Funct Integr Genomics (2015) 15:17–26
DOI 10.1007/s10142-014-0395-y

ORIGINAL PAPER

# Xylem sap in cotton contains proteins that contribute to environmental stress response and cell wall development

**Zhiyong Zhang · Wanwan Xin · Sufang Wang · Xin Zhang · Haifang Dai · Runrun Sun · Taylor Frazier · Baohong Zhang · Qinglian Wang**

Received: 17 June 2014 / Revised: 18 August 2014 / Accepted: 19 August 2014 / Published online: 28 August 2014


**Abstract** The xylem sap of a plant is primarily responsible for transporting molecules from the underground root system to the aboveground parts of the plant body. In order to understand the role that roots play in cotton growth and development, the components present in xylem sap must be elucidated. In this study, we used a shotgun HPLC-ESI-MS/MS proteomics approach to identify 455 peptides from the xylem sap of field-grown cotton plants at peak blooming stage. Of these peptides, 384 (84.4 %) were found to be secreted proteins and 320 (70.3 %) had special molecular functions. Based on Gene Ontology (GO) analysis, 348 peptides were annotated in terms of molecular function, biological process, and cellular localization, with 46.9 and 45.1 % being related to catalytic activity and binding activity, respectively. Many xylem sap-containing proteins were predicted to be involved in different phases of xylem differentiation including cell wall metabolism, secondary cell wall development and patterning, and programmed cell death. The identification of starch and sucrose hydrolyzing enzymes implicated the interaction between roots and aboveground parts on the aspect of carbohydrate metabolism. Many of the proteins identified in this study are involved in defense mechanisms including pathogen-related proteins, such as peroxidases, chitinases, and germin-like proteins, proteases involved in disease resistance, and phytoalexin phenylpropanoid synthesis-related proteins. The majority of identified signaling proteins were fasciclin-like arabinogalactan proteins and kinases. The results of this study provide useful insight into the communication mechanisms between cotton roots and the rest of the cotton plant.

**Keywords** Cotton · Xylem sap · Proteomics · Defense · Cell wall

## Introduction

Roots have unique and important roles in plants. For example, roots are responsible for obtaining, as well as distributing, water and mineral nutrients essential for plant growth and development. Roots are also responsible for anchoring the plant in the soil. Recent studies have demonstrated that an assortment of organic matter, including proteins, is synthesized in the roots and that these metabolites are then transported to the upper parts of the plant through the xylem vessels. These transported metabolites include hormones, amino acids, proteins, and both oligo- and polysaccharides and they have been shown to perform a variety of different functions (Satoh 2006; Alvarez et al. 2008; Floerl et al. 2008; Ligat et al. 2011).

In 1991, Biles and Abeles observed that proteins were present in xylem sap (Biles and Abeles 1991). Since then, xylem sap-containing proteins have been investigated in several plant species, including *Brassica napus* (Buhtz et al. 2004; Kehr et al. 2005; Floerl et al. 2008), *Brassica oleracea* (Buhtz et al. 2004; Ligat et al. 2011), *Cucumis sativus* (Sakuta and Satoh 2000; Oda et al. 2003; Buhtz et al. 2004), *Cucurbita maxima* (Buhtz et al. 2004), *Solanum lycopersicum* (Rep et al. 2002, 2003; Houterman et al. 2007), *Glycine max* (Djordjevic et al. 2007; Subramanian et al. 2009), *Zea mays* (Alvarez et al. 2006, 2008), *Malus domestica* (Biles and Abeles 1991),

**Electronic supplementary material** The online version of this article (doi:10.1007/s10142-014-0395-y) contains supplementary material, which is available to authorized users.

Z. Zhang · W. Xin · S. Wang · X. Zhang · H. Dai · R. Sun · B. Zhang · Q. Wang (✉)
School of Life Science and Technology, Henan Institute of Science and Technology, Hualan Road, 453003 Xinxiang, China
e-mail: wangql985@163.com

R. Sun · T. Frazier · B. Zhang
Department of Biology, East Carolina University, Greenville, NC 27858, USA

*Prunus persica* (Biles and Abeles 1991), *Pyrus communis* (Biles and Abeles 1991), and a poplar hybrid (*Populus trichocarpa*×*Populus deltoides*) (Dafoe and Constabel 2009). The results of these studies identified several types of proteins that are transported in the xylem, including chitinases, glycosyl hydrolases, peroxidases, proteases, lectins, pathogenesis-related proteins (PR-proteins), and cell wall structural proteins such as glycine-rich proteins (GRPs). Additional studies of these xylem sap-containing proteins will further elucidate the communication mechanisms between different parts of the plant and ultimately reveal more root contributions to plant development.

More and more evidence has shown that xylem sap-containing proteins play an important role in plant development; these proteins have been found to circulate between plant roots and the rest of the plant body (Dafoe and Constabel 2009; Agrawal et al. 2010; Ligat et al. 2011). The xylem fluid aids in the plant's interaction with its environment by delivering materials to the cells and providing intracellular and intercellular communications (Agrawal et al. 2010). During these processes, xylem vessels also deliver proteins from the roots to the leaves for cell wall repair, lignification, or defense purposes (Alvarez et al. 2006, 2008). Hormones found in xylem sap, such as cytokinins and abscisic acid (ABA), have been shown to play significant roles in regulating leaves' senescence (Wang et al. 2012). Despite these advances, the roles of xylem sap proteins involved in the signaling of root to shoot are still unknown. Therefore, it is necessary to study the composition of xylem sap in the field in order to further investigate the expression profiles and signaling roles of xylem sap-containing proteins.

Although there have been several studies in other plant species aimed at identifying xylem sap-containing peptides, no research has been performed in cotton. Cotton is the world's leading fiber crop and one of the most important economic cash crops. Often times, it can be observed that cotton plants will senesce early due to nutrient deficiency, with symptoms particularly evident in the leaves (Zhang et al. 2007). Cotton plant senescence may originate from root senescence at an early stage and then be transported to other parts of the plant through the xylem. Therefore, understanding the composition of xylem sap-containing constituents will help determine the metabolic and chemical role that roots play in plant growth and development.

