An up-to-date workflow for plant (phospho)proteomics identifies differential drought-responsive phosphorylation events in maize leaves

Elisabeth Stes[a,b,c,d,#], Lam Dai Vu[a,b,c,d,#], Michiel Van Bel[a,b], Hilde Nelissen[a,b], Dirk Inzé[a,b], Frederik Coppens[a,b], Kris Gevaert[c,d,*], Ive De Smet[a,b,*,†]

[a]Department of Plant Systems Biology, VIB, 9052 Ghent, Belgium

[b]Department of Plant Biotechnology and Bioinformatics, Ghent University, 9052 Ghent, Belgium

[c]Medical Biotechnology Center, VIB, 9000 Ghent, Belgium

[d]Department of Biochemistry, Ghent University, 9000 Ghent, Belgium

ABSTRACT

Protein phosphorylation is one of the most common post-translational modifications (PTMs), which can regulate protein activity and localization, as well as protein–protein interactions in numerous cellular processes. Phosphopeptide enrichment techniques enabled plant researchers to acquire insight in phosphorylation-controlled signaling networks in various plant species. Most phosphoproteome analyses of plant samples still involve stable isotope labeling, peptide fractionation, and demand lots of mass spectrometry (MS) time. Here, we present a simple workflow to probe, map and catalogue plant phosphoproteomes, requiring relatively low amounts of starting material, no labeling, no fractionation, and no excessive analysis time. Following optimization of the different experimental steps on *Arabidopsis thaliana* samples, we transferred our workflow to maize, a major monocot crop, to study signaling upon drought stress. In addition, we included normalization to protein abundance to identify true phosphorylation changes. Overall, we identified a set of new phosphosites in both *Arabidopsis thaliana* and maize, some of which are differentially phosphorylated upon drought. All data are available via ProteomeXchange with identifier PXD003634, but to provide easy access of the whole scientific community to our model plant and crop datasets, we created an online database, Plant PTM Viewer (bioinformatics.psb.ugent.be/webtools/ptm_viewer/), where all phosphosites identified in our study can be consulted.

Key Words: Phosphoproteomics, maize, Arabidopsis, drought stress, database

INTRODUCTION

The balanced action of protein kinases and phosphatases determines a proteome's phosphorylation status. Protein phosphorylation may transiently modify protein properties, such as enzymatic activity, subcellular localization, protein structure and stability, and interactions with other proteins. As such, many cellular signaling processes, such as transmembrane signaling, intracellular amplification of signals and cell cycle control, occur via reversible protein phosphorylation (1). In plants, phosphorylation-mediated signaling is of central importance in various physiological processes, including hormone signaling and stress responses (2). However, only a limited number of kinases and phosphatases (and their targets) have been studied in different levels of detail (3).

Mass spectrometry (MS)-based proteomics became an essential tool for studying protein phosphorylation and has enabled the identification of numerous phosphorylation sites on plant proteins (3). Nevertheless, studies of phosphorylation events remain challenging, due to their dynamic nature and the sub-stoichiometric levels of phosphorylated proteins. Therefore, at least some level of enrichment for phosphorylation sites is needed and this is best done at the peptide level to maximize the identification of phosphosites. The most productive approach is based on metal (ion) chelation. By exploiting the interaction between negatively charged phosphate groups and positively charged metal ions or metal oxides, immobilized metal affinity chromatography (IMAC) and metal oxide affinity chromatography (MOAC) methods, respectively, represent efficient ways to enrich phosphopeptides from complex mixtures. Enrichment with TiO_2 beads became a routine method in plant proteomics studies in recent years (4-12). In an attempt to maximally cover phosphoproteomes, the majority of phosphoproteomics approaches make use of peptide fractionation methods, such as strong cation exchange chromatography, hydrophilic interaction chromatography or

reversed-phase chromatography (7, 13-17). These however result in far more LC-MS/MS measurement time per sample to be analyzed and also require large(r) amounts of starting material.

With the ever-increasing number of phosphorylation sites being identified – for nearly every human cellular protein a phosphosite has been reported (18) – the functionality of these post-translational modifications (PTMs) is questioned. Crowdedness effects are hypothesized to give rise to non-functional transfer of a phosphate group by kinases upon encounter of a random protein (19, 20). Merely profiling phosphorylation sites will hence likely lead to the large scale identification of nonfunctional PTMs. To discriminate these 'noisy' phosphosites from sites with regulatory significance, experiments where differential conditions are compared are vital. Obviously, this requires assessing dynamics in the phosphoproteome via quantitative methods. Methodologies for the quantitative analysis of phosphoproteomes in plants are most frequently based on stable isotope labeling, like $^{15}N/^{14}N$ metabolic labeling of proteins during plant growth (4, 11, 16, 21, 22) or post-metabolic labeling of peptides with iTRAQ (9, 12, 23, 24). As labeling imposes limitations on the number of conditions that can be monitored, label-free methods represent a practical alternative. Two label-free methods, spectral counting and precursor ion intensity-based quantification, have been applied in plant phosphoproteome strategies (8, 10, 11, 14, 25-27). However, label-free approaches often suffer from quantitative incompleteness due to stochastic data acquisition (MS/MS sequencing) leading to numerous missing values in the dataset, which – to some extent – can be avoided by matching data between LC-MS(/MS) runs.

Although missing in most published plant phosphoproteome studies [some exceptions are (12, 15, 16)], parallel and in depth investigation of the overall proteome is recommended for normalization of quantitative PTM studies. To determine if phosphopeptide changes are the result of true phosphorylation changes or rather general abundance changes of the

phosphoprotein, phosphopeptide levels need to be normalized to overall protein abundances. Ideally, such changes in overall protein levels should be derived from an analysis of non-phosphopeptides of the same sample (28).

Agricultural plants, such as maize, routinely face drought stress, which is one of the worst environmental hazards that impacts crop productivity (29, 30). Some crop cultivars are known to better withstand abiotic stress, but these responses are dynamic and complex and often a genetic basis is hard to find (31). As plants remodel their proteome in response to stress, drought-adaptive traits are likely to be reflected at the proteome level (32). Moreover, as a universal biochemical signal in cells, protein phosphorylation controls stress responses, transmitting stress signals from the cell surface to the nucleus (33).

Taken together, commonly used strategies in plant phosphoproteomics involve tedious labeling approaches and fractionation steps, which are time consuming, expertise demanding and negatively affect reproducibility, robustness and throughput. Here, we present a label-free quantitative workflow for quick and reproducible phosphoproteome analysis of plant tissue, requiring only small sample amounts and no costly expert software for data analysis and integrating steps (such as normalization) that are not yet standard in plant phosphoproteomics. We applied our workflow to maize, a major monocot crop, to study signaling upon drought stress, and we identified a set of new phosphosites in maize, some of which are differentially phosphorylated upon drought.

EXPERIMENTAL SECTION

Plant growth

Seedlings of *A. thaliana* (ecotype Columbia) were grown on vertically-held plates with half-strength Murashige and Skoog medium solidified with 0.8% agar at 22°C in continuous light.

Four days post germination (dpg), the plants were transferred to 10 μM 1-naphthaleneacetic acid (NAA)-containing plates. Roots were harvested 5 dpg. Maize plants (inbred line B104) were grown in soil in a growth chamber with controlled relative humidity (55%) and temperature (24°C), in a 16h/8h (day/night) cycle. Drought was induced by lowering the soil water capacity to 62.5% relative to that of the well-watered control plants. 21 days after sowing, the first 4 cm of growing leaf 7 was harvested.

Protein Extraction and Tryptic Digestion

Plant material was harvested in three biological replicates. One g of fresh weight material was flash-frozen in liquid nitrogen, and manually ground into a fine powder with a pestle and mortar. Proteins were extracted in homogenization buffer containing 50 mM Tris-HCl buffer (pH 8), 0.1 M KCl, 30% sucrose, 5 mM EDTA, and 1 mM DTT in milliQ water, and the appropriate amounts of the Complete protease inhibitor mixture and the PhosSTOP phosphatase inhibitor mixture (both from Roche) were added. The samples were sonicated on ice and centrifuged at 4°C for 15 min at 2,500×*g* to remove debris. Supernatants were collected and a methanol/chloroform precipitation was carried out by adding 3, 1 and 4 volumes of methanol, chloroform and water, respectively. Samples were centrifuged for 10 min at 5,000×*g*, and the aqueous phase was removed. After addition of 4 volumes methanol, the proteins were pelleted via centrifugation for 10 min at 2,500×*g*. Pellets were washed with 80% acetone and re-suspended in 6 M guanidinium hydrochloride in 50 mM triethylammonium bicarbonate (TEAB) buffer (pH 8). Alkylation of cysteines was carried out by adding a combination of tris(carboxyethyl)phosphine (TCEP, Pierce) and iodoacetamide (Sigma-Aldrich) to final concentrations of 15 mM and 30 mM respectively, and the reaction was allowed for 15 min at 30°C in the dark. Before digestion, the samples were buffer exchanged on Illustra NAP columns (GE Healthcare Life Sciences) to 50 mM TEAB buffer

(pH 8) and the protein concentration was measured using the Bio-Rad Protein Assay. One mg of the proteins was pre-digested with EndoLysC (Wako Chemicals) for 4 h, followed by a digestion with trypsin overnight (Promega Trypsin Gold, mass spectrometry grade), both digestions occurring at 37°C at an enzyme-to-substrate ratio of 1:100 (w:w). The digest was acidified to pH ≤ 3 with trifluoroacetic acid (TFA) and desalted with SampliQ C18 SPE cartridges (Agilent) according to the manufacturer's guidelines. The eluates were split into two and dried in a vacuum centrifuge. One half of the samples served for proteome analyses and were re-dissolved in 30 µL of 2% (v/v) acetonitrile and 0.1% (v/v) TFA right before LC-MS/MS analysis.

