

Compositional differences between conventional Chinese and genetically modified Roundup Ready soybeans

Yimiao Xia^A, Fusheng Chen^{ID A,B}, Kunlun Liu^A, Lifan Zhang^A, Xiaojie Duan^A, Xin Zhang^A, and Zhenya Zhu^A

^ACollege of Food Science and Technology, Henan University of Technology, Zhengzhou 450001, China.

^BCorresponding author. Email: fushengc@haut.edu.cn

Abstract. Glyphosate-tolerant genetically modified (GM) soybeans (*Glycine max* (L.) Merr.), known commercially as Roundup Ready soybeans, dominate oil consumption and are partly used for protein intake in China. Chemical composition of soybean seed determines its nutritional value, its processing suitability for various protein products, and market decisions. We conducted a compositional comparison of eight GM and 16 conventional Chinese representative soybean varieties. Crude protein, crude fat, moisture, ash, carbohydrate, crude fibre, amino acid and fatty acid contents of the different soybean genotypes were compared and analysed. The GM soybeans had the highest oil concentration but poorer quality, whereas conventional soybeans from the Huanghuaihai region of China showed significantly higher protein, total amino acid, essential amino acid and oleic acid contents, and lower n-6 : n-3 ratio and carbohydrate content, which suggested superior nutritional value. Principal component analysis indicated that protein, carbohydrates and amino acids (except tryptophan, methionine, tyrosine, histidine and proline) contributed most to distinguishing GM soybeans from conventional Chinese soybeans. Differences among the GM and conventional soybeans collected from two major producing regions in China can help to guide manufacturing processes and market decisions with respect to soybeans. High protein and amino acid content in conventional Chinese soybeans mean the potential to expand and improve the International Life Sciences Institute Crop Composition Database used for safety assessment of GM soybean.

Additional keywords: difference analysis, nutrients, nutritional composition, substantial equivalence, transgenic.

Received 6 January 2019, accepted 26 April 2019, published online 24 June 2019

Introduction

Soybeans (*Glycine max* (L.) Merr.) originated in China and serve as a vital source of nutrition for human health (Rizzo and Baroni 2018). Genetically modified (GM) soybeans have been commercially cultivated in America since 1996 and their use has spread globally, with GM soybean becoming the highest yielding soybean. Because of increasing domestic consumption demand, Chinese soybean imports have continued to show sustained growth. In 2017, China imported ~93.50 Mt, accounting for more than 60% of total global soybean imports by quantity, whereas Chinese soybean production was only 12.90 Mt in the same year (USDA 2018). Furthermore, planting of GM soybean was forbidden in China, and most of the imported soybeans have the glyphosate-tolerance characteristic. Therefore, GM soybean has a predominant role in Chinese soybean consumption.

In China, GM soybean is mainly used for oil extraction, with residual soybean meal used as animal feed. GM soybean has also been used in the manufacturing of soybean protein products (Wang and Zhu 2016; Song *et al.* 2018). For soybean products, high protein, essential amino acid (EAA), cysteine, methionine, fat and oleic acid contents, and a high n-3 : n-6 fatty acid ratio, are positively correlated with superior nutrition, whereas high fibre

and carbohydrate contents are negatively correlated with nutritional performance (Hoffman and Falvo 2004; Krishnan 2005; Rizzo and Baroni 2018). The biomass composition of soybean products such as refined oil and meal for animal feed can be predicted from the soybean-seed profile (Karr-Lilienthal *et al.* 2004; Amaral *et al.* 2008). Moreover, the quality attributes of traditional soybean food, such as tofu and soymilk, are reportedly usually affected by soybean-seed biomass composition (Yang and James 2013; James and Yang 2014; Ma *et al.* 2015). Soybean-seed composition characteristics are also essential to making market decisions (McClure *et al.* 2017). Therefore, understanding differences in compositional characteristics between Chinese traditional soybean and imported GM soybean can inform development of the soybean market and guide further processing efforts, with the aims of using soybean more efficiently and developing soybean foods of high nutritional value.

Many comparisons of the compositions of GM soybean and conventional soybean from the United States and Brazil have been conducted (Padgett *et al.* 1996; Harrigan *et al.* 2007, 2010; Zhou *et al.* 2011; Bohn *et al.* 2014; Galão *et al.* 2014). The chemical compositions of soybeans from Brazil, China and the United States have also been compared without discussing

soybean transgenic characters (Grieshop and Fahey 2001; Karr-Lilienthal *et al.* 2004). However, specific research aimed at supplying data for a comparison of GM soybean and conventional Chinese counterparts has rarely been reported.

Consequently, this research aimed to determine compositional differences between conventional Chinese soybean and imported glyphosate-tolerant GM soybean. Soybeans from the northeast single-cropping spring-planting eco-region (NEC region) and the Huanghuaihai double-cropping, spring- and summer-planting eco-region (HHC region) of China account for ~35% and 30% of total Chinese soybean production, respectively (Gai and Wang 2001; NBSC 2017). Accordingly, soybean samples for this work were collected from these two typical soybean-producing regions as representative of conventional Chinese soybean samples. Principal component analysis (PCA) was performed to determine the most important variables for differentiating GM and traditional Chinese soybeans. The potential contribution of adding Chinese conventional soybean composition data to the International Life Sciences Institute Crop Composition Database (ILSI-CCDB), which is used for GM soybean safety assessment, is also discussed.

Materials and methods

Biological material

Eight soybean varieties were collected from the NEC region: Huajiang4, Kenfeng22, Suinong43, Suinong50, Heinong66, 4139, Suizhongzuo40 and Heinong68 (numbered 1–8, respectively). Eight soybean varieties were also collected from the HHC region: Zhonghuang13, Jidou12, Zheng120, Zheng9525, He14, Zhoudou23, Xu9302 and Zhoudou22 (numbered 9–16). GM soybean seeds were obtained from Henan Sunshine Oils and Fats Group, Zhengzhou, and Henan Shennong Feed Technology Company, Zhengzhou. Of these seeds, four types were imported from United States and the other four from Brazil. An immunochromatography test strip assay showed that 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) was positive for glyphosate resistance for all GM soybean seeds. Therefore, the GM soybeans were identified as Roundup Ready (Monsanto, St. Louis, MO, USA), and they were numbered 17–24, respectively.

Nutritional analyses

Seed samples were first ground in a 1095 Knifetec sample mill (FOSS, Höganäs, Sweden) with liquid nitrogen to avoid oil loss, and then passed through a 60-mesh sieve (0.250 mm) and stored at –20°C for further analysis.

