Rural Development Administration
National Institute of Crop Science

Multivariate statistical analysis using FT-IR spectrum data of Soybean Core Collection in Korea

Seung-Yeob Song[1]*, Eunji Suh[1], Mi Ja Lee[1], Han Gyeol Lee[1] and Woo Duck Seo[1]

[1]Crop Foundation Research Division, National institute of Crop Science, Rural Development Administration, Wanju 55365, Republic of Korea

*Corresponding author: Tel. +82-63-238-5305, e-mail: s2y337@korea.kr

Abstract

To determine whether FT-IR spectral analysis based on multivariate analysis for whole cell extracts can be used to discriminate between Soybean plants seed at the metabolic level, leaves of 383 soybean core collection plants were subjected to Fourier transform infrared(FT-IR) spectroscopy. FT-IR spectral data from leaves were analyzed by principal component analysis (PCA), partial least square discriminant analysis (PLS-DA) and hierarchical clustering analysis (HCA). FT-IR spectra confirmed typical spectral differences between the frequency regions of 1,700 - 1,500, 1,500 - 1,300 and 1,100 - 950 cm^{-1}, respectively. These spectral regions reflect the quantitative and qualitative variations of amide I, II from amino acids and proteins (1,700 - 1,500cm^{-1}), phosphodiester groups from nucleic acid and phospholipid (1,500 - 1,300cm^{-1}) and carbohydrate compounds (1,100 - 950cm^{-1}). PCA revealed separate clusters that corresponded to their species relationship. Thus, PCA could be used to distinguish between soybean with different metabolite contents. PLS-DA showed similar metabolite contents of soybean. Further more these metabolic discrimination systems could be used for the rapids election and classification of useful soybean cultivars.

Materials and Methods

I. Material

Plants : Soybean(perilla seed 2019, 2020)
used by 20mg in powder form.

II. Methods

Whole cell extracts were prepared from freeze-dried perilla powder (20 mg) in 1.5 mL Eppendorf tubes by using 20% (v/v) methanol (200 μL)

The samples were mixed vigorously by vortexing and inverting and were then incubated in a 50°C water bath for 20 min and additionally vortexed every 10 min

The mixture was centrifuged at 13,000rpm for 15 min at room temperature, and the supernatants were transferred into a new tube. A second centrifugation step was performed at 13,000rpm for 1 min to remove the debris.

FT-IR spectrum and spectral data processing

FT-IR spectrometer (Tensor II, Bruker) deuterated triglycine sulfate (DTGS) detector was used for infrared measurements 5 ul of prepared whole cell extracts were loaded on 384-well ZnSe plate on hot plate pre-warmed at 37°C. After drying each sample HTS-XT (Bluker Optics GmbH, Ettlingen, Germany). FT-IR spectra were spectra were acquired with OPUS program (ver. 7.0)

PCA (Principal component analysis) and PLS-DA (Partial least squar es discriminant analysis) was the analysis of FT-IR spectrum data (18 00 – 800cm^{-1}) by using the R (ver 3.3.2) program