There are several approaches for identifying and studying xylem sap-containing proteins, two of which include two-dimensional electrophoresis (2-DE)/liquid chromatography–mass spectrometry/mass spectrometry (LC-MS/MS) (Alvarez et al. 2006, 2008) and sodium dodecyl sulfate-polyacrylamide gel electrophoresis/liquid chromatography–mass spectrometry/mass spectrometry (SDS-PAGE/LC-MS/MS) (Ligat et al. 2011). Although both 2-DE/LC-MS/MS and SDS-PAGE/LC-MS/MS have been widely used, there are several disadvantages for these methods. For instance, these methods can be time-consuming because they require gel electrophoresis. Additionally, the 2-DE/LC-MS/MS method cannot detect proteins with extreme molecular masses, isoelectric points, and if the proteins are hydrophobic or unsolvable, such as integral membrane proteins. Recently, HPLC-MS/MS shotgun analysis has emerged as an effective approach for comprehensive proteomic analysis (Ishihama et al. 2005; Liu et al. 2010). HPLC-MS/MS has been proven to be an alternative technology capable of identifying hundreds of proteins from a single sample, including proteins with a very low level of expression that are usually undetectable by gel staining. Shotgun proteomics relies on protein separation after the proteolytic digestion and takes advantage of MS/MS to infer the amino acid sequences of individual peptides. Compared to the 2-DE/LC-MS/MS proteomics approach, the shotgun HPLC-MS/MS method is highly efficient and labor saving. Here, in this study, we employed HPLC-MS/MS shotgun analysis to identify and functionally analyze xylem sap-containing proteins in cotton.

## Materials and methods

### Cotton genotype and its cultivation

*Gossypium hirsutum* L. cv Baimian 1, bred by the Henan Institute of Science and Technology (HIST), was used in this study. Baimian 1 is a long-season cultivar, which has been widely adopted in China. Cotton plants were cultivated at the HIST Research Station (Xinxiang, Henan, China). The soil conditions of the experimental field were as described in a previous report (Zhang et al. 2013). Cotton seeds were sowed on May 1, 2012 and seedlings emerged at 5 days after sowing. The emerging day was recorded as 0 day after emerging (DAE). The planting density of the field was 45,000 plants per square hectometer with an inter-row spacing of 1 m and an intra-row spacing of 0.22 m.

### Xylem sap collection

At the 100 DAE, during the peak blooming stage of cotton, the aboveground part of the cotton plants was cut at the stems approximately 5 cm above ground level. Because of the "root pressure", xylem sap was pressed out of the stems and collected from each cut cotton plant in the early evening. The collected xylem sap was immediately frozen in liquid nitrogen, transported to the laboratory, and stored at −80 °C. The xylem sap was collected from 12 individual healthy plants. These 12 individual samples were then pooled to form 3 biological replicates for proteomics analysis.

### Protein preparation

The xylem sap was thawed and filtered through 0.2 μm cellulose acetate filters. Sap-containing proteins were then precipitated with 10 %*w/v* TCA. The precipitated proteins were collected by centrifugation for 20 min at 4,000×*g* at 4 °C. The protein pellet was then washed with acetone, the supernatant was discarded, and the pellet was air dried. Protein concentrations in the xylem sap and extracts were determined using the BCA Protein Assay Kit (Pierce, Rockford, IL, US).

### In-solution digestion

Protein digestion was performed according to the filter aided sample preparation (FASP) procedure as described by Wisniewski et al. (2009). Briefly, the protein pellet (about 30 μg) was solubilized in 30 μl SDT buffer (4 % SDS, 100 mM detergent (DTT), 150 mM Tris–HCl pH 8.0) at 90 °C for 5 min. The DTT and other low molecular weight components were removed using 200 μl of UA buffer (8 M Urea, 150 mM Tris–HCl pH 8.0) by repeated ultrafiltration (Microcon units, 30 kD). Next, 100 μl of 0.05 M iodoacetamide dissolved in UA buffer was added to block the reduced cysteine residues and the samples were incubated for 20 min in the dark. The filter was then washed three times with 100 μl UA buffer followed by two washes with 100 μl 25 mM $NH_4HCO_3$. Finally, the protein suspension was digested overnight at 37 °C with 2 μg trypsin (Promega) dissolved in 40 μl 25 mM $NH_4HCO_3$. The resulting peptides were collected as a filtrate.

### LC-MS/MS

Ettan™ MDLC system (GE Healthcare) was employed for desalting, separating, and detecting the tryptic peptide mixtures. In this system, samples were desalted on a reverse phase (RP) trap column (Zorbax 300 SB C18, Agilent Technologies), and then separated on a RP column (150 μm i.d., 100 mm length, Column Technology Inc., Fremont, CA). Two mobile phases were used: mobile phase A utilized HPLC-grade water containing 0.1 % formic acid and mobile phase B used HPLC-grade acetonitrile containing 0.1 % formic acid. A total of 20 μg of the tryptic peptide mixtures were injected into the columns. The flow rate was adjusted to 2 μL/min by using a linear gradient of 4–50 % buffer B for 50 min, 50–100 % buffer B for 4 min, and 100 % buffer B for 6 min. A LTQ Velos (Thermo Scientific) linear ion trap, equipped with a micro-spray interface, was connected to the LC system for detecting peptides eluted from the columns. Data-dependent MS/MS spectra were obtained simultaneously. Each scan cycle consisted of one full scan mass spectrum (m/z 300–1,800) followed by 20 MS/MS events of the most intense ions with the following dynamic exclusion settings: repeat count 2, repeat duration 30 s, and exclusion duration 90 s.

### Data analysis

Using the BioworksBrowser rev. 3.1 program (Thermo Electron, San Jose, CA), MS/MS spectra were automatically searched against the target/reverse Uniprot Eudicotyledons database (containing 897,683 sequences, downloaded on April 12, 2013). The results of the protein identification were extracted from SEQUEST out files with BuildSummary.

The peptides were described to be tryptic and up to two missed cleavages were allowed. Carbamidomethylation of cysteines were treated as a fixed modification, whereas oxidation of methionine residues was considered a variable modification. The mass tolerance allowed for the precursor ions and the fragment ions was 2.0 and 0.8 Da, respectively. Protein identification criteria were based on the false discovery rate (FDR) value. The filter was set to FDR ≤0.01 at the peptide and protein levels.

### GO annotation of identified proteins

To investigate the functions of the xylem sap-containing proteins, we employed gene ontology (GO) analysis to classify all of the identified proteins in terms of cellular component, molecular function, and biological process. Briefly, identified proteins (peptide sequences) were aligned against the GO protein database for GO term classification. Biological function of an individual protein was based on either GO annotation of the most similar proteins from other di-cotyledon species or based on literature for easy interpretation. Annotated proteins were clustered into general biological categories that were identified on the basis of common terminology used in literatures. Proteins that contained a signal peptide and were predicted to be secreted were identified using signalP (www.cbs.dtu.dk/services/SignalP/). SecretomeP was used to identify secreted proteins that did not contain a signal peptide (www.cbs.dtu.dk/services/SecretomeP).