Soybean seeds were dried to a constant weight in a DZF vacuum oven (Yongguangming Medical Equipment Co., Beijing) at 105°C for 3 h. Moisture content was determined from the weight loss. The Kjeldahl method was used to analyse the nitrogen content (Kjeltec 2300 Analyzer Unit; FOSS), and the corresponding crude protein content was calculated by multiplying by a factor of 6.25. Total oil was extracted with petroleum ether (bp 30–60°C) for 6.5 h in a 2050 Soxtec auto extraction unit (FOSS), with the extraction temperature set to 75°C, and soaking, solvent reflux, evaporation and drying times of 2.5 h, 2.5 h, 59 min and 30 min, respectively. Crude fat content

was determined gravimetrically from the residual crude fat after drying. Ash content was determined gravimetrically from residual remaining ash after combustion at 560°C for 6 h. Carbohydrate content was calculated by using Eqn 1, in which all components were fresh weight-derived data (OECD 2012):

$$\text{Carbohydrate\%} = 100 - (\text{crude protein\%} + \text{crude fat\%} + \text{ash\%} + \text{moisture\%}) \quad (1)$$

Crude fibre content was investigated with a Fibretec 2010 system (FOSS) according to manufacturer's instructions. Data were converted from fresh weight to dry weight by using the moisture content.

Amino acid analysis

The amino acid composition was analysed according to the methods of Nielsen and Hurrell (1985) and Liu (1994). Samples of 20–25 mg were used in this work. Samples were subjected to acid hydrolysis, and the contents of 16 amino acids were determined: threonine (Thr), valine (Val), methionine (Met), isoleucine (Ile), leucine (Leu), phenylalanine (Phe), lysine (Lys), aspartate (Asp), serine (Ser), glutamate (Glu), glycine (Gly), alanine (Ala), tyrosine (Tyr), histidine (His), arginine (Arg) and proline (Pro). Quantities of Asp and Glu included the amino acid and its corresponding acid. Samples were also subjected to alkaline hydrolysis for determination of the tryptophan (Trp) content.

Fatty acid analysis

Fatty acid analysis was performed via the protocol of Galão *et al.* (2014). The fatty acid composition was calculated and expressed by using the area normalisation method.

Statistical analyses

Measurements were performed as three replicates. One-way analysis of variance (ANOVA) was used to analyse the differences in nutrients among soybean groups, and multiple comparisons of means were conducted by the least significance difference (l.s.d.) test, with significant differences determined at $P = 0.05$. PCA was used to investigate the nutrition variables that contributed most to differentiating NEC, HHC and GM soybeans. SPSS software version 21.0 (IBM, Armonk, NY, USA) was used for data analysis.

Results and discussion

Comparison of the general nutritional characteristics of NEC, HHC and GM soybeans

The GM soybeans had nutrient compositions different from those of NEC and HHC soybeans (Table 1). The GM samples contained significantly higher moisture, fat and fibre contents than conventional Chinese soybeans, whereas protein and carbohydrate contents showed the opposite trend. Ash contents did not differ among groups. The higher oil contents of GM soybeans facilitated soybean oil extraction, which is important considering that almost all imported soybeans are used for oil processing in China. HHC soybeans had significantly higher protein content than NEC and GM

Table 1. Nutrient composition of different soybean genotypes

Results expressed as means \pm standard deviation. NEC soybean is conventional soybean from the northeast single-cropping spring-planting eco-region of China; HHC soybean is conventional soybean from the Huanghuaihai double-cropping spring- and summer-planting eco-region of China; GM soybean is Roundup Ready genetically modified soybean. Within rows, means with the same letter are not significantly different at $P=0.05$. EAAs, Essential amino acids (threonine–tryptophan in the first column of the Table); n.s., no significant difference

