scientific reports

Check for updates

OPEN

Metabonomics analysis of flavonoids in seeds and sprouts of two Chinese soybean cultivars

Weiwei Bi[1,2,3,6]✉, Guixing Zhao[2,6], Yutong Zhou[4], Xiaoyu Xia[2], Jinsheng Wang[2], Guangjin Wang[2], Shuwen Lu[5], Wenjin He[2], Taifei Bi[3] & Jinrong Li[2]

A popular food in China, soybean seeds and sprouts contained many biologically active substances which are beneficial to the human body, such as flavonoids. Northeast of China is the main producing area of soybean. The experimental materials came from the main soybean producing areas in Northeast China, this study compared flavonoids of two China cultivars of soybeans, Heinong52(HN52) and Heinong71(HN71). Here, we also considered the effects of germination on the chemical profile of flavonoids. Using a LC–ESI–MS/MS system, 114 differential flavonoid metabolites were identified. A total of 18 metabolites were significantly different between the two soybean varieties before germination, of which 14 were up-regulated and 4 were down-regulated. After germination, 33 significantly different metabolites were found in the two soybean sprouts, of which 19 were up-regulated and 14 were down-regulated. These experimental results revealed significant up-regulation of metabolites in soybean sprouts compared with soybean seeds, thus suggesting that soybean germination may increase content of flavonoid metabolites. There are six main pathways for the synthesis of flavonoids: isoflavonoid biosynthesis, flavonoid biosynthesis, flavone and flavonol biosynthesis, biosynthesis of secondary metabolites, and biosynthesis of phenylpropanoids. Soybean seeds lack flavone and flavanol biosynthesis and develop the capacity for this biosynthetic pathway after germination as sprouts. Isoflavonoid biosynthesis is the most abundantly utilized pathway.

China is the first country to grow and domesticate the soybean(*Glycine max*(L.) Merr.), with a cultivation history of at least 4,000 years. Nowadays, soybean is gaining popularity in many countries largely as a vegetable protein and oil source, owing to its bean composition of approximately 40% protein and 20% oil[1]. Soybean is an important nutritional component of diets and is used in many foods, such as soybean oil, soybean sprout, paste, soymilk, and tofu[2]. Mostly present in plant leaves and fruits, flavonoids primarily exist as glycosides combining with sugars. An important phenolic secondary metabolite in plants, flavonoids can support plant disease resistance, and in the human body have roles such as anti-oxidation, anti-cancer, and anti-aging[3,4]. Certain major flavonoids, such as flavonols, flavonol glycosides, isoflavones, chalcones, anthocyanins and procyanidins have been previously isolated and identified from soybeans[5,6]. Additionally, recent research showed that pinto and black beans as excellent dietary sources of natural antioxidants and thus may serve as disease preventative and healthy foods[7]. Moreover, a recent review analyzing epidemiological studies in the last ten years and the flavonoid content of diets concluded that dietary flavonoids may prevent a range of different types of cancer[8]. In fact, free phenolic compounds, such as flavonols and isoflavones, have been found to have with immunomodulatory and cellular antiproliferative functions[9].

The sprouting process induces a variety of biochemical changes in soybean seeds, leading to the accumulation of primary and secondary metabolites[10]. Some studies have found that germination and sprout formation of legumes are accompanied by a significant increase in total flavonoids content with antioxidant potential[11–13]. Generally, content of genistein derivatives is about two times higher than that of daidzein derivatives in raw soybeans[14–16]. Germination can induce isoflavone profile changes in soybeans including decreases of β-glycosides

[1]Heilongjiang Academy of Agricultural Sciences Post-Doctoral Station, Nangang District, Harbin 150086, People's Republic of China. [2]Soybean Research Institute, Heilongjiang Academy of Agricultural Sciences, Nangang District, Harbin 150086, People's Republic of China. [3]Colleage of Life Sciences, Yulin University, Yuyang District, Yulin 719000, People's Republic of China. [4]School of Hydraulic and Electric Power, Heilongjiang University, Nangang District, Harbin 150086, People's Republic of China. [5]Food Processing Research Institute, Heilongjiang Academy of Agricultural Sciences, Nangang District, Harbin 150086, People's Republic of China. [6]These authors contributed equally: Weiwei Bi and Guixing Zhao ✉email: weiwebi5321@163.com

and increases of both 6′-O-malonyl-β-glucosides and aglycones[17–19]. During germination, a 13% increase of total isoflavone content was found by Kim, et al[20]. One comparative study of the chemical profiles of three Egyptian cultivars of fava beans; Sakha 3, Nubaria 3, and Giza 843, focused on the effects of germination on the chemical profiles of phenolic compounds and saponins and characterized 65 metabolites based on UV spectra, accurate MS, and MS/MS data. Germination was found to dramatically increase the quantities of flavonoids and saponins to which biological activities are attributed[20]. A previous study has suggested that germination could enhance the content of isoflavonesin soybeans[21]. Accordingly, studies based on metabolomics have applied combined techniques, such as chromatographic separation (UHPLC and/or gas chromatography) and mass spectrometry, for comprehensive metabolic analysis of plant substrates[22,23]. Recent technological developments in the field of metabolomics, especially as a result of the widespread use of LC-high resolution MS have considerably improved accuracy and sensitivity of metabolite detection[24]. For example, non-target LC–MS-Orbitrap metabolomics methods were used to compare and analyze the biologically active metabolites of frozen, boiled, and canned chickpeas, lentils, and white beans[25]. The main purposes of this study were to: (1) compare the flavonoid contents in soybean seeds and sprouts grown under dark conditions for 3 days at 24 °C (2) analyze the metabolism of flavonoids in soybean sprouts.

Materials and methods

Samples preparation. According to our previous detection of isoflavones in 12 main soybean varieties grown in Heilongjiang, China, two varieties with high isoflavone content were selected for experimentation[26]. The two soybean cultivars Heinong52 (HN52), and Heinong71(HN71) used in this experiment were cultivated at the Harbin Soybean Research Institute (Modern Agriculture Demonstration Zone of Heilongjiang Academy of Agricultural Sciences, Minzhu Township, Daowai District, Harbin; 126.65°E, 45.78°N), Heilongjiang Academy of Agricultural Sciences, Harbin, China(statement as follow). Soybean seeds were harvested from three replicates of each cultivar for a cropping year and stored at room temperature until analysis of flavonoids compounds. Briefly, high-quality soybean grains were selected, rinsed with clean water, soaked in distilled water at 24 °C for 4 h, spread in a single layer on a tray, covered with gauze, and placed in an artificial climate box (HPG-1600H, Harbin Donglian Electronic Technology Development Co., Ltd. Harbin, China) at about 24 °C and 80% humidity to germinate for 3 days in the dark. The growth chamber was set up to spray water for 3 min every 1 h. After germination, approximately 20 g FW(fresh weight) of soybean sprouts (HNS52,HNS71) per cultivar with three replicates every sampling day were randomly collected from containers, and freeze-dried at − 50 °C for 20 h. The freeze-dried sprouts and seeds were crushed using a mixer mill (MM 400, Retsch) with a zirconia bead for 1.5 min at 30 Hz. 100 mg powder was weighed and extracted overnight at 4 °C with 1.0 mL 70% aqueous methanol. Following centrifugation at 10,000 g for 10 min, the extracts were absorbed (CNWBOND Carbon-GCB SPE Cartridge, 250 mg, 3 mL; ANPEL, Shanghai, China, www.anpel.com.cn/cnw) and filtrated (SCAA-104,0.22 μm pore size; ANPEL, Shanghai, China, http://www.anpel.com.cn/) LC–MS analysis(contained "2.2 HPLC Conditions" and "2.3 ESI-Q TRAP-MS/MS") .

HPLC conditions. The sample extracts were analyzed using a LC–ESI–MS/MS system (HPLC, Shim-pack UFLC SHIMADZU CBM30A system, www.shimadzu.com.cn/; MS, Applied Biosystems 4500 Q TRAP, www.appliedbiosystems.com.cn/). We used the following steps to analyze, HPLC: column, Waters ACQUITY UPLC HSS T3 C18 (1.8 μm, 2.1 mm*100 mm); solvent system, water (0.04% acetic acid): acetonitrile (0.04% acetic acid); gradient program,100:0 V/V at 0 min, 5:95 V/V at 11.0 min, 5:95 V/V at 12.0 min, 95:5 V/V at 12.1 min, 95:5 V/V at 15.0 min; flow rate, 0.40 mL/min; temperature, 40 °C; injection volume: 5μL. The effluent was alternatively connected to an ESI-triple quadrupole-linear ion trap (QTRAP)-MS[27].