## Results

### Identification and annotation of peptides obtained from the xylem sap of cotton

Based on the shotgun HPLC-ESI-MS/MS proteomic analysis, a total of 455 peptides were identified from proteins extracted from cotton xylem sap. Of these 455 peptides, 16 (3.5 %) were evident at the protein level, 121 (26.6 %) were evident at the transcript level, 70 (15.4 %) were inferred from homology,

and 248 (54.5 %) were predicted. Correspondingly, 13, 101, 70, and 164 were GO-annotated for each level of identified protein. A total of 348 xylem sap-containing proteins were identified with certain functions based on GO annotation (Table 1). Of these 348 proteins, 302, 262, and 118 were annotated with molecular function, biological process, and cellular component, respectively. Among them, 73 proteins were classified with all three annotations (Fig. 1).

## Characteristics of the xylem sap-containing proteins: molecular mass and isoelectric points

The molecular mass of the cotton xylem sap-containing proteins varied with a range of 1.39 [unknown protein 1 (fragment)] to 332.46 kDa [putative uncharacterized protein At2g17930 (fragment)]. However, the molecular masses of the majority [365 (80.2 %)] of proteins were between 10 and 80 kDa in size (Fig. 2). The isoelectric point (pI) for the identified proteins also varied from 4.34 (putative uncharacterized protein) to 12.01 (putative uncharacterized protein); however, the pI for the majority of identified proteins was between 4 and 10. Interestingly, there were almost the same numbers of proteins with an acidic pI (pI<7) as a basic pI (pI≥7) (Fig. 3). These results suggest that cotton xylem sap contains a plethora of different proteins with regards to size and charge.

## Molecular functions of the xylem sap-containing proteins

The xylem sap-containing proteins were clustered into 12 catalogues of molecular functions. The two richest of these catalogues included the catalytic activity (46.9 %) and binding activity (45.1 %) groups (Fig. 4a). Based on their specific functions, the proteins within the catalytic activity group were further classified into nine subgroups (Fig. 4b). The majority of these proteins were related to hydrolase activity (44.1 %) and oxido-reductase activity (36.0 %) (Fig. 4c). Identification of enzymes involved in redox reactions was expected as many oxidation reactions occur in the extracellular matrix and function to construct complex polymer networks that include carbohydrates, aromatic compounds, and structural proteins

**Table 1** Summary of proteins identified from the xylem saps of the field-grown cotton plants at the peak blooming stage

| Evidence for the existence of a protein | Total number | Annotation number |
|---|---|---|
| Evidence at protein level | 16 | 13 |
| Evidence at transcript level | 121 | 101 |
| Inferred from homology | 70 | 70 |
| Predicted | 248 | 164 |
| Total | 455 | 348 |

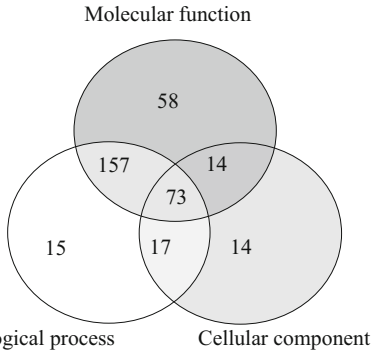


**Fig 1** Gene ontology annotations of proteins identified by the shotgun ESI-MS approach in the xylem sap of field-grown cotton plants at the peak blooming stage

(Passardi et al. 2004). Therefore, our results suggest that the hydrolysis and oxidation-reduction of carbohydrates and proteins is extremely active in the xylem and that these processes may play a role in xylem differentiation and secondary cell wall remodeling.

Among the binding activity proteins, 15.8 % of proteins were for large organic molecules including DNA, RNA, chromatin, and carbohydrates, and 44.9 % were for smaller ions including copper, zinc, manganese, iron, and calcium (Fig.4b). The identification of xylem sap-containing proteins with binding activity provides support for long-distance transport of organic molecules and nutrients from the roots to other parts of the plant.

## Potential localization of the xylem sap-containing proteins

SignalP and SecretomeP were employed to identify secreted proteins in the xylem sap that did and did not contain a signal peptide. Of the identified xylem sap-containing proteins, 160 (35.2 %) contained signal peptides and 224 (49.2 %) did not have a signal peptide, designating these as non-classical secreted proteins. Overall, secreted proteins constituted 84.4 % of all identified proteins in the xylem sap of cotton (Fig. 5).

Of the 42 proteins classified as CC-ECS proteins, 25 (59.5 %) had a signal peptide whereas 15 (35.7 %) were found to be non-classical secreted proteins. Therefore, the majority (95.2 %) of proteins localized in the extracellular space, according to GO annotation, were also secreted proteins. Of the 76 proteins characterized as CC-NECS, 10 (13.2 %) contained a signal peptide and 44 (57.9 %) were identified as non-classical secreted proteins (Fig. 6).

## Classification of known proteins predicted to be secreted proteins

The cotton xylem sap-containing proteins that were predicted to being secreted proteins (Table 2; supplementary Table S1) were classified into seven of the nine functional classes

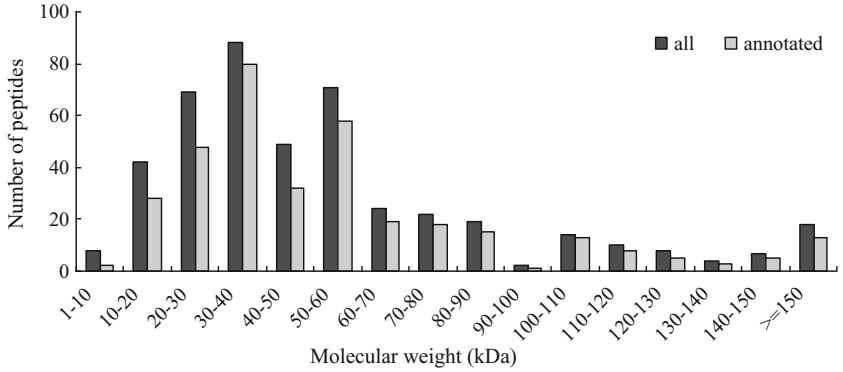


**Fig 2** Distribution of molecular weight for all (**a**) and annotated proteins (**b**) identified by the shotgun ESI-MS approach in the xylem sap of field-grown cotton at the peak blooming stage

previously defined for *Arabidopsis thaliana* (Albenne et al. 2013). These seven classes of proteins included proteins acting on carbohydrates (22.3 %), oxido-reductases (29.3 %), proteases (7.5 %), proteins related to lipid metabolism (3.2 %), proteins involved in signaling (7.5 %), proteins with domains interacting with carbohydrates/proteins/DNA (9.0 %), miscellaneous proteins having diverse functions (20.2 %), and proteins of unknown function (1.1 %). No structural proteins were identified in these analyses. Among the 42 proteins acting on carbohydrates, two glycoside hydrolase (GH) families were well represented: 7, 7, 5, and 4 proteins belonged to GH13, GH17, GH18, and GH16 protein families, respectively. More than half of the identified oxido-reductases that contained interacting domains were peroxidases (31) and transcription factors (9). Other homologous oxido-reductase proteins included multicopper oxidases, copper binding proteins (8), and P450 family proteins (4). According to molecular function annotation by Gene Ontology (GO), all identified peroxidases belong to the haem-peroxidase family. A total of 19 identified peroxidases were classified as Class III peroxidases (EC1.11.1.7). The proteases identified in this study were classified into three individual classes: serine proteases (7), aspartic proteases (6), and cysteine proteases (1). In the signaling function class, half of all proteins were fasciclin-like arabinogalactan proteins (FLAs) (9) but only one was an arabinogalactan protein (AGP). The numbers of proteins with function in lipid metabolism varied.

