Data in Brief 32 (2020) 106326

Contents lists available at ScienceDirect

Data in Brief

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

Data Article

Proteomics data of SNF1-related protein kinase 2.4 interacting proteins revealed by immunoprecipitation-mass spectrometry

Tong Zhang[a], Jacqueline D. Schneider[a], Chuwei Lin[a], Jin Koh[d], Sixue Chen[a,b,c,d,*]

[a] Department of Biology, University of Florida, Gainesville, FL 32610, USA
[b] Genetics Institute, University of Florida, Gainesville, FL 32610, USA
[c] Plant Molecular and Cellular Biology Program, University of Florida, Gainesville, FL 32610, USA
[d] Interdisciplinary Center for Biotechnology Research, University of Florida, Gainesville, FL 32610, USA

ARTICLE INFO

Article history:
Received 13 July 2020
Revised 5 September 2020
Accepted 16 September 2020
Available online 20 September 2020

Keywords:
Immunoprecipitation-mass spectrometry
Proteomics
SNF1-related protein kinase 2.4
Protein-protein interaction

ABSTRACT

Identification of kinase substrates is a prerequisite for elucidating the mechanism by which a kinase transduces internal or external stimuli to cellular responses. Conventional methods to profile this type of protein-protein interaction typically deal with one kinase-substrate pair at a time. Mass spectrometry-based proteomics, on the other hand, can determine putative kinase-substrate pairs at a large-scale in an unbiased manner. In this study, we identified the interacting partners of SNF1-related protein kinase 2.4 (SnRK2.4) via immunoprecipitation coupled with mass spectrometry. Proteins from stable transgenic Arabidopsis plants overexpressing a FLAG-tagged SnRK2.4 (cloned from *Brassica napus*) were pulled down using an anti-FLAG antibody. The protein components were then identified by mass spectrometry. In parallel, proteins from wild type plants were also analyzed, providing a list of nonspecific binding proteins that were further removed from the candidate SnRK2.4-interacting protein list. Our data identified over 30 putative SnRK2.4 interacting

* Corresponding author at: Department of Biology, Genetics Institute, Plant Molecular and Cellular Biology Program, Interdisciplinary Center for Biotechnology Research, University of Florida, Gainesville, FL 32610, USA.
E-mail address: schen@ufl.edu (S. Chen).

https://doi.org/10.1016/j.dib.2020.106326

partners, which included many key players in stress responses, transport, and cellular metabolic processes.

Specifications Table

Subject	Plant molecular biology
Specific subject area	Cellular proteomics
Type of data	Table, image, figure
How data were acquired	LC-MS/MS analysis, using an Orbitrap Fusion Tribrid Mass Spectrometer (Thermo Scientific)
Data format	Raw data and analyzed data
Parameters for data collection	Proteins from both the wild type and transgenic plants overexpressing a FLAG-tagged SnRK2.4 were immunoprecipitated with an anti-FLAG antibody. The resulting proteins were identified by LC-MS/MS.
Description of data collection	A data-dependent acquisition method was used for MS data collection
Data source location	Gainesville, FL, United States
Data accessibility	The data on putative SnRK2.4-interacting proteins are available with this article. Raw data files and search results are available through ProteomeXchange (http://www.proteomexchange.org/) with the identifier PXD017499.
Direct link to the dataset	https://www.ebi.ac.uk/pride/archive/projects/PXD017499
Related research article	Mengmeng Zhu; Tong Zhang; Wei Ji; Cecilia Silva-Sanchez; Wen-yuan Song; Sarah M. Assmann; Alice C. Harmon; Sixue Chen Redox regulation of a guard cell SNF1-related protein kinase in *Brassica napus*, an oilseed crop Biochemical Journal (2017), 474(15):2585–2599. doi: 10.1042/BCJ20170070.