Phosphopeptide Enrichment

The dried eluates were resuspended in 100 µl of loading solvent (80% acetonitrile, 5% TFA) and incubated with 1 mg MagReSyn® Ti-IMAC microspheres for 20 min at room temperature. The microspheres were next washed once with wash solvent 1 (80% acetonitrile, 1% TFA, 200 mM NaCl) and two times with wash solvent 2 (80% acetonitrile, 1% TFA). The bound phosphopeptides were eluted with three volumes (80 µl) of a 1% NH_4OH solution, followed immediately by acidification to pH ≤ 3 with formic acid. Prior to MS analysis, the samples were vacuum-dried and re-dissolved in 50 µL of 2% (v/v) acetonitrile and 0.1% (v/v) TFA.

Mass Spectrometry

Each sample was analyzed twice (i.e. in two technical replicates) via LC−MS/MS on an Ultimate 3000 RSLC nano LC (Thermo Fisher Scientific) in-line connected to a Q Exactive mass spectrometer (Thermo Fisher Scientific). The sample mixture was first loaded on a trapping column (made in-house, 100 µm internal diameter (I.D.) × 20 mm, 5 µm beads C18

Reprosil-HD, Dr. Maisch, Ammerbuch-Entringen, Germany). After flushing from the trapping column, the sample was loaded on an analytical column (made in-house, 75 µm I.D. × 150 mm, 3 µm beads C18 Reprosil-HD, Dr. Maisch). Peptides were loaded with loading solvent A (0.1% TFA in water) and separated with a linear gradient from 98% solvent A' (0.1% formic acid in water) to 55% solvent B′ (0.1% formic acid in water/acetonitrile, 20/80 (v/v)) in 170 min at a flow rate of 300 nL/min. This was followed by a 5 min wash reaching 99% solvent B'. The mass spectrometer was operated in data-dependent, positive ionization mode, automatically switching between MS and MS/MS acquisition for the 10 most abundant peaks in a given MS spectrum. The source voltage was 3.4 kV, and the capillary temperature was 275°C. One MS1 scan (m/z 400−2000, AGC target 3×10^6 ions, maximum ion injection time 80 ms) acquired at a resolution of 70000 (at 200 m/z) was followed by up to 10 tandem MS scans (resolution 17500 at 200 m/z) of the most intense ions fulfilling predefined selection criteria (AGC target 5×10^4 ions, maximum ion injection time 60 ms, isolation window 2 Da, fixed first mass 140 m/z, spectrum data type: centroid, underfill ratio 2%, intensity threshold 1.7xE4, exclusion of unassigned, 1, 5-8, >8 charged precursors, peptide match preferred, exclude isotopes on, dynamic exclusion time 20 s). The HCD collision energy was set to 25% Normalized Collision Energy and the polydimethylcyclosiloxane background ion at 445.120025 Da was used for internal calibration (lock mass).

Data Analysis

For the Arabidopsis samples, MS/MS spectra were searched against a Uniprot database containing *A. thaliana* sequences (34,509 entries, version November, 2014) with the MaxQuant software (version 1.5.3.8). For the maize samples, the searches were done against a *Zea mays* database downloaded from PLAZA Monocots 3.0 (34) containing sequences (39,305 entries, version 2014) with the MaxQuant software (version 1.5.0.30). For all

searches, a precursor mass tolerance was set to 20 ppm for the first search (used for nonlinear mass re-calibration) and set to 4.5 ppm for the main search. Trypsin was selected as enzyme setting. Cleavages between lysine/arginine-proline residues and up to two missed cleavages were allowed. Carbamidomethylation of cysteine residues was selected as a fixed modification and oxidation of methionine residues was selected as variable modification. For the samples enriched for phosphopeptides phosphorylation of serine, threonine and tyrosine residues were set as additional variable modifications. The false discovery rate for peptide and protein identifications was set to 1%, and the minimum peptide length was set to 7. The minimum score threshold for both modified and unmodified peptides was set to 30. The MaxLFQ algorithm allowing label-free quantification (35) and the 'Matching Between Runs' feature were enabled. All mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (36) partner repository with the dataset identifier PXD003634. For the quantitative maize proteome and phosphoproteome analyses, the 'ProteinGroups' and 'Phospho(STY)sites' output files, respectively, generated by the MaxQuant search was loaded into Perseus, the data analysis software available in the MaxQuant package. Only proteins or phosphosites which were quantified in at least two of the three biological replicates of at least one sample were retained. Log2 transformed protein LFQ intensities or phosphosites intensities were centered by subtracting the median of the entire set of protein/phosphosite intensities per sample. A two-sample test with $p<0.05$ was carried out to test the differences between groups. The statistically significant hits were then Z-scored and clustered into groups by a hierarchical clustering analysis based on Euclidean distance.

To identify novel phosphosites (not previously reported ones), we compared our data to the PhosPhAt 4.0 full dataset of experimentally identified phosphosites (37) (data from

04.04.2016) and to the retrieved database of phosphosites identified in maize seed and leaf tissues from the Maize Protein Atlas (17, 38) (data from 04.04.2016).

Normalization of phosphoproteome data

We accounted for protein expression changes to allow proper interpretation of the maize quantitative phosphoproteomics data. After log2 transformation and centralization, the intensities of the phosphosites were normalized by subtracting the log2 transformed and centralized LFQ intensities of the corresponding proteins. The latter dataset of protein LFQ intensities resulted from the parallel protein expression study of all maize samples.

Gene Ontology Analysis

GO enrichment analysis was performed in the PLAZA 3.0 workbench (34). For the proteome dataset, 234 proteins with significant changes in abundance were analyzed, using the dataset of 2299 identified proteins as background model. A FDR cutoff ≤ 0.04 was used to score significantly overrepresented or depleted GO terms. For the phosphoproteome dataset, the identified proteins in each conditions were used for the enrichment, and the whole theoretical proteome (based on the genome annotation of *Z. mays*) was used at background. P-value cutoff was set at 0.04 and only terms enriched in either condition were presented (Table Sxxx).

Motif-X analysis

The Motif-X algorithm (39) was used to extract significantly enriched amino acid motifs surrounding the identified phosphosites. The sequence window was limited to 13 amino acids and foreground peptides were pre-aligned with the phosphosite centered. *Zea mays* proteome

data set from PLAZA was used as the background database. The occurrence threshold was set at the minimum of 20 peptides and the P-value threshold was set at $< 10^{-6}$.

STRING analysis of protein-protein interaction networks

Protein-protein interactions were analyzed by STRING (http://string-db.org/) (40), using the sequences of differentially phosphorylated proteins and proteins with significant abundance changes as input. The required confidence score was set as > 0.700 for highly confident interactions (STRING protein-protein interaction prediction is based on data available for genomic homology, gene fusion, occurrence in the same metabolic pathways, co-expression, experiments, database and text mining. A combined score is calculated based on the score of all the methods that were used for the protein-protein interaction prediction. The higher the score is, the more confident the interaction). The results were visualized using the Cytoscape package.

Pubmed search

A Pubmed search (www.ncbi.nlm.nih.gov/pubmed/) was performed on 31/03/2016 using 'maize proteomics 2015' or 'proteomics arabidopsis 2016' to identify relevant papers (only research papers with the correct focus were retained).

RESULTS AND DISCUSSION

Optimized quantitative workflow for proteomics and phosphoproteomics in plants

To facilitate efficient proteome analyses of plants, we developed a simple workflow, which maximizes the coverage and reproducibility of protein and phosphorylation site quantification

in single LC-MS/MS runs, thus without requiring peptide fractionation steps (**Figure 1**). To optimize the pipeline, we used the fully sequenced and well-annotated model plant *Arabidopsis thaliana* (see next section for results).

First, we provide a brief overview of the key steps in the protocol and the improvements that were introduced step-by-step to robustly survey plant proteomes and phosphoproteomes (more details can be found in the **Experimental Section**). To reproducibly capture a comprehensive spectrum of proteins, we opted for a protein precipitation approach. Plant tissue is known to be more challenging for proteome analyses than yeast and mammalian cells, with plant cells holding low protein contents and high concentrations of compounds that hinder preparation of proteome samples (e.g. polysaccharides, phenolic compounds, lipids and secondary metabolites) (41). From plant material that was grinded into powder, proteins were extracted with a sucrose buffer, containing protease and phosphatase inhibitors. The extract was subsequently purified through a chloroform/methanol precipitation step and the pelleted proteins were reconstituted in a buffer containing guanidinium hydrochloride. Cysteine disulfide bonds were reduced with tris(2-carboxyethyl)phosphine hydrochloride (TCEP-HCl), allowing the alkylation reaction with iodoacetamide to simultaneously take place (42). Next, we pre-digested the proteins with endoproteinase-LysC for 4 hours, followed by a full digestion with trypsin for 14 hours. The pre-digestion step was previously shown to substantially improve the proteolytic efficiency of trypsin (43). The resulting peptides were desalted, and split into two. One part was used for the proteome analysis, leaving 500 µg of digest material as input for phosphopeptide enrichment. We opted for a Ti^{4+}-IMAC-based method, as it was found to perform extremely well in terms of reproducibility and provides even greater selectivity and sensitivity than the more commonly used TiO_2 chromatography (44, 45). Both the proteome and phosphoproteome samples were analyzed by 3 hour gradients on a quadrupole Orbitrap instrument [Q Exactive (46)].

