Food Research International 106 (2018) 842–852

Contents lists available at ScienceDirect

Food Research International

journal homepage: www.elsevier.com/locate/foodres

Metabolomics approach for understanding geographical dependence of soybean leaf metabolome

Dae-Yong Yun[a,1], Young-Gyu Kang[b,1], Eun-Hee Kim[c], Myoyeon Kim[b], Nok-Hyun Park[b], Hyang-Tae Choi[b], Gun Hee Go[d], John Hwan Lee[b], Jun Seong Park[b,*], Young-Shick Hong[a,*]

[a] Division of Food and Nutrition, Chonnam National University, Yongbong-ro, Buk-gu, Gwangju 500-757, Republic of Korea
[b] Applied Technology & Research Division, R&D Center, AmorePacific Corporation, Yongin-si, Gyeonggi-do 446-729, Republic of Korea
[c] Protein Structure Group, Korea Basic Science Institute, Cheongwon-Gu, Cheongju-Si, Chungbuk 363-883, Republic of Korea
[d] Osulloc Tea R&D Center, Osulloc Farm Co., Ltd., Jeju 699-820, Republic of Korea

ARTICLE INFO

Keywords:
Metabolomics
Metabolism
NMR
Soybean leaves
Glycine max
Environmental stress

ABSTRACT

The soybean plant (*Glycine max*) is widely used as an ingredient in various foods, nutraceuticals and cosmetics, due to their diverse bioactive compounds. Their metabolic compositions are likely affected by environmental conditions during growth. To investigate the influence of different environmental conditions on the metabolite composition of soybean leaves, we cultivated soybean (*G. max* Sinhwa) in the southernmost island and volcanic region of Korea, and in the central section and limestone region of the Korean peninsula. Comprehensive metabolite variations of their leaves were analyzed through ^{1}H NMR-based metabolomics approach. With marked differences in soil compositions and climatic conditions between the two growing areas, differences in accumulations of pinitol and diverse flavonoids were noted between the soybean leaves, reflecting the distinct metabolism of soybean plants for physiological adaptation toward different environmental conditions. Therefore, the current study highlights the geographical dependences of diverse soybean leaf metabolites for developing biofunction-enhanced soybean products.

1. Introduction

The soybean plant is used as a major nutrition source of protein, lipid and carbohydrate, and its seeds, particularly, are processed into various foods, such as tofu, soybean milk and oil. It is well-known that the abundant phytochemicals, typically flavonoids, contribute to many of the health promoting attributes of soybean. Moreover, soybean is regarded as a nutraceutical ingredient in food and beauty industries (Isanga & Zhang, 2008). Many articles have reported that soybean leaves contain diverse bioactive compounds, which differ from those in soybean seeds, and have flavonoids and pterocarpans, such as kaempferol and its glycosides, and coumestrol. These bioactive compounds in soybean leaves contribute to alleviating various diseases and metabolic disorders. For example, Choi et al. (2014) reported that soybean leaf extracts play an important role in improving blood glucose levels, insulin resistance, adiposity and dyslipidemia in pre-diabetic adults. The beneficial effects of soybean leaf extracts on diabetes, obesity and glucose homeostasis have also been reported in animal experiments (Li et al., 2015). Since diverse compounds in soybean leaves likely contribute to health promoting effects, it is important to better understand the perturbations of the compounds according to cultivation conditions of soybean plants including cultivars and environmental factors. For example, the ways to increase the synthesis of the bioactive compounds in plants could provide an important information in industries producing bioactive-related and cosmetic products.

A recent previous study revealed a distinct flavonoid metabolism between cultivated and semi-wild soybean, together with a wide range of soybean leaf metabolites, and reported the strong dependence of the soybean leaf metabolome on its cultivar (Yun et al., 2016). Moreover, intrinsic metabolism of various tea (*Camellia sinensis*) cultivars that are rich in taste, epigallocatechin gallate (EGCG) and epigallocatechin-3-*O*-(3-*O*-methyl)-gallate (EGCG3"Me) have been reported (Ji et al., 2017). These studies highlight their metabolic phenotypes through comprehensive metabolite profiling and provide a useful information for assessing and developing a new plant cultivar. Therefore, further study of how environmental conditions affect the chemical or metabolite compositions in soybean plants may lead to a better understanding of the intrinsic metabolism of soybean plants grown under different

* Corresponding authors.
E-mail addresses: superbody@amorepacific.com (J.S. Park), chtiger@jnu.ac.kr (Y.-S. Hong).
[1] These authors equally contributed to this work.

https://doi.org/10.1016/j.foodres.2018.01.061
Received 26 November 2017; Received in revised form 22 January 2018; Accepted 24 January 2018

environmental conditions.