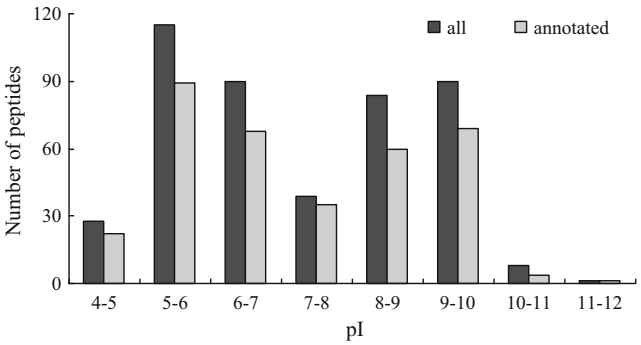


**Fig 3** Distribution of isoelectric points (pI) for all (**a**) and annotated proteins (**b**) identified by the shotgun ESI-MS approach in the xylem sap of field-grown cotton plants at the peak blooming stage

## Discussion

### Potential origin of xylem sap-containing proteins

It is important to elucidate the origin of xylem sap-containing proteins. Since xylem vessels are generated from a modification of programmed cell death, xylem sap-containing proteins may be secreted from the surrounding cells, including the tracheids (Biles and Abeles 1991; Alvarez et al. 2006). It appears that a number of cotton xylem sap-containing proteins may participate in the process of xylem vessel differentiation. These proteins include xylem cysteine proteins (XCPs), cysteine protease (Cp) proteins, proteins regulating ROP GTPase activity (Rop guanine nucleotide exchange factor, ADP-ribosylation factor GTPase-activating protein), proteins related with cytoskeleton movement (kinesin-like protein and formin-like protein 3), and transcription factors such as WRKY, MYB, and NAC (Table 2; supplementary Table S1). XCPs have been found to be involved in programmed cell death (PCD) of xylem cells (Bollhöner et al. 2012) and were previously detected in xylem saps of *B. napus* (Kehr et al. 2005). Correspondingly, WRKY, MYB, and NAC transcription factor families were demonstrated to regulate secondary cell wall formation (Ulker and Somssich 2004; Salazar et al. 2013). Kinesin-like protein and formin-like protein were shown to be involved in microtubule-based movement and actin nucleation of cytoskeletal components (actins and microtubles), as well as participate in patterned secondary cell wall formation (Oda and Hasezawa 2006). A recent study has demonstrated that ROP GTPases participate in spatial organization of xylem cell walls (Oda and Fukuda 2013). The activity of Rho GTPases is regulated by cycling between GDP-bound inactive and GTP-bound active forms, in which guanine nucleotide exchange factors (GEFs) and GTPase-activating proteins predominantly function to

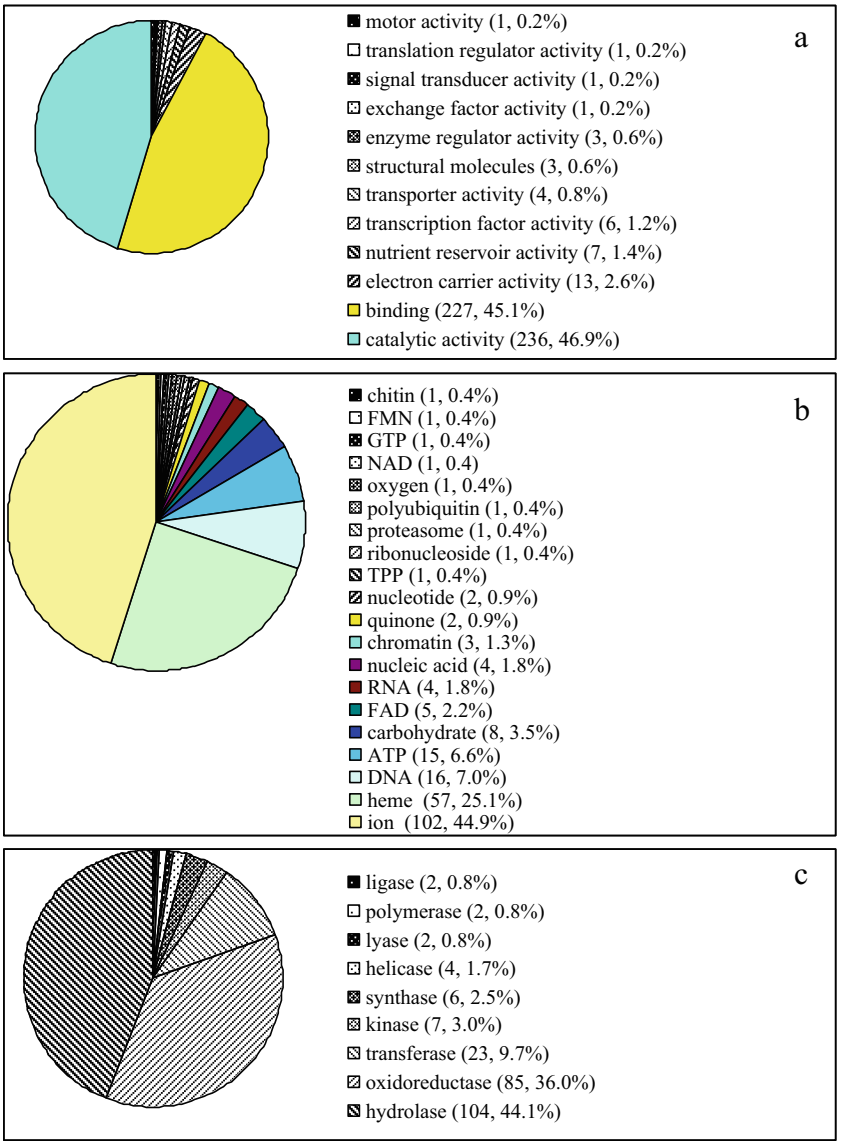


**Fig 4** GO annotation of the proteins identified by the shotgun ESI-MS approach. **a** Numbers and percentages of the annotated proteins in 12 groups of molecular functions. **b** Numbers and percentages of the annotated proteins with catalytic activity. **c** Numbers and percentages of the annotated proteins with binding activity. As one protein may have multiple GO terms, the total of GO terms in each category or subcategory is more than the total number of proteins

promote activation of the GTPases and to terminate the signaling event.