Value of the Data

- This data provides putative interacting proteins of SnRK2.4, information of which is crucial toward understanding SnRK2.4-mediated signaling.
- The list of candidate proteins serves as a foundation for further detailed studies to validate protein-protein interactions between SnRK2.4 and other proteins.
- This data set provides the comprehensive characterization of the SnRK2.4 interactome, which can be used to compare with the interactomes of other SnRKs in plants.
- This dataset demonstrates the utility of immunoprecipitation-mass spectrometry (IP-MS) to screen potential interacting proteins in the context of kinase signaling networks.
- The method has broad applications in studying protein complexes and interaction networks in different organisms and biological context.

1. Data Description

The overall experimental approach consisted of the purification of SnRK2.4-interacting proteins by immunoprecipitation, followed by identification via mass spectrometry-based proteomics (Fig. 1). Arabidopsis plants overexpressing a FLAG-tagged SnRK2.4 were used, making it possible to pull down SnRK2.4 and its associated proteins by an anti-FLAG antibody. Processing of protein samples from the wild type (WT) plants in the same manner provided a robust way to eliminate nonspecific binding proteins (See Methods for detail). We define specific interacting proteins with our protein of interest as those that are only identified from the transgenic

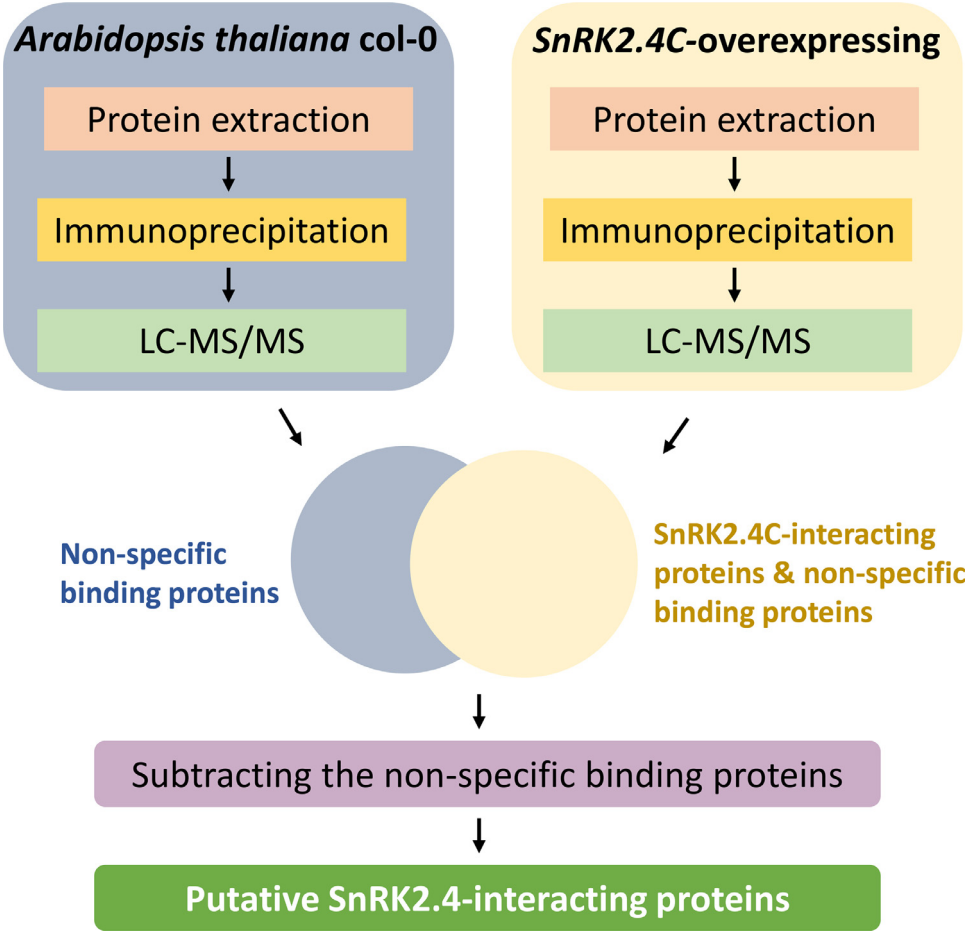

Fig. 1. IP-MS workflow. Note that samples from the wild type (top left panel) and transgenic plants overexpressing FLAG-tagged SnRK2.4 were processed in parallel. Following protein extraction, immunoreaction with anti-FLAG antibody, and LC-MS/MS analysis, the resulting proteins lists were compared. Nonspecific proteins identified from the wild type plants were subtracted, providing a clean list of SnRK2.4-interacting proteins.

