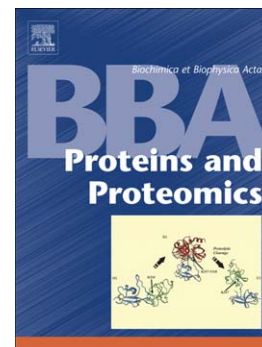# Accepted Manuscript

Plant fluid proteomics: Delving into the xylem sap, phloem sap and apoplastic fluid proteomes

Jorge Rodríguez-Celma, Laura Ceballos-Laita, Michael A. Grusak, Javier Abadía, Ana-Flor López-Millán

Please cite this article as: Jorge Rodríguez-Celma, Laura Ceballos-Laita, Michael A. Grusak, Javier Abadía, Ana-Flor López-Millán, Plant fluid proteomics: Delving into the xylem sap, phloem sap and apoplastic fluid proteomes, *BBA - Proteins and Proteomics* (2016), doi: 10.1016/j.bbapap.2016.03.014

Plant fluid proteomics: delving into the xylem sap, phloem sap and apoplastic fluid proteomes✫

Jorge Rodríguez-Celma [a], Laura Ceballos-Laita [b], Michael A. Grusak [c], Javier Abadía [b], Ana-Flor López-Millán [b,c,*]

[a] *University of East Anglia / John Innes Centre, Norwich Research Park, Norwich NR4 7UH, United Kingdom*

[b] *Department of Plant Nutrition, Aula Dei Experimental Station, Consejo Superior de Investigaciones Científicas (CSIC), P.O. Box 13034, E-50080 Zaragoza, Spain*

[c] *USDA-ARS Children's Nutrition Research Center, Department of Pediatrics, Baylor College of Medicine, 1100 Bates Street, Houston, TX 77030, U.S.A.*

✫*This article is part of a special issue titled: Plant Proteomics: a bridge between fundamental processes and crop production*

[*] *Corresponding author at: USDA-ARS Children's Nutrition Research Center, Department of Pediatrics, Baylor College of Medicine, 1100 Bates Street, Houston, TX 77030, U.S.A.*

E-mail address: lopezmil@bcm.edu (A.-F. López-Millán)

Abstract

The phloem sap, xylem sap and apoplastic fluid play key roles in long and short distance transport of signals and nutrients, and act as a barrier against local and systemic pathogen infection. Among other components, these plant fluids contain proteins which are likely to be important players in their functionalities. However, detailed information about their proteomes is only starting to arise due to the difficulties inherent to the collection methods. This review compiles the proteomic information available to date in these three plant fluids, and compares the proteomes obtained in different plant species in order to shed light into conserved functions in each plant fluid. Inter-species comparisons indicate that all these fluids contain the protein machinery for self-maintenance and defense, including proteins related to cell wall metabolism, pathogen defense, proteolysis, and redox response. These analyses also revealed that proteins may play more relevant roles in signaling in the phloem sap and apoplastic fluid than in the xylem sap. A comparison of the proteomes of the three fluids indicates that although functional categories are somewhat similar, proteins involved are likely to be fluid-specific, except for a small group of proteins present in the three fluids, which may have a universal role, especially in cell wall maintenance and defense.

Keywords: Apoplastic fluid, Phloem sap, Proteome, Xylem sap

1. Introduction

The vascular system of land plants is essential for long- and short-distance transport and distribution of nutrients and signals. The vascular transport system consists of the xylem and phloem conduits, with the apoplast compartment acting as an interface between them and cells and also between cells. The xylem is composed of long treachery elements and shorter cells that are dead at maturity and have primary and secondary cell walls; these elements are connected together forming long tubes known as xylem vessels. The xylem mainly transports water and mineral nutrients taken up by the roots from the soil to the aerial part of the plant, although it also contains a number of other organic compounds, including carboxylates, hormones, amino acids, peptides and proteins [1]. Transport in the xylem is passive and driven by shoot transpiration, which causes a negative pressure that pulls the xylem fluid from the roots, and/or by root pressure, which can occur after soil solution enters the root by osmosis, and causes the sap to move towards the shoot [2]. A major role of the xylem is to provide water and nutrients to the shoots, but it also plays an important role in the root to shoot signaling system.

In contrast, the phloem conduit is a living tissue, whose major role is to transport the photosynthates from photosynthetically active source to sink tissues. The conducting portion of the phloem is composed of two cell types, the sieve elements and the companion cells. At maturity, sieve elements are arranged to form the sieve tube; sieve elements lack a nucleus and contain pores at their ends, forming sieve plates that allow sap flow. The companion cells are metabolically active and usually contain a large number of ribosomes and mitochondria [1]. These cells are in close association with the sieve elements and support their function [3, 4]. The exchange of molecules between these two cells occurs via plasmodesmata [3]. While transport in xylem sap is unidirectional (upwards), transport of the phloem sap is multidirectional and driven

3

by positive hydrostatic pressures created by the gradients in sugar concentration between source and sink tissues. The phloem sap is rich in sugars, but also contains inorganic ions, other organic compounds including proteins and small signaling molecules such as hormones, systemic wound signals and mRNAs [3-5]. Furthermore, the phloem sap is also a conduit for the trafficking of pathogens and foreign compounds such as herbicides and other xenobiotics [4, 6].

The apoplast is the free diffusional space outside the plasma membrane and comprises the cell wall matrix and the fluid in the intercellular spaces. The apoplast occupies less than 5% of the plant tissue volume in aerial organs [7, 8] and the root cortex [9] and its composition reflects the exchange between xylem and phloem compartments and the export and import of molecules by individual cells [7]. Therefore, small changes in these fluxes could produce large differences in the solute composition of the apoplast. The apoplast plays important roles in a wide array of physiological processes, including water and nutrient trafficking [10], plant defense, cell interactions, cell wall maintenance, and in the transduction of environmental and developmental signals [11-15].

The study of plant fluid dynamics can be approached by means of non-destructive methodologies such as those used to measure sap flow velocity and pressure or to visualize vessel structures (reviewed and discussed in [16]). Other approaches include the use of isotope labelling to study fluxes of specific compounds [17]. However, studies of the chemical composition, including their proteomes, often require the use of harsh techniques to collect these plant fluids. While information about the chemical composition of these plant fluids is ample with regard to minerals, sugars and small organic compounds such as carboxylates and amino acids [18], information about their proteomes has arisen more recently. Proteomic studies can provide useful information about processes occurring in plant fluids and also to target proteins

4

putatively involved in them for future studies. Most of the proteomic studies in plants were initially carried out with whole tissue samples [19, 20], and later on with subcellular organelles such as the mitochondria [21], chloroplast [22], and different membrane systems [23-25], as well as with cell walls [26, 27] and plant fluids (see references in Table 1).

A major constraint for the proteomic analysis of plant fluids is the limited amount of sample that can be obtained. In the case of the xylem sap, another constraint is the low protein concentration, which is only in the ng μl^{-1} range. An additional constraint is an inherent consequence of the sampling methodologies: the presence of proteins not intrinsic to these fluids, which can be considered either as contaminants or as the result of protein-protein interactions occurring *in vivo* or during isolation (see below). Methodological progress in the proteomic field has shed light not only into the protein composition of these plant fluids, but also into their functionality. This includes the development of high-throughput technologies such as shotgun proteomics and LC-MS/MS that allow for the analysis of small sample volumes and for the detection of low abundance proteins, as well as the constant improvement in plant proteome databases, which have led to the construction of several mapping datasets (see Table 1 for references).

The aim of this review is to provide a general overview of the proteomic studies carried out in the three plant fluids mentioned. We have considered as outside the scope of this review, which is mainly focused into describing these plant fluid proteomes in healthy plants, a number of differential proteomic studies focused on the specific effects of biotic stresses, including pathogen systemic dissemination and plant defense mechanisms.

2. Plant fluid collection

The main limitations for obtaining reliable plant fluid proteomes have been the technical difficulties in the collection of plant fluids, which usually imply a disruption of the vascular tissues, and the small volumes and sometimes dubious purity of the samples obtained.

Xylem sap is frequently obtained by de-topping plant shoots, cleaning the cut area, and subsequently collecting the fluid that bleeds out, but this requires plant species with sufficient root pressure and turgid stems. This can be done with some dicots, including *Solanum lycopersicum*, *Brassica napus* and *B. oleracea*, and with most of the monocots (*Oryza sativa*, *Zea mays*, etc.) [28]. The collection of phloem sap is commonly achieved by bleeding after performing small incisions in the peduncles of plant species with slow sieve element wound closure. However, there are only a limited number of "phloem bleeder" species, mostly including members of the *Curcubitaceae*, *Brassicaceae* and the genus *Lupinus* [29]. Therefore, usual collection techniques limit the use of model plants such as *Arabidopsis thaliana* for the study of plant fluids, and force the use of non-canonical model plant species, thus delaying progress in the knowledge of these proteomes. The limited number of plant species that can be studied may be not relevant for mineral or metabolomic studies, but it is crucial for proteomic techniques, where the use of non-sequenced species decreases the number of confidently identified proteins.

Amongst the three plant fluids, xylem sap can usually be obtained in sufficient quantities; however, the protein concentration in the xylem sap is generally low, thus limiting the proteomic approaches to be used [30]. Volume yields collected by de-topping shoots of hydroponically-grown healthy plants can reach several ml h^{-1} of acceptably pure xylem sap [31]. The mapping studies of the xylem sap proteome published to date have used root pressure as a means to obtain xylem sap; protein yields ranged from 5 to 12 ng µl^{-1} (Table 1). Although none of these studies assessed the possible sap contamination by cytoplasmic proteins by experimental work, *in silico*

prediction of localization as secreted or intracellular proteins was used in *B. oleracea* to assess the quality of the xylem sap proteome [32]. Several other methods to obtain xylem sap have been published, and their advantages and disadvantages have been reviewed recently [18, 33].

As commented above, the collection of phloem sap is carried out using "phloem bleeder" species, making direct incisions in the phloem vessels located in peduncles, or alternatively by stylectomy using aphids or other sucking insects [33]. The latter technique requires very specialized expertise, provides low volumes of sample (in the nl range), is time-consuming and is only achievable with some plant-insect combinations [29]. Advantages and disadvantages of sampling methods for phloem sap and for purity assessment have been critically reviewed [29, 33]. The phloem sap proteomes published so far have been obtained by collecting exuded sap from direct incisions in *B. napus* [24, 34, 35], *Cucurbita maxima* and in two *Lupin* species [36-38], and by aphid stylectomy in *O. sativa* [30]. With the direct incision method volumes collected were in the several hundred µl range [24, 35, 36], whereas protein yields ranged from 0.2 µg µl^{-1} found in rice and *Lupinus* species [30, 36, 38] to 35-60 µg µl^{-1} in *Curcubitaceae* [29, 39].

