

Research Article

IDENTIFICATION AND FUNCTIONAL ANNOTATION OF APOPLASTIC PHOSPHOPROTEINS OF *HIPPOPHAE RHAMNOIDES* SEEDLINGS

Ravi Gupta and Renu Deswal*

Molecular Plant Physiology and Proteomics Laboratory, Department of Botany, University of Delhi, Delhi-110007, India

Abstract: Phosphorylation is a reversible switch that regulates the biological activities of the proteins. Although there are ample of reports on the plant phosphoproteome analysis, phosphorylation status of apoplastic proteins has not been investigated profoundly. Here a shotgun proteomics approach was used to identify the phosphoproteins from the apoplast of the *Hippophae rhamnoides* (Seabuckthorn). A total of 123 phosphoproteins were identified using an SYNAPT G2 quadrupole time-of-flight mass spectrometer (Q-ToF-MS). Functional annotation of the identified phosphoproteins using PANTHER, Gene ontology, and KEGG programs showed that the majority of proteins were associated with the transporter, nucleic acid binding and amino acid metabolic activities. Prediction of secretory nature of the identified proteins using SignalP and SecretomeP servers showed that 56 % of the proteins were secretory, while rest of the 44 % of the proteins were non-secretory. PhosPhAt 4.0 detected 534 putative phosphorylation sites in the 75 unique *Arabidopsis* annotated proteins, wherein 195 (36%) were on the serine residue, 196 (37%) were on the threonine residue and 143 (27%) were detected on the tyrosine residue. Taken together, our results provide the first insight into the phosphorylation-mediated regulation of apoplastic proteins by cellular processes, which would be helpful in an in-depth understanding of the apoplastic signaling.

Keywords: Seabuckthorn; Phospho-proteome; Apoplast

Note : Coloured Figures and Supplementary Information available on Journal Website in "Archives" Section

Introduction

Phosphorylation is one of the most important and well-studied post-translational modifications (PTM) in plants, which regulates protein activity, their subcellular localization and stability (Novak *et al.*, 2010; Maures *et al.*, 2011). Phosphorylation is a reversible PTM, where addition and removal of phosphate groups to the specific amino acids of the proteins are catalyzed by protein kinases and phosphatases, respectively. In *Arabidopsis*, more than a thousand genes are predicted to encode protein kinases, while more than 100

genes encode phosphatases, suggesting a vast regulation of plant proteins by phosphorylation. It is estimated that more than one-third of the eukaryotic proteins undergo phosphorylation at one or more time points, representing that phosphorylation is the most abundant PTM in the eukaryotes. Given the key importance of phosphorylation in regulation of the cellular processes, plenty of reports have been published in the last decade on the analysis of the phosphoproteome of plants in response to growth and development, biotic and abiotic stress conditions (Rampitsch and Bykova, 2012). These reports have demonstrated the role of phosphorylation in fine tuning of any vital processes viz. cell signaling, cell trafficking, growth and development, responses to environment and many other factors (Schulze,

Corresponding Author: Renu Deswal
E-mail: rdeswal@botany.du.ac.in

Received: September 23, 2016

Accepted: November 28, 2016

Published: December 14, 2016

2010). However, the phosphorylation of RuBisCO, a major high-abundance protein in the leaves of green plants remains the major impediment for in-depth analysis of plant phosphoproteome (Slade *et al.*, 2014). Being highly phosphorylated, RuBisCO masks the identification of the other low abundant phosphorylated proteins, which may be of prime importance (Gupta and Kim, 2015). It has been previously shown that some of the key players of biological processes such as transcription factors, receptor kinases, and protein phosphatases are low abundance in nature, which can be masked by the RuBisCO (Slade *et al.*, 2014; Gupta *et al.*, 2015b). RuBisCO depletion and analysis of subcellular proteome are two major strategies, which are being used for the identification of RuBisCO free proteome (Gupta *et al.*, 2015b). It was shown that the depletion of the RuBisCO led to significant enrichment of the S-nitrosylated proteins in *Brassica juncea*, suggesting that the depletion of RuBisCO can lead to the identification of novel low abundance proteins (Sehrawat *et al.*, 2013), however, the impact of RuBisCO depleteome in phosphoproteome analysis has not been investigated yet. There have been reports on the analysis of phosphoproteome of subcellular organelles including plasma membrane (Whiteman *et al.*, 2008a), mitochondria (Ito *et al.*, 2009), chloroplast (Reiland *et al.*, 2009), vacuolar membrane (Whiteman *et al.*, 2008b) and nucleus (Kumar *et al.*, 2014; Bigeard *et al.*, 2014). Analysis of phosphoproteins of chickpea seedlings led to the identification of 107 putative phosphoproteins, primarily involved in protein folding, signaling and gene regulation, DNA replication, repair and modification, and metabolism (Kumar *et al.*, 2014). Recently, Bigeard and co-workers analyzed the phosphorylation of the chromatin-associated proteins to get a deeper insight into the nuclear phosphoproteins, which led to the identification of 198 phosphoproteins in *Arabidopsis thaliana* (Bigeard *et al.*, 2014). The identified proteins were mainly involved in the chromatin remodeling, transcriptional regulation, and RNA processing (Bigeard *et al.*, 2014). These reports suggest that subproteome analysis can indeed provide an in-depth overview of the post-translational events in plants.

Despite the phosphoproteome analysis of various subcellular organelles, lesser reports have been published so far specifically on the analysis of the phosphoproteome of the apoplast (Pechanova *et al.*, 2010; Casasoli *et al.*, 2008). Apoplastic proteins are not only involved in the gene regulation but are also the crucial components of the signal perception, signal transduction and cell to cell communication. Therefore, analysis of phosphoproteome of the apoplast can provide a better understanding and regulation of the above-mentioned processes. One of the major challenges in the apoplast proteome analysis is the isolation of pure apoplastic proteins without cytoplasmic contamination. Previously, we developed an efficient protocol for the isolation of apoplastic proteins from *Hippophae rhamnoides* (seabuckthorn) seedlings with increased protein yield and negligible intracellular contamination (Gupta and Deswal, 2012; Gupta and Deswal, 2014). *H. rhamnoides* is a Himalayan shrub, which can face multiple stress conditions efficiently including low temperature, drought and high UV-stress (Juvany and Munné-Bosch, 2015), thus offers a good platform to understand the abiotic stress signaling in plants. Although the genome of the seabuckthorn has not been sequenced, attempts have been made to analyze its transcriptome and proteome in search of cold-responsive transcripts/proteins (Ghangal *et al.*, 2012; Gupta and Deswal, 2012). Initially, Ghangal and coworkers used an expressed sequence tag (EST)-based approach to search the differentially expressed genes in response to cold/freezing stress (Ghangal *et al.*, 2012). Subsequently, Gupta and Deswal employed two-dimensional electrophoresis (2DE) approach and identified 32 cold/freezing-induced apoplastic proteins (Gupta and Deswal, 2012). Recently a total of 11922 differentially expressed genes including 6539 up-regulated and 5383 down-regulated genes were identified using DeepSAGE based approach (Chaudhary and Sharma, 2015). Because of the unique stress tolerance capacity, the number of studies on the seabuckthorn is increasing rapidly.

In this study, we report the identification and functional annotation of apoplastic phosphoproteins of seabuckthorn seedlings. We used a shotgun proteomics approach to identify

the phosphoproteins in the apoplast, which were functionally annotated using KEGG, Gene Ontology, and PANTHER programs. A putative interaction network was also predicted to provide a deeper insight into the functional roles of the identified phosphoproteins.

Materials and Methods

Plant Growth Conditions and Isolation of apoplastic proteins

Seabuckthorn seeds were germinated as described previously (Gupta and Deswal, 2012) and 20 days old seedlings were used in this study. Apoplastic proteins were isolated using a vacuum infiltration protocol as described previously (Gupta and Deswal, 2012). In brief, seedlings were cut into small segments of about 2 cm and washed thoroughly with deionized water to remove cytoplasmic proteins from the cut ends. Segments were then incubated with ice-cold apoplastic protein extraction buffer containing 20 mM ascorbic acid, 20 mM calcium chloride and 1 mM sodium orthovanadate and 1 mM sodium fluoride for 20 min under vacuum. Finally, apoplastic proteins were extracted by centrifuging the vacuum infiltrated extraction buffer at 4000 g (Gupta and Deswal, 2012; Gupta and Deswal, 2014).