plants, not from the WT. In this study, we identified 110 unique proteins from the SnRK2.4-overexpressing plants. Among them, 73 were also identified in the WT, which were considered as non-specific binding proteins. After removing these proteins, the remaining 37 proteins were further analyzed.

Quantification of protein intensities showed good correlation among the four independent IP-MS analyses of SnRK2.4-overexpressing plants (average Pearson correlation $r = 0.93$, Fig. 2). Table 1 lists putative SnRK2.4-interacting proteins, which are enriched primarily in stress responses such as response to abiotic stress and response to endogenous stimulus, as well as in transport and cellular metabolic processes (Fig. 3). This supports the notion that SnRK2 protein kinases are key regulators in stress responses [1,2]. Of note, we identified ARM repeat superfamily protein (AT4G36550, a protein known to mediate protein-protein interaction [3]) and heat shock protein 60-2 (AT2G33210) as heat-shock proteins had been shown to function downstream of SnRKs-mediated signaling [4,5]. Photosynthesis, another significantly over-represented biological process, is closely related to stress response. Thus, our data suggests a potential role

Table 1
Information of putative SnRK2.4 interacting proteins.

Accession	Protein description	Replicates (log_2 intensities)				Subcellular localization
AT3G47470	light-harvesting chlorophyll-protein complex I subunit A4	26.14	26.45	26.36	26.04	chloroplast/membrane/nucleus
AT5G12860	dicarboxylate transporter 1	24.32	24.13	24.6	24.43	chloroplast/membrane/mitochondrion
AT5G66760	succinate dehydrogenase 1–1	23.8	23.76	23.9	24.25	mitochondrion
AT4G36550	ARM repeat superfamily protein	26.09	26.36	26.08	25.73	nucleus/cytoplasma/mitochondrion
AT3G08940	light harvesting complex photosystem II	26.9	27.45	27.35	27.53	membrane/chloroplast/nucleus
AT5G54270	light-harvesting chlorophyll B-binding protein 3	25.37	25.73	26.07	26.04	membrane/chloroplast
AT5G25920	hypothetical protein	26.16	27.22	26.94	26.57	mitochondrion
AT1G61520	PSI type III chlorophyll a/b-binding protein	28.69	28.4	27.58	28.22	chloroplast/membrane/nucleus
AT5G53350	CLP protease regulatory subunit X	25.27	25.73	25.41	24.57	mitochondrion
AT3G63410	S-adenosyl-L-methionine-methyltransferase	23.26	23.17	24.07	24.11	chloroplast/membrane/cytosol
AT1G29930	chlorophyll A/B binding protein 1	24.07	23.83	24.5	23.18	chloroplast/membrane/nucleus
AT3G13300	transducin/WD40 repeat-like superfamily protein	27.03	27.09	28.36	27.43	nucleus/cytosol
AT1G31330	photosystem I subunit F	26.09	26.57	26.93	27.6	chloroplast/membrane/nucleus
AT1G60940	SNF1-related protein kinase 2.10	31.78	32.64	32.02	31.05	plasma membrane/nucleus/cytosol
AT4G38630	regulatory particle non-ATPase 10	24.18	22.74	22.9	22.75	membrane/nucleus/cytosol
ATCG00580	photosystem II reaction center protein E	23.31	24.1	23.46	22.23	membrane/chloroplast/nucleus
AT5G16030	mental retardation GTPase activating protein	25.29	25.21	25.15	23.64	nucleus
AT1G11910	aspartic proteinase A1	27.04	25.35	25.14	26.49	vacuole/plasmodesma/cytosol
AT3G18780	actin 2	27.18	26.28	27.11	25	cytosol/membrane/cytoskeleton/nucleus
AT2G05100	photosystem II light harvesting complex protein 2.1	25.48	26.26	25.41	27.74	membrane/chloroplast/nucleus
AT5G14040	phosphate transporter 3	23.6	23.82	23.27	25.71	chloroplast/mitochondrion/membrane
ATCG00130	ATPase F subunit	26.23	25.6	24.02	27.16	membrane/chloroplast/nucleus