The most common methodology used to collect apoplastic fluid is vacuum infiltration-centrifugation (VIC), which can be used with roots, leaves and stems [40-43]. Briefly, the tissue is infiltrated with a buffer while applying a vacuum and the infiltrated intercellular fluid (considered as a diluted apoplastic fluid) is thereafter collected by centrifugation. A study of the variables in this method, including the pH and composition of the buffer, as well as applied centrifugal force and time, indicated that centrifugal forces below 1000 x*g* do not lead to significant contamination, whereas the composition and pH of the solution affect the metabolite composition of the fluid collected [43]. From a proteomic point of view, a detailed study of the

7

influence of the composition of the infiltration solution on the number and nature of proteins collected indicated that the lowest number of intracellular protein contaminants and the largest number of extracellular proteins was obtained in *Z. mays* using 100 mM Na phosphate buffer, whereas water alone led to the highest contamination [44]. In Arabidopsis rosette leaves, the efficiency of several VIC solutions for the collection of apoplastic proteins was also assessed, with 0.2 M $CaCl_2$ yielding the largest number of solubilized proteins among the salts studied [40]. Buffers with different ionic strengths have been applied in order to extract proteins more tightly or loosely bound to the cell wall [45]. Following VIC, approximately 100 µL of apoplastic fluid per g of tissue can be obtained without significant cytosolic contamination (references in Table 1). Direct centrifugation without infiltration can also be applied to obtain apoplastic fluid, leading to similar yields but in general implying a slightly higher contamination [46].

Another major limitation for the study of plant fluid proteomes is their purity. The collection methods, although optimized, cannot completely avoid intracellular fluid leakage. The major source of contamination for these three fluids consists of soluble cytosolic proteins; since the concentration of proteins in the cytosol is usually much higher than in the intercellular space, a relatively small cytosolic leakage (e.g., the rupture of a few cells) could lead to markedly altered protein profiles. Several methods have been used to assess the purity of the plant fluids. Among the xylem mapping studies reviewed (Table 1), purity was assessed only in one study by means of *in silico* prediction of secreted proteins [32]. However, the low protein yields obtained and the fact that only a small amount of cytosolic proteins were identified may suggest an acceptable level of purity. Common phloem sap purity assessment tests are based in the measurement of sugar concentrations as well as the reducing to non-reducing sugar ratio, the presence of

8

Rubisco, and the measurement of certain enzymatic activities as cytosolic contaminant markers [24, 34, 36, 38, 47]. A thorough assessment of purity was performed in the study of the *B. napus* proteome [34], where the purity of the phloem was studied using four independent techniques: sugar composition, western blot and RT-PCR of Rubisco, as well as the comparison of the 2-DE protein maps of the phloem sap with those from full leaf extracts. Given the harsher techniques used for the collection of apoplastic fluid, this plant fluid is more prone to contain products of cytosolic leakage. Apoplastic fluid contamination is usually assessed by measuring the activity of cytosolic enzymes, such as malate dehydrogenase or hexose-phosphate isomerase, as markers [45, 46]. When the extraction buffer used in VIC is not compatible with the enzymatic activity assays, antibody detection of cytosolic contamination has also been used as an alternative [48]. The contamination level measured by these techniques was below 5% in all the studies reviewed. However, the meta-analysis of all the proteins identified in these apoplastic fluid proteomes indicates that approximately 10% of the proteins identified are known to be cytosolic, suggesting that these purity control methods may underestimate the presence of cytoplasmic components.

3. Approach for plant fluid proteome comparisons

Given the difficulties to compare proteomes from different plant species and also the fact that some of the published proteomes have a certain degree of duplicity (e.g., containing the same protein identified in different databases, plant species, or in different gel spots), we have produced a "non-redundant proteome" for the three plant fluids reviewed (included in Supplementary Tables 1-3). BLAST searches of each dataset were performed against the TAIR10 database, the closest Arabidopsis orthologue (E-values lower than 10^{-30}) was selected and duplicates deleted, thus obtaining a unique AGI number per protein. Functional classification of these three non-redundant proteomes was then carried out using the Gene Ontology (GO)

9

annotation included for these proteins in the TAIR database, and the GO categories were manually checked and grouped. The classification includes categories as follows: polysaccharide metabolism, proteolysis, oxido-reductases, general metabolism (including lipid, amino acid, nucleotide and miscellaneous metabolism), regulation, plant defense, unknown, and miscellaneous (including cytoskeleton related, nutrient reservoir, cell division, protein synthesis, photosynthesis, and transport). When the GO annotation was too general, references or functional domains included in the TAIR entry were used as additional criteria. These non-redundant proteomes of xylem sap, phloem sap and leaf apoplastic fluid are presented in Supplementary Tables 1, 2 and 3, respectively.

The subcellular localization database for Arabidopsis proteins (SUBA3, http://suba3.plantenergy.uwa.edu.au) was used to predict protein subcellular localization in the proteome sets. SUBA3 combines manual literature curation of large scale subcellular proteomics, fluorescent protein visualization and protein-protein interaction (PPI) datasets with subcellular targeting output results from 22 prediction software packages [49]. The SUBA3 output includes a consensus location (SUBAcon, Supplementary Table 4), based on Bayesian probabilities, which allows for an automated evaluation of large data sets. In this study, a protein was considered as extracellular when SUBAcon indicated extracellular location. A summary of these results is shown in Supplementary Table 4. Although *in silico* prediction is a powerful tool to analyze large data sets, the existence of non-canonical secretory proteins and the limitations of the prediction algorithms make recommendable using alternative experimental methods to verify the true extracellular nature of proteins classified as such. It should be noted that all three redundant proteomes contain a significant amount of intracellular proteins (43, 99 and 57% in the xylem sap, phloem sap and leaf apoplastic fluid, respectively), which likely reflect some

10

degree of cytoplasmic contamination. Their unequivocal assignment to the respective proteomes deserves further studies.

4. The xylem sap proteome

The xylem sap proteome has been studied in eight plant species, including two monocots, *Z. mays* [50] and *O. sativa* [30], two Brassicaceae species, *Brassica oleracea* [32] and *Brassica napus* [51, 52], two Curcubitaceae, *Cucurbita maxima* and *Cucumis sativus* [51], the legume *G. max* [52, 53] and the Malvaceae *Gossypium hirsutum* [54]. These mapping studies used the 'root pressure' method [31] for xylem sap collection in most species, and 'root pressure' applying vacuum in *G. max* [53]. Different proteomic techniques (1-DE [51, 53, 55], 2-DE [32, 50, 52, 55] and 2D-LC [30, 32, 54]) were used to obtain the xylem sap proteome. The number of xylem sap proteins identified by 1-DE was low and ranged from 14 to 24, whereas the number of proteins identified by 2-DE and 2D-LC ranged from 69 to 154 and from 118 and 455, respectively. In *B. oleracea,* both the xylem sap proteome and N-glycoproteome were studied, the latter one using Concanavalin A affinity chromatography followed by LC-MS/MS [32]. This study reported 81 glycoproteins in the xylem sap, from which 25 were probably low abundance ones, since they were not detected in the direct xylem sap proteome analysis [32].

The comparison of the reviewed xylem sap proteomes led to the assignment of 354 unique AGI proteins (Supplementary Table 1), with 57% of them predicted as extracellular (Supplementary Table 4). Given that the number of proteins obtained in the 1-DE studies was low, only the non-redundant proteomes obtained by 2D techniques in *O. sativa*, *Z. mays*, *B. oleracea*, *B. napus* and *G. hirsutum* were compared in order to identify putative proteins conserved between species. The overlap between them was limited, with no common proteins found in the five species compared and only ten proteins identified in at least four of them (Table

2). Interestingly, this small subset contains only proteins related to the polysaccharide metabolism (5), oxido-reductases (3), plant defense (1) and proteolysis (1) categories (Table 2), suggesting that these are conserved functional categories of the xylem sap proteome and that these specific proteins may have a species-independent role in the xylem sap. However, it should be noted that given the differences in resolution of the proteomes studied and the low coverage in some of the plant species this hypothesis needs further support. Among these ten proteins, nine were predicted by SUBA3con to be extracellular (Table 2).

The functional classification of the 354 non-redundant proteins of the xylem sap proteome is presented in Figure 1. The major functional categories found were polysaccharide metabolism (28%), general metabolism (22%), oxido-reductases (14%) and proteolysis (12%). Many (approximately 65%) of the enzymes identified in the polysaccharide metabolism category are cell wall hydrolases (glucanases, galactosidases and xylosidases) involved in primary cell wall degradation and cell wall remodeling [56, 57]. This degradation is needed for the formation of secondary cell walls [58]. This group of enzymes can also play a role in modifying the cell wall surface in conjunction with pectic enzymes and hemicellulases [59, 60]. Lignin and secondary cell wall provide structural support for xylem vessels, being essential for its function [61]. Lignin deposition requires polymerization and de-polymerization of monolignols from the phenylpropanoid metabolism, and these reactions involve hydrogen peroxide [62].

The second most important functional group found is general metabolism, including proteins related to carbohydrate (22), lipid (20), nucleotide (seven) and amino acid metabolisms (ten) (Supplementary Table 1). There are several possible explanations for the presence of such a large number of metabolic enzymes in xylem sap. The first plausible explanation would be that they are cytosolic contaminants due to the sampling process. The fact that most of them are not

12

present in the overlaps (Table 2) and that only 47% of them are predicted to be extracellular (Supplementary Tables 1 and 4) would argue in favor of this explanation. Another possibility would be that at least some of these enzymes are needed for the production of cellulose and lignin; these include carbohydrate metabolism enzymes, phenylalanine ammonia-lyase and 4-coumarate-CoA ligase [57]. Further studies are needed to clarify the presence and functions of these proteins in the xylem sap.

Proteins classified as oxido-reductases include several enzymes of the ascorbate-glutathione cycle, Cu,Zn-SOD and several redox enzymes such as thioredoxin H and glutaredoxin (Supplementary Table 1). A large number of peroxidases (26 of the 50 identified proteins that belong to the oxido-reductase category) were also found in the studies reviewed (Supplementary Table 1), and although most authors in the studies assigned them to redox response, they could well be associated also to cell wall processes and lignification. Approximately 90% of these peroxidases were predicted to be extracellular (Supplementary Tables 1 and 4). The presence of proteins involved in the response to redox homeostasis in the xylem sap proteome of non-stressed plants supports the existence of a constitutive basal defense mechanism, which may be especially useful given that xylem cells do not have the protein machinery to synthesize proteins when submitted to a stress situation.

Proteolysis related proteins accounted for 12% of the non-redundant xylem proteome. Most of the proteases, including pepsins, subtilisins, and serine and cysteine peptidases are known to have an extracellular localization, and accordingly 63% of them were classified as extracellular. In addition to a role in the degradation of damaged proteins for the maintenance of xylem sap, extracellular proteases may also constitute a basal defense strategy against pathogen infection [63, 64]. Plant defense related proteins accounted for 5% of the non-redundant xylem proteome.