LIT and triple quadrupole (QQQ) scans were acquired using a triple quadrupole-linear ion trap mass spectrometer (Q TRAP), API 4500 Q TRAP LC/MS/MS System, equipped with an ESI Turbo Ion-Spray interface, operating in a positive ion mode and controlled by Analyst 1.6.3 software (AB Sciex). The parameters of ESI source operation were as follows: ion source, turbo spray; source temperature 550 °C; ion spray voltage (IS)5500 V; ion source gas I (GSI), gas II(GSII), curtain gas (CUR) were set at 55, 60, and 25.0 psi, respectively; the collision gas(CAD) was high. Instrument tuning and mass calibration were performed with 10 and 100 μmol/L polypropylene glycol solutions in QQQ and LIT modes, respectively. QQQ scans were acquired as MRM experiments with collision gas (nitrogen) set to 5 psi. DP and CE for individual MRM transitions was done with further DP and CE optimization. A specific set of MRM transitions were monitored for each period according to the metabolites eluted within this period[28–30]. Similarly, Valente et al. calculated the metabolic profile and antioxidant activity of 7 European broad bean seed varieties[31] and their pods (by-products) using HPLC–DAD-MS/MS-based methods and found the existence of 105 phenolic compounds, alkaloids, jasmonic acids, and organic acids[32].

Statistics analysis. For all experiments, three independent assays were carried out. One-way analysis of variance was used to study the differences between means, with a significant level at $P < 0.05$. Data analysis was carried out with SPSS (Windows version 12.0, SPSS Inc., Chicago,IL, USA). All data are presented as mean ± standard error of means.

Results and discussion

Principal component analysis. Principal component analysis (PCA) was performed on metabolic data to visualize the biological variability of soybean seeds and sprouts. These results indicated that the dominating source of variance was the differential flavonoid content in the two soybean varieties across all germination stages (PC1). From Fig. 1A, the PC1 of 4 groups and mix was 58.6%. The PC1 of two soybean seeds was 77.25%. The PC1 of two soybean seeds and sprouts were74.46% and 81.16% respectively. The PC1 of all samples was

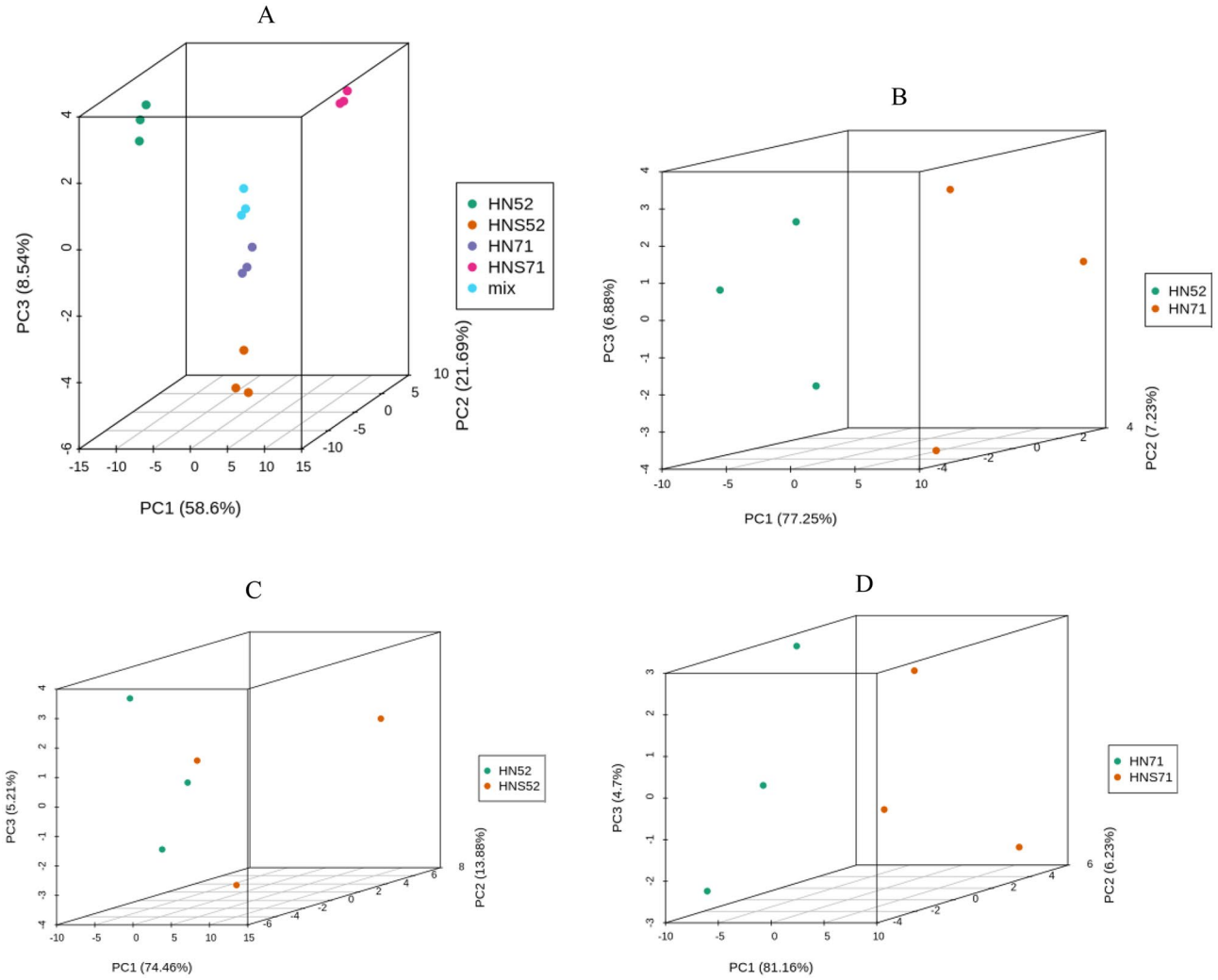

Figure 1. PCA score chart of mass spectrum data of each group of samples and quality control samples. Mix: quality control samples, (**A**): all samples of PCA. (**B**): Two soybean seeds of PCA; (**C**):Soybean and sprouts of PCA(HN52 and HNS52); (**D**): Soybean and sprouts of PCA(HN71 and HNS71).

greater than 50%. These results indicated that the grouping of samples is reasonable. Through PCA of samples (including quality control samples MIX), in order to have a preliminary understanding of the overall metabolic differences between samples in each group and the degree of variability between samples within the group. The PCA results show the trend of metabolome separation between groups, suggesting whether there is a difference in metabolome between sample groups. The X axis represents the first principal component, and the Y axis represents the second principal component. Grouping principal component analysis: before doing difference analysis, first perform principal component analysis on the grouped samples for difference comparison to observe the degree of variability between the difference groups and the samples within the group.

Differential metabolic analysis of flavonoids. From Table 1, a total of 114 flavonoid metabolites were detected in the two soybean varieties and their sprouts. These consisted of 41 isoflavones, 26 flavonoids, 14 flavonols, 10chalcones, 6 dihydroflavones, 5 tannins, 4 dihydroflavonols, 3 carbonosides, 2 flavonoids, 1 flavanol, 1 proanthocyanidin, and 1anthocyanin. These results indicate that the flavonoids in soybeans and sprouts are mainly composed of isoflavones (35.96%). 13 substances detected in this study were the same as those detected in Egyptian cultivars of chickpea in a previous study[30]. These were $C_{13}H_{16}O_{10}$ (6-O-Galloyl-glucose), $C_{13}H_{16}O_{10}$(3-O-Galloyl-glucose),$C_{13}H_{16}O_{10}$ (2-O-Galloyl-glucose), $C_{27}H_{30}O_{16}$(Kaempferol-3-O-neohesperidoside), $C_{27}H_{30}O_{16}$ (5,7-Dihydroxy-4-methoxyflavone-3-O-xylose-(1–6)-glucose), $C_{21}H_{22}O_{11}$ (Aromadendrin-7-O-glucoside), $C_{27}H_{30}O_{16}$ (Quercetin-3-O-robinobioside), $C_{27}H_{30}O_{17}$ (Quercetin-3-O-sophoroside (Baimaside)), $C_{33}H_{40}O_{21}$ (Kaempferol-3-O-sophoroside -7-O- glucoside) , $C_{27}H_{32}O_{15}$ (Eriodictyol-7-O-Rutinoside (Eriocitrin)), $C_{27}H_{30}O_{16}$ (Quercetin-3-O-rutinoside (Rutin)), $C_{28}H_{32}O_{17}$ (2'-Hydoxy, 5-methoxy Genistein-4',7- O-diglucoside), $C_{27}H_{30}O_{15}$ (Eriodictyol-7-O-Rutinoside (Eriocitrin)), respectively. This indicated that these flavonoids may be universal in beans.