(*continued on next page*)

Table 1 (*continued*)

Accession	Protein description	Replicates (log_2 intensities)				Subcellular localization
AT5G66880	sucrose nonfermenting 1(SNF1)-related protein kinase 2.3	26.08	28.2	25.87	24.7	nucleus/cytosol
AT4G33950	Protein kinase superfamily protein	24.6	26.56	26.09	23.29	nucleus/cytosol
AT3G01390	vacuolar membrane ATPase 10	19.69	19.77	18.38	22.18	Golgi apparatus/vacuole/cytosol
AT3G03530	non-specific phospholipase C4	25.85	23.61	26.49	27.36	plasma membrane/cytosol
AT2G46620	P-loop containing nucleoside triphosphate hydrolase	22.62	26.6	25.5	26.01	plasma membrane/cytosol
ATCG00720	photosynthetic electron transfer B	22.67	20.42	20.07	24.08	chloroplast/membrane/nucleus
AT2G28900	outer plastid envelope protein 16-1	23.83	20.13	21.83	24.45	chloroplast/vacuole/mitochondrion
AT3G09660	minichromosome maintenance 8	26.98	26.36	27.13	22.64	nucleus
AT4G11010	nucleoside diphosphate kinase 3	25.64	21.52	25.56	26.38	mitochondrial/chloroplast
AT5G46800	Mitochondrial substrate carrier family protein	19.83	23.4	24.83	24.02	cytosol/membrane
AT5G35200	ENTH/ANTH/VHS superfamily protein	20.24	24.98	23.71	25.39	plasma membrane/Golgi/mitochondrion
AT2G20260	photosystem I subunit E-2	21.28	20.68	21.22	25.96	chloroplast/membrane/cytosol
AT5G52090	P-loop containing nucleoside triphosphate hydrolase	24.4	19.6	18.68	22.32	nucleus/cytosol
AT2G33210	heat shock protein 60-2	18.38	22.85	23.42	25.17	mitochondrion/cytosol/chloroplast