13

Many plant defense proteins, such as thaumatins and chitinases, have been found in the xylem sap of different plant species [51, 65], and they are consistently found in the proteomic studies reviewed. This already established defense system found in plant fluids is thought to create a hostile environment to prevent pathogen infection [48, 66].

In summary, from a proteomic point of view the bulk xylem sap composition is relatively simple. It contains proteins involved in maintaining the xylem conduit architecture, as well as a battery of proteins, including redox and plant defense ones in addition to proteases, which constitutes a basic defense barrier that provides a rapid and local first response against possible stress situations and pathogen infection.

5. The phloem sap proteome

A large number of proteomic studies have focused on deciphering the phloem sap proteome. Initial studies used 1-DE approaches and yielded a limited number of proteins [41, 67]. More recently, the development of high throughput proteomic methods, such as two dimensional electrophoresis (2-DE) and especially LC-MS/MS, have allowed a more thorough characterization of the phloem sap proteome. Phloem sap in these studies was obtained either by using stylectomy in *O. sativa* [30] or by sap exudation after incision in *Ricinus communis* [68], *B. napus* [24, 34], *C. maxima* [37, 39], *C. sativus* [38], *Populus trichocarpa* x *Populus deltoides* [69], *Lupinus albus* [38] and *Lupinus texensis* [36]. The number of phloem sap proteins identified by 2-DE was similar, irrespective of the gel size, and ranged from 45 in *C. maxima* to 140 in *B. napus*, with 86 and 54 identified in *L. albus* and *L. texensis*, respectively. The number of phloem proteins identified by LC-MS/MS was larger, and included 1121 proteins in *C. maxima* and 107 in *O. sativa*.

14

Comprehensive comparisons of the phloem sap proteomes of these different plant species have already been performed [36, 37]. These inter-species comparisons were approached by using the phloem protein data sets from *C. maxima*, *O. sativa*, *B. napus* and *R. communis* to perform BLAST searches against *A. thaliana* and by creating a "total non-redundant phloem proteome" containing 688 proteins (Fig. 2A in [37]). A similar approach was used to compare the phloem proteome of *L. texensis* to the previously reported non-redundant proteome and found only 12 new protein species [36]. From these, two proteins were previously described in *L. albus* [38] and the remaining ten novel protein species had non-redundant identifiers, but had functions similar to other protein species already described in the phloem of other plant species (Fig. 2C in [36]). The comparison of the reviewed phloem sap proteomes led to the assignment of 720 unique AGI proteins (Figure 2; Supplementary Tables 2 and 4). Comparisons of the non-redundant phloem proteomes previously described [36, 37] revealed that approximately 60 proteins were found in the phloem sap of at least two plant species, with 13 of them being consistently detected in at least four of the five plant species considered (Table 3). Among these 13 proteins, none were predicted to be extracellular (Table 3). Proteins in this list were representative of the most abundant functional categories found in the phloem sap proteome (see below).

The functional classification and cellular localization of the non-redundant phloem sap proteome (Figure 2; Supplementary Tables 2 and 4) clearly reflects the complex nature of this fluid, which results from the exchange between companion cells and sieve elements along the translocation pathway. On one hand, categories such as oxido-reductases (5%), plant defense (3%), and proteolysis (12%) are enriched in this fluid, as also seen in the xylem sap and apoplastic fluid (see references in Table 1). These categories provide a basal defense mechanism

against pathogens and stress or are involved in the maintenance and stability of the sieve elements. On the other hand, other enriched categories such as general metabolism (28%) and miscellaneous (24%) [the latter including the sub-categories protein synthesis (9%) and macromolecular trafficking (8%)] reflect exchange processes occurring between the metabolically active companion cells and the enucleate sieve elements (Figure 2 and Supplementary Table 2). The category of regulation, including RNA binding, is also enriched (12%) in the phloem sap proteome and may represent proteins that are being translocated (references in Table 1). Furthermore, *in silico* prediction of the localization of the non-redundant phloem proteome indicated that 99% of them are intracellular (Supplementary Tables 2 and 4), suggesting that the contribution of companion cells to the phloem sap proteome is very significant, or alternatively that there may be some contamination from other cell types.

Proteins in the oxido-reductase category typically found in phloem sap include enzymes of the ascorbate-glutathione cycle, glutaredoxins and gluthatione transfereases, with thioredoxin being one of the most abundant proteins in this group in most plant species (Supplementary Table 2). Defense related proteins include pathogenesis related proteins, LRR disease resistance proteins and several hormone responsive proteins.

In addition to extracellular proteases such as cysteine or serine proteases (16 of the 26 phloem proteases were predicted to be extracellular), components of the proteasome 26S were also found in the non-redundant proteome, although none of them were predicted as extracellular. Furthermore, the *C. maxima* proteome, the most detailed proteome described so far, contains all the components to form a functional 26S proteasome as well as a large set of proteins related to ubiquitin related processes [37], suggesting that the sieve tube system has retained the capacity for proteolysis *via* de ubiquitin-proteasome 26S pathway. The presence of protein degrading

16

enzymes in sieve elements could be related to protein turnover or also constitute a defense mechanism. Cell structural components such as actin and profilin and actin-depolymerizing factors are also common proteins in the non-redundant phloem sap proteome (23 cytoskeleton related proteins) (Supplementary Table 2 and references in Table 1). In *R. communis*, these proteins are constantly delivered into the sieve elements and are mobile in the translocation stream [67], but their specific role in sieve elements remains to be determined.

A general overview of the proteins identified in different plant species in the general metabolism category indicates the existence of a wide array of metabolic processes in phloem sap, including carbohydrate metabolism, synthesis and breakdown of organic-, amino- and fatty-acids as well as secondary metabolites and nucleotides (Supplementary Table 2, references in Table 1), suggesting that these processes might still be active in sieve elements. However, only 1.5% (three out of the 200 proteins in this category; Supplementary Table S2) were classified as intracellular, highlighting the special nature of the phloem sap proteome and raising doubts about their possible origin.

A large array of molecules related to regulation (12%) has been described in the phloem sap proteome. The regulation category would include those proteins involved in protein modification and folding and those acting as mobile signals for systemic signaling. The first group includes chaperones such as peptidyl-prolyl isomerases, several heat shock proteins and proteins involved in phosphorylation such as protein kinases, and in Ca-dependent signaling such as annexins, calmodulins and C2 domain containing proteins (Supplementary Table 2, references in Table 1). Overall, these proteins are involved in signal processing, protein phosphorylation, folding and unfolding of proteins for protein trafficking between companion cells and sieve elements and microRNA regulation of gene expression [38]. Among putative mobile signals, the flowering

17

locus FT is ubiquitously found in all plant species (Table 3) [30, 34]. The presence of extracellular RNA binding proteins such as cyclophilins (one of them, cyclophilin 19, although predicted as intracellular, was found in all species compared here, Table 3), lectins, and ~~the~~ glycine rich proteins in the phloem sap is consistent with RNA transport in phloem as ribonucleoprotein complexes, which are most likely involved in long distance signaling. Secreted lectins have been shown to translocate RNAs [70]. Many of the glycine-rich proteins have been associated with stress and hormone treatments [71], supporting their possible role in RNA homeostasis and signaling [72, 73], whereas some of the glycine-rich proteins (such as AT2G21660 and AT3G03773 found in the non-redundant phloem proteome, Supplementary Table 2) are possibly involved in the formation of secondary walls [74].

Finally, the non-redundant phloem proteome contains 67 proteins related to protein synthesis and only eight were structural components of the ribosome (Supplementary Table 2); in addition, none of the proteins in this sub-category were predicted as extracellular (Supplementary Table 2). Most of these proteins come for the high-throughput analysis of the *C. maxima* proteome [37], which contains around 100 proteins involved in t-RNA aminoacylation, translation initiation, elongation, and termination, but very few components of the ribosomal complexes, which prompted the authors to suggest the existence of protein synthesis in the sieve element. A limited number of translation initiation factors have been detected in several plant species, with one of them, the intracellular eiF-5A, being found in all species compared (Table 3). Given that some of them bind RNA, including eiF-5A, a role in RNA transport in phloem has also been suggested for these proteins in phloem sap [34]. The large number of this set of proteins in *C. maxima* could reflect a larger contribution of companion cells in this proteome, the high sensitivity of the proteomic approach used in this study, or to the specific characteristics of the

18

phloem sap in *Curcubitaceae* species, which is mostly composed of extra-fascicular phloem sap and may have a peculiar composition [35, 75].

6. The leaf apoplastic fluid proteome

The proteome of the leaf apoplastic fluid has been described in six plant species so far (Table 1). This fluid was collected by the VIC technique using buffers of different ionic strengths in four of them: *Z. mays* [44], *Medicago truncatula* [45], *Vitis vinifera* [48] and *A. thaliana* [40]. In *Populus deltoides*, the leaf apoplastic fluid was collected using a pressure chamber [66] and in *Beta vulgaris* it was obtained by direct leaf centrifugation [46]. Protein yields were only reported in three studies, including sugar beet apoplastic fluid obtained by direct leaf centrifugation (400-800 µg protein mL^{-1}; [46]) and grapevine and *M. truncatula* apoplastic fluids obtained by VIC (6 and 50 µg protein g^{-1} FW, respectively) [45, 48]. 2-DE analysis identified a similar number of protein spots: 328, 220, 306, and 203 in the apoplastic fluid of *Z. mays, M. truncatula, V. vinifera* and *B. vulgaris,* respectively [44-46, 48]; whereas the combination of 2-DE and 2D-LC in *P. deltoides* [66] and of 2-DE and sepharose fractioning followed by 1-DE in *A. thaliana* [40] yielded a somewhat lower number of apoplastic proteins: 144 and 133, respectively. However, it should be noted that redundancy existed in the 2-DE approaches, and the number of unique proteins identified was in the one hundred range (Table 1). In addition to these mapping studies there is a significant number of studies on the effect of biotic stresses in the leaf apoplast proteome, which as stated above are out of the scope of this review [76-81].

The comparison of the leaf apoplastic fluid proteomes led to the assignment of 373 unique AGI proteins (Supplementary Table 3), with 43% being predicted as extracellular (Supplementary Table 4). Although no AGIs were found in all plant species reviewed, the overlap between these leaf apoplastic fluid proteomes was significant, with 77 non-redundant

AGIs identified in at least two plant species (Supplementary Table 3) and a total of 25 proteins found in at least three plant species (Table 4). This small subset contains 22 extracellular proteins related to polysaccharide metabolism (13), oxido-reductases (three), proteolysis (three), plant defense (two), and general metabolism (one) categories, suggesting that these are conserved functionalities of the leaf apoplastic fluid (Table 4). Two chloroplast proteins were found among the non-redundant proteins present in at least two plant species (Table 4) and this underlines the difficulty of obtaining pure apoplastic fluid. Within this subset, *Z. mays*, *P. deltoides* and *V. vinifera* were the plant species containing the most proteins in common (Table 4).