Index	Formula	Compounds	Class	CAS
pme0309	$C_8H_8O_5$	3-O-Methylgallic Acid	Tannin	3934-84-7
pme0355	$C_{15}H_{10}O_4$	Daidzein	Isoflavones	486-66-8
mws0902	$C_{15}H_{12}O_4$	Liquiritigenin	Dihydroflavone	578-86-9
pme3217	$C_{15}H_{12}O_4$	Isoliquiritigenin	Chalcones	961-29-5
mws0037	$C_{16}H_{12}O_4$	Formononetin (7-Hydroxy-4'-methoxyisoflavone)	Isoflavones	485-72-3
mws0063	$C_{15}H_{10}O_5$	Genistein*	Isoflavones	446-72-0
pme3261	$C_{15}H_{10}O_5$	6-Hydroxydaidzein*	Isoflavones	17,817-31-1
mws0912	$C_{15}H_{10}O_5$	2'-Hydroxydaidzein*	Isoflavones	7678-85-5
pmp000344	$C_{15}H_{10}O_5$	3',4',7-Trihydroxyflavone*	Flavonoid	2150-11-0
Lmmp005125	$C_{16}H_{14}O_4$	Pinostrobin Chalcone	Chalcones	18956-15-5
mws4060	$C_{16}H_{14}O_4$	Echinatin	Chalcones	34221-41-5
pme1397	$C_{15}H_{11}O_5$ +	Pelargonidin	Anthocyanins	7690-51-9
mws0914	$C_{15}H_{12}O_5$	Pinobanksin*	Dihydroflavonol	548-82-3
pme0376	$C_{15}H_{12}O_5$	Naringenin (5,7,4'-Trihydroxyflavanone)*	Dihydroflavone	480-41-1
pme3250	$C_{16}H_{12}O_5$	Biochanin A	Isoflavones	491-80-5
pme3233	$C_{16}H_{12}O_5$	Calycosin*	Isoflavones	20575-57-9
mws0908	$C_{16}H_{12}O_5$	Glycitein*	Isoflavones	40957-83-3
mws0062	$C_{15}H_{10}O_6$	Isoluteolin (Orobol)(5,7,3',4'-tetrahydroxyisoflavone)	Isoflavones	480-23-9
pme0088	$C_{15}H_{10}O_6$	Luteolin (5,7,3',4'-Tetrahydroxyflavone)	Flavonoid	491-70-3
mws1094	$C_{15}H_{12}O_6$	Dihydrokaempferol	Dihydroflavonol	480-20-6
Lmdp005525	$C_{17}H_{14}O_5$	Afrormosin	Isoflavones	550-79-8
Lmqn003210	$C_{17}H_{14}O_5$	5-Hydroxy-6,7-dimethoxyflavone	Flavonoid	740-33-0
Zmhp003514	$C_{16}H_{12}O_6$	6,7,8-Tetrahydroxy-5-methoxyflavone*	Flavonoid	–
mws0058	$C_{16}H_{12}O_6$	Diosmetin (5,7,3'-Trihydroxy-4'-methoxyflavone)	Flavonoid	520-34-3
pmp000001	$C_{16}H_{12}O_6$	Hispidulin (5,7,4'-Trihydroxy-6-methoxyflavone)*	Flavonoid	1447-88-7
pmp000804	$C_{19}H_{20}O_4$	Isobavachalcone D	Chalcones	–
pmp000348	$C_{20}H_{18}O_4$	Kanzonol D	Other Flavonoids	155,233-20-8
Lmfn000604	$C_{13}H_{16}O_{10}$	6-O-Galloyl-glucose*	Tannin	13,186-19-1
Hmln000659	$C_{13}H_{16}O_{10}$	3-O-Galloyl-glucose*	Tannin	–
Hmln000873	$C_{13}H_{16}O_{10}$	2-O-Galloyl-glucose	Tannin	–
pmp000350	$C_{20}H_{16}O_5$	Glabrone	Isoflavones	60008-02-8
pmp001223	$C_{20}H_{18}O_5$	Psoralenol	Isoflavones	70522-30-4
pmp000352	$C_{20}H_{18}O_5$	Licoflavone C	Flavonoid	72357-31-4
pmp000351	$C_{20}H_{18}O_5$	Eurycarpin A	Isoflavones	166547-20-2
pmp000637	$C_{20}H_{20}O_5$	Sophoraflavanone B	Dihydroflavone	53846-50-7
Xmgp006913	$C_{20}H_{20}O_5$	2,4,2',4'-tetrahydroxy-3'-prenylchalcone	Chalcones	–
pmn001389	$C_{21}H_{20}O_5$	Gancaonin G	Isoflavones	126716-34-5
pmp000355	$C_{20}H_{18}O_6$	Licoflavonol	Flavonols	60197-60-6
pmp000811	$C_{20}H_{18}O_6$	Morachalcone C	Chalcones	1304549-24-3
pmp000360	$C_{21}H_{22}O_5$	3-Hydroxylicochalcone A	Chalcones	–
pmp000361	$C_{21}H_{22}O_5$	Licochalcone D	Chalcones	144506-15-0
pmp001224	$C_{20}H_{20}O_6$	Brosimacutin G	Chalcones	350,221-50-0
Cmsn000894	$C_{14}H_{18}O_{11}$	7-O-Galloyl-D-sedoheptulose	Tannin	–
mws0055	$C_{20}H_{20}O_7$	Tangeretin	Flavonols	481-53-8
pmp000417	$C_{21}H_{20}O_9$	Daidzein-4'-O-glucoside	Isoflavones	–
pme1587	$C_{21}H_{20}O_9$	Daidzein-7-O-glucoside(Daidzin)	Isoflavones	552-66-9
mws1597	$C_{21}H_{20}O_9$	Puerarin	Isoflavones	3681-99-0
pmp000384	$C_{21}H_{22}O_9$	Isoliquiritin*	Chalcones	5041-81-6
pmp000383	$C_{21}H_{22}O_9$	Liquiritigenin-4'-O-Glucoside (Liquiritin)*	Dihydroflavone	551-15-5
pmp000647	$C_{25}H_{28}O_6$	Kushenol E	Other Flavonoids	99,119-72-9
pme3504	$C_{22}H_{22}O_9$	Formononetin-7-O-glycoside (Ononin)	Isoflavones	486-62-4
HJN089	$C_{21}H_{20}O_{10}$	Sophoricoside	Isoflavones	152-95-4
pmp000413	$C_{21}H_{20}O_{10}$	Genistein-8-C-glucoside*	Isoflavones	66,026-80-0
mws1434	$C_{21}H_{20}O_{10}$	Apigenin-6-C-glucoside (Isovitexin)*	Flavonoid carbonoside	29,702-25-8
Lmlp005572	$C_{21}H_{20}O_{10}$	Galangin-7-O-glucoside*	Flavonoid	–
mws0072	$C_{21}H_{20}O_{10}$	Apigenin-5-O-glucoside*	Flavonoid	28,757-27-9
Continued				