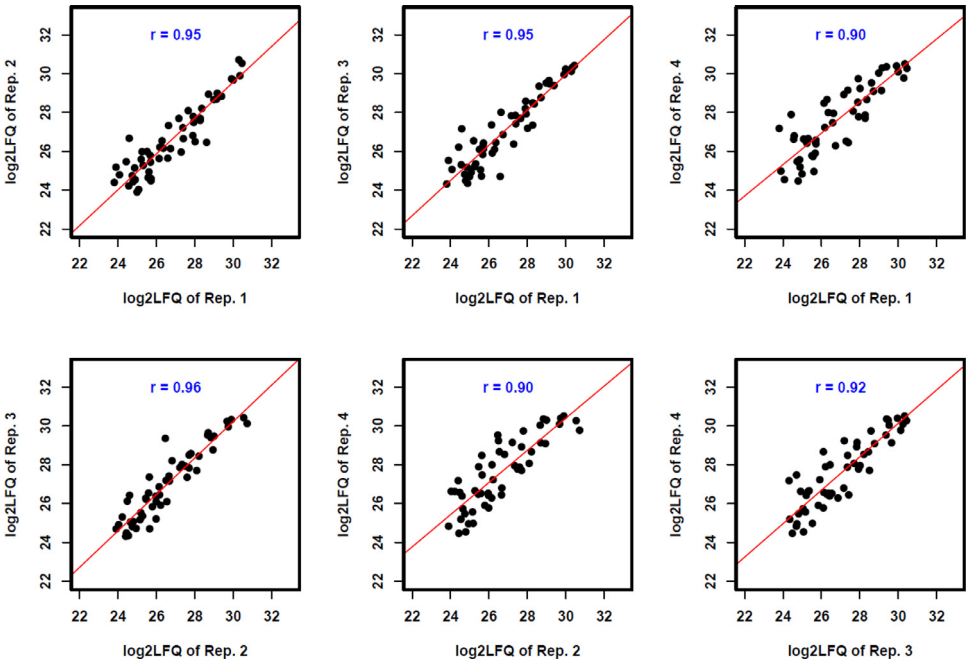

Fig. 2. Quantification metrics IP-MS. Correlation of relative protein abundance, expressed as log2 transformed label-free quantification (LFQ) intensities, from four biological replicates were shown. In each plot, the Pearson correlation value (r, shown in blue) and the best linear fit line (shown in red) were indicated. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

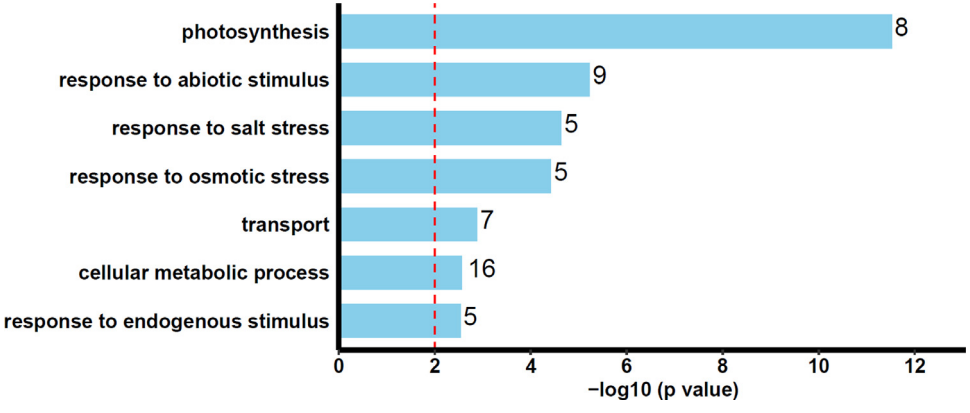

Fig. 3. Over-represented biological processes enriched from putative SnRK2.4-interacting proteins. Gene ontology of the protein lists was performed with agriGO, and statistically over-represented biological processes ($p < 0.01$, indicated by a vertical red line) were shown. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

of SnRK2.4 in balancing stress response and growth. In addition, most of the putative interacting proteins in Table 1 share subcellular localization with the SnRK2.4, which is found in nucleus, cytosol and mitochondria. Lastly, the coupling of genetic engineering and immunoprecipitation makes the approach described here easy to implement without the need to generate protein

specific antibody. Thus, it can be readily applied to study the interactome of other proteins to broaden the search for protein-protein interactions.

2. Experimental Design, Materials and Methods

2.1. Plant materials

Both wild type (WT) *Arabidopsis thaliana* col-0 (obtained from Arabidopsis Biological Resource Center) and transgenic plants overexpressing a *Brassica napus SnRK2.4-1C* were used in this study. The transgenic plants were generated by transferring a FLAG-tagged *SnRK2.4-1C* into the WT *A. thaliana* as described previously [6,7]. Seeds were soaked in 50% bleach for disinfection and were then sown on half-strength Murashige and Skoog medium supplemented with $1 \times$ MS vitamins (Caisson, UT). Following stratification at 4 °C for 2 days in the dark, seeds were allowed to germinate and grow in a growth chamber at 22/20 °C with 8/16 h light cycle (day/night). Two-week-old seedlings were used for proteomics experiments.