Index	NEC soybean		HHC soybean		GM soybean		ANOVA <i>P</i> -value
	Range	Mean \pm s.d.	Range	Mean \pm s.d.	Range	Mean \pm s.d.	
Moisture (g 100 g ⁻¹ fresh weight)	5.67–7.31	6.63 \pm 0.58a	6.18–7.32	6.58 \pm 0.43a	6.57–8.95	7.71 \pm 0.88b	0.04
<i>Broad compositional analysis (g 100 g⁻¹ dry weight)</i>							
Crude protein	37.59–42.10	40.10 \pm 1.69a	42.00–48.57	44.63 \pm 2.12b	37.95–42.22	40.54 \pm 1.44a	0.00006
Crude fat	18.69–21.31	20.15 \pm 1.00a	16.99–20.75	18.50 \pm 1.44b	19.86–24.75	22.63 \pm 1.77c	0.00005
Ash	4.67–5.18	4.98 \pm 0.19a	4.37–5.50	5.09 \pm 0.36a	4.91–5.88	5.26 \pm 0.43a	n.s.
Crude fibre	4.82–5.75	5.32 \pm 0.31a	4.76–6.19	5.37 \pm 0.49a	4.90–6.46	5.77 \pm 0.55b	0.04
Carbohydrates	32.35–36.22	34.78 \pm 1.26a	28.62–35.82	31.78 \pm 2.35b	29.25–33.99	31.57 \pm 1.67b	0.003
<i>Amino acid content (g 100 g⁻¹ fresh weight)</i>							
Threonine	1.48–1.67	1.58 \pm 0.07a	1.67–1.83	1.75 \pm 0.06b	1.48–1.76	1.63 \pm 0.08a	0.0004
Valine	1.98–2.19	2.08 \pm 0.08a	2.21–2.43	2.32 \pm 0.09b	1.96–2.29	2.17 \pm 0.10a	0.0004
Methionine	0.19–0.54	0.43 \pm 0.11a	0.32–0.58	0.48 \pm 0.09a	0.20–0.55	0.39 \pm 0.13a	n.s.
Isoleucine	1.91–2.12	1.99 \pm 0.07a	2.09–2.37	2.22 \pm 0.10b	1.92–2.26	2.13 \pm 0.11b	0.0003
Leucine	3.03–3.34	3.17 \pm 0.13a	3.40–3.76	3.56 \pm 0.13b	3.03–3.61	3.39 \pm 0.18c	0.0001
Phenylalanine	1.98–2.23	2.08 \pm 0.10a	2.25–2.63	2.46 \pm 0.13b	1.89–2.44	2.25 \pm 0.18c	0.00009
Lysine	2.53–2.78	2.65 \pm 0.08a	2.68–3.13	2.91 \pm 0.14b	2.47–2.87	2.68 \pm 0.14a	0.0005
Tryptophan	0.27–0.33	0.31 \pm 0.02a	0.27–0.34	0.32 \pm 0.02a	0.30–0.33	0.32 \pm 0.01a	n.s.
Aspartic acid	4.56–5.03	4.82 \pm 0.17a	5.21–5.95	5.51 \pm 0.28b	4.63–5.33	5.00 \pm 0.26a	0.00003
Serine	1.81–2.11	1.98 \pm 0.10a	2.08–2.48	2.21 \pm 0.12b	1.91–2.22	2.06 \pm 0.10a	0.001
Glutamic acid	7.56–8.38	7.99 \pm 0.27a	8.60–9.48	8.91 \pm 0.34b	7.44–8.82	8.34 \pm 0.44a	0.0001
Glycine	1.71–1.88	1.78 \pm 0.06a	1.81–2.09	1.98 \pm 0.09b	1.65–1.95	1.83 \pm 0.09a	0.0002
Alanine	1.73–1.92	1.81 \pm 0.07a	1.78–2.12	2.02 \pm 0.11b	1.69–2.02	1.90 \pm 0.10a	0.001
Tyrosine	1.46–1.71	1.54 \pm 0.09a	1.60–1.91	1.74 \pm 0.11b	1.48–1.81	1.71 \pm 0.11b	0.002
Histidine	1.21–1.60	1.45 \pm 0.11a	1.41–1.76	1.64 \pm 0.11b	1.26–1.58	1.47 \pm 0.11a	0.006
Arginine	2.74–3.27	2.98 \pm 0.18a	3.26–3.98	3.55 \pm 0.24b	2.88–3.26	3.10 \pm 0.12a	0.00002
Proline	2.04–2.50	2.30 \pm 0.17a	1.81–2.83	2.32 \pm 0.35a	1.73–2.26	2.01 \pm 0.19b	0.04
Sum of EAAs	13.74–14.90	14.29 \pm 0.51a	15.15–16.88	16.03 \pm 0.63b	13.45–15.80	14.96 \pm 0.77c	0.0001
<i>Fatty acid content (%)</i>							
16:0, palmitic acid	10.19–11.62	10.92 \pm 0.52ab	10.22–11.97	11.08 \pm 0.55a	10.00–10.88	10.43 \pm 0.35b	0.04
18:0, stearic acid	4.16–5.36	4.58 \pm 0.40a	3.54–4.47	3.91 \pm 0.33b	3.86–4.19	3.98 \pm 0.14b	0.0004
18:1, oleic acid	18.91–23.11	20.75 \pm 1.43a	22.00–30.38	25.93 \pm 3.01b	22.63–24.70	23.68 \pm 0.86c	0.0002
18:2, n-6, linoleic acid	52.31–55.33	53.67 \pm 0.91a	47.77–53.81	51.00 \pm 2.14b	53.19–55.99	54.42 \pm 1.18a	0.0004
18:3, n-3, linolenic acid	7.95–10.23	9.02 \pm 0.92a	5.88–9.17	7.07 \pm 1.16b	6.22–6.84	6.42 \pm 0.24b	0.00002
20:0, arachidic acid	0.38–0.48	0.43 \pm 0.04a	0.29–0.45	0.38 \pm 0.05b	0.34–0.41	0.38 \pm 0.03b	0.02
20:1, eicosenoic acid	0.15–0.21	0.18 \pm 0.02a	0.18–0.26	0.21 \pm 0.02b	0.22–0.26	0.23 \pm 0.01c	0.0001
22:0, behenic acid	0.35–0.49	0.44 \pm 0.04a	0.26–0.58	0.42 \pm 0.10a	0.39–0.52	0.46 \pm 0.05a	n.s.
n-6 : n-3	5.29–6.79	6.00 \pm 0.59a	5.87–8.46	7.34 \pm 0.95b	7.96–8.99	8.48 \pm 0.38c	0.000002
Sum saturated	15.63–17.06	16.38 \pm 0.50a	14.82–16.40	15.78 \pm 0.61b	14.78–15.65	15.25 \pm 0.35c	0.0007
Sum monounsaturated	19.06–23.29	20.93 \pm 1.45a	22.19–30.59	26.14 \pm 3.02b	22.88–24.92	23.91 \pm 0.85c	0.0002
Sum polyunsaturated	60.46–64.41	62.69 \pm 1.35 ^a	54.45–62.99	58.08 \pm 3.05 ^b	59.52–62.25	60.84 \pm 1.16 ^a	0.0007

soybeans, corresponding to the higher EAA and total amino acid contents in HHC soybeans. In addition, low fibre and carbohydrate contents characterised HHC soybeans as a superior protein source (Grieshop and Fahey 2001; Rizzo and Baroni 2018). Chinese soybeans have been previously shown to have higher protein and lower lipid contents than soybeans from Brazil and the United States, and a negative correlation has been observed between soybean oil and protein contents (Grieshop and Fahey 2001; Karr-Lilienthal *et al.* 2004; Qin *et al.* 2014; Song *et al.* 2016). Those results are reinforced by the present study. Methionine, which contains sulfur, is critical for animal health (Krishnan 2005). However, Table 1 shows that mass concentration differences of Met were not significant ($P > 0.05$) among NEC, HHC and GM soybeans. Therefore, the total protein and EAAs

contents were crucial for evaluating the nutritional value of soybeans, with the higher protein and EAAs contents in HHC soybeans predicting a better nutrient performance. Furthermore, the higher protein content of HHC soybeans indicated limited applicability for production of soymilk (Ma *et al.* 2015) and natto (Yoshikawa *et al.* 2014).

The NEC soybeans had both the highest average saturated fatty acid (SFA) and linolenic acid contents. High SFA content has negative effects on human health, and linolenic acid can promote oil oxidation. By contrast, oleic acid improves the nutritional value and oxidative stability, with the oleic acid content of HHC soybeans shown to be significantly higher than of NEC and GM soybeans. Therefore, HHC soybeans had superior nutrient composition compared with NEC

soybeans. Furthermore, a high omega-6 (linoleic acid) intake and low omega-3 (oleic acid) intake presents an increased risk of obesity (Bonafe *et al.* 2011; Carbonera *et al.* 2014; Bøhn *et al.* 2014). A lower n-6 : n-3 ratio was observed in NEC and HHC soybeans; therefore, conventional soybean varieties were more likely to perform better than GM varieties with respect to consumer health.

