

ダイズプロテオームデータベース： 電気泳動を基盤にした手法と質量分析計を基盤にした手法 から得られたタンパク質データの統合

○坂田克己1)、大柳一2)、杉本和彦3)、小松節子4)

1)前橋工科大学、2)King Abdullah University of Science and Technology、3)次世代作物開発研究セ、4)筑波大学



© BY-NC-ND Licensed under a Creative Commons 表示 4.0 国際ライセンス (cc) 17

Abstract

The Soybean Proteome Database (SPD) stores data on soybean proteins obtained with gel-based and gel-free proteomic techniques. The database was constructed to provide information on proteins for functional analyses. The majority of the data is focused on soybean (*Glycine max* 'Enrei'). The growth and yield of soybean are strongly affected by environmental stresses such as flooding. The database was originally constructed using data on soybean proteins separated by two-dimensional polyacrylamide gel electrophoresis, which is a gel-based proteomic technique. Since 2015, the database has been expanded to incorporate data obtained by label-free mass spectrometry-based quantitative proteomics, which is a gel-free proteomic technique. Here, the portions of the database consisting of gel-free proteomic data are described. The gel-free proteomic database contains 39,212 proteins identified in 63 sample sets, such as temporal and organ-specific samples of soybean plants grown under flooding stress or non-stressed conditions. In addition, data on organellar proteins identified in mitochondria, nuclei, and endoplasmic reticulum are stored. Furthermore, the database integrates multiple omics data such as genomics, transcriptomics, metabolomics, and proteomics. The SPD database is accessible at <http://proteome.dc.affrc.go.jp/Soybean/>.

Introduction

Soybean is an important food crop that is rich in vegetable oil and protein, as well as phytochemicals such as isoflavones and phenolic compounds. Soybean has a complex genome because multiple neopolyploid speciation events are superimposed on the paleopolyploid genome. To provide information on proteins for functional analyses, the Soybean Proteome Database (SPD) was originally constructed using data for soybean proteins separated by two-dimensional polyacrylamide gel electrophoresis, which is classified as a gel-based proteomic technique (Ohyang et al., 2012). In addition, the data were integrated with the information from transcriptomics, proteomics, and metabolomics (Sakata et al., 2009). The database has been improved by incorporation of data from label-free mass spectrometry (MS)-based quantitative proteomics, which is classified as a gel-free proteomic technique. This technique is directly applied to peptide samples after digestion of proteins for liquid chromatography (LC)-MS. In this report, the portions of the SPD derived from gel-free proteomics are described. Although soybean is a widely cultivated crop, its growth and yield are markedly affected by adverse environmental conditions, especially flooding stress. The SPD also stores information on the flooding-responsive proteins, which were analyzed in temporal and organ-specific protein profiles of soybean grown in the presence or absence of flooding stress in the present study (Komatsu et al., 2017).

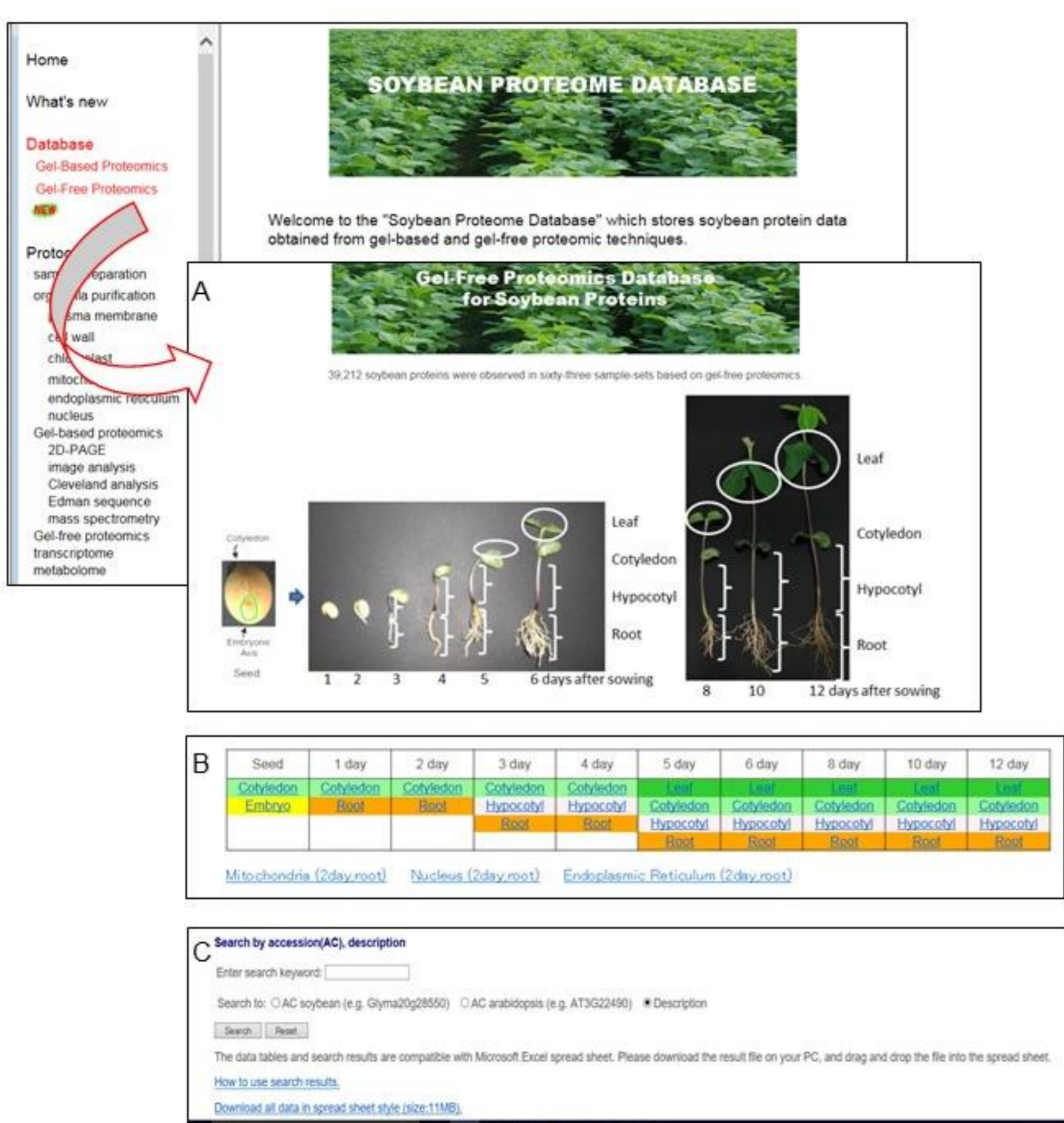


Fig. 1. Web presentation of SPD.
After opening the Soybean Proteome Database homepage (<http://proteome.dc.affrc.go.jp/Soybean/>), the lefthand menu includes options for "Gel-based Proteomics" and "Gel-free Proteomics". When "Gel-free Proteomics" is selected, images of soybean plants (A) and a list of organs or organelles over a sequence of time points (B) are displayed. The "Enter Search keyword" box allows the user to search for protein information and "Download all data in spread sheet style" is used to download all data (C).

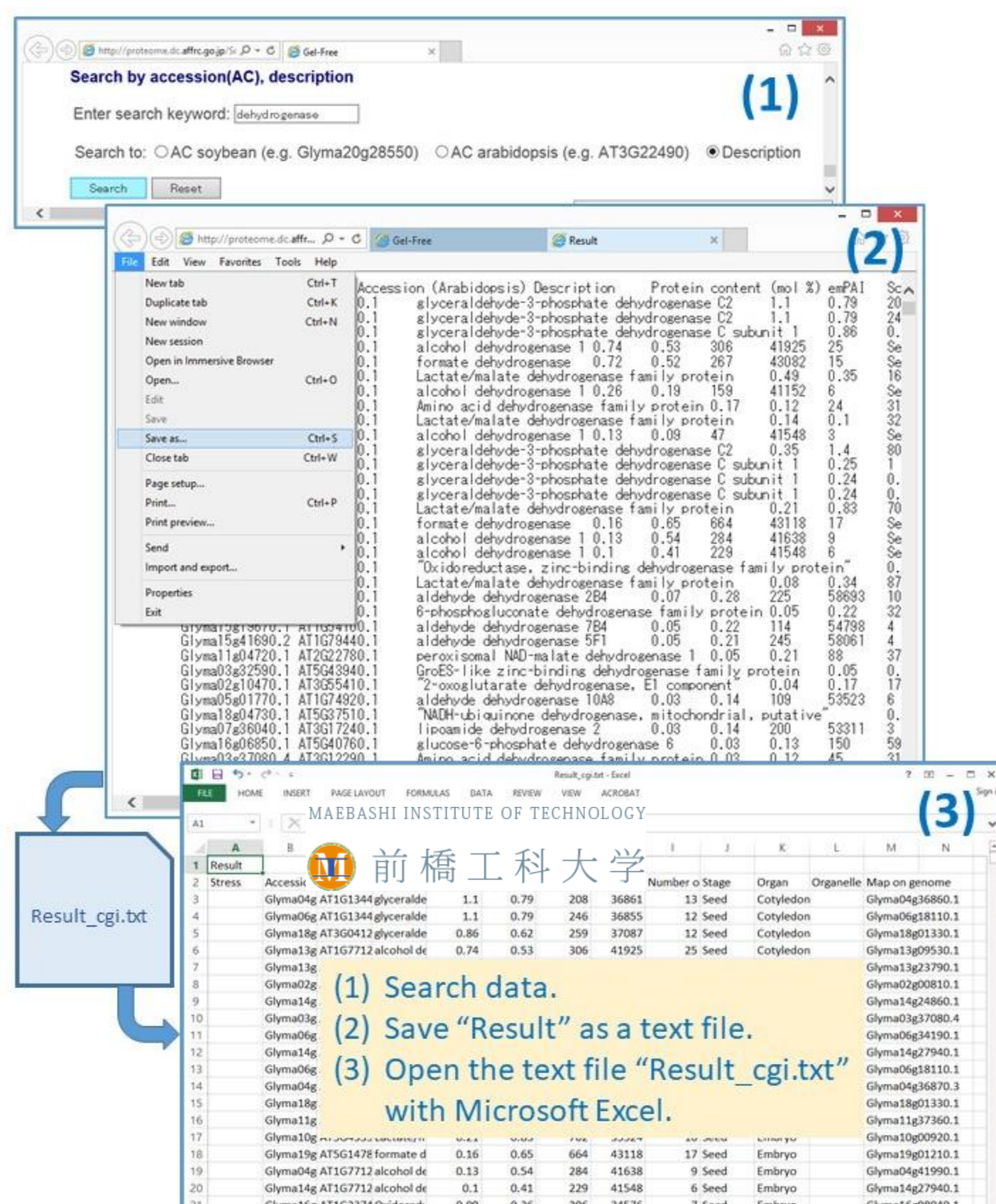


Fig. 2. How to use search result.
The explanation for "How to use search result" is in the bottom of first page of "Gel-free Proteomics Database for Soybean Proteins". (1) Search data, (2) Save "Result" as a text file, and then (3) Open the file "Result.cgi.txt" with Microsoft Excel.

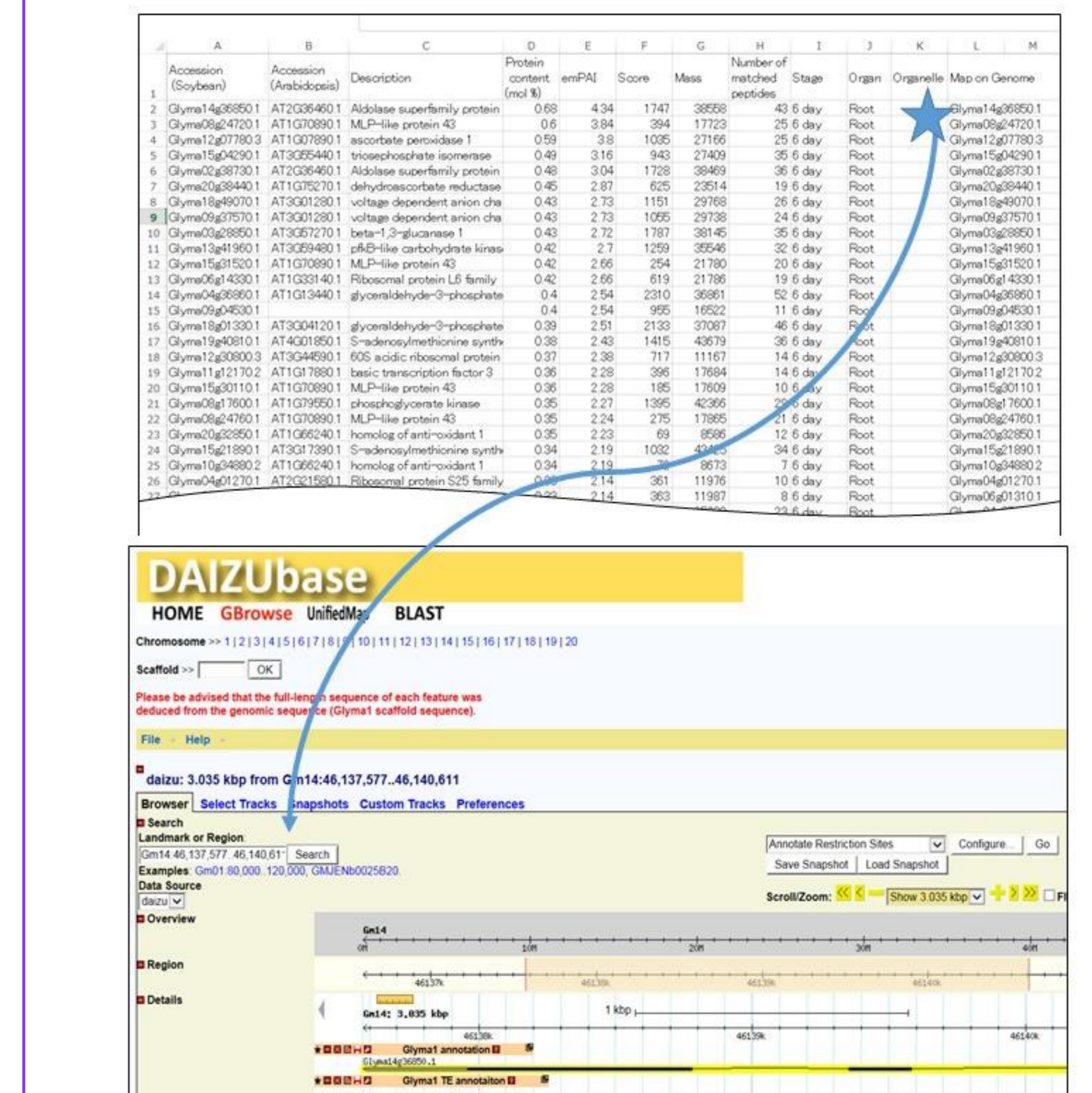


Fig. 3. Example of a search for soybean protein information in SPD.
When an image of a specific soybean plant or an organ or organelle is selected, a list of proteins is displayed. The data tables and search results are compatible with Microsoft Excel spreadsheet. The result file can be downloaded onto a personal computer, and dragged/dropped the file into a spreadsheet. When one of the proteins is selected, information on its chromosomal position from DAIZUbase (<http://daizu.dna.affrc.go.jp/>) is shown.

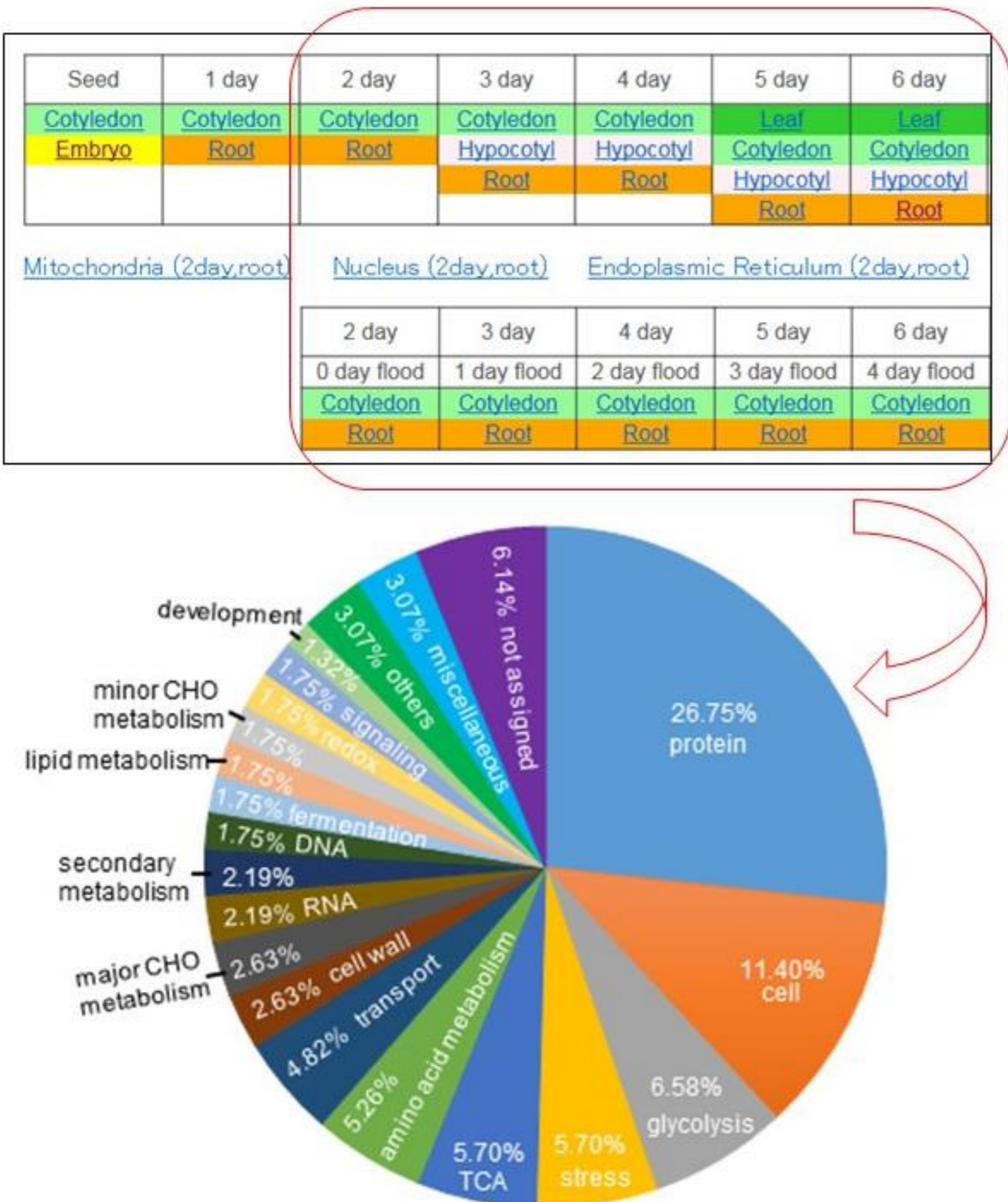


Fig. 4. Application of SPD for functional categorization of proteins in the soybean root identified in a time-dependent manner.
The Soybean Proteome Database stores data for organs of soybean plants grown under flooding stress with time course manner. Using these data, a set of 228 proteins significantly changed in abundance in the root under flooding stress, compared with those of non-treated soybean, were detected and analyzed using MapMan bin codes. The proteins used in this analysis are listed in Table in Ref 3.

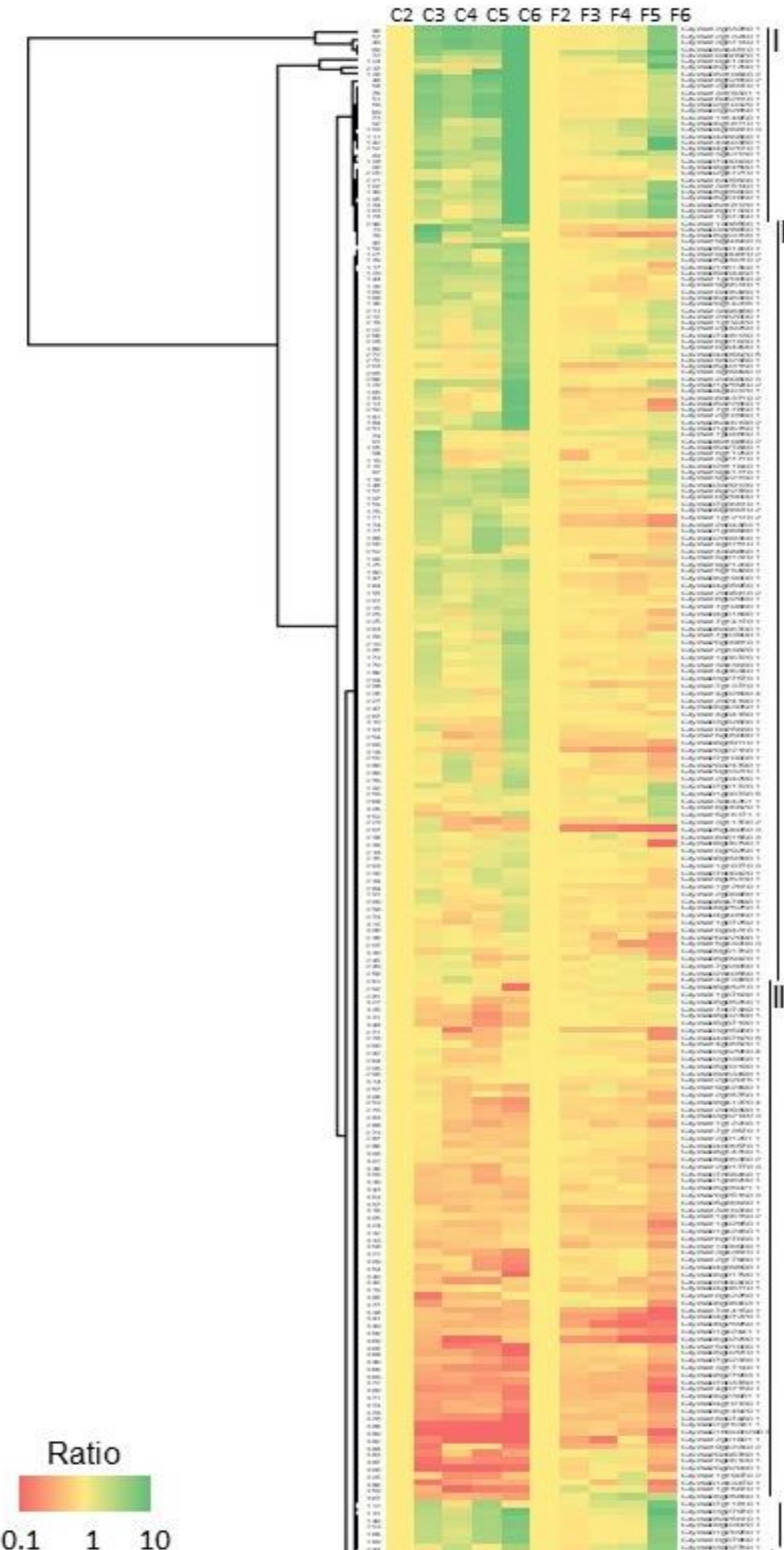


Fig. 5. Application of SPD for cluster analysis of significantly changed proteins in soybean during exposure to flooding stress.
The same set of 228 proteins used for functional categorization was subjected to cluster analysis using a centroid linkage method based on a Euclidean distance metric.

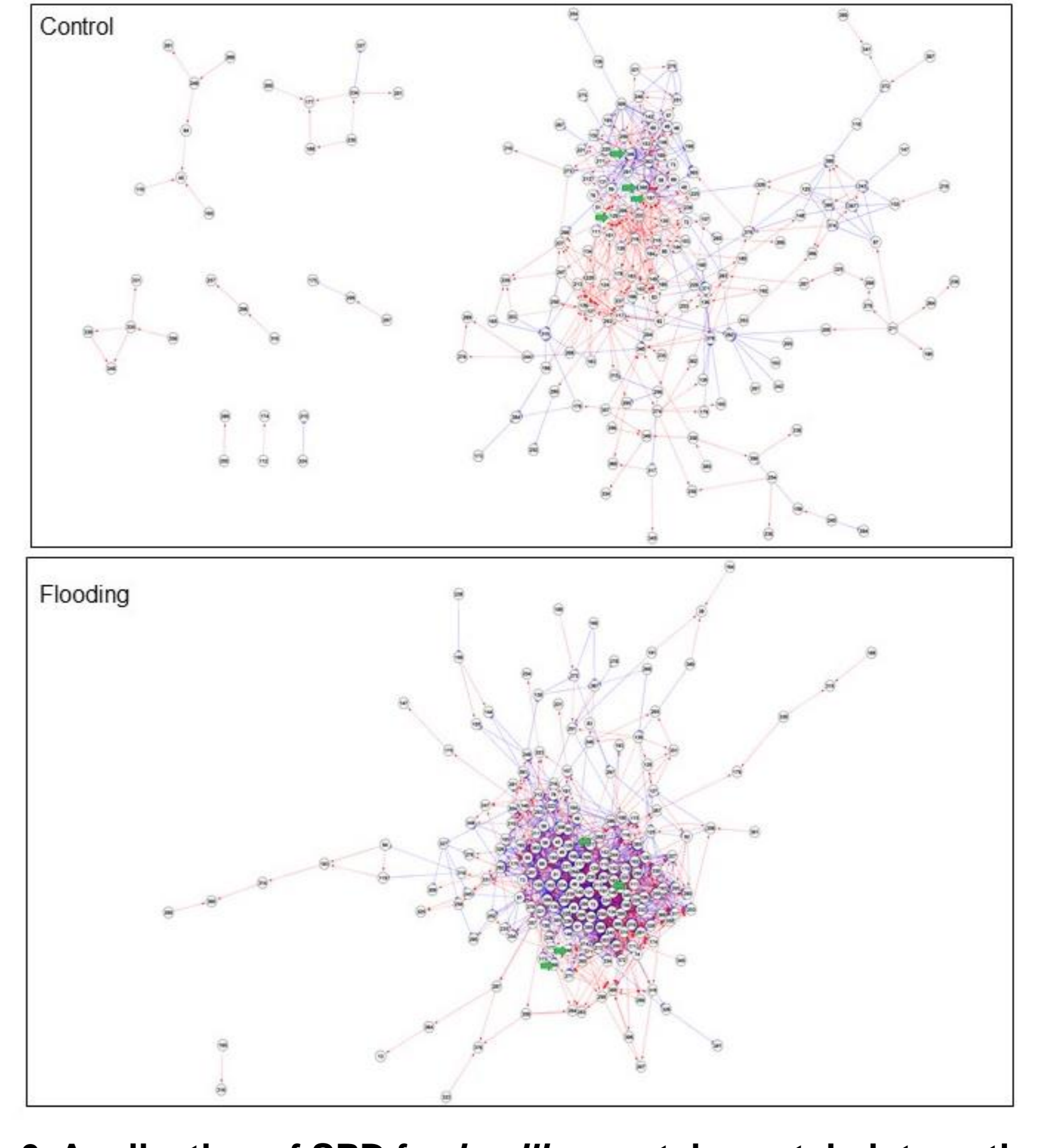


Fig. 6. Application of SPD for in silico protein-protein interaction analysis of proteins with significant changes in abundance.
A set of 228 proteins that changed in abundance under flooding stress compared with the starting point of treatment, which was 2-day-old soybean seedlings, were analyzed for in silico protein-protein interactions using the S-system differential equation.

Conclusion

1. SPD stores data obtained from both gel-based and gel-free proteomic techniques.
2. SPD stores different organs of soybean grown without or with stress in a time-dependent manner.
3. Organellar proteins identified in mitochondria, nuclei, and endoplasmic reticulum are stored
4. Based on functional analyses, these data are useful for analyses of biological mechanisms in soybean.

<http://proteome.dc.affrc.go.jp/Soybean/>

References

- 1) Sakata, K., Ohyang, H., Nobori, H., Nakamura, T., Hashiguchi, A., Nanjo, Y., Mikami, Y., Yunokawa, H., Komatsu, S. Soybean proteome database: a data resource for plant differential omics. *J. Proteome Res.* **8**, 3539-3548 (2009).
- 2) Ohyang, H., Sakata, K., Komatsu, S. Soybean Proteome Database 2012: update on the comprehensive data repository for soybean proteomics. *Front. Plant Sci.* **3**, 110 (2012).
- 3) Komatsu S, Wang X, Yin X, Nanjo Y, Ohyang H, Sakata K. Integration of gel-based and gel-free proteomic data for functional analysis of proteins through Soybean Proteome Database. *J Proteomics* 163:52-66 (2017).