2.2. Immunoprecipitation and proteomic sample preparation

Seedlings were pulled off from the agar plate with tweezers and immediately frozen in liquid nitrogen. Samples were then ground into fine power with a mortar and pestle in the presence of liquid nitrogen. Next, proteins were extracted in 50 mM Tris–HCl, pH 7.4, 150 mM NaCl, 1 mM EDTA, 1% Triton X-100 supplemented with protease and phosphatase inhibitor cocktails (Thermo Scientific, IL). Insoluble cellular debris were removed by centrifugation at 10,000 g for 10 min, and proteins in the supernatant were quantified with the Bradford method assay.

Immunoprecipitation and the subsequent sample preparation for proteomics were performed according to previous descriptions [8]. Briefly, 10 mg of proteins were incubated with 300 µL of anti-FLAG M2 resin (Sigma, MO) for 2 h at 4 °C with gentle agitation. Nonspecific binding proteins were removed by washing the resin three times with 50 mM Tris–HCl and 150 mM NaCl (pH 7.4). Subsequently, FLAG-tagged SnRK2.4 and its associated proteins were eluted with 150 ng/μL 3 X FLAG peptide (Sigma, MO). The eluted samples were cleaned by a short-run SDS-PAGE (30 min at 100 V), which effectively removed salts and the 3 X FLAG peptide that would otherwise compromise protein identification. Protein bands of interest were cut from the gel and were then digested with trypsin. Specifically, three fractions were collected for each sample: the major band corresponding to SnRK2.4, proteins above this band, and proteins below this band. The resulting peptides from each fraction were further desalted by Ziptip prior to liquid chromatography mass spectrometry analysis. Four biological replicates were conducted from both the WT and the transgenic plants, thus yielding 24 samples for proteomic identification (8 samples X 3 fractions/sample = 24).

2.3. Liquid chromatography mass spectrometry

On-line peptide separation and mass spectrometry data acquisition were performed on an EASY-nLC 1000 (Thermo Scientific, CA) coupled with an Orbitrap Fusion Tribrid (Thermo Scientific, Germany) as described previously [9]. Briefly, peptides were eluted with a reverse phase C18 column at 350 nL/min with a linear increase of buffer B (0.1% formic acid in acetonitrile) from 2 to 30% in 60 min. The eluted peptides were subjected to a data-dependent acquisition scheme where a full MS scan was followed by MS/MS scans of the most intense precursor ions within a cycle of 2 s. Full MS scans were performed in the Orbitrap at 350–1800 m/z at a resolution of 120,000 with automatic gain control (AGC) of 4E5 and maximum ion injection time of 50 ms. Most intense precursor ions were selected by the quadrupole at an isolation window

1.3 m/z, and then fragmented by higher-energy collisional dissociation at collision energy of 35. Full MS/MS scans were conducted in the Orbitrap at AGC of 1E4 and maximum ion injection time of 35 ms.

2.4. Data analysis

Analysis of IP-MS data was performed as previously described [8,10]. All raw data were searched with MaxQuant [11,12] against the TAIR10 protein database (35,386 entries, downloaded on October 2017). Gel pieces from the same biological replicates were assigned as fractions. Key searching parameters included: trypsin as the enzyme for digestion with a maximum missed cleavage sites of 2; Carbamidomethyl on cysteine residues (+ 57.021464 Da) as a fixed modification; oxidation of methionine (+ 15.994915 Da) and acetylation of protein N termini (+ 42.010565 Da) as dynamic modifications; mass tolerance for MS1 and MS/MS spectra being 4.5 ppm and 0.5 Da, respectively. A 1% false discovery rate was applied for both peptide and protein identification. In addition, at least two unique peptides were required for confident protein identification. The match between run and label-free quantification features [11] was also enabled.