PCA of crude protein, fat and fibre, and moisture, ash and carbohydrate

Two principal components (PCs) were selected to differentiate NEC, HHC and GM soybean samples. The contribution rates of PC1 and PC2 were 31.42% and 26.88%, respectively. The NEC and HHC soybean varieties partially overlapped, and GM soybean varieties were well distinguished from the conventional soybeans (Fig. 1a). Clearly, from the PC1 dimension, HHC soybean scores were overall lower than NEC soybean scores, which were in turn lower than GM soybean scores. From the PC2 dimension, NEC soybean scores were overall lowest, and the GM soybean scores were highest. All score differences contributed to the distribution distinction shown in Fig. 1a.

Each PC score was calculated by using a formula consisting of standardised data for original nutrition indices, and summarised the variability of the original data (Peng *et al.* 2014; Galão *et al.* 2014). The importance of each index to the PC was evaluated by using loading values. Figure 1b implies that crude protein contributed most to PC1, whereas carbohydrates contributed most to PC2. Given that ~60% of the variance in the original data was explained by PC1 and PC2, crude protein and carbohydrates were the most important variables differentiating NEC, HHC and GM soybean samples compared with other indices.

PCA of amino acid data

The average amino acid PC1 score increased from NEC to GM to HHC soybeans, whereas the PC2 scores of the three groups were equal (Fig. 2a). The groups NEC, GM and HHC were

distinguished from each other to some extent from the PC1 axis. Furthermore, PC1 explained ~70% of variance in the original data and was therefore the main contributor to distribution differences among the NEC, GM and HHC soybean groups. The loading graphic (Fig. 2b) indicated that Asp, Ser, Glu, Gly, Ala, Arg and EAAs (except Trp and Met) mainly represented PC1. Accordingly, Asp, Ser, Glu, Gly, Ala, Arg and EAAs (except Trp and Met) contributed most to differences among NEC, GM and HHC soybean samples.

Because all 17 amino acids reported in this study contributed positively to PC1, the PC1 scores showed obvious positive correlations with the mass fraction of amino acids. The PC1 scores in Fig. 2a demonstrate that the mass fraction of amino acids in soybean samples decreased in the order HHC > GM > NEC soybeans, in agreement with the potential nutritional quality in soybean-derived food intake.

PCA of fatty acid data

According to Fig. 3a, the NEC soybean group was well distinguished from the HHC and GM soybean groups, whereas the latter two groups overlapped. PC1, rather than PC2, played a vital role in distinguishing the NEC group. The loading graphic (Fig. 3b) showed that C18:3, C18:1 and C20:1 were the most important components of PC1. Therefore, C18:3, C18:1 and C20:1 explained the separation of NEC soybeans from HHC and GM soybean samples with respect to fatty acid content.

For visual assessment of the fatty acid saturation of NEC, HHC and GM soybean samples, SFA, monounsaturated fatty acid (MUFA) and polyunsaturated fatty acid (PUFA) contents were set as independent variables to conduct PCA among these three groups. From the score graphic (Fig. 3c), the PC1 scores were ranked in the order HHC < GM < NEC soybeans; the PC2 scores for GM soybean samples were lower than those of the conventional soybeans. The loading graphic (Fig. 3d) showed that PC1 was mainly negatively correlated with MUFA and positively correlated with PUFA, whereas PC2 was mostly positively correlated with SFA. Accordingly, the MUFA

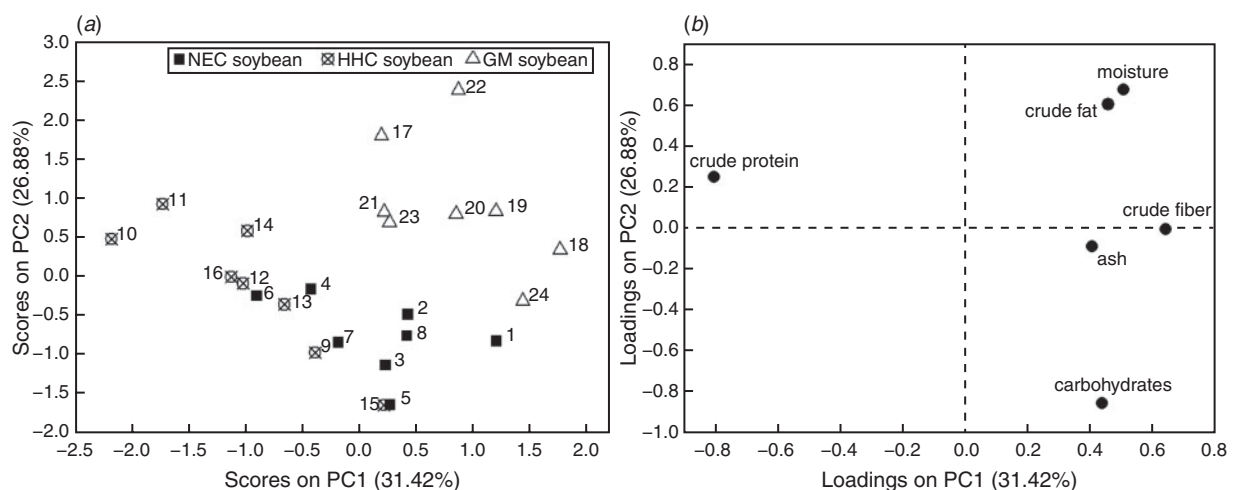


Fig. 1. (a) Score and (b) loading plots of soybeans based on broad compositional data. NEC soybean is conventional soybean from the northeast single-cropping spring-planting eco-region of China; HHC soybean is conventional soybean from the Huanghuaihai double-cropping spring- and summer-planting eco-region of China; GM soybean, Roundup Ready genetically modified soybean.