Functional classification of this non-redundant leaf apoplastic fluid proteome showed that general metabolism was the most abundant category (27%) followed by polysaccharide metabolism (21%), proteolysis (12%), oxido-reductase (11%), plant defense (7%), and regulation (7%) (Figure 3). Some of the general metabolism proteins (70%) have an intracellular localization and most likely reflect some degree of cytoplasmic contamination; this is also clear for the 23 photosynthesis related proteins present in this non-redundant proteome (Supplementary Table 3). These proteins were found in all the studies regardless of their purity assessment, but percentages were higher when pressure chamber (18%) or direct leaf centrifugation (25%) were used, suggesting that leaf VIC may be better at preventing leakage of proteins from the cytoplasm [43, 44]. The overlapping degree of these metabolism-related proteins between species was low (Supplementary Table 3), suggesting that this category might be over-represented in the non-redundant proteome that compiles proteins from all plant species. Some of the proteins in this group, such as the six lipid transfer proteins and eight extracellular lipases, are indeed annotated as secretory proteins and their presence in the apoplastic fluid may suggest an active role of this fluid in the maintenance of the leaf cell plasma membranes and/or

in cuticle biogenesis [82]. The presence of proteins related to this array of metabolic pathways deserves further studies to clarify the intrinsic metabolic processes occurring in the apoplast. As described for xylem sap, polysaccharide related proteins accounted for a large proportion of the leaf apoplastic fluid proteome (21%), with glucosidases and galactosidases being the two larger families present (Supplementary Table 3). Typical cell wall proteins such as α-mannosidase, chitinase, β-1,3-D-glucosidase, xyloglucan, heparanase, endotransglucosylase-hydrolase and α- and β-galactosidases (Supplementary Table 3) have been identified as prominent constituents in different studies, playing various roles in cell wall modification during plant development [44, 83]. The presence of a high number of hydrolases in this non-redundant proteome confirms that the apoplast is a highly glycolytic compartment [11].

Plant defense and proteolysis accounted for approximately 19% of the leaf apoplastic fluid proteome. The plant defense category contains a large number of pathogenesis related proteins (five; four of them extracellular), and disease resistance proteins containing LRR domains (seven; five of them extracellular) (Supplementary Table 3), which frequently increase in xylem sap and apoplastic fluids when plants experience biotic stresses [76, 84-87] and, as commented for the xylem sap, constitute a basal plant defense mechanism in non-stressed conditions. The disease resistance proteins with LRR domains may also be involved in signaling [40]. Other proteins in this category included extracellular osmotin-like (AT4G11650) and a basic secretory protein (AT2G15130) (Supplementary Table 3). The proteolysis-related proteins included 24 extracellular peptidases, such as serine-carboxypeptidases, cysteine and aspartic proteases, subtilisin-like serine proteases, and four proteinase inhibitors, most of them similar to cystatin (Supplementary Table 3). Four proteins related to the proteasome, although not predicted as extracellular, were also found in the non-redundant apoplastic fluid proteome; while it is clear

that the apoplastic fluid is enriched in extracellular proteases, the presence of proteasome components was only detected in the sugar beet apoplastic fluid proteome [46]. It has been hypothesized that the proteolytic composition of the apoplastic fluid is species-dependent [44], and therefore the presence of some components of the proteasome in sugar beet may be species-specific or reflect a possible intracellular contamination.

Among the proteins classified as oxido-reductases (11%), 19 were predicted to be extracellular. These included a large number of peroxidases (ten), cupro-redoxins (three), and FAD-binding proteins (three), among others (Supplementary Table 3). Furthermore, three peroxidases were found in three or more plant species (Table 4) suggesting that their presence is highly conserved in this fluid. Five thioredoxins were also found in the non-redundant apoplastic fluid proteome, although only one of them was predicted as extracellular. Thioredoxins are also abundant in the phloem sap proteome [88] and, in addition to redox homeostasis, some thioredoxins have been associated with inter-cellular communication and may be required for redox signaling [89] (Supplementary Table 3).

Interestingly, as it occurs in phloem sap, the apoplastic fluid also contains proteins involved in regulation including RNA binding (7%), although only two proteins were predicted as extracellular. This set of extracellular proteins included an auxin-independent growth promoter-like protein (AT5G63390) and an homolog of CLAVATA 1b (a putative receptor kinase) (Supplementary Table S3). The later one was probably identified as a proteolytic fragment belonging to the extracellular domain [40]. The presence of these proteins in the apoplast may suggest a role of this compartment in signaling that deserves further study.

In summary, the functional categorization of the apoplastic fluid proteome indicates that a large part of this proteome is involved in basal defense and in the maintenance of the cell wall

22

and possibly the leaf cell plasma membrane. The functional categorization also reveals a possible role in regulation and points out two candidates which deserve further study. However, it should be kept in mind that 57% of the non-redundant apoplastic fluid proteome is predicted to be intracellular and therefore some of these enriched functional categories may arise from cytoplasmic contamination and deserve additional studies.

7. The root apoplastic fluid proteome

To the best of our knowledge, a comprehensive description of the root apoplastic fluid proteome has not been published yet. There are two studies in rice describing the effects of H_2O_2 and salt stress on the proteome of the root apoplastic fluid [90, 91]. Both studies used VIC and 2-DE to obtain 400 and 100 spots using 24 and 13 cm IPG strips, respectively. Although mapping was not performed in these studies, proteins identified as responsive to NaCl and H_2O_2 belong to two of the categories, defense and redox, described in the leaf apoplastic fluid proteome.

8. Conclusion and outlook

The overlap of the non-redundant plant fluid proteomes (Figure 4) indicates that xylem and apoplastic fluid share more proteins between them (117) than with the phloem sap (40 and 38 proteins in common in each comparison, respectively), highlighting the peculiar composition of the phloem sap proteome. Although the overlap in unique identifiers is somewhat limited, there are functional categories, including polysaccharide metabolism, defense, proteolysis and oxido-reductases that occur in the three fluids, altogether accounting for large percentages of their proteomes. These categories indicate that plant fluids contain the machinery needed for the maintenance of xylem vessels, phloem sieve elements and cell walls, and for basal defense processes. It is also interesting to note that a significant number of proteins in these categories, such as peroxidases or proteases, have dual roles in both defense and structural maintenance. The

23

relatively low number of common identifiers may indicate that the specific proteins in charge of these functions are fluid-specific. On the other hand, the category regulation is enriched to a higher extent in phloem sap (Fig. 2) than in xylem sap (Fig. 1) and apoplastic fluid (Fig. 3), highlighting the role of phloem sap in long distance signaling and communication. The comparison of the non-redundant proteomes reveals 14 proteins commonly found in the three plant fluids (Table 5, Fig. 4), among them, four of them being predicted as extracellular. These included three extracellular glycosidase hydrolases from the families 17, 18 and 38 as well as peroxidase 54, indicating a possible universal role of these proteins in cell wall maintenance and defense. Functional classification of the remaining ten intracellular proteins indicates that plant fluid proteomes share proteins related to general metabolism (three proteins), amino acid metabolism (two), oxido-redeductases (two), as well as in nucleotide metabolism (nucleoside diphosphate kinase 1) and photosynthesis (small subunits 3B and 1A of Rubisco). These intracellular proteins may reflect intracellular contamination and their presence in these fluids needs to be further addressed. The pairwise overlaps indicate that xylem and apoplastic fluids share a large number of polysaccharide metabolism related proteins (41 proteins, accounting for 35% of the shared proteins), whereas phloem and xylem or apoplast share proteins related to general metabolism and oxido-reductase categories, accounting for approximately 50% of the common proteins (Fig. 4).

Extracellular proteins accounted for 57, 43 and 1% of the xylem sap, leaf apoplastic fluid and phloem sap proteomes, respectively. These data clearly distinguish the phloem sap proteome from those of the leaf apoplastic fluid and xylem sap. When specifically looking at peroxidases as an example, the phloem sap non-redundant proteome only contained four peroxidases, with one of them being extracellular, whereas xylem sap and apoplastic fluids contained a larger

24

number of peroxidases, most of them being extracellular. These results also point out that significant percentages of these non-redundant proteomes are predicted to be intracellular, and therefore additional experiments are needed to ascertain the presence and possible role of these proteins in plant fluids.

The compilation of the studies reviewed has allowed us to obtain a composite picture of each plant fluid proteome and has clarified major functions occurring in these compartments. The comparison between these three plant fluid proteomes highlights their differences and commonalities, thus providing hints into possible functionalities specific and shared by them and has pointed out several candidates for which their functions in these compartments deserve further studies (i.e., proteins presented in Tables 2, 3, 4 and 5, especially the extracellular ones).

However, there are still many gaps that need to be addressed to obtain a complete picture. On one hand, the use of the high-throughput and more sensitive LC-MS/MS proteomic approaches can increase the inventory of these proteomes. On the other hand, ProteoMiner protein enrichment technologies [92, 93] may help to reveal low abundance proteins such as those involved in signaling that might be under-represented in protein extracts, but that still play important roles in the long-distance communication system. Indeed, this technique has been recently used to study root cell wall proteins in Arabidopsis allowing one to identify low abundance proteins never described before in cell walls [94]. These studies need to be performed in a variety of plant species to reveal fluid specific and species-specific processes occurring in these plant fluids. Furthermore, given the high enrichment of these plant fluids in glycoproteins, specific studies of this sub-proteome and progress in bioinformatics tools to predict plant-specific glycoproteins will provide further insights into these plant fluids.

Given the invasive nature of the methods used for the collection of these fluids, clear criteria and well-established protocols to assess their purity are needed, including the establishment of thresholds for contamination and the use of data validation via complementary techniques such as Western blot, immuno- and GFP-localization techniques. Although the coverage of plant fluid proteomes has improved considerably, its integration with other high-throughput techniques is still lacking. Proteomic data combined with small RNA sequencing and metabolomics using a systems biology approach could provide a holistic view of plant fluids and provide new insights into their functions. Finally, while there is no doubt that some proteins found in phloem and xylem sap are being translocated, this information is lacking for other proteins and experiments are needed to clearly demonstrate that these proteins are being unloaded into the target tissues. This is especially relevant for putative signaling proteins for which, in addition to delivery, a function in the sink tissue needs to be proven.

Acknowledgements

References

[1] W.J. Lucas, A. Groover, R. Lichtenberger, K. Furuta, S.R. Yadav, Y. Helariutta, X.Q. He, H. Fukuda, J. Kang, S.M. Brady, J.W. Patrick, J. Sperry, A. Yoshida, A.F. López-Millán, M.A. Grusak, P. Kachroo, The plant vascular system: evolution, development and functions, J. Integr. Plant Biol., 55 (2013) 294-388.

[2] D.B. Fisher, Long-distance transport, in: B. Buchanan, W. Gruissem, R. Jones (Eds.) Biochemistry and molecular biology of plants, American Society of Plant Physiologists, Rockville, MD, USA, 2000, pp. 729-784.

[3] K.J. Oparka, S.S. Cruz, The great escape: phloem transport and unloading of macromolecules, Annu. Rev. Plant Physiol. Plant Mol. Biol., 51 (2000) 323-347.

[4] A.J.E. Van Bel, The phloem, a miracle of ingenuity, Plant Cell Environ., 26 (2003) 125-149.

[5] G.A. Thompson, A. Schulz, Macromolecular trafficking in the phloem, Trends Plant Sci., 4 (1999) 354-360.

[6] O. Fiehn, Metabolic networks of *Cucurbita maxima* phloem, Phytochemistry, 62 (2003) 875-886.

[7] D. Parkhurst, Stereological methods for measuring internal leaf structural variables, Am. J. Bot., 69 (1982) 31–39.

[8] S.J. Steudle E, Lüttge U, Water-relation parameters of individual mesophyll cells of *Kalanchöe daigremontiana*, Plant Physiol., 66 (1980) 1155–1163.

[9] D.B. Vakhmistrov, Localization of the free space in the barley roots., Russ. J. Plant Physiol., 14 (1967) 397-404.

[10] B. Sattelmacher, K.-H. Mühling, K. Pennewiß, The apoplast - its significance for the nutrition of higher plants, Z. Pflanz. Bodenkunde, 161 (1998) 485-498.

[11] K.-J. Dietz, A. Sauter, K. Wichert, D. Messdaghi, W. Hartung, Extracellular β-glucosidase activity in barley involved in the hydrolysis of ABA glucose conjugate in leaves, J. Exp. Bot., 51 (2000) 937-944.

[12] C.M. Geilfus, A. Mithofer, J. Ludwig-Muller, C. Zorb, K.H. Muehling, Chloride-inducible transient apoplastic alkalinizations induce stomata closure by controlling abscisic acid distribution between leaf apoplast and guard cells in salt-stressed *Vicia faba*, New Phytol., 208 (2015) 803-816.

[13] T. Hoson, Apoplast as the site of response to environmental signals, J. Plant Res., 111 (1998) 167-177.

[14] N. Sakurai, Dynamic function and regulation of apoplast in the plant body, J. Plant Res., 111 (1998) 133-148.

[15] K. Ohyama, H. Shinohara, M. Ogawa-Ohnishi, Y. Matsubayashi, A glycopeptide regulating stem cell fate in *Arabidopsis thaliana*, Nat. Chem. Biol., 5 (2009) 578-580.

[16] H.K. Kim, J. Park, I. Hwang, Investigating water transport through the xylem network in vascular plants, J. Exp. Bot., 65 (2014) 1895-1904.

[17] I. Orera, J.A. Rodríguez-Castrillón, M. Moldovan, J.I. García-Alonso, A. Abadía, J. Abadía, A. Álvarez-Fernández, Using a dual-stable isotope tracer method to study the uptake, xylem

transport and distribution of Fe and its chelating agent from stereoisomers of an Fe(III)-chelate used as fertilizer in Fe-deficient Strategy I plants, Metallomics, 2 (2010) 646-657.

[18] A. Álvarez-Fernández, P. Diaz-Benito, A. Abadía, A.F. López-Millán, J. Abadía, Metal species involved in long distance metal transport in plants, Front. Plant Sci., 5 (2014) 105.

[19] J.V. Jorrin-Novo, J. Pascual, R. Sanchez-Lucas, M.C. Romero-Rodriguez, M.J. Rodriguez-Ortega, C. Lenz, L. Valledor, Fourteen years of plant proteomics reflected in Proteomics: moving from model species and 2DE-based approaches to orphan species and gel-free platforms, Proteomics, 15 (2015) 1089-1112.

[20] A.F. López-Millán, M.A. Grusak, A. Abadía, J. Abadía, Iron deficiency in plants: an insight from proteomic approaches, Front. Plant Sci., 4 (2013) 254.

[21] S. Huang, R.P. Jacoby, A.H. Millar, N.L. Taylor, Plant mitochondrial proteomics, Methods Mol. Biol., 1072 (2014) 499-525.

[22] M. Ferro, S. Brugiere, D. Salvi, D. Seigneurin-Berny, M. Court, L. Moyet, C. Ramus, S. Miras, M. Mellal, S. Le Gall, S. Kieffer-Jaquinod, C. Bruley, J. Garin, J. Joyard, C. Masselon, N. Rolland, AT_CHLORO, a comprehensive chloroplast proteome database with subplastidial localization and curated information on envelope proteins, Mol. Cell. Proteomics, 9 (2010) 1063-1084.

[23] G. Ephritikhine, M. Ferro, N. Rolland, Plant membrane proteomics, Plant Physiol. Biochem., 42 (2004) 943-962.

[24] E. Gutierrez-Carbonell, G. Lattanzio, A. Albacete, J.J. Rios, J. Kehr, A. Abadía, M.A. Grusak, J. Abadía, A.F. López-Millán, Effects of Fe deficiency on the protein profile of *Brassica napus* phloem sap, Proteomics, (2015).

[25] A. Marmagne, M.A. Rouet, M. Ferro, N. Rolland, C. Alcon, J. Joyard, J. Garin, H. Barbier-Brygoo, G. Ephritikhine, Identification of new intrinsic proteins in *Arabidopsis* plasma membrane proteome, Mol. Cell. Proteomics, 3 (2004) 675-691.

[26] C. Albenne, H. Canut, E. Jamet, Plant cell wall proteomics: the leadership of Arabidopsis thaliana, Front. Plant Sci., 4 (2013) 111.

[27] J.K. Rose, S.J. Lee, Straying off the highway: trafficking of secreted plant proteins and complexity in the plant cell wall proteome, Plant Physiol., 153 (2010) 433-436.

[28] U. Schurr, Xylem sap sampling-new approaches to an old topic, Trends Plant Sci., 3 (1998) 293-298.

[29] C.A. Atkins, P.M. Smith, C. Rodriguez-Medina, Macromolecules in phloem exudates- a review, Protoplasma, 248 (2011) 165-172.

[30] T. Aki, M. Shigyo, R. Nakano, T. Yoneyama, S. Yanagisawa, Nano scale proteomics revealed the presence of regulatory proteins including three FT-Like proteins in phloem and xylem saps from rice, Plant Cell Physiol., 49 (2008) 767-790.

[31] J.Q. Goodger, R.E. Sharp, E.L. Marsh, D.P. Schachtman, Relationships between xylem sap constituents and leaf conductance of well-watered and water-stressed maize across three xylem sap sampling techniques, J. Exp. Bot., 56 (2005) 2389-2400.

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

[33] J. Kehr, M. Rep, Protein extraction from xylem and phloem sap, Methods Mol. Biol., 355 (2007) 27-35.

[34] P. Giavalisco, K. Kapitza, A. Kolasa, A. Buhtz, J. Kehr, Towards the proteome of *Brassica napus* phloem sap, Proteomics, 6 (2006) 896-909.

[35] C. Zhang, X. Yu, B.G. Ayre, R. Turgeon, The origin and composition of cucurbit "phloem" exudate, Plant Physiol., 158 (2012) 1873-1882.

[36] G. Lattanzio, S. Andaluz, A. Matros, J.J. Calvete, J. Kehr, A. Abadía, J. Abadía, A.F. López-Millán, Protein profile of *Lupinus texensis* phloem sap exudates: searching for Fe- and Zn-containing proteins, Proteomics, 13 (2013) 2283-2296.

[37] M.K. Lin, Y.J. Lee, T.J. Lough, B.S. Phinney, W.J. Lucas, Analysis of the pumpkin phloem proteome provides insights into angiosperm sieve tube function, Mol. Cell. Proteomics, 8 (2009) 343-356.

[38] C. Rodriguez-Medina, C.A. Atkins, A.J. Mann, M.E. Jordan, P.M. Smith, Macromolecular composition of phloem exudate from white lupin (*Lupinus albus* L.), BMC Plant Biol., 11 (2011) 36.

[39] C. Walz, P. Giavalisco, M. Schad, M. Juenger, J. Klose, J. Kehr, Proteomics of curcurbit phloem exudate reveals a network of defence proteins, Phytochemistry, 65 (2004) 1795-1804.

[40] G. Boudart, E. Jamet, M. Rossignol, C. Lafitte, G. Borderies, A. Jauneau, M.-T. Esquerré-Tugayé, R. Pont-Lezica, Cell wall proteins in apoplastic fluids of *Arabidopsis thaliana* rosettes: Identification by mass spectrometry and bioinformatics, Proteomics, 5 (2005) 212-221.

[41] V. Dani, W.J. Simon, M. Duranti, R.R. Croy, Changes in the tobacco leaf apoplast proteome in response to salt stress, Proteomics, 5 (2005) 737-745.

[42] S. Husted, J.K. Schjoerring, Apoplastic pH and ammonium concentration in leaves of *Brassica napus* L, Plant Physiol., 109 (1995) 1453-1460.

[43] G. Lohaus, K. Pennewiss, B. Sattelmacher, M. Hussmann, K. Hermann Muehling, Is the infiltration-centrifugation technique appropriate for the isolation of apoplastic fluid? A critical evaluation with different plant species, Physiol. Plant., 111 (2001) 457-465.

[44] K. Witzel, M. Shahzad, A. Matros, H.-P. Mock, K.H. Mühling, Comparative evaluation of extraction methods for apoplastic proteins from maize leaves, Plant Methods C7 - 48, 7 (2011) 1-11.

[45] N.C. Soares, R. Francisco, C.P. Ricardo, P.A. Jackson, Proteomics of ionically bound and soluble extracellular proteins in *Medicago truncatula* leaves, Proteomics, 7 (2007) 2070-2082.

[46] L. Ceballos-Laita, E. Gutierrez-Carbonell, G. Lattanzio, S. Vázquez, B. Contreras-Moreira, A. Abadía, J. Abadía, A.F. López-Millán, Protein profile of *Beta vulgaris* leaf apoplastic fluid and changes induced by Fe deficiency and Fe resupply, Front. Plant Sci., 6 (2015) 145.

[47] D.G. Mendoza-Cozatl, E. Butko, F. Springer, J.W. Torpey, E.A. Komives, J. Kehr, J.I. Schroeder, Identification of high levels of phytochelatins, glutathione and cadmium in the

phloem sap of *Brassica napus*. A role for thiol-peptides in the long-distance transport of cadmium and the effect of cadmium on iron translocation, Plant J., 54 (2008) 249-259.

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

[49] S.K. Tanz, I. Castleden, C.M. Hooper, M. Vacher, I. Small, H.A. Miller, SUBA3: a database for integrating experimentation and prediction to define the SUBcellular localization of proteins in *Arabidopsis*, Nucl. Acids. Res., 41 (2012) 1185-1191.

[50] S. Alvarez, J.Q.D. Goodger, E.L. Marsh, S. Chen, V.S. Asirvatham, D.P. Schachtman, Characterization of the maize xylem sap proteome, J. Proteome Res., 5 (2006) 963-972.

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

[52] J. Kehr, A. Buhtz, P. Giavalisco, Analysis of xylem sap proteins from *Brassica napus*, BMC Plant Biol., 5 (2005) 11.

[53] M.A. Djordjevic, M. Oakes, D.X. Li, C.H. Hwang, C.H. Hocart, P.M. Gresshoff, The *Glycine max* xylem sap and apoplast proteome, J. Proteome Res., 6 (2007) 3771-3779.

[54] Z. Zhang, W. Xin, S. Wang, X. Zhang, H. Dai, R. Sun, T. Frazier, B. Zhang, Q. Wang, Xylem sap in cotton contains proteins that contribute to environmental stress response and cell wall development, Funct. Integr. Genomics, 15 (2015) 17-26.

[55] H.B. Krishnan, S.S. Natarajan, J.O. Bennett, R.C. Sicher, Protein and metabolite composition of xylem sap from field-grown soybeans (*Glycine max*), Planta, 233 (2011) 921-931.

[56] G.S. Ross, E.A. Wegrzyn T Fau - MacRae, R.J. MacRae Ea Fau - Redgwell, R.J. Redgwell, Apple beta-galactosidase. Activity against cell wall polysaccharides and characterization of a related cDNA clone, Plant Physiol., 106 (1995) 521-528.

[57] W. Van den Ende, B. De Coninck, A. Van Laere, Plant fructan exohydrolases: a role in signaling and defense?, Trends Plant Sci., 9 (2004) 523-528.

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

[59] Z. Minic, Physiological roles of plant glycoside hydrolases, Planta, 227 (2008) 723-740.

[60] M.T. Numan, N.B. Bhosle, Alpha-L-arabinofuranosidases: the potential applications in biotechnology, J. Ind. Microbiol. Biotechnol., 33 (2006) 247-260.

[61] C.K. Boyce, M.A. Zwieniecki, G.D. Cody, C. Jacobsen, S. Wirick, A.H. Knoll, N.M. Holbrook, Evolution of xylem lignification and hydrogel transport regulation, Proc. Natl. Acad. Sci. USA, 101 (2004) 17555-17558.

[62] W. Boerjan, J. Ralph, M. Baucher, Lignin biosynthesis, Annu. Rev. Plant. Biol., 54 (2003) 519-546.

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

[64] Y. Xia, H. Suzuki, J. Borevitz, J. Blount, Z. Guo, K. Patel, R.A. Dixon, C. Lamb, An extracellular aspartic protease functions in *Arabidopsis* disease resistance signaling, EMBO J., 23 (2004) 980-988.

[65] M. Rep, H.L. Dekker, J.H. Vossen, A.D. de Boer, P.M. Houterman, C.G. de Koster, B.J. Cornelissen, A tomato xylem sap protein represents a new family of small cysteine-rich proteins with structural similarity to lipid transfer proteins, FEBS Lett., 534 (2003) 82-86.

[66] O. Pechanova, C.-Y. Hsu, J. Adams, T. Pechan, L. Vandervelde, J. Drnevich, S. Jawdy, A. Adeli, J. Suttle, A. Lawrence, T. Tschaplinski, A. Seguin, C. Yuceer, Apoplast proteome reveals that extracellular matrix contributes to multistress response in poplar, BMC Genomics, 11 (2010) 674.

[67] C. Schobert, M. Gottschalk, D.R. Kovar, C.J. Staiger, B.C. Yoo, W.J. Lucas, Characterization of *Ricinus communis* phloem profilin, RcPRO1, Plant Mol. Biol., 42 (2000) 719-730.

[68] A. Barnes, J. Bale, C. Constantinidou, P. Ashton, A. Jones, J. Pritchard, Determining protein identity from sieve element sap in *Ricinus communis* L. by quadrupole time of flight (Q-TOF) mass spectrometry, J. Exp. Bot., 55 (2004) 1473-1481.

[69] N.J. Dafoe, A. Zamani, A.K. Ekramoddoullah, D. Lippert, J. Bohlmann, C.P. Constabel, Analysis of the poplar phloem proteome and its response to leaf wounding, J. Proteome Res., 8 (2009) 2341-2350.

[70] G. Gómez, V. Pallás, A long-distance translocatable phloem protein from cucumber forms a ribonucleoprotein complex in vivo with Hop stunt viroid RNA, J. Virol., 78 (2004) 10104-10110.

[71] T. Nomata, Y. Kabeya, N. Sato, Cloning and characterization of glycine-rich RNA-binding protein cDNAs in the moss *Physcomitrella patens*, Plant Cell Physiol., 45 (2004) 48-56.

[72] W.J. Lucas, B.C. Yoo, F. Kragler, RNA as a long-distance information macromolecule in plants, Nat. Rev. Mol. Cell. Biol., 2 (2001) 849-857.

[73] B.C. Yoo, F. Kragler, E. Varkonyi-Gasic, V. Haywood, S. Archer-Evans, Y.M. Lee, T.J. Lough, W.J. Lucas, A systemic small RNA signaling system in plants, Plant Cell, 16 (2004) 1979-2000.

[74] A. Mangeon, R.M. Junqueira, G. Sachetto-Martins, Functional diversity of the plant glycine-rich proteins superfamily, Plant Signal. Behav., 5 (2010) 99-104.

[75] R. Turgeon, K. Oparka, The secret phloem of pumpkins, Proc. Natl. Acad. Sci. USA, 107 (2010) 13201-13202.

[76] S. Floerl, C. Druebert, A. Majcherczyk, P. Karlovsky, U. Kues, A. Polle, Defence reactions in the apoplastic proteome of oilseed rape (*Brassica napus* var. *napus*) attenuate *Verticillium longisporum* growth but not disease symptoms, BMC Plant Biol., 8 (2008) 129.

[77] S. Floerl, A. Majcherczyk, M. Possienke, K. Feussner, H. Tappe, C. Gatz, I. Feussner, U. Kues, A. Polle, *Verticillium longisporum* infection affects the leaf apoplastic proteome, metabolome, and cell wall properties in *Arabidopsis thaliana*, PLoS ONE, 7 (2012) e31435.

[78] C. Goulet, C. Goulet, M.-C. Goulet, D. Michaud, 2-DE proteome maps for the leaf apoplast of *Nicotiana benthamiana*, J. Proteomics, 10 (2010) 2536-2544.

[79] S.G. Kim, Y. Wang, K.H. Lee, Z.-Y. Park, J. Park, J. Wu, S.J. Kwon, Y.-H. Lee, G.K. Agrawal, R. Rakwal, S.T. Kim, K.Y. Kang, In-depth insight into in vivo apoplastic secretome of rice-*Magnaporthe oryzae* interaction, J. Proteomics, 78 (2013) 58-71.

[80] M. Petriccione, A.M. Salzano, I. Di Cecco, A. Scaloni, M. Scortichini, Proteomic analysis of the *Actinidia deliciosa* leaf apoplast during biotrophic colonization by *Pseudomonas syringae* pv. *actinidiae*, J. Proteomics, 101 (2014) 43-62.

[81] M. Shenton, T. Berberich, M. Kamo, T. Yamashita, H. Taira, R. Terauchi, Use of intercellular washing fluid to investigate the secreted proteome of the rice-*Magnaporthe* interaction, J. Plant Res., 125 (2012) 311-316.

[82] A.L. Girard, F. Mounet, M. Lemaire-Chamley, C. Gaillard, K. Elmorjani, J. Vivancos, J.L. Runavot, B. Quemener, J. Petit, V. Germain, C. Rothan, D. Marion, B. Bakan, Tomato GDSL1 is required for cutin deposition in the fruit cuticle, Plant Cell, 24 (2012) 3119-3134.

[83] J. Zhang, H. Ma, J. Feng, L. Zeng, Z. Wang, S. Chen, Grape berry plasma membrane proteome analysis and its differential expression during ripening, J. Exp. Bot., 59 (2008) 2979-2990.

[84] I. Islas-Flores, C. Alcocer-Alvarez, Y.A. Sánchez-Rodríguez, B. Canto-Canché, Recovery of active pathogenesis-related enzymes from the apoplast of *Musa acuminata* infected by *Mycosphaerella fijiensis*, Afr. J. Biotechnol., 14(23) (2015) 1970-1981.

[85] A. Johansson, J. Staal, C. Dixelius, Early responses in the *Arabidopsis-Verticillium longisporum* pathosystem are dependent on NDR1, JA- and ET-associated signals via cytosolic NPR1 and RFO1, Mol. Plant Microbe In., 19 (2006) 958-969.

[86] M.H.A.J. Joosten, P.J.G.M. De Wit, Identification of several pathogenesis-related proteins in tomato leaves inoculated with *Cladosporium fulvum* (syn. *Fulvia fulva*) as 1,3-β-glucanases and chitinases, Plant Physiol., 89 (1989) 945-951.

[87] B.S. Watson, Z. Lei, R.A. Dixon, L.W. Sumner, Proteomics of *Medicago sativa* cell walls, Phytochemistry, 65 (2004) 1709-1720.

[88] Y. Ishiwatari, C. Honda, I. Kawashima, S. Nakamura, H. Hirano, S. Mori, T. Fujiwara, H. Hayashi, M. Chino, Thioredoxin h is one of the major proteins in rice phloem sap, Planta, 195 (1995) 456-463.

[89] L. Meng, J.H. Wong, L.J. Feldman, P.G. Lemaux, B.B. Buchanan, A membrane-associated thioredoxin required for plant growth moves from cell to cell, suggestive of a role in intercellular communication, Proc. Natl. Acad. Sci. USA, 107 (2010) 3900-3905.

[90] L. Zhang, L.-H. Tian, J.-F. Zhao, Y. Song, C.-J. Zhang, Y. Guo, Identification of an apoplastic protein involved in the initial phase of salt stress response in rice root by two-dimensional electrophoresis, Plant Physiol., 149 (2009) 916-928.

[91] L. Zhou, S.A. Bokhari, C.-J. Dong, J.-Y. Liu, Comparative proteomics analysis of the root apoplasts of rice seedlings in response to hydrogen peroxide, PLoS ONE, 6 (2011) e16723.

[92] A. Frohlich, F. Gaupels, H. Sarioglu, C. Holzmeister, M. Spannagl, J. Durner, C. Lindermayr, Looking deep inside: detection of low-abundance proteins in leaf extracts of *Arabidopsis* and phloem exudates of pumpkin, Plant Physiol., 159 (2012) 902-914.

[93] P.G. Righetti, E. Boschetti, The ProteoMiner and the FortyNiners: searching for gold nuggets in the proteomic arena, Mass Spectrom. Rev., 27 (2008) 596-608.

[94] H. Nguyen-Kim, H.S. Clemente, T. Balliau, M. Zivy, C. Dunand, C. Albenne, E. Jamet, *Arabidopsis thaliana* root cell wall proteomics: Increasing the proteome coverage using a combinatorial peptide ligand library and description of unexpected Hyp in peroxidase amino acid sequences, Proteomics, 16 (3) (2016) 491-503.

Table 1. Summary of proteomic bibliography reviewed. Functional classification is depicted as reported in the original publications.

Reference	Plant Species	Technique	Material	Separation	Identified	Functional Classification
Buhtz et al., 2004	*Brassica napus* *B. oleracea* *Curcubita maxima* *Cucumis sativus*	Stem de-top	Xylem	1-DE	14	Redox response, plant defense, proteolysis
Kehr et al., 2005	*Brassica napus*	Stem de-top	Xylem	2-DE	69	Plant defense, cell wall
Alvarez et al., 2006	*Zea mays*	Stem de-top	Xylem	2-DE	154	Cell wall, plant defense, proteolysis
Djordjevic et al., 2007	*Glycine max*	Stem de-top	Xylem	1&2DE, 2D-LC	24	Proteolysis, cell wall, redox stress, metabolism
Aki et al., 2008	*Orzya sativa*	Stem de-top	Xylem	2D-LC	118	Metabolism, cell wall, pathogenesis related, redox stress
		Stylectomy	Phloem	1DE followed by 2D-LC	107	Metabolism, pathogenesis related, redox stress, signal transduction
Ligat et al., 2011	*B. oleracea*	Stem de-top	Xylem	1DE followed by 2D-LC	189	Carbohydrate metabolism, proteolysis, cell wall, oxido-reductases
Krishnan et al., 2011	*G. max*	Stem de-top	Xylem	2-DE	38	
Zhang et al., 2015	*Gossypium hirsutum*	Stem de-top	Xylem	LC	455	Carbohydrate metabolism, cell wall, stress response, signal transduction
Boudart et al., 2005	*Arabidopsis thaliana*	VIC	Leaf apoplast	2-DE, 1-DE	93	Cell wall, defense related, protein interaction, proteinases
Soares et al., 2007	*Medicago truncatula*	VIC and IB	Leaf apoplast	2-DE	81	Defense, redox, transport, cell wall, pathogenesis related
Pechanova et al., 2010	*Populus deltoides*	Pressure chamber	Leaf apoplast	2-DE, 2D-LC	144	Cell wall, stress defense, proteolysis

34

Witzel et al., 2011	*Z. mays*	VIC	Leaf apoplast	2-DE	67	Cell wall, defense, transport
Delaunois et al., 2013	*Vitis vinifera*	VIC	Leaf apoplast	2-DE	89	Defense, cell wall, proteolysis
Ceballos-Laita et al., 2015	*Beta vulgaris*	Leaf centrifugation	Leaf apoplast	2-DE	164	Stress and defense, cell wall, metabolism
Barnes et al., 2004	*Ricinus communis*	Stem puncture	Phloem	2-DE	18	Sugar metabolism, redox regulation, chaperones
Walz et al., 2004	*C. sativus, C. maxima*	Stem puncture	Phloem	1-DE, 2-DE	45	Redox response, proteinase inhibitors, signaling, defense proteins
Giavalisco et al., 2006	*B. napus*	Stem puncture	Phloem	2-DE and 1-DE	140	Redox stress, signaling, structural, RNA binding, metabolism
Dafoe et al., 2009	*Populus trichocarpa x P. deltoides*	Phloem flow from a cut stem into solution	Phloem	2-DE	48	Metabolism, signaling, stress, structural
Lin et al., 2009	*C. maxima*	Stem puncture	Phloem	2D-LC	1121	Embryo development, ubiquitination, proteolysis, RNA binding, metabolism
Rodriguez-Medina et al., 2011	*Lupinus albus*	Stem puncture	Phloem	2-DE	86	Metabolism, protein modification, redox regulation, stress and defense, structural components
Lattanzio et al., 2013	*Lupinus texensis*	Stem puncture	Phloem	2-DE and 1-DE	54	Protein modification, metabolism, redox stress, cell wall

35

Table 2. The consistent xylem sap proteome. The list contains ten non-redundant *Arabidopsis*-referenced proteins present in at least four of the following xylem sap proteomes: maize (*Z. mays*), rice (*O. sativa*), rapeseed (*B. oleracea*), napus (*B. napus)* and cotton (*G. hirsutum*). Accession identifiers are those reported for *Z. mays* (Alvarez et al., 2005), *O. sativa* (Aki et al., 2008), *B. oleracea* (Ligat et al., 2011), *B. napus* (Buhtz et al., 2004; Kehr et al., 2005) and *G. hirsutum* (Zhang et al., 2015). Symbols and annotation correspond to the TAIR10 description from the BLAST-matched *Arabidopsis* proteins. *Localization according to the SUBA3 consensus prediction.

Z. mays	*O. sativa*	*B. oleracea*	*B. napus*	*G. hirsutum*	AGI	Symbol	Extracellular*	Annotation
gi|13398412	Os12g0128700	TC43276		Q76LU3	AT3G10740	ARAF1	Y	arabinoxylan arabinofuranohydrolase isoenzyme
gi|2109457	Os03g0418000	TC15913		A7LAB9	AT3G12500	CHI-B	Y	chitinase
gi|30692538	P0644B06.40	TC25741/TC46151 4	AT3G61490		AT3G61490		Y	polygalacturonase-like protein
gi|50931079	Os05g0104200		AAM94869	Q2HYU7	AT5G06860	PGIP1	Y	putative polygalacturonase inhibitor
gi|7451375	Os01g0860500	TC17043		A9XTK9	AT5G24090	CHIA	Y	acidic endochitinase
gi|7442163	Os12g0629700	TC54095 4	AAN23104		AT4G11650	OSM34	Y	thaumatin-like pathogenesis- related protein
gi|28400794	Os04g0651000	TC35223 7	gi|15239075	Q7XYR7	AT5G05340	PRX52	Y	peroxidase
gi|7433034	Os03g0235000	TC55805 4	AT5G19890	B9T8I2	AT5G19890	PRX59	Y	peroxidase (EC 1.11.1.7)
gi|50940455		TC58934 4	AT4G33420	F4YAW4	AT4G33420	PRX47	N	putative peroxidase 47 precursor
gi|21593457	OsI_030760	ES901600/TC56831/TC462494	AT1G20160		AT1G20160	CRSP	Y	subtilisin-like serine protease

36

Table 3. The consistent phloem sap proteome. The list contains 13 non-redundant *Arabidopsis*-referenced proteins present in at least four of the following phloem sap proteomes: pumpkin (*C. maxima*), rice (*O. sativa*), rape (*B. napus*), *Lupinus texensis* and *Lupinus albus*. Accession identifiers are those reported for *C. maxima* (Lin et al., 2008), *O. sativa* (Aki et al., 2008), *B. napus* (Giavalisco et al., 2006), *L. texensis* (Lattanzio et al., 2013) and *L. albus* (Rodriguez-Medina et al., 2011). Symbols and annotation correspond to the TAIR10 description from the BLAST-matched *Arabidopsis* proteins. *Localization according to the SUBA3 consensus prediction.

C. maxima	*O. sativa*	*B. napus*	*L. texensis*	*L. albus*	Symbol	Extracellular*	Annotation
FG226997_3	gi|115458768	AT3G04120		CAI83772	GAPC-1	N	glyceraldehyde-3-phosphate dehydrogenase subunit 1
FG227596_2	gi|115481260	AT2G36530		CAB75428	ENO2	N	enolase 2
FG227348_3	gi|115447399	AT5G03300		XP_002531678	ADK2	N	adenosine kinase 2
FG227154_3	gi|115461741	AT1G75270		AAL71857	DHAR2	N	dehydroascorbate reductase 2
gil26985219	gi|115456247b	AT5G28540	giI1143427		HSC70-1	N	heat shock cognate 70-kDa protein 1
FG227112_2	gi|115443875	AT2G16600	gi|6014890	O49886	ATCYP19-1	N	cyclophilin 19
FG227050_1	gi|115465581	AT5G20020	giI4336905		RAN2	N	RAS-related GTP-binding nuclear protein 2
FG227803_2	gi|115489714	AT2G21660		BAF34340	ATGRP7	N	cold, circadian rhythm, and RNA binding 2
gil157346142	gi|115461424	AT3G52560		NP_566968	UEV1D-4	N	ubiquitin E2 variant 1D-4
FG227133_5	gi|115473151	AT1G69410	giI13094963		EIF-5A	N	eukaryotic translation initiation factor 5A-1
gil115343291	gi|115468204	AT1G65480		BAJ33494	FT	N	flowering locus T
FG227096_3	gi|115455697	AT3G46000	giI7339501		ADF2	N	actin-depolymerizing factor 2
gil118481204	gi|115454971	AT5G09810	giI9965319	AAD03741	ACT7	N	actin7

37

Table 4. The consistent apoplastic fluid proteome. The list contains 25 non-redundant *Arabidopsis*-referenced proteins present in at least three of the six plant species for which the leaf apoplast proteome has been described. Accession identifiers are those reported for *A. thaliana* (Boudart et al., 2005), *B. vulgaris* (Ceballos-Laita et al., 2015), *M. truncatula* (Soares et al., 2007), *Z. mays* (Witzel et al., 2011), *V. vinifera* (Delaunois et al., 2013) and *P. deltoides* (Pechanova et al., 2010). Protein identifiers (AGI), symbols and description correspond to the TAIR10 annotation from the BLAST-matched *Arabidopsis* proteins. *Localization according to the SUBA3 consensus prediction.

