the Jilin and Liaoning provinces (JL, 24 cultivars), and the Huang-Huai-Hai Rivers Valley (HH, 25 cultivars); these regions cover a wide geographic area between N32° and N50° [19]. The maturity group (MG) of these cultivars is 00-IV. The soybean cultivars, their respective regions of origin, maturity groups and their years of release are presented in Table 1. The soybean seeds were obtained from the breeders or from the National Genebank at the Institute of Crop Science, Chinese Academy of Agricultural Sciences (CAAS).

Field trials were conducted in 2011 at four locations, representing the production environments of each region (Heihe, Harbin, Tieling and Beijing in the NH, MSH, JL and HH regions, respectively). The planting locations (latitude and longitude) and sowing dates are shown in Table 2. At each site, the plants were grown in a randomized complete block design with three replications. Each plot consisted of six rows that were 5 m long, with an inter-row spacing of 0.65 m. The plants were thinned to a uniform stand of 30 plants m⁻² after emergence. A basal fertilizer was applied [150 kg ha⁻¹ (NH₄) H₂PO₄, 75 kg ha⁻¹ urea, and 40 kg ha⁻¹ K₂SO₄] prior to planting. Weeds and pests were routinely controlled. Plants were hand-harvested after the growth stage R8 [23]. A 500-g

Table 1 Regions of origin and release years of widely grown soybean cultivars

Region	Cultivar	Year of release	Maturity group	Region	Cultivar	Year of release	Maturity group
NH ^a	Kexi283	1956	0	JL	Jilin3	1963	II
NH	Heilongjiang41	1958	0	JL	Jilin6	1963	I
NH	Fengshou10	1966	00	JL	Tiefeng3	1967	I
NH	Heihe3	1966	00	JL	Tiefeng18	1973	II
NH	Heihe51	1967	00	JL	Jilin13	1976	II
NH	Fengshou12	1969	0	JL	Jiunong9	1976	I
NH	Fengshou17	1977	00	JL	Kaiyu3	1976	I
NH	Beifeng2	1983	000	JL	Kaiyu8	1980	II
NH	Fengshou19	1985	0	JL	Jilin20	1985	II
NH	Heihe9	1990	0	JL	Changnong4	1985	I
NH	Beifeng9	1995	0	JL	Kaiyu10	1989	II
NH	Heihe19	1998	0	JL	Changnong5	1990	I
NH	Heihe27	2002	0	JL	Jilin30	1993	II
NH	Fengshou24	2003	0	JL	Tiefeng29	1997	III
NH	Heihe38	2005	0	JL	Jiunong22	1999	III
NH	Beidou5	2006	0	JL	Jilin47	1999	II
NH	Heihe43	2007	0	JL	Tiefeng31	2001	III
MSH ^a	Jinyuan2	1941	0	JL	Jiyu57	2001	I
MSH	Zhi2	1958	00	HH ^a	Pamanqing	1929	IV
MSH	Jingshanpu	1958	0	HH	Pingdingshi	1940	IV
MSH	Dongnong4	1959	0	HH	Niumaohuang	1950	IV
MSH	Hejiao8	1962	0	HH	Pingdinghuang	1951	IV
MSH	Hejiao6	1963	I	HH	Xinhuangdou	1952	IV
MSH	Heinong16	1970	0	HH	Juxuan23	1963	IV
MSH	Suinong3	1973	0	HH	Qihuang10	1966	III
MSH	Hefeng22	1974	0	HH	Fengshouhuang	1970	IV
MSH	Heinong26	1975	I	HH	Henanzaofeng1	1971	IV
MSH	Suinong4	1981	0	HH	Wenfeng7	1971	IV
MSH	Hefeng25	1984	0	HH	Yanhuang1	1973	IV
MSH	Heinong33	1988	0	HH	Yuejing5	1975	IV
MSH	Suinong8	1989	0	HH	Youbian30	1983	IV
MSH	Heinong35	1990	0	HH	Jidou4	1984	III
MSH	Suinong10	1994	I	HH	Ludou4	1985	III
MSH	Hefeng35	1994	0	HH	Yudou2	1985	IV
MSH	Suinong14	1996	I	HH	Zhongdou19	1987	IV
MSH	Heinong43	2002	0	HH	Kefeng6	1988	III
MSH	Heinong44	2002	0	HH	Jidou7	1990	III
MSH	Hefeng45	2003	0	HH	Ludou11	1995	III
MSH	Hefeng50	2006	0	HH	Yudou22	1997	IV
JL ^a	Huangbaozhu	1923	II	HH	Zhonghuang13	2001	III
JL	Fushou	1930s	I	HH	Zheng92116	2001	IV
JL	Xiaojinhuang	1941	I	HH	Jidou12	2001	III
JL	Fengdihuang	1943	II	HH	Handou5	2004	III
JL	Jiti1	1956	I				
JL	Jiti5	1956	0				

^a HH Huang-Huai-Hai Rivers Valley region, JL Jilin and Liaoning region, MSH Midsouth Heilongjiang region, NH Northern Heilongjiang region

Table 2 The latitude, longitude and sowing dates of experimental sites in four regions

Region	Location	Latitude (°)	Longitude (°)	Sowing date
NH ^a	Heihe, Heilongjiang	50.14 N	127.29 E	10 May
MSH ^a	Harbin, Heilongjiang	45.44 N	126.36 E	13 May
JL ^a	Tieling, Liaoning	42.28 N	123.51 E	28 Apr
HH ^a	Shunyi, Beijing	40.13 N	116.65 E	14 June

^a HH Huang-Huai-Hai Rivers Valley region, JL Jilin and Liaoning province S, MSH Midsouth Heilongjiang province, NH Northern Heilongjiang province

seed sample was collected from each plot. The seeds were ground and passed through a 60-mesh sieve in a FOSS 1093 Cyclotec 126 sample mill (FOSS Tecator, Höganäs, Sweden).

Seed Composition Analysis

Thirty-five components of soybean seed constituents consisting of eight groups were analyzed, including seed nutritional components [saccharides, fiber, mineral elements and vitamins (tocopherol and carotenoid)] and seed bioactive constituents (isoflavones, saponins, phospholipids and lectins). The methodology for analyzing each seed constituent is described below.

Starch Analysis

The starch content was determined according to the method described in the National Standard for Crop Seed Starch Measurement of China [24].

Sugar Analysis

Sugar content was determined according to the method described in the National Standard for Soybean Sugar Measurement [25]. Sugar was extracted by adding 1.0 g soybean flour to 10 mL 70% ethanol solution, with shaking at 200 rpm for 10 h and then centrifugation at 4000 rpm for 10 min. The clear aliquot was filtered through a 0.45- μ m filter. The analysis was performed using an Alltech Prevail Carbohydrate ES (4.6 \times 250 mm, 5 μ m; Alltech, Connecticut, USA). The mobile phases were solvent A (acetonitrile solution) and solvent B (water) (solvent A/B = 60/40), with a flow rate of 1 mL/min. The standards for sucrose, raffinose and stachyose were purchased from Sigma-Aldrich (Sigma-Aldrich Corp., St. Louis, MO, USA). The sucrose, raffinose, and stachyose content was obtained from standard curves and converted to a percentage of dry base.

Crude Fiber Analysis

A 1.0-g sample of flour was de-fatted with diethyl ether three times (each time for 6 min) and then analyzed using an ANKOM fiber analyzer (ANKOM Technology, Macedon, NY, USA). Sodium sulfite was added, and the sample was washed with water to neutral pH, after which NaOH was added and washed with water to neutral. Next, the sample was washed with acetone, dried to a constant weight at 130 °C in a drying oven, and then ashed to a constant weight at 525 °C in a muffle furnace.

Mineral Elements Analysis

The mineral elements [copper (Cu), iron (Fe), manganese (Mn), calcium (Ca), and magnesium (Mg)] were determined using flame atomic absorption spectrometry [26].

Tocopherol and Carotenoid Analysis

Extraction and quantification of tocopherol was achieved using a previously described method, with minor modifications [27]. A 1.0-g sample of flour was suspended in 6 mL of hexane and 0.2 mL butylated hydroxytoluene (BHT; 0.05 g/L), subjected to ultrasonication for 5 min, mixed by shaking for 20 min in a 60 °C water bath, and centrifuged at 4600 rpm for 10 min. The clear aliquot was filtered through a 0.45- μ m filter. Analysis was carried out using a Waters 2695 liquid chromatograph (Waters Corporation, Milford, MA, USA) with a Phenomenex Luna silica column (5 μ m, 250 \times 4.6 mm; Phenomenex, Torrance, CA, USA). The mobile phases included hexane, isopropanol, acetonitrile and acetic acid (97.6/0.8/0.8/0.8), with a flow rate of 0.7 mL/min and an injection volume of 10 μ L. Standards of α -tocopherol (α -Toc), β -tocopherol (β -Toc), γ -tocopherol (γ -Toc) and δ -tocopherol (δ -Toc) were purchased from Sigma-Aldrich (Sigma-Aldrich Corp., St. Louis, MO, USA). The content of α -Toc, β -Toc, γ -Toc, δ -Toc and total tocopherol was calculated using standard curves and expressed in units of mg/g.

The extraction and quantification of carotenoid was conducted using high-performance liquid chromatography (HPLC), as previously described, with minor modifications [28]. Carotenoid was extracted by suspending a 1.0-g sample of flour in 10 mL acetone (acetone/methanol = 7/3) and 0.15 mL BHT (0.05 g/L) for 5 min, and subjecting to ultrasonication for 30 min and centrifugation at 4000 rpm for 10 min. The clear aliquot was filtered through a 0.45- μ m filter. The analysis was conducted using a Shimadzu LC-20A (Shimadzu Corporation, Kyoto, Japan) with YMC Carotenoid (5 μ m, 250 \times 4.6 mm). Solvent A in the mobile phase was methyl tert-butyl ether and triethylamine (1000:0.5), and Solvent B was acetonitrile, methanol and

triethylamine (750:250:0.5), with an injection volume of 30 μ L. The solvent system (% solvent A/% solvent B) was run for 0–20 min (0–55/100–4), 20–25 min (55–0/45–100) and 25–35 min (0/100). Standards of lutein, β -carotene, and carotenoid were purchased from Sigma-Aldrich (Sigma-Aldrich Corp., St. Louis, MO, USA). Lutein, β -carotene and carotenoid content was calculated from standard curves and expressed in units of mg/g.