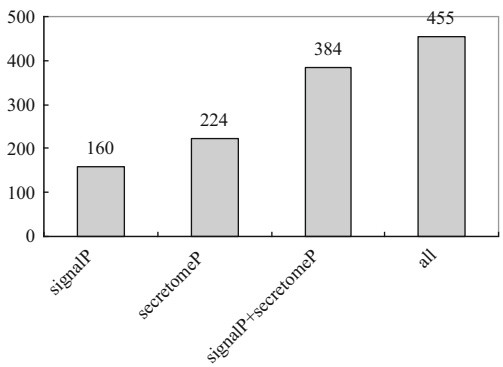


**Fig 5** Number of proteins with signal peptides predicted with signalP, leaderless signal peptide but secrete proteins predicted with secretomeP and both

The xylem sap-containing proteins may also originate from secretions of adjacent xylem parenchyma or pericycle cells (Satoh 2006). The proteins are, as for other apoplastic proteins, most likely mediated by an amino-terminal signal peptide (Nielsen et al. 1997; Bendtsen et al. 2004). In the classical secretory pathway, proteins are synthesized in the endoplasmic reticulum (ER) before passing through the dictyosomes and entering one of two fates: (1) insertion by vesicles in the plasma membrane or (2) secretion into the extracellular space. Among the secreted proteins, some can be linked or anchored to the cell membrane. Varying proportions of apoplastic proteins were predicted with a signal peptide sequence in different organisms such as grapevine (66 %) (Delaunois et al. 2013), soybean (65 %) (Subramanian et al. 2009), *Arabidopsis* (47 %) (Casasoli et al. 2008), and rice (37 %) (Cho et al. 2009). In this study, we found that 35.2 % of the cotton xylem sap-containing proteins contained signal peptides.

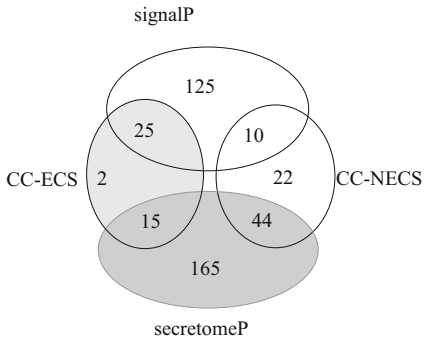


**Fig 6** Comparison among proteins with GO annotation of cellular component (*CC*) as extracellular space (*ECS*) protein or non-extracellular space (*NECS*), signalP (according to www.cbs.dtu.dk/services/SignalP/, with signal peptide), and secretomeP (according to www.cbs.dtu.dk/services/SecretomeP/, without *signal peptide*). Extracellular space includes cell wall, extracellular region, apoplast, anchored to plasma membrane according to GO annotation about cellular component. Non-extracellular space included cytosol, cell or organelle membrane, integral to membrane, and organelles like nucleus, endoplasmic reticulum, and mitochondria

Proteins without an N-terminal leader signal peptide that were identified in the apoplastic fluid (AF) were named leaderless secreted proteins (LSPs). These proteins could be secreted by a non-classical secretory mechanism, as described in yeast and bacteria (Agrawal et al. 2010; Bendtsen et al. 2004, 2005; Nombela et al. 2006). LSPs potentially possess dual functions depending on the localization of the proteins either inside or outside of the plant cell. Part of the predicted cytoplasmic proteins found in the AF could be actively translocated into the extracellular space (Delaunois et al. 2013). These proteins may have other functions; for example, heat shock proteins may play a role in pathogen elicitation (Chivasa et al. 2005). The presence of LSPs can be predicted by SecretomeP analysis (Bendtsen et al. 2004).

The ratio of LSPs to the total proteins in the apoplast was associated with plants under stress or non-stress conditions. Under normal conditions, the leaf contained 15 % of LSPs based on an analysis using the vacuum-infiltration-centrifugation method (Delaunois et al. 2013). Under stress conditions, the plant usually generated more LSPs. One study shows that *Arabidopsis* apoplasts contained more than 55 % of LSPs when the plants were treated with salicylic acid (Cheng et al. 2009). In our study, LSPs constituted 49.2 % of 84.4 % of all identified secreted proteins. These results suggest that the cotton plants may have been facing certain kind of abiotic or biotic stress.

## The predicted roles of xylem sap-containing proteins

Plant cell walls are primarily composed of polysaccharides (cellulose), cross linking glycans (hemicellulose), pectins, and

**Table 2** List of 16 representative proteins discovered in this study contributing to cell wall development as well as stress response

| Accession no. | Protein | Organism | PC | UPC | CP (%) | MW | PI | PE | SP | LSP | Category |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Proteins acting on carbohydrate and cell wall development | | | | | | | | | | | |
| Q8VZ56 | Alpha-amylase 1 | *Arabidopsis thaliana* | 1 | 1 | 1.9 | 47,377.6 | 5.6 | 2 | N | Y | CHME |
| Q1HGA7 | Cell wall invertase (fragment) | *Populus* sp. UG-2006 | 1 | 1 | 3.4 | 43,454.1 | 6.7 | 2 | N | Y | CHME |
| P93154 | Chitinase | *Gossypium hirsutum* | 29 | 7 | 33.5 | 28,806.4 | 5.9 | 4 | Y | | PD[a] |
| P93153 | 1,3-beta-glucanase | *Gossypium hirsutum* | 3 | 2 | 8.2 | 37,644.9 | 5.2 | 2 | Y | | CW/PD[a] |
| Q94CD8 | Glucan endo-1,3-beta-glucosidase 4 | *Arabidopsis thaliana* | 1 | 1 | 2.4 | 54,416.0 | 5.9 | 1 | Y | | CW/PD[a] |
| Q7XAS3 | Beta-D-glucosidase | *Gossypium hirsutum* | 13 | 8 | 21.7 | 68,444.7 | 8.7 | 1 | Y | | CW |
| Q9SVE5 | Expansin-like A2 | *Arabidopsis thaliana* | 3 | 2 | 8.7 | 28,642.7 | 8.5 | 2 | Y | | CW[a] |
| G7IC11 | Callose synthase | *Medicago truncatula* | 1 | 1 | 14.6 | 11,110.8 | 6.3 | 4 | N | Y | SR[a] |
| Proteins relating to environmental stresses | | | | | | | | | | | |
| Q7XYR7 | Class III peroxidase (pod7) | *Gossypium hirsutum* | 17 | 9 | 38.2 | 35,420.8 | 9.3 | 2 | Y | | SR |
| I1T4H9 | Bacterial-induced peroxidase | *Gossypium mustelinum* | 6 | 2 | 9.8 | 35,687.6 | 9.5 | 3 | Y | | SR |
| Q6TDS6 | Secretory laccase | *Gossypium arboreum* | 45 | 8 | 22.8 | 63,342.6 | 6.2 | 2 | Y | | PD[a] |
| B9SRN1 | Cytochrome P450, putative | *Ricinus communis* | 1 | 1 | 2.4 | 57,462.6 | 8.6 | 3 | N | Y | PD[a] |
| A9XTK6 | Arabinogalactan protein 2 | *Gossypium hirsutum* | 2 | 2 | 15.64 | 25,612.04 | 9.27 | 2 | Y | | SR[a]/ST[a] |
| A9XTK7 | Fasciclin-like arabinogalactan protein 2 | *Gossypium hirsutum* | 1 | 1 | 6.79 | 28,306.3 | 6.5 | 2 | Y | | SR[a]/ST[a] |
| G7KT20 | WRKY transcription factor | *Medicago truncatula* | 1 | 1 | 2.1 | 37,274.0 | 5.8 | 4 | N | Y | SR/GE[a] |
| G7ZWS1 | Resistance protein RGC2 (fragment) | *Medicago truncatula* | 1 | 1 | 2.1 | 104,745.0 | 5.7 | 4 | N | Y | PD[a] |