Phosphoprotein enrichment

Phosphoproteins were enriched using TALON® PMAC Phosphoprotein Enrichment Kit, Clonetech, as per manufacturer's protocol. In brief, extracted apoplastic proteins were desalted and de-lipidated using a methanol-chloroform procedure and the obtained pellet was dissolved in Buffer-A. Proteins concentration of the isolated proteins was quantified using 2D Quant kit (GE Healthcare, Uppsala, Sweden). A total of 4 mg of apoplastic proteins were loaded onto the Phosphate Metal Affinity Chromatography (PMAC) column which was pre-equilibrated with Buffer-A and incubated at 4 °C for 30 min with constant shaking. After incubation, the column was washed with Buffer-A and phosphoproteins were eluted from the column using Buffer-B (20 mM sodium phosphate in 500 mM KCl).

In-solution Trypsin digestion and Mass spectrometry

Enriched phosphoproteins from three biological replicates were pooled together, desalted using methanol-chloroform procedure and the obtained pellet was dissolved in lysis buffer containing 6 M Urea in 50 mM Ammonium bicarbonate (ABC). Proteins in the lysis buffer were quantified using 2D-quant kit and 100 µg of the proteins were used for in-solution trypsin digestion. Prior to digestion, proteins were reduced using 0.5 mM dithiothreitol (DTT) in 50 mM ABC by incubating at 37 °C for 1 hr and then alkylated using 1 mM iodoacetamide in 50 mM ABC at RT for 1 hr. Reduced and alkylated proteins were then diluted to 10 fold to reduce the concentration of Urea to 0.6 M. Proteins were then subjected to trypsin digestion using 2 µg of trypsin (Trypsin gold, Promega) by incubating at 37 °C overnight. Trypsin digestion was stopped using 1 % formic acid. The samples were dried under vacuum and reconstituted in 0.1% v/v formic acid and 2% v/v ACN. Nanoscale LC separation of the tryptic digested peptides was carried out using a nanoAcquity Ultra pressure liquid chromatography (UPLC) system (Waters). The first-dimensional separation was carried out on a Waters Acquity UPLC system with a photodiode array detector (Waters, Milford, MA) using a 250 mm × 2.1 mm Hypersil GOLD aQ 5 µm C18 column (Thermo, CA). Digested peptides were redissolved in 40 µL of mobile phase A (25 mM ammonium formate (NH_4FA) aqueous buffer, pH 7.5). Mobile phase B was 25 mM NH_4FA in water/acetonitrile (1:9). The gradient elution was performed by 0-10% B (0-80 min) and 10-35% B (80-90 min), and fractions were collected every 1 min. The fractions were pooled as required.

Analysis of the eluted tryptic peptides was performed using a Synapt G2 Q-TOF (quadrupole time-of-flight) mass spectrometer (Waters) equipped with a nano-lock spray source (Waters) fitted with a pico-tip emitter (New Objective) operated at a capillary voltage of approximately 3 kV. The MS data obtained from the LC-MSE analysis were processed and searched using the ProteinLynx Global Server (PLGS) version 2.5 (Waters, Manchester, UK). The obtained MS-MS

spectra were searched against Viridiplanteae database (UniProt, released in October 2013) and only high confidence identifications ($p < 0.01$) were considered. For protein identification following search parameters were used: for mass accuracy “automatic” setting was used (approximately 10 ppm for precursor ions and 25 ppm for product ions); a minimum of one peptide match per protein, one missed cleavage site was allowed; carbamidomethylation of cysteine was set as fixed modification; and methionine oxidation along with the phosphorylation on Serine, Threonine, or Tyrosine was set as variable modification. Protein identification data was exported as .csv files and imported into Microsoft Excel.

Prediction of the extracellular localization and phosphorylation sites in the identified Proteins

For predicting the secretory nature of the identified phosphoproteins, complete amino acid sequences were downloaded in FASTA format and were used for the prediction using SignalP (ver. 4.1) and SecretomeP (ver. 2.0) servers. For the prediction of the phosphorylation sites, PhosPhAt (ver. 4.0) server was used.

Functional annotation and interactome prediction of the identified proteins

Orthologs of the identified phosphoproteins were searched in *Arabidopsis* using Phytozome 10.3 (http://phytozome.jgi.doe.gov/pz/portal.html#linfo?alias=Org_Athaliana), which is linked to the TAIR10 (The *Arabidopsis* Information Resource). The orthologous proteins were functionally annotated using Kyoto Encyclopedia of Genes and Genomes (KEGG), Gene Ontology and PANTHER programs. For analyzing the interaction between the identified phosphoproteins, BAR (The Bio-Analytic Resource for Plant Biology, <http://bar.utoronto.ca/>) program was used.

Results and Discussion

Here, we used a shot-gun proteomics approach to identify the phosphoproteins in the apoplast of seabuckthorn. Apoplastic proteins were isolated and phosphoproteins were enriched using a PMAC phosphoprotein enrichment kit

(Figure 1). The PMAC resin is highly specific to the phosphate groups, thus allowing the binding of the phosphorylated proteins only. Other non-phosphorylated proteins and contaminants pass through in the flow-through and wash fractions. The efficiency and specificity of the PMAC have already been shown using a Western Blotting approach (Kinoshita-Kikuta *et al.*, 2006). For the identification of the phosphoproteins, phosphopeptide enrichment can also be performed. In the case of phosphopeptide enrichment, the identification of the respective protein relies only on one specific phosphopeptide, rather than the multiple peptides (Wolschin and Weckwerth, 2008). However, it is always better to identify proteins with multiple peptide spectral match (PSM) than with only one PSM which will lead to inaccuracy. Moreover, this condition would be very stringent for the identification of proteins from an unsequenced plant like seabuckthorn. Therefore, here we used phosphoprotein enrichment instead of phosphopeptide enrichment for analyzing the phosphoproteome of seabuckthorn apoplast (Figure 1).

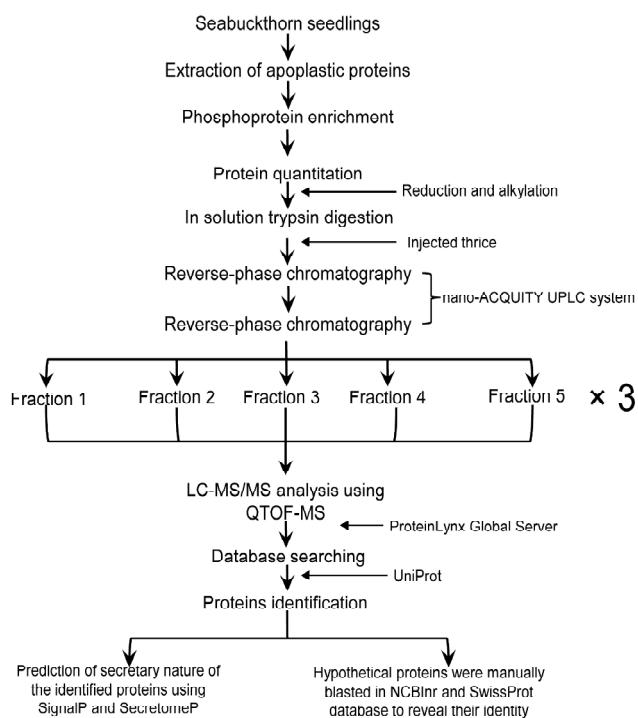


Figure 1: Flow chart of the methodology used for the enrichment and identification of phosphoproteins in seabuckthorn apoplast.

As compared to the phosphopeptide enrichment, phosphoproteins enrichment leads to the generation of multiple peptides thus increases the sample complexity, which can be a major hurdle for phosphoproteome analysis. Therefore, in order to reduce the sample complexity, here we used nanoACQUITY UPLC System with 2D Technology for peptide separation. This system expands the use of sub-2-micron particles to achieve high peak capacity separations. This system effectively uses 2D-UPLC for better chromatographic resolution of the complex proteomic samples by using a dual reversed-phase approach, unlike conventional 2D-LCs which first uses a strong cation exchange followed by reversed-phase separations. The drawback of strong cation exchange approach is the use of salt-containing buffers which causes ionization background and fouling problems when these enter the mass spectrometer. Moreover, strong cation exchange separations are solely based on the charge of the peptide, therefore, strong cation exchange often results in poor chromatographic resolution with the peptides appearing in multiple fractions, causing data interpretation difficult. In contrast to the conventional LCs, the nanoACQUITY UPLC system employs a two reversed-phase separations at high and low pH, thus providing high-resolution separations in both the directions instead of one low-resolution (typically strong cation exchange) and one high-resolution separation (typically reverse phase). UPLC provides better reproducibility, larger dynamic range and dramatically less fraction-to-fraction carryover for protein analyses (Wu *et al.*, 2006) (Figure 1).

