Analysis of the Pumpkin Phloem Proteome Provides Insights into Angiosperm Sieve Tube Function*

Ming-Kuem Lin‡§¶, Young-Jin Lee¶§**, Tony J. Lough‡‡§§, Brett S. Phinney¶, and William J. Lucas‡¶¶

Increasing evidence suggests that proteins present in the angiosperm sieve tube system play an important role in the long distance signaling system of plants. To identify the nature of these putatively non-cell-autonomous proteins, we adopted a large scale proteomics approach to analyze pumpkin phloem exudates. Phloem proteins were fractionated by fast protein liquid chromatography using both anion and cation exchange columns and then either in-solution or in-gel digested following further separation by SDS-PAGE. A total of 345 LC-MS/MS data sets were analyzed using a combination of Mascot and X!Tandem against the NCBI non-redundant green plant database and an extensive Cucurbit maxima expressed sequence tag database. In this analysis, 1,209 different consensi were obtained of which 1,121 could be annotated from GenBank™ and BLAST search analyses against three plant species, Arabidopsis thaliana, rice (Oryza sativa), and poplar (Populus trichocarpa). Gene ontology (GO) enrichment analyses identified sets of phloem proteins that function in RNA binding, mRNA translation, ubiquitin-mediated proteolysis, and macromolecular and vesicle trafficking. Our findings indicate that protein synthesis and turnover, processes that were thought to be absent in enucleate sieve elements, likely occur within the angiosperm phloem translocation stream. In addition, our GO analysis identified a set of phloem proteins that are associated with the GO term “embryonic development ending in seed dormancy”; this finding raises the intriguing question as to whether the phloem may exert some level of control over seed development. The universal significance of the phloem proteome was highlighted by conservation of the phloem proteome in species as diverse as monocots (rice), eudicots (Arabidopsis and pumpkin), and trees (poplar). These results are discussed from the perspective of the role played by the phloem proteome as an integral component of the whole plant communication system. Molecular & Cellular Proteomics 8:343–356, 2009.

In the plant kingdom, the vascular system is composed of xylem, which serves to deliver water and mineral nutrients from the soil to the aerial regions of the plant, and phloem, which delivers sugars and amino acids to developing organs. In the phloem system of flowering plants, the delivery of nutrients occurs through a highly specialized sieve tube system composed of files of nucleate companion cells that support the operation of their enucleate sieve elements (1). Recent studies have shown that the phloem translocation stream also serves as a conduit for the long distance delivery of proteins as well as mRNA, small RNA, and viral nucleic acids (2–16). These studies suggested that such phloem-mobile macromolecules may well participate in and coordinate developmental and physiological events at the whole plant level (17–20).

Using grafting techniques, the translocation of specific RNA molecules and proteins from the stock into the recipient scion tissues was shown to correlate with changes in developmental and physiological phenotypes (8–12). A wide variety of proteins are also transported into and selectively delivered by the phloem translocation stream (4, 7, 20–31). For example, FLOWERING LOCUS T (FT),1 identified as a component of the florigenic signal in the angiosperms, is translocated through the phloem to the shoot apex where it induces flowering (11–15). Other studies have established that certain proteins involved in stress, defense, and an antioxidant defense system are also present within the phloem translocation stream (16, 29, 32). However, current knowledge on the process(es) and regulation of macromolecular trafficking within the phloem as well as the capacity of these signaling molecules to act within the context of a whole plant signaling network remains limited.

A proteomics approach has been used to identify phloem proteins from several plant species. To date, proteomics anal-

1 The abbreviations used are: FT, FLOWERING LOCUS T; Cm, C. maxima; Cmo, C. moschata; GO, gene ontology; NCBI, National Center for Biotechnology; PTB, poly pyrimidine tract-binding protein; RBP, RNA-binding protein; TAIR, The Arabidopsis Information Resource; FDR, false discovery rate; EST, expressed sequence tag; BLAST, Basic Local Alignment Search Tool; AX, anion column exchange; CX, cation column exchange; FTL2, FT-LIKE2; Rubisco, ribulose-1,5-bisphosphate carboxylase; TPR, TOPOLESS-RELATED; E3, ubiquitin-protein ligase; TF, transcription factor; E-value, Expect value.
yses of phloem exudate collected from pumpkin (Cucurbita maxima) (26), castor bean (Ricinus communis) (27) rape (Brassica napus) (28), cucumber (Cucumis sativus) (29), and rice (Oryza sativa) (30) have led to the identification of 17, 18, 103, 16, and 107 proteins, respectively. Despite this progress, the number of proteins identified remains insufficient to allow a comprehensive analysis of the underlying cellular processes that are taking place within this unique enucleate sieve tube system.

The cucurbits represent a model plant system on which to examine the contents of the phloem translocation stream as analytical quantities of phloem sap can be collected from an incision made into their vascular system. This feature makes the cucurbits a powerful experimental system for biochemical analysis of macromolecules that are translocated by the phloem. In the present study, we carried out a large scale proteomics analysis of phloem sap collected from pumpkin. A total of 1,209 non-redundant proteins were identified from 345 LC-MS/MS data sets searched against a combination of comprehensive analysis of the underlying cellular processes that are taking place within this unique enucleate sieve tube system.

**EXPERIMENTAL PROCEDURES**

**Phloem Sap Collection and Protein Separation—**C. maxima cv. Big Max (pumpkin) plants were grown in an insect-free greenhouse, and 8-week-old plants were used to collect phloem sap as described previously (3, 7, 12). Aliquots of phloem sap (30 ml total including equal volumes of phloem sap and collection buffer (100 mM Tris, pH 7.5, 10 mM EDTA, 5 mM EGTA, 10% (v/v) glycerol, 1% (v/v) 2-mercaptoethanol, and protease inhibitors (Complete, Roche Applied Science)) were first dialyzed against buffer A (50 mM Tris, pH 7.5, 1 mM EDTA, and 30 mM 2-mercaptoethanol) for anion column exchange (AX) chromatography or buffer B (20 mM Na2HPO4/NaH2PO4, pH 7.0, 1 mM EDTA, 10% glycerol, and 1% 2-mercaptoethanol) for cation column exchange (CX) chromatography. Phloem proteins were then separated by AX and CX chromatography using HiTrap Q and HiTrap SP columns, respectively, connected to a fast protein liquid chromatography system (Amersham Biosciences) (7, 12). Fast protein liquid chromatography-fractionated phloem proteins were then separated by 12% SDS-PAGE. Gels were then systematically cut into sections (2–5 mm in width), and the proteins were prepared for MS analysis using standard reduction, alkylation, and tryptic digest procedures (33). For some low abundance cation exchange and anion exchange fractions in-solution digestion was also performed (Fig. 1).

Additional data sets, acquired from studies on pumpkin and Cucurbita moschata (squash) phloem sap were incorporated into a total phloem sap data set. These included pyrimidine trap-binding proteins (PTBs), RNA-binding proteins, and CmPP16-, eIF-5A-, and FT-LIKE2 (FTL2)-interacting proteins. Samples of phloem sap were collected as above, and then interacting proteins were isolated by co-immunoprecipitation using the ProFound Co-Immunoprecipitation kit (Pierce) for CmPP16, eIF-5A, FTL2, and PTBs following the manufacturer’s protocol.