Index	Formula	Compounds	Class	CAS
mws0895	$C_{21}H_{20}O_{10}$	Genistein-7-O-Glucoside (Genistin)*	Isoflavones	529-59-9
pme1611	$C_{21}H_{22}O_{10}$	Isohemiphloin*	Flavonoid carbonoside	3682-02-8
mws1179	$C_{21}H_{22}O_{10}$	Naringenin-7-O-glucoside (Prunin)*	Dihydroflavone	529-55-5
HJN090	$C_{21}H_{22}O_{10}$	Butin-7-O-glucoside*	Flavonoid	–
pmp000550	$C_{22}H_{22}O_{10}$	Calycosin-7-O-glucoside	Isoflavones	20,633-67-4
HJN091	$C_{22}H_{20}O_{10}$	Prunetin-4'-O-glucoside	Isoflavones	154-36-9
pme3400	$C_{22}H_{22}O_{10}$	Biochanin A-7-O-glucoside (Sissotrin)	Isoflavones	5928-26-7
pmp000415	$C_{22}H_{22}O_{10}$	3'-Methoxydaidzin	Isoflavones	200,127-80-6
mws0894	$C_{22}H_{22}O_{10}$	Glycitin	Isoflavones	40,246-10-4
mws1172	$C_{22}H_{22}O_{10}$	Trifolirhizin (Maackiain-3-O-glucoside)	Isoflavones	6807-83-6
Lmlp003531	$C_{21}H_{20}O_{11}$	Luteolin-3'-O-glucoside*	Flavonoid	5154-41-6
pme2459	$C_{21}H_{20}O_{11}$	Luteolin-7-O-glucoside (Cynaroside)*	Flavonoid	5373-11-5
Xmyp005654	$C_{21}H_{20}O_{11}$	Kaempferol-4'-O-glucoside*	Flavonoid	–
mws1361	$C_{21}H_{22}O_{11}$	Astilbin	Dihydroflavonol	29838-67-3
Lmtn002796	$C_{21}H_{22}O_{11}$	Aromadendrin-7-O-glucoside	Flavonoid	28189-90-4
Lmlp005236	$C_{21}H_{22}O_{11}$	Dihydrokaempferol-3-O-glucoside	Dihydroflavonol	1049-08-8
Zmdp004370	$C_{23}H_{22}O_{10}$	6"-O-Acetyldaidzin	Isoflavones	71385-83-6
Lmdp003994	$C_{23}H_{24}O_{10}$	6,4'-Dimethoxyisoflavone-7-O-glucoside (Wistin)	Isoflavones	19046-26-5
mws0091	$C_{21}H_{20}O_{12}$	Quercetin-3-O-glucoside (Isoquercitrin)	Flavonols	482-35-9
mws0061	$C_{21}H_{20}O_{12}$	Quercetin-3-O-galactoside (Hyperin)	Flavonols	482-36-0
pmp000191	$C_{23}H_{22}O_{11}$	6"-O-Acetylgenistin*	Isoflavones	73566-30-0
pmp000581	$C_{23}H_{22}O_{11}$	Apigenin-7-O-(6"-acetyl)glucoside*	Flavonoid	72741-92-5
pmp000192	$C_{24}H_{24}O_{11}$	Acetylglycitin	Isoflavones	134859-96-4
Lmmp003903	$C_{23}H_{22}O_{12}$	Kaempferol-3-O-(2"-acetyl)glucoside	Flavonols	–
Lmmn003398	$C_{23}H_{22}O_{12}$	Kaempferol-3-O-(6"-acetyl)glucoside	Flavonols	–
Lmgp004829	$C_{23}H_{24}O_{12}$	5,7,4'-Trihydroxy-6,8-dimethoxyisoflavone-7-O-galactosides	Isoflavones	–
pmp000193	$C_{24}H_{22}O_{12}$	6"-O-Malonyldaidzin	Isoflavones	124590-31-4
Zmdp004305	C28H24O9	Daidzein-7-O-(2"-benzoyl)rhamnoside	Flavonoid	-
Zmdp005767	$C_{25}H_{24}O_{12}$	Formononetin-7-O-(6"-Malonyl)glucoside	Isoflavones	-
Zmdp004112	$C_{24}H_{22}O_{13}$	Genistein-7-O-(6"-malonyl)glucoside	Isoflavones	-
pmp000194	$C_{24}H_{22}O_{13}$	6"-O-Malonylgenistin	Isoflavones	51011-05-3
pmp000195	$C_{25}H_{24}O_{13}$	6"-O-Malonylglycitin	Isoflavones	137705-39-6
Lmmp003817	$C_{24}H_{22}O_{14}$	Kaempferol-3-O-(6"-malonyl)glucoside*	Flavonols	–
pmp000587	$C_{24}H_{22}O_{14}$	Luteolin-7-O-(6"-malonyl)glucoside*	Flavonoid	–
Lmdp004892	$C_{24}H_{22}O_{14}$	Kaempferol-3-O-(6"-malonyl)galactoside*	Flavonols	–
pmb0608	$C_{25}H_{24}O_{14}$	Chrysoeriol-7-O-(6"-malonyl)glucoside	Flavonoid	–
Zmdp003677	$C_{26}H_{28}O_{13}$	Daidzein-7-O-Glucoside-4'-O-Apioside	Flavonoid	108069-01-8
Wmkn002777	$C_{27}H_{30}O_{13}$	7-Hydroxy-3"-methoxy-isoflavone-7-primeveroside	Isoflavones	–
Lmnp202580	$C_{26}H_{28}O_{14}$	Apigenin-8-C-(2"-xylosyl)glucoside	Flavonoid carbonoside	–
pmp000411	$C_{26}H_{28}O_{14}$	Apigenin-7-O-(2"-glucosyl)arabinoside	Flavonoid	–
mws0836	$C_{30}H_{26}O_{12}$	Procyanidin B1	Proanthocyanidins	20315-25-7
pmp000414	$C_{27}H_{30}O_{14}$	Puerarin-4'-O-glucoside	Isoflavones	117047-08-2
pmp001079	$C_{27}H_{30}O_{15}$	Luteolin-7-O-neohesperidoside (Lonicerin)	Flavonoid	25694-72-8
mws1073	$C_{27}H_{30}O_{15}$	Apigenin-6,8-di-C-glucoside	Flavonoid	23666-13-9
Lmjp002867	$C_{27}H_{30}O_{15}$	Kaempferol-3-O-neohesperidoside	Flavonols	32602-81-6
Wmkp002590	$C_{27}H_{30}O_{15}$	5,7-Dihydroxy-4-methoxyflavone-3-O-xylose-(1-6)-glucose	Flavonoid	–
mws1519	$C_{27}H_{32}O_{15}$	Eriodictyol-7-O-Rutinoside (Eriocitrin)	Dihydroflavone	13463-28-0
Lmmp000897	$C_{30}H_{26}O_{14}$	Gallocatechin-Gallocatechin	Flavanols	–
pmn001583	$C_{27}H_{30}O_{16}$	Quercetin-3-O-robinobioside*	Flavonols	52525-35-6
Zmxp003107	C27H30O16	Luteolin-7,3'-di-O-glucoside*	Flavonoid	52187-80-1
mws0059	$C_{27}H_{30}O_{16}$	Quercetin-3-O-rutinoside (Rutin)*	Flavonols	153-18-4
Lmhp003217	$C_{28}H_{32}O_{16}$	2'-Hydoxy,5-methoxyGenistein-O-rhamnosyl-glucoside	Isoflavones	–
pme1540	$C_{28}H_{32}O_{16}$	Isorhamnetin-3-O-neohesperidoside	Flavonols	55033-90-4
Lmtp003677	$C_{27}H_{30}O_{17}$	Quercetin-3-O-sophoroside (Baimaside)	Flavonols	18609-17-1
Lmhp002800	$C_{28}H_{32}O_{17}$	2'-Hydoxy,5-methoxyGenistein-4',7-O-diglucoside	Isoflavones	–
Hmqp003435	$C_{29}H_{30}O_{17}$	Apigenin-7-O-(2"-O-apiosyl)(6"-Malonyl)glucoside	Flavonoid	–
Continued				

Index	Formula	Compounds	Class	CAS
Hmlp003185	$C_{30}H_{32}O_{18}$	Luteolin-7-O-(6”-malonyl)glucoside-5-O-rhamnoside	Flavonoid	–
Lmwp003888	$C_{33}H_{40}O_{21}$	Kaempferol-3-O-sophoroside-7-O-glucoside	Flavonols	–

Table 1. Flavonoids in soybeans seeds and sprouts.

Metabolomic data is massive and multidimensional, so it is imperative to combine univariate and multivariate statistical analyses, to analyze the data from a multivariate perspective according to the characteristics of the data, and to accuratelymine differential metabolites. This study has three biological replicates. The method of combining fold change ≥ 2 and OPLS-DA model VIP (The VIP value represents the influence of the difference between the corresponding metabolites in the classification and discrimination of each group of samples in the model. It is generally believed that the metabolites with VIP ≥ 1 are significantly different.) ≥ 1 was used to screen for differential metabolites. Shown in Fig. 2A, there were a total of 18 significantly different metabolites between the two soybean varieties, of which 14 were up-regulated and 4 were down-regulated. After germination, the significantly different metabolites in the two soybean sprouts were 33, of which 19 were up-regulated and 14 were down-regulated (Fig. 2B). Moreover, there were a total of 27 significantly different metabolites between HN52 and HNS52 cultivars, of which 26 were up-regulated and 1 were down-regulated (Fig. 2C). Additionally, there were a total of 25 significantly different metabolites between HN71 and HNS71, of which 21 were up-regulated and 4 were down-regulated (Fig. 2D). These experimental results showed that both differential and up-regulated metabolites increased in soybean sprouts compared to soybean seeds, thus indicating that soybean germination can increase content of flavonoid metabolites. From a qualitative comparison, that the germination process clearly has an impact on the number of characterized metabolites, especially flavonoid compounds. In fact, the biological activity of the plant substrate is mainly due to these metabolites[33–37].

Each point in the volcano map represents a metabolite. The abscissa represents the logarithm of the quantitative difference of a certain metabolite in the two samples; the ordinate represents the VIP value. The larger the absolute value of the abscissa, the greater the multiple difference in the expression amount between the two samples; the larger the ordinate value, the more significant the differential expression, and the more reliable the differentially expressed metabolites screened. The green dots in the figure represent down-regulated differentially expressed metabolites, the red dots represent up-regulated differentially expressed metabolites, and black represents detected but not significantly different metabolites.

As shown in Fig. 3, the flavonoids of the two varieties have a balanced distribution of the differential metabolite multiples, and the differential metabolites multiples decrease after germination. This indicated that the flavonoid metabolism of different varieties of soybeans while germinating was highly similar. The flavonoid metabolites in the same variety of soybean seeds and sprouts differ greatly in multiples, showing that flavonoid metabolism is especially active during soybean germination, and that germination is a way to accumulate flavonoids. During the growth process, small sprouts may be damaged by microbial and environmental pressures or nutrient deficiency. Therefore it is necessary for plants to synthesize secondary metabolites (such as flavonoid compounds) through different metabolic pathways to develop defense mechanisms[38–40].