Identification of the phosphoproteins

Phosphoprotein enrichment followed by shotgun proteomics approach led to the identification of 123 unique phosphoproteins in the apoplast of seabuckthorn seedlings (Supplementary Table 1). The molecular weights and pIs of the identified proteins ranged from 3.5 kDa to 142 kDa and 3.4 to 12.15, respectively. Using this shotgun proteomics approach, we have identified 16 proteins with extreme high and low molecular weight (<10 and >100), which would have been generally lost in the gel-based proteomics approaches (Gupta *et al.*, 2015b). Among the

identified proteins, a maximum number of proteins (18%) had molecular weights in the range of 21 to 30 kDa (Figure 2A), while more than 50% of the identified proteins had a pI value less than 7.0 (Figure 2B).

A major portion of the identified proteins reflects leaderless secretory proteins

In order to identify the secretory nature of the identified apoplastic phosphoproteins, SignalP and SecretomeP servers were used. While SignalP predicts the extracellular localization of proteins based on the presence of N-terminal signal peptide, SecretomeP predicts based on the presence or absence of a membrane anchoring domain (Bendtsen *et al.*, 2004). Results of SignalP and SecretomeP showed that 56% of the identified proteins were secretory; while rest 44% of the proteins was not secretory. Out of the total 56% secretory proteins, only 22% were found to have an N-terminal signal peptide and thus are supposed to be secreted via classical Golgi-endoplasmic reticulum pathway (Figure 2C, Supplementary Table 1). Rest 34% of the secretory proteins lacked a signal peptide and is supposed to be secreted by non-classical protein secretion pathway(s). Previously, it was shown that the percentage of the non-secretary proteins in apoplast vary in response to tissue type and stress conditions and can constitute up to 80% of the total apoplastic proteins (Agrawal *et al.*, 2010). These proteins may be pumped to the apoplast by unknown mechanism(s) and could play crucial functions, which are yet to be discovered. These proteins may have different functions in the symplast and apoplast and are well known as leaderless secretory proteins (LSPs). Hexokinase, Presenilin, malate dehydrogenase and superoxide dismutase are some of the well characterized LSPs in plants (Huberts and van der Klei, 2010; Gupta *et al.*, 2015a). Of these, superoxide dismutase and malate dehydrogenase have been identified in the apoplast though none of these contain neither a signal peptide nor a membrane anchoring domain (Cheng and Williamson, 2010).

Prediction of phosphorylation sites

Specific phosphorylation sites of the identified proteins were predicted using PhosPhAt 4.0

Table 1
List of identified phosphoproteins from seabuckthorn apoplast

<i>Functional classification</i>	<i>Accession</i>	<i>Organism</i>	<i>Description</i>	<i>m/zV (Da)</i>	<i>pI (pH)</i>	<i>PLGS Score</i>	<i>Peptides</i>	<i>Theoretical Peptides</i>	<i>Loc Peptides</i>
Cell wall associated	G7J699	<i>Medicago truncatula</i>	Arabinogalactan peptide	6129	5.0962	247.6694	1	1	SP
Cell wall associated	Q41072	<i>Pinnus taeda</i>	Arabinogalactan like protein	26585	6.0791	175.4299	2	7	SP
Cell wall associated	Q43682	<i>Vigna unguiculata</i>	Extensin-like-protein	23011	6.7134	62.3883	2	12	NS
Cell wall associated	Q9FL53	<i>Arabidopsis thaliana</i>	Fasciclin-like arabinogalactan protein 21	38677	6.4819	55.3427	2	22	SP
Cell wall associated	M8C0S0	<i>Aegilops tauschii</i>	Callose synthase 7	62467	9.6958	23.2959	2	33	NC
Defence/Stress related	EOY19635.1	<i>Theobroma cacao</i>	MATH and LRR domain-containing protein PFIE0570w, putative isoform 2	31102	4.2056	83.3723	1	18	NC
Defence/Stress related	G3BGS8	<i>Gladiolus grandiflorus</i>	Chalcone synthase 1	43058	5.5649	58.4415	1	29	NS
Defence/Stress related	B0LQT3	<i>Vitis vinifera</i>	MLO like protein	57845	9.3267	52.1679	1	40	NS
Defence/Stress related	Q652P9	<i>Oryza sativa</i> subsp. <i>Japonica</i>	Germin-like protein 9-3	21666	5.8301	49.7595	2	11	SP
Defence/Stress related	G7J7U2	<i>Medicago truncatula</i>	Pentatricopeptide repeat-containing protein	47860	7.1997	31.1185	1	31	NC
Defence/Stress related	B3ITB8	<i>Zingiber zerumbet</i>	Sesquiterpene synthase 3	26828	4.9351	28.7754	1	25	NS
Defence/Stress related	B6THF7	<i>Zea mays</i> (Maize)	Glucan endo-1,3-beta-glucosidase A6	52395	9.6138	25.2891	1	29	SP
Defence/Stress related	B9S164	<i>Ricinus communis</i>	WRKY transcription factor, putative	30896	7.9731	23.7221	1	17	NC
Defence/Stress related	M8AQ15	<i>Aegilops tauschii</i>	Putative LRR receptor-like serine/threonine-protein kinase	140911	5.2983	23.6121	3	80	SP
Defence/Stress related	Q5JMK2	<i>Oryza sativa</i> subsp. <i>Japonica</i>	Putative xyلانase inhibitor	47624	6.0615	21.8103	1	22	SP
Defence/Stress related	XP_004231703.1	<i>Solanum lycopersicum</i>	PREDICTED: probable disease resistance protein At4g27220-like	87522	5.436	21.4435	2	62	NS
Defence/Stress related	AAM94867.1	<i>Brassica napus</i>	Polygalacturonase inhibitor protein	65513	5.4097	19.7164	1	45	SP
Gene regulation	XP_004250825.1	<i>Solanum lycopersicum</i>	PREDICTED: pre-mRNA-processing protein 40A-like	5609	10.0913	189.0878	1	2	NS
Gene regulation	Q9BAZ0	<i>Stanhopea cirrhata</i>	Maturase K Fragment	61006	9.9404	131.6491	2	34	NS

contd. table 1

Functional classification	Accession	Organism	Description	mNW (Da)	pI (pH)	PLGS Score	Peptides	Theoretical Peptides	Loc
Gene regulation	G7JMT9	<i>Medicago truncatula</i>	Zinc finger CCHC domain containing protein	29234	10.355	127.3774	1	17	NS
Gene regulation	A7UQU9	<i>Medicago truncatula</i>	Homeodomain-like	47369	7.9131	91.6946	1	29	NS
Gene regulation	B9RNS4	<i>Ricinus communis</i>	Zinc finger protein, putative	28874	7.7578	67.3689	1	20	NS
Gene regulation	XP_004968688.1	<i>Setaria italica</i>	PREDICTED: retrotransposon-like protein 1-like	23493	3.4424	34.7904	1	17	NS
Gene regulation	A9T9D3	<i>Physcomitrella patens</i> subsp <i>patens</i>	Histone H2A	14586	10.8149	30.7958	1	11	NC
Gene regulation	K4JBT1	<i>Zea mays</i> (Maize)	MYB-related transcription factor	32380	9.3984	27.6456	1	23	NS
Gene regulation	B9RFG0	<i>Ricinus communis</i>	Splicing factor, arginine/ serine-rich 2, map interacting protein, putative	104869	4.6567	25.8013	2	69	NS
Gene regulation	B9RDD7	<i>Ricinus communis</i>	DNA binding protein, putative	33460	9.8965	23.6549	1	21	NS
Gene regulation	B8PZJ5	<i>Hordeum vulgare</i> var. <i>distichum</i>	C-repeat binding factor 6 (Predicted protein)	46984	4.7622	22.6526	1	22	NS
Gene regulation	D7KHS5	<i>Arabidopsis lyrata</i> subsp <i>lyrata</i> (Lyre-leaved rock-cress)	Aminoacyl-tRNA synthetase family	84751	4.5015	22.2606	2	61	NC
Gene regulation	NP_001152256.1	<i>Zea mays</i> (Maize)	Transposon protein	35144	6.854	22.1497	1	27	SP
Gene regulation	A7WPA2	<i>Oryza sativa</i> subsp. <i>Indica</i>	Putative pre-mRNA-splicing factor cwc-22	101069	5.3672	22.0485	2	84	NS
Gene regulation	A7J9L7	<i>Larix x marschallii</i>	APETALA2 L1	58250	6.4585	20.6324	1	39	NS
Gene regulation	G7IRL0	<i>Medicago truncatula</i>	Zinc finger protein CONSTANS-like protein	19187	4.1411	19.8234	1	10	NC
Metabolism	A2XGF8	<i>Oryza sativa</i> subsp. <i>Indica</i>	Reticulon-like protein	25180	9.6724	105.4794	1	15	NC
Metabolism	EOY23908.1	<i>Theobroma cacao</i>	NADP-specific glutamate dehydrogenase, putative	28646	5.0098	86.1053	1	14	NC
Metabolism	B9SDH2	<i>Ricinus communis</i>	Cytochrome P450 putative	59747	9.3794	60.3027	2	39	SP
Metabolism	F6LR33	<i>Dendrobium officinale</i>	Sucrose phosphate synthase	118819	5.8975	44.1611	2	77	NS
Metabolism	B9RA05	<i>Ricinus communis</i>	ATP binding protein, putative	13140	9.772	43.5551	1	5	SP