Results

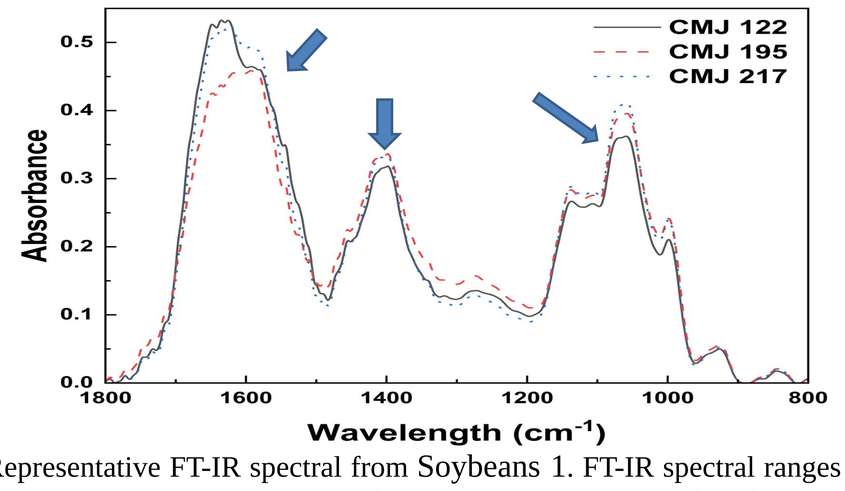

Fig. 1. Representative FT-IR spectral from Soybeans 1. FT-IR spectral ranges showed quantitative information of protein/amide I, II (1500-1700cm^{-1}), phosphodiester group (1300-1500cm^{-1}), and sugar compound (950-1100cm^{-1}).

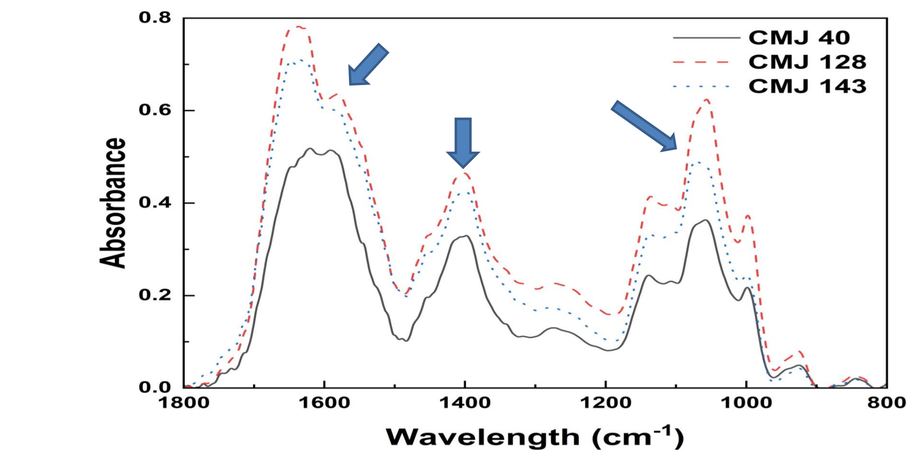

Fig. 2. Representative FT-IR spectral from Soybean 2. FT-IR spectral ranges showed quantitative information of protein/amide I, II (1500-1700cm^{-1}), phosphodiester group (1300-1500cm^{-1}), and sugar compound (950-1100cm^{-1}).

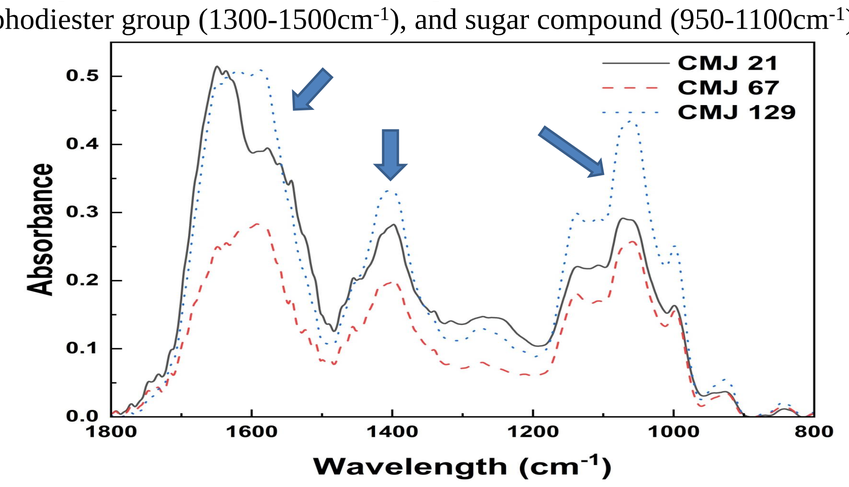

Fig. 3. Representative FT-IR spectral from Soybean 3. FT-IR spectral ranges showed quantitative information of protein/amide I, II (1500-1700cm^{-1}), phosphodiester group (1300-1500cm^{-1}), and sugar compound (950-1100cm^{-1}).