As shown in Fig. 4, there were total 33 differential metabolites, of which 12 identical metabolites in HN52 vs HNS52 and HN52 vs HN71. There were total 45 differential metabolites, of which 13 metabolites were shared in HN71 vs HNS71 and HNS52 vs HNS71. These results indicated that flavonoid metabolites increased in germinated soybeans. For example, Wu et al. discussed the effect of germination on the antioxidant potential and isoflavonoid content of chickpeas and observed increases in both antioxidant activity and isoflavonoids content with germination[41]. Additionally, Ayet et al. revealed that germination increased the content of soya saponins in lentils[42,43]. Moreover, Mekky et al. revealed that germination increased the quantities of flavonoids, phenolic acids and saponins in fava beans[30].

Functional annotation and enrichment analysis of differential metabolites. In the homepage of the KEGG website, we directly enter the keyword "Flavonoids" and the top 4 pathways are all pathways involving flavonoid metabolites, including:flavonoid synthesis, anthocyanin synthesis, isoflavone synthesis, flavonoids and flavonols synthesis. Using the KEGG database[44] to annotate and display the differential metabolites identified in these samples, 6 metabolite were significantly up-regulated in the isoflavone synthesis pathway, 1 metabolite was significantly down-regulated, and 13 metabolites did not change significantly (Fig. 5).

There are six primary metabolic pathways for differential metabolism of flavonoids; isoflavonoid biosynthesis, flavonoid biosynthesis, flavone and flavonol biosynthesis, secondary metabolite biosynthesis, and phenylpropanoid biosynthesis (Fig. 6). Unlike soybean seeds, soybean sprouts have flavone and flavonol biosynthesis. Isoflavonoid biosynthesis accounts for most of the pathways, and the proportions in different samples are 66.67% in HN52 vs HN71 (Fig. 6A), 41.67% in HNS52 vs HNS71 (Fig. 6B), 54.55% in HN52 vs HNS52 (Fig. 6C), and 40% in HN71 vs HNS71 (Fig. 6D), respectively. Another important pathway is the biosynthesis of secondary metabolites, and the proportions in different samples are 33.33% in HN52 vs HN71 (Fig. 6A, 16.67% in HNS52 vs HNS71 (Fig. 6B), 54.55% in HN52 vs HNS52 (Fig. 6C), 70% in HN71 vs HNS71 (Fig. 6D), respectively). Metabolic pathway data explained why isoflavones were higher in different metabolites.

As shown in Fig. 7, that the larger size dot indicated isoflavonoid biosynthesis and biosynthesis of secondary metabolites representing the number of these metabolites enriched. The red point represented isoflavonoid

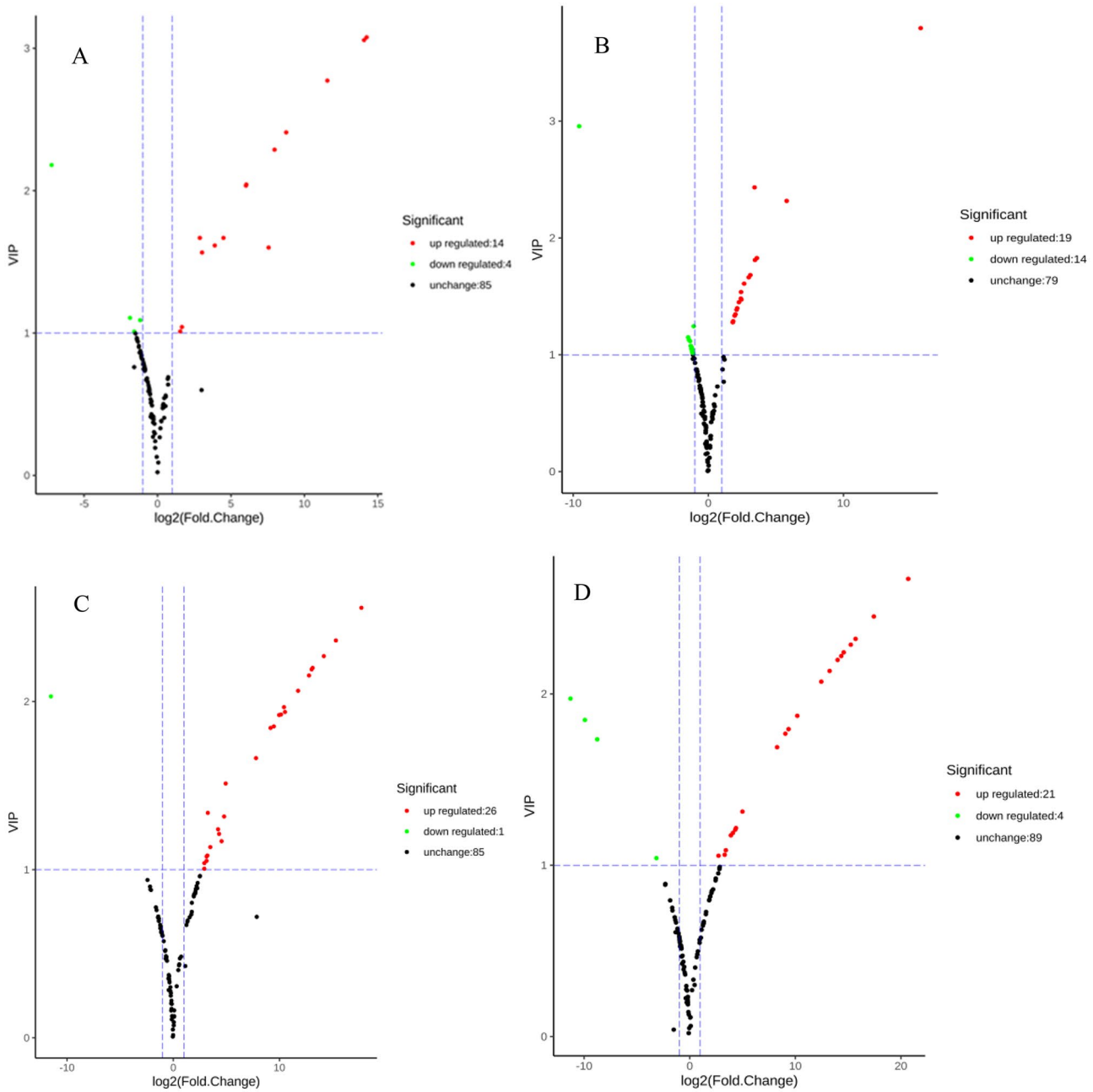

Figure 2. Volcano Map of Differential Metabolites. (**A**): HN52 vs HN71; (**B**): HNS52 vs HNS71; (**C**): HN52 vs HNS52; (**D**): HN71 vs HNS71.

biosynthesis in Fig. 7C, indicating that isoflavonoid was more significantly enriched. The redder points indicating isoflavonoid biosynthesis in Fig. 7A, showed that isoflavonoids were more significantly enriched.

Conclusion

In this study, two soybean varieties with high isoflavone content and their soybean sprouts in Heilongjiang Province, China were analyzed with flavonoid metabolome analysis. 114 differential metabolites were detected which contained 41 isoflavones and 26 flavonoids. These experimental results revealed that the differentially expressed and up-regulated metabolites in soybean sprouts increased compared to soybean seeds. The flavonoid metabolites in the same variety of soybean seeds and sprouts differ greatly in multiples, indicating that flavonoid metabolism is very active during soybean germination, and germination may be a way to accumulate flavonoids. Metabolic pathway data explained why isoflavones were higher than different metabolites. This study detected and analyzed the changes of flavonoid metabolites after soybean germination, providing data reference for the research of flavonoids, and providing ideas for future research on plant secondary metabolism.