*CHME* carbohydrate metabolism, *CP* cover percent, *CW* cell wall synthesis/modification, *GE* gene expression, *LSP* leaderless secretory protein, *MW* molecular weight, *N* no, *PC* peptide count, *PD* pathogenesis defense, *PE* protein existence level, *PI* isoelctric poin, *SR* stress response, *ST* signal transduction, *UPC* unique peptide count, *Y* yes

[a] category according to literatures but not to gene ontology

some proteins (Cosgrove 1997a, 1997b). These compounds form a complex interactive network known as the extracellular matrix (ECM) (Bacic et al. 1988; Carpta and Gibeaut 1993). Most ECM proteins belong to large families that include enzymes (such as hydrolases, proteases, glycosidases, peroxidases, and esterases), expansins, cell wall-associated kinases, and hydroxyproline (Hyp)-rich glycoproteins (The Arabidopsis Genome Initiative 2000). These compounds are involved in many biological processes, such as reinforcement or restructuring of cell wall architecture, protection from pathogen attack and abiotic stresses, signaling, and metabolism of apoplastic compounds (Jamet et al. 2008). Glycoside hydrolase (GH) enzymes are important in not only cell wall metabolism, but also the biosynthesis of glycans, plant defense response, signaling, and the mobilization of storage reserves (Minic 2008).

*Starch and sucrose metabolism*

Starch accumulation in plants peaks at the time of first anthesis. Simultaneously, the level of starch in the roots of field-grown cotton quickly declines after flowering. This indicates that roots could be a significant source of carbohydrates for boll development and that they could possibly participate in a source/sink relationship (Taliercio et al. 2009). In this study, we detected several carbohydrate metabolizing enzymes in cotton xylem sap. For instance, we identified enzymes related to starch metabolism including alpha-amylase, isoamylase, pentatricopeptide repeat-containing protein, and alpha-glucosidase. Sucrose hydrolases such as cell wall invertase, acid beta-fructofuranosidase, and beta-fructofuranosidase were also present. These results suggest that cotton roots may play a role in carbohydrate metabolism in all parts of the plant.

*Signaling: arabinogalactan proteins response to stress*

Arabinogalactan proteins (AGPs) are a class of Hyp-rich glycoproteins that are highly glycosylated and are abundant in the plant cell wall and plasma membrane (Guan and Nothnagel 2004; Bacic et al. 1988; Gaspar et al. 2001). The binding of AGPs has been shown to trigger wounding-like responses such as callose synthesis (Guan and Nothnagel 2004), which indicates that AGPs may serve as signaling molecules in cotton responses to various biotic and abiotic stresses. The fasciclin-like AGPs (FLAs) are one major subclass of AGPs. FLAs represent a complex multigene family of proteoglycans and are regulated by both development and stress (Tan et al. 2012). In this study, 9 FLAs were identified, indicating that these proteins may play a role in root response to stress and xylem development. Some of the FLAs are likely to be attached to the plasma membrane through a glycosylphosphatidylinositol (GPI) anchor and may interact with receptor-like kinases, such as cell wall-associated kinases (Gens et al. 2000). In the cotton xylem sap, both FLAs and receptor-like protein kinases were identified. Therefore, it is possible that the FLAs found in cotton xylem sap are involved in this type of interaction, as well as in interactions with other signaling molecules.

*Chitin recognition and signaling*

Activation of plant defense responses relies on the recognition of pathogen-associated molecular patterns (PAMPs) such as microbial N-acetylglucosamine (GlcNAc)-containing glycans, and other pathogen signatures that are not inherent to the plants themselves. LysM proteins mediate either direct or indirect recognition of such structures (reviewed by Gust et al. 2012). Plant LysM proteins are involved in defense signaling pathways against fungal attack (reviewed by Buist et al. 2008). Chitin, the major component of fungal cell walls, insect exoskeletons, and crustacean shells, is a polymer of GlcNAc and a known PAMP. In *Arabidopsis*, the recognition of chitin can induce gene expression and defense responses through LysM-mediated signaling pathways (Wan et al. 2008). In this study, we identified a peptidoglycan-binding LysM domain-containing related protein in the xylem sap of cotton plants. This suggests that the cotton plants may have been exposed to fungal or insect pathogens and that chitin recognition and signaling pathways had already been activated.

*Pathogenesis-related (PR) proteins*

Pathogenesis-related (PR) proteins are a group of well-studied proteins that are associated with plant defense. According to their properties and functions, PR-proteins are currently categorized into 17 families, including β-1,3-glucanases, chitinases, thaumatin-like proteins, peroxidases, ribosome-inactivating proteins, defensins, thionins, nonspecific lipid transfer proteins, oxalate oxidases, and oxalate-oxidase-like proteins (van Loon and van Strien 1999). Among these PR-proteins, chitinases and β-1,3-glucanases are two important classes of hydrolytic enzymes that are up-regulated in many plant species after infection by different types of pathogens. The amount of these two proteins significantly increases upon fungal pathogen detection, and both classes play a role in plant defense by degrading the fungal cell wall. Many plant defense-related proteins were also identified from the xylem sap of field-grown cotton plants.

*Proteases and plant defense*

Secreted proteases have been reported in the xylem stream of several plant species (Buhtz et al. 2004; Alvarez et al. 2006; Djordjevic et al. 2007). Proteases exhibit a broad spectrum of physiological roles, including response to biotic and abiotic stresses (Shabab et al. 2008; Xia 2004; Van der Hoorn and Jones 2004). XCP1 and XCP2 may remain in the xylem sap

even after the death of the xylem vessel elements, potentially as a safety mechanism against eventual pathogen invasion. Additionally, these proteases could be responsible for the slow degradation of secreted proteins.

*Cell wall modification*

Several proteins identified in the cotton xylem sap are functionally associated with cell wall metabolism (Table 2; supplementary Table S1). Many of these proteins are glycosyl hydrolases that are involved in breaking down the cell wall. The identified proteins included beta-glucosidase (GH family 1), endo-1,4-beta-glucanase (GH family 5), endo-1,3-beta-glucanase (GH family 17), polygalacturonase (GH family 28), alpha-xylosidases (GH family 31), and arabinosidase (GH family 51). Secondary cell wall formation in tracheary elements is simultaneously coupled with the degradation of primary cell walls; this process is accompanied with the activity of many cell wall degrading enzymes (Demura et al. 2002).

In summary, this study identified 455 proteins from the xylem sap of field-grown cotton plants. These proteins may have multiple functions during cotton growth and development. They may also play an important role in cotton response to biotic and abiotic environmental stresses. The discovery of secreted proteins and proteins involved in transport provide useful insight into the communication mechanisms between cotton roots and the rest of the plant body.

**Acknowledgment** This research was supported by the National Natural Science Foundation of China (31271648) and the Science and Technology Innovation Talents Project of Henan Province of China (114100510008). This research was also supported by the National High Technology Research and Development Program of China (Program 863) (2011AA10A102), the Industry Technology System of Cotton in Henan China (S2013-07) and the National Genetically Modified Organisms Breeding Technology Major Projects of China (2014ZX08005-002).

## References

Agrawal GK, Jwa N-S, Lebrun M-H, Job D, Rakwal R (2010) Plant secretome: unlocking secrets of the secreted proteins. Proteomics 10(4):799–827

Albenne C, Canut H, Jamet E (2013) Plant cell wall proteomics: the leadership of arabidopsis thaliana. Frontiers in Plant Science 4:1–17

Alvarez S, Goodger JQD, Marsh EL, Chen S, Asirvatham VS, Schachtman DP (2006) Characterization of the maize xylem sap proteome. J Proteome Res 5:963–972

Alvarez S, Marsh EL, Schroeder SG, Schactman DP (2008) Metabolomic and proteomic changes in the xylem sap of maize under drought. Plant Cell Environ 31:325–340

Bacic A, Harris PJ, Stone BA (1988) Structure and function of plant cell walls. In: Preiss J (ed) The biochemistry of plants. Academic, New York, pp 297–371

Bendtsen JD, Jensen LJ, Blom N, Von Heijne G, Brunak S (2004) Feature-based prediction of non-classical and leaderless protein secretion. Protein Eng Des Selection 17(4):349–356

Bendtsen JD, Kiemer L, Fausboll A, Brunak S (2005) Non-classical protein secretion in bacteria. BMC Microbiol 5(1):58

Biles CL, Abeles FB (1991) Xylem sap proteins. Plant Physiol 96(2): 597–601

Bollhöner B, Prestele J, Tuominen H (2012) Xylem cell death: emerging understanding of regulation and function. J Exp Bot. doi: 10.1093/jxb/err438

Buhtz A, Kolasa A, Arlt K, Walz C, Kehr J (2004) Xylem sap protein composition is conserved among different plant species. Planta 219(4):610–618

Buist G, Steen A, Kok J, Kuipers OP (2008) LysM, a widely distributed protein motif for binding to (peptido) glycans. Mol Microbiol 68(4): 838–847

Carpta NC, Gibeaut DM (1993) Structural models of primary cell walls in flowering plants: consistency of molecular structure with the physical properties of walls during growth. Plant J 3(1):1–30

Casasoli M, Spadoni S, Lilley KS, Cervone F, De Lorenzo G, Mattei B (2008) Identification by 2-D DIGE of apoplastic proteins regulated by oligogalacturonides in arabidopsis thaliana. Proteomics 8(5): 1042–1054

Cheng FY, Blackburn K, Lin YM, Goshe MB, Williamson JD (2009) Absolute protein quantification by LC/MSE for global analysis of salicylic acid-induced plant protein secretion responses. J Proteome Res 8(1):82–93

Chivasa S, Simon WJ, Yu XL, Yalpani N, Slabas AR (2005) Pathogen elicitor-induced changes in the maize extracellular matrix proteome. Proteomics 5(18):4894–4904

Cho WK, Chen XY, Chu H, Rim Y, Kim S, Kim ST, Kim SW, Park ZY, Kim JY (2009) Proteomic analysis of the secretome of rice calli. Physiol Plant 135(4):331–341

Cosgrove DJ (1997a) Assembly and enlargement of the primary cell wall in plants. Annu Rev Cell Dev Biol 13:171–201

Cosgrove DJ (1997b) Relaxation in a high-stress environment: the molecular bases of extensible cell walls and cell enlargement. Plant Cell 9(7):1031–1041

Dafoe NJ, Constabel CP (2009) Proteomic analysis of hybrid poplar xylem sap. Phytochemistry 70(7):856–863

Delaunois B, Colby T, Belloy N, Conreux A, Harzen A, Baillieul F, Clément C, Schmidt J, Jeandet P, Cordelier S (2013) Large-scale proteomic analysis of the grapevine leaf apoplastic fluid reveals mainly stress-related proteins and cell wall modifying enzymes. BMC Plant Biol 13:24

Demura T, Tashiro G, Horiguchi G, Kishimoto N, Kubo M, Matsuoka N, Minami A, Nagata-Hiwatashi M, Nakamura K, Okamura Y, Sassa N, Suzuki S, Yazaki J, Kikuchi S, Fukuda H (2002) Visualization by comprehensive microarray analysis of gene expression programs during transdifferentiation of mesophyll cells into xylem cells. Proc Natl Acad Sci 99(24):15794–15799

Djordjevic MA, Oakes M, Li DX, Hwang CH, Hocart CH, Gresshoff PM (2007) The glycine max xylem sap and apoplast proteome. J Proteome Res 6(9):3771–3779

Floerl S, Druebert C, Majcherczyk A, Karlovsky P, Kües U, Polle A (2008) Defence reactions in the apoplastic proteome of oilseed rape (brassica napus var. napus) attenuate verticillium longisporum growth but not disease symptoms. BMC Plant Biology 8:129

Gaspar YM, Johnson KL, McKenna JA, Bacic A, Schultz CJ (2001) The complex structures of arabinogalactan-proteins and the journey towards understanding. Plant Mol Biol 47: 161–176

Gens JS, Fujiki M, Pickard BG (2000) Arabinogalactan proteins and wall associated kinases in a plasmalemmal reticulum with specialized vesicles. Protoplasma 212:115–134