Second, peptide identification and quantification are important steps following LC-MS/MS analysis. We chose a label-free quantitation approach over a labeling method, as it is cost-effective, does not restrict the numbers of samples that can be compared, and can span several orders of magnitude of protein concentrations (35). In most label-free studies of plant phosphoproteomes, the raw data are analyzed by a combination of expensive expert peptide identification software, like Proteome Discoverer or Mascot, and in-house developed algorithms to facilitate label free quantification (10, 11, 26, 27) (**Supplementary Table S1**). Here, peptide identification was carried out by the freely available and easy-to-use software package MaxQuant (47). Simultaneously, the label-free quantitation is carried out by MaxQuant, in an ion intensity-based manner (35). The missing value issue, due to stochastic peptide sequencing inherent to mass spectrometry, was tackled by using the "match between runs" feature in MaxQuant, which can transfer MS/MS identifications between measurements based on a peptide retention time correlation approach (35).

Taken together, compared to published methods for (phospho)proteomics in plants (12, 14-16, 23), we reduced the number of sample preparation steps, MS time and data analysis complexity, due to the lack of labeling, gel-based steps and pre-fractionation steps and the introduction of MaxQuant in our workflow. With respect to the latter, this facilitates and standardizes data analysis, but does (not yet) seem to be routinely integrated in plant proteomics (**Supplementary Table S1**).

Validating the optimized workflow on **Arabidopsis thaliana** *roots*

To validate our complete (phospho)proteomics pipeline we used 100 mg *Arabidopsis* roots of five-day old seedlings, which were harvested in three independent biological replicates and yielded 1 mg proteins per sample. All samples were analyzed twice via nano-LC-MS/MS

using three hour gradients. We first analyzed the non-enriched samples, amounting to a total of 18 hours of MS time and leading to the cumulative identification of 34,216 unique peptides (with an estimated false discovery rate of 1%) that could be mapped on 4,903 protein groups (**Supplementary Table S2**). The latter can be defined as protein entries distinguishable on the basis of identified peptides (48). Via the MaxLFQ algorithm, 4,847 of those could be quantified in at least one biological replicate and 2,992 in all replicates. Seeing that label-free methods are very replicate dependent, reproducibility of the chromatographic separation must be very high. The data from the replicate experiments clearly show highly accurate quantitative reproducibility with an average Pearson correlation of 0.978 (**Supplementary Figure S1**).

A common challenge for plant proteomics studies is the difficulty of isolating proteins from the different subcellular organelles with sufficient efficiency. Membrane proteins represent an additional hurdle, as their large size and hydrophobicity render them difficult to isolate. To obtain a wide-ranging snapshot of cellular signaling processes it is vital to capture proteins from not only the cytosol, but also from membranes and organelles. GO analysis shows that the applied protocol extracted proteins from cytoplasm, nucleus, plasma membrane and other organelles (**Supplementary Table S3)**. This evidences that our approach is not limited by particular experimental difficulties and recovers proteins from all subcellular membranes and organelles.

Next, we monitored the phosphorylation events in the *Arabidopsis* samples. The total of six LC-MS/MS runs of the Ti^{4+}-IMAC enriched samples resulted in the identification of 1,051 unique phosphopeptides, corresponding to 1,331 phosphosites on 706 protein groups, at an estimated false discovery rate of 1%, for both peptide-spectrum match and protein (**Supplementary Table S4)**. The vast majority of these sites occurred on serine and threonine residues (90.3% and 9.1%, respectively), whereas phosphotyrosines accounted for less than

1% of the identified sites. This is in agreement with other reports (9, 13). Accurate site localization (probability > 0.75) was achieved for 799 of these phosphosites on 552 proteins. Of the 1,331 unique phosphosites, we could accurately quantify 1,022 and 711 in at least one and in all biological replicates, respectively. To evaluate the quality of the experiment, we assessed the correlation of all phosphopeptide intensities between the three biological replicates. An average Pearson's correlation of 0.818 illustrates the high reproducibility of the phosphopeptide enrichment strategy (**Supplementary Figure S1**). All the identified *Arabidopsis* phosphosites were used to search against the PhosPhAt 4.0 full dataset of experimentally identified phosphosites (37). This resulted in 169 phosphosites (13% of the dataset) uniquely identified in our study (**Supplementary Table S5**).

In summary, we have experimental evidence that our workflow successfully detects a large portion of the (phospho)proteome and can thus be applied to understanding biological processes.

Applying the (phospho)proteomics workflow to maize leaves under drought stress

Following the validation of our pipeline in *Arabidopsis* roots, we applied our workflow to a monocot crop under stress. Given the importance of drought-related research (33, 49-51), we profiled the proteome and phosphoproteome of maize leaves subjected to (severe) drought stress. Since the growth zone of the maize leaf determines to a great extent the final leaf length (52) and drought affects cell division and cell expansion in the growth zone of the maize leaf (53), we harvested the growth zone (4 basal centimeters) of the growing leaf 7 of 21 day old plants grown under drought conditions and control plants (three independent biological replicates were collected for both conditions). The drought was applied by preventing irrigation upon sowing and when the soil water content reached 62.5% of that of

the well-watered controls, the plants were maintained at the respective watering regime by daily watering. At the moment leaf 7 appeared, the effects of the drought were quantified by measuring the final leaf length of the youngest fully grown leaf, leaf 4. This leaf showed a significant length reduction compared to the control (**data not shown**), supporting that the applied drought affected leaf growth.

Proteome and phosphoproteome data were obtained for growth zones of maize leaves as described above, further emphasizing the importance of moving away from gel-based approaches, also in crops where this is not standard yet (**Supplementary Table S1**). All biological samples were analyzed twice by nanoLC-MS/MS using three hour gradients.

The maize leaf proteome under drought stress

In the non-enriched samples, a total of 22,093 peptides were identified originating from 4,409 protein groups. 4,361 protein groups could be accurately quantified, of which 2,299 in at least two of the three biological replicates (**Supplementary Table S6**). The data from the replicate experiments show quantitative reproducibility with an average Pearson correlation of 0.856 and 0.892 for control and drought samples, respectively (**Supplementary Figure S2**). Statistical testing ($p<0.05$) pinpointed 234 of these proteins to be differently regulated, with 156 up- and 78 down-regulated proteins upon drought stress (**Figure 2A and Supplementary Table S7-S8**). GO analysis of these proteins showed an overrepresentation of proteins involved in carbohydrate metabolism and chromatin remodeling (**Supplementary Table S9**), in agreement with previous studies of drought responses in crops (54-56). Furthermore, amongst the proteins with increased abundance we observed proteins associated with water deprivation, like lipoxygenase (GRMZM2G015419; (57)), keto reductase family 4 member C9 (GRMZM2G059314; (58)), fructose-biphosphate aldolase (GRMZM2G057823; (59)) and

protein-L-isoaspartate methyltransferase (GRMZM2G423027; (60)). Interestingly, BRI1-associated receptor kinase (BAK1) (GRMZM2G089819), known for its role in plant defense responses and brassinosteroid signaling (61, 62), was found to be down regulated in water deprived conditions. The brassinosteroid receptors, BRIs, to which BAK1 binds upon brassinosteroid induction, play an important role in maize leaf growth (63). Moreover, brassinosteroid signaling has previously been linked to abiotic stress (64-66) and brassinosteroid application is reported to improve drought tolerance in wheat and the resurrection grass *Sporobolus stapfianus* (67).

Analysis of protein-protein interactions between the 234 proteins with significant abundance changes resulted in a network of 141 proteins and 277 interactions, whereas each interaction has a combined score of all prediction methods > 0.7 (see **Experimental Section**) (**Figure 3)**. The network is approximately centralized around the DNA topoisomerase II (GRMZM2G021270/PLAZA identifier ZM05G37510). From the resulted network, we identified different groups of interaction between proteins involved in different cellular processes. These included the categories DNA/chromatin organization, photosynthesis and glucose metabolism, of which the corresponding GO terms were enriched in the dataset. Further, a small cluster of proteins involved in protein folding, including a member of the heat shock protein HSP70 family (GRMZM2G415007/PLAZA identifier ZM04G41380), three chaperone proteins belonging to the Clp protease family (GRMZM2G110023/PLAZA identifier ZM01G09650; GRMZM2G123922/PLAZA identifier ZM10G15640; GRMZM2G162968/ PLAZA identifier ZM09G19730) and the subunit β of the chaperonin containing T-complex (AC215201.3_FG005/PLAZA identifier ZM06G23100), was identified. It is known that the control of protein folding state is crucial for the survival of plants during abiotic stress (68).