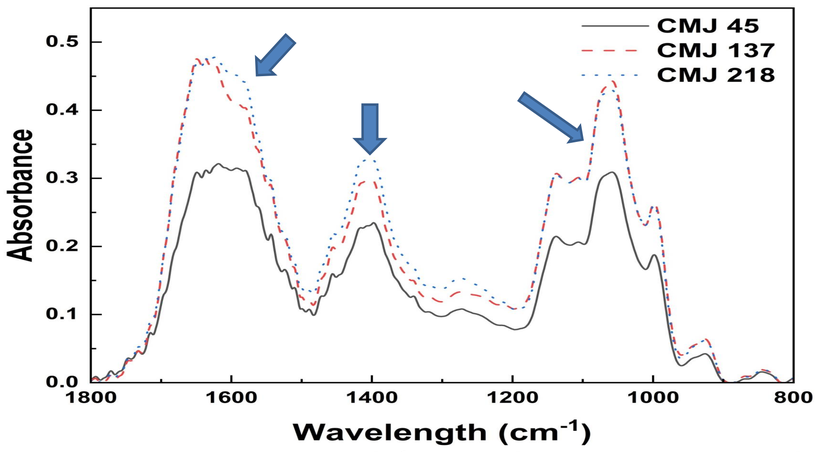

Fig. 3. Representative FT-IR spectral from Soybean 4. FT-IR spectral ranges showed quantitative information of protein/amide I, II (1500-1700cm^{-1}), phosphodiester group (1300-1500cm^{-1}), and sugar compound (950-1100cm^{-1}).

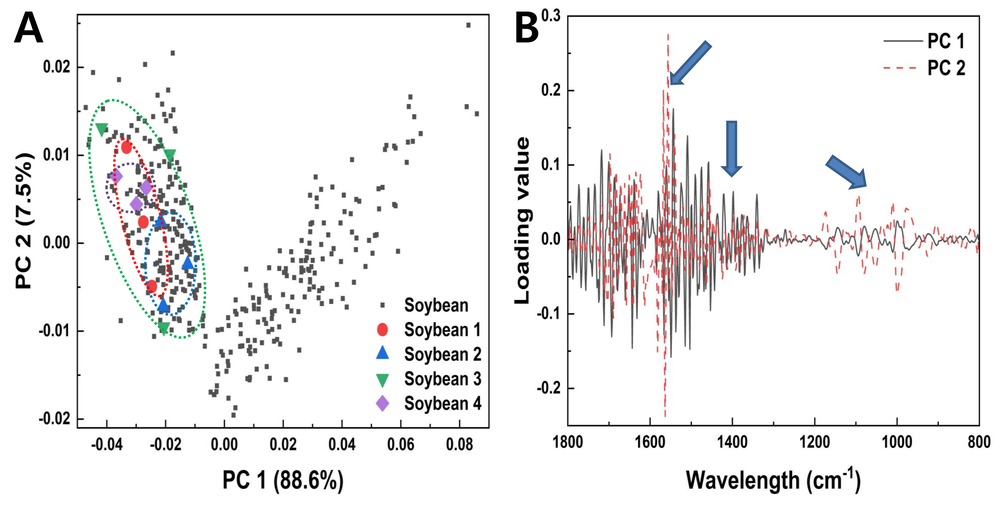

Fig. 4. . PCA score plot (A) and Loading value plot (B) of PCA analysis from FT-IR data of soybean. Dot rectangles represent significant FT-IR spectral region for metabolic discrimination of each soybean 1-4.