contd. table 1

<i>Functional classification</i>	<i>Accession</i>	<i>Organism</i>	<i>Description</i>	<i>m/zV (Da)</i>	<i>pI (pH)</i>	<i>PLGS Score</i>	<i>Peptides</i>	<i>Theoretical Peptides</i>	<i>Loc</i>
Metabolism	M4FEZ0	<i>Brassica rapa</i> subsp <i>pekinensis</i>	Malate dehydrogenase	35735	8.5342	39.7773	2	22	NC
Metabolism	Q85GD3	<i>Weigela florida</i>	NADH dehydrogenase subunit F Fragment	58110	8.8066	37.765	1	20	SP
Metabolism	A9RJX3	<i>Physcomitrella patens</i> subsp <i>patens</i>	Reticulon-like protein	24618	9.7251	32.2724	1	15	NS
Metabolism	M1C5Y9	<i>Ricinus communis</i>	Aldose 1 epimerase	37030	6.0601	29.8198	1	22	NC
Metabolism	O04866	<i>Alnus glutinosa</i>	Acetylornithine aminotransferase	48801	8.3481	28.1374	1	37	NS
Metabolism	B9R7Y3	<i>Ricinus communis</i>	ATP binding protein, putative	42327	7.0796	26.6486	1	30	NC
Metabolism	NP_195146.1	<i>Arabidopsis thaliana</i>	D-3-phosphoglycerate dehydrogenase	24927	6.1787	25.5989	1	15	NS
Metabolism	B9S6M2	<i>Ricinus communis</i>	Adenosine diphosphatase, putative (EC 3.6.1.5)	35225	9.3721	25.254	2	22	NC
Metabolism	G7ISN0	<i>Medicago truncatula</i>	Glycosyl transferase, family 8	75945	9.186	20.0668	2	53	SP
Metabolism	B9IE14	<i>Populus trichocarpa</i>	Galactosyltransferase family protein	76760	7.0649	18.5148	2	54	NC
Photosynthesis	I1SZM7	<i>Sagittaria guayanensis</i>	Ribulose bisphosphate carboxylase large chain Fragment	42956	6.5684	57.1812	1	29	NC
Photosynthesis	P41600	<i>Pinus thunbergii</i>	Photosystem II reaction center protein Ycf12	3505	7.0151	723.4255	1	3	-
Protein metabolism	A9UFX8	<i>Vitis vinifera</i>	Cysteine protease	41436	5.48	48.5512	2	20	SP
Protein metabolism	A9NLJ2	<i>Picea sitchensis</i>	Proteasome subunit beta type	25122	5.0288	37.6108	1	21	NC
Protein metabolism	G7K6H0	<i>Medicago truncatula</i>	E3 ubiquitin-protein ligase synoviolin	56828	6.6035	36.6534	1	40	SP
Protein metabolism	C5XZM5	<i>Aegilops tauschii</i>	E3 ubiquitin-protein ligase SINA-like protein 4	84439	5.7979	35.9717	2	51	SP
Protein metabolism	Q06H22	<i>Arachis hypogaea</i>	Ubiquitin-conjugating enzyme 9499	4.5923	30.2799	1	6	NC	
Protein metabolism	B6UA54	<i>Zea mays</i> (Maize)	Ubiquitin-protein ligase	40080	8.0171	20.3725	1	25	NS
ROS-regulation	G7J5C2	<i>Medicago truncatula</i>	NAD(P)H-quinone oxidoreductase chain	55121	6.7222	63.7087	1	25	NS
ROS-regulation	E7D6J0	<i>Phlox drummondii</i>	Anthocyanin synthase Fragment	14685	5.6514	55.3105	1	9	NC

contd. table 1

Functional classification	Accession	Organism	Description	mNW (Da)	pI (pH)	PLGS Score	Peptides	Theoretical Peptides	Loc
ROS-regulation	A9XUK7	<i>Albizia julibrissin</i>	NADH plastoquinone oxidoreductase subunit 5 Fragment	77665	8.7671	51.6066	1	28	SP
ROS-regulation	D7MVG4	<i>Arabidopsis lyrata</i> subsp. <i>lyrata</i> (Lyre-leaved rock-cress)	Oxidoreductase	34538	8.9385	40.0362	1	25	NS
ROS-regulation	A1E9S8	<i>Sorghum bicolor</i>	NADPH quinone oxidoreductase subunit K chloroplastic	25465	9.082	35.8397	1	13	NC
ROS-regulation	H6VLE8	<i>Salvia miltiorrhiza</i>	3 hydroxy 3 methylglutaryl coenzyme A reductase 3	60457	5.6309	25.5086	1	24	NC
ROS-regulation	G7KRU3	<i>Medicago truncatula</i>	GDP-L-fucose synthase	35627	6.6665	22.2157	1	33	NC
Signal transduction	D7LPP6	<i>Arabidopsis lyrata</i> subsp. <i>lyrata</i> (Lyre-leaved rock-cress)	Calmodulin-binding protein	20508	4.3843	383.0636	2	15	NC
Signal transduction	B9S727	<i>Ricinus communis</i>	Protein binding protein, putative	91488	8.9692	207.6122	4	80	NS
Signal transduction	EOY06955.1	<i>Theobroma cacao</i>	Phosphatase methyltransferase 1	8528	9.4395	112.734	1	7	NC
Signal transduction	O65268	<i>Arabidopsis thaliana</i>	Putative threonine aspartate Two-component response regulator APR1	42599	6.0059	99.0023	3	25	NS
Signal transduction	G7JCG7	<i>Medicago truncatula</i>	Probable cytokinin riboside 5'-monophosphate phosphoribohydrolase LOG2 (EC 3.2.2.n1)	35261	5.8916	60.2794	1	18	NS
Signal transduction	B9F166	<i>Oryza sativa</i> subsp. <i>japonica</i>	26766	6.1055	56.9411	3	18	NC	
Signal transduction	Q0GZQ0	<i>Armenia euchroma</i>	Calnexin Fragment	18677	4.9263	50.6654	1	10	NS
Signal transduction	G7LC60	<i>Medicago truncatula</i>	TSL-kinase interacting protein	59408	8.7129	45.3741	1	42	NC
Signal transduction	Q9LVQ9	<i>Arabidopsis thaliana</i>	Protein kinase ATN1-like protein (protein kinase family protein)	45489	7.9805	38.0627	1	35	NS
Signal transduction	XP_003551570.1	<i>Glycine max</i>	PREDICTED: receptor-like protein 12-like	24100	4.2773	34.1636	1	23	SP
Signal transduction	K7V133	<i>Zea mays</i> (Maize)	Putative leucine-rich repeat receptor-like protein kinase family protein	68624	7.4941	30.9461	1	43	SP

contd. table 1

Functional classification	Accession	Organism	Description	<i>mNW</i> (Da)	<i>pI</i> (<i>pH</i>)	PLGS Score	Peptides	Theoretical Peptides	Loc
Signal transduction	Q5Z6A8	<i>Oryza sativa</i> subsp. <i>japonica</i>	Calcineurin B-like	14474	12.1538	30.765	1	12	NC
Signal transduction	Q948J4	<i>Oryza sativa</i> subsp. <i>japonica</i>	Cryptochrome 1b	80299	5.4507	30.3041	1	52	NS
Signal transduction	G7K578	<i>Medicago truncatula</i>	Indole-3-acetic acid-amido synthtase GH3 3	20260	10.3286	26.3139	1	14	NS
Signal transduction	G7KHQ6	<i>Medicago truncatula</i>	Auxin induced protein 5NC4	39983	9.3911	24.6377	2	21	NC
Signal transduction	G7IKL2	<i>Medicago truncatula</i>	Xenotropic and polytopic retrovirus receptor	45855	8.2632	23.5976	2	24	NC
Signal transduction	F4HXPO	<i>Arabidopsis thaliana</i>	Putative ADP ribosylation factor GTPase activating protein AGD14	71515	5.2749	22.5128	1	32	NC
Signal transduction	Q6Z294	<i>Oryza sativa</i> subsp. <i>japonica</i>	Putative ankyrin repeat protein family	85576	7.5176	21.1118	1	56	NS
Signal transduction	M8BIU9	<i>Aegilops tauschii</i>	Putative serine/theronine-protein kinase	32708	7.5249	20.5905	2	23	NS
Signal transduction	G7L9Z1	<i>Medicago truncatula</i>	Rac GTPase activating protein	55274	5.0581	19.7804	1	40	NC
Transport	Q940Q3	<i>Arabidopsis thaliana</i>	Zinc transporter ZTP29	31028	9.4863	111.7472	1	10	SP
Transport	B6T3V3	<i>Zea mays</i> (Maize)	Mitochondrial import receptor subunit TOM22	9495	10.04	79.3082	2	12	NS
Transport	XP_004974274.1	<i>Setaria italica</i>	PREDICTED; mitochondrial import receptor subunit TOM9-2	23761	5.8125	68.1975	2	18	NS
Transport	B9RKKN2	<i>Ricinus communis</i>	Sugar transporter, putative (EC 1.3.1.74)	51549	8.1943	62.5486	1	24	NS
Transport	C1FHE3	<i>Micromonas sp strain RCC299</i>	Resistance modulation cell division superfamily	139059	6.6709	37.1817	1	85	SP
Transport	B9RJS5	<i>Ricinus communis</i>	Purine transporter, putative	54868	9.0908	34.3031	2	28	NS
Transport	XP_004983525.1	<i>Setaria italica</i>	PREDICTED; plastidal glycolate/glycerate translocator 1, chloroplastic-like	54899	9.5845	31.2998	3	28	NC
Transport	Q851R0	<i>Oryza sativa</i> subsp. <i>japonica</i>	Nodulin family protein, putative, expressed (Putative nodule-specific protein)	64881	8.1987	30.7082	3	28	NS