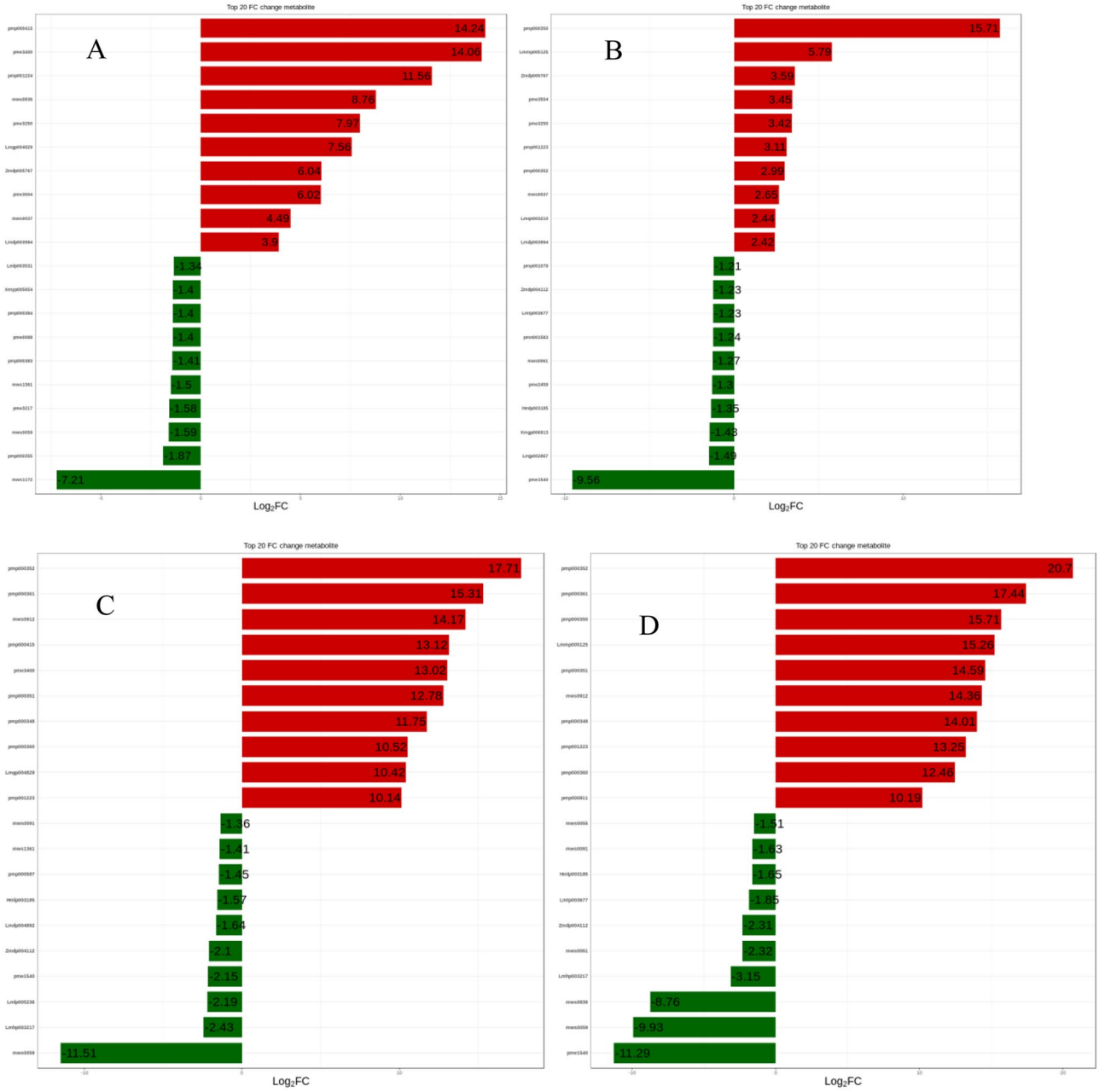

Figure 3. Analysis of the fold difference of metabolites between different samples. (**A**): HN52 vs HN71; (**B**): HNS52 vs HNS71; (**C**): HN52 vs HNS52; (**D**): HN71 vs HNS71.

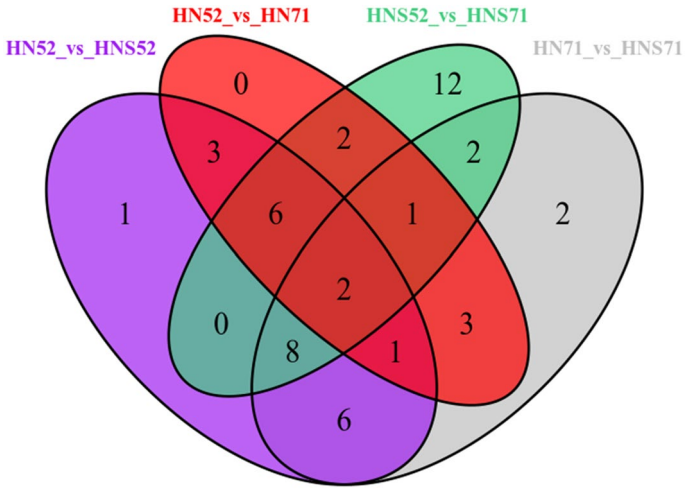

Figure 4. Venn diagram showing the relationships between different groups of metabolites.

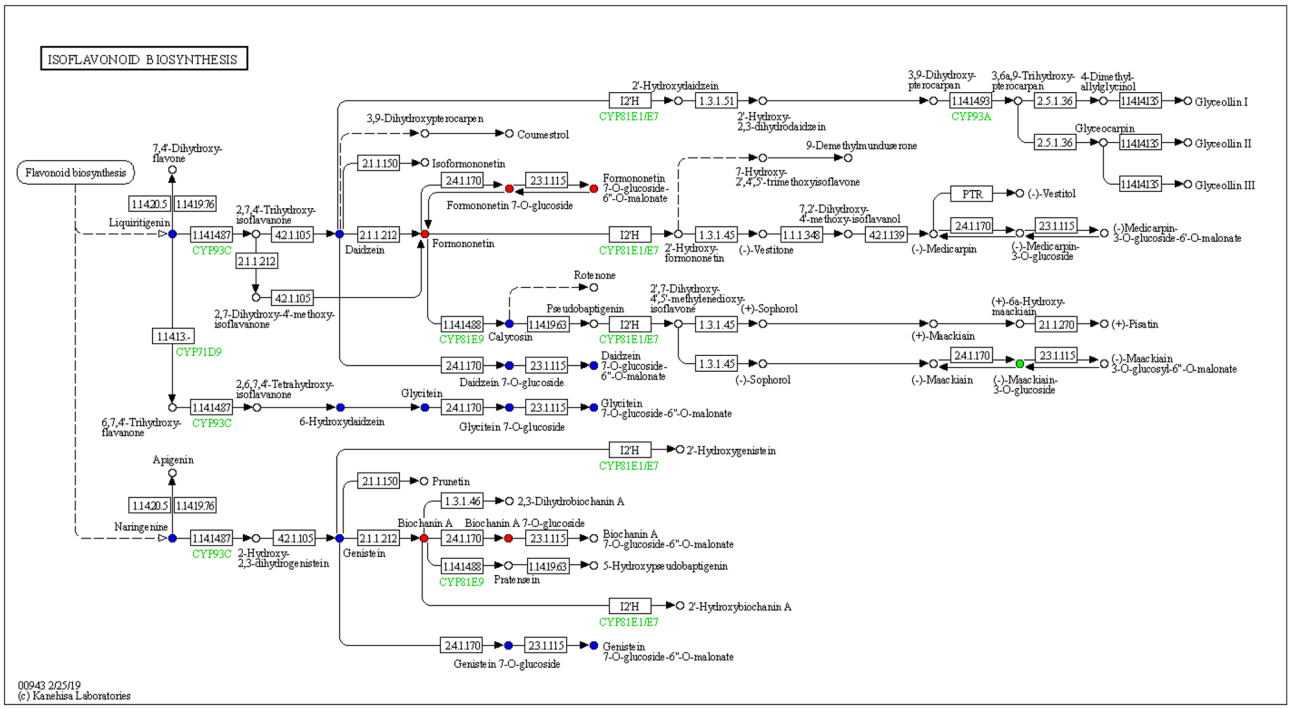

Figure 5. Differential metabolite KEGG pathway diagram. Note: Red indicates that the metabolite content is significantly up-regulated in the experimental group, and blue indicates that the metabolite is detected but not significantly changed. Green indicates that the metabolite content is significantly down-regulated in the experimental group.

