



ELSEVIER

Contents lists available at ScienceDirect

## Data in Brief

journal homepage: [www.elsevier.com/locate/dib](http://www.elsevier.com/locate/dib)

## Data Article

# Datasets of mung bean proteins and metabolites from four different cultivars



Akiko Hashiguchi<sup>a,\*</sup>, Wei Zhu<sup>b</sup>, Jingkui Tian<sup>b</sup>,  
Setsuko Komatsu<sup>c,d,\*\*</sup>

<sup>a</sup> Faculty of Medicine, University of Tsukuba, Tsukuba 305-8577, Japan

<sup>b</sup> College of Biomedical Engineering & Instrument Science, Zhejiang University, Hangzhou 310027, China

<sup>c</sup> National Institute of Crop Science, National Agriculture and Food Research Organization, Tsukuba 305-8518, Japan

<sup>d</sup> Faculty of Life and Environmental Sciences, University of Tsukuba, Tsukuba 305-8572, Japan

## ARTICLE INFO

## Article history:

Received 12 May 2017

Accepted 28 June 2017

Available online 4 July 2017

## Keywords:

Mung bean

Health-promoting food

Nutritional/nutraceutical value

Metabolic network

Metabolomics

Proteomics

## ABSTRACT

Plants produce a wide array of nutrients that exert synergistic interaction among whole combinations of nutrients. Therefore comprehensive nutrient profiling is required to evaluate their nutritional/nutraceutical value and health promoting effect. In order to obtain such datasets for mung bean, which is known as a medicinal plant with heat alleviating effect, proteomic and metabolomic analyses were performed using four cultivars from China, Thailand, and Myanmar. In total, 449 proteins and 210 metabolic compounds were identified in seed coat; whereas 480 proteins and 217 metabolic compounds were detected in seed flesh, establishing the first comprehensive dataset of mung bean for nutraceutical evaluation.

© 2017 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

DOI of original article: <http://dx.doi.org/10.1016/j.bbapap.2017.05.006>

Abbreviations: LC, liquid chromatography; MS, mass spectrometry; GC, gas chromatograph; TOF, time-of-flight

\* Corresponding author.

\*\* Corresponding author at: Faculty of Medicine, University of Tsukuba, Tsukuba 305-8577, Japan.

E-mail addresses: [hashiguchi.akiko.ge@un.tsukuba.ac.jp](mailto:hashiguchi.akiko.ge@un.tsukuba.ac.jp) (A. Hashiguchi), [komatsu.setsuko.fu@u.tsukuba.ac.jp](mailto:komatsu.setsuko.fu@u.tsukuba.ac.jp) (S. Komatsu).

<http://dx.doi.org/10.1016/j.dib.2017.06.051>

2352-3409/© 2017 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

## Specifications Table

Subject area	Biology
More specific subject area	Plant Science
Type of data	Figure, Tables
How data was acquired	Gel-free/label-free proteomic analysis using data-dependent acquisition mode on a LTQ Orbitrap mass spectrometer coupled to an Ultimate 3000 nanoLC system. Metabolic analysis using an Agilent 7890 GC system coupled with a Pegasus HT TOF-MS.
Data format	Filtered
Experimental factors	Comparison of metabolic pathways in seed coat and flesh of mature mung bean
Experimental features	Whole-protein and metabolic compounds of mung bean coat and flesh were identified from four cultivars.
Data accessibility	Datasets are directly provided with this article.

## Value of the data

- The data represent comprehensive repository of protein and primary metabolites contained in seed coat and flesh of mung bean, a medicinal plant with heat alleviating activity.
- This experimental design allows multiomics analysis of metabolic pathways of seed coat and flesh of mung bean.
- Differences in metabolic pathways and bioactive compound-containing pattern between seed coat and flesh were revealed using this data as described in [1].

## 1. Data

The data here represent different omics approaches to understand the mung bean metabolic pathways and compound-containing pattern in seed coat and flesh. The dataset is associated with the research article in BBA Proteins and Proteomes entitled “Proteomics and metabolomics-driven pathway reconstruction of mung bean for nutraceutical evaluation” and contains eight lists of proteins and two lists of metabolites obtained from four cultivars originated from different habitats (Fig. 1).

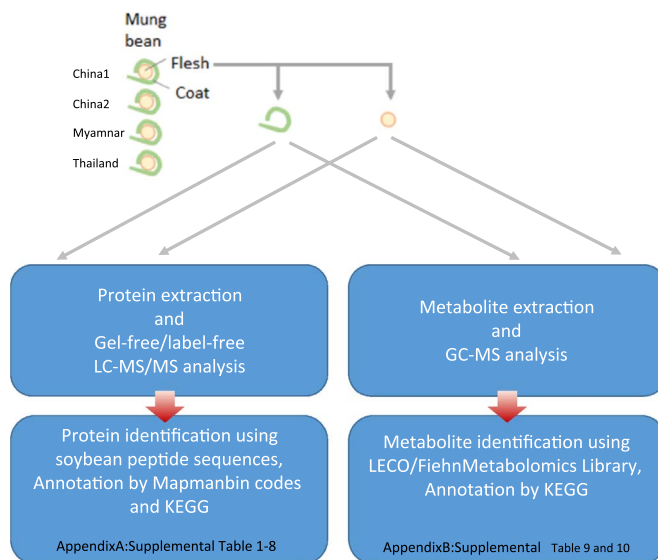
## 2. Experimental design, materials and methods

### 2.1. Plant materials

Mung beans (*Vigna radiata* (L.) R. Wilczek) from different habitats in Asian countries were purchased from local supermarkets in Tokyo and Yokohama, Japan. These cultivars were referred as China 1, China 2, Thailand, Myanmar, respectively, according to their habitats. For protein and metabolite extraction, mung bean seeds was soaked in milliQ water to separate coat and flesh.