contd. table 1

Functional classification	Accession	Organism	Description	mNW (Da)	pI (pH)	PLGS Score	Peptides	Theoretical Peptides	Loc
Transport	B9IAM2	<i>Populus trichocarpa</i>	Lipid-binding serum glycoprotein	48342	6.1831	27.5955	2	33	SP
Transport	C9WB11	<i>Stellaria longipes</i>	ABC transporter Fragment	25237	9.6401	26.7881	1	12	NS
Transport	F4JMN3	<i>Arabidopsis thaliana</i>	SNARE associated Golgi protein family	33386	5.6704	23.2836	1	19	NC
Transport	F4HQ84	<i>Arabidopsis thaliana</i>	Sec34-like protein	88720	5.7832	23.0703	1	52	NC
Transport	B6TYM7	<i>Zea mays</i> (Maize)	Folate/biopterin transporter family protein	58458	9.3457	21.9014	1	26	NC
Unknown	B4FHE2	<i>Zea mays</i> (Maize)	Uncharacterized protein	19234	6.646	322.7742	2	13	NS
Unknown	G7KV91	<i>Melilotus truncatula</i>	High mobility group protein	16443	4.8208	173.0444	1	13	NS
Unknown	M1DTQ9	<i>Solanum tuberosum</i>	Uncharacterized protein	17698	4.4663	125.0425	2	23	NS
Unknown	C1EFF4	<i>Micromonas sp strain RCC299</i>	Predicted protein	19581	4.396	124.2918	1	12	NS
Unknown	M4C7D2	<i>Brassica rapa</i> subsp <i>pekinensis</i>	Uncharacterized protein	53901	9.3208	70.3352	1	32	NC
Unknown	A8HM88	<i>Chlamydomonas reinhardtii</i>	Predicted protein	11418	10.3345	66.6237	1	9	NS
Unknown	M0XWT0	<i>Hordeum vulgare</i>	Uncharacterized protein	33073	9.2461	51.4121	2	20	SP
Unknown	Q6Z4B6	<i>Oryza sativa</i> subsp <i>japonica</i>	Os07g0510600 protein	16914	11.8345	49.6805	1	11	NC
Unknown	D7LDQ5	<i>Arabidopsis lyrata</i> subsp <i>lyrata</i>	Putative uncharacterized protein	47887	8.6147	46.3273	2	37	NS
Unknown	B9SWA6	<i>Ricinus communis</i>	Putative uncharacterized protein	13942	3.9609	45.8141	1	14	SP
Unknown	M4E7W6	<i>Brassica rapa</i> subsp <i>pekinensis</i>	Uncharacterized protein	20135	4.6934	39.9279	1	16	NS
Unknown	I0YK31	<i>Coccomyxa subellipsoidea</i>	Uncharacterized protein	118442	10.4546	39.8281	2	74	NS
Unknown	K3XLK6	<i>Setaria italica</i>	Uncharacterized protein	24930	5.7715	37.8957	1	21	NS
Unknown	R7W205	<i>Aegilops tauschii</i>	Uncharacterized protein	29019	4.9585	32.7335	1	26	NS
Unknown	C1MVX7	<i>Micromonas pusilla</i> strain CCMP1545	Predicted protein	41333	4.3184	27.3655	1	36	NS
Unknown	B9N71	<i>Populus trichocarpa</i>	Predicted protein	15766	8.1313	24.9082	1	12	NC

contd. table 1

<i>Functional classification</i>	<i>Accession</i>	<i>Organism</i>	<i>Description</i>	<i>mNW (Da)</i>	<i>pI (pH)</i>	<i>PLGS Score</i>	<i>Peptides</i>	<i>Theoretical Peptides</i>	<i>Loc</i>
Unknown	M4DB35	<i>Brassica rapa</i> subsp <i>pekinensis</i>	Uncharacterized protein	9649	9.8232	24.7878	1	6	NC
Unknown	B9GNW4	<i>Populus trichocarpa</i>	Dem family protein	142100	5.5664	23.6589	2	91	NS
Unknown	M0SLM0	<i>Musa acuminata</i> subsp <i>malaccensis</i>	Uncharacterized protein	16144	8.7085	23.5581	1	9	NC
Unknown	B8AQL2	<i>Oryza sativa</i> subsp <i>indica</i>	Putative uncharacterized protein	59866	4.8296	22.3831	1	38	NC
Unknown	Q8S0U3	<i>Oryza sativa</i> subsp <i>japonica</i>	Putative uncharacterized protein	22433	6.1831	20.544	1	10	SP
Unknown	Q9M016	<i>Arabidopsis thaliana</i>	Putative uncharacterized protein B1144D11 16	64911	5.2134	20.3447	1	49	NS
Unknown	D8UK90	<i>Volvox carteri</i>	Putative uncharacterized protein	24757	4.5454	19.662	1	13	SP
Unknown	A2YRB7	<i>Oryza sativa</i> subsp <i>indica</i>	Putative uncharacterized protein	102567	9.9858	13.1314	2	84	NS
Unknown	K4BHK8	<i>Solanum lycopersicum</i>	Uncharacterized protein	26437	10.1089	318.7148	2	18	NC
Unknown	Q53MP7	<i>Oryza sativa</i> subsp <i>japonica</i>	Putative uncharacterized protein	7766	3.668	237.6175	1	8	NC
Unknown	J3MW18	<i>Oryza brachyantha</i>	Uncharacterized protein	42439	4.812	231.6535	1	19	SP