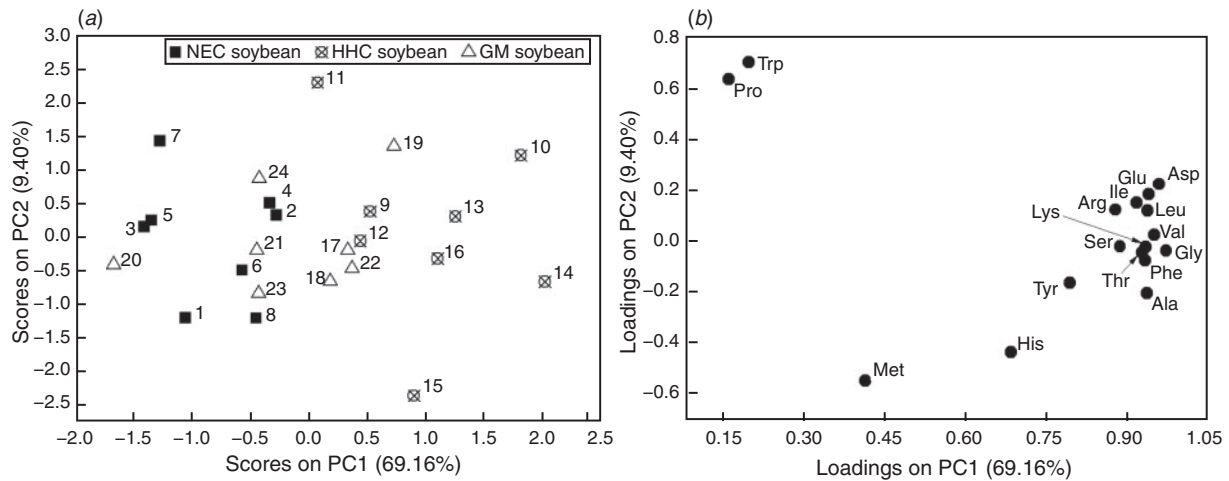


Fig. 2. (a) Score and (b) loading plots of soybeans based on amino acid data. NEC soybean is conventional soybean from the northeast single-cropping spring-planting eco-region of China; HHC soybean is conventional soybean from the Huanghuaihai double-cropping spring- and summer-planting eco-region of China; GM soybean, Roundup Ready genetically modified soybean. Thr, Threonine; Val, valine; Met, methionine; Ile, isoleucine; Leu, leucine; Phe, phenylalanine; Lys, lysine; Trp, tryptophan; Asp, aspartic acid; Ser, serine; Glu, glutamic acid; Gly, glycine; Ala, alanine; Tyr, tyrosine; His, histidine; Arg, arginine; Pro, proline.

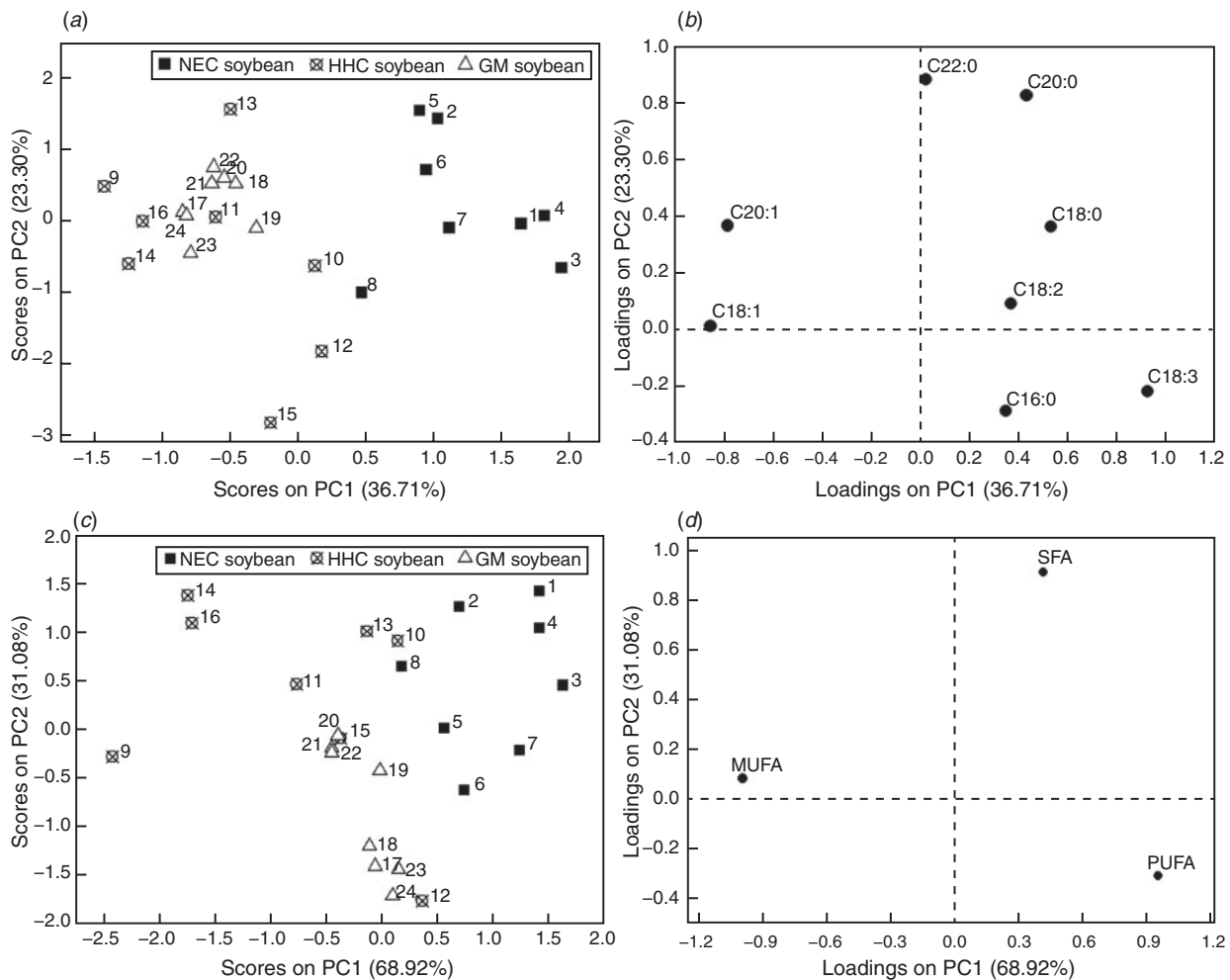
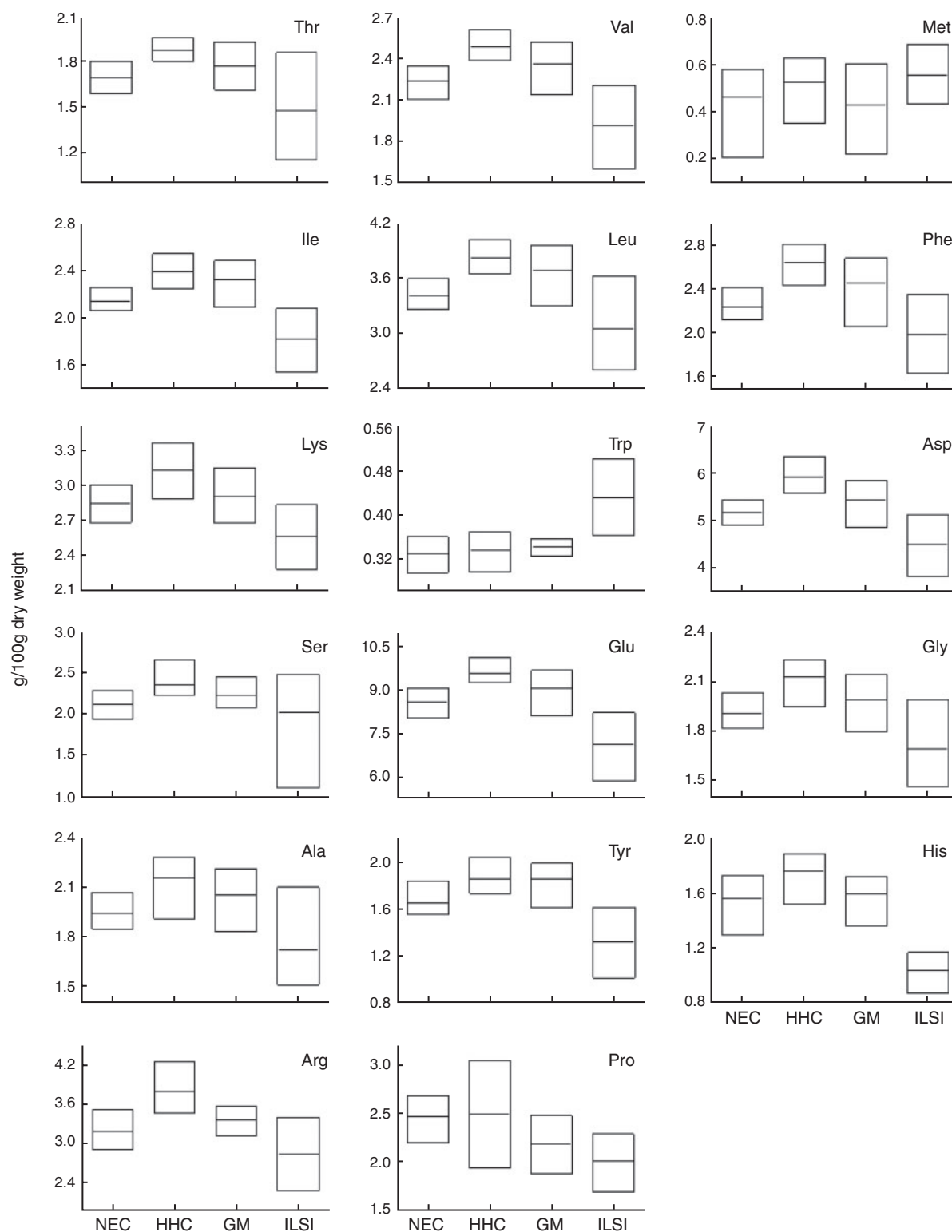