Isoflavone Analysis

Isoflavone content was determined by HPLC according to a previous method, with minor modifications [29]. The extraction of isoflavone was conducted with 0.4 g soybean flour in 10 mL 70% methanol solution, which was mixed by shaking at 240 rpm for 8 h and then centrifuged at 4000 rpm for 10 min. The clear aliquot was filtered through a 0.45- μ m filter. The analysis was performed using an Apollo C18 column (5 μ m, 25 cm \times 4.6 mm; Alltech, Connecticut, USA). The mobile phases comprised solvent A (0.1% iced acetic acid acetonitrile solution) and solvent B (0.1% iced acetic acid), with a flow rate of 1 mL/min and an injection volume of 20 μ L. The solvent system (% solvent A/% solvent B) was conducted for 0–40 min (15–30/85–70%). The standards of isoflavone components, including daidzin (D), glycitin (GL), genistin (G), malonyldaidzin (MD) and malonylglycitin (MGL), were provided by Sigma-Aldrich (Sigma-Aldrich Corp., St. Louis, MO, USA). The content of various components (D, GL, G, MD and MGL) and the total isoflavone content were calculated from standard curves and expressed in units of mg/g.

Saponin Analysis

The saponin content was determined using HPLC according to a previous method, with minor modifications [30]. A total of 0.70 g soybean flour was placed in 10 mL of a 70% methanol solution and mixed by shaking at 240 rpm for 8 h, followed by centrifugation at 4000 rpm for 10 min. The clear aliquot was filtered through a 0.45- μ m filter. The analysis was performed using a YMC-Pack ODS-AM C18 column (5 μ m, 25 cm \times 4.6 mm; YMC, Kyoto, Japan). The mobile phases were solvent A (0.05% iced acetic acid acetonitrile solution) and solvent B (0.05% iced acetic acid), with a flow rate of 1 mL/min and an injection volume of 20 μ L. The solvent system (% solvent A/% solvent B) was applied for 0–12 min (15–30/85–70%), 12–16 min (30–50/70–50%) and 16–28 min (50/50%). A standard of saponin I (sap I) was prepared by our lab, and it was also used as standard for the other four saponin components [saponin II (sap II), saponin α g (sap α g), saponin β g (sap β g) and saponin $\beta\alpha$ (sap $\beta\alpha$)]. We calculated the content of saponin

components (I, II, α g, β g and $\beta\alpha$) and the total saponin content from standard curves and expressed them in mg/g.

Phospholipid Analysis

Extraction and quantitation of phospholipids was conducted according to the methodology described by Lesnfsky *et al.* [31]. Phospholipid extraction consisted in placing 0.5 g of soybean flour in 10 mL 70% Folch solution (chloroform: isopropanol = 2:1) and using ultrasonic extraction for 30 min, after which the sample was centrifuged at 4000 rpm for 10 min. Next, 5 mL of Folch solution was added to the 10-ml clear aliquot, and 4 mL 0.58% NaCl solution was added, mixed and centrifuged at 2000 rpm for 2 min. Four milliliters of 50% methanol was added to the bottom layer, mixed and centrifuged at 2000 rpm for 2 min, after which 2 ml of the bottom layer solution was absorbed, dried and added with water to 2 mL for analysis. The analysis was performed using a Luna 5- μ M silica column (250 mm \times 4.6 mm; Phenomenex, Torrance, CA, USA). The mobile phases were solvent A (chloroform–methanol–water–ammonium = 75:24:0.5:0.5) and solvent B (chloroform–methanol–water–ammonium = 55:39:5.5:0.5), with a flow rate of 1 mL/min. The solvent system (% solvent A/% solvent B) comprised 1–25 min (100–0/0–100%), 25–29 min (0/100%), 29–31 min (0–100/100–0%) and 31–36 min (100/0%). Standards of phosphatidylcholine (PC), phosphatidylethanolamine (PE) and phosphatidylinositol (PI) were purchased from Sigma-Aldrich (Sigma-Aldrich Corp., St. Louis, MO, USA). The content of PE, PI and PC was calculated from standard curves and expressed in units of mg/g.

Bowman–Birk Inhibitor (BBI) and Lectin Analysis

BBI and lectin were extracted and determined according to the method described by Anta *et al.* [32]. A 1.0-g sample of flour was suspended in 10 mL acetate (pH 4.5, 0.1 M) buffered saline, and ultrasonication was applied for 5 min. Next, the solution was centrifuged at 3430 rpm for 30 min. The clear aliquot was filtered through a 0.45- μ m filter. A Waters Alliance HPLC separator (Waters Corporation, Milford, MA, USA) equipped with an Apollo C18 column [250 \times 4.6 mm, 5 μ m, W.R. Grace & Co./Alltech, Connecticut, USA] was employed, with column temperature of 60 $^{\circ}$ C and UV detection at 260 nm. The mobile phases were as follows: (A) 0.1% glacial acetic acid in water, and (B) 0.1% glacial acetic acid in acetonitrile. The gradient program with a flow rate of 1.0 mL/min was as follows: 25% B for 2 min; 25–60% B for 16 min; 60–100% B for 1 min; 100% B for 5 min; 100–25% B for 1 min; 25–25% B for 5 min; and an injection volume of 20 μ L. Standards of the trypsin-chymotrypsin inhibitor (Bowman–Birk

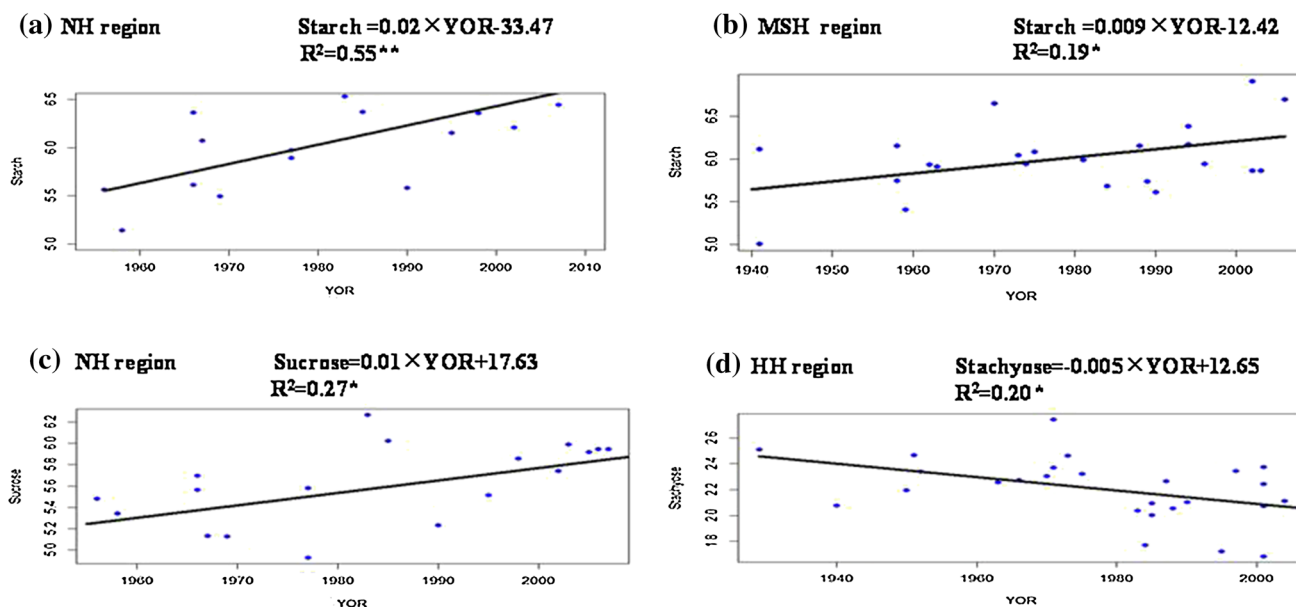


Fig. 1 Changes in carbohydrate saccharide/oligosaccharide and starch content with the year of cultivar release. Changes in **a** starch content (%) in the NH region, **b** starch content (%) in the MSH region, **c** sucrose content (%) in the NH region, and **d** stachyose con-

tent (%) in the HH region with the year of cultivar release. The equations and their corresponding R^2 values are shown. * and ** indicate linear correlation coefficients significant at the 0.05 and 0.01 probability levels, respectively. Abbreviations are shown in the text

inhibitor) and lectin were purchased from Sigma-Aldrich (Sigma-Aldrich Corp., St. Louis, MO, USA). The content of BBI and lectin was calculated from standard curves and expressed in mg/g.

Data Analysis

An analysis of variance for each of the 35 seed constituent traits was computed using the PROC MIXED procedure in the SAS software program, version 9.22 [33], where the region and cultivar nested in region were fixed variables, and replication was a random variable. The cultivar means were calculated based on the least squares means (LS means) across replications per cultivar. To obtain the annualized estimate of genetic change, a linear regression was performed between a cultivar trait mean and the respective year of release (YOR) in each region. To compare the cultivar means across regions, boxplots were drawn to show the distribution and variation in seed composition among the four regions. Pearson correlation coefficients among 35 seed compositional components were calculated with the PROC CORR procedure in SAS [34]. To determine the essential constituent contribution to the total variation, principal component analysis (PCA) was performed using the FactoMineR package in the R environment, with default settings using 35 constituents [35]. First, two principal components (DIM1 and DIM2) were selected, and an

individual factor map and variables factor map were then drawn.

Results

Analysis of Variance

The effects of the fixed variables region and cultivar nested in region were significant for all seed constituents (Table S1), whereas the replication factor was significant for only PE and PC. According to the mean square (data not shown), the factor region contributed to the largest proportion of variation for all traits, while the factor cultivar nested in region was the second largest source of variation for all traits. The residual mean square for each trait was minor compared with that of the fixed variables, suggesting that this randomized complete block design performed well, with negligible error.