**LC-MS/MS—**Three different LC-MS/MS systems were used for the current phloem proteomics study. Initial experiments were performed using a Thermo-Finnigan LCQ Deca XP Plus (Thermo, San Jose, CA) ion trap mass spectrometer coupled on line with a Michrom Paradigm HPLC system (Michrom BioResources, Auburn, CA). Data were collected for PAGE gel bands from three representative cation and anion exchange fractions (Fig. 1): CX E3, E5, and E7 and AX E5, E9, and E13–15 gel bands cut into 20–30 sections. Enrichment of RNA-binding proteins and CmPP16-, eIF-5A- and FTL2-interacting proteins as well as PTBs was also performed with the LCQ Deca XP Plus LC-MS/MS system. A second set of experiments was performed using a Thermo-Finnigan LTQ (Thermo) ion trap mass spectrometer coupled on line with an Eksigent Nano-LC-2D system (Eksigent, Dublin, CA). Overall coverage was obtained by analyzing the in-gel digestion of the AX fractions E3–4, E5–6, and E7–8 and AX fractions E5–6, E7–8, E9–10, E11–12 and the in-solution digestion of the remaining low complexity fractions. An LTQ-FT (Thermo) high resolution hybrid mass spectrometer coupled with a nanoACQUITY UPLC system (Waters, Milford, MA) was also used to acquire data for the PTB- and FTL2-interacting protein experiments.

Details of our LC-MS/MS setup were reported previously for the LCQ Deca XP Plus, LTQ, and LTQ-FT LC-MS/MS systems (12, 34, 35). In brief, samples were loaded onto the C18 trap column, then eluted, and separated by either a homemade or commercial reverse phase column using a 0.3–1 μl/min flow rate. An MS survey scan for the m/z range of either 400–1,400 (for the LCQ Deca XP Plus or LTQ) or 300–1,400 (for the LTQ-FT) and MS/MS spectra for the three (Deca XP Plus), five (LTQ-FT), or 10 (LTQ) most intense ions from the survey scan were acquired during the gradient. An isolation mass window of 3.0 Da (Deca XP Plus), 2.0 Da (LTQ), and 2.0 Da (LTQ-FT) was used for the precursor ion selection, and a normalized collision energy of 35% was used for the fragmentation. Dynamic exclusion for a duration of 3 min was used to acquire MS/MS spectra from low intensity ions. In LTQ-FT experiments, the preview mode for FTMS master scan was turned on, and singly charged ions were rejected for MS/MS.

**Data Search and Analyses—**A total of 345 LC-MS/MS data sets (234 Deca XP Plus, 66 LTQ, and 45 LTQ-FT data sets) were used in this study. Tandem mass spectra were extracted by Mascot distiller (Matrix Science, London, UK; version 2.1.1.0) with default parameters of either LTQ no zoom.otp or LTQ_FT.otp. AgriGenesis BioSciences Ltd., Auckland, New Zealand, performed the initial database search with the 234 Deca XP Plus and nine LTQ LC-MS/MS data sets. This was carried out with Mascot (version 1.9.0) against their proprietary cucurbit EST database (conceptual translation in all six potential reading frames of 154,170 translated sequences; 17,941,293 residues). Search parameters were as follows: trypsin digestion, one missed cleavage allowance, carbamidomethylation on Cys, variable oxidation on Met, precursor tolerance of 1 Da, and fragment tolerance of 0.5 Da.

A second database search was performed at the University of California-Davis Proteomics Facility with Mascot (version 2.2.0) and XITandem (version 2008.01.01.1) against primary pumpkin phloem ESTs plus green plant sequences (total of 493,804 entries): 2,876 ESTs resulting from the above Mascot analysis were translated into six different reading frames and combined with the NCBI non-redundant green plant database (476,548 sequences made from all protein sequences in the Viridiplantae (Taxonomy identifier 33090) taxonomic group downloaded from NCBI on April 3, 2008). The search parameters were as follows: trypsin digestion, one missed cleavage allowance, carbamidomethylation on Cys, variable oxidation on Met, precursor tolerance of 2.5 Da, and fragment tolerance of 0.6 Da. An XITandem search was also performed in parallel with a parent ion tolerance of 2 Da and fragment ion mass tolerance of 0.40 Da. For the data set obtained with the LTQ-FT, a parent tolerance of 10 ppm was used for both searches.

Scaffold (version Scaffold-2_00_00_final, Proteome Software Inc., Portland, OR) was used to validate MS/MS-based peptide and protein identifications. Peptide identifications were accepted if they could be established at greater than 95% probability as specified by.
the Peptide Prophet algorithm (36). Protein identifications were accepted if they could be established at greater than 99.0% probability and contained at least two unique identified peptides. Protein probabilities were assigned by the Protein Prophet algorithm (37). Proteins that contained similar peptides and could not be differentiatated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony (37). The identified cucurbit ESTs were submitted to the NCBI dbEST database, and corresponding accession numbers are herein provided. Suffixes added to NCBI dbEST accession numbers, _"1" for example, indicate their translation frames.

**BLAST Search and Gene Ontology (GO) Classification** — The BLAST program (version 2.2.18) downloaded from NCBI was used for search analyses and annotation of all phloem proteins. Two sets of BLAST searches were performed. First, pumpkin phloem protein sequences were mapped against three representative plant species. Subject sequences included those from *Arabidopsis thaliana* downloaded from TAIR (version TAIR7; release date, April 25, 2007), *O. sativa* downloaded from The Institute for Genomic Research (release 5.0; release date, January 24, 2007), and draft sequences of *Populus trichocarpa* downloaded from the United States Department of Energy Joint Genome Institute (version 1.1; release date, December 2004). Second, phloem proteins reported for rape (28), rice (30), and castor bean (27) were BLAST searched against *Arabidopsis* sequences to derive corresponding *Arabidopsis* non-redundant phloem proteins. Unless otherwise stated, BLASTP analyses were conducted using subject sequences described above. GO analysis was performed for BLAST-matched *Arabidopsis* accession entries of pumpkin phloem proteins based on their TAIR GO categories (ATH_GO_GOSLIM.20080621.txt). Fisher’s exact test was used for GO enrichment analysis using GOSSIP (version 1.4.1) (38) through BLAST2GO (version 2.2.2) (39) with all TAIR7 entries used as a reference.

Raw mass spectrometric data and the Scaffold result file are stored at Tranche, a public repository for sharing scientific data. Instructions to download the files are available upon request. A viewer version of the Scaffold program can be downloaded (free) from Proteome Software Inc.

**RESULTS AND DISCUSSION**

**Phloem Sap from the Pumpkin Enucleate Sieve Tube System Contains Over 1,000 Proteins**—Phloem sap proteins from pumpkin were identified by LC-MS/MS, following two-dimen- 

sional separation, using either cation or anion exchange chromatography in combination with SDS-PAGE. Polyacrylamide gel fragments were then processed for MS analysis. Additional phloem sap-derived proteins obtained from independent data sets (PTB-binding, RNA-binding, and CmPP16-, CmOFL2- and eIF-5A-interacting proteins) were also included in the total data set described herein. Phloem sap proteins in anion and cation exchange eluants were analyzed by LC-MS/MS. A combination of a proprietary cucurbit EST database and the NCBI non-redundant green plant database was used to analyze these phloem proteins by Mascot and X!Tandem search engines, and the results were compiled with Scaffold.

A total of 345 LC-MS/MS data sets were processed in which 84,931 MS/MS spectra were assigned to 8,859 unique peptides with a 95% probability. The identified peptides were then assigned to 1,209 consensi using rigid filtering criteria to a required threshold of 99% protein probability that included assignment of at least two unique peptides. Through application of these criteria, our overall false positive is expected to be 1 or less. Supplemental Table I provides the full list of the 1,209 consensi identified in this study as well as the BLAST search results against *Arabidopsis*, rice, and poplar database subject sequences.