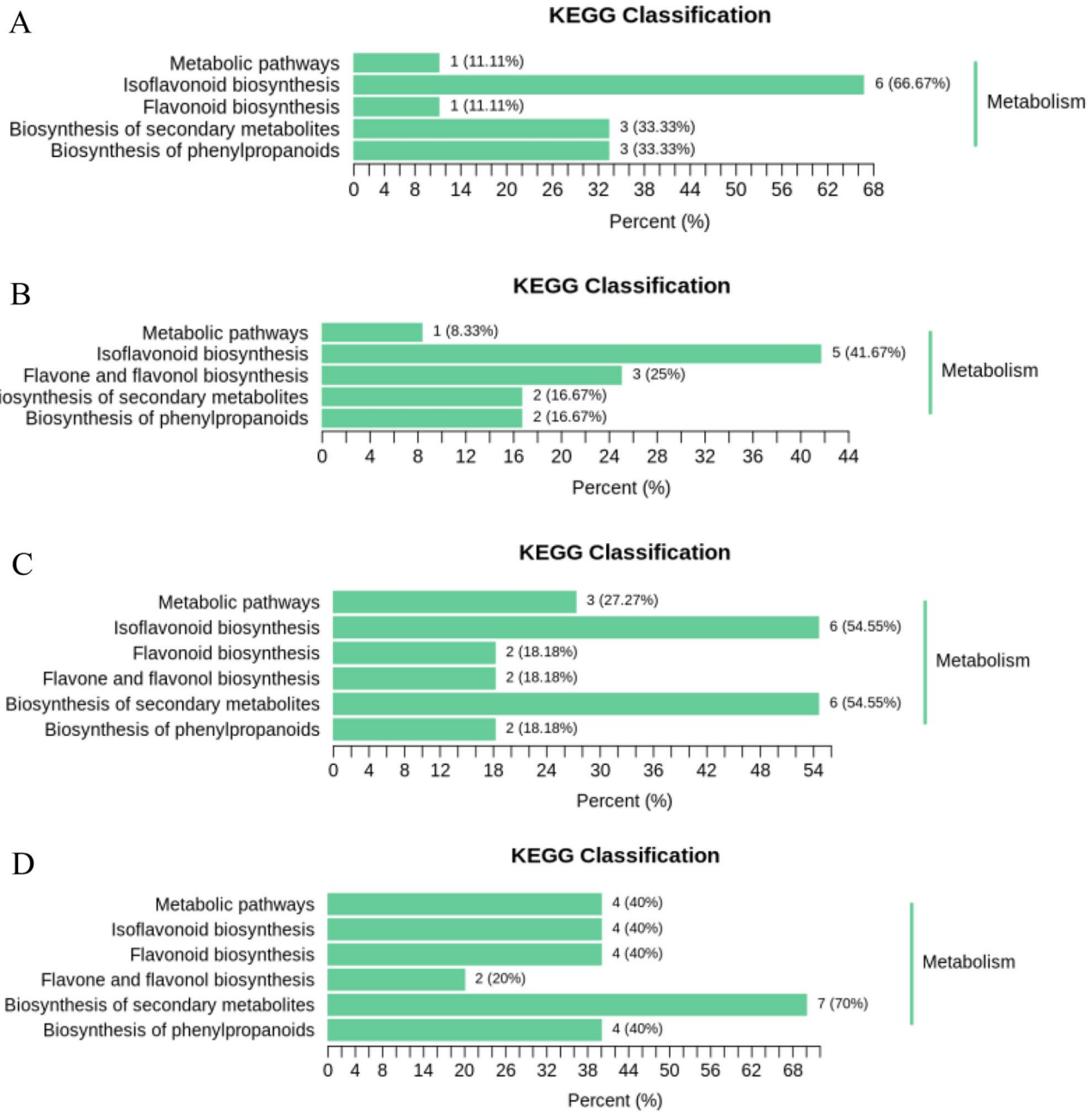

Figure 6. Differential metabolite KEGG classification chart. (**A**): HN52 vs HN71; (**B**): HNS52 vs HNS71; (**C**): HN52 vs HNS52; (**D**): HN71 vs HNS71. Note: The ordinate is the name of the KEGG metabolic pathway, and the abscissa is the number of metabolites annotated to the pathway and its proportion to the total number of metabolites annotated.

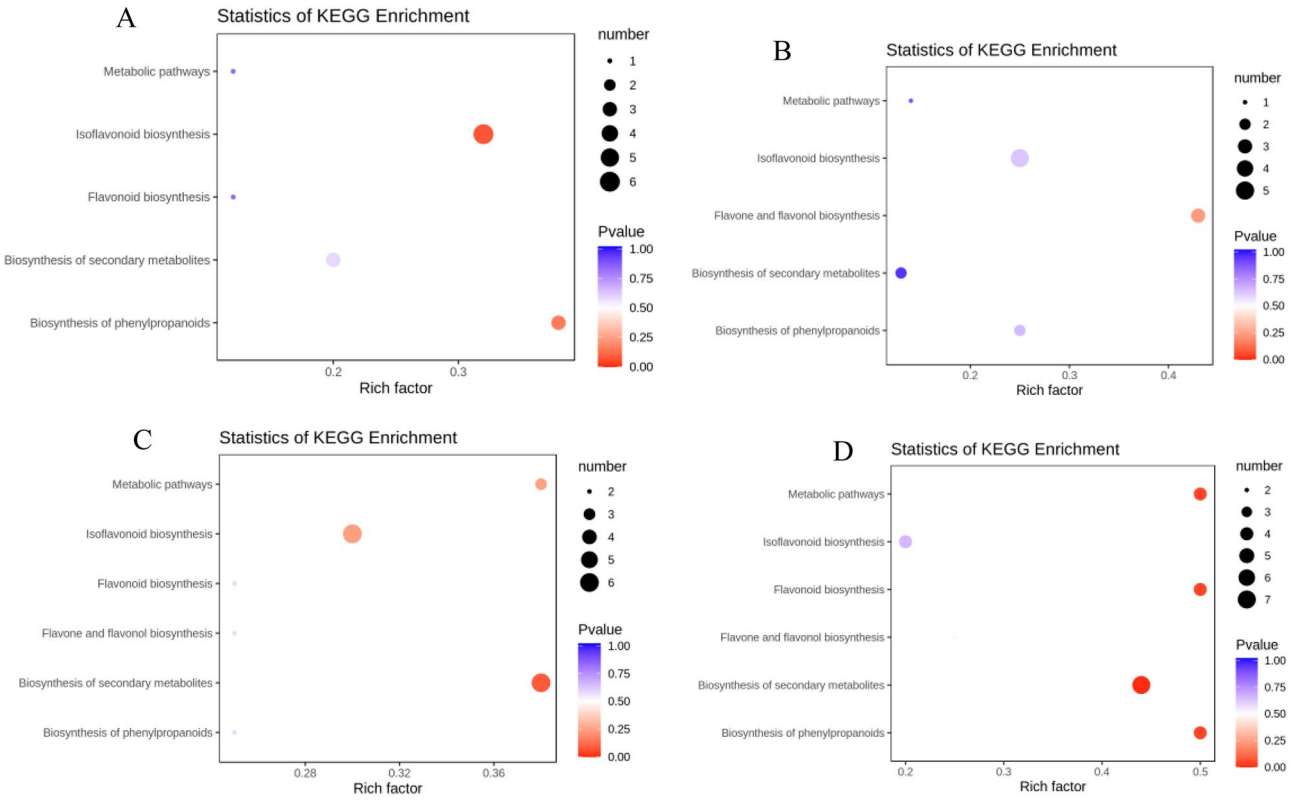

Figure 7. Differential metabolite KEGG enrichment map. (**A**): HN52 vs HN71; (**B**): HNS52 vs HNS71; (**C**): HN52 vs HNS52; (**D**): HN71 vs HNS71. Note: The abscissa represents the rich factor corresponding to each channel, the ordinate is the channel name, and the color of the point is p value. The more red, the more significant the enrichment. The size of the dot represents the number of different metabolites enriched.

Received: 11 July 2021; Accepted: 21 March 2022
Published online: 01 April 2022

References

1. Messina, M. J. Legumes and soybeans: overview of their nutritional profiles and health effects. *Am J. Clin Nutr.* **70**, 439s–450s (1999).
2. Kim, E. H., Chung, J. I., Chi, H. Y., Kim, J. A. & Chung, I. M. Analysis of phenolic compounds and isoflavones in soybean seeds (*Glycine max* (L.) Merill) and sprouts grown under different conditions. *Eur. Food Res. Technol.* **222**, 201–208 (2006).
3. Beninger, C. W. & Hosfield, G. L. Antioxidant activity extracts, condensed tannin fractions, and pure flavonoids from *Phaseolus vulgaris* L. seed coat color genotypes. *J. Agr. Food Chem.* **51**, 7879–7883 (2003).
4. Lee, C. H. *et al.* Relative antioxidant activity of soybean isoflavones and their glycosides. *Food Chem.* **90**, 735–741 (2005).
5. Aparicio-Fernandez, X., Yousef, G. G., Loarca-Pina, G., Mejia, E. & Lila, M. A. Characterization of polyphenolics in the seed coat of black jamapa bean (*Phaseolus vulgaris* L.). *J. Agr. Food Chem.* **53**, 4615–4622 (2005).
6. Beninger, C. W. *et al.* Changes in polyphenols of the seed coat during the after-darkening process in pinto beans (*Phaseolus vulgaris* L.). *J. Agr. Food Chem.* **53**, 7777–7782 (2005).
7. Diaz-Batalla, L., Widholm, J. M., Fahey, G. C., Castano-Tostado, E. & Paredes-Lopez, O. Chemical components with health implications in wild and cultivated Mexican common bean seeds (*Phaseolus vulgaris* L.). *J. Agr. Food Chem.* **54**(6), 2045–2052 (2006).
8. Rodríguez-García, C., Sánchez-Quesada, C. & Gaforio, J. J. Dietary flavonoids as cancer chemopreventive agents: an updated review of human studies. *Antioxidants.* **8**, 137 (2019).
9. Du, H., Huang, Y. & Tang, Y. Genetic and metabolic engineering of isoflavonoid biosynthesis. *Appl. Microbiol. Biot.* **86**(5), 1293–1312 (2010).
10. Chon, S. U. Total polyphenols and bioactivity of seeds and sprouts in several legumes. *Curr. Pharm. Design.* **19**(34), 6112–6124 (2013).
11. Kim, M. A. & Kim, M. J. Isoflavone profiles and antioxidant properties in different parts of soybean sprout. *J. Food Sci.* **85**, 689–695 (2020).
12. Cevallos-Casals, B. A. & Cisneros-Zevallos, L. Impact of germination on phenolic content and antioxidant activity of 13 edible seed species. *Food Chem.* **119**(4), 1485–1490 (2010).
13. Gharachorloo, M., Ghiassi Tarzi, B. & Baharinia, M. The effect of germination on phenolic compounds and antioxidant activity of pulses. *J. AM Oil Chem. Soc.* **90**(3), 407–411 (2013).
14. Ebert, W. A., Chang, C. H., Yan, M. R. & Yang, R. Y. Nutritional composition of mungbean and soybean sprouts compared to their adult growth stage. *Food Chem.* **237**(15), 15–22 (2017).
15. Kim, T. J. *et al.* Effects of a soaking-fermentation-drying process on the isoflavone and γ-aminobutyric acid contents of soybean. *Food Sci. Biotechnol.* **16**, 83–89 (2007).
16. Lee, J. H. *et al.* Isoflavone characterization and antioxidant activity of Ohio soybean. *J. Agr. Food Chem.* **52**, 2647–2651 (2004).
17. Sakthivelu, G. *et al.* Isoflavone composition, phenol content, and antioxidant activity of soybean seeds from India and Bulgaria. *J. Agr. Food Chem.* **56**, 2090–2095 (2008).