The maize leaf phosphoproteome under drought stress

Ti^{4+} IMAC enrichment of maize phosphopeptides led to the detection of a total of 980 unique phosphosites on 686 phosphopeptides, which could be mapped on 536 phosphoproteins (**Supplementary Table S10**). The data from the replicate experiments show quantitative reproducibility with an average Pearson correlation of 0.887 and 0.856 for control and drought samples, respectively (**Supplementary Figure S3**). The number of identified phosphoproteins lies in between the range published in recent maize phosphoproteomics studies: 282 phosphoproteins in (23), 858 in (9), 2,852 in (14) and 3,557 in (17). Important to note is that the latter two studies fractionated the enriched phosphopeptides via extensive SCX chromatography, hereby greatly increasing MS analysis time per sample to two days (14, 17). In our work, six hours analysis time was used per sample, hence yielding a relative high number of phosphoprotein identifications.

Overall, the majority (97.4%) was mono-phosphorylated peptides, while around 2.6% of the phosphopeptides carried two phosphorylated residues. There were 84.0% phosphoserine, 15.2% phosphothreonine and 0.8% phosphotyrosine containing peptides identified, sharing a similar distribution pattern to other maize phosphoproteomics studies (9, 14, 17, 23). All the identified maize phosphosites were searched against the retrieved set of phosphosites identified in maize seed from the Atlas of Maize Proteotypes and a dataset of phosphosites garnered from different developmental zones of maize leaves (17, 38). This resulted in 359 phosphosites (37% of the dataset) uniquely identified in our study (**Supplementary Table S11**).

Overrepresentation of amino acid motifs surrounding the identified phosphosites were analyzed using Motif-X (**Table 1**). Phosphorylated tyrosine sites were excluded from the analysis due to their low abundance in the dataset. Similarly to other studies in *Arabidopsis*

and other monocots (69-72), [sP] is the most enriched motif for the S-phosphorylation as well as its phosphorylated threonine counterpart [tP] for the T-phosphorylation dataset. Further, 20 peptides are enriched with the proline-rich motif [sxSP]. Peptides containing the proline-directed [sP] and [tP] motifs are suggested to be substrates for MAP-kinases (MAPK), sucrose non-fermenting1-related protein kinase 2 (SnRK2), receptor-like kinases (RLKs), AGC family protein kinases PKA, PKG and PKC, CDKs (cyclin-dependent kinases), calcium-dependent protein kinases (CDPKs) and STE20-like kinases (SLKs) (69). Only one common acidic motif – [sDxE] – resulted from the analysis, belonging to 22 peptides that might be potential substrates for casein kinase II (CKII) and CDPKs. Further, three basophilic motifs are overrepresented in the dataset, [Rxxs] and the subtype [RSxs], which are recognized by MAPK kinases (MKKs), and [Kxxs], which is targeted by PKA and PKC. No specific protein kinases are found for the T-phosphorylation motif [tS].

Differential analysis of phosphorylation sites

Earlier differential phosphoproteomics of maize leaf tissue, identifying differences between stress conditions, lack normalization to the protein abundance (9, 15, 23). Here, thanks to our extensive analysis, we can simultaneously take into account protein and phosphorylation site profiles.

In total, 615 phosphosites on 445 phosphoproteins were quantified, of which 536 phosphosites in at least two biological replicates of one condition. Taking into account that differences in protein levels can influence the outcome of the differential phosphorylation data, we set out to normalize the intensities of the phosphopeptides to the protein intensities. For 224 phosphosites, matching proteins were quantified in the proteome experiment allowing normalization. This result demonstrated that phosphopeptide enrichment facilitated the

identification of low abundance proteins, of which non-phosphorylated peptides were likely missed in the proteome scans due to different dynamic ranges and crowdedness. A two sample test ($p<0.05$) on the normalized phosphosites intensities showed that 18 of those were differentially regulated by drought (**Figure 2B and Supplementary Table S12**). Some of those phosphosites, S470 on HISTONE DEACETYLASE 6 (GRMZM2G457889) (8.3-fold down regulated upon drought) and S247 stem-specific protein TSJT1 (GRMZM2G169671) (2.5-fold up regulated upon drought), are mapped on proteins previously shown to be regulated during drought signaling (73, 74). In mammalian systems, phosphorylation of HISTONE DEACETYLASE 6 was shown to correlate with enzyme activity and consequent tubulin deacetylation and microtubule destabilization (75). A similar mechanism could take place during drought signaling, as plant microtubules are known to function as sensors for abiotic stress (76).

As protein levels for many phosphosites could not be inferred, because no non-phosphorylated peptides of the corresponding proteins were detected, we also subjected the not-normalized phosphosites dataset to a two sample test ($p<0.05$). Based on the phosphoproteome data alone, we found 44 phosphosites to be statistically significant between two conditions, of which 32 were up-regulated upon drought stress and 12 down-regulated (**Figure 2C and Supplementary Table S13-14**). Interestingly, four microtubule-associated proteins were found to be differentially phosphorylated upon drought stress: MICROTUBULE-ASSOCIATED PROTEIN 70-2 (GRMZM2G017525), DYNEIN LIGHT CHAIN 1 (GRMZM2G472231), MICROTUBULE-ASSOCIATED PROTEIN (GRMZM2G026309), and KINESIN-LIKE PROTEIN KIN12A (GRMZM2G034828). Microtubule-associated proteins are involved in microtubuli organization and their binding affinity to microtubules is known to be controlled via phosphorylation (77, 78). Furthermore, the phosphorylation status of a putative MAP kinase superfamily protein

(GRMZM2G044557) and a PROTEIN PHOSPHATASE 2C 64 (GRMZM2G107565; GRMZM2G021610) changed upon water deprivation. In stressed conditions, a SPS1-related proline-alanine-rich protein kinase (GRMZM2G413544; GRMZM2G073399) was found to be less phosphorylated. This serine/threonine kinase is a part of the Sterile 20 (Ste20)-related kinase family that is conserved across the fungi, plant and animal kingdom (79) and its mammalian homologs are known to act as a mediator of stress-activated signals (80). Up to our knowledge this protein has not been related to drought stress.

Because of the small dataset of significantly regulated phosphosites, no GO classes showed a significant over- or underrepresentation versus the dataset of all identified phosphorylation sites ($p < 0.05$). Alternatively, we investigated which GO classes were only enriched in the control or the drought stressed samples versus a background set of all maize proteins. This analysis showed that pathways involved in sodium transport, immune response and chromatin silencing are exclusively overrepresented in the drought samples (**Supplementary Table S15**).

A crop PTM database

Comprehensive information of *Arabidopsis* protein phosphorylation sites can be found at the PhosPHAt database (37) and P3DB (81). However, information on detected phosphorylation sites in crop species is only scarcely available in these resources. A very specific database, Atlas of Maize Proteotypes, holds proteomics data for maize seed tissue and can also be consulted to query phosphorylation sites, but only in limited, seed-specific datasets (38). To serve as a general tool for PTMs in all plants, we generated a searchable database for plant protein posttranslational modifications, called *Plant PTM viewer* (bioinformatics.psb.ugent.be/webtools/ptm_viewer/). This database will function as an

important resource for future functional studies of plant protein posttranslational modifications, including protein phosphorylation.

CONCLUSIONS

Phosphoproteomics workflows traditionally involve tedious labeling and fractionation steps for comprehensive quantitative analysis of phosphorylation profiles. Here, we present a streamlined and reproducible platform for quantitative phosphoproteomics, which (1) does not require specialized equipment nor expert software and can be easily implemented in any molecular biology lab with access to a mass spectrometer, (2) involves limited sample prep time due to the lack of labeling, fractionation and gel steps, and requires relatively low amounts of starting material [which varies a lot across studies (17, 82, 83)] needed because of the straightforward and sensitive pipeline, (3) does not require excessive MS analysis time as the samples are not fractionated into multiple fractions, and (4) can be applied to model plants, such as *A. thaliana*, and economically important crops, such as maize. Wherever possible, protein levels inferred from non-phosphorylated peptides should be used to normalize phosphopeptide intensities, so that true phosphorylation events rather than changes in phosphoprotein amounts can be monitored. This has so far been poorly implemented in plant phosphoprotemics, but is absolutely essential when reporting differential changes in the phosphoproteome.

As a consequence of climate changes, drought stress has become a severe limiting factor in plant growth and productivity throughout the world (84, 85). Showcased in drought-stressed maize leaves, our workflow enabled the in-depth quantitative comparison of phosphorylation patterns. Finally, the data generated, comprising novel (phospho)protein candidates implicated in drought stress signaling, contributes to our understanding of the

molecular and cellular mechanisms utilized by crops to survive unfavorable environmental conditions.

ASSOCIATED CONTENT

SUPPORTING INFORMATION

Figure S1. Pearson correlation coefficient for Arabidopsis proteome and phosphoproteome data.

Figure S2. Pearson correlation coefficient for maize proteome data.

Figure S3. Pearson correlation coefficient for maize phosphoproteome data.

Table S1. Pubmed search.

Table S2. Identified proteins in whole proteome analysis in Arabidopsis root

Table S3. GO categorization in Arabidopsis proteome dataset

Table S4. Total phosphosites identified in Arabidopsis roots

Table S5. List of phosphosites in Arabidopsis uniquely identified in this study

Table S6. Identified proteins in whole proteome analysis in maize leaves under drought and control conditions

Table S7. Maize proteins after filtering and Student's t-test

Table S8. Significantly different protein groups upon drought stress in maize leaves ($p<0.05$).