We next examined the phloem protein database for possible contamination during sample collection. Ribulose-1,5-bisphosphate carboxylase (Rubisco) is the most highly abundant protein in green plant tissues, and because it is present in cells that surround the sieve elements, it was used as a test for contaminating proteins. In an earlier study on proteins collected from rape phloem some seven unique peptide sequences of Rubisco were detected (28). Consistent with these findings, we also detected unique Rubisco peptides (index 817 in supplemental Table I; RbcL). However, our data set was far more extensive, and yet we could detect only three unique peptides, indicating that our sampling of pumpkin phloem sap was likely essentially free from protein contamination from neighboring cell types.

The cucurbit EST and NCBI protein sequences used to analyze our LC-MS/MS data sets included a large number of redundant entries that could have resulted in redundant protein assignments. However, redundancy in the 1,209 consensi should be minimal as proteins without uniquely identifying peptides were placed into a single consensus. Of the 1,209 consensi identified in our study, 349 were found to possess at least two accession numbers (supplemental Table I). For example, index number 352 was assigned to 211 accession numbers of ubiquitin or ubiquitin-related proteins. This is because ubiquitin has a high degree of sequence homology throughout the plant kingdom. Indeed meaningful annotations could be obtained for 1,121 (or 93%) of these 1,209 identified consensi by using either NCBI annotations or BLAST-matched *Arabidopsis* or rice annotations. Furthermore it is interesting that many of the entries for which clear NCBI annotations were not found (i.e. unnamed or hypothetical proteins) had identified accession numbers from *Vitis vinifera* (206 of a total of 260), suggesting that there must be good sequence homologies in phloem proteins between grape and pumpkin, two plants that are vines.

The most abundant phloem proteins are presented in Table I; 45 proteins were identified based on the presence of minimum unique peptides and sequence coverage of either six and 50% or 10 and 40%, respectively. As expected, CmPP1 and CmPP16-2 were among the most abundant proteins in terms of the number of unique peptides and overall sequence coverage, respectively. These two proteins appear as *dark blue* bands in Fig. 1 (CmPP16-2, E3–E5 of the anion exchange fractions; CmPP1, E3–E5 of the cation exchange fractions). Here we should mention that to minimize detection of the highly abundant CmPP1, we excluded the E3–E5 gel bands in the first data set. For the second data set, we added the major peptide peaks for CmPP1 into an exclusion
TABLE I

Major phloem proteins were defined as those having either a minimum number of unique peptides and sequence coverage of 10 and 40% or six and 50%.

| Index | Accession nos. | Molecular mass | No. unique peptides | Coverage | Arabidopsis accession no. | BLAST E-value\(^a\) | Annotation\(^b\) |
|-------|----------------|----------------|---------------------|----------|---------------------------|----------------------|-----------------|
| 1     | gi:25090875    | 15.6           | 19                  | 93       | AT3G55470.2               | 2.00e−33             | 16-kDa phloem protein 2 (C. maxima) |
| 2     | gi:9937311     | 42.8           | 33                  | 76       | AT1G47710.1               | 1.00e−118            | Phloem serpin-1 (C. maxima)       |
| 3     | FG227486_3, FG227487_3, FG227491_3, FG227794_3 | 24.8 | 15                  | 75       |                           |                      | Dimeric phloem-specific lectin PP2 (C. maxima) |
| 4     | gi:18069       | 26.3           | 22                  | 74       |                           |                      |                               |
| 5     | gi:1402347     | 16.8           | 7                   | 69       | AT2G27030.3               | 7.00e−81             | Calmodulin-2 (Capsicum annuum)     |
| 6     | gi:1753099     | 95.4           | 95                  | 68       |                           |                      | Phloem filament protein; PP1 (C. maxima) |
| 7     | gi:68349004    | 16.5           | 11                  | 64       | AT3G55470.2               | 1.00e−35             | 16-kDa phloem-protein 1 (C. maxima) |
| 8     | gi:111481204, gi:14774746 | 41.7 | 22                  | 63       | AT5G09810.1               | 0                    | Actin (P. trichocarpa)            |
| 9     | FG227053_2     | 45.2           | 25                  | 62       | AT5G36230.1               | 1.00e−175            | eIF4-γ/eIF5/eIF2-ε domain-containing protein |
| 10    | FG227644_1     | 31.6           | 22                  | 58       | AT5G09550.1               | 1.00e−135            | RAB GDP-dissociation inhibitor    |
| 11    | FG227107_1     | 38.8           | 16                  | 58       | AT3G14290.1               | 1.00e−132            | PAE2 (20 S proteasome α subunit E2); peptidase binding |
| 12    | FG226976_3     | 55.9           | 31                  | 57       | AT1G30580.1               | 0                    | GTP binding                        |
| 13    | FG227623_2     | 23.7           | 10                  | 56       | AT3G60820.2               | 1.00e−103            | PBF1 (20 S proteasome β subunit F1); peptidase |
| 14    | gi:14582160    | 36.3           | 24                  | 55       | AT5G20080.1               | 1.00e−108            | Cytochrome b, reductase PP36 (C. maxima) |
| 15    | FG227373_4, FG227754_4 | 25.7 | 15                  | 55       | AT1G79690.1               | 2.00e−76             | ATNUDT3 (A. thaliana Nudix hydrolase homolog 3); hydrolase |
| 16    | FG227359_3     | 26.1           | 11                  | 55       |                           |                      |                               |
| 17    | FG227489_1     | 22.7           | 9                   | 53       | AT5G42190.1               | 2.00e−67             | ASK2 (Arabidopsis SKP1-LIKE 2); ubiquitin-protein ligase |
| 18    | FG227533_3     | 20.4           | 6                   | 53       |                           |                      |                               |
| 19    | FG227347_3     | 37.6           | 15                  | 52       | AT2G18110.1               | 6.00e−78             | Elongation factor 1-β, putative   |
| 20    | FG227054_1, FG227705_3 | 51.6 | 30                  | 51       |                           |                      |                               |
| 21    | FG227366_6, FG227751_6 | 24.2 | 12                  | 51       |                           |                      |                               |
| 22    | FG227297_3     | 23.6           | 9                   | 51       | AT1G56700.2               | 6.00e−80             | Pyrrolidine-carboxylate peptidase family protein |
| 23    | FG227133_5, FG227599_5 | 29.5 | 16                  | 49       | AT1G13950.1               | 3.00e−82             | EIF-5A (eukaryotic translation initiation factor 5A) |
| 24    | FG227261_1     | 30.2           | 13                  | 49       | AT3G24170.2               | 1.00e−116            | ATGR1; glutathione-disulfide reductase |
| 25    | FG227282_2, FG227747_2 | 28.2 | 10                  | 49       | AT5G06730.1               | 3.00e−64             | Peroxidase, putative            |
| 26    | FG227348_3     | 25.3           | 10                  | 49       | AT5G03300.1               | 2.00e−97             | ADK2 (ADENOSINE KINASE 2); kinase |
| 27    | FG227407_6     | 28.1           | 14                  | 47       | AT3G53110.1               | 1.00e−108            | LOS4 (low expression of osmotically responsive genes 1); ATP-dependent helicase |
| 28    | FG226986_3     | 21.8           | 11                  | 47       | AT3G12050.1               | 2.00e−66             | Aha1 domain-containing protein    |
| 29    | FG226986_3     | 21.8           | 11                  | 47       | AT3G12050.1               | 2.00e−66             | Aha1 domain-containing protein    |
| 30    | gi:3914449     | 47.5           | 28                  | 45       | AT1G53750.1               | 0                    | 26 S protease regulatory subunit 7 (Prunus persica) |
Other well characterized phloem proteins, such as CmPP16-1, the serpin, CmPS-1, and the dimeric phloem-specific lectin CmPP2, were also detected as major proteins. Some major proteins, such as calmodulin, eukaryotic initiation factors, 20 S proteasome subunits, etc., are noteworthy and will be discussed later. Among the 45 proteins listed in Table I, six consensi were detected exclusively from ESTs having no significant BLAST matches (H11021/H11023) or annotation and one consensus with BLAST match to an unknown Arabidopsis entry. An understanding of the functions performed by these proteins must await further detailed studies.