Temporal Trends in Seed Nutritional Composition

Carbohydrate Saccharides/Oligosaccharides and Starch

The genetic changes in seed composition were computed by regressing the concentrations of the seed constituents on the respective year of cultivar release in each region. The concentration of starch presented significant and positive

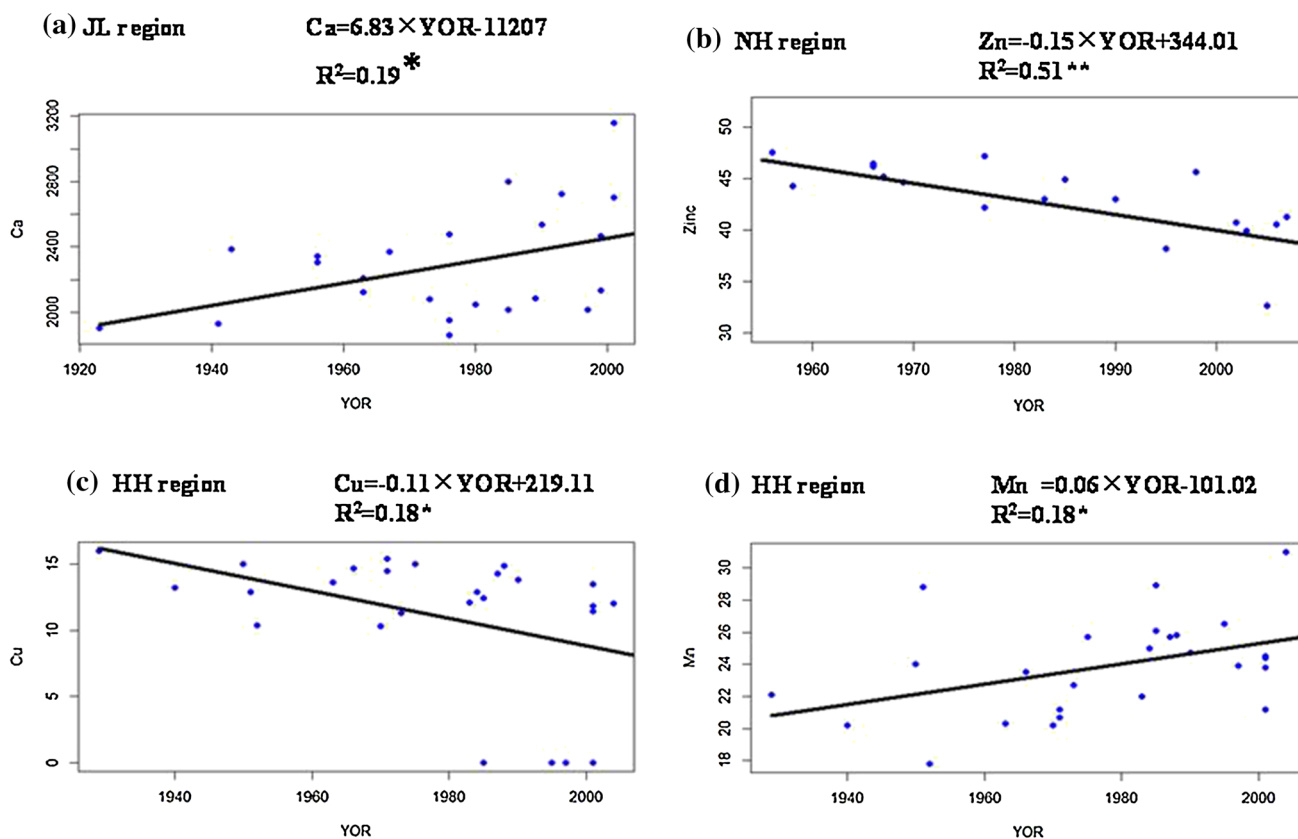


Fig. 2 Changes in mineral element content with the year of cultivar release. Change in **a** Zn content ($\mu\text{g/g}$) in the NH region, **b** Cu content ($\mu\text{g/g}$) in the HH region, **c** Mn content ($\mu\text{g/g}$) in the HH region, and **d** Ca content ($\mu\text{g/g}$) in the JL region with the year of cultivar

release. The equations and their corresponding R^2 values are shown. * and ** indicate linear correlation coefficients that are significant at the 0.05 and 0.01 probability levels, respectively. Abbreviations are shown in the text

relationships with YOR in the NH and MSH regions, with an increase of 0.02 and 0.009% per year, respectively (Fig. 1a and b). However, there were no significant changes in oligosaccharides, with the exception of sucrose and stachyose, in any of the regions. A positive relationship between sucrose and YOR, and a negative relationship between stachyose and YOR, were observed in the NH and HH regions (Fig. 1c and d), with corresponding rates of 0.01 and 0.005% per year, respectively, over the nine decades, indicating that the changes in starch and carbohydrate saccharides/oligosaccharides that occurred during the breeding process were moderate.

Mineral Elements

Ca and Mg are macro-mineral elements, and Zn, Mn and Cu are micro-elements. In the JL region, Ca exhibited a positive relationship with YOR, with an increase of $6.84 \mu\text{g g}^{-1}$ per year (Fig. 2a). Significant and decreasing trends were found for Zn concentrations in the NH region and Cu in the HH region, and an increasing trend for Mn in the HH region, over the nine decades (Fig. 2b–d). The annual decreases for

Zn and Cu were 0.15 and $0.11 \mu\text{g g}^{-1}$, respectively, and the annual increase for Mn was $0.06 \mu\text{g g}^{-1}$.

Tocopherol and Carotenoid

A significant increasing trend of tocopherol concentration was found in the JL region, from $87.67 \mu\text{g g}^{-1}$ for cv. Jiti 5 released in 1956, to $413.15 \mu\text{g g}^{-1}$ for cv. Tiefeng 29 released in 1997. The annual increase was $1.49 \mu\text{g g}^{-1}$, which translates to an increase of 1.70% per year over the past nine decades (Fig. 3a). There was a significant increasing trend for α -Toc concentration in the MSH region and for $\beta + \gamma$ -Toc and δ -Toc concentrations in the JL region (Fig. 3b–d), with increases of 0.29, 0.74 and $0.52 \mu\text{g g}^{-1}$ per year, respectively. The highest concentrations of α -Toc, $\beta + \gamma$ -Toc and δ -Toc found in these regions were 75.84 (cv. Hefeng 50, 2006), 217.51 (cv. Jiunong 9, 1976) and 137.13 (cv. Tiefeng 29, 1997) $\mu\text{g g}^{-1}$, and their lowest concentrations were 22.41 (cv. Jinyuan 2, 1941), 45.29 (cv. Jiti 5, 1956) and 22.65 (cv. Jiti 5, 1956) $\mu\text{g g}^{-1}$, respectively. For carotenoids, there was only a significant and negative relationship between lutein and YOR in the NH region ($0.14 \mu\text{g g}^{-1}$ per year, $R^2 = 0.29$; Fig. 3e).

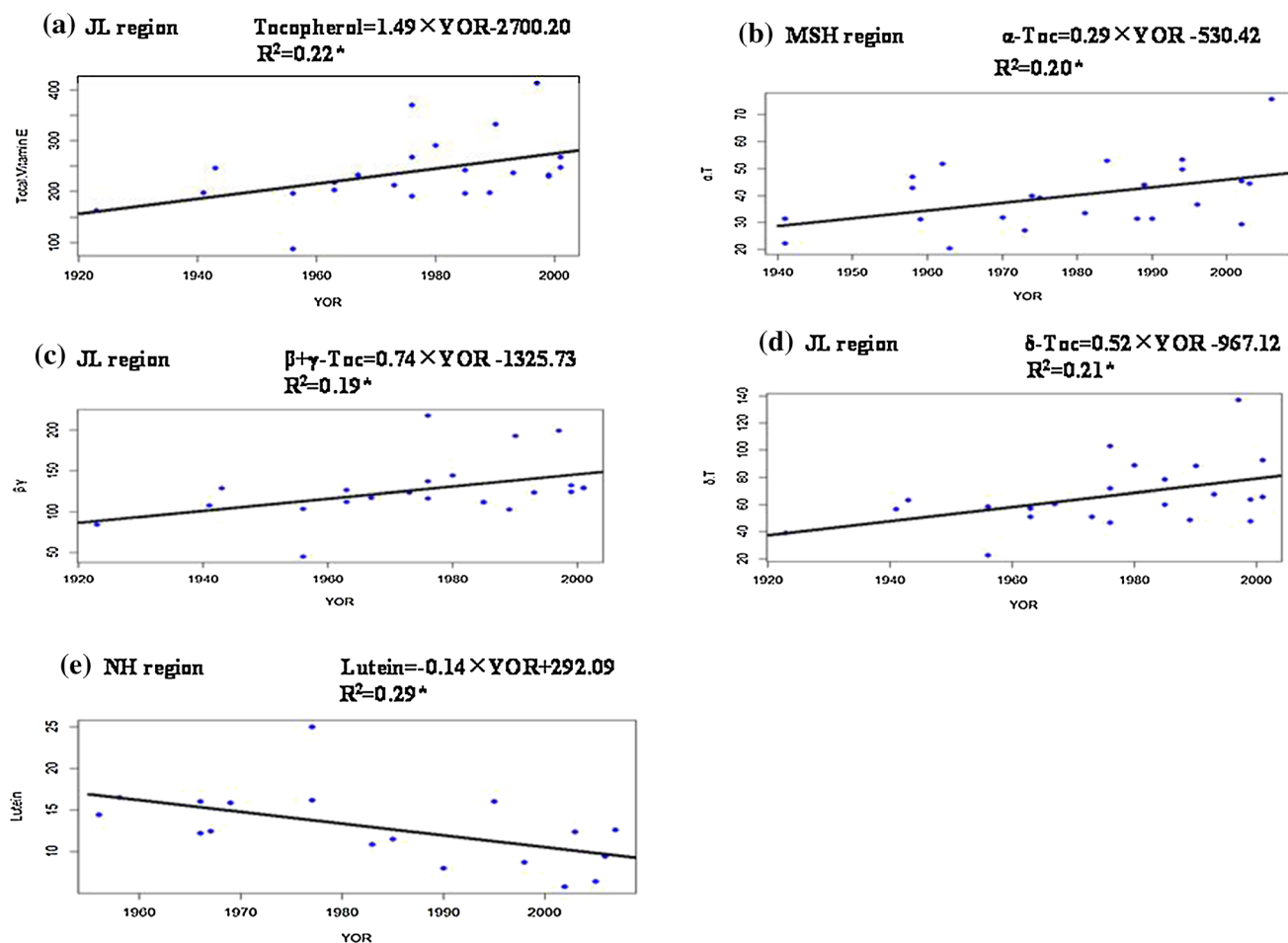


Fig. 3 Changes in the tocopherol and carotenoid content with the year of cultivar release. Changes in **a** tocopherol content ($\mu\text{g/g}$) in the JL region, **b** α -Toc content ($\mu\text{g/g}$) in the MSH region, **c** $\beta + \gamma$ -Toc content ($\mu\text{g/g}$) in the JL region, **d** δ -Toc content ($\mu\text{g/g}$) in the JL region and **e** lutein content ($\mu\text{g/g}$) in the NH region with the year of

cultivar release. The equations and their corresponding R^2 values are shown. * and ** indicate linear correlation coefficients that are significant at the 0.05 and 0.01 probability levels, respectively. Abbreviations are shown in the text

Temporal Trends in Seed Bioactive Constituents

Isoflavone and Its Components

The increases in total isoflavone content were significant in the NH and MSH regions (0.03 and $0.02 \mu\text{g g}^{-1}$ per year, respectively; Fig. 4a and b), with the lowest total isoflavone content of 1.51 (cv. Heilongjiang 41, 1958) and 1.35 (Jinyuan 2, 1941), and the highest content of $3.62 \mu\text{g g}^{-1}$ (Heihe 38, 2005) and $3.19 \mu\text{g g}^{-1}$ (Heinong 33, 1988), respectively, during the nine-decade period. The changing patterns of five isoflavone individual forms were different in various regions (Fig. 4c–m). Among these, daidzin and malonyldaidzin concentrations were positively correlated with YOR in the NH and MSH regions, but negatively correlated with YOR in the HH region. Genistin and malonylgenistin were positively associated with YOR in the NH

and MSH regions. Glycitin was found to be significantly decreased over the nine decades in the JL region. The changing rates of the five isoflavone components ranged between 0.001 and $0.01 \mu\text{g g}^{-1}$ per year.