Table S9. GO enrichment analysis of significantly differential maize proteins, with all identified maize proteins as background dataset

Table S10. Total phosphosites identified in maize leaves under control and drought conditions

Table S11. List of phosphosites in maize leaves uniquely identified in this study

Table S12. Significantly different phosphosites upon drought stress in maize leaves ($p<0.05$), after normalization to protein levels.

Table S13. Maize phosphosites after filtering and Student's t-test

Table S14. Significantly different phosphosites upon drought stress in maize leaves ($p<0.05$), without normalization to protein levels.

Table S15. GO classes exclusively overrepresented in one of the two samples, either control or drought, with the whole maize annotated genome as background set.

AUTHOR INFORMATION

Corresponding Author

† E-mail: ive.desmet@psb.vib-ugent.be; Phone 003293313930.

Author Contributions

E.S., D.I., K.G., and I.D.S. designed research; E.S., L.D.V., M.V.B, H.N. and F.C. performed research; E.S. and L.D.V. analyzed data; E.S., L.D.V., K.G., and I.D.S. wrote the paper. All authors have given approval to the final version of the manuscript.

# or * These authors contributed equally.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

E.S. was a Postdoctoral Fellow of the Research Foundation-Flanders. L.D.V. is the recipient of a VIB International PhD program fellowship.

REFERENCES

1. Cohen, P., The origins of protein phosphorylation. *Nat Cell Biol* **2002,** 4, (5), E127-30.
2. Xu, J.; Zhang, S., Mitogen-activated protein kinase cascades in signaling plant growth and development. *Trends Plant Sci* **2015,** 20, (1), 56-64.
3. Li, J.; Silva-Sanchez, C.; Zhang, T.; Chen, S.; Li, H., Phosphoproteomics technologies and applications in plant biology research. *Front Plant Sci* **2015,** 6, 430.
4. Benschop, J. J.; Mohammed, S.; O'Flaherty, M.; Heck, A. J.; Slijper, M.; Menke, F. L., Quantitative phosphoproteomics of early elicitor signaling in Arabidopsis. *Mol Cell Proteomics* **2007,** 6, (7), 1198-214.
5. Fan, S.; Meng, Y.; Song, M.; Pang, C.; Wei, H.; Liu, J.; Zhan, X.; Lan, J.; Feng, C.; Zhang, S.; Yu, S., Quantitative phosphoproteomics analysis of nitric oxide-responsive phosphoproteins in cotton leaf. *PLoS One* **2014,** 9, (4), e94261.
6. Stecker, K. E.; Minkoff, B. B.; Sussman, M. R., Phosphoproteomic Analyses Reveal Early Signaling Events in the Osmotic Stress Response. *Plant Physiol* **2014,** 165, (3), 1171-1187.
7. Yang, F.; Melo-Braga, M. N.; Larsen, M. R.; Jorgensen, H. J.; Palmisano, G., Battle through signaling between wheat and the fungal pathogen Septoria tritici revealed by proteomics and phosphoproteomics. *Mol Cell Proteomics* **2013,** 12, (9), 2497-508.
8. Reiland, S.; Finazzi, G.; Endler, A.; Willig, A.; Baerenfaller, K.; Grossmann, J.; Gerrits, B.; Rutishauser, D.; Gruissem, W.; Rochaix, J. D.; Baginsky, S., Comparative phosphoproteome profiling reveals a function of the STN8 kinase in fine-tuning of cyclic electron flow (CEF). *Proc Natl Acad Sci U S A* **2011,** 108, (31), 12955-60.
9. Wu, L.; Hu, X.; Wang, S.; Tian, L.; Pang, Y.; Han, Z.; Wu, L.; Chen, Y., Quantitative analysis of changes in the phosphoproteome of maize induced by the plant hormone salicylic acid. *Sci Rep* **2015,** 5, 18155.
10. Qiu, J.; Hou, Y.; Tong, X.; Wang, Y.; Lin, H.; Liu, Q.; Zhang, W.; Li, Z.; Nallamilli, B. R.; Zhang, J., Quantitative phosphoproteomic analysis of early seed development in rice (Oryza sativa L.). *Plant Mol Biol* **2015**.
11. Hoehenwarter, W.; Thomas, M.; Nukarinen, E.; Egelhofer, V.; Rohrig, H.; Weckwerth, W.; Conrath, U.; Beckers, G. J., Identification of novel in vivo MAP kinase substrates in Arabidopsis thaliana through use of tandem metal oxide affinity chromatography. *Mol Cell Proteomics* **2013,** 12, (2), 369-80.
12. Roitinger, E.; Hofer, M.; Kocher, T.; Pichler, P.; Novatchkova, M.; Yang, J.; Schlogelhofer, P.; Mechtler, K., Quantitative phosphoproteomics of the ataxia telangiectasia-mutated (ATM) and ataxia telangiectasia-mutated and rad3-related (ATR) dependent DNA damage response in Arabidopsis thaliana. *Mol Cell Proteomics* **2015,** 14, (3), 556-71.
13. Nakagami, H.; Sugiyama, N.; Mochida, K.; Daudi, A.; Yoshida, Y.; Toyoda, T.; Tomita, M.; Ishihama, Y.; Shirasu, K., Large-scale comparative phosphoproteomics identifies conserved phosphorylation sites in plants. *Plant Physiol* **2010,** 153, (3), 1161-74.
14. Marcon, C.; Malik, W. A.; Walley, J. W.; Shen, Z.; Paschold, A.; Smith, L. G.; Piepho, H. P.; Briggs, S. P.; Hochholdinger, F., A high-resolution tissue-specific proteome and phosphoproteome atlas of maize primary roots reveals functional gradients along the root axes. *Plant Physiol* **2015,** 168, (1), 233-46.
15. Bonhomme, L.; Valot, B.; Tardieu, F.; Zivy, M., Phosphoproteome dynamics upon changes in plant water status reveal early events associated with rapid growth adjustment in maize leaves. *Mol Cell Proteomics* **2012,** 11, (10), 957-72.

16. Zhang, H.; Zhou, H.; Berke, L.; Heck, A. J.; Mohammed, S.; Scheres, B.; Menke, F. L., Quantitative phosphoproteomics after auxin-stimulated lateral root induction identifies an SNX1 protein phosphorylation site required for growth. *Mol Cell Proteomics* **2013,** 12, (5), 1158-69.
17. Facette, M. R.; Shen, Z.; Bjornsdottir, F. R.; Briggs, S. P.; Smith, L. G., Parallel proteomic and phosphoproteomic analyses of successive stages of maize leaf development. *Plant Cell* **2013,** 25, (8), 2798-812.
18. Sharma, K.; D'Souza, R. C.; Tyanova, S.; Schaab, C.; Wisniewski, J. R.; Cox, J.; Mann, M., Ultradeep human phosphoproteome reveals a distinct regulatory nature of Tyr and Ser/Thr-based signaling. *Cell Rep* **2014,** 8, (5), 1583-94.
19. Lienhard, G. E., Non-functional phosphorylations? *Trends Biochem Sci* **2008,** 33, (8), 351-2.
20. Niu, S.; Wang, Z.; Ge, D.; Zhang, G.; Li, Y., Prediction of functional phosphorylation sites by incorporating evolutionary information. *Protein Cell* **2012,** 3, (9), 675-90.
21. Kline, K. G.; Barrett-Wilt, G. A.; Sussman, M. R., In planta changes in protein phosphorylation induced by the plant hormone abscisic acid. *Proc Natl Acad Sci U S A* **2010,** 107, (36), 15986-91.
22. Qing, D.; Yang, Z.; Li, M.; Wong, W. S.; Guo, G.; Liu, S.; Guo, H.; Li, N., Quantitative and Functional Phosphoproteomic Analysis Reveals that Ethylene Regulates Water Transport via the C-Terminal Phosphorylation of Aquaporin PIP2;1 in Arabidopsis. *Mol Plant* **2015**.
23. Hu, X.; Wu, L.; Zhao, F.; Zhang, D.; Li, N.; Zhu, G.; Li, C.; Wang, W., Phosphoproteomic analysis of the response of maize leaves to drought, heat and their combination stress. *Front Plant Sci* **2015,** 6, 298.
24. Lan, P.; Li, W.; Wen, T. N.; Schmidt, W., Quantitative phosphoproteome profiling of iron-deficient Arabidopsis roots. *Plant Physiol* **2012,** 159, (1), 403-17.
25. Engelsberger, W. R.; Schulze, W. X., Nitrate and ammonium lead to distinct global dynamic phosphorylation patterns when resupplied to nitrogen-starved Arabidopsis seedlings. *Plant J* **2012,** 69, (6), 978-95.
26. Wang, P.; Xue, L.; Batelli, G.; Lee, S.; Hou, Y. J.; Van Oosten, M. J.; Zhang, H.; Tao, W. A.; Zhu, J. K., Quantitative phosphoproteomics identifies SnRK2 protein kinase substrates and reveals the effectors of abscisic acid action. *Proc Natl Acad Sci U S A* **2013,** 110, (27), 11205-10.
27. Xue, L.; Wang, P.; Wang, L.; Renzi, E.; Radivojac, P.; Tang, H.; Arnold, R.; Zhu, J. K.; Tao, W. A., Quantitative measurement of phosphoproteome response to osmotic stress in arabidopsis based on Library-Assisted eXtracted Ion Chromatogram (LAXIC). *Mol Cell Proteomics* **2013,** 12, (8), 2354-69.
28. Wu, R.; Dephoure, N.; Haas, W.; Huttlin, E. L.; Zhai, B.; Sowa, M. E.; Gygi, S. P., Correct interpretation of comprehensive phosphorylation dynamics requires normalization by protein expression changes. *Mol Cell Proteomics* **2011,** 10, (8), M111 009654.
29. McKersie, B., Planning for food security in a changing climate. *J Exp Bot* **2015,** 66, (12), 3435-50.
30. Shanker, A. K.; Maheswari, M.; Yadav, S. K.; Desai, S.; Bhanu, D.; Attal, N. B.; Venkateswarlu, B., Drought stress responses in crops. *Funct Integr Genomics* **2014,** 14, (1), 11-22.
31. Kole, C.; Muthamilarasan, M.; Henry, R.; Edwards, D.; Sharma, R.; Abberton, M.; Batley, J.; Bentley, A.; Blakeney, M.; Bryant, J.; Cai, H.; Cakir, M.; Cseke, L. J.; Cockram, J.; de Oliveira, A. C.; De Pace, C.; Dempewolf, H.; Ellison, S.; Gepts, P.; Greenland, A.; Hall, A.; Hori, K.; Hughes, S.; Humphreys, M. W.; Iorizzo, M.; Ismail, A. M.; Marshall, A.;