Universal Role of Phloem Proteins Implicated by Comparison of Phloem Proteomes between Species—Unfortunately only a limited number of phloem proteins have been identified from plant species, including the cucurbits (3–5, 7, 12, 16, 21, 22, 25, 26), rice (30), rape (28), and castor bean (27). Purification techniques of analytical quantities of phloem sap have utilized two-dimensional or one-dimensional gel separations, resulting in the identification of some 100 proteins. In any event, we compared our pumpkin phloem proteome with that derived from these three species to ascertain whether similar proteins are present and, by extrapolation, which biochemical pathways and processes may be held in common across multiple species.

This comparison of the phloem proteome of multiple species was complicated because of issues related to and difficulties in handling homology between species where comparative genomic data sets are poorly developed or absent. As a solution to this situation, we adopted an approach whereby phloem proteins from each species were compared with A. thaliana as a suitable reference species through use of BLAST analyses. For each phloem protein data set, a corresponding "Arabidopsis non-redundant phloem proteome" was deduced using the following procedure. All Arabidopsis accession numbers with BLAST E-values of less than \(1 \times 10^{-30}\) were saved for each protein. If any of these Arabidopsis accession numbers overlapped between two or more entries, they were not included in the "Arabidopsis non-redundant phloem proteome".

| Index | Accession nos. | Molecular mass | No. unique peptides | Coverage | Arabidopsis accession no. | BLAST E-value | Annotation |
|-------|----------------|----------------|---------------------|----------|--------------------------|---------------|------------|
| 31    | gi147844475    | 47.5           | 18                  | 45       | AT3G05530.1              | 0             | RPT5A (regulatory particle triple-A 5A); ATPase/calmodulin binding |
| 32    | FG227127_4, FG227729_4 | 38.4           | 13                  | 44       | AT1G76960.1              | 1.00E−12      | ATNUDT3 (A. thaliana Nudix hydrolase homolog 3); hydrolase |
| 33    | FG227327_2, FG227740_2 | 24.1           | 12                  | 44       | AT5G11500.1              | 4.00E−10      | Unknown |
| 34    | FG227037_1     | 55.2           | 26                  | 43       | AT3G02360.1              | 0             | 6-Phosphogluconate dehydrogenase family protein |
| 35    | FG227097_3     | 25.0           | 11                  | 43       | AT3G18060.1              | 1.00E−10      | Transducin family protein/WD-40 repeat family protein |
| 36    | FG227681_4     | 37.8           | 11                  | 43       | AT5G61780.1              | 1.00E−12      | Tudor domain-containing protein/nuclease family protein |
| 37    | FG227175_3, FG227620_3 | 42.7           | 17                  | 42       | AT4G20904.1              | 0             | RPT2A (regulatory particle triple-A 2A); ATPase |
| 38    | FG227339_3     | 28.3           | 13                  | 42       | AT1G70840.1              | 2.00E−13      | Bet v I allergen family protein |
| 39    | FG227382_1, FG227743_1 | 26.5           | 11                  | 42       | AT2G34690.1              | 1.00E−10      | ACD11 (ACCELERATED CELL DEATH 11) |
| 40    | gi157347550    | 89.6           | 42                  | 41       | AT3G53520.1              | 0             | Cell division cycle protein 48, putative/CDC48, putative |
| 41    | FG227803_2     | 28.6           | 13                  | 41       | AT2G21660.2              | 3.00E−35      | ATGRP7 (COLD, CIRCADIAN RHYTHM, AND RNA BINDING 2); RNA binding |
| 42    | FG227098_1, FG227719_1 | 31.0           | 11                  | 41       | AT4G31750.1              | 1.00E−14      | Protein phosphatase 2C, putative/PP2C, putative |
| 43    | FG227058_3     | 31.9           | 10                  | 41       | AT3G08590.2              | 1.00E−118     | 2,3-Bisphosphoglycerate-independent phosphoglycerate mutase, putative/phosphoglyceromutase, putative |
| 44    | FG227611_3     | 52.3           | 16                  | 40       | AT4G13930.1              | 0             | SHM4 (SERINE HYDROXYMETHYLTRANSFERASE 4); glycine hydroxymethyltransferase |
| 45    | FG227361_2     | 24.6           | 12                  | 40       | AT5G14680.1              | 1.00E−79      | Universal stress protein (USP) family protein |

\(^a\) Arabidopsis BLAST score higher than \(1 \times 10^{-30}\) are not shown.

\(^b\) Annotations are from GenBank\textsuperscript{TM} or from BLAST-matched Arabidopsis annotations. Note that when a GenBank annotation is used the species name is attached.
considered to be redundant, and only a single representative entry was retained as non-redundant.

Based on this approach, 636, 63, 69, and 16 Arabidopsis (reference) non-redundant phloem proteins were identified for pumpkin (this study), rice (30), rape (28), and castor bean (27) phloem proteins as query sequences, respectively. Supplemental Table II lists these non-redundant phloem proteins. Fig. 2A provides a Venn diagram comparing the non-redundant Arabidopsis reference proteins for these species. Castor bean had only 16 non-redundant entries, and the extent of its overlap with the other three species is shown in parentheses in each category. The most interesting proteins were those that were consistently detected in the different species. Table II lists 18 common phloem proteins present in pumpkin, rice, and rape.

Proteins such as ascorbate peroxidase, adenosine kinase 2, glutathione dehydrogenase, glyceraldehyde-3-phosphate dehydrogenase, malate dehydrogenase, and phosphopyruvate hydratase likely reflect the underlying biochemistry that is occurring within sieve elements. Cytoskeletal components like actin7 and ADF6, an actin-binding protein; chaperones like the heat shock cognate 70-1; and the cyclophilin SQUINT as well as a GTP-binding protein, ubiquitin, and an ubiquitin-protein ligase reflect the presence of basic cellular components within the highly reduced enucleate sieve elements. Detection of ATGRP7, an RNA-binding protein, and elf-5A is consistent with earlier reports on the presence of mRNA within the pumpkin phloem translocation stream (3, 4, 9, 28). The presence of TCTP (translationally controlled tumor protein) in the phloem sap is intriguing because in animals this protein plays a role in cell proliferation (40). The presence of FT in the phloem translocation stream of rice, pumpkin, and rape is consistent with its role as a long distance signaling molecule in floral induction (11–15). Interestingly the phloem-specific lectin CmPP2 was not included in Table II because the BLAST E-value was slightly higher than 1e-30. However, CmPP2 (index 44 in supplemental Table I) is abundant in pumpkin phloem sap, and this finding is consistent with previous reports for the cucurbits (26, 29), rape (PP2; At4g19840) (28), and rice (ricin B lectin domain-containing protein; gi|115474137|gi|115452789|gi|115434012) (30). The cucurbit lectin has been implicated in long distance RNA trafficking (5).