Fig. 3. (a, c) Score and (b, d) loading plots of soybeans based on fatty acid data. NEC soybean is conventional soybean from the northeast single-cropping spring-planting eco-region of China; HHC soybean is conventional soybean from the Huanghuaihai double-cropping spring- and summer-planting eco-region of China; GM soybean, Roundup Ready genetically modified soybean. 16:0, Palmitic acid; 18:0, stearic acid; 18:1, oleic acid; 18:2, linoleic acid; 18:3, linolenic acid; 20:0, arachidic acid; 20:1, eicosenoic acid; 22:0, behenic acid; SFA, saturated fatty acid; MUFA, monounsaturated fatty acid; PUFA, polyunsaturated fatty acid.



Soybean origin regions, genotypes, and data from ILSI

Fig. 4. Comparison of maximums, minimums and means of amino acids in GM and conventional soybean cultivars. NEC soybean is conventional soybean from the northeast single-cropping spring-planting eco-region of China; HHC soybean is conventional soybean from the Huanghuaihai double-cropping spring- and summer-planting eco-region of China; GM soybean, Roundup Ready genetically modified soybean; ILSI, International Life Sciences Institute. Thr, Threonine; Val, valine; Met, methionine; Ile, isoleucine; Leu, leucine; Phe, phenylalanine; Lys, lysine; Trp, tryptophan; Asp, aspartic acid; Ser, serine; Glu, glutamic acid; Gly, glycine; Ala, alanine; Tyr, tyrosine; His, histidine; Arg, arginine; Pro, proline. Data were converted from fresh weight to dry weight basis by using the given moisture level.

contents of different soybean groups were in the order HHC > GM > NEC, and PUFA contents showed the opposite order. SFA content was lowest in the GM soybeans, and SFA contents of the two Chinese conventional soybean groups were almost the same. SFA allowed GM soybean to be distinguished from NEC and HHC soybeans, but failed to distinguish the NEC and HHC soybean groups. By contrast, MUFA and PUFA more efficiently differentiated the NEC, HHC and GM soybean samples.

Comparison of the SFA, MUFA and PUFA contents by one-way ANOVA is shown in Table 1. The results were consistent with the PCA conclusions; however, PCA had the advantage of allowing visualisation of the sample distribution in the score graphic and verifying the most important variable to differentiate the original data.

Effect of compositional differences on GM soybean safety assessment

A concept of substantial equivalence has been proposed and accepted worldwide as a guideline to assess the safety of novel foods derived from biotechnology (OECD 1998; FAO/WHO 2000; Kuiper *et al.* 2002; Alba *et al.* 2010; Kitta 2013). When the nutritional composition of novel foods falls within the range of the corresponding index of conventional counterparts, substantial equivalence is verified. Therefore, obtaining high-quality and comprehensive compositional information for natural varieties is of great importance. The ILSI-CCDB plays an important role in building compositional databases of conventionally bred crops and provides open access to global visitors (OECD 2012). For soybean, the field trial locations of the ILSI were distributed in South America (Argentina, Brazil, Chile), North America (Canada, Puerto Rico, United States) and the European Union (France, Italy, Romania) (Sult *et al.* 2016). However, China, as the biological origin and a major growing region of soybean, has yet to be included in this database. Objective reasons might have contributed to this exclusion, and the potential contribution made by compiling Chinese soybean compositional-characteristic data into the ILSI-CCDB will be discussed below.

Although significant differences were observed with respect to moisture, crude protein, crude fat, crude fibre and carbohydrate contents among NEC, HHC and GM soybean groups (Table 1), certain target contents in GM soybean fall within the range of corresponding indices supplied by the ILSI-CCDB. For example, the GM soybean crude protein content (37.95–42.22%) was within the range 29.51–46.60% for protein in the ILSI-CCDB. Therefore, GM soybeans tested in this study were substantially equivalent to conventional soybeans based on these data.