Since four biological replicates were used here, a protein is defined as identified in either the WT or the transgenic plants if it is quantified in at least two biological replicates (i.e., missing values <=2). SnRK2.4-interacting proteins were defined as these that are uniquely identified in the transgenic plants but not in the WT controls. The quantification correlation among biological replicates was performed with log2 transformed label-free quantification (LFQ) protein intensities using Pearson correlation in the base R [13]. Gene ontology of the putative SnRK2.4-interacting proteins was performed using agriGO [14] with the results in *Supplemental Table 1*. The R script code for plotting can be found in *Supplemental code*.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper

Acknowledgments

This research was supported by grants from the National Science Foundation (MCB 0818051 to S. Chen). Ning Zhu and Dr. Joe Collins at University of Florida were acknowledged for technical assistance and invaluable discussions, respectively.

Supplementary Materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.dib.2020.106326.

References

[1] A. Kulik, I. Wawer, E. Krzywińska, M. Bucholc, G. Dobrowolska, SnRK2 protein kinases–key regulators of plant response to abiotic stresses, OMICS 15 (2011) 859–872.
[2] T. Zhang, S. Chen, A.C. Harmon, Protein phosphorylation in stomatal movement, Plant Signal Behav. 9 (2014) e972845.
[3] Y. Mudgil, S.-H. Shiu, S.L. Stone, J.N. Salt, D.R. Goring, A large complement of the predicted Arabidopsis ARM repeat proteins are members of the U-box E3 ubiquitin ligase family, Plant Physiol. 134 (2004) 59–66.

[4] M. Clément, N. Leonhardt, M.-J. Droillard, I. Reiter, J.-L. Montillet, B. Genty, C. Laurière, L. Nussaume, L.D. Noël, The cytosolic/nuclear HSC70 and HSP90 molecular chaperones are important for stomatal closure and modulate abscisic acid-dependent physiological responses in Arabidopsis, Plant Physiol. 156 (2011) 1481–1492.
[5] P. Jacob, H. Hirt, A. Bendahmane, The heat-shock protein/chaperone network and multiple stress resistance, Plant Biotechnol. J. 15 (2017) 405–414.
[6] M. Zhu, T. Zhang, W. Ji, C. Silva-Sanchez, W.Y. Song, S.M. Assmann, A.C. Harmon, S. Chen, Redox regulation of a guard cell SNF1-related protein kinase in Brassica napus, an oilseed crop, Biochem. J. 474 (2017) 2585–2599.
[7] T. Ma, M.J. Yoo, T. Zhang, L. Liu, J. Koh, W.Y. Song, A.C. Harmon, W. Sha, S. Chen, Characterization of thiol-based redox modifications of Brassica napusSNF1-related protein kinase 2.6-2C, FEBS Open Bio 8 (2018) 628–645.
[8] T. Zhang, J.D. Schneider, C. Lin, S. Geng, T. Ma, S.R. Lawrence, C.P. Dufresne, A.C. Harmon, S. Chen, MPK4 phosphorylation dynamics and interacting proteins in plant immunity, J. Proteome Res. 18 (2019) 826–840.
[9] T. Zhang, S. Chhajed, J.D. Schneider, G. Feng, W.-.Y. Song, S. Chen, Proteomic characterization of MPK4 signaling network and putative substrates, Plant Mol. Biol. 101 (2019) 325–339.
[10] T. Zhang, J.D. Schneider, N. Zhu, S. Chen, Identification of MAPK substrates using quantitative phosphoproteomics, Methods Mol. Biol. 1578 (2017) 133–142.
[11] J. Cox, M.Y. Hein, C.A. Luber, I. Paron, N. Nagaraj, M. Mann, Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ, Mol. Cell Proteom. 13 (2014) 2513–2526.
[12] S. Tyanova, T. Temu, J. Cox, The MaxQuant computational platform for mass spectrometry-based shotgun proteomics, Nat. Protoc. 11 (2016) 2301–2319.
[13] R.C. TeamR: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, 2017.
[14] T. Tian, Y. Liu, H. Yan, Q. You, X. Yi, Z. Du, W. Xu, Z. Su, AgriGO v2.0: a GO analysis toolkit for the agricultural community, 2017 update, Nucl. Acids Res. 45 (2017) W122–W129.