Metabolome is an assemblage of metabolites and plant metabolome is considerably affected by environmental or external stress, developmental signals and genotype (Kim, Choi, & Verpoorte, 2011). Thus, metabolomics can provide important information about comprehensive systemic metabolite changes under various conditions. Nuclear magnetic resonance (NMR) spectroscopy has been prominently used for plant metabolomics analysis because it facilitates high-throughput and impartial analysis of the diverse metabolites present in a plant, as well as structural elucidation of new metabolites (Kim, Choi, & Verpoorte, 2010). Examples include the study of green tea metabolite dependence on geography (Lee et al., 2015), metabolite profiling of *Ipomoea aquatica* at different growth stages (Lawal et al., 2015), and the investigation of *Astragali radix* of different ages and geographical origin (Zheng et al., 2015).

Therefore, the present study uses a ^{1}H NMR-based metabolomics approach to explore comprehensive metabolite variations and potential bioactivity in soybean plants cultivated in two distinctly different environments. One environment is the Yeongwol area which is located in the central part of the Korean peninsula and is a representative limestone region, and the other is the Jeju island area which is located in the southernmost part of Korea and is a volcanic region.

2. Materials and methods

2.1. Chemicals

Coumestrol, apigenin, apigetrin (apigenin-7-*O*-glucoside), astragalin (kaempferol-3-*O*-glucoside), gallic acid, Folin Ciocalteu's phenol reagent, catechin hydrate, 2,2-diphenyl-2-picrlhydrazyl (DPPH), 2,2′-azino-bis-(3-ethylbenzothiazoline-6-sulfonic acid) (ABTS) and potassium persulfate were purchased from Sigma Chemical Co. (St. Louis, MO, USA). For the NMR analysis, methanol-*d*4 (CD_3OD, 99.8% 4H) and deuterium oxide (D_2O, 99.9% 2H) were also purchased from Sigma Chemical Co. (St. Louis, MO, USA). Daidzin, genistin and malonylgenistin were purchased from Wako Pure Chemical Industries, Ltd. (Osaka, Japan). Sodium carbonate (Na_2CO_3), sodium nitrite ($NaNO_2$), aluminum chloride ($AlCl_3$), sodium hydroxide (NaOH) and ascorbic acid were purchased from Junsei (Tokyo, Japan). HPLC grade solvents of ethyl alcohol (EtOH) and methanol (MeOH) were purchased from E. Merck Co. (Darmstadt, Germany) and J.T.Baker® (Center Valley, PA, USA), respectively.

2.2. Soil analysis

Triplicate soil samples were obtained from Dosun-dong, Seogwiposi, Jeju-do (Jeju; 33° 27′ 30.53″N, 126° 47′ 43.05″E) and Mungok-ri, Buk-myeon, Yeongwol-gun, Gangwon-do (Yeongwol; 37° 24′ 86.33″N, 128° 41′ 03.96″E) in the Republic of Korea in 2015. The soil samples were analyzed for physicochemical properties. Organic matter (O.M.) was determined by the method of Walkley and Black (1934). Total nitrogen (N) was measured by the Kjeldahl method (Bremner, 1960). Soil salinity (NaCl) and electrical conductivity (EC), and pH were determined in a 1:5 soil-to-water solution using EC and pH meters, respectively. Cation exchange capacity (CEC) was determined by the 1-N ammonium acetate (pH 7.0) method (Schollenberger & Simon, 1945). Available phosphate (Avail. P_2O_5) was determined by the Lancaster test, using UV/VIS spectrophotometer. Exchangeable potassium, calcium, magnesium and sodium ions were extracted with 1-N ammonium acetate (pH 7.0), and their concentrations were analyzed by inductively coupled plasma optical emission spectrometry (ICP-OES; Optima 5300DV, PerkinElmer, Waltham, MA, USA). The soil texture was determined by USDA soil taxonomy.