Furthermore, a compositional comparison was conducted between the ILSI-CCDB and data from NEC and HHC soybean varieties. The highest protein content was found in HHC soybean (48.57%), which was above the corresponding range given by the ILSI-CCDB (29.51–46.60%). Similar results for crude protein content were also found previously in soybeans from the HHC (46.82%), the northwest spring-planting region in China (48.96%), and the south cropping region in China (49.78%) (Song *et al.* 2016). Therefore, crude protein data for

Chinese soybean cultivars would probably contribute to expanding and improving the ILSI-CCDB.

For individual amino acids, the maximum contents of Thr, Val, Ile, Leu, Phe, Lys, Asp, Glu, Gly, Ala, Tyr, His, Arg and Pro, and the minimum contents of Met, Trp and His, in the GM soybean cultivars (Fig. 4) were beyond the range of corresponding indices in the ILSI-CCDB, indicating the absence of substantial equivalence in GM soybeans. However, when amino acid content data for NEC and HHC soybean groups were added to the ILSI-CCDB, the substantial equivalence of GM soybeans was verified.

Many reports have emphasised that soybean seed composition is affected by abiotic and biotic factors simultaneously (Padgett *et al.* 1996; Kuiper *et al.* 2002; Harrigan *et al.* 2010; Galão *et al.* 2014; Song *et al.* 2016; McClure *et al.* 2017). Compositional variations among NEC, HHC and GM soybeans reported in this study were partly caused by different environment factors. Whether the soybean compositional variation caused by different environment factors was greater than variation caused by biotechnology-driven genetic modification should be evaluated in future studies. However, China, as the origin and a major producing country of conventional soybean, should be included in the field trial locations used by the ILSI to make the soybean data more comprehensive and to improve the ILSI-CCDB further.

Conclusions

Compositional analyses reported in this study showed that indices such as moisture, crude fat and crude fibre contents in GM soybean varieties were significantly higher than those in NEC and HHC soybean varieties. On the other hand, crude protein contents in HHC soybeans and carbohydrate contents in NEC soybeans were higher than in GM soybeans. Amino acid concentration showed a variation trend similar to crude protein, with the highest total amino acid contents (except Met and Trp) observed in HHC soybean. The highest linoleic content and n-6 : n-3 ratio were observed in GM soybeans, which predicted poorer nutritional value. PCA showed that crude protein, carbohydrates, Asp, Ser, Glu, Gly, Ala, Arg and EAAs (except Met and Trp) contributed most to differentiating NEC, HHC and GM soybeans, and that MUFA and C18:3 allowed NEC soybeans to be efficiently distinguished from HHC and GM soybeans. The compositional characteristics of Chinese soybeans would expand the ILSI-CCDB, and their inclusion is highly recommended.

Conflicts of interest

The authors declare no conflicts of interest.

Acknowledgements

The authors thank Professor Lianzhou Jiang from Northeast Agricultural University, Professor Yongzhan Zheng from Henan Academy of Agricultural Sciences, Henan Sunshine Oils and Fats Group, and Henan Shennong Feed Technology Company greatly for provision of research materials. This research was supported by the National Natural Science Foundation of China (Nos. 21676073, 21376064), Key Program of Natural Science Foundation of He'nan Province (No. 162300410046), and the Outstanding Science and Technology Innovation Team of Henan Province (No. 5 in 2015).