18. Phommalth, S., Hwang, Y. H., Jeong, Y. S. & Kim, Y. H. Isoflavone composition within each structural part of soybean seeds and sprouts. *J. Crop. Sci. Biotechnol.* **11**, 57–62 (2008).
19. Quinhone Júnior, A. & Ida, E. Profile of the contents of different forms of soybean isoflavones and the effect of germination time on the secompounds and the physical parameters in soybean sprouts. *Food Chem.* **166**, 173–178 (2015).
20. Yuan, J. P., Liu, Y. B., Peng, J., Wang, J. H. & Liu, X. Changes of isoflavone profile in the hypocotyls and cotyledons of soybeans during dry heating and germination. *J. Agr. Food Chem.* **57**, 9002–9010 (2009).
21. Kim, W. J., Lee, H. Y., Won, M. H. & Yoo, S. H. Germination effect of soybean on its contents of isoflavones and oligosaccharides. *Food Sci. Biotechnol.* **14**, 498–502 (2005).
22. Mekky, R. H. *et al.* Comparative metabolite profiling and antioxidant potentials of seeds and sprouts of three Egyptian cultivars of *Vicia faba* L.. *Food Res. Int.* **136**, 35–37 (2020).
23. Abu-Reidah, I. M., Arráez-Román, D., Warad, I., Fernández-Gutiérrez, A. & Segura-Carretero, A. UHPLC/MS2-based approach for the comprehensive metabolite profiling of bean (*Vicia faba* L) by-products: a promising source of bioactive constituents. *Food Res. Int.* **93**, 87–96 (2017).
24. Ammar, S., Contreras, M. D. M., Belguith-Hadrich, O., Bouaziz, M. & Segura-Carretero, A. New insights into the qualitative phenolic profile of *Ficus carica* L. fruits and leaves from Tunisia using ultra-high-performance liquid chromatography coupled to quadrupole-time-of-flight mass spectrometry and their antioxidant activity. *RSC Adv.* **5**(26), 20035–20050 (2015).
25. Alseekh, S., Perez de Souza, L., Benina, M. & Fernie, A. R. The style and substance of plant flavonoid decoration; towards defining both structure and function. *Phytochem.* **174**, 112347 (2020).
26. Bi, W. W. *et al.* Comprehensive analysis on soybean sprout characteristics of main soybean varieties in Heilongjiang Province. *Soybean Sci.* **40**(02), 209–216 (2021).
27. Torres-Fuentes, C., Contreras, M. D. M., Recio, I., Alaiz, M. & Vioque, J. Identification and characterization of antioxidant peptides from chickpea protein hydrolysates. *Food Chem.* **180**, 194–202 (2015).
28. Mohseni-Mehran, S. & Golshani, B. Simultaneous determination of levodopa and carbidopa from fava bean, green peas and green beans by high performance liquid gas chromatography. *J. Clin. Diagn. Res. JCDR.* **7**(6), 1004 (2013).
29. Abu-Reidah, I. M., Contreras, M. D. M., Arráez-Román, D., Fernández-Gutiérrez, A. & Segura-Carretero, A. UHPLC-ESI-QTOF-MS-based metabolic profiling of *Vicia faba* L. (Fabaceae) seeds as a key strategy for characterization in foodomics. *Electrophoresis* **35**(11), 1571–1581 (2014).
30. Mekky, R. H., Abdel-Sattar, E., Segura-Carretero, A. & Contreras, M. D. M. Phenolic compounds from sesame cake and antioxidant activity: a new insight for Agri-Food residues' significance for sustainable development. *Foods.* **8**(10), 432 (2019).
31. Valente, I. M. *et al.* Unravelling the phytonutrients and antioxidant properties of European *Vicia faba* L. seeds. *Food Res. Int.* **116**, 888–896 (2019).
32. Mekky, R. H. *et al.* Profiling of phenolic and other compounds from Egyptian cultivars of chickpea (*Cicer arietinum* L.) and antioxidant activity: a comparative study. *RSC Adv.* **5**(23), 17751–17767 (2015).
33. Rodríguez-Pérez, C., Zengin, G., Segura-Carretero, A., Lobine, D. & Mahomoodally, M. F. Chemical fingerprint and bioactivity evaluation of *Globularia orientalis* L. and Globularia trichosantha Fisch. C. A. Mey. using non-targeted HPLC-ESI-QTOF-MS approach. *Phytochem. Anal.* **30**(2), 237–252 (2019).
34. Abouzed, T. K. *et al.* Red onion scales ameliorated streptozotocin-induced diabetes and diabetic nephropathy in Wistar rats in relation to their metabolite fingerprint. *Diabetes Res. Clin. Pract.* **140**, 253–264 (2018).
35. Gurfinkel, D. M. & Rao, A. V. Soyasaponins: The relationship between chemical structure and colon anticarcinogenic activity. *Nutr. Cancer* **47**(1), 24–33 (2003).
36. Hwang, E.-S. Antioxidant and quinone reductase activity of soyasaponins in hepa1c1c7 mouse hepatocarcinoma cells. *Prev. Nutr. Food Sci.* **22**(4), 300 (2017).
37. Megías, C., Cortés-Giraldo, I., Girón-Calle, J., Alaiz, M. & Vioque, J. Characterization of Vicia (Fabaceae) seed water extracts with potential immunomodulatory and cell antiproliferative activities. *J. Food Biochem.* **42**(5), e12578 (2018).
38. Fabiola, A. G.-O. *et al.* Profile analysis and correlation across phenolic compounds, isoflavones and antioxidant capacity during germination of soybeans (*Glycine max* L.). CyTA. *J. Food.* **15**(4), 516–524 (2017).
39. Dueñas, M., Hernández, T., Estrella, I. & Fernández, D. Germination as a process to increase the polyphenol content and antioxidant activity of lupin seeds (*Lupinus angustifolius* L.). *Food Chem.* **117**, 599–607 (2009).
40. Michalak, A. Phenolic compounds and their antioxidant activity in plants growing under heavy metal stress. *Pol. J. Environ. Stud.* **15**, 523–530 (2006).
41. Wu, Z. *et al.* Germination dramatically increases isoflavonoid content and diversity in Chickpea (*Cicer arietinum* L.) seeds. *J. Agr. Food Chem.* **60**(35), 8606–8615 (2012).
42. Valente, I. M. *et al.* Profiling of phenolic compounds and antioxidant properties of European varieties and cultivars of *Vicia faba* L. pods. *Phytochemistry* **152**, 223–229 (2018).
43. Ayet, G. *et al.* Effect of germination, under different environmental conditions, on saponins, phytic acid and tannins in lentils (*Lens culinaris*). *J. Sci. Food Agr.* **74**(2), 273–279 (1997).
44. Kanehisa, M. & Goto, S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**(1), 27–30 (2000).

Acknowledgements

This study was subsidized by Heilongjiang Province Postdoctoral Science Foundation (LBH-Z16184) and funded by Project of Harbin Science and Technology Bureau (2017RAQYJ034), College Project of Heilongjiang Academy of Agricultural Sciences (2017BZ14, 2020YYYF038) and Agricultural Science and Technology Innovation Leapfrog Project of Heilongjiang Academy of Agricultural Sciences (HNK2019CX01-14-5, HNK2019CX19-09).

Author contributions

W.B., G.Z., and Y.Z., X.X. wrote the main manuscript text and jinsheng Wang, guangjin Wang prepared Figs. 1–3, S.L., W.H. and T.B. prepared Figs. 4–7, Jintong Li prepared Table 1. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to W.B.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.