Genes Encoding Pumpkin Phloem Proteins Are Conserved in Other Species—To determine the extent to which the pumpkin phloem proteins are conserved across plant spe-

---

**Fig. 1.** Flow diagram of the procedures used to collect and analyze the protein content of the pumpkin (C. maxima) phloem translocation stream. A, 8-week-old pumpkin plants were used as a source for angiosperm phloem sap. B, mature source leaves (petioles) or stems were excised from the plant, and phloem sap that exuded from the cut surface was collected and transferred into collection buffer. C and D, proteins contained in the pumpkin phloem sap were separated by anion and cation exchange chromatography, respectively. E, in-gel and/or in-solution digestion was used to prepare tryptic peptides that were then subjected to LC-MS/MS analysis. The resultant MS/MS spectra were searched against a proprietary cucurbit EST and the NCBI green plant databases.
cies, we next used the phloem proteome to perform a BLAST search against the Arabidopsis, poplar, and rice genomic databases. Arabidopsis and rice were chosen as representatives of the eudicot and monocot species, respectively, and because their genomes are now well annotated. Poplar was chosen as it represents a woody species. For this BLAST analysis, a minimum E-value of $1 \times 10^{-30}$ was used. The data presented in Fig. 2B demonstrate that a high degree of conservation was detected between these three plant genomes and the pumpkin phloem proteome. Of the 1,209 entries in the pumpkin phloem proteome, some 1,087 could be matched with all three species, and 928 were non-redundant. It is interesting to note that only 75 entries in the pumpkin phloem proteome could not be matched with any of these three databases. Some of these proteins, such as PP1 (gi 1753099), which was the most abundant protein in pumpkin phloem sap, may well be unique to cucurbit species.

Entries with BLAST matches to the same protein are potentially redundant especially when all three BLAST matches are consistently the same. However, if entries have BLAST matches to different proteins in any species they are probably non-redundant, although they may be highly homologous. Combining all non-redundant entries in Fig. 2B (values in parentheses in red) yielded a total of 969 proteins that are represented across at least two species. Adding the 75 entries with no BLAST hits into this number gives a total of 1,044 entries of the 1,209 pumpkin phloem proteins that are expected to be non-redundant. Similar results were obtained when we used an E-value of $1 \times 10^{-10}$ as the cutoff. In any event, based on these analyses, we identified about 1,000 non-redundant pumpkin phloem proteins that are likely conserved among vascular plants.

GO Classifications of Identified Phloem Proteins—GO of identified pumpkin phloem proteins was evaluated for BLAST-matched Arabidopsis genes on the basis of the TaIR7 GO categories. GO analysis through BLAST is often performed by sequence homology matching to protein motifs. Accordingly an E-value cutoff of $1 \times 10^{-10}$ was used instead of the above used $1 \times 10^{-30}$; this value was still lower than the commonly used cutoff of $1 \times 10^{-6}$ (39). Fisher’s exact test was performed using GOSSIP (38) to assess the significance of overrepresentation of specific GO terms with the total Arabidopsis genes being used as reference.

Our analysis identified a total of 194 GO terms that were enriched using a cutoff of false discovery rate (FDR) and familywise error rate of 1 and 5%, respectively (supplemental Table III). Enrichment of certain GO terms can be attributed to the contribution from other GO terms that are enriched at the lower hierarchy. Consequently a directed acyclic graph approach was used to provide a clearer understanding of which GO terms were specifically enriched and how these affected other GO terms through upper hierarchies. As an example, Fig. 3 presents a part of a directed acyclic graph in the aspect of “biological process” that is related with the GO term “embryonic development ending in seed dormancy.” Surprisingly this GO term was enriched at a very low FDR ($<1 \times 10^{-8}$), and several other GO terms at upper hierarchies, such as “seed development,” “embryonic development,” and “reproductive structure development” were enriched as a result. Interestingly photomorphogenesis GO categories were also enriched but at a somewhat lower probability (Fig. 3).

The 39 Arabidopsis genes corresponding to the GO term embryonic development ending in seed dormancy are listed in supplemental Table IV. A number of these proteins represent...
enzymes involved in housekeeping functions, such as those that participate in protein turnover (e.g. FUS6, RPN12, SAE2, ASK2, and UBPI4). However, other proteins, like TOPELESS-1 (TPL-1), TOPELESS-RELATED 2 (TPR2), and TPR4 are involved in the regulation of apical embryonic fate (41, 42). Thus, their presence in the enucleate sieve tube system is most certainly intriguing and may implicate the phloem in control over plant embryonic and seed development. The presence of CELL DIVISION CONTROL 2 (CDC2), MULTI-COPY SUPPRESSOR OF IRA1 (MSI1), and the transcription activator EMBRYO DEFECTIVE 1374 (EMB1374) adds further weight to the notion that the phloem translocation stream may influence the genetic programming of developing tissues and organs, such as seeds. The vascular cambium could also be a target for these regulatory proteins.

A very large scale Arabidopsis proteomics study was recently reported in which almost half of the total genes for this species were identified (i.e. 13,029 of a total of 27,029 genes) (43). Interestingly the GO term embryonic development enriched specifically in the seed. As mentioned above, we also detected a high level of enrichment for a set of phloem proteins within this GO category (supplemental Table III), suggesting that it could serve as a biomarker for both the seed and the phloem. Of significance to this conclusion, other GO categories reported as biomarkers by Baerenfaller et al. (43) were not enriched in the phloem proteome.

**Ubiquitin-mediated Proteolysis**—Enriched GO terms were identified associated with ubiquitin-mediated proteolysis (supplemental Figs. 1 and 2); note that some 116 proteins involved in proteasome-mediated protein degradation (supplemental Table VI) were detected in the pumpkin phloem sap. As illustrated in Fig. 4, the pumpkin phloem sap contains all the components of the 26 S proteasome with the exception of Rpn4. In this regard it is important to note that Rpn4 has not been reported previously in plants, and furthermore, a significant BLAST match of the Saccharomyces cerevisiae Rpn4 sequence could not be identified against the entire non-redundant green plant protein database. Thus, it would appear that the enucleate sieve tube system of pumpkin contains all the components to form a functional 26 S proteasome.
It has long been known that the cucurbit phloem sap contains ubiquitin (44, 45), and inspection of Fig. 4 reveals that the pumpkin phloem sap also contains various ubiquitin-activating, ubiquitin-conjugating, and ubiquitin ligase enzymes. With respect to the E3 type ubiquitin ligase systems, it is clear that a number of components remain to be identified from within the phloem sap. Future experiments will use co-immunoprecipitation in an effort to enrich for components of the homologous to the E6-AP carboxyl terminus (HECT), U-box, and multisubunit Really interesting new gene (RING) finger type E3 complexes. Irrespective of the fact that some of these components have yet to be identified, the presence of the detected 116 phloem proteins associated with ubiquitin-mediated proteolysis (supplemental Table VI) offers strong support for the hypothesis that the functional, enucleate sieve tube system has retained the capacity for proteolysis. The important questions posed by these results relate to the process by which the 26 S proteasome is assembled in sieve elements and whether these complexes are anchored in some way to the plasma membrane or to the cytoskeleton to limit their mobility within the phloem translocation stream.

**Machinery for Protein Synthesis Present in Pumpkin Phloem Sap**—One of the most unexpected discoveries from the analysis of the pumpkin phloem proteome was the identification of some 100 proteins involved in protein synthesis (supplemental Table VI). As the sieve elements that comprise the sieve tube system are enucleate in their functional state, it has long been held that protein synthesis does not take place in this translocation system (17, 20, 28). However, as the components involved in tRNA aminoacylation for protein translation, translation initiation, elongation, and termination represented highly enriched GO categories (supplemental Fig. 3), this hypothesis must now be challenged. Although only a small number of ribosomal proteins were identified (Fig. 5), the majority of the remaining machinery required for protein synthesis does appear to be present in the pumpkin phloem sap. The 80 and 60 S ribosomal complexes may well be anchored in some way to the parietal endoplasmic reticulum, thereby limiting their release into the phloem sap that is collected following an incision into the vascular system. This could well account for their absence from the phloem sap.