Guan Y, Nothnagel EA (2004) Binding of arabinogalactan proteins by yariv phenylglycoside triggers wound-like responses in arabidopsis cell cultures. Plant Physiol 135(3):1346–1366

Gust AA, Willmann R, Desaki Y, Grabherr HM, Nürnberger T (2012) Plant LysM proteins: modules mediating symbiosis and immunity. Trends Plant Sci 17(8):495–502

Houterman PM, Speijer D, Dekker HL, DE Koster CG, Cornelissen BJC, Rep M (2007) The mixed xylem sap proteome of fusarium oxysporum-infected tomato plants. Mol Plant Pathol 8(2):215–221

Ishihama Y, Sato T, Tabata T, Miyamoto N, Sagane K, Nagasu T, Oda Y (2005) Quantitative mouse brain proteomics using culture-derived isotope tags as internal standards. Nat Biotechnol 23(5):617–621

Jamet E, Albenne C, Boudart G, Irshad M, Canut H, Pont-Lezica R (2008) Recent advances in plant cell wall proteomics. Proteomics 8(4):893–908

Kehr J, Buhtz A, Giavalisco P (2005) Analysis of xylem sap proteins from Brassica napus. BMC Plant Biology 5:11

Ligat L, Lauber E, Albenne C, Clemente HS, Valot B, Zivy M, Pont-Lezica R, Arlat M, Jamet E (2011) Analysis of the xylem sap proteome of Brassica oleracea reveals a high content in secreted proteins. Proteomics 11(9):1798–1813

Liu JS, Zheng SC, Liu L, Li L, Feng QL (2010) Protein profiles of the midgut of spodoptera litura Larvae at the sixth instar feeding stage by shotgun ESI-MS approach. J Proteome Res 9:2117–2147

Minic Z (2008) Physiological roles of plant glycoside hydrolases. Planta 227(4):723–740

Nielsen H, Engelbrecht J, Brunak S, von Heijne G (1997) Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng 10(1):1–6

Nombela C, Gil C, Chaffin WL (2006) Non-conventional protein secretion in yeast. Trends Microbiol 14(1):15–21

Oda Y, Fukuda H (2013) Spatial organization of xylem cell walls by ROP GTPases and microtubule-associated proteins. Curr Opin Plant Biol 16(6):743–748

Oda Y, Hasezawa S (2006) Cytoskeletal organization during xylem cell differentiation. J Plant Res 119(3):167–177

Oda A, Sakuta C, Masuda S, Mizoguchi T, Kamada H, Satoh S (2003) Possible involvement of leaf gibberellins in the clock-controlled expression of XSP30, a gene encoding a xylem sap lectin, in cucumber roots. Plant Physiol 133(4):1779–1790

Passardi F, Penel C, Dunand C (2004) Performing the paradoxical: how plant peroxidases modify the cell wall. Trends Plant Sci 9:534–540

Rep M, Dekker HL, Vossen JH, de Boer AD, Houterman PM, Speijer D, Back JW, de Koster CG, Cornelissen BJC (2002) Mass spectrometric identification of isoforms of PR proteins in xylem sap of fungus-infected tomato. Plant Physiol 130(2):904–917

Rep M, Dekker HL, Vossen JH, de Boer AD, Houterman PM, de Koster CG, Cornelissen BJC (2003) A tomato xylem sap protein represents a new family of small cysteine-rich proteins with structural similarity to lipid transfer proteins. FEBS Lett 534:82–86

Sakuta C, Satoh S (2000) Vascular tissue-specific gene expression of xylem sap glycine-rich proteins in root and their localization in the walls of metaxylem vessels in cucumber. Plant Cell Physiol 41(5): 627–638

Salazar MM, Nascimento LC, Camargo ELO, Gonçalves DC, Neto JL, Marques WL, Teixeira PJPL, Mieczkowski P, Mondego JMC, Carazzolle MF, Deckmann AC, Pereira GAG (2013) Xylem transcription profiles indicate potential metabolic responses for economically relevant characteristics of Eucalyptus species. BMC Genomics 14:201–213

Satoh S (2006) Organic substances in xylem sap delivered to aboveground organs by the roots. J Plant Res 119(3):179–187

Shabab M, Shindo T, Gu C, Kaschani F, Pansuriya T, Chintha R, Harzen A, Colby T, Kamoun S, van der Hoorn RAL (2008) Fungal effector protein AVR2 targets diversifying defense-related Cys proteases of tomato. Plant Cell 20(4):1169–1183

Subramanian S, Cho U-H, Keyes C, Yu O (2009) Distinct changes in soybean xylem sap proteome in response to pathogenic and symbiotic microbe interactions. BMC Plant Biology 9: 119

Taliercio EW, Romano G, Scheffler J, Ayre BG (2009) Expression of genes associated with carbohydrate metabolism in cotton stems and roots. BMC Plant Biology 9:11

Tan L, Showalter AM, Egelund J, Hernandez-Sanchez A, Doblin MS, Antony B (2012) Arabinogalactan-proteins and the research challenged for these enigmatic plant cell surface proteoglycans. Front Plant Sci 3:1–10

The Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant arabidopsis thaliana. Nature 408: 796–815

Ulker B, Somssich IE (2004) WRKY transcription factors: from DNA binding towards biological function. Curr Opin Plant Biol 7:491–498

Van der Hoorn RAL, Jones JDG (2004) The plant proteolytic machinery and its role in defence. Curr Opin Plant Biol 7(4):400–407

Van Loon LC, van Strien EA (1999) The families of pathogenesis related proteins, their activities, and comparative analysis of PR-1 type proteins. Physiol Mol Plant Pathol 55(2):85–97

Wan J, Zhang XC, Neece D, Ramonell KM, Clough S, Kim SY, Stacey MG, Stacey G (2008) A LysM receptor-like kinase plays a critical role in chitin signaling and fungal resistance in arabidopsis. The Plant Cell 20(2):471–481

Wang Y, Li B, Du MW, Eneji AE, Wang BM, Duan LS, Li ZH, Tian XL (2012) Mechanism of phytohormone involvement in feedback regulation of cotton leaf senescence induced by potassium deficiency. J Exp Bot 63(16):5887–5901

Wisniewski JR, Zougman A, Nagaraj N, Mann M (2009) Universal sample preparation method for proteome analysis. Nat Methods 6(5):359–362

Xia Y (2004) Proteases in pathogenesis and plant defence. Cell Microbiol 6(10):905–913

Zhang ZY, Tian XL, Duan LS, Wang BM, He ZP, Li ZH (2007) Differential responses of conventional and Bt-transgenic cotton to potassium deficiency. J Plant Nutr 30:659–670

Zhang ZY, Zhang X, Wang SF, Xin WW, Tang JX, Wang QL (2013) Effect of mechanical stress on cotton growth and development. PLoS ONE 8(12):e82256. doi:10.1371/journal.pone.0082256