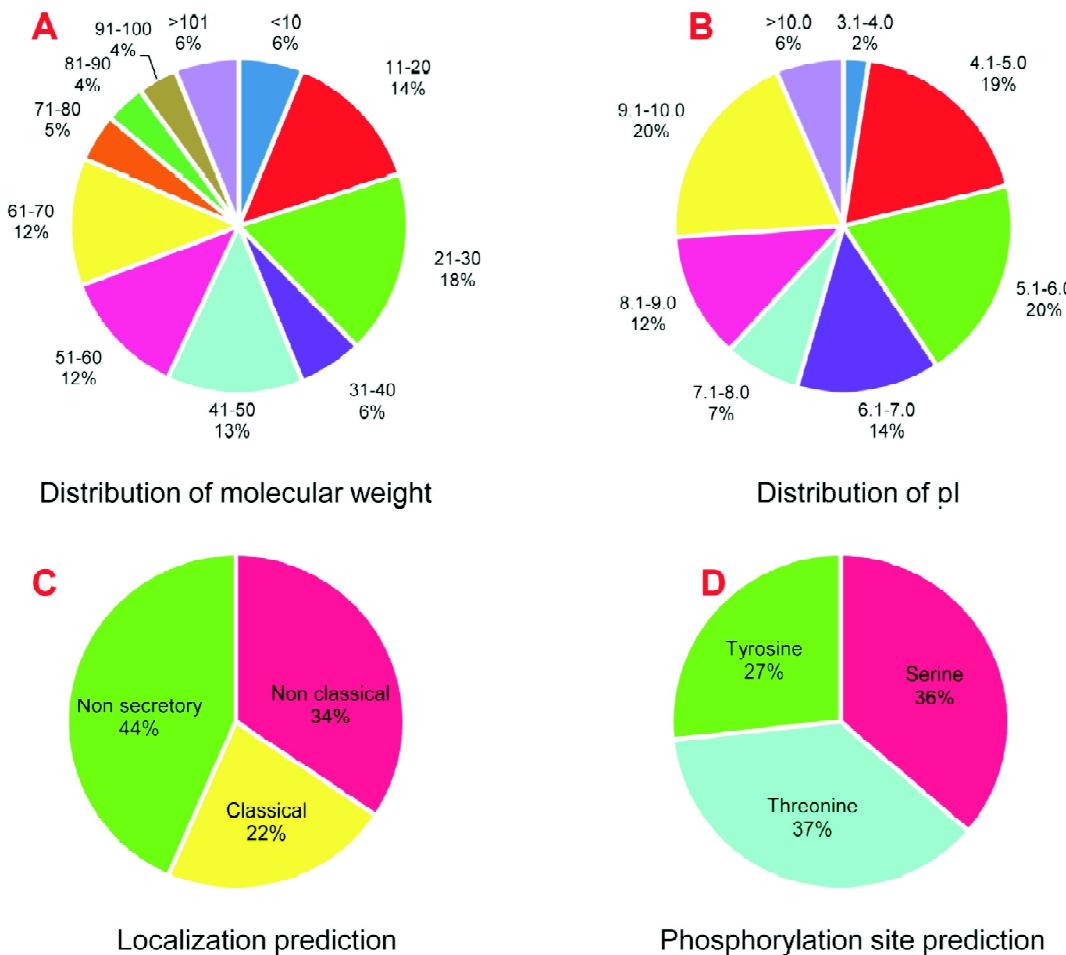


Figure 2: Pie charts showing distribution of molecular weight (A), pI (B), localization prediction (C) and phosphorylation site prediction (D) results. For localization prediction, SignalP and SecretomeP servers were used while for the prediction of phosphorylation site, PhosPhAt database was used.

server (Heazlewood *et al.*, 2008). For this, first, the orthologs of the identified phosphoproteins were searched in against the *Arabidopsis* database in Phytozome. Out of the total 123 identified phosphoproteins, 79 were significantly ($p<0.05$) mapped with the *Arabidopsis* proteins (Supplementary Table 2), representing 75 unique proteins, which were used to predict the putative phosphorylation sites using PhosPhAt server (Supplementary Table 3). The PhosPhAt server predicts the phosphorylation motives based on the MS validated results. Phosphorylation sites prediction scores ≥ 0.5 were considered as significant hits. Using this criterion, a total of 551 putative phosphorylation sites were predicted in the 75 unique *Arabidopsis* annotated proteins wherein 195 (36%) were predicted on the serine residues, 196 (37%) were on threonine and 143 (27%) were on tyrosine residues (Figure 2D,

Supplementary Table 3). A maximum of 29 putative phosphorylation sites were predicted in one protein- “zinc knuckle (CCHC-type) family protein” (AT5G49400) of which 22 were on serine, 4 were on threonine and 3 were on tyrosine residues. Zinc knuckle (CCHC-type) family protein are a member of zinc binding proteins, which are involved in a variety of biological functions including gene transcription, translation, mRNA trafficking, cytoskeleton organization, epithelial development, cell adhesion, protein folding, chromatin remodelling and zinc sensing (Laity *et al.*, 2001). Of the total 75, *Arabidopsis* annotated proteins, 22 (~ 30%) were experimentally validated as phosphoproteins (Supplementary Table 3) while the phosphorylation of the rest of the 70% proteins needs to be confirmed. As these proteins showed a phosphorylation sites prediction scores

≥ 0.5 , it is highly likely that these proteins undergo phosphorylation at one or more time points in the plant's life cycle, however, this needs further investigations.

Functional categorization of the identified phosphoproteins

For analyzing the putative functions of the identified proteins, PANTHER, Gene Ontology and KEGG analyses were carried out on the 75 *Arabidopsis* annotated proteins. PANTHER analysis showed that these identified proteins belonged to 15 protein classes including calcium binding protein (3.1%), chaperone (3.1%), cytoskeletal protein (3.1%), enzyme modulator (6.3%), extracellular matrix protein (3.1%), kinase (3.1%), ligase (3.1%), nucleic acid binding (18.8%), oxidoreductase (6.3%), receptor (3.1%), signaling molecule (3.1%), transcription factor (6.3%), transfer/carrier protein (3.1%), transferase (12.5%), transporter (18.8%), suggesting range of proteins to be regulated by phosphorylation in the apoplast (Figure 3A). Clearly, proteins related to the nucleic acid binding, transporter and transferase activity were the most abundant protein classes. Based on the molecular functions of the identified proteins, GO analysis clustered identified phosphoproteins into eight categories including binding (23.7%), catalytic activity (28.9%), enzyme regulator activity (5.3%), nucleic acid binding transcription factor activity (5.3%), protein binding transcription factor activity (2.6%), receptor activity (2.6%), structural molecular activity (15.8%) and transporter activity (15.8%) (Figure 3B). Similar to the PANTHER analysis, GO analysis also showed that the proteins with the transporter and transcription factor activity were highly abundant. Moreover, GO analysis also showed that proteins with the binding and catalytic activity were the most abundant groups of the identified proteins.

To get even a deeper insight into the metabolic pathways associated with the identified proteins, PANTHER pathway, and KEGG analyses were carried out. Results of these analyses showed that these proteins were involved in many vital biological processes including circadian clock system, DNA replication, amino acid biosynthesis, ionotropic glutamate receptor

pathway and ubiquitin-proteasome pathway (Figure 3C and 3D. KEGG pathway mapper mapped the identified proteins in two major pathways; cysteine and methionine metabolic pathways (Figure 4) and nicotinate and nicotinamide metabolic pathways (Supplementary Figure 1). Of the total 75 unique orthologs of the *Arabidopsis* proteins, 37 were mapped in the cysteine and methionine metabolic pathway (Figure 4), suggesting the key involvement of this pathway in the regulation of apoplastic functions. One of the key byproducts of this pathway is ethylene and interestingly, enzymes of ethylene biosynthetic pathways were identified as phosphorylated proteins in this study. Ethylene is a well-known signaling molecule, which is produced during an array of biological processes including biotic and abiotic stress. 1-aminocyclopropane-1-carboxylate (ACC) oxidase is an enzyme, which catalyzes the final step of ethylene biosynthesis. Recently, it was reported that ACC oxidase undergo autophosphorylation *in vitro* and promotes phosphorylation of some apple fruit proteins in a ripening-dependent manner (Dilley *et al.*, 2013).

Interactome analysis of the identified proteins

To extend the functional analysis of the identified proteins, an interaction network of the identified proteins was constructed using *Arabidopsis* Interaction Viewer (AIV) program (ver 2.0). AIV consists of 70944 predicted and 36329 confirmed *Arabidopsis* interacting proteins. Based on the published and experimentally established interaction only, BAR program detected 299 interactions among the 79 *Arabidopsis* mapped proteins, indicating that the identified proteins could be involved in the similar kind of functions and shares similar localization (Supplementary Table 4). BAR program also predicts the localization and functions of the interacting proteins using MapMan analysis. The two major proteins, identified by interactome analysis were signaling and cell wall associated (marked with number 10 and 30, Figure 5). Interestingly, the majority of the key proteins identified in the interactome analysis were transporters including glucose-6-phosphate/phosphate translocator 2 (At1g61800), glutamine dumper 5 (At5g24920), MATE efflux family protein (At1g71140), proline

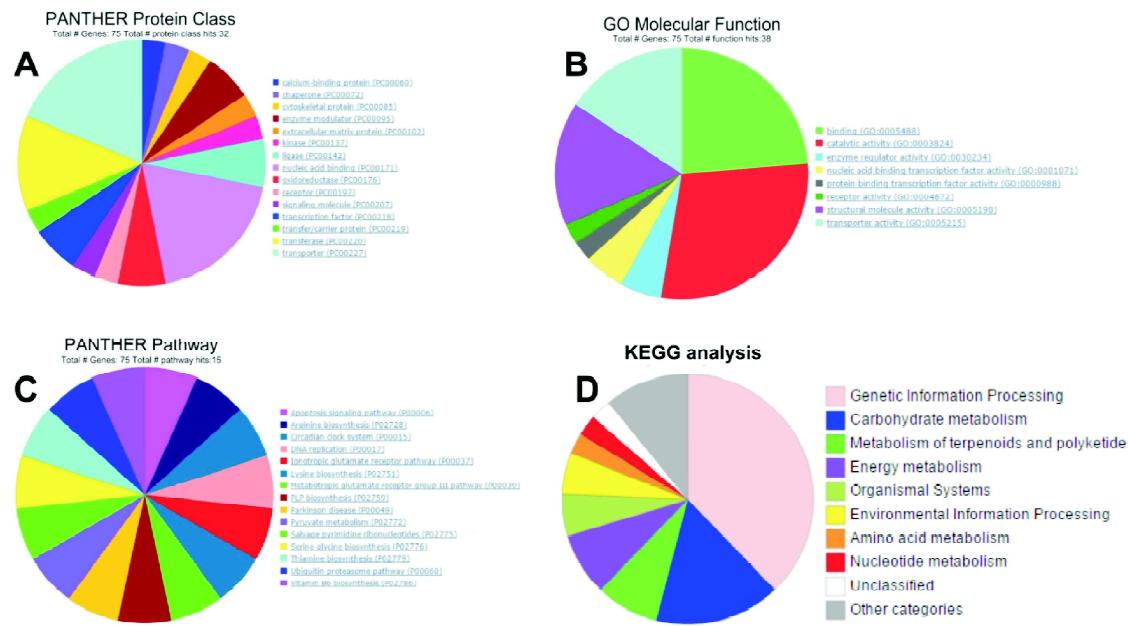


Figure 3: Functional annotation of the identified phosphoproteins using PANTHER (A and C), gene ontology (GO) (B) and KEGG (D) databases.