References

- Alba R, Phillips A, Mackie S, Gillikin N, Maxwell C, Brune P, Ridley W, Fitzpatrick J, Levine M, Harris S (2010) Improvements to the International Life Sciences Institute Crop Composition Database. *Journal of Food Composition and Analysis* **23**, 741–748. doi:10.1016/j.jfca.2010.03.007
- Amaral JS, Casal S, Costa J, Mafra I, Oliveira MBPP (2008) Effect of refining on the fatty acid, sterol and tocopherols compositions of soybean oil from GM seeds. *Applied Surface Science* **254**, 7893–7896.
- Bohn T, Cuhra M, Traavik T, Sanden M, Primicerio R (2014) Compositional differences in soybeans on the market: glyphosate accumulates in roundup ready GM soybeans. *Food Chemistry* **153**, 207–215. doi:10.1016/j.foodchem.2013.12.054
- Bonafe EG, Aguiar AC, Boroski M, Monteiro J, Souza NE, Matsushita M, Visentainer J (2011) Quantification of EPA and DHA in seafood of the south coast of Brazil. *Nutrition & Food Science* **41**, 401–411. doi:10.1108/00346651111181958
- Carbonera F, Bonafe EG, Martin CA, Montanher PF, Ribeiro RP, Figueiredo LC, Almeida VC, Visentainer JV (2014) Effect of dietary replacement of sunflower oil with perilla oil on the absolute fatty acid composition in Nile tilapia (GIFT). *Food Chemistry* **148**, 230–234. doi:10.1016/j.foodchem.2013.10.038
- FAO/WHO (2000) Safety aspects of genetically modified foods of plant origin. Report of a Joint FAO/WHO Expert Consultation on Foods Derived from Biotechnology, Geneva, Switzerland. Food and Agriculture Organization of the United Nations, Rome. Available at: <http://agris.fao.org/openagris/search.do?recordID=US201300048645> (accessed 29 May 2019).
- Gai JY, Wang YS (2001) A study on the varietal eco-regions of soybeans in China. *Scientia Agricultura Sinica* **34**, 139–145.
- Galão OF, Carrão-Panizzi MC, Gontijo Mandarino JM, Santos Júnior OO, Maruyama SA, Figueiredo LC, Bonafe EG, Visentainer JV (2014) Differences of fatty acid composition in Brazilian genetic and conventional soybeans (*Glycine max* (L.) Merrill) grown in different regions. *Food Research International* **62**, 589–594. doi:10.1016/j.foodres.2014.04.025
- Grieshop CM, Fahey GC (2001) Comparison of quality characteristics of soybeans from Brazil, China, and the United States. *Journal of Agricultural and Food Chemistry* **49**, 2669–2673. doi:10.1021/jf0014009
- Harrigan GG, Ridley WP, Riordan SG, Nemeth MA, Sorbet R, Trujillo WA, Breeze ML, Schneider RW (2007) Chemical composition of glyphosate-tolerant soybean 40-3-2 grown in Europe remains equivalent with that of conventional soybean (*Glycine max* L.). *Journal of Agricultural and Food Chemistry* **55**, 6160–6168. doi:10.1021/jf0704920
- Harrigan GG, Glenn KC, Ridley WP, Kough JJ (2010) Assessing the natural variability in crop composition. *Regulatory Toxicology and Pharmacology* **58**, S13–S20. doi:10.1016/j.yrtph.2010.08.023
- Hoffman JR, Falvo MJ (2004) Protein—which is best? *Journal of Sports Science & Medicine* **3**, 118–130.
- James AT, Yang A (2014) Influence of globulin subunit composition of soybean proteins on silken tofu quality. 2. Absence of 11SA4 improves the effect of protein content on tofu hardness. *Crop & Pasture Science* **65**, 268–273. doi:10.1071/CP13399
- Karr-Lilienthal LK, Grieshop CM, Merchen NR, Mahan DC, Fahey GC (2004) Chemical composition and protein quality comparisons of soybeans and soybean meals from five leading soybean-producing countries. *Journal of Agricultural and Food Chemistry* **52**, 6193–6199. doi:10.1021/jf049795+
- Kitta K (2013) Availability and utility of crop composition data. *Journal of Agricultural and Food Chemistry* **61**, 8304–8311. doi:10.1021/jf400777v
- Krishnan HB (2005) Engineering soybean for enhanced sulfur amino acid content. *Crop Science* **45**, 454–461. doi:10.2135/cropsci2005.0454
- Kuiper HA, Kleter GA, Noteborn HP, Kok EJ (2002) Substantial equivalence—an appropriate paradigm for the safety assessment of genetically modified foods? *Toxicology* **181–182**, 427–431. doi:10.1016/S0300-483X(02)00488-2
- Liu HJ (1994) Determination of amino acids by precolumn derivatization with 6-aminoquinolyl-N-hydroxysuccinimidyl carbamate and high-performance liquid chromatography with ultraviolet detection. *Journal of Chromatography A* **670**, 59–66. doi:10.1016/0021-9673(94)80280-7
- Ma L, Li B, Han F, Yan S, Wang L, Sun J (2015) Evaluation of the chemical quality traits of soybean seeds, as related to sensory attributes of soymilk. *Food Chemistry* **173**, 694–701. doi:10.1016/j.foodchem.2014.10.096
- McClure T, Cocuron JC, Osmark V, Mchale LK, Alonso AP (2017) Impact of environment on the biomass composition of soybean (*Glycine max*) seeds. *Journal of Agricultural and Food Chemistry* **65**, 6753–6761. doi:10.1021/acs.jafc.7b01457
- National Bureau of Statistics of China (2017) National data. <http://data.stats.gov.cn/easyquery.htm?cn=C01>.
- Nielsen HK, Hurrell RF (1985) Tryptophan determination of food proteins by h.p.l.c. after alkaline hydrolysis. *Journal of the Science of Food and Agriculture* **36**, 893–907. doi:10.1002/jsfa.2740360920
- OECD (1998) Report of the OECD Workshop on the Toxicological and Nutritional Testing of Novel Foods. Aussois, France. Organisation for Economic Co-operation and Development, Paris. Available at: <http://agris.fao.org/agris-search/search.do?recordID=US201300051383> (accessed 29 May 2019).
- OECD (2012) Revised consensus document on compositional considerations for new varieties of soybean [*Glycine max* (L.) Merr]: key food and feed nutrients, anti-nutrients, toxicants and allergens. Series on the Safety of Novel Foods and Feeds, No. 25. Organisation for Economic Co-operation and Development, Paris. Available at: [http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=ENV/JM/MONO\(2012\)24&docLanguage=En](http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=ENV/JM/MONO(2012)24&docLanguage=En) (accessed 23 May 2019).
- Padgett SR, Taylor NB, Nida DL, Bailey MR, Macdonald J, Holden LR, Fuchs RL (1996) The composition of glyphosate-tolerant soybean seeds is equivalent to that of conventional soybeans. *The Journal of Nutrition* **126**, 702–716. doi:10.1093/jn/126.3.702
- Peng X, Li X, Shi X, Guo S (2014) Evaluation of the aroma quality of Chinese traditional soy paste during storage based on principal component analysis. *Food Chemistry* **151**, 532–538. doi:10.1016/j.foodchem.2013.11.095
- Qin P, Song W, Yang X, Sun S, Zhou X, Yang R, Li N, Hou W, Wu C, Han T, Ren G (2014) Regional distribution of protein and oil compositions of soybean cultivars in China. *Crop Science* **54**, 1139–1146. doi:10.2135/cropsci2013.05.0314
- Rizzo G, Baroni L (2018) Soy, soy foods and their role in vegetarian diets. *Nutrients* **10**, 43–93. doi:10.3390/nu10010043
- 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. doi:10.1021/acs.jafc.6b00008
- Song J, Song Q, Wang D, Zhang F (2018) Monitoring the prevalence of genetically modified soybeans in tofu in Chengdu, China using real-time and conventional PCR. *Journal of Food Composition and Analysis* **67**, 172–177. doi:10.1016/j.jfca.2018.01.014
- Sult T, Barthet VJ, Bennett L, Edwards A, Fast B, Gillikin N, Launis K, New S, Rogers-Szuma K, Sabbatini J, Srinivasan JR, Tilton GB, Venkatesh TV (2016) Report: Release of the International Life Sciences Institute Crop Composition Database Version 5. *Journal of Food Composition and Analysis* **51**, 106–111. doi:10.1016/j.jfca.2016.05.002
- USDA (2018) Oilseeds: world markets and trade. United States Department of Agriculture Foreign Agricultural Service, Washington, DC. Available

- at: <https://apps.fas.usda.gov/psdonline/app/index.html#/app/downloads> (accessed 29 May 2019).
- Wang C, Zhu Y (2016) Investigation of transgenic soybean components in soybean from an area of China. *Journal of the Science of Food and Agriculture* **96**, 3169–3172. doi:[10.1002/jsfa.7495](https://doi.org/10.1002/jsfa.7495)
- Yang A, James AT (2013) Effects of soybean protein composition and processing conditions on silken tofu properties. *Journal of the Science of Food and Agriculture* **93**, 3065–3071. doi:[10.1002/jsfa.6140](https://doi.org/10.1002/jsfa.6140)
- Yoshikawa Y, Chen P, Zhang B, Scaboo A, Orazaly M (2014) Evaluation of seed chemical quality traits and sensory properties of natto soybean. *Food Chemistry* **153**, 186–192. doi:[10.1016/j.foodchem.2013.12.027](https://doi.org/10.1016/j.foodchem.2013.12.027)
- Zhou J, Berman KH, Breeze ML, Nemeth MA, Oliveira WS, Braga DP, Berger GU, Harrigan GG (2011) Compositional variability in conventional and glyphosate-tolerant soybean (*Glycine max* L.) varieties grown in different regions in Brazil. *Journal of Agricultural and Food Chemistry* **59**, 11652–11656. doi:[10.1021/jf202781v](https://doi.org/10.1021/jf202781v)