Mayes, S.; Nguyen, H. T.; Ogbonnaya, F. C.; Ortiz, R.; Paterson, A. H.; Simon, P. W.; Tohme, J.; Tuberosa, R.; Valliyodan, B.; Varshney, R. K.; Wullschleger, S. D.; Yano, M.; Prasad, M., Application of genomics-assisted breeding for generation of climate resilient crops: progress and prospects. *Front Plant Sci* **2015,** 6, 563.

32. Kosova, K.; Vitamvas, P.; Urban, M. O.; Klima, M.; Roy, A.; Prasil, I. T., Biological Networks Underlying Abiotic Stress Tolerance in Temperate Crops--A Proteomic Perspective. *Int J Mol Sci* **2015,** 16, (9), 20913-42.

33. Wu, X.; Gong, F.; Cao, D.; Hu, X.; Wang, W., Advances in crop proteomics: PTMs of proteins under abiotic stress. *Proteomics* **2016,** 16, (5), 847-65.

34. Proost, S.; Van Bel, M.; Vaneechoutte, D.; Van de Peer, Y.; Inze, D.; Mueller-Roeber, B.; Vandepoele, K., PLAZA 3.0: an access point for plant comparative genomics. *Nucleic Acids Res* **2015,** 43, (Database issue), D974-81.

35. Cox, J.; Hein, M. Y.; Luber, C. A.; Paron, I.; Nagaraj, N.; Mann, M., Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol Cell Proteomics* **2014,** 13, (9), 2513-26.

36. Vizcaino, J. A.; Csordas, A.; Del-Toro, N.; Dianes, J. A.; Griss, J.; Lavidas, I.; Mayer, G.; Perez-Riverol, Y.; Reisinger, F.; Ternent, T.; Xu, Q. W.; Wang, R.; Hermjakob, H., 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res* **2016,** 44, (D1), D447-56.

37. Durek, P.; Schmidt, R.; Heazlewood, J. L.; Jones, A.; MacLean, D.; Nagel, A.; Kersten, B.; Schulze, W. X., PhosPhAt: the Arabidopsis thaliana phosphorylation site database. An update. *Nucleic Acids Res* **2010,** 38, (Database issue), D828-34.

38. Walley, J. W.; Shen, Z.; Sartor, R.; Wu, K. J.; Osborn, J.; Smith, L. G.; Briggs, S. P., Reconstruction of protein networks from an atlas of maize seed proteotypes. *Proc Natl Acad Sci U S A* **2013,** 110, (49), E4808-17.

39. Chou, M. F.; Schwartz, D., Biological sequence motif discovery using motif-x. *Curr Protoc Bioinformatics* **2011,** Chapter 13, Unit 13 15-24.

40. Szklarczyk, D.; Franceschini, A.; Wyder, S.; Forslund, K.; Heller, D.; Huerta-Cepas, J.; Simonovic, M.; Roth, A.; Santos, A.; Tsafou, K. P.; Kuhn, M.; Bork, P.; Jensen, L. J.; von Mering, C., STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res* **2015,** 43, (Database issue), D447-52.

41. Rose, J. K.; Bashir, S.; Giovannoni, J. J.; Jahn, M. M.; Saravanan, R. S., Tackling the plant proteome: practical approaches, hurdles and experimental tools. *Plant J* **2004,** 39, (5), 715-33.

42. Herbert, B. R.; Molloy, M. P.; Gooley, A. A.; Walsh, B. J.; Bryson, W. G.; Williams, K. L., Improved protein solubility in two-dimensional electrophoresis using tributyl phosphine as reducing agent. *Electrophoresis* **1998,** 19, (5), 845-51.

43. Glatter, T.; Ludwig, C.; Ahrne, E.; Aebersold, R.; Heck, A. J.; Schmidt, A., Large-scale quantitative assessment of different in-solution protein digestion protocols reveals superior cleavage efficiency of tandem Lys-C/trypsin proteolysis over trypsin digestion. *J Proteome Res* **2012,** 11, (11), 5145-56.

44. Zhou, H.; Ye, M.; Dong, J.; Corradini, E.; Cristobal, A.; Heck, A. J.; Zou, H.; Mohammed, S., Robust phosphoproteome enrichment using monodisperse microsphere-based immobilized titanium (IV) ion affinity chromatography. *Nat Protoc* **2013,** 8, (3), 461-80.

45. de Graaf, E. L.; Giansanti, P.; Altelaar, A. F.; Heck, A. J., Single-step enrichment by Ti4+-IMAC and label-free quantitation enables in-depth monitoring of phosphorylation dynamics with high reproducibility and temporal resolution. *Mol Cell Proteomics* **2014,** 13, (9), 2426-34.

46. Michalski, A.; Damoc, E.; Hauschild, J. P.; Lange, O.; Wieghaus, A.; Makarov, A.; Nagaraj, N.; Cox, J.; Mann, M.; Horning, S., Mass spectrometry-based proteomics using Q Exactive, a high-performance benchtop quadrupole Orbitrap mass spectrometer. *Mol Cell Proteomics* **2011,** 10, (9), M111 011015.
47. Cox, J.; Neuhauser, N.; Michalski, A.; Scheltema, R. A.; Olsen, J. V.; Mann, M., Andromeda: a peptide search engine integrated into the MaxQuant environment. *J Proteome Res* **2011,** 10, (4), 1794-805.
48. Cox, J.; Mann, M., MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* **2008,** 26, (12), 1367-72.
49. Budak, H.; Hussain, B.; Khan, Z.; Ozturk, N. Z.; Ullah, N., From Genetics to Functional Genomics: Improvement in Drought Signaling and Tolerance in Wheat. *Front Plant Sci* **2015,** 6, 1012.
50. Gilbert, M. E.; Medina, V., Drought Adaptation Mechanisms Should Guide Experimental Design. *Trends Plant Sci* **2016**.
51. Claeys, H.; Inze, D., The agony of choice: how plants balance growth and survival under water-limiting conditions. *Plant Physiol* **2013,** 162, (4), 1768-79.
52. Nelissen, H.; Rymen, B.; Jikumaru, Y.; Demuynck, K.; Van Lijsebettens, M.; Kamiya, Y.; Inze, D.; Beemster, G. T., A local maximum in gibberellin levels regulates maize leaf growth by spatial control of cell division. *Curr Biol* **2012,** 22, (13), 1183-7.
53. Avramova, V.; Sprangers, K.; Beemster, G. T., The Maize Leaf: Another Perspective on Growth Regulation. *Trends Plant Sci* **2015,** 20, (12), 787-97.
54. Rampitsch, C.; Bykova, N. V., The beginnings of crop phosphoproteomics: exploring early warning systems of stress. *Front Plant Sci* **2012,** 3, 144.
55. Benesova, M.; Hola, D.; Fischer, L.; Jedelsky, P. L.; Hnilicka, F.; Wilhelmova, N.; Rothova, O.; Kocova, M.; Prochazkova, D.; Honnerova, J.; Fridrichova, L.; Hnilickova, H., The physiology and proteomics of drought tolerance in maize: early stomatal closure as a cause of lower tolerance to short-term dehydration? *PLoS One* **2012,** 7, (6), e38017.
56. Hao, P.; Zhu, J.; Gu, A.; Lv, D.; Ge, P.; Chen, G.; Li, X.; Yan, Y., An integrative proteome analysis of different seedling organs in tolerant and sensitive wheat cultivars under drought stress and recovery. *Proteomics* **2015,** 15, (9), 1544-63.
57. De La Fuente, G. N.; Murray, S. C.; Isakeit, T.; Park, Y. S.; Yan, Y.; Warburton, M. L.; Kolomiets, M. V., Characterization of genetic diversity and linkage disequilibrium of ZmLOX4 and ZmLOX5 loci in maize. *PLoS One* **2013,** 8, (1), e53973.
58. Sengupta, D.; Naik, D.; Reddy, A. R., Plant aldo-keto reductases (AKRs) as multi-tasking soldiers involved in diverse plant metabolic processes and stress defense: A structure-function update. *J Plant Physiol* **2015,** 179, 40-55.
59. Riccardi, F.; Gazeau, P.; Jacquemot, M. P.; Vincent, D.; Zivy, M., Deciphering genetic variations of proteome responses to water deficit in maize leaves. *Plant Physiol Biochem* **2004,** 42, (12), 1003-11.
60. Thapar, N.; Kim, A. K.; Clarke, S., Distinct patterns of expression but similar biochemical properties of protein L-isoaspartyl methyltransferase in higher plants. *Plant Physiol* **2001,** 125, (2), 1023-35.
61. Vert, G., Plant signaling: brassinosteroids, immunity and effectors are BAK ! *Curr Biol* **2008,** 18, (20), R963-5.
62. Yamada, K.; Yamashita-Yamada, M.; Hirase, T.; Fujiwara, T.; Tsuda, K.; Hiruma, K.; Saijo, Y., Danger peptide receptor signaling in plants ensures basal immunity upon pathogen-induced depletion of BAK1. *EMBO J* **2015**.