Saponin and Its Components

There was no noticeable changing trend in total saponin, indicating that saponin content has been stable over the past nine decades. There were consistent changing trends in saponin components by specific region (Fig. 5a–f). Sap I, sap II and sap αg were positively associated with YOR in the NH region, whereas they were negatively correlated with YOR in the HH region. The annual increases in sap I, sap II and sap αg were 0.003 , 0.004 and 0.008 mg g^{-1} in the NH region, respectively, whereas they were 0.003 , 0.001 and 0.004 mg g^{-1} in the HH region, respectively.

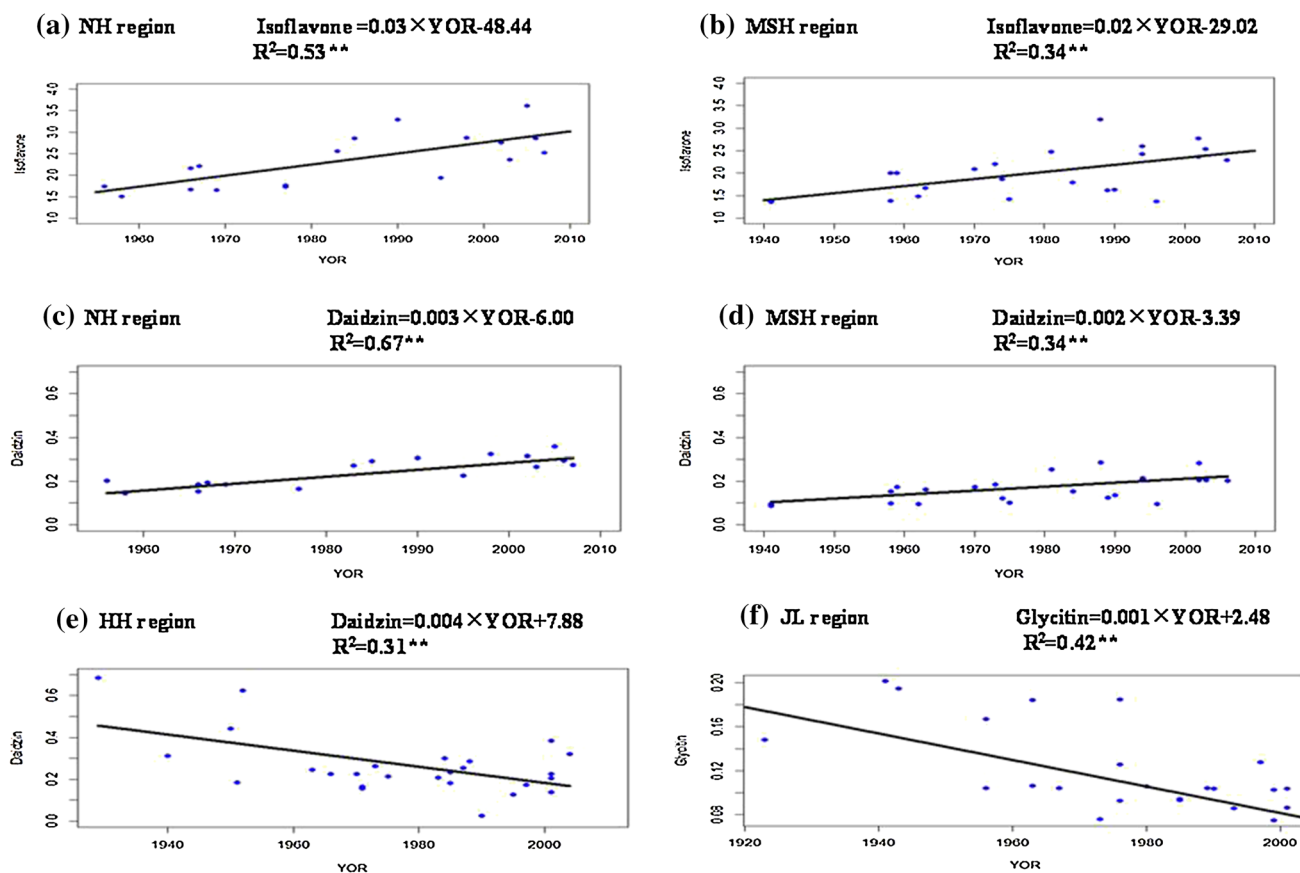


Fig. 4 Changes in the isoflavone and isoflavone components content with the year of cultivar release. Changes in **a** isoflavone content (mg/g) in the NH region, **b** isoflavone content (mg/g) in the MSH region, **c** daidzin content (mg/g) in the NH region, **d** daidzin content (mg/g) in the MSH region, **e** daidzin content (mg/g) in the HH region, **f** glycitin content (mg/g) in the JL region, **g** genistin content (mg/g) in the NH region, **h** genistin content (mg/g) in the MSH region, **i** malonyldaidzin (mg/g) in the NH region, **j** malonyldaidzin (mg/g) in

the MSH region, **k** malonyldaidzin (mg/g) in the HH region, **l** malonylgenistin content (mg/g) in the NH region, and **m** malonylgenistin content (mg/g) in the MSH region, with the year of cultivar release. The equations and their corresponding R^2 values are shown. * and ** indicate linear correlation coefficients that are significant at the 0.05 and 0.01 probability levels, respectively. Abbreviations are shown in the text

Lectin and Phospholipids

The content of lectin and phospholipids was found to increase with YOR. Lectin increased linearly with YOR at an annual rate of 0.03 mg g^{-1} in the NH region (Fig. 6a). PE was found to increase with YOR in the JL region at an annual rate of 0.02 mg g^{-1} . PC also increased with YOR, at a rate of 0.02 mg g^{-1} per year, in the NH region (Fig. 6b and c).

The changing patterns of each trait were different across different regions, and a consistent trend across regions, or a single trend in each region when a trait was significant only in a specific region, is summarized as follows: with the genetic improvements over the nine-decade period, modern soybean cultivars have higher levels of isoflavone, tocopherol, starch, sucrose, Ca, Mn, phospholipids and lectin, but lower levels of Zn, Cu and lutein than older cultivars.

The Spatial Distribution of Seed Constituents

The regional means of seed compositional components were computed and compared across four regions, and significant differences were found with respect to all constituents ($P < 0.01$; Table S1). The distribution, variation and outliers of each variable within each region are shown in the boxplots in Fig. 7.

The regional distribution of isoflavone and its components (daidzin, malonylglycitin and glycitin) was found at higher levels in cultivars planted at lower latitudes (JL and HH regions) compared to higher-latitude (NH and MSH regions) cultivars. This is in agreement with the conclusion reached in another study, where lower isoflavone content was present in early- rather than late-maturing soybean cultivars [22]. In the current study, the frost-free period in the low-latitude HH and JL regions is

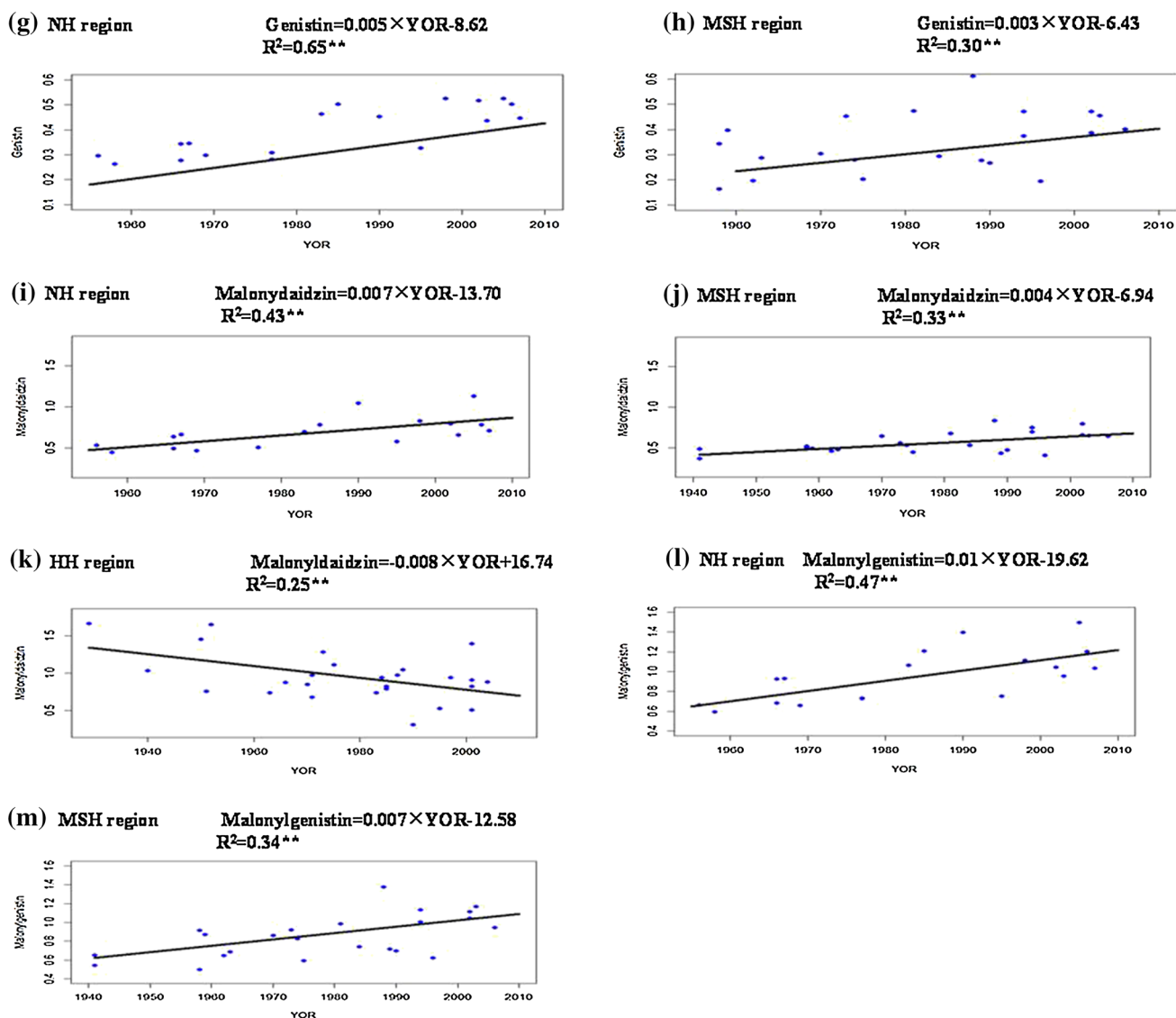


Fig. 4 continued

longer than that in the high-latitude regions, and cultivars grown in the lower-latitude regions are later-maturing than those in the higher-latitude regions (NH and MSH regions); therefore, a longer growing period can allow for a greater accumulation of isoflavones in the seeds. In contrast, with regard to the components genistin and malonyldaidzin, cultivars in the HH region had the highest levels, followed by those in the NH, MSH and JL regions. The range of isoflavone and isoflavone components was larger in the HH region than in the other regions.