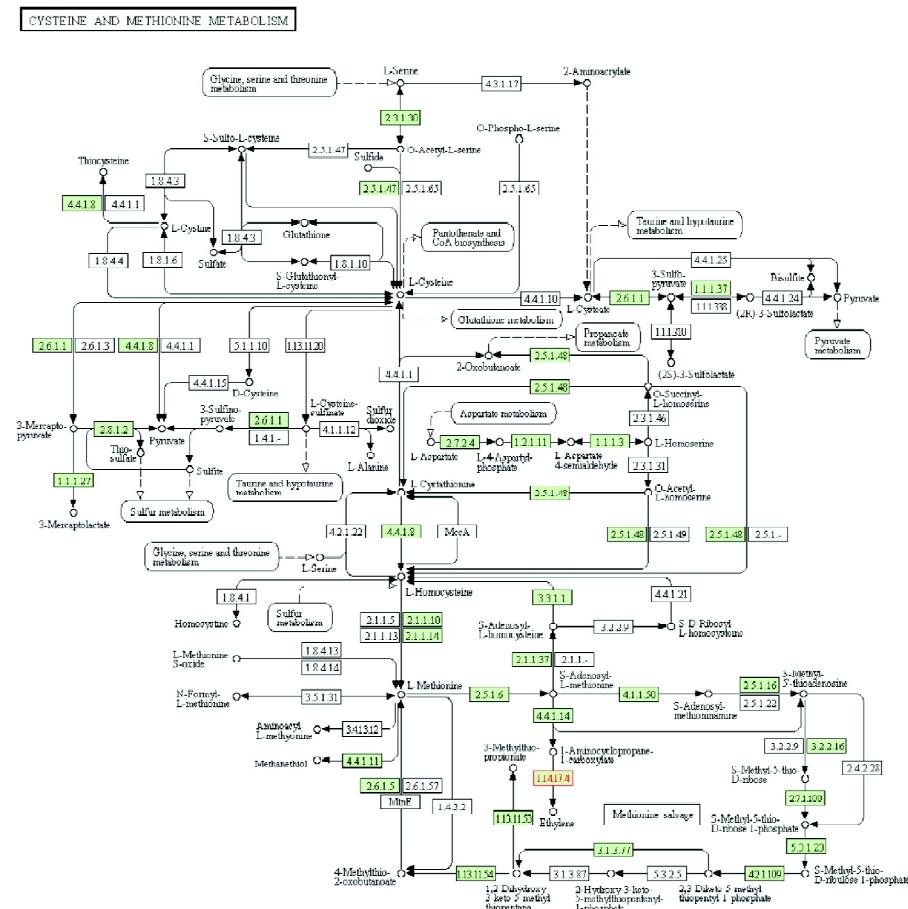


Figure 4: KEGG pathway analysis showing enrichment of phosphoproteins in cysteine and methionine, especially ethylene biosynthesis pathways.

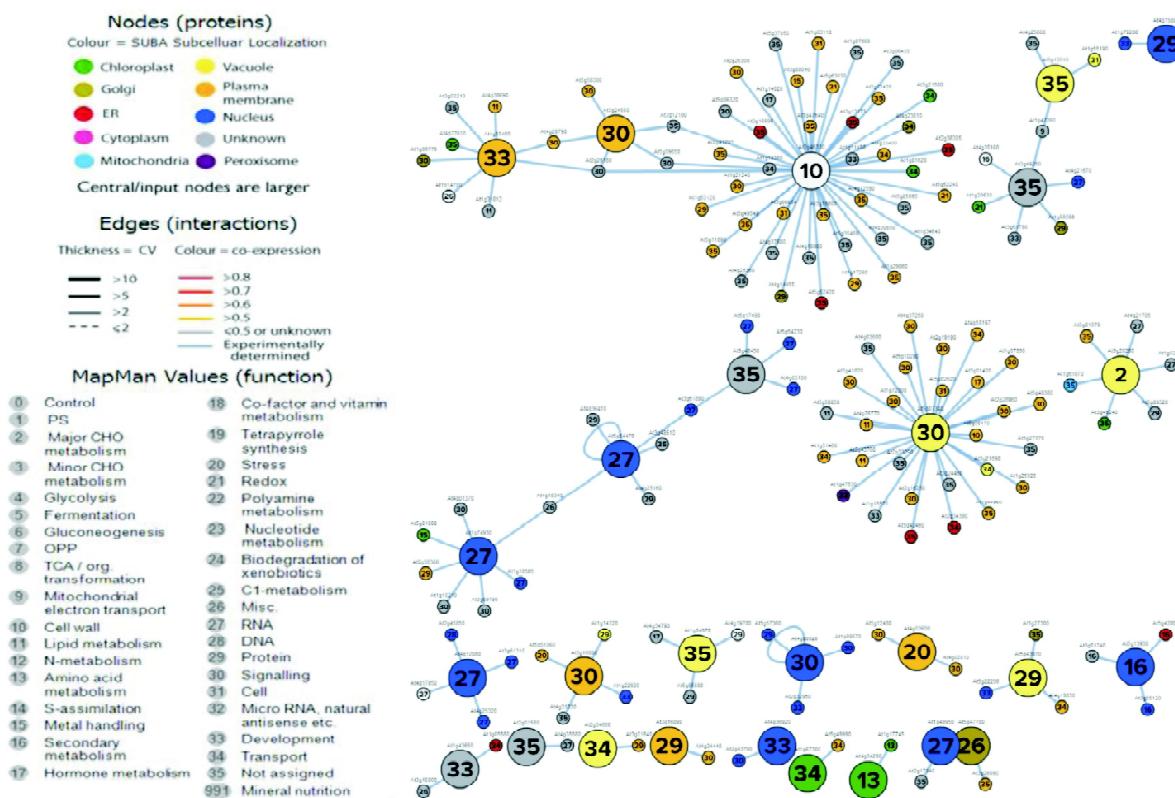


Figure 5: Protein interactome analysis using identified phosphoproteins. For construction of protein interactome map, orthologs of the identified proteins were searched in *Arabidopsis* genome and the interactome map was made using *Arabidopsis* interaction viewer program of “The Bio-Analytic Resource for Plant Biology” (BAR).

transporter 2 (At3g55740) and phosphate transporter 4;1 (At2g29650). Other key proteins identified included receptor kinase 3 (At4g21380), RING/U-box superfamily protein (At3g07120), galactose oxidase (At3961590), nucleosome assembly protein 1;2 (At2g19480), DUF247 (At2g36430), DUF607 (At2g23790), zinc knuckle (CCHC-type) family protein (AT5G49400) and pectin lyase-like superfamily protein (At2g43880). Of these key proteins, phosphorylation of only receptor kinase 3, RING/U-box superfamily protein and zinc knuckle (CCHC-type) family protein has been reported while the phosphorylation of rest of the proteins still needs experimental evidence.

Conclusions

This study provides the first comprehensive phosphoproteomic analysis of the apoplastic proteins of *H. rhamnoides*. Using a shotgun proteomics approach, 123 phosphoproteins were

identified, which were further identified as the orthologs of 75 unique *Arabidopsis* proteins. Our results showed that the identified phosphoproteins were mainly transporters, receptors, and transferases, indicating that phosphorylation and de-phosphorylation of these proteins may possibly regulate the cell transport and signaling. Moreover, KEGG analysis also indicates that the proteins related to the ethylene biosynthesis are also modulated by phosphorylation, indicating a fine tuning of ethylene biosynthesis by phosphorylation. Phosphorylation site prediction/analysis showed that phosphorylation of only 30% of the identified proteins has been reported till date, highlighting the need for enrichment or identification of novel low abundant phosphoproteins. Taken together, our results provide the first insight of the phosphorylation of the apoplastic proteins which would be helpful in achieving a thorough understanding of the apoplastic signaling in plants.