It has been hypothesized that proteins in the sieve elements are synthesized in the neighboring companion cells before being transported through the specialized plasmodesmata that connect these two cell types (17–20). Recent studies have identified situations in which both protein and the encoding mRNA in polyadenylated form are present in the phloem sap (3, 4); in other cases, only the protein has been detected in the phloem translocation stream (21, 22, 47). Thus, our identification of a large component of the protein...
machinery required for protein synthesis in conjunction with earlier studies in which individual proteins and their mRNA were detected (3, 4) offers support for the hypothesis that some proteins in the phloem translocation stream are actually synthesized in the enucleate sieve tube system.

Pumpkin Phloem Sap Contains a Set of RNA-binding Proteins—An increasing number of RNA-binding proteins (RBPs) have been detected in the phloem translocation stream (4, 5, 7, 20, 28, 30), and their presence is consistent with the notion that RNA transport in the phloem plays a role in the coordination of developmental and physiological events at the whole plant level (3, 4, 6, 8–10). These phloem-mobile RBPs are thought to form ribonucleoprotein complexes with the unique population of mRNA present in the translocation stream (20).

Based on the annotations for the proteins identified in this study, at least 82 non-redundant Arabidopsis (reference) proteins were detected that have RNA binding properties (supplemental Table V). A number of these RBPs have been identified previously, including CmPP16-1/-2 (4, 48), glycine-rich RNA-binding protein (23, 26), eIF-5A (23, 28), and CmPP1/Cc1-F-box protein; ECS, Elongin C-Cul2-SOCS box; ECV, SCF-like E3 ubiquitin ligase complex.
Fig. 5. Machinery for protein synthesis is present in pumpkin phloem sap. The diagram illustrates the 80 S ribosomal system and the key steps involved in mRNA translation. Pumpkin phloem proteins involved in protein synthesis are presented in supplemental Table VI. Graphic and box representation of proteins identified in the pumpkin phloem sap and involved in protein synthesis including aminoacyl-tRNA synthetases, ribosomal proteins, translation initiation, elongation, and termination factors is shown. Listed in green boxes are proteins identified in this work but not shown, whereas purple boxes represent those detected and shown. This figure was drawn based on a scheme for animals (46); CBP, cap-binding protein; PABP, poly(A)-binding protein; EPA, ribosome exit, peptidyl, and amino acyl binding sites.

romolecular and vesicle trafficking (supplemental Fig. 5 and Table III) and membrane dynamics (supplemental Fig. 6 and Table III). Inspection of these data sets indicated the presence of a number of proteins belonging to endoplasmic reticulum-to-Golgi vesicle-mediated transport (GO:0048193) as well as to the vacuolar protein sorting-associated protein (VPS) family. Such proteins were unexpected as it was generally considered that, with the exception of the specialized endoplasmic reticulum (17), during sieve element maturation the endomembrane system and vacuole were degraded. Identification of transport proteins, such as the A, B2, C, D1, E1, F, and H subunits of the tonoplast ATPase, provide additional evidence that mature functional sieve elements may well have retained small vacuoles. Proteins associated with coated membranes and clathrin-coated vesicles (supplemental Fig. 6) indicate that mature sieve elements may also have retained the capacity to recycle components of their plasma membrane. Future studies will be required to test for the presence of functional Golgi, endosomes, and vacuoles as well as to explore the roles they may play in sieve tube function.

Given that, at maturity, the nucleus of individual sieve elements has undergone complete degradation, it is interesting to note that the phloem proteome was found to contain a number of proteins, such as importin-α, importin-β, and NTF2B, involved in nuclear import and export (supplemental Fig. 5). Whether these proteins represent contaminants from either neighboring companion cells or maturing sieve elements or serve some function in long distance signaling will require further investigation.
Pumpkin Sieve Tube System Contains Many Stress-related Proteins—Our GO enrichment analysis identified significant numbers of phloem proteins that are involved in various aspects of abiotic stress (supplemental Table III). For example, some 66 proteins were present in the GO category "response to abiotic stimulus" (GO:0009628), and 78 proteins were associated with temperature stress (GO:0009266, GO:0009408, and GO:0009409). Detection of these stress-related proteins is consistent with earlier reports on the characterization of phloem proteins (28, 32, 49). Interestingly in comparison with earlier reports (29), enrichment of pathogen-related proteins was not detected in our pumpkin phloem proteome. This may reflect the fact that plants were propagated under insect- and pathogen-free conditions. Although we used well watered conditions, pumpkin plants were often slightly wilted during the middle of the day. This may account for the presence of proteins associated with temperature-related stress. The role of these proteins in the functioning of the sieve tube system or as components of the long distance signaling system will be evaluated in future studies.

Sieve Tube System as a Specialized Metabolic Compartment—It is of interest that the pumpkin phloem sap appeared to contain a very broad range of enzymes. Studies on the phloem of poppy have revealed that enzymes specifically located within companion cells, sieve elements, and laticifers (specialized cells that produce latex) appear to act coordinately in certain biosynthetic pathways (50–52). Thus, there is the definite possibility that similar biochemical pathways may well be restricted to either the sieve tube system or the companion cell-sieve element complex. A detailed characterization of the enzymes identified in the phloem sap will be required to test this hypothesis.

Enzymes involved in glycolysis are represented by highly enriched GO categories (supplemental Fig. 4). Here it is interesting to note that although the phloem translocation stream has often been considered to carry only non-reducing sugars, such as sucrose, sugar alcohols, and raffinose family sugars, the pumpkin phloem sap contains a number of enzymes that function in hexose/monosaccharide metabolism. This finding is consistent with earlier reports that the cucurbit phloem sap was found to contain appreciable hexose levels (53, 54). Indeed a recent study examining the phloem sap collected from a wide range of plant species revealed that reducing sugars may be a common component of the fixed carbon present in or being delivered through the translocation stream (55).

Another interesting set of enzymes included superoxide dismutase, ascorbate peroxidases, glutathione peroxidases, monodehydroascorbate reductase, glutathione reductases, thioredoxin, and thioredoxin reductases, all of which are involved in redox regulation and antioxidation responses. The presence of these proteins in the pumpkin phloem proteome is consistent with previous reports (29, 32).

Failure to Detect Low Abundance Phloem Proteins—In a previous study, we detected CmFTL1 and CmFTL2 in the phloem translocation stream of flowering pumpkin plants (12). However, only CmFTL2 is shown in our phloem proteome (supplemental Table I). As FT proteins are small (20 kDa) and highly conserved, the majority of the CmFTL1 sequence overlaps with that of CmFTL2, making it difficult to detect unique peptides. Because of its low abundance, the unique CmFTL1 peptide, VIGDVVDSFSR, was not detected in any of the 345 regular LC-MS/MS data sets collected in the present study. This prototypic peptide for CmFTL1 was detected earlier by a multiple reaction monitoring type experiment (12).

The absence of the CmFTL1 from the pumpkin phloem proteome indicates that with the current instrumentation used very low abundance proteins may not be fully represented. Thus, the challenge ahead will be to develop approaches that will allow us to identify these missing components of the phloem proteome. Examples will be components missing from the ubiquitin-proteolysis pathway (Fig. 4) and translation initiation factors (Fig. 5) that were not detected. Here it may be of interest to note that transcription factors (TFs), such as BTF3b, HAP5A, HAP5B, MBF1B, and two predicted TFs (supplemental Table I), were also identified in the phloem sap. However, the number was low and may well be an underestimate of the true population of TFs that move in the phloem translocation stream. Consequently the true size of the phloem proteome will most certainly be greater than the 1,209 proteins identified in the current study.