Oligosaccharide components (sucrose and raffinose), carotenoid and carotenoid components (lutein and β -carotene), and phospholipid components (PC and PI) exhibited higher concentrations in higher-latitude regions (NH and MSH) than in lower-latitude regions (JL and

HH), whereas stachyose, lectin, saponin and saponin components (sap β g and sap β α) and the mineral element Zn showed an opposite trend with latitude (Table S2). There were decreasing trends from north to south in terms of tocopherol and tocopherol components ($\beta + \gamma$ -Toc and δ -Toc) and sap α g content. There was no relationship to latitude observed in the distribution of other components, including starch, fiber, BBI, PE, α -TOC, and sap I and II content, and all mineral elements with the exception of Zn.

In sum, the spatial distribution of most seed nutritional and bioactive constituents, including isoflavone, tocopherol, carotenoid, saponin, and oligosaccharides, is influenced by climatic conditions (temperature and rainfall) and exhibits a relationship with latitude. Consequently,

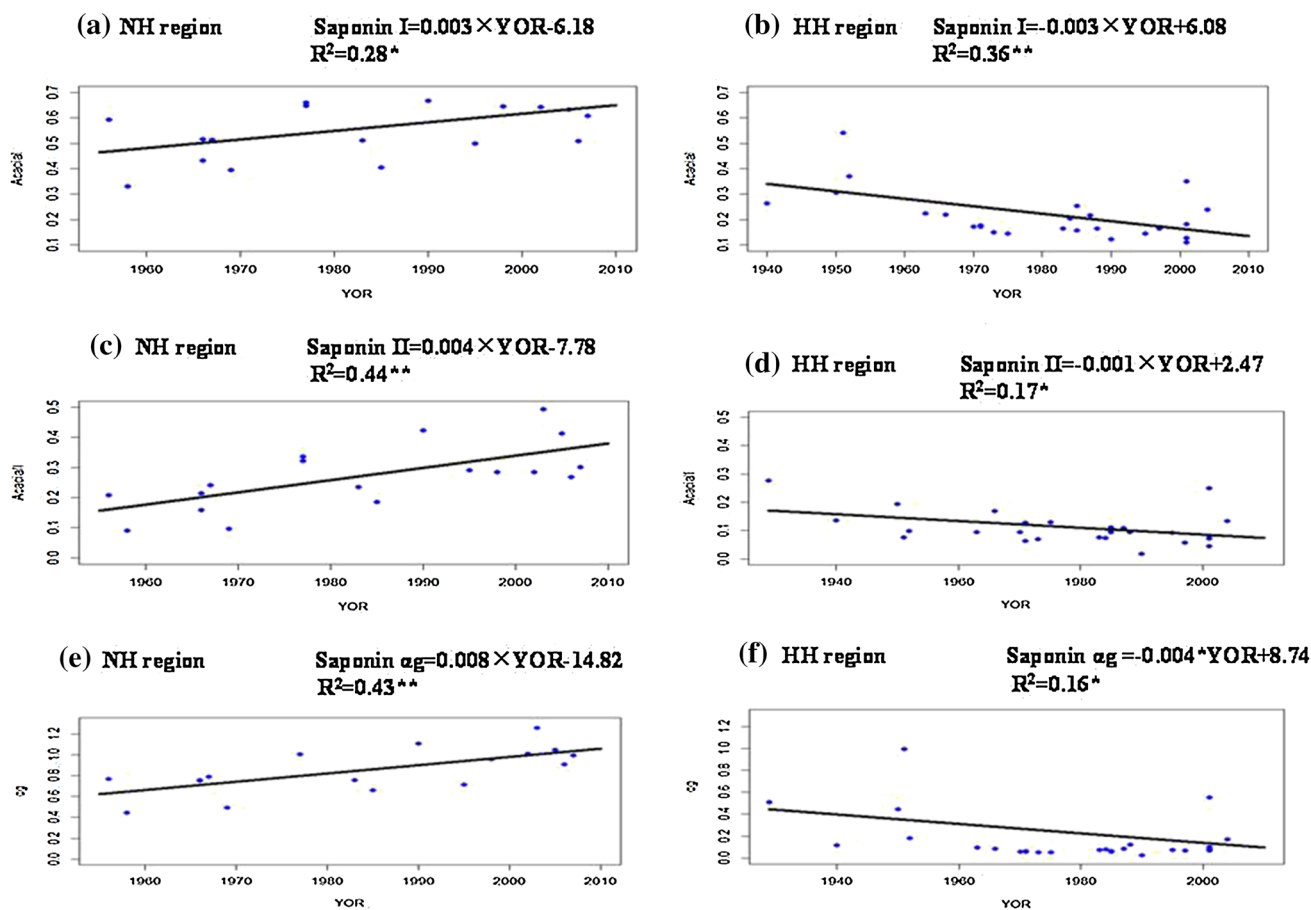


Fig. 5 Changes in saponin and saponin component content with the year of cultivar release. Changes in **a** saponin I content (mg/g) in the NH region, **b** saponin I content (mg/g) in the HH region, **c** saponin II content (mg/g) in the NH region, **d** saponin II content (mg/g) in the HH region, **e** saponin α g content (mg/g) in the NH region, and **f**

saponin α g content (mg/g) in the HH region with the year of cultivar release. The equations and their corresponding R^2 values are shown. * and ** indicate linear correlation coefficients that are significant at the 0.05 and 0.01 probability levels, respectively. Abbreviations are shown in the text

the environmental conditions in the high-latitude regions (NH and MSH regions) are relatively favorable for the production of tocopherol and its components ($\beta + \gamma$ -Toc and δ -Toc), carotenoid and its components (lutein and β -carotene), oligosaccharides (sucrose and raffinose) and phospholipids (PC and PI). Likewise, low-latitude regions (JL and HH regions) can provide a suitable environment for the synthesis of stachyose, isoflavones and isoflavone components (daidzin, malonylglycitin and glycitin), saponin and saponin components (sap β g and sap β α), lectin, and the mineral element Zn.

Correlations Among Seed Constituents

The coefficients of the relationship between different seed constituents were computed and compared. The relationship between the total content and the content of the components was generally positive and significant (Table 3). For example, isoflavone content was positively correlated

with all of its components (daidzin, glycitin, genistin, malonyldaidzin and malonylglycitin). Among these, the highest ($r = 0.77$) and lowest ($r = 0.15$) correlation coefficients were exhibited between the total content and daidzin, and between the total content and glycitin, respectively, suggesting that daidzin is the major contributor to total isoflavone concentration.

The pairwise correlations between the components of the same constituent were generally significant. The isoflavone components were largely positively correlated with one another. The saponin components were negatively correlated with one another, with the exception of sap I and sap II, sap β g and sap β α , Sap α g and sap I, and sap α g and sap II. The mineral elements were generally significantly and positively correlated with one another (Mn, Mg, Ca, Cu and Fe). The PE and PC were negatively correlated (-0.31 , $P = 0.004$). Sucrose and stachyose were negatively correlated with each other (-0.29 , $P = 0.007$), while BBI and lectin were positively correlated (0.31 , $P = 0.003$). The

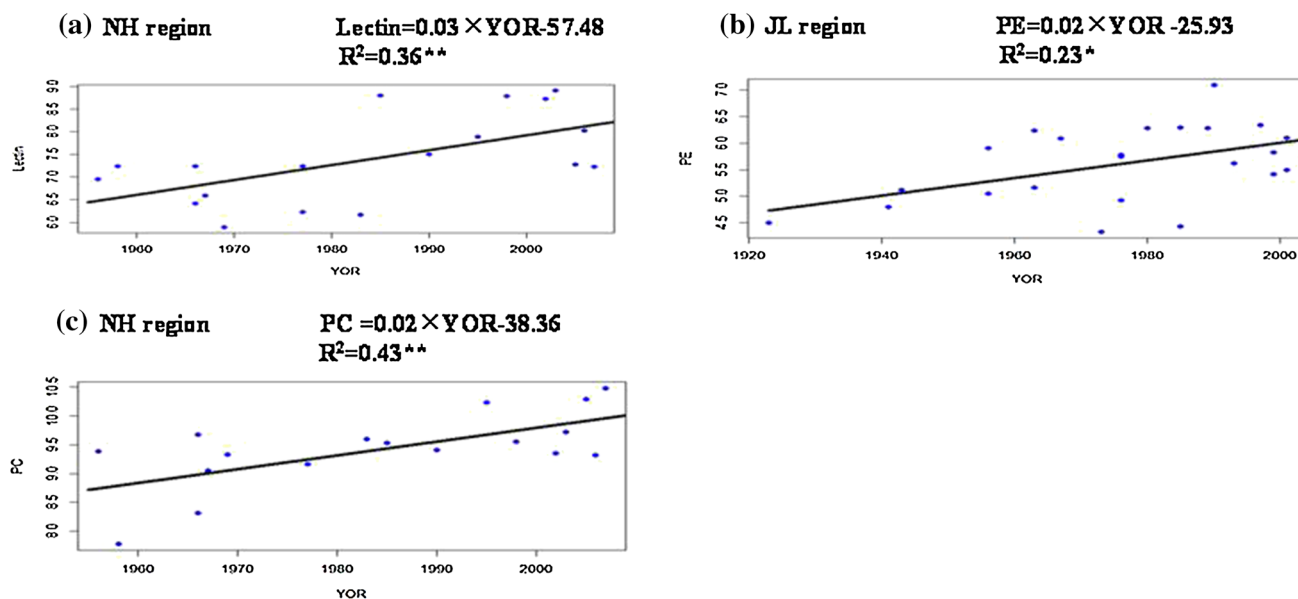


Fig. 6 Changes in lectin and phospholipid content with the year of cultivar release. Changes in **a** lectin content (mg/g) in the NH region, **b** PE content (mg/g) in the JL region and **c** PC content ($\mu\text{g/g}$) in the NH region with the year of cultivar release. The equations and their

corresponding R^2 values are shown. * and ** indicate linear correlation coefficients that are significant at the 0.05 and 0.01 probability levels, respectively. Abbreviations are shown in the text

tocopherol components α -Toc, $\beta + \gamma$ -Toc and δ -Toc were positively correlated. Finally, the carotenoid components lutein and β -carotene were positively correlated with each other ($0.30, P = 0.005$).