Acknowledgement

This work was supported by a grant (BT/PR10799/NDB/51/171/2008) from Department of Biotechnology, Government of India and a research grant provided by University of Delhi to RD. We thank Dr. Virendra Singh for providing us the seeds of *H. rhamnoides*.

Abbreviations

PMAC, phosphate metal affinity chromatography; PTM, post translational modification; GO, gene ontology; KEGG, kyoto encyclopedia of genes and genomes; UPLC, ultra pressure liquid chromatography.

Conflict of Interest

The authors do not have any conflict of interest with the contents of this manuscript.

References

- Agrawal, G. K., Jwa, N. S., Lebrun, M. H., Job, D., and Rakwal, R. (2010). Plant secretome: Unlocking secrets of the secreted proteins. *Proteomics* 10, 799–827.
- Bendtsen, J. D., Jensen, L. J., Blom, N., Von Heijne, G., and Brunak, S. (2004). Feature-based prediction of non-classical and leaderless protein secretion. *Protein Eng. Des. Sel.* 17, 349–356.
- Bigeard, J., Rayapuram, N., Bonhomme, L., Hirt, H., and Pflieger, D. (2014). Proteomic and phosphoproteomic analyses of chromatin-associated proteins from *Arabidopsis thaliana*. *Proteomics* 14, 2141–2155.
- Casasoli, M., Spadoni, S., Lilley, K.S., Cervone, F., Lorenzo, G.D. and Mattei, B. (2008). Identification by 2-D DIGE of apoplastic proteins regulated by oligogalacturonides in *Arabidopsis thaliana*. *Proteomics* 8, 1042–1054.
- Chaudhary, S., and Sharma, P. C. (2015). DeepSAGE based differential gene expression analysis under cold and freeze stress in seabuckthorn (*Hippophae rhamnoides* L.). *PLoS One* 10.
- Cheng, F., and Williamson, J. D. (2010). Is there leaderless protein secretion in plants? *Plant Signal. Behav.* 5, 129–131.
- Dilley, D. R., Wang, Z., Kadirjan-Kalbach, D. K., Verteridis, F., Beaudry, R., and Padmanabhan, K. (2013). 1-Aminocyclopropane-1-carboxylic acid oxidase reaction mechanism and putative post-translational activities of the ACCO protein. *AoB Plants* 5.
- Ghangal, R., Raghuvanshi, S., and Sharma, P. C. (2012). Expressed sequence tag based identification and expression analysis of some cold inducible elements in seabuckthorn (*Hippophae rhamnoides* L.). *Plant Physiol. Biochem.* 51, 123–8.
- Gupta, R., and Deswal, R. (2012). Low temperature stress modulated secretome analysis and purification of antifreeze protein from *Hippophae rhamnoides*, a himalayan wonder plant. *J. Proteome Res.* 11, 2684–2696.
- Gupta, R., and Deswal, R. (2014). Refolding of β-stranded class I chitinases of *Hippophae rhamnoides* enhances the antifreeze activity during cold acclimation. *PLoS One* 9, e91723.
- Gupta, R., and Kim, S. T. (2015). Depletion of RuBisCO Protein Using the Protamine Sulfate Precipitation Method. *Proteomic Profiling Methods Protoc.*, 225–233.
- Gupta, R., Lee, S. E., Agrawal, G. K., Rakwal, R., Park, S., Wang, Y., et al. (2015a). Understanding the plant-pathogen interactions in the context of proteomics-generated apoplastic proteins inventory. *Front. Plant Sci.* 6, 352.
- Gupta, R., Wang, Y., Agrawal, G. K., Rakwal, R., Jo, I. H., Bang, K. H., et al. (2015b). Time to dig deep into the plant proteome: a hunt for low-abundance proteins. *Front. Plant Sci.* 6.
- Heazlewood, J. I., Durek, P., Hummel, J., Selbig, J., Weckwerth, W., Walther, D., et al. (2008). PhosPhAt : A database of phosphorylation sites in *Arabidopsis thaliana* and a plant-specific phosphorylation site predictor. *Nucleic Acids Res.* 36.
- Huberts, D. H. E. W., and van der Klei, I. J. (2010). Moonlighting proteins: An intriguing mode of multitasking. *Biochim. Biophys. Acta - Mol. Cell Res.* 1803, 520–525.
- Ito, J., Taylor, N. L., Castleden, I., Weckwerth, W., Millar, A. H., and Heazlewood, J. L. (2009). A survey of the *Arabidopsis thaliana* mitochondrial phosphoproteome. *Proteomics* 9, 4229–4240.
- Juvany, M., and Munné-Bosch, S. (2015). Sex-related differences in stress tolerance in dioecious plants: A critical appraisal in a physiological context. *J. Exp. Bot.* 66, 6083–6092.
- Kinoshita-Kikuta, E., Kinoshita, E., Yamada, A., Endo, M., and Koike, T. (2006). Enrichment of phosphorylated proteins from cell lysate using a novel phosphate-affinity chromatography at physiological pH. *Proteomics* 6, 5088–5095.
- Kumar, R., Kumar, A., Subba, P., Gayali, S., Barua, P., Chakraborty, S., et al. (2014). Nuclear phosphoproteome of developing chickpea seedlings (*Cicer arietinum* L.) and protein-kinase interaction network. *J. Proteomics* 105, 58–73.
- Laity, J. H., Lee, B. M., and Wright, P. E. (2001). Zinc finger proteins: New insights into structural and functional diversity. *Curr. Opin. Struct. Biol.* 11, 39–46.
- Maures, T. J., Su, H.-W., Argetsinger, L. S., Grinstein, S., and Carter-Su, C. (2011). Phosphorylation controls a dual-function polybasic nuclear localization sequence in the adapter protein SH2B1α to regulate its cellular function and distribution. *J. Cell Sci.* 124, 1542–1552.
- Novak, B., Kapuy, O., Domingo-Sananes, M. R., and Tyson, J. J. (2010). Regulated protein kinases and phosphatases in cell cycle decisions. *Curr. Opin. Cell Biol.* 22, 801–808.
- Pechanova, O., Hsu, C.Y., Adams, J.P., Pechan, T., Vandervelde, L., Drnevich, J., Jawdy, S., Adeli, A., Suttle, J.C., Lawrence, A.M., Tschaplinski, T.J., Séguin, A. and Yuceer, C. (2010). Apoplast proteome reveals that extracellular matrix contributes to multi stress response in poplar. *BMC Genomics*, 11:674, 1–22.

- Rampitsch, C., and Bykova, N. V. (2012). The beginnings of crop phosphoproteomics: exploring early warning systems of stress. *Front. Plant Sci.* 3, 1–15.
- Reiland, S., Messerli, G., Baerenfaller, K., Gerrits, B., Endler, A., Grossmann, J., *et al.* (2009). Large-scale *Arabidopsis* phosphoproteome profiling reveals novel chloroplast kinase substrates and phosphorylation networks. *Plant Physiol.* 150, 889–903.
- Schulze, W. X. (2010). Proteomics approaches to understand protein phosphorylation in pathway modulation. *Curr. Opin. Plant Biol.* 13, 280–287.
- Sehrawat, A., Abat, J. K., and Deswal, R. (2013). RuBisCO depletion improved proteome coverage of cold responsive S-nitrosylated targets in *Brassica juncea*. *Front. Plant Sci.* 4, 342.
- Slade, W. O., Werth, E. G., Chao, A., and Hicks, L. M. (2014). Phosphoproteomics in photosynthetic organisms. *Electrophoresis* 35, 3441–3451.
- Whiteman, S. A., Nühse, T. S., Ashford, D. A., Sanders, D., and Maathuis, F. J. M. (2008a). A proteomic and phosphoproteomic analysis of *Oryza sativa* plasma membrane and vacuolar membrane. *Plant J.* 56, 146–156.
- Whiteman, S. A., Serazetdinova, L., Jones, A. M. E., Sanders, D., Rathjen, J., Peck, S. C., *et al.* (2008b). Identification of novel proteins and phosphorylation sites in a tonoplast enriched membrane fraction of *Arabidopsis thaliana*. *Proteomics* 8, 3536–3547.
- Wolschin, F., and Weckwerth, W. (2008). "Phosphoproteins: Where are we Today?", in *Plant Proteomics* (John Wiley & Sons, Inc.), 419–442.
- Wu, Y., Engen, J. R., and Hobbins, W. B. (2006). Ultra performance liquid chromatography (UPLC) further improves hydrogen/deuterium exchange mass spectrometry. *J. Am. Soc. Mass Spectrom.* 17, 163–167.