Conclusions—Three major issues in obtaining a large scale comprehensive characterization of a particular proteome are its quantity, complexity, and dynamic range of proteins in the sample. The present study revealed the phloem proteome to be at least 10 times larger in size than that of previously reported studies, identifying over 1,000 proteins. The success of this study was because of (a) the availability of analytical quantities of pumpkin phloem sap, (b) large scale multidimensional separation of proteins to reduce overall complexity, and (c) enrichment of certain proteins through use of co-immunoprecipitation approaches.

The phloem proteome likely forms an integral component of the whole plant communication system (20). The universal significance of the proteome is highlighted by conservation of the phloem proteome in species as diverse as the monocots (rice), eudicot (Arabidopsis and pumpkin) species, and trees (poplar). The presence of RNA-binding proteins within the phloem underscores an integrated system, including not only proteins but also transcripts (3, 4, 9, 56), small RNA (7, 57–59), and sugars as well as other small molecules. New paradigms are emerging through elucidation of the proteome. The possibility of both protein synthesis and turnover within the sieve tube system is highlighted by identification of significant components of the 26 S proteasome and protein translation machineries. Perhaps the most exciting possibility is the potential for the phloem to play not only a role in delivery of sugars to developing seeds but also a regulatory role through the iden-
titification of proteins thought previously to be seed-specific regulators of seed development (43).

Acknowledgments—We thank Eriko Miura and Byung-Kook Ham for assistance with the elf-5a and CmPP16 co-immunoprecipitation experiments.

¶ This work was supported by a University of California Davis intramural award (to W. J. L.); United States Department of Energy, Division of Energy Biosciences Grant DE-FG02-94ER20134 (to W. J. L.); National Science Foundation Grant IOS-07-15513 (to W. J. L.); and a grant from the New Zealand Foundation for Research Science and Technology (to T. J. L.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The on-line version of this article (available at http://www.mcponline.org) contains supplemental material.

§ Both authors contributed equally to this work.

¶¶ To whom correspondence should be addressed. Tel.: 530-752-5410; E-mail: wjlucas@ucdavis.edu.

Present address: School of Chinese Medicine Resource, College of Pharmacy, China Medical University, No. 91 Hsueh-Shih Rd., Taichung, Taiwan 40020, Republic of China.

** Present address: Dept. of Chemistry, Iowa State University, Ames, IA 50011.

§§ Present address: Pure Power Technology Ltd., P. O. Box 8928, Symonds St., Auckland 1150, New Zealand.

REFERENCES

1. Knoblauch, M., and van Bel, A. J. E. (1998) Sieve tubes in action. Plant Cell 10, 35–50

2. Voinnet, O., Vain, P., Angell, S., and Baulcombe, D. C. (1998) Systemic spread of sequence-specific transgene RNA degradation in plants is initiated by localized introduction of ectopic promoterless DNA. Cell 95, 177–187

3. Ruiz-Medrano, R., Xocostenile-Cázaras, B., and Lucas, W. J. (1999) Phloem long-distance transport of CmNACP mRNA: implications for supracellular regulation in plants. Development 126, 4405–4419

4. Xocostenile-Cázaras, B., Xiang, Y., Ruiz-Medrano, R., Wang, H. L., Monzer, J., Yoo, B. C., McFarland, K. C., Franceschi, V. R., and Lucas, W. J. (1999) Plant paralog to viral movement protein that potentiates transport of mRNA into the phloem. Science 283, 94–98

5. Gomez, G., Torres, H., and Pallas, V. (2005) Identification of translocatable RNA-binding phloem proteins from melon, potential components of the long-distance RNA transport system. Plant J. 41, 107–116

6. Foster, T. M., Lough, T. J., Emerson, S. J., Lee, R. H., Bowman, J. L., Forster, R. L., and Lucas, W. J. (2002) A surveillance system regulates selective entry of RNA into the shoot apex. Plant Cell 14, 1497–1508

7. Yoo, B. C., Krabler, F., Varkonyi-Gasic, E., Haywood, V., Archer-Evans, S., Lee, Y. M., Lough, T. J., and Lucas, W. J. (2004) A systemic small RNA signaling network in plants. Plant Cell 16, 1979–2000

8. Kim, M., Canio, W., Kessler, S., and Sinha, N. (2001) Developmental changes due to long-distance movement of a homeobox fusion transcript in tomato. Science 293, 287–290

9. Haywood, V., Yu, T. S., Huang, N. C., and Lucas, W. J. (2005) Phloem long-distance trafficking of GIBBERELLIC ACID-SENSITIVE RNA regulates leaf development. Plant J. 42, 49–68

10. Banerjee, A. K., Chatterjee, M., Yu, Y., Suh, S. G., Miller, W. A., and Hannapel, D. J. (2006) Dynamics of a mobile RNA of potato involved in a long-distance signaling pathway. Plant Cell 18, 3443–3457

11. Corbesier, L., Vincent, C., Jang, S., Fornara, F., Fan, Q., Searle, I., Giakoussis, A., Farrona, S., Gissot, L., Turnbull, C., and Coupland, G. (2007) Phloem protein movement contributes to long-distance signaling in floral induction of Arabidopsis. Science 316, 1030–1033

12. Lin, M. K., Belanger, H., Lee, Y. J., Varkonyi-Gasic, E., Taoka, K., Miura, E., Xocostenile-Cázaras, B., Gendler, K., Jongensen, R. A., Phinney, B., Lough, T. J., and Lucas W. J. (2007) FLOWERING LOCUS T protein may act as the long-distance flowering signal in the cucurbits. Plant Cell 19, 1488–1506

13. Matthieu, J., Warthmann, N., Kuttner, F., and Schmid, M. (2007) Export of FT protein from phloem companion cells is sufficient for floral induction in Arabidopsis. Curr. Biol. 17, 1055–1060

14. Tamaki, S., Matsuo, S., Wong, H. L., Yoki, S., and Shimamoto, K. (2007) Phloem protein is a mobile flowering signal in rice. Science 316, 1033−1036

15. Jaeger, K. E., and Wigge, P. A. (2007) FT protein acts as a long-range signal in Arabidopsis. Curr. Biol. 17, 1050–1054

16. Golecki, B., Schulz, A., Carstens-Behrens, U., and Kollmann, R. (1998) Evidence for graft transmission of structural phloem proteins or their precursors in heterografts of Cucurbitaceae. Planta 200, 630–640

17. Oparka, K. J., and Turgeon, R. (1999) Sieve elements and companion cells—traffic control centers of the phloem. Plant Cell 11, 739–750

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

19. Lucas, W. J., and Lee, J. Y. (2004) Plasmodesmata as a supracellular control network in plants. Nat. Rev. Mol. Cell Biol. 5, 712–726

20. Lough, T. J., and Lucas, W. J. (2006) Integrative plant biology: role of phloem long-distance macromolecular trafficking. Annu. Rev. Plant Biol. 57, 203–232

21. Xocostenile-Cázaras, B., Ruiz-Medrano, R., and Lucas, W. J. (2000) Proteinolytic processing of CmPP36, a protein from the cytoplasmic b2 subunit family, is required for entry into the phloem translocation pathway. Plant J. 24, 735–747

22. Aoki, K., Krabler, F., Xocostenile-Cázaras, B., and Lucas, W. J. (2002) A subclass of plant heat shock cognate 70 chaperones carries a motif that facilitates trafficking through plasmodesmata. Proc. Natl. Acad. Sci. U. S. A. 99, 16342–16347

23. Aoki, K., Suzuki, N., Fujimaki, S., Doehmeh, N., Hayashi, H., Yamaya, T., and Sakakibara, H. (2005) Destination-selective long-distance phloem proteins. Plant Cell 17, 1801–1814