Significant relationships were also observed between different nutritional and bioactive components. Starch was significantly and positively correlated with isoflavone ($0.21, P = 0.047$), most mineral elements [Fe ($0.39, P = 0.0001$), Mn ($0.34, P = 0.001$) and Ca ($0.25, P = 0.02$)], but was negatively correlated with BBI ($-0.36, P = 0.0006$). Fiber content was positively correlated with saponin ($0.23, P = 0.03$) and mineral elements [Cu ($0.27, P = 0.01$) and Ca ($0.35, P = 0.0009$)]. Saponin content was positively correlated with fiber ($0.23, P = 0.03$) and mineral elements [Fe ($0.41, P < 0.0001$), Cu ($0.30, P = 0.004$), Mn ($0.26, P = 0.01$), Mg ($0.22, P = 0.04$)], but negatively correlated with carotenoid [β -Carotene ($-0.21, P = 0.045$) and carotenoid ($-0.23, P = 0.03$)]. The content of BBI and lectin was negatively correlated with phospholipids and vitamins (tocopherols and carotenoids) and their components. TE was positively correlated with carotenoid ($0.31, P = 0.004$), but negatively correlated with Cu ($-0.25, P = 0.02$). Carotenoid was negatively correlated with isoflavone and its components, all mineral elements, saponin, BBI and lectin, and TE.

We also investigated the relationships between yield, seed protein levels, seed oil content and the seed constituents. The yield was positively correlated with the content of starch ($0.30, P = 0.005$) and TE ($0.26, P = 0.01$),

whereas yield was negatively correlated with the content of Cu ($-0.43, P < 0.0001$) and stachyose ($-0.23, P = 0.03$). Protein content was positively correlated with the content of fiber ($0.25, P = 0.02$) and Cu ($0.37, P = 0.0003$), but negatively correlated with the content of starch ($-0.28, P = 0.008$), Ca ($-0.22, P = 0.04$) and Mn ($-0.36, P = 0.0006$). Oil content was positively correlated with the content of the mineral elements Zn ($0.23, P = 0.03$), Mn ($-0.41, P < 0.0001$) and Ca ($0.26, P = 0.01$), whereas it was negatively correlated with Cu content ($-0.41, P < 0.0001$). This indicates that the indirect selection of seed constituents may occur through selecting for yield, seed proteins and oil content. In short, knowledge of the correlations among different constituents is valuable for efforts to increase multiple desirable constituents simultaneously, without an increase in undesirable elements.

Principal Component Analysis of Seed Composition

A principal component analysis indicated that the first two components could explain 21.12% (Dim 1) and 15.34% (Dim 2), respectively, and together accounted for 36.46% of the total variation in seed constituent traits (Fig. 8a). β -carotene, carotenoid, sap αg and sap βg exhibited a high correlation with Dim1, with positive loadings of β -carotene, carotenoid and sap αg , whereas sap βg was negatively loaded on Dim 1. The content of Fe, Mn, glycinin and saponin showed a high and positive correlation with

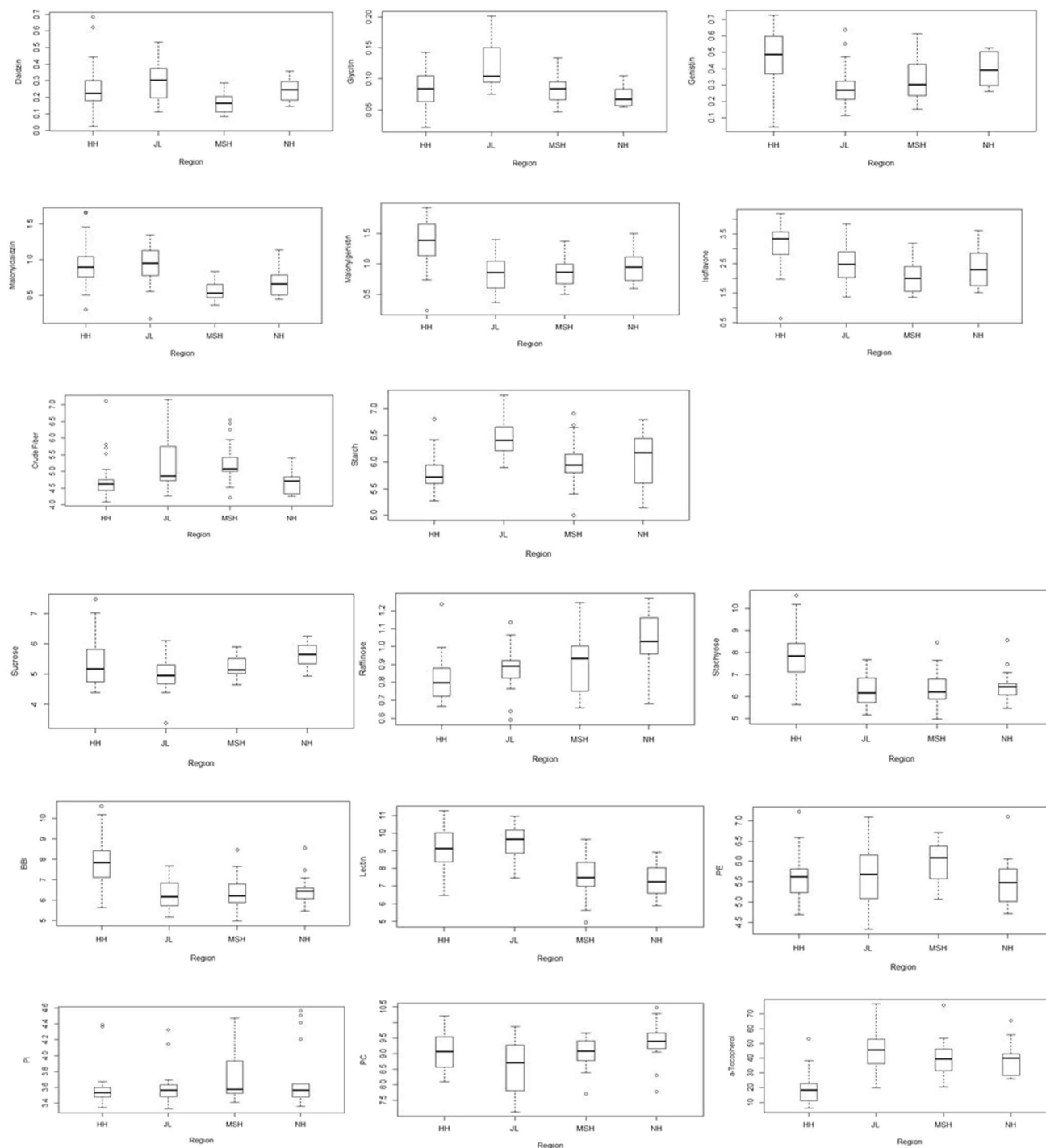


Fig. 7 Means of seed constituents of widely grown soybean cultivars in different regions. The *box* represents data between the 25th and 75th percentiles, *black lines* inside the boxes represent the mean, *error bars* indicate the 90th and 10th percentiles, and *dots* represent

outliers. *HH* Huang-Huai-Hai Rivers Valley, *JL* Jilin and Liaoning region, *MSH* Midsouth Heilongjiang region, *NH* Northern Heilongjiang region

Dim 2. This suggests that Dim 1 can represent the carotenoid and saponin components, while Dim 2 can reflect the mineral elements, saponin and isoflavone components. Collectively, carotenoid, saponin, mineral elements and

isoflavone components played an essential contributory role in the total variation among the 89 cultivars. Among these, the mineral elements Fe and Mn, with the largest loadings, are the most important constituents for distinguishing these

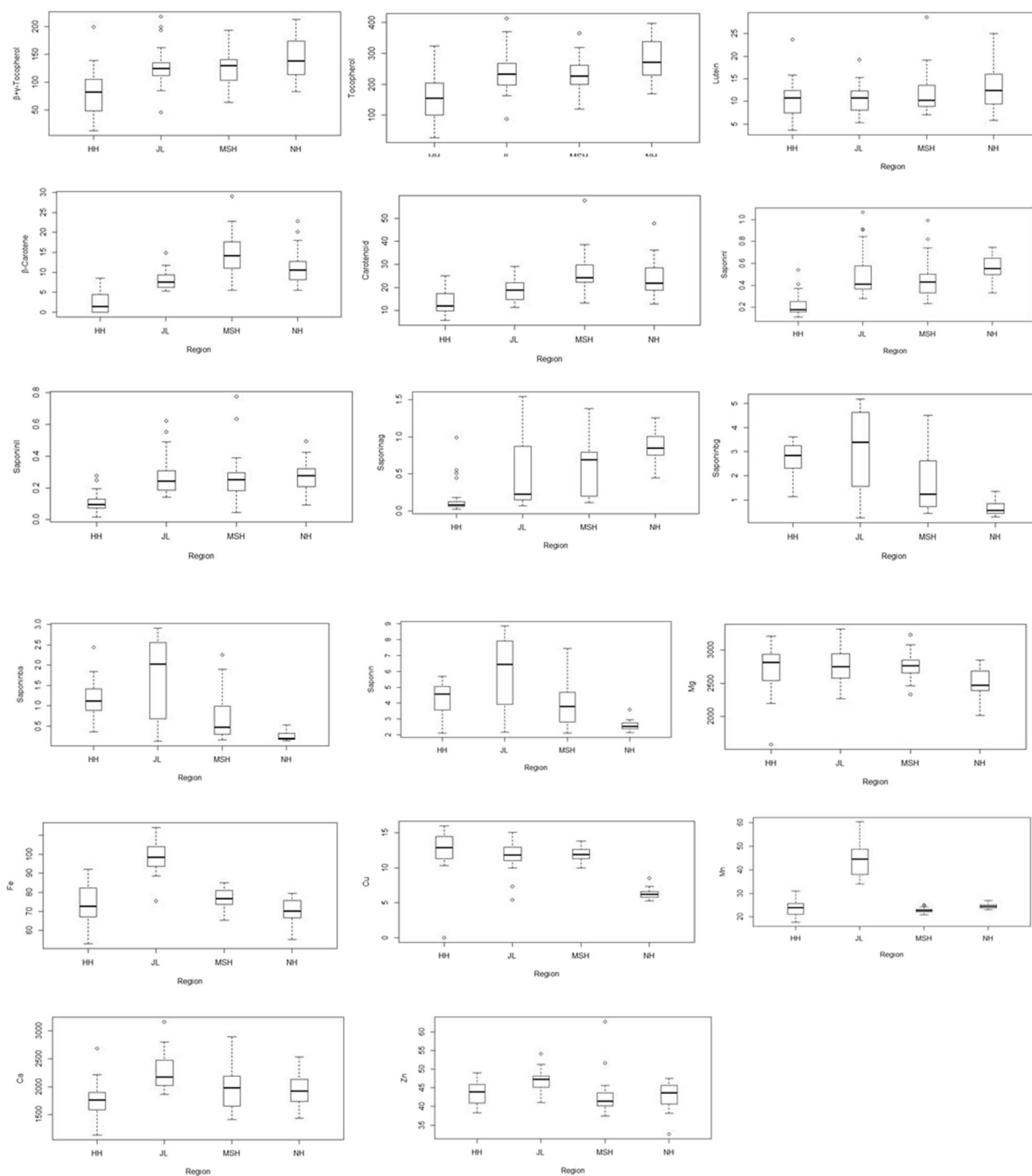


Fig. 7 continued

genotypes. In addition, PCA indicated negative associations between isoflavone and carotenoid, and malonylglycitin and carotenoid, as expressed by the position at nearly 180° , which is consistent with the results from the Pearson correlation coefficients. Together, Dim 1 and Dim 2 divided

the widely grown cultivars according to their regions of origin (NH, MSH, JL and YHH) (Fig. 8b). The subgroups of HH, JL, NH and MSH were dispersed along the DIM 1 and DIM 2 coordinates, and NH and MSH overlapped due to their similar environmental conditions.

