Complementary Proteome and Transcriptome Profiling in Phosphate-deficient Arabidopsis Roots Reveals Multiple Levels of Gene Regulation*

Ping Lan‡, Wenfeng Li‡, and Wolfgang Schmidt‡§¶

Phosphate (P) deficiency impairs plant growth and productivity in many agricultural ecosystems, causing severe reductions in crop yield. To uncover novel aspects in acclimation to P starvation, we investigated the correlation between P deficiency-induced changes in transcriptome and proteome profiles in Arabidopsis roots. Using exhaustive tandem mass spectrometry-based shotgun proteomics and whole-genome RNA sequencing to generate a nearly complete catalog of expressed mRNAs and proteins, we reliably identified 13,298 proteins and 24,591 transcripts, subsets of 356 proteins and 3106 mRNAs were differentially expressed during P deficiency. Most dramatic changes were noticed for genes involved in P acquisition and in processes that either liberate P or bypass P/ATP-consuming metabolic steps, for example during membrane lipid remodeling and glycolytic carbon flux. The concordance between the abundance of mRNA and its encoded protein was generally high for highly up-regulated genes, but the analysis also revealed numerous discordant changes in mRNA/protein pairs, indicative of divergent regulation of transcription and post-transcriptional processes. In particular, a decreased abundance of proteins upon P deficiency was not closely correlated with changes in the corresponding mRNAs. In several cases, up-regulation of gene activity was observed solely at the protein level, adding novel aspects to key processes in the adaptation to P deficiency. We conclude that integrated measurement and interpretation of changes in protein and transcript abundance are mandatory for generating a complete inventory of the components that are critical in the response to environmental stimuli. Molecular & Cellular Proteomics 11: 10.1074/mcp.M112.020461, 1156–1166, 2012.

Bottom-up, high-throughput profiling of transcripts or proteins is a powerful method to analyze changes in biological processes, e.g., during development or upon environmental perturbations. Methods examining gene activity have improved dramatically during the past few years. Quantitative gel-free shotgun proteomics based on tandem mass spectrometry is now possible with high resolution, yielding detailed protein expression maps (1, 2). Similarly, the technologies for whole transcriptome sequencing (RNA-seq) ameliorates some of the caveats of DNA microarrays, permitting deep coverage of transcriptomic landscapes and detection of expression changes with a high dynamic range. Computational algorithms to analyze and display the data in a biologically meaningful way have become available. The task of assigning functions to the annotated genes/proteins, however, remains a major challenge. As opposed to a reductionist approach (i.e., studying one gene/protein at a time), systems biology attempts to address the system as a whole in order to understand the interdependence and dynamics of its components and to predict cellular behavior. Deciphering functional gene networks that mediate biological processes in concert will reveal a more holistic view of the processes under study and will provide an approach to manipulate the fitness and performance of plants. Because of the relatively low cost and ease of application, transcriptional profiling has been adopted as the method of choice by many labs to interrogate the adaptation of plants to environmental signals. A large set of genome-wide data is available regarding transcriptional changes induced by P deficiency, derived from microarray studies using roots, shoots and seedlings of Arabidopsis (3–10) and also other species such as wild mustard (11), rice (12), maize (13), tomato (14) and bean (15). However, biological function is chiefly carried out by proteins, and determination of protein abundance is mandatory to fully understand the function of a system. Generally, because of post-translational turnover and alternate translation efficiency, mRNA abun-
dance is not a reliable proxy of protein abundance, and modest congruency of the two levels has been reported except for high abundance transcripts and “molecular machines” (16–22). Much has been learned regarding Pi deficiency responses from proteomic approaches (23–25). Despite these efforts, data sets on Pi deficiency-induced changes in the proteome remained somewhat patchy, because of a skew toward high-abundant proteins and work carried out on different species. The parallel analysis and interpretation of disparate “omics” data sets integrate different levels of the cellular response and allows for the uncovering of regulatory mechanisms beyond transcription. A recent study addressing the regulation of transcript and protein abundance in mammalian cells highlights the importance of translational control for the cellular abundance of protein (26), underlining the importance of an integrated view on gene activity. Parallel analysis of transcriptome and proteome expression data is challenging because of biological (e.g. different mRNA and protein stability) and technical biases (e.g. differences in resolution of transcriptomic and proteomic profiles). Because proteomic and transcriptomic data are often collected separately, it remains often unclear whether discordant mRNA and protein expression merely reflects experimental noise or represents biological meaningful post-transcriptional regulation.

Phosphorous, mainly taken up as phosphate (Pi)1 by plants, is an essential macronutrient of crucial importance in signaling, metabolism and photosynthesis. Because of its tendency to form complexes with soil cations, the low bioavailability of Pi often limits plant growth. In natural ecosystem, the availability of and demand for Pi are major determinants for the composition of plant communities. Phosphate deficiency is a major cause of severe yield losses in crops and poor quality of edible plant parts. Low Pi availability is often corrected by the use of fertilizers but which are associated with environmental damages and substantial costs. To cope with low Pi availability, plants have evolved a plethora of adaptive processes, that vary among species and cultivars. A prerequisite for developing Pi-efficient germplasm is a thorough understanding of the mechanisms that control cellular Pi homeostasis. Proteomic and transcriptomic profiling studies have uncovered several robustly changed processes in Pi-deficient plants, including the remodeling of lipid metabolism, changes in glycolytic carbon flux, alterations in root development, and signaling pathways (27–30). Based on microarray experiments, the expression of ~1000 genes was estimated to be differentially expressed upon Pi deficiency. How tightly these changes correspond to changes of the proteomic profiles is not known.

In an attempt to identify nodes that are important for the function of roots under conditions of Pi starvation, we cataloged differences in the abundance of mRNAs and proteins by integrated profiling of gene activity using RNA-seq and high-resolution quantitative iTRAQ proteomics. Transcriptomic and proteomic expression profiles were remarkably well correlated, but our analysis also revealed several exceptions indicative of different control mechanisms that dictate gene activity. We argue that this “molecular compendium” of gene and protein expression data provides an informative readout about the physiological state of Pi-deficient roots, uncovering several novel aspects in the metabolism of Pi-deficient plants and provides a comprehensive reference map of gene activity in Arabidopsis roots.

MATERIALS AND METHODS

Plant Growth—Arabidopsis (Arabidopsis thaliana L.) plants were grown in a growth chamber on an agar-based medium as previously described (31). Seeds of the Col-0 accession were obtained from the Arabidopsis Biological Resource Center (ABRC, Ohio State University, Columbus). Seeds were surface sterilized by immersing them in 5% (v/v) NaOCl for 5 min and 70% ethanol for 7 min, followed by four rinses in sterile water. Seeds were placed onto Petri dishes and kept for 1 d at 4 °C in the dark, after the plates were transferred to a growth chamber and grown at 21 °C under continuous illumination (50 μmol m−2 s−1, Philips TL lamps). The medium was composed of (mM): KNO3 (5), MgSO4 (2), Ca (NO3)2 (2), KH2PO4 (2.5), (µM): H3BO3 (70), MnCl2 (14), ZnSO4 (1), CuSO4 (0.5), NaCl (10), Na2MoO4 (0.2) and 40 μM FeEDTA, solidified with 0.3% Phytagel (Sigma-Aldrich). Sucrose (43 mm) and 4.7 mm Mes were included and the pH was adjusted to 5.5. After 10 days of precultivation, plants were transferred either to fresh agar media without phosphate or to fresh control media and grown for 3 days. The lower concentration of potassium because of the absence of KH2PO4 in the Pi-free media was corrected by the addition of KCl.

Protein Extraction—Roots from control plants and Pi-deficient plants (13-day-old) were ground in liquid nitrogen and suspended in 10× volume of precooled acetone (−20 °C) containing 10% (w/v) trichloroacetic acid and 0.07% (v/v) 2-mercaptoethanol. Proteins were then precipitated for 2 h at −20 °C after thorough mixing. Proteins were collected by centrifuging at 35,000 × g (JA-20 108 rotor; Beckman J2-HS) at 4 °C for 30 min. The supernatant was carefully removed, and the protein pellets were washed twice with cold acetone containing 0.07% (v/v) 2-mercaptoethanol and 1 mM phenylmethylsulfonyl fluoride and a third time with cold acetone without 2-mercaptoethanol. Protein pellets were dried by lyophilization and stored at −80 °C or immediately extracted using protein extraction buffer composed of 6 m urea, 50 mM triethylammonium bicarbonate, pH 8.5, and 2% (3-[3-cholamidopropyl]dimethylammonio])propanesulfonate for 1 h at 6 °C under constant shaking. Protein extracts were centrifuged at 19,000 × g for 20 min at 10 °C. The supernatant was then collected, and the protein concentration was determined using a protein assay kit (Pierce, Rockford, IL).

In-solution Trypsin Digestion and iTRAQ Labeling—Total protein (100 μg) was reduced by adding dithiorthreitol to a final concentration of 10 mM and incubated for 1 h at room temperature. Subsequently, iodoacetamide was added to a final concentration of 40 mM, and the mixture was incubated for 1 h at room temperature in the dark. Then, dithiorthreitol (10 mM) was added to the mixture to consume any free iodoacetamide by incubating the mixture for 1 h at room temperature in the dark. Proteins were then diluted by 50 mM triethylammonium.

1 The abbreviations used are: Pi, phosphate; SQDG, sulfoquinovosyldiacylglycerol.

Molecular & Cellular Proteomics 11.11 1157
bicarbonate and 1 mM CaCl₂ to reduce the urea concentration to less than 0.6 M and digested with 40 μg of modified trypsin (Promega, Madison, WI) at 37 °C overnight. The resulting peptide solution was acidified with 10% trifluoroacetic acid and desalted on a C18 solid-phase extraction cartridge.

Desalted peptides were then labeled with iTRAQ reagents (Applied Biosystems, Foster City, CA) according to the manufacturer’s instructions. Control samples (proteins extracted from roots of control plants) were labeled with reagent 114; samples from Pi-deficient roots were labeled with reagent 117. Three independent biological experiments with two technical repeats each were performed. The reaction was allowed to proceed for 1 h at room temperature. Subsequently, treated and control peptides were combined and further fractionated offline using high-resolution cation-exchange chromatography (PolySulfoethyl A, 5 μm, 200-Å bead). In total, 35 fractions were collected and combined into 16 final fractions. Each final fraction was lyophilized in a centrifugal speed vacuum concentrator. Samples were stored at −80 °C.

LC-MS/MS Analysis—Liquid chromatography was performed on a nanoACQUITY UPLC System (Waters, Milford, MA) coupled to an LTQ Orbitrap Velos hybrid mass spectrometer (Thermo Scientific, Waltham, MA) equipped with a PicoView nanospray interface (New Objective, Woburn, MA). Peptide mixtures were loaded onto a 75 μm × 250 mm nanoACQUITY UPLC BEH130 column packed with C18 resin (Waters, Milford, MA) and separated using a segmented gradient in 120 min from 5 to 40% solvent B (95% acetonitrile with 0.1% formic acid) at a flow rate of 300 nl/min. Solvent A was 0.1% formic acid in water. The samples were maintained at 4 °C in the autosampler. The LTQ Orbitrap was operated in the positive ion mode with the following acquisition cycle: a full scan (m/z 350−1600) recorded in the Orbitrap analyzer at resolution R 60,000 was followed by MS/MS of the 10 most intense peptide ions with collision-induced dissociation high energy collision-induced dissociation acquisition of the same precursor ion. Collision-induced dissociation was done with collision energy of 35%. high energy collision-induced dissociation-generated ions were detected in the Orbitrap with collision energy of 45%.

Database Search—Two search algorithms, Mascot (version 2.2.06, Matrix Science) and SEQUEST, which is integrated in Proteome Discoverer software (version 1.3.0.339, Thermo Scientific), were used to simultaneously identify and quantify proteins (supplemental Fig. S2). Searches were made against the Arabidopsis protein database (TAIR10 20110103, 27416 sequences; ftp://ftp.arabidopsis.org/home/tair/Sequences/blast_datasets/TAIR10_blastsets/TAIR10_pep20110103 representative gene model) concatenated with a decoy database containing the randomized sequences of the original database. Peak list data (MGF) files used for database searches were generated from Xcalibur raw files using a program in the MassMatrix conversion tools (v. 1.4, http://www.massmatrix.net). The protein sequences in the database were searched with trypsin digestion at both ends and two missed cleavages allowed, fixed modifications of carbamidomethylation at Cys, variable modifications of oxidation at Met and iTRAQ 4plex at Tyr; peptide tolerance was set at 0.8 D. iTRAQ 4plex was chosen for quantification during the search simultaneously. The search results were passed through additional filters, peptide confidence more than
For protein quantitation, only unique peptides were used to quantify proteins. These filters resulted in a false discovery rate of less than 1% after decoy database searches were performed. For biological repeats, spectra from the two technical repeats were combined into one file and searched. MS/MS spectra of single peptide-based identifications are deposited in the public proteome database Tranche (ProteomeCommons.org) under the entry name TfZmsyU6n.

**FIG. 2.** Concordance between changes in the abundance of mRNA and its encoded protein. A–E, Correlation between protein and transcript fold-changes upon P deficiency for all transcript/protein pairs (A), up-regulated transcripts (B), down-regulated transcripts (C), significantly up-regulated transcripts (D) and significantly down-regulated transcripts (E). F, Pearson correlations of the comparisons shown in A–E. * indicates significant correlation at $p < 0.01$.

**FIG. 3.** Detection of mRNAs and their cognate proteins. A, Dependence of protein detection on the number of reads of the corresponding transcript. B, Overrepresented gene ontology categories (cellular component) of detected and nondetected proteins.

95% ($p < 0.05$), before exporting the data. For protein quantitation, only unique peptides were used to quantify proteins. These filters resulted in a false discovery rate of less than 1% after decoy database searches were performed. For biological repeats, spectra from the two technical repeats were combined into one file and searched. MS/MS spectra of single peptide-based identifications are deposited in the public proteome database Tranche (ProteomeCommons.org) under the entry name TfZmsyU6n.

**Statistical Analysis**—Proteins identified and quantified by at least in two biological repeats were considered to further analyze the abundance change in response to P deficiency using a method described by (32). In brief, the log2 ratios of the 7026 quantified proteins overlapping in at least two biological repeats were calculated and analyzed for normal distribution. For a given protein in one biological repeat, the ratio was calculated as the reverse log2 of the median of the log2 value of all peptide ratios and averaged across the
biological replicates. Next, mean and S.D. were calculated and 95% confidence (Z score) was used to select those proteins whose distribution was far from the main distribution. For the down-regulated proteins, the confidence interval was $-0.263053843 (0.052103724, \text{mean ratio of the 7026 proteins})$ corresponding to a protein ratio of 0.83. Similarly, for the up-regulated proteins, the confidence interval was calculated (mean ratio $+1.96 \times \text{S.D.}$), corresponding to a protein ratio of 1.29. Protein ratios outside this range were defined as being significantly different at $p = 0.05$.

**RNA Sequencing**—Total RNA was extracted from roots grown under control and Pi-deficient conditions using RNeasy Plant Mini kit (Qiagen). Equal amounts of RNA collected from three independent experiments and used for sequencing. cDNA libraries for sequencing were prepared from 5 μg of total RNA following protocols provided by the instrument manufacturer (Illumina). The cDNA libraries were enriched by 15 cycles of PCR amplification. The resulting cDNA libraries were sequenced on a single lane per sample of an Illumina Genome Analyzer II. The RNAseq and data collection followed published protocols (33). The length of the cDNA library ranged from 250 to 300 bp with a 5’-adapter of 58 bp and a 3’-adapter of 63 bp at both ends. The fragment length of the cDNA ranged from 129 to 179 bp.

**Computing Differentially Expressed Genes**—Adapters were trimmed from reads and ~20.0 M 80-mers reads for each sample from treated and untreated roots were obtained in three replicates. Reads were aligned to the TAIR10 genome using the BLAT program (34) with minimum 95% identity, but only the highest identity for each read was considered for mapping. Multireads were distributed in proportion to the number of unique and splice reads recorded at similar loci using Enhanced Read Analysis of Gene Expression (ERANGE) strategy (33). RPKM values (reads per 1Kbps of exon model per million mapped reads) were computed as described in (33). A transcript was defined as present when it was detected with at least five reads in at least two experiments within one growth type (i.e. either treated or untreated). A gene was defined as differentially expressed if the $p$ value for the comparison of the RPKM values between samples from treated and untreated plants was $< 0.05$ based on Student’s $t$ test.

**RESULTS AND DISCUSSION**

**Changes in the Root Proteome Upon Pi Starvation**—Phosphate deficiency-induced changes in the proteome of Arabidopsis roots were quantitatively cataloged using the iTRAQ technology. Proteins were extracted, digested in solution, and iTRAQ-labeled peptides were analyzed by liquid chromatography combined with tandem mass spectroscopy on an LTQ Orbitrap with higher-energy collisional dissociation (HCD) and collision-induced dissociation capabilities. This combination allows for an accurate identification of the peptides and pre-
cise quantification of the iTRAQ label with a wide scan range. Two search algorithms, Mascot and SEQUEST, were used to identify proteins. Using this strategy, we identified 57,153 unique peptides from 1,534,861 spectra, corresponding to 17,007 proteins (10,794 proteins/protein groups) in the three experiments (Fig. 1A, supplemental Data Sets S1, S2, S3). A subset of 13,298 proteins was identified by at least two peptides or in two experiments, 10,256 proteins were identified in all three experiments (Fig. 1A). Robustness of the analysis is supported by the nearly identical SCX chromatograms of the three samples (supplemental Fig. S1). Pi deficiency-induced changes in protein abundance values showed a normal distribution (Fig. 1B), from which cutoff values for proteins with significantly different expression at \( p < 0.05 \) were calculated (Pi-deficient/control plants ratio of \( \geq 1.29 \) for induced proteins and \( \leq 0.83 \) for repressed proteins). Using this criterion, 356 proteins from a total subset of 7,026 quantified proteins was defined as differentially expressed upon Pi deficiency (Fig. 1C; supplemental Data Set S4). The highest increase in abundance was observed for the high-affinity Pi transporter PHT1;4, the purple acid phosphatase ACP5 (PAP17), the transcription factor SPX1, and for several proteins with functions in lipid metabolism, including the nonspecific phospholipase NPC4 and the UDP sulfoquinovosylosyl pyruvate: UDP-sulfoquinovosyl pyruvate: UDP-sulfoquinovosylosyl; Pyr: pyruvate; Cho: choline; EA: ethanolamine.

Comparison of Cognate Gene and Protein Expression—To compare changes in protein abundance with alterations in transcript levels, we conducted RNA-seq experiments with triplicate mRNA samples extracted from Pi-sufficient and Pi-deficient roots. Approximately 20 million reads per sample were acquired on the Illumina GAIIx platform and aligned to the TAIR10 release of the Arabidopsis genome. More than 21,000 transcripts were detected in roots from both growth types in all three experiments (Fig. 1D). Replicate experiments showed a generally high overlap of the detected transcripts. With more stringent criteria (\( \geq \) five reads in all three runs), the number of transcripts was ca. 20,000 in both growth types (Fig. 1D), a subset of 3,106 genes was differentially expressed between control and Pi-deficient

![Membrane lipid remodeling upon Pi deficiency](image)

Fig. 5. Membrane lipid remodeling upon Pi deficiency. Differentially expressed genes/proteins are framed in boxes, values on top of the boxes indicate 117/114 ratios of quantifiable proteins. The weight of the arrows is proportional to the abundance of the transcript in RPKM on a log2 scale. Black arrows show the constitutive level, red (up-regulated) arrows indicate transcript changes upon Pi deficiency, and green circles Pi-releasing reactions. Glu: glucose; Glu1P: glucose-1-phosphate; UDP-Glu: UDP-glucose; UDP-SQ: UDP-sulfoquinovosyl; Pyr: pyruvate; Cho: choline; EA: ethanolamine.
plants at \( p < 0.05 \) (supplemental Data Set S5). Transcripts were detected for 97.7% of the proteins (Fig. 1E).

To compare the proteome with the transcriptome, we first matched all differentially expressed mRNAs with quantifiable proteins. We then generated a second list comparing proteins with their cognate mRNAs by sorting the proteins according to their \( \Delta\text{Pi}/\Pi \) ratio. We observed a positive correlation of \( r = 0.51 \) when all significantly changed mRNAs with a cognate protein were considered regardless of the direction of the change (Figs. 2A, 2F). Restricting the analysis to pairs in which the mRNA was up-regulated, markedly increased the correlation (\( r = 0.58 \); Figs. 2B, 2F). No correlation between transcript and protein abundance was observed for transcripts with decreased abundance upon Pi deficiency (Fig. 2C). For protein/mRNA pairs in which the protein was significantly up-regulated (117/114 \( > 1.29 \)), the highest positive correlation \( (r = 0.61) \) between the two levels was calculated. Interestingly, a negative correlation was observed between protein and mRNAs for significantly down-regulated proteins (Figs. 2E, 2F). This indicates that, contrary to expectations, the vast majority of the down-regulated proteins was associated with a higher abundant rather than with a lower abundant transcript. This correlation was, however, not significant at \( p < 0.01 \).

Two major factors restricted the number of identified proteins relative to the detection of their cognate transcripts: the number of mRNA copies in the sample and the subcellular localization of the protein. Comparing the number of reads recorded for transcripts corresponding to identified and not identified proteins, a transition is reached at around 100 reads, above which the products of the majority of transcripts was detected (Fig. 3A). As anticipated, proteins tightly associated with membranes are underrepresented in the pool of identified proteins relative to the predicted proteome (Fig. 3B).

A Protein-Protein Interaction Network Reveals New Players in Pi Homeostasis—To uncover functional aspects associated with the identified proteins, we constructed a protein-protein network based on predicted or verified interologs and compared protein abundance with the transcript level. Using all differentially expressed proteins as input, a network comprising 43 proteins was generated that can be divided into four subclusters (Fig. 4). The largest subcluster consists of structural constituents of ribosomes and proteins involved in ribosomal biogenesis which showed, with one exception, de-
Fig. 7. Effects of phosphate deficiency on the abundance of transcripts and proteins engaged in glucose metabolism. Differentially expressed genes/proteins are framed in boxes, values on top of the boxes indicate 117/114 ratios of quantifiable proteins. The weight of the arrows is proportional to the abundance of the transcript in RPKM on a log2 scale. Gray arrows show the constitutive level, red (up-regulated)
increased abundance upon Pi deficiency. Notably, none of the corresponding transcripts was differentially expressed. A small subcluster was comprised of histones and proteins involved in nucleosome assembly, pointing to a potential importance of histone composition for Pi homeostasis. A third subcluster contains several RNA-binding proteins, some of which have been attributed to RNA processing or to the regulation of meiosis. This cluster also contains the PHOSPHATE TRANSPORTER TRAFFIC FACILITATOR 1 (PHF1), which is required for the transport of high-affinity phosphate transporters from the ER to the plasma membrane in Arabidopsis and rice (36, 37). Interestingly, PHF1 is connected to the network by a single edge via the binding partner of the target of rapamycin (TOR) kinase RAPTOR1B. TOR is conserved among eukaryotes and important for the stimulation of cell growth in response to nutrients and may link growth, nutrient availability, and translation efficiency (38). Both PHF1 and RAPTOR1B are up-regulated at the protein level, but only PHF1 transcript levels were increased upon Pi deficiency. The fourth subcluster contains, among other proteins, three ubiquitin conjugases, and SmD3, a core protein of small nuclear ribonucleoprotein with a role in mRNA splicing (39).

Remodeling of Lipid Membrane Composition Massively Recycles Pi—In order to maintain membrane integrity and functionality, the lipid composition of membranes is remodeled under Pi-deficient conditions by replacing phospholipids with the galactolipid digalactosyldiacylglycerol and a sulfolipid, sulfogalactosyldiacylglycerol (SQDG) (40, 41). In general, our data support existing models for membrane remodeling under Pi-deficient conditions but add several novel aspects. Degradation of phospholipids by the nonspecific phospholipase NCP4, the synthesis of SQDG via SQD1 and SQD2, and the hydrolyzation of glycerophosphodiester by GDPD1 to form glycerol-3-phosphate (G3P), were strongly up-regulated at both the protein and RNA level (Fig. 5). Inorganic phosphate is further liberated by the phosphatidate phosphohydrolase PAH1 via the generation of diacylglycerol from phosphatidic acid. Diacylglycerol may be channeled into the synthesis of MGDG and SQDG. The synthesis of SQDG is supported by APR1-mediated induction of adenosine 5-phosphosulphate reduction to sulfate. Notably, the vacuolar sulfate effluxer SULT4A1 was found to be up-regulated exclusively at the protein level, supposedly to provide sulfate for SQDG biosynthesis. Four acyl carrier proteins, cofactors of fatty acid bio-synthesis, were induced by Pi starvation solely at the protein level. The function of ACPs in Pi homeostasis has not been established, but it is reasonable to speculate that their concerted induction is indicative of increased de novo fatty acid biosynthesis. Three G3P transporters targeted to the plasma membrane, G3PP1, G3PP2, and G3PP3, were up-regulated at the transcript level. G3P is a substrate for purple acid phosphatases, which are localized in the cytoplasm or are secreted. The expression of several purple acid phosphatases was induced upon Pi deficiency, in particular ACP5 which was strongly increased in abundance both at the transcript and protein level. Interestingly, ACP5 is closely co-expressed with GDPD1 (Fig. 6), suggesting a concerted action of the two proteins. ACP5 is localized at the cell surface, making it tempting to speculate that G3P is shuttled through the plasma membrane where Pi is liberated by secreted ACP5. A recently described phosphoethanolamine/phosphocholine phosphatase, C (42), was also strongly induced upon Pi deficiency. It has been assumed that PEPC1 is involved in the liberation of Pi from PLs via formation of phosphocholine and phosphoethanolamine (42).

**P-deficiency-induced Alterations in Carbon Flux**—Under Pi-limiting conditions, glycolysis is compromised by the large decrease in ATP and Pi. Alternative reactions that can bypass ATP- or Pi-consuming reactions are thought to be activated upon Pi deficiency to maintain carbon flux (43, 28). Our data support the importance of some of these reactions. For example, PPI-dependent phosphofructokinase, catalyzing the conversion of Fru6P to Fru1,6P and PPPC1/PPC2, mediating the bypass of pyruvate kinase via OOA/malate production, were highly expressed constitutively and further induced upon Pi deficiency (Fig. 7). Sucrose breakdown by invertase was reduced, possibly because of a large surplus of sugar imported from the leaves. Contrary to what might be expected, three genes encoding the (irreversible) synthesis of Suc6P from UDP-Glu, SPS1F, SPS2F, and SPS4F, were significantly up-regulated upon Pi deficiency, yielding one Pi liberated from Fru6P and another one liberated during the synthesis of sucrose. Strongly up-regulated, both at the protein and transcript level, was also the formation of Pi via PPA4 and PPA5, associated with the production of UDP-Glu from Glu1P. Sucrose synