SWATH-MS-Based Proteomics: Strategies and Applications in Plants

Article · October 2020

DOI: 10.1016/j.tibtech.2020.09.002

5 authors, including:

Mo-Xian Chen
The Chinese University of Hong Kong
102 PUBLICATIONS **801** CITATIONS
SEE PROFILE

Youjun Zhang
Max Planck Institute of Molecular Plant Physiology
37 PUBLICATIONS **506** CITATIONS
SEE PROFILE

Ying-Gao Liu
Shandong Agricultural University
36 PUBLICATIONS **1,061** CITATIONS
SEE PROFILE

Fu-Yuan Zhu
Nanjing Forestry University
35 PUBLICATIONS **459** CITATIONS
SEE PROFILE

Trends in Biotechnology

CellPress

Forum

SWATH-MS-Based Proteomics: Strategies and Applications in Plants

Mo-Xian Chen,[1,2,3] Youjun Zhang,[4,5] Alisdair R. Fernie,[4,5] Ying-Gao Liu,[6,*] and Fu-Yuan Zhu[1,2,*]

Most applied proteomic approaches require labeling steps. Recent technological advances provide an alternative label-free proteomics approach: SWATH-MS. This powerful tool is now widely used in animal studies but has drawn far less attention in plants. Here we summarize how this promising technology can be applied to facilitate functional analysis in plant research.

SWATH-MS (sequential window acquisition of all theoretical fragment ion spectra mass spectrometry) (Box 1) is an emerging profiling technique for quantifying proteome dynamics. The SWATH-MS approach is characterized by a data-independent acquisition (DIA) method followed by a novel targeted data extraction approach [1]. Briefly, complete fragment ion maps of the entire set of all peptide precursor ions without preselection (MS1 full scan) are obtained in SWATH mode. A series of 32 MS2 scans with 25 *m/z* isolation windows, recording the chromatographic elution traces of peak groups, are subsequently introduced into target analysis. This analysis uses MS2-based quantification methods to analyze the resulting fragment ion datasets according to their independent intensities, which provide continuous information for protein quantification (see Figure I in Box 1). Hence, the quantitative analyses of peptides in SWATH-MS are supported by extracted ion chromatograms at both the MS1 and MS2 levels. Furthermore, results identified by SWATH-MS have greater reproducibility and consistency, coupled with broader coverage and higher detection sensitivity of the target proteome, giving more target proteins or pathways for further functional validation [2]. These advantages of SWATH-MS result from retrospectively targeting fragmentation maps to monitor peptides of interest, as well as extendable spectra and virtual libraries. In combination, these properties contribute to a superior ability to carry out proteomic quantification in a single profiling experiment, especially with higher sensitivity for identifying low-abundance proteins. Additional strategies have been developed to accurately identify peptides with substantial post-translational modifications (PTMs) [3], including ubiquitination, glycosylation, phosphorylation, and methylation.

SWATH-MS has been used widely to study biological processes in animal systems [2], but emerging cases using this technology have only recently been reported in plants. We believe that this powerful technology should be introduced to the plant research community. To this end, we provide a brief introduction of this quantitative proteomic technique, give an overview of its general experimental and analytical pipeline, and provide suggestions for its future application in plant research.

Challenges and Recommendations for Implementing the SWATH-MS Protocol in Plant Research

Plants have always been a crucial resource for humans and other nonphotosynthetic organisms. Animal studies have been applying SWATH-MS-based proteomics for years [2], but the low accuracy and narrow identification range of traditional shotgun proteomics hindered the discovery of genuine regulators in plant functional studies. In addition, a number of technical difficulties must be overcome in order to transfer this technology to plant research (Figure 1A). For example, sample preparation is te initial step for all label-free-based proteomics, including SWATH-MS. In order to obtain deep, high-quality coverage of the proteome, proper extraction methods for various tissues from different plant species are required. Recently, a sodium deoxycholate method has been developed as a cheap and effective approach for plant total protein extraction followed by SWATH-MS detection [4]. This method had both higher reproducibility and detection efficiency for protein identification [4].

Besides protein preparation, instrumentation, data acquisition, and a downstream data-processing pipeline are also crucial factors in determining the quantification accuracy of SWATH-MS proteomics and need to be optimized for plant research (Figure 1B) [5]. By combining the advantages of high throughput and reproducibility from shotgun proteomics and targeted

Box 1. History, Principles, and Characteristics of SWATH-MS Proteomics

During the past few decades, mass spectrometry (MS)-based proteomics have become valuable tools for unraveling the functional significance of both prokaryotic and eukaryotic organisms. In a traditional discovery-based proteomics approach, proteins are quantified by two consecutive steps: a survey scan (MS1) and a second step (MS2) for the fragmentation of a selected precursor ion (Figure I). SWATH-MS is an effective label-free approach in which highly variable MS2 spectra are observed and quantified in a targeted manner according to an established assay spectrum library. Compared with other DIA approaches, such as MS^E, Shot-gun collision-induced dissociation, all-ion fragmentation, and Shotgun SWATH, which harbors multiple precursor isolation windows, enables the feasibility of spectral library-based targeted analysis of DIA data and reduces the complexity of MS/MS spectra as well as inferring ions. Therefore, it greatly enhances the number of identified proteins and improves the accuracy and precision of quantification. Compared with other label-based proteomics, SWATH-MS has greater throughput capacity and has the advantage of reducing cost, given the lack of labeling steps [2,4].

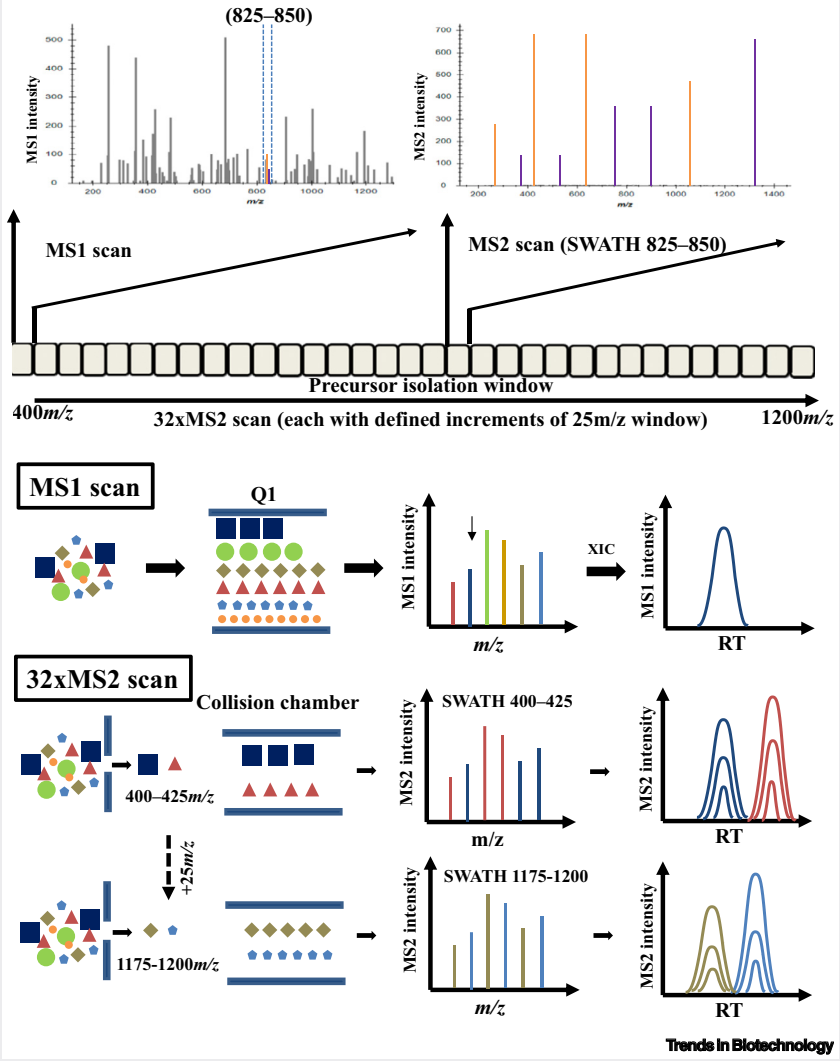


Figure I. Principle of SWATH-MS Technology. Generally, SWATH-MS measurement employs a TripleTOF or Orbitrap mass spectrometer, and SWATH data acquisition is typically conducted by a single precursor ion (MS1) scan coupled with product ion (MS2) scans of multiple precursor isolation windows. The full MS1 scan detects all precursor ions without preselection at the beginning, followed by 32 MS2 scans with defined increments of 25 *m/z* units, from 400 *m/z* to 1200 *m/z*. In this hypothetical example, two eluting peptide species from MS1 scan in the mass range of 825–850 *m/z* are detected (orange and purple), and the corresponding MS2 scan within a precursor isolation window of 825–850 *m/z* represents a complete fragment ion map of both peptide species. Abbreviations: RT, retention time; SWATH-MS, sequential window acquisition of all theoretical fragment ion spectra mass spectrometry; XIC, Extracted ion chromatogram.

quantitative proteomics, respectively, SWATH-MS can accurately quantify a large number of proteins within a single profiling experiment. In particular, the quantification accuracy is affected by parameters used in data acquisition, including the precursor mass range, width of isolation window, and MS2 accumulation and cycle time. In addition, software using different identification algorithms used for protein database searches, including PeakView, Spectronaut, and ProteinPilot [6,7], will affect the final identification and quantification of proteins. However, detailed manual comparisons of the output of these software packages remain necessary. Comprehensive identification with multiple software packages is recommended for high coverage in targeted proteomics experiments for plant studies.

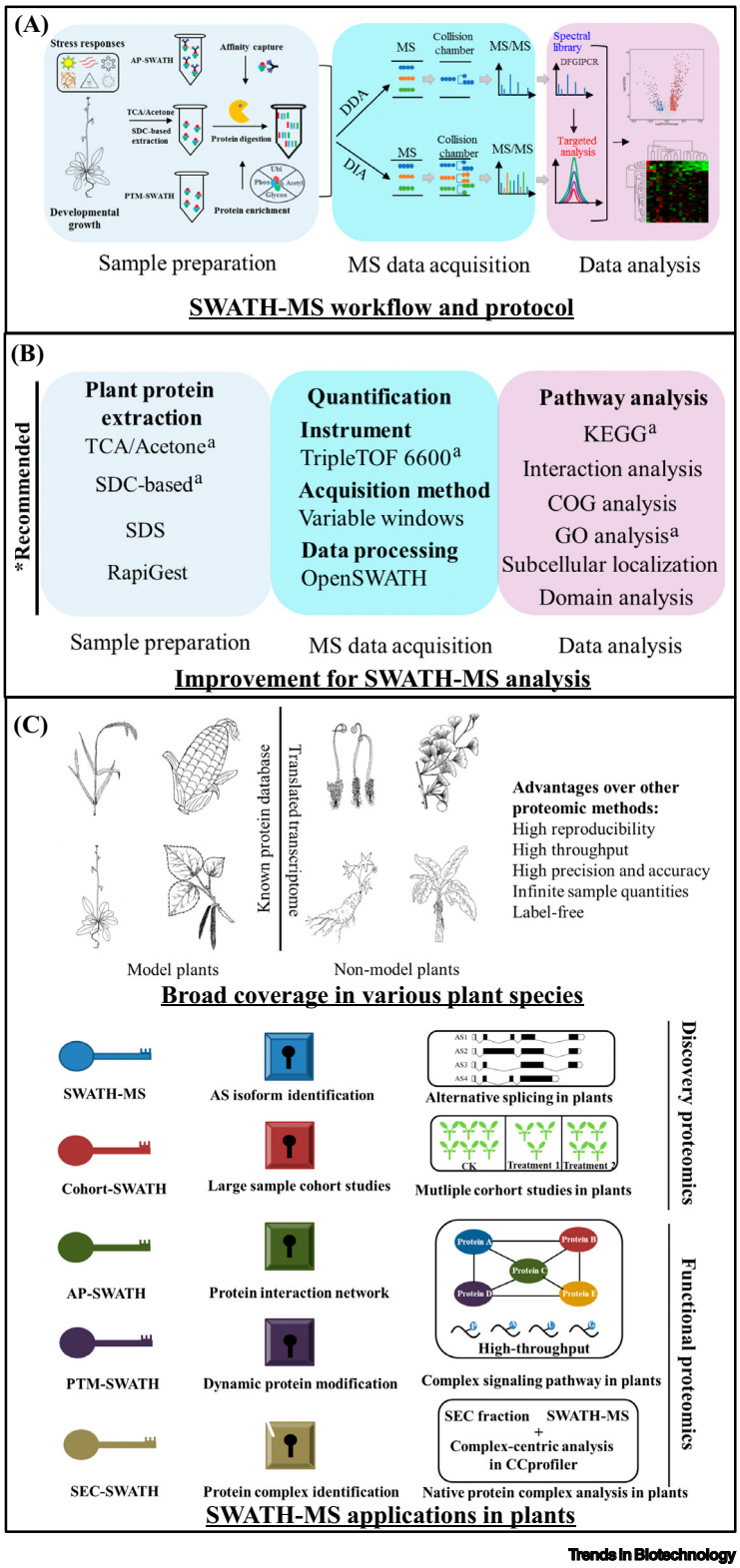
A third important issue is the downstream bioinformatic analytical pipeline (Figure 1B). Given the polyploid nature of some plant species, proper use of statistical methods to identify differentially abundant proteins (DAPs) and a well-annotated database to conduct pathway enrichment analysis are pivotal to evaluate the biological significance of changes in proteome dynamics during, for example, plant development or stress responses. The superior ability of SWATH-MS for DAP identification leads to a two- to threefold increase in identified DAPs [7,8]. Thus, this larger DAP dataset will substantially benefit subsequent pathway analysis and help to unravel novel pathways or regulators in plants [7,8]. Furthermore, conventional pathway enrichment analysis can be carried out using databases available online, such as Kyoto Encyclopedia of Genes and Genomes, Gene Ontology, OneOmics, and others. A manually curated database for a particular plant species or integration of several datasets will greatly enhance the identification of pivotal components from a DAP dataset and is therefore highly recommended at this step [9,10]. The discovery of novel regulators or pathways from the SWATH-MS approach is a key step in functionally facilitating the identification and subsequent *in planta* validation using genetic or biochemical methods.

Extending the Application of SWATH-MS Proteomics by Combining It with Biochemical Approaches

SWATH-MS has been used to characterize a number of model and non-model plant proteomes, including apple, *Arabidopsis*, rice, maize, barley, and others. Early studies using SWATH-MS-based proteomics

demonstrated novel pathways involving the lead detoxification process [7] or nitrogen starvation responses [8], providing examples of the use of quantitative proteomic data in the elucidation of biological function in plants. Furthermore, SWATH-MS-based proteomics approaches within particular subcellular compartments or with PTMs have revealed the spatial organization of the plant proteome and the dynamic change of PTMs during rice germination [11], respectively.

Given the advantages of this technique [12], we anticipate that SWATH-MS proteomics can be applied for a variety of large-scale profiling studies in plants, moving from model plant species to diverse plants without reference genome annotation (Figure 1C). SWATH-MS can be carried out on plant species lacking a genome sequence by using its transcriptome data to construct a virtual protein library. This proteogenomic approach further expands the use of SWATH-MS in plant research [9,10]. Furthermore, in the field of discovery proteomics, alternative splicing is an emerging research area related to post-transcriptional regulation. SWATH-MS could be applied to specifically identify peptides translated from splicing junctions [9], thus providing protein-based evidence for these splicing isoforms. In addition, large sample cohort studies can be facilitated using a SWATH-MS method because of its unlimited sample size within a particular analysis. Moreover, SWATH-MS proteomic data could be correlated with metabolomic datasets to study proteins with enzyme-like activities. Moreover, in the field of functional proteomics, protein interaction networks provide important evidence to reveal the molecular function of a target protein. Traditional screening methods such as yeast two-hybrid or coimmunoprecipitation coupled with MS identification suffer from a high false-positive rate and low throughput, respectively. However, affinity purification coupled SWATH-MS (AP-SWATH) can rapidly



(See figure legend at the bottom of the next page.)

quantify protein interaction dynamics at scale [13]. Notably, SWATH-MS-based proteomics could quantify PTM peptides when followed by an additional enrichment process of target PTM peptides. PTMs, including phosphorylation, *N*-glycosylation, acetylation, and succinylation, have been reported using proteomic approaches. However, SWATH-MS-based PTM identification is high in accuracy and coverage, enabling comparisons between nonmodified and modified peptides to reveal the proportion of modified proteins [14]. In addition, native protein complexes could be monitored in a modified SWATH-MS protocol by applying size exclusion chromatography [15], providing valuable information about these large protein complexes *in planta*.

Concluding Remarks and Future Perspectives

SWATH-MS proteomics was developed a decade ago and includes protein libraries and individual digitized sample records for proteomic investigation. However, this advanced technology has lagged behind the animal field; until recently [7], the SWATH-MS method had not been used widely in large-scale plant studies. This might be due to its dependence on specific MS instrumentation and skilled, specialized operators, unlike labeling-based proteomic approaches, which can be used with a variety of mass spectrometers. Furthermore, the dynamic range, sensitivity, and scan speed of a particular instrument further constrain the total number of samples that can be analyzed in SWATH-MS detection experiments. However, the robustness of SWATH-MS has been demonstrated recently by comparing several major labeling-based proteomic approaches that are currently used for plant proteome research [2] (Figure 1C). The systematic workflow uses a simplified pipeline to carry out such experiments for SWATH-MS in plants. It also emphasizes its versatile use in plant research hotspots as well as broad coverage across non-model plants due to its excellent reproducibility and accuracy.

Emerging approaches connected with SWATH-MS, such as AP- and PTM-SWATH, may benefit from protocol optimization to render them more suitable for plant research, thus identifying functional membrane or cell wall proteins and complex signaling pathways in plants. Some optimization strategies might include more efficient sample preparation reagents or advanced software analysis. At the same time, upgrades to current mass spectrometers or more powerful database search engine algorithms will enhance the performance of SWATH-MS in future plant research. We anticipate that a renewed focus on this technology as a powerful tool in plant research will reveal novel protein regulators and pathways that will represent useful leads for biotechnology and agriculture alike.

Acknowledgments

This work was supported by the by the National Natural Science Foundation of China (NSFC31701341), Natural Science Foundation of Jiangsu Province (SBK2020042924), the Nanjing Forestry University project funding (GXL2018005), Natural Science Foundation of Shandong Province (BS2015NY002), the Funds of Shandong 'Double Top' Program, and the National Natural Science Foundation of China (NSFC81401561 and 91535109). We thank Dr Chen Xi from SpecAlly Life Technology Co., Ltd., and Wuhan Institute of Biotechnology, Wuhan, China, and Dr Zhong Chuan-Qi from Xiamen University for their critical comments on this article. A.R.F. and Y.Z. thank the Max Planck Society and the EU Horizon 2020 Research and Innovation Programme, project PlantaSYST (SGA-CSA nos. 664621 and 739582 under FPA no. 664620) for supporting their research.

Author Contributions

M.X.C., Y.-G.L., and F.-Y.Z. designed and drafted the article. Y.-G.L., Y.Z., and A.R.F. critically reviewed and revised the manuscript.

[1]Co-Innovation Center for Sustainable Forestry in Southern China, College of Biology and the Environment, Nanjing Forestry University, Nanjing 210037, Jiangsu Province, China
[2]International Cultivar Registration for Osmanthus, College of Biology and the Environment, Nanjing Forestry University, Nanjing 210037, Jiangsu Province, China
[3]CAS Key Laboratory of Quantitative Engineering Biology, Shenzhen Institute of Synthetic Biology, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China
[4]Center for Plant Systems Biology and Biotechnology, 4000 Plovdiv, Bulgaria
[5]Max-Planck-Institut für Molekulare Pflanzenphysiologie, Am Mühlenberg 1, 14476 Potsdam-Golm, Germany
[6]State Key Laboratory of Crop Biology, College of Life Science, Shandong Agricultural University, Taian, Shandong, China

*Correspondence:
liuyg@sdau.edu.cn (Y.-G. Liu) and fyzhu@njfu.edu.cn (F.-Y. Zhu).

https://doi.org/10.1016/j.tibtech.2020.09.002

Figure 1. Strategies and Applications for SWATH-MS-Based Proteomics in Plant Research. (A) Schematic representation of SWATH-MS workflow and protocol, including its three major steps: sample preparation, mass spectrometric data acquisition, and data analysis. (B) Potential improvement of SWATH-MS data analysis and integration. (C) How combining other technologies with SWATH-MS-based proteomics could be used to study the processes of plant developmental growth or stress responses. Compared with other labeling-based proteomic techniques, no further peptide-labeling steps need to be performed. Due to mass window scanning, more than eight samples, with three to seven repeats of each sample, can be monitored at the same time in a single experiment. Abbreviations: AP, affinity purification; AS, alternative splicing; COG, Clusters of Orthologous Groups; DDA, data-dependent analysis; DIA, data-independent analysis; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; PTM post-translational modification; SDC, sodium deoxycholate; SDS, sodium dodecyl sulfate; SEC, size exclusion chromatography; SWATH-MS, sequential window acquisition of all theoretical fragment ion spectra mass spectrometry; TCA, trichloroacetic acid.

References

1. Gillet, L.C. *et al.* (2012) Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. *Mol. Cell. Proteomics* 11, O111.016717
2. Ludwig, C. *et al.* (2018) Data-independent acquisition-based SWATH-MS for quantitative proteomics: a tutorial. *Mol. Syst. Biol.* 14, e8126
3. Keller, A. *et al.* (2016) Opening a SWATH window on posttranslational modifications: automated pursuit of modified peptides. *Mol. Cell. Proteomics* 15, 1151–1163
4. Scheerlinck, E. *et al.* (2015) Minimizing technical variation during sample preparation prior to label-free quantitative mass spectrometry. *Anal. Biochem.* 490, 14–19
5. Wang, J. *et al.* (2015) MSPLIT-DIA: sensitive peptide identification for data-independent acquisition. *Nat. Methods* 12, 1106–1108
6. Li, S. *et al.* (2017) Optimization of acquisition and data-processing parameters for improved proteomic quantification by sequential window acquisition of all theoretical

fragment ion mass spectrometry. *J. Proteome Res.* 16, 738–747
7. Zhu, F. *et al.* (2016) SWATH-MS quantitative proteomic investigation reveals a role of jasmonic acid during lead response in *Arabidopsis*. *J. Proteome Res.* 15, 3528–3539
8. Zhu, F. *et al.* (2018) SWATH-MS quantitative proteomic investigation of nitrogen starvation in *Arabidopsis* reveals new aspects of plant nitrogen stress responses. *J. Proteome* 187, 161–170
9. Zhu, F.Y. *et al.* (2017) Proteogenomic analysis reveals alternative splicing and translation as part of the abscisic acid response in *Arabidopsis* seedlings. *Plant J.* 91, 518–533
10. Chen, M. *et al.* (2020) Full-length transcript-based proteogenomics of rice improves its genome and proteome annotation. *Plant Physiol.* 182, 1510–1526
11. Zhang, H. *et al.* (2016) Analysis of dynamic protein carbonylation in rice embryo during germination through AP-SWATH. *Proteomics* 16, 989–1000
12. Aebersold, R. and Mann, M. (2016) Mass-spectrometric exploration of proteome structure and function. *Nature* 537, 347–355
13. Caron, E. *et al.* (2017) Precise temporal profiling of signaling complexes in primary cells using SWATH mass spectrometry. *Cell Rep.* 18, 3219–3226
14. Schmidlin, T. *et al.* (2016) Assessment of SRM, MRM(3), and DIA for the targeted analysis of phosphorylation dynamics in non-small cell lung cancer. *Proteomics* 16, 2193–2205
15. Heusel, M. *et al.* (2019) Complex-centric proteome profiling by SEC-SWATH-MS. *Mol. Syst. Biol.* e8438, 15