Table 3 Pearson correlation coefficients of the seed constituents of 89 widely grown soybean cultivars

Table with 38 columns (Protein, Oil, Starch, etc.) and 38 rows showing Pearson correlation coefficients between various constituents. Values range from -0.88 to 0.8888.

*, **, and *** indicate significance at the 0.05, 0.01 and 0.001 levels, respectively. For abbreviations, see text.

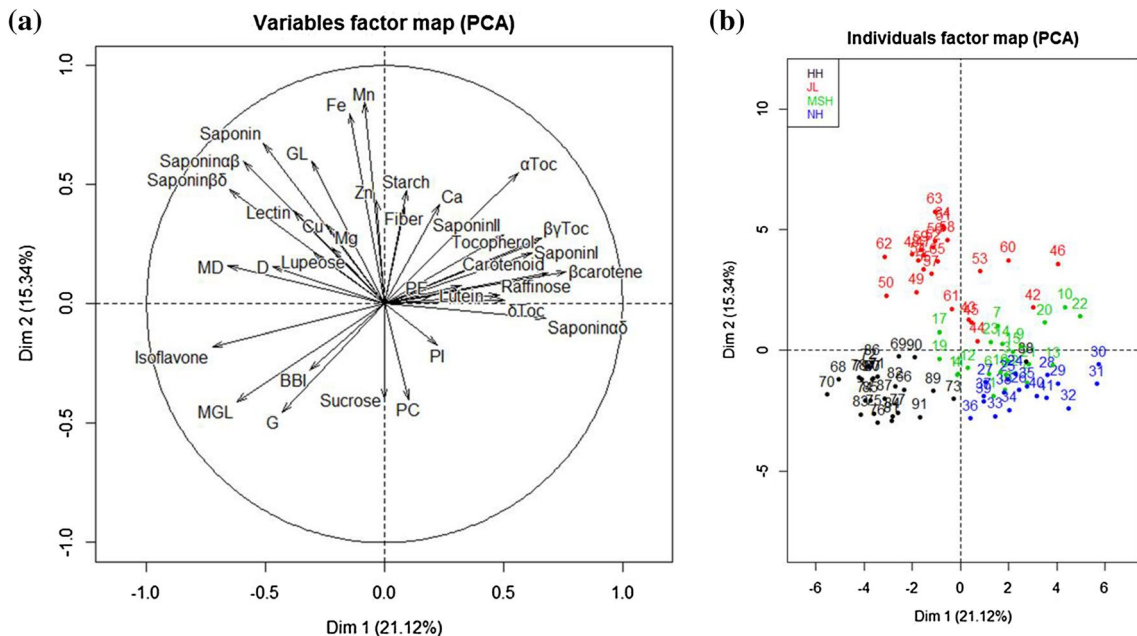


Fig. 8 Principal components 1 and 2 (DIM1 and DIM2) were determined from the set of 89 widely grown soybean cultivars based on the main seed constituents. a Variables factor map for PCA. b Individual factor map for PCA. Blue dots indicate cultivars from the

Northern Heilongjiang, green from the Midsouth Heilongjiang subgroup, red the Jilin and Liaoning subgroup, and black the Huang-Huai-Hai Rivers Valley subgroup. The numbers in the plot correspond to the cultivar numbers listed in Table 1

Discussion

Soybean seeds have myriad nutritive constituents. A total of 33% of soybean seed components are carbohydrates [13], which are highly associated with desiccation tolerance and stability. Among these, 16.6% are soluble sugars. Sucrose and stachyose are the main saccharides in soybean seeds [36]. Stachyose affects the efficiency of soybean utilization as human food and animal feed, and can cause diarrhea [37]. In this study, we found an elevated level of sucrose and a reduced level of stachyose in the NH and HH regions, respectively. This may be because high levels of

sucrose and low levels of stachyose can improve the flavor, digestibility and absorption of soy-based foods [2].

Tocopherols are lipid-soluble compounds: α-Toc is the most active vitamin E isomer, followed by β-Toc, γ-Toc and δ-Toc [38]. However, in soybeans, the most abundant TE component, γ-Toc, is relatively inactive [39]. The α-Toc content in soybeans is lower than that in other oil crops such as safflower and sunflower [40]. Thus, TE activity must be increased by elevating the content of α-Toc. A previous study indicated that individual and total TE content had relatively low heritability. The Satt376 and Satt266 markers have been reported to be significantly associated

with α -Toc content [41], indicating that these markers can be used in marker-assisted selection to facilitate the development of high-TE or α -Toc cultivars. In the short-season region of Canada, a series of 16 cultivars representing seven decades of cultivar release reflected significant increases in α -Toc and δ -Toc concentrations [6], which is consistent with our results. Given the positive correlation between oil and tocopherol content, the increase in tocopherol content in the present study may be ascribed to the indirect selection for high oil content by plant breeders over time.

Soybeans are an essential source of dietary minerals as well [2]. We found that the concentration of Ca increased with cultivar improvement in the JL region. Because Ca concentration is positively correlated with seed quality traits (protein and oil content) [42], the increase in Ca concentration may be attributed to the indirect selection of seed traits. However, Zn and Cu are less concentrated in modern cultivars than in old cultivars, and this may be attributed to the negative correlation between their concentration and the seed yield per plant and seed size [20]. The positive relationships among mineral elements found in this study were confirmed in another study [20]. The content of mineral elements is affected by climatic and environmental conditions; for example, studies have found that Fe content increased with wet weather [43], Ca content decreased with drought stress during seed formation and filling [42], and Ca and Mn content decreased with an increase in daytime temperatures from 30 to 35 °C during the R1–R5 stages [20].

Soybeans also provide a rich source of biologically active and health-beneficial phytochemicals, including isoflavones [44], saponins [30], phospholipids [45] and lectins [46]. The antioxidant and anti-carcinogenic activity of soybeans may be due in part to the presence of these bioactive constituents. Isoflavone is among the most important health-beneficial components in soybean seeds, due to its antioxidant and estrogenic properties [44]. In the present study, we found a significant increasing trend of isoflavone content in the NH and MSH regions, indicating that the health value of soybeans is improved by long-term genetic breeding. The accumulation of isoflavones is influenced by environmental conditions [47], and higher isoflavone concentration has been reported to accumulate during the seed maturation stage at lower temperatures [48]. The positive correlation between total isoflavone and isoflavone content was also supported by the evidence from another study with MG 0 to VI soybean cultivars [22]. In the current study, isoflavone was negatively correlated with oil content, but was positively correlated with protein content. Similar results have been reported with MG 0 to VI soybean cultivars in China [22].

Saponins have significant health-promoting effects, including anticarcinogenic, antihepatotoxic and anti-HIV activity, and cholesterol reduction, along with adverse

sensory characteristics [9, 10]. Our results indicate that the saponin component increased in the NH region but decreased in the HH region. Given the positive relationship between saponin concentration and oil content, and the negative relationship between saponin and protein content, the saponin component was indirectly selected with the improved oil content in the high-oil soybean-producing region (NH region) and with the elevated protein content in the high-protein soybean-producing region (HH region).

The membrane lipids included mainly phospholipids, which have been found to be at very low levels in soybeans [45]. We found an increase in the phospholipid component of PE and PC during the breeding progression, indicating that the bioactivity of soybeans was improved through genetic modifications. Lectins are ubiquitous proteins, and demonstrate anti-cancer characteristics [46]. In this study, we found that the lectin content increased over nine decades of breeding, indicating that genetic improvements played a positive role in elevating the health value of soybeans.

The relationship between yield, seed quality and seed composition also played an essential role in the selection of seed constituents. Because yield is the primary target of soybean breeding, the increased content of starch and TE and the decreased content of Cu and stachyose may be explained by the positive effects on yield associated with starch (0.30, $P = 0.005$) and TE (0.26, $P = 0.01$) content, but negative effects associated with Cu (-0.43 , $P < 0.0001$) and stachyose (-0.23 , $P = 0.03$) content. Likewise, oil content was highly and positively correlated with starch, Fe, Mn, Ca, Zn, saponin, lectin and carotenoid content, suggesting that simultaneous selection for these components was possible. The increase in tocopherol, starch, Mn and Ca content in the present study may be attributed to the indirect selection for high oil content by plant breeders over time. Interestingly, protein content was negatively correlated with starch, saponin, Fe, Zn, Mn, Ca, tocopherol and lectin content; therefore, compromise and balance among different components should be taken into consideration in genetic breeding.

In the current study, 89 widely grown cultivars represented 84 years of breeding, and their regions of origin covered a wide geographic area between N32° and N50°. A total of 35 nutritional and bioactive components were assessed, including some rarely analyzed components, and an overall temporospatial profile of soybean seed constituents was generated. Despite the relatively short 1-year observation period, the genetic progression was clear from the supporting evidence based on a comprehensive analysis of multiple seed constituents across a large set of widely grown cultivars over an extended time frame in a broad geographic area, indicating that data collection was sufficient.

In summary, the results pertaining to the genetic progression in seed composition, as well as the correlations among constituents, have substantial implications for soybean breeding in China, and even worldwide. Breeders can make an effort to pyramid various nutritional and bioactive constituents simultaneously and concordantly through conventional and molecular approaches based on the aforementioned conclusions. In addition, the spatial distribution of seed constituents can provide insight into potential regions for production of different desired components. In short, it is conceivable that soybean quality, along with the alimentary and pharmaceutical value of the product, can be continuously improved through effective breeding practices. A temporospatial database of soybean seed components is important for evaluating the nutritional and bioactive characteristics and health-promoting functions of soybean cultivars.