63. Kir, G.; Ye, H.; Nelissen, H.; Neelakandan, A. K.; Kusnandar, A. S.; Luo, A.; Inze, D.; Sylvester, A. W.; Yin, Y.; Becraft, P. W., RNA Interference Knockdown of BRASSINOSTEROID INSENSITIVE1 in Maize Reveals Novel Functions for Brassinosteroid Signaling in Controlling Plant Architecture. *Plant Physiol* **2015,** 169, (1), 826-39.
64. Marshall, A.; Aalen, R. B.; Audenaert, D.; Beeckman, T.; Broadley, M. R.; Butenko, M. A.; Cano-Delgado, A. I.; de Vries, S.; Dresselhaus, T.; Felix, G.; Graham, N. S.; Foulkes, J.; Granier, C.; Greb, T.; Grossniklaus, U.; Hammond, J. P.; Heidstra, R.; Hodgman, C.; Hothorn, M.; Inze, D.; Ostergaard, L.; Russinova, E.; Simon, R.; Skirycz, A.; Stahl, Y.; Zipfel, C.; De Smet, I., Tackling drought stress: receptor-like kinases present new approaches. *Plant Cell* **2012,** 24, (6), 2262-78.
65. Janeczko, A.; Oklestkova, J.; Pociecha, E.; Koscielniak, J.; Mirek, M., Physiological effects and transport of 24-epibrassinolide in heat-stressed barley. *Acta Physiologiae Plantarum* **2011,** 33, (4), 1249-1259.
66. Anjum, S. A.; Wang, L. C.; Farooq, M.; Hussain, M.; Xue, L. L.; Zou, C. M., Brassinolide Application Improves the Drought Tolerance in Maize Through Modulation of Enzymatic Antioxidants and Leaf Gas Exchange. *Journal of Agronomy and Crop Science* **2011,** 197, (3), 177-185.
67. Bajguz, A.; Hayat, S., Effects of brassinosteroids on the plant responses to environmental stresses. *Plant Physiology and Biochemistry* **2009,** 47, (1), 1-8.
68. Ghosh, D.; Xu, J., Abiotic stress responses in plant roots: a proteomics perspective. *Front Plant Sci* **2014,** 5, 6.
69. van Wijk, K. J.; Friso, G.; Walther, D.; Schulze, W. X., Meta-Analysis of Arabidopsis thaliana Phospho-Proteomics Data Reveals Compartmentalization of Phosphorylation Motifs. *Plant Cell* **2014,** 26, (6), 2367-2389.
70. Zhang, M.; Lv, D.; Ge, P.; Bian, Y.; Chen, G.; Zhu, G.; Li, X.; Yan, Y., Phosphoproteome analysis reveals new drought response and defense mechanisms of seedling leaves in bread wheat (Triticum aestivum L.). *J Proteomics* **2014,** 109, 290-308.
71. Lv, D. W.; Li, X.; Zhang, M.; Gu, A. Q.; Zhen, S. M.; Wang, C.; Li, X. H.; Yan, Y. M., Large-scale phosphoproteome analysis in seedling leaves of Brachypodium distachyon L. *BMC Genomics* **2014,** 15, 375.
72. Hou, Y.; Qiu, J.; Tong, X.; Wei, X.; Nallamilli, B. R.; Wu, W.; Huang, S.; Zhang, J., A comprehensive quantitative phosphoproteome analysis of rice in response to bacterial blight. *BMC Plant Biol* **2015,** 15, 163.
73. Sircar, S.; Parekh, N., Functional characterization of drought-responsive modules and genes in Oryza sativa: a network-based approach. *Front Genet* **2015,** 6, 256.
74. Kim, J. M.; Sasaki, T.; Ueda, M.; Sako, K.; Seki, M., Chromatin changes in response to drought, salinity, heat, and cold stresses in plants. *Front Plant Sci* **2015,** 6, 114.
75. Pugacheva, E. N.; Jablonski, S. A.; Hartman, T. R.; Henske, E. P.; Golemis, E. A., HEF1-dependent Aurora A activation induces disassembly of the primary cilium. *Cell* **2007,** 129, (7), 1351-63.
76. Nick, P., Microtubules, signalling and abiotic stress. *Plant J* **2013,** 75, (2), 309-23.
77. Beck, M.; Komis, G.; Muller, J.; Menzel, D.; Samaj, J., Arabidopsis homologs of nucleus- and phragmoplast-localized kinase 2 and 3 and mitogen-activated protein kinase 4 are essential for microtubule organization. *Plant Cell* **2010,** 22, (3), 755-71.
78. Komis, G.; Illes, P.; Beck, M.; Samaj, J., Microtubules and mitogen-activated protein kinase signalling. *Curr Opin Plant Biol* **2011,** 14, (6), 650-7.
79. Gagnon, K. B.; Delpire, E., Molecular physiology of SPAK and OSR1: two Ste20-related protein kinases regulating ion transport. *Physiol Rev* **2012,** 92, (4), 1577-617.

80. Johnston, A. M.; Naselli, G.; Gonez, L. J.; Martin, R. M.; Harrison, L. C.; DeAizpurua, H. J., SPAK, a STE20/SPS1-related kinase that activates the p38 pathway. *Oncogene* **2000,** 19, (37), 4290-7.
81. Yao, Q.; Ge, H.; Wu, S.; Zhang, N.; Chen, W.; Xu, C.; Gao, J.; Thelen, J. J.; Xu, D., P(3)DB 3.0: From plant phosphorylation sites to protein networks. *Nucleic Acids Res* **2014,** 42, (Database issue), D1206-13.
82. Pi, E.; Qu, L.; Hu, J.; Huang, Y.; Qiu, L.; Lu, H.; Jiang, B.; Liu, C.; Peng, T.; Zhao, Y.; Wang, H.; Tsai, S. N.; Ngai, S.; Du, L., Mechanisms of Soybean Roots' Tolerances to Salinity Revealed by Proteomic and Phosphoproteomic Comparisons Between Two Cultivars. *Mol Cell Proteomics* **2016,** 15, (1), 266-88.
83. Choudhary, M. K.; Nomura, Y.; Wang, L.; Nakagami, H.; Somers, D. E., Quantitative Circadian Phosphoproteomic Analysis of Arabidopsis Reveals Extensive Clock Control of Key Components in Physiological, Metabolic, and Signaling Pathways. *Mol Cell Proteomics* **2015,** 14, (8), 2243-60.
84. Cairns, J. E.; Sanchez, C.; Vargas, M.; Ordonez, R.; Araus, J. L., Dissecting maize productivity: ideotypes associated with grain yield under drought stress and well-watered conditions. *J Integr Plant Biol* **2012,** 54, (12), 1007-20.
85. Harrison, M. T.; Tardieu, F.; Dong, Z.; Messina, C. D.; Hammer, G. L., Characterizing drought stress and trait influence on maize yield under current and future conditions. *Glob Chang Biol* **2014,** 20, (3), 867-78.

FIGURE LEGENDS

Figure 1. Workflow and key improved steps.

Figure 2. Differential (phospho)proteome response upon drought stress in maize leaves. Heat map showing average log2 values of MaxLFQ intensities of the significantly differentially expressed proteins (A), phosphosite intensities of the significantly regulated phosphosites after normalization to protein levels (B), or phosphosite intensities of the significantly regulated phosphosites without correction for protein abundance (C). The log2 values of the intensities were Z-scored for graphical representation.

Figure 3. Protein-protein interaction networks resulted from STRING analysis of maize proteins with significant abundance changes. Cytoscape was used for visualization. Subnetworks with fewer than 6 interactors are excluded from the representation. Nodes in red represent upregulated proteins, in green downregulated proteins upon drought. Interaction groups are indicated with black circles. The PLAZA identifiers for maize proteins were used because of space constraints.