24. Itaya, A., Ma, F., Qi, Y., Matsuda, Y., Zhu, Y., Liang, G., and Ding, B. (2002) Plasmodesmas-mediated selective protein trafficking between “symplastically isolated” cells probed by a viral movement protein. Plant Cell 14, 2071–2083

25. Yoo, B. C., Lee, J. Y., and Lucas, W. J. (2002) Analysis of the complexity of protein kinases within the phloem sieve tube system. Characterization of Cucurbita maxima calmodulin-like domain protein kinase 1. J. Biol. Chem. 277, 15325–15332

26. Haebel, S., and Kehr, J. (2001) Matrix-assisted laser desorption/ionization time of flight mass spectrometry: An invaluable technique for proteome analysis. Proteomics 1, 24–30

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

28. Giavalisco, P., Kapitza, K., Kolasa, A., Buhtz, A., and Kehr, J. (2006) Towards the proteome of Brassica napus phloem sap. Proteomics 6, 896–909

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

30. Aki, T., Shigyo, M., Nakano, R., Yoneyama, T., and Yanagisawa, S. (2008) N-terminal sequence of a protease isolated from the phloem of mature sieve tubes. Plant J. 51, 189–197

31. Rosenfeld, J., Capdevielle, J., Guillemot, J. C., and Ferrara, P. (1992) In-gel digestion of proteins for internal sequence analysis after one- or two-dimensional gel electrophoresis. Anal. Biochem. 203, 173–179

32. Lee, Y. J., Rice, R. H., and Lee, Y. M. (2006) Proteome analysis of human hair by two dimensional liquid chromatography coupled with tandem mass spectrometry (2-D LC-MS/MS). From protein identification to post-

Molecular & Cellular Proteomics 8.2 355
translational modification. Mol. Cell. Proteomics 5, 789–800
35. Lee, Y. J., Lachner, L. L., Nunnari, J. M., and Phinney, B. S. (2007) Shotgun cross-linking analysis for studying quaternary and tertiary protein structures. J. Proteome Res. 6, 3908–3917
36. Keller, A., Nesvizhskii, A. I., Kolker, E., and Aebersold, R. (2002) Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. Anal. Chem. 74, 5383–5392
37. Nesvizhskii, A. I., Keller, A., Kolker, E., and Aebersold, R. (2003) A statistical model for identifying proteins by tandem mass spectrometry. Anal. Chem. 75, 4646–4658
38. Blüthgen, N., Brand, K., Cajavec, B., Swat, M., Herzel, H., and Beule, D. (2005) Biological profiling of gene groups utilizing gene ontology. Genome Informatics 16, 106–115
39. Götz, S., García-Gómez, J. M., Terol, J., Williams, T. D., Nagaraj, S. H., Nueda, M. J., Robles, M., Talón, M., Dopazo, J., and Conesa, A. (2008) High-throughput functional annotation and data mining with the Blast2GO suite. Nucleic Acids Res. 36, 3420–3435
40. Hinojosa-Moya, J., Xoconostle-Cázares, B., Piedra-Ibarra, E., Mendez-Tenorio, A., Lucas, W. J., and Ruiz-Medrano, R. (2008) Phylogenetic and structural analysis of translationally controlled tumor proteins. J. Mol. Evol. 66, 472–483
41. Long, J. A., Ohno, C., Smith, Z. R., and Meyerowitz, E. M. (2006) TOPESS regulates apical embryonic fate in Arabidopsis. Science 312, 1520–1523
42. Kieffer, M., Stern, Y., Cook, H., Clerici, E., Maulbetsch, C., Laux, T., and Davies, B. (2006) Analysis of the transcription factor WUSCHEL and its functional homologue in Antirrhinum reveals a potential mechanism for their roles in meristem maintenance. Plant Cell 18, 560–573
43. Baerenfaller, K., Grossmann, J., Grobei, M. A., Hull, R., Hirsch-Hoffmann, M., Yalovsky, S., Zimmermann, P., Grossniklaus, U., Gruissem, W., and Baginsky, S. (2008) Genome-scale proteomics reveals Arabidopsis thaliana gene models and proteome dynamics. Science 320, 938–941
44. Schobert, C., Grossmann, P., Gottschalk, M., Komor, E., Peczvaradi, A., and Zurnieden, U. (1995) Sieve-tube exudate from Ricinus communis L. seedlings contains ubiquitin and chaperones. Planta 196, 205–210
45. Schobert, C., Baker, L., Szeredkenyi, J., Grossmann, P., Komor, E., Hayashi, H., Chino, M., and Lucas, W. J. (1998) Identification of immunologically related proteins in sieve-tube exudate collected from monocotyledonous and dicotyledonous plants. Planta 206, 245–252
46. Mathews, M. B., Sonenberg, N., and Hershey, J. W. B. (2007) Translational Control in Biology and Medicine, p. 9, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
47. Ishiwatari, Y., Fujiwara, T., McFarland, K. C., Nemoto, K., Hayashi, H., Chino, M., and Lucas, W. J. (1998) Rice phloem thioredoxin h has the capacity to mediate its own cell-to-cell transport through plasmodesmata. Planta 205, 12–22
48. Taoka, K.-i., Ham, B.-K., Xoconostle-Cázares, B., Rojas, M. R., and Lucas, W. J. (2007) Reciprocal phosphorylation and glycosylation recognition motifs control NCAPPl interaction with pumpkin phloem proteins and their cell-to-cell movement. Plant Cell 19, 1866–1884
49. Hayashi, H., Fukuda, A., Suzui, N., and Fujimaki, S. (2000) Proteins in the sieve element-companion cell complexes: their detection, localization and possible functions. Aust. J. Plant Physiol. 27, 489–496
50. Bird, D. A., Franceschi, V. R., and Facchin, P. J. (2003) A tale of three cell types: alkaloid biosynthesis is localized to sieve elements in opium poppy. Plant Cell 15, 2626–2635
51. Samanani, N., Alcantara, J., Bourgault, R., Zulak, K. G., and Facchin, P. J. (2006) The role of phloem sieve elements and laticifers in the biosynthesis and accumulation of alkaloids in opium poppy. Plant J. 47, 547–563
52. Ziegler, J., and Facchin, P. J. (2008) Alkaloid biosynthesis: metabolism and trafficking. Annu. Rev. Plant Biol. 59, 735–769
53. Richardson, P. T., and Baker, D. A. (1982) The chemical composition of cucurbit vascular exudates. J. Exp. Bot. 33, 1239–1247
54. Mitchell, D. E., Gadus, M. V., and Madore, M. A. (1992) Patterns of assimilate production and translocation in muskmelon (Cucumis melo L.). I. Diurnal patterns. Plant Physiol. 99, 959–965
55. van Bel, A. J. E., and Hess, P. H. (2008) Hexoses as phloem transport sugars: the end of a dogma? J. Exp. Bot. 59, 261–272
56. Roney, J. K., Khatibi, P. A., and Westwood, J. H. (2007) Cross-species translocation of mRNA from host plants into the parasitic plant dodder. Plant Physiol. 143, 1037–1043
57. Buhnt, A., Springer, F., Chappell, L., Baulcombe, D. C., and Kehr, J. (2008) Identification and characterization of small RNAs from the phloem of Brassica napus. Plant J. 53, 739–749
58. Pant, B. D., Buhtz, A., Kehr, J., and Scheible, W. R. (2008) MicroRNA399 is a long-distance signal for the regulation of plant phosphate homeostasis. Plant J. 53, 731–738
59. Lin, S. I., Chiang, S. F., Lin, W. Y., Chen, J. W., Tseng, C. Y., Wu, P. C., and Chiou, T. J. (2008) Regulatory network of microRNA399 and PHO2 by systemic signaling. Plant Physiol. 147, 732–746