### 2.2. Gel-free/label-free proteomic analysis

The coat and flesh was ground to powder in liquid nitrogen using a mortar and pestle and transferred to an acetone solution containing 10% trichloroacetic acid and 0.07% 2-mercaptoethanol. The proteins were extracted as described in [1]. After enrichment with methanol and chloroform to



**Fig. 1.** Flowchart of experiments and data processing.

remove any detergent, proteins were digested into tryptic peptides with trypsin and lysyl endopeptidase (Wako, Osaka, Japan). Peptide identification was performed by nanoliquid chromatography (LC) MS/MS with a nanospray LTQ Orbitrap mass spectrometer (Thermo Fisher Scientific, San Jose, CA, USA) coupled to an Ultimate 3000 nanoLC system (Dionex). Full-scan mass spectra were acquired in the mass spectrometer over 400–1500  $m/z$  with a resolution of 30,000 in a data dependent mode as previously described [2,3]. Proteins were identified by Mascot search engine (version 2.5.1, Matrix Science, London, UK) of a soybean peptide database (55,787 sequences) constructed from the Phytosome (version 9.1; <http://www.phytosome.net/soybean>). The acquired raw data files were processed by Proteome Discoverer software (version 1.4.0.0288; Thermo Fisher Scientific). Peptides with a percolator ion score of more than 13 ( $p < 0.05$ ) alone were used for analysis. Proteins with single matched peptide were also taken into consideration. Protein abundance in mol% was calculated based on the empAI value [4]. Appendix file A contained the datasets with identified proteins from seed coat and flesh of four different cultivars. In total, 449 and 480 proteins were identified in seed coat and flesh, respectively, as described in [1].

### 2.3. Metabolomic analysis

Metabolites were extracted from dried sample powder with 0.4 mL extraction liquid (V methanol:V water=3: 1) using ball mill (JXFSTPRP-24, Shanghai Superscience Technology, Shanghai, China) as described in [1]. Gas chromatograph time-of-flight (GC TOF)-MS analysis was performed according to [5] using an Agilent 7890 GC system (Agilent, Palo Alto, CA, USA) coupled with a Pegasus HT TOF-MS (LECO, St Joseph, MI, USA). Peak analysis was performed by Chroma TOF 4.3X software (LECO) and LECO-Fiehn Rtx5 [6]. The similarity value obtained from the LECO/Fiehn Metabolomics Library was used for the evaluation of the accuracy of the discriminating compound identification is reliable. If the similarity is less than 200, the compound is defined as an “analyte”. The compound with a similarity between 200 and 700 is considered as a putative annotation. Appendix file B contains identified metabolic compounds from seed coat and flesh of four different cultivars.

#### 2.4. Annotation of proteins and metabolites with KEGG identifiers

Proteins were categorized based on function using MapMan bin codes [7]. For pathway mapping, identifiers in the Kyoto Encyclopedia of Genes and Genomes (KEGG) database [7] (<http://www.genome.jp/kegg/>) were retrieved from MapMan system or KEGG COMPOUND search.

#### Acknowledgement

This work was financially supported by JSPS KAKENHI Grant number 15H00474.

#### Transparency document. Supplementary material

Transparency data associated with this article can be found in the online version at <http://dx.doi.org/10.1016/j.dib.2017.06.051>.

#### Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at <http://dx.doi.org/10.1016/j.dib.2017.06.051>.

#### Appendix B. Supplementary material

Supplementary data associated with this article can be found in the online version at <http://dx.doi.org/10.1016/j.dib.2017.06.051>.

#### References

- [1] A. Hashiguchi, W. Zhu, J. Tian, S. Komatsu, Proteomics and metabolomics-driven pathway reconstruction of mung bean for nutraceutical evaluation *Biochim. Biophys. Acta*, 1865, 2017, 1057–1066.
- [2] Y. Nanjo, L. Skultety, L. Uváčková, K. Klubicová, M. Hajduch, S. Komatsu S, Mass spectrometry-based analysis of proteomic changes in the root tips of flooded soybean seedlings, *J. Proteome Res.* 11 (2012) 372–385.
- [3] A. Hashiguchi, K. Hitachi, W. Zhu, J. Tian, K. Tsuchida, S. Komatsu, Mung bean (*Vigna radiata* (L.)) coat extract modulates macrophage functions to enhance antigen presentation: a proteomic study, *J. Proteom.* 161 (2007) 26–37.
- [4] K. Shinoda, M. Tomita, Y. Ishihama, emPAI Calc—for the estimation of protein abundance from large-scale identification data by liquid chromatography–tandem mass spectrometry, *Bioinformatics* 26 (2010) 576–577.
- [5] B. Yang, X. Wang, C. Gao, M. Chen, Q. Guan, J. Tian, S. Komatsu, Proteomic and metabolomic analyses of leaf from *Clematis terniflora* DC. Exposed to high-level ultraviolet-B irradiation with dark treatment, *J. Proteome Res.* 15 (2016) 2643–2657.
- [6] T. Kind, G. Wohlgemuth, D.Y. Lee, Y. Lu, M. Palazoglu, S. Shahbaz, O. Fiehn, FiehnLib—mass spectral and retention index libraries for metabolomics based on quadrupole and time-of-flight gas chromatography/mass spectrometry, *Anal. Chem.* 81 (2009) 10038–10048.
- [7] M. Kanehisa, Y. Sato, M. Kawashima, M. Furumichi, M. Tanabe, KEGG as a reference resource for gene and protein annotation, *Nucleic Acids Res.* 44 (2016) D457–D462.