TABLES

Motif	Motif score	Foreground matches	Foreground size	Background matches	Background size	Fold Increase
. . . . . .S	16.00	210	784	63261	975876	4.13
. . . . . .S. SP. . .	13.24	20	415	6516	850139	6.29
. . . . . .SD. E . . .	19.40	22	437	3450	853589	12.46
. . .R . .S.	16.00	98	535	52077	905666	3.19
. . .RS. S.	22.98	39	574	6949	912615	8.92
. . .K . .S.	6.33	43	395	39892	843623	2.30
. . . . . .TP.	6.60	27	140	35595	556423	3.01
. . . . . .TS.	6.01	29	113	50857	520828	2.63

Table 1. Motif-X analysis for overrepresented phosphorylation motifs of all identified phosphosites in maize leaves.

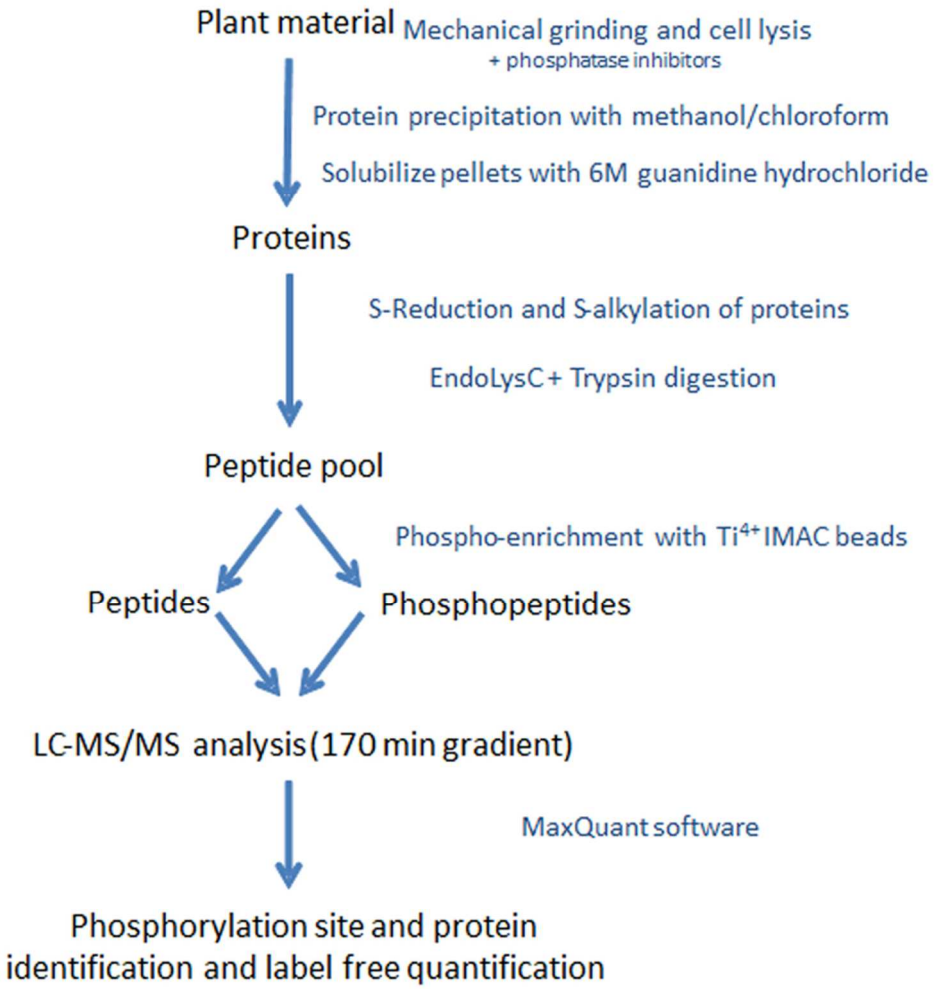

Figure 1. Workflow and key improved steps.
76x78mm (300 x 300 DPI)

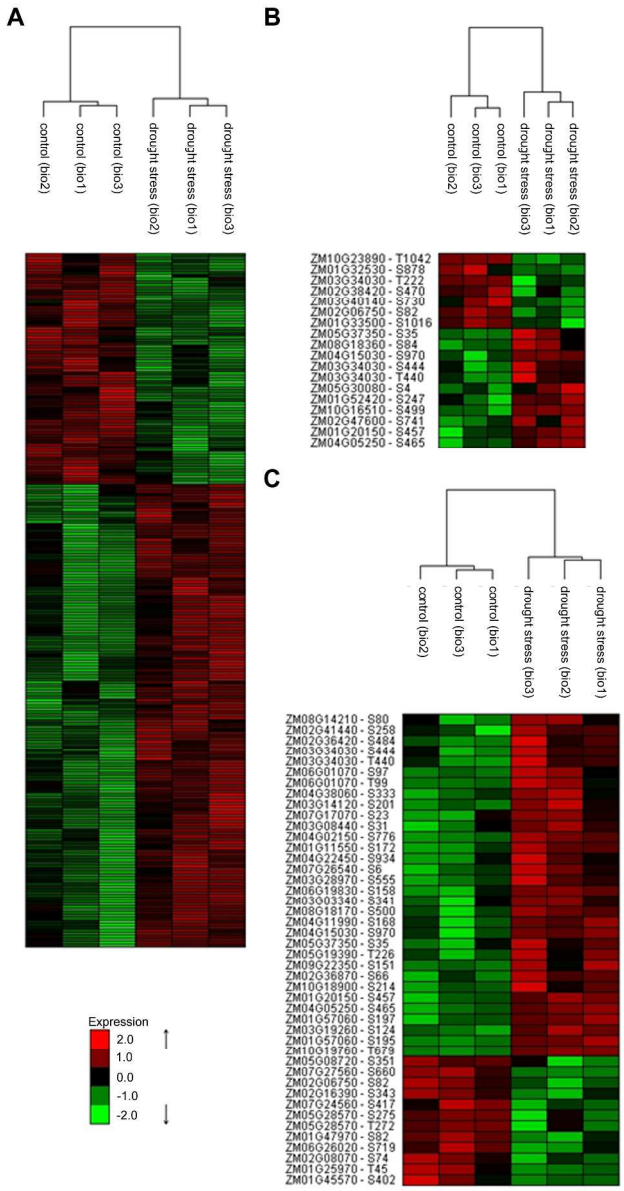

Figure 2. Differential (phospho)proteome response upon drought stress in maize leaves.
Heat map showing average log2 values of MaxLFQ intensities of the significantly differentially expressed proteins (A), phosphosite intensities of the significantly regulated phosphosites after normalization to protein levels (B), or phosphosite intensities of the significantly regulated phosphosites without correction for protein abundance (C). The log2 values of the intensities were Z-scored for graphical representation.

195x366mm (300 x 300 DPI)

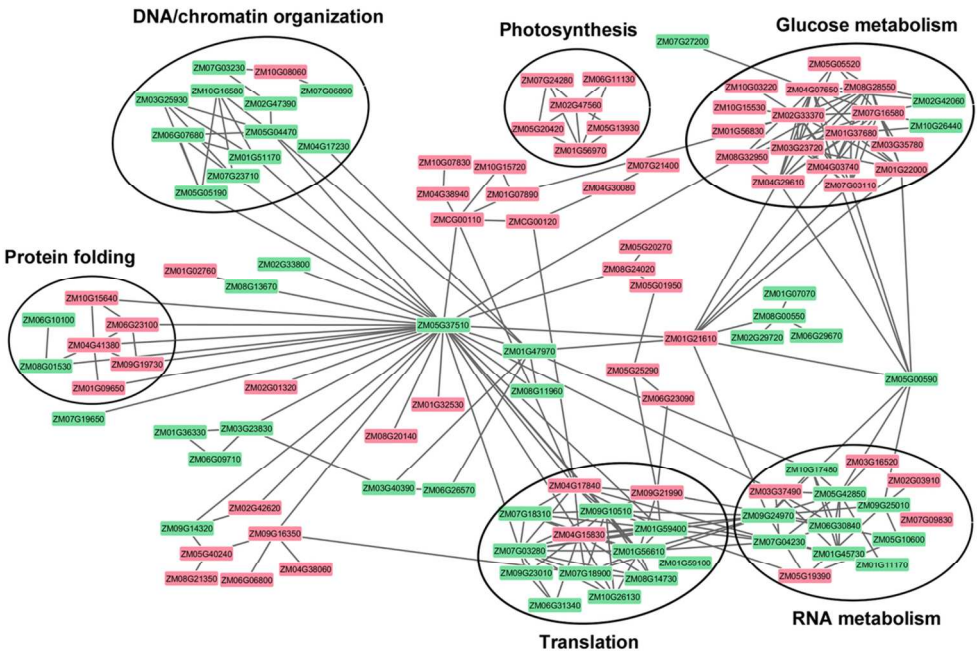

Figure 3. Protein-protein interaction networks resulted from STRING analysis of maize proteins with significant abundance changes. Cytoscape was used for visualization. Subnetworks with fewer than 6 interactors are excluded from the representation. Nodes in red represent upregulated proteins, in green downregulated proteins upon drought. Interaction groups are indicated with black circles. The PLAZA identifiers for maize proteins were used because of space constraints.
106x70mm (300 x 300 DPI)