A. thaliana	*B. vulgaris*	*M. truncatula*	*Z. mays*	*V. vinifera*	*P. deltoides*	AGI	Symbol	Extracellular*	description
AT5G08370			gi|242039251 gi|29468168		Q84VQ7	AT5G08370	AGAL2	Y	α-galactosidase 2
AT3G57240		AF435088		gi|225441379 gi|225441375 gi|163914215 gi|225441373		AT3G57240	BG3	Y	β-1,3-glucanase 3
AT2G28470				gi|157337481 gi|157332401 gi|157329180	Q93X57	AT2G28470	BGAL8	Y	β-galactosidase
AT5G13980	A9PG55		gi|219887203 gi|241935426 gi|115485699	gi|157343878	Q8LPJ3	AT5G13980	BXL4	Y	α-mannosidase
AT5G64570	V4SY44		gi|146762261 gi|223975771 gi|241939802	gi|157355824 gi|157350003	Q2MCJ5	AT5G64570	XYL4	Y	β-D-xylosidase 4
AT4G16260	B9GI31	AF239617 O23473 Q6S9W0 Q8GT15 Q9ZP12		gi|170243	Q56AP0	AT4G16260		Y	putative beta-1,3-endoglucanase
			gi|238908875 gi|195623744 gi|241943398 gi|108707861 gi|77548573	gi|225440254	Q7X9G7	AT3G10740	ASD1	Y	α-L-arabinofuranosidase 1

			gi|125535042	gi|225426168	Q84LI7	AT3G61490		Y	polygalacturonase-like
			gi|255640568 gi|2204224 gi|22240	gi|147810287	Q2MK92	AT5G08380	AGAL1	Y	α-galactosidase 1
			gi|195615700 gi|308081951	gi|157360089	Q9FZP1	AT5G34940	GUS3	Y	glucuronidase 3
		AAD34596 CAA64868 Q42428 Q7X9F6 Q9SDY6	gi|226507679 gi|241939007gi|2340043	gi|10880381		AT3G12500	PR-3	Y	basic chitinase B
	P42820 P36910		gi|413937682 gi|116329	gi|15213852 gi|116333 gi|266324 gi|225434076 gi|157353734	Q9M2U5	AT3G54420	EP3	Y	endochitinase EP3
	Q8LST3	P29024	gi|7687414	gi|225462669	Q09Y38	AT5G24090	CHIA	Y	acidic endochitinase
AT5G20630	I3SGS4	CAC34417	gi|226504688	gi|225444754		AT5G20630	GER3	Y	germin-like protein subfamily 3 member 3
AT5G67360	XP_006466502		gi|125544798 gi|226508174 gi|195647626 gi|223947903	gi|1771160 gi|157335622 gi|157345245	Q8RVJ7	AT5G67360	SBT1.7	Y	subtilisin-like protease SBT1.7
	D7SXW6	AY185207		gi|225429295	Q9XIY9	AT2G15220	BSP	Y	basic secretory protein family
	E0CQB3		gi|242048242	gi|157335112	Q8LAE1	AT1G20160	SBT5.2	Y	serine-type endopeptidase
	XP_008235895		gi|226533158	gi|157348245	Q8L7B2	AT4G12910	SCPL20	Y	serine carboxypeptidase-like 20
	Q38745	O04364	gi|226505234	gi|225426801 gi|225426793	Q8GUQ2	AT4G11650	OSM34	Y	osmotin
			gi|257644672	gi|223635592	Q9ZNZ5	AT1G05260	RCI3	Y	peroxidase

H9BQP8	AAC14127		gi|134684	A9PJW9	AT2G28190	SOD2	N	superoxide dismutase [Cu-Zn] 2
		gi|125597507	gi|157355447 gi|225459180	Q58GF4	AT5G05340	PRX52	Y	peroxidase 52
XP_007014796		gi|226493671	gi|225435616	Q42517	AT5G19890	PRX59	Y	peroxidase 59
	CAA26709	gi|130280		P00299	AT1G20340	PETE2	N	plastocyanin major isoform
B0L802	AW559699	gi|157482855			AT1G06680	OEE2	N	oxygen-evolving enhancer protein 2-1

Table 5. List of the 14 proteins present in the proteomes of xylem sap, phloem sap and apoplastic fluid. Symbols and description corresponds to TAIR10 annotation. *Localization according to the SUBA3 consensus prediction.

AGI	Symbol	Extracellular*	Annotation
AT2G36530	ENO2	N	enolase 2
AT3G04120	GAPC1	N	cytosolic GADPH-1
AT5G43330	MDH2	N	malate dehydrogenase
AT5G17920	MS1	N	methionine synthase
AT4G31990	ASP5	N	aspartate aminotransferase
AT4G09320	NDK1	N	nucleoside diphosphate kinase 1
AT5G13980		Y	glycosyl hydrolase family 38 protein
AT4G16260		Y	putative β-1,3-endoglucanase
AT5G24090	CHIA	Y	chitinase A (class III)
AT1G07890	APX1	N	L-ascorbate peroxidase 1
AT1G11840	GLX1	N	lactoylglutathione lyase
AT5G38410	RBCS3B	N	Rubisco small subunit 3B
AT1G67090	RBCS1A	N	Rubisco small chain 1A
AT5G06730	PRX54	Y	Peroxidase 54

Figure Legends

Figure 1- Functional classification of the non-redundant xylem sap proteome. The 354 *Arabidopsis* non-redundant proteins obtained from the BLAST of species-specific protein sets against that of *A. thaliana* were classified based on their gene ontology (GO) biological process and TAIR annotations.

Figure 2- Functional classification of the non-redundant phloem sap proteome. The 720 Arabidopsis non-redundant proteins obtained from the BLAST of species-specific protein sets against that of *A. thaliana* were classified based on their gene ontology (GO) biological process and TAIR annotations.

Figure 3- Functional classification of the non-redundant leaf apoplast proteome. The 373 Arabidopsis non-redundant proteins obtained from the BLAST of species-specific protein sets against that of *A. thaliana* were classified based on their gene ontology (GO) biological process and TAIR annotations.

Figure 4- Overview of the similarities among xylem sap, phloem sap and leaf apoplastic fluid proteomes. The Venn diagram shows the overlap of the non-redundant plant fluid proteomes provided in Supplementary Tables S1-3. Functional analysis is presented for each of the pair-wise overlaps. The 14 overlapping proteins among the three fluids are listed in Table 5.
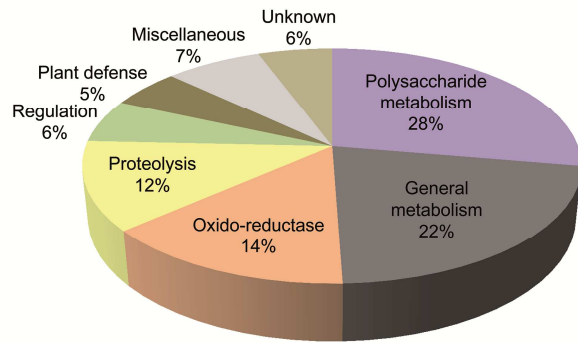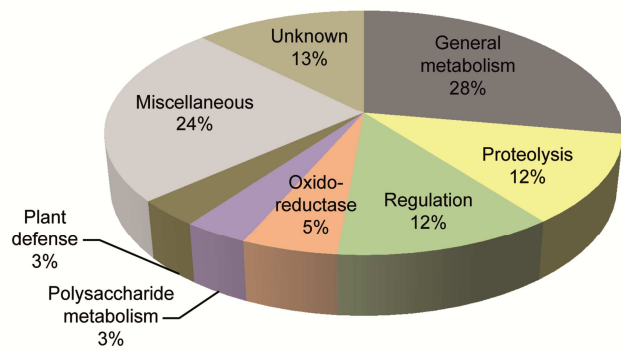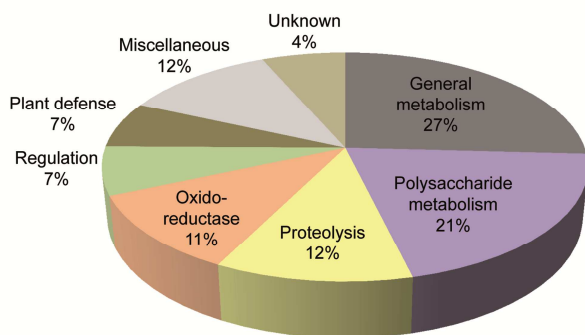
43

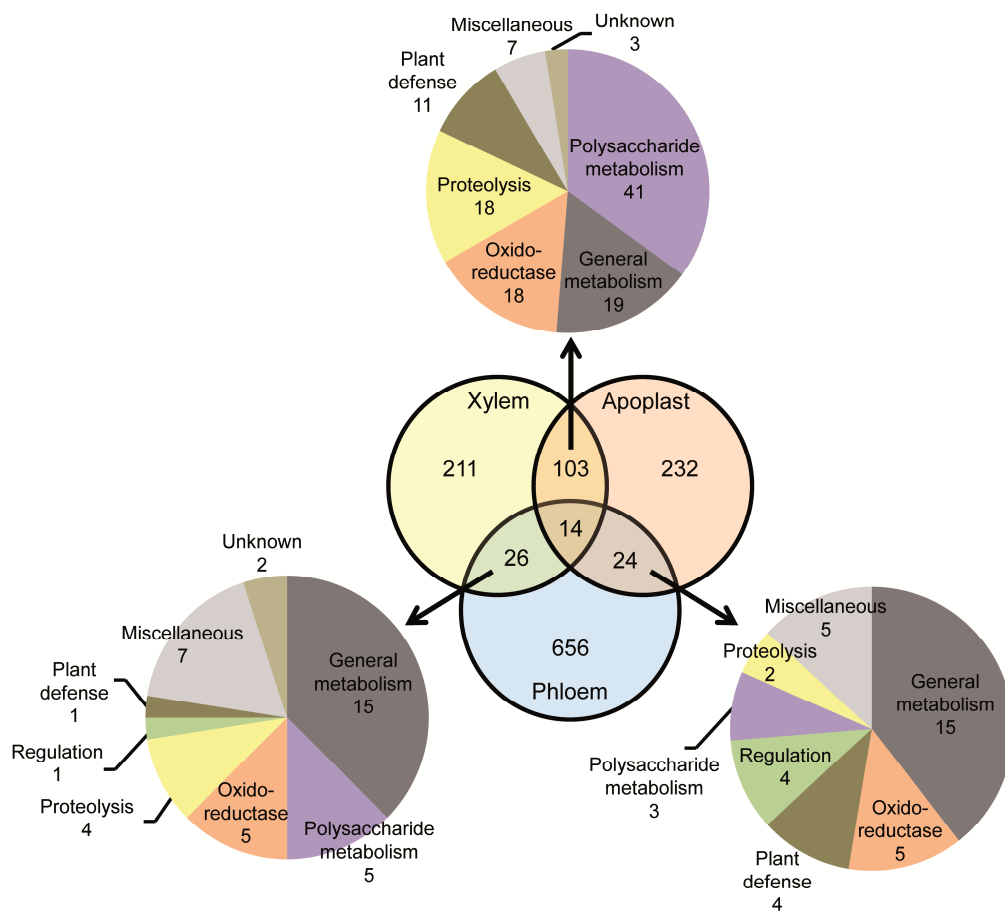Figure 1

Figure 2

45

Figure 3

Figure 4

Highlights

- Review of the phloem sap, xylem sap and apoplastic fluid proteomes
- An inter-species comparisons of plant fluid proteomes has been drawn
- An inter-fluid proteome comparison has been carried out
- Structure-maintenance and defense are predominant functionalities in all plant fluids
- Xylem sap and apoplastic fluid proteomes are more similar between them than with phloem sap