Acknowledgements We thank the researchers at the CARS experimental stations for providing samples, and Drs. Lijuan Qiu and Zhangxiong Liu (Institute of Crop Science, the Chinese Agricultural Academy of Sciences, China) for providing soybean germplasm.

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflict of interest.

References

- Berman KH, Harrigan GG, Riordan SG, Nemeth MA, Hanson C, Smith M, Sorbet R, Zhu E, Ridley WP (2010) Compositions of seed, forage and processed fractions from insect-protect soybean MON 87701 are equivalent to those of conventional soybean. *J Agric Food Chem* 57:11360–11369
- Wang TL, Domoney C, Hedley CL, Casey R, Grusak MA (2003) Can we improve the nutritional quality of legume seeds. *Plant Physiol* 131:886–891
- Qin P, Song W, Yang X, Sun S, Zhou X, Yang R, Li N, Hou W, Wu C, Han T, Ren G (2014) Regional distribution of protein and oil compositions of soybean cultivars in China. *Crop Sci* 54:1139–1146
- Xu B, Yuan S, Chang S (2007) Comparative studies on the antioxidant activities of nine common food legumes against copper-induced human low density lipoprotein oxidation in vitro. *J Food Sci* 72:S522–S527
- Yoshida H, Takagi S, Lenaga H, Tsuchiya C (1998) Regional distribution of tocopherols and fatty acids within soybean seeds. *J Oil Fat Ind* 75(7):767–774
- Morrison MJ, Cober ER, Frégeaureid JA, Seguin P (2015) Changes in lutein and tocopherol concentrations in soybean cultivars released across seven decades in the short-season region. *Crop Sci* 55:312–319
- Chen YW, Huang SC, Shiau SL, Lin J (2005) Bowman-Birk inhibitor abates proteasome function and suppresses the proliferation of MCF7 breast cancer cells through accumulation of MAP kinase phosphatase-1. *Carcinogenesis* 26(7):1296–1306
- Anthony MS, Clarkson TB, Hughes CL, Morgan TM, Burke GL (1996) Soybean isoflavones improve cardiovascular risk factors without affecting the reproductive system of peripubertal rhesus monkeys. *J Nutr* 126:43–50
- Potter SM (1996) Overview of proposed mechanisms for the hypocholesterolemic effect of soy. *J Nutr* 125:606S–611S
- Kinjo J, Imagire M, Udayama M, Arao T, Nohara T (1998) Structure-hepatoprotective relationships study of soyasaponins I–IV having soyasapogenogenol B as aglycone. *Planta Med* 64:233–236
- Prasad KN, Kumar A, Kochupillai V, Cole WC (1999) High doses of multiple antioxidant vitamins: essential ingredients in improving the efficacy of standard cancer therapy. *J Am Coll Nutr* 18:13–25
- Grusak MA (2002) Enhancing mineral content in plant food products. *J Am Coll Nutr* 21:178S–183S
- Hymowitz T, Newell CA (1981) Taxonomy of the genus *Glycine*, domestication and uses of soybeans. *Econ Bot* 35:272–288
- Wang L (2007) The origin, evolution and dissemination of soybean. In: Wang L, Guo Q (eds) *Contemporary soybean research in China*. Jindun, Beijing
- Wang C, Sun S, Jin S, Li W, Wu C, Hou W, Han T (2013) Genetic diversity analysis of widely-planted soybean varieties from different decades and major production regions in China. *Acta Agron Sin* 39:1917–1921
- Ustun A, Allen FL, English BC (2001) Genetic progress in soybean of the US Midsouth. *Crop Sci* 41:993–998
- Morrison MJ, Voldeng HD, Cober ER (2000) Agronomic changes from 58 years of genetic improvement of short-season soybean cultivars in Canada. *Agron J* 92:780–784
- Wang Z, Kang B, Deng S, Liu W, Yao T (1993) Analysis of changing trends of widely-grown soybean cultivars released in different years in Jilin province. *J Jilin Agric Univ* 15(4):92–95
- Hu G (2013) Soybean production in Northeast China. In: Hu G (ed) *Comprehensive soybean producing techniques for decreasing input and increasing income*. China Agric, Beijing
- Gibson LR, Mullen RE (2001) Mineral concentrations in soybean seed produced under high day and night temperature. *Can J Plant Sci* 81(4):595–600
- Song W, Yang R, Wu T, Wu C, Sun S, Zhang S, Jiang B, Tian S, Liu X, Han T (2016) Analysing the effects of climate factors on soybean protein, oil contents and compositions by extensive and high-density sampling in China. *J Agric Food Chem* 64(20):4121–4130
- Zhang J, Ge Y, Han F, Li B, Yan S, Sun J, Wang L (2014) Isoflavone content of soybean cultivars from maturity group 0 to VI grown in Northern and Southern China. *J Am Oil Chem Soc* 91:1019–1028
- Fehr WR, Caviness CE (1977) *Stages of Soybean Development*. Special Report 80, Cooperation Extension Service, Agriculture and home economic experiment station. Iowa State University, Ames, Iowa, US, pp 1–11
- GB/T 5006-1985 (1985) Method for the determination of crude starch in cereals seeds. Standardization Administration of the People's Republic of China, National Standard of the People's Republic of China, Beijing
- GB/T 22491-2008 (2008) Method for the determination of oligo-saccharide in soybean seeds. Standardization Administration of the People's Republic of China, National Standard of the People's Republic of China, Beijing
- GB/T 14609-2008 (2008) Method for the determination of copper, iron, zinc, calcium and magnesium in cereal and crop seeds. Standardization Administration of the People's Republic of China, National Standard of the People's Republic of China, Beijing

27. Seguin P, Turcotte P, Tremblay G, Pageau D, Liu W (2009) Tocopherols concentration and stability in early maturing soybean genotypes. *Agron J* 101:1153–1159
28. Seguin P, Tremblay G, Pageau D, Liu W, Turcotte P (2011) Soybean lutein concentration: impact of crop management and genotypes. *Crop Sci* 51:1151–1160
29. Sun J, Sun B, Han F, Yan S, Yang H, Kikuchi A (2011) Rapid HPLC method for determination of 12 isoflavone components in soybean seeds. *Agric Sci China* 10:70–77
30. Yang X, Dong C, Ren G (2011) Effect of soyasaponins-rich extract from soybean on acute alcohol-induced hepatotoxicity in mice. *J Agric Food Chem* 59(4):1138–1144
31. Lesnefsky EJ, Stoll MK, Minkler PE, Hoppel CL (2000) Separation and quantitation of phospholipids and lysophospholipids by high-performance liquid chromatography. *Anal Biochem* 285:246–254
32. Anta L, Luisa MM, Garcia MC (2010) Simultaneous and rapid determination of the anticarcinogenic proteins Bowman-Birk inhibitor and lectin in soybean crops by perfusion RP-HPLC. *J Chromatogr A* 1217:7138–7143
33. PROC MIXED Statement. SAS/STAT(R) 13.2 User's Guide (2014) http://support.sas.com/documentation/cdl/en/statug/67523/HTML/default/viewer.htm#statug_mixed_syntax01.htm. Accessed 5 Aug 2014
34. The CORR Procedure: PROC CORR Statement. Base SAS(R) 9.2 Procedures Guide: Statistical Procedures (2013) http://support.sas.com/documentation/cdl/en/procstat/63104/HTML/default/viewer.htm#procstat_corr_sect004.htm. Accessed 17 Dec 2013
35. Husson F, Josse J, Le S, Mazet J (2013) FactoMineR: multivariate exploratory data analysis and data mining with R, R package version 1.24
36. Kim SL, Berhow MA, Kim JT, Chi HY, Lee SJ, Chung IM (2006) Evaluation of soyasaponin, isoflavone, protein, lipid, and free sugar accumulation in developing soybean seeds. *J Agric Food Chem* 54:10003–10010
37. Qiu D, Vuong T, Valliyodan B, Shi H, Guo B, Shannon JG, Nguyen HT (2015) Identification and characterization of a stachyose synthase gene controlling reduced stachyose content in soybean. *Theor Appl Genet* 128:2167–2176
38. Kamal-Eldin A, Appelqvist LA (1996) The chemistry and antioxidant properties of tocopherols and tocotrienols. *Lipids* 31:671–701
39. Dwiyantri MS, Ujji A, Thuy LTB, Yamda T, Kitamura K (2007) Genetic analysis of high α -tocopherol content in soybean seeds. *Breed Sci* 57:23–28
40. Bramley PM, Elmadfa I, Kafatos A, Kelly FJ, Manios Y, Roxborough HE, Schuch W, Sheehy PJA, Wagner KH (2000) Vitamin E. *Sci Food Agric* 80:913–938
41. Li H, Liu H, Han Y, Wu X, Teng W, Liu G, Li W (2010) Identification of QTL underlying vitamin E contents in soybean seed among multiple environments. *Theor Appl Genet* 120:1405–1413
42. Smiciklas KD, Mullen RE, Carlson RE, Knapp AD (1989) Drought-induced stress effect on soybean seed calcium and quality. *Crop Sci* 29(6):1519–1523
43. List GR, Evans CD, Warner K, Beal RE, Kwolek WF, Black LT, Moulton KJ (1977) Quality of oil from damaged soybeans. *J Am Oil Chem Soc* 54:8–14
44. Lee J, Renita M, Fioritto RJ, St Martin SK, Schwartz SJ, Vodovotz Y (2004) Isoflavone characterization and antioxidant activity of ohio soybeans. *J Agric Food Chem* 52:2647–2651
45. Liu K, Brown EA (1996) Fatty acid compositions in newly differentiated tissues of soybean seedlings. *J Agric Food Chem* 44:1395–1398
46. Panda PK, Mukhopadhyay S, Behera B, Bhol CS, Dey S, Das DN, Sinha N, Bissoyi A, Pramanik K, Maiti TK, Bhutia SK (2014) Antitumor effect of soybean lectin mediated through reactive oxygen species-dependent pathway. *Life Sci* 111(s1–2):27–35
47. Wang H, Murphy P (1994) Isoflavone composition of American and Japanese soybeans in Iowa: effects of variety, crop year, and location. *J Agric Food Chem* 42:1674–1677
48. Tsukamoto C, Shimada S, Igita K, Kudou S, Kakubun M, Okubo K, Kitamura K (1995) Factors affecting isoflavone content in soybean seeds: change in isoflavone saponins, and composition of fatty acids at different temperature during seed development. *J Agric Food Chem* 43:1184–1192