2.3. Soybean leaf samples

The soybean leaves used in this study were from the *Glycine max* (*G. max*) Sinhwa cultivar cultivated in the Jeju and Yeongwol areas. Ten samples of these leaves were collected at 10 different parcels of land from the two areas, thus, providing 10 biological replications. The leaves were at the R4/R5 stage defined as rapid pod growth and beginning development of the seed pod, and the R7 stage defined as full maturity. All fresh soybean leaf samples were stored at −80 °C until extracted or processed. The frozen leaves were ground with a mortar and pestle under liquid nitrogen. The ground samples were transferred into a plastic tube using a spatula, kept in a deep freezer for 24 hours (h), and then freeze dried for 48 h. Extracts were prepared from each freeze-dried sample, and were analyzed by NMR, the aluminum chloride colorimetric assay, the Folin–Ciocalteu assay, DPPH- and ABTS-radical scavenging assays and HPLC experiments. The analytical methods for DPPH- and ABTS-radical scavenging activities, and total phenolic (TPC) and flavonoid contents (TFC) of soybean leaf extract were described in our previous study (Yun et al., 2016).

2.4. ^{1}H NMR spectroscopic analysis of soybean leaf extracts

Metabolite extraction of soybean leaves for ^{1}H NMR analysis was carried out according to Kim et al. (2010) A 10 mg freeze-dried sample was dissolved in a mixture of methanol-d_4 (CD_3OD, 490 μL) and deuterium water (D_2O, 210 μL) in a 1.5 mL Eppendorf (EP) tube. The mixture was sonicated at room temperature for 20 min to extract soybean leaf metabolites and then centrifuged at 13,000 rpm at room temperature for 15 min. The supernatant of each soybean leaf extract (550 μL) was transferred into 5 mm NMR tubes. CD_3OD in the supernatant provides a field frequency lock, and sucrose was used as a chemical shift reference (^{1}H, δ 5.40). ^{1}H NMR spectra were acquired on a Bruker Avance 700 spectrometer (Bruker Biospin, Rheinstetten, Germany), operating at 700.40 MHz ^{1}H frequency and a temperature of 298 K, using a cryogenic triple-resonance probe and a Bruker automatic injector. Signal assignment for representative samples was facilitated by two-dimensional (2D) total correlation spectroscopy (TOCSY), heteronuclear multiple bond correlation (HMBC), heteronuclear single quantum correlation (HSQC), and spiking experiments.

2.5. NMR data processing and multivariate statistical analysis

All NMR spectra were manually corrected for phase and baseline distortions and then converted to ASCII format. The ASCII format files were imported into MATLAB (R2010b; The Mathworks Inc., Natick, USA). The spectra were normalized to a total integral to avoid dilution effects of samples. After calibrating to glucose (^{1}H, δ 5.23), the spectra were further aligned using the *i*coshift method (Savorani, Tomasi, & Engelsen, 2010). The full-resolution ^{1}H NMR spectra, without spectrum bucketing or binning, were used for the multivariate statistical analysis, after excluding several regions corresponding to residual water (δ 4.5–5.2 ppm) and methanol (δ 3.28–3.34 ppm) prior to the normalization and spectrum alignment. The resulting data sets were then imported into SIMCA-P version 12.0 (Umetrics, Umeå, Sweden), and a mean-centered scaling method was applied for multivariate statistical analysis. The principal component analysis (PCA), an unsupervised pattern recognition method, was initially performed in order to examine intrinsic variation in the data set. A supervised pattern recognition method, orthogonal projection on latent structure-discriminant analysis (OPLS-DA) was used to extract maximum information on discriminant compounds from the data (Bylesjo et al., 2006). OPLS-DA provides a way of removing systematic variation from an input data set X (compounds or metabolites), which is not correlated with the response set Y (discriminant classes). Hotelling's T^2 regions, shown as an ellipse in the scores plot, defined the 95% confidence interval of the modeled variation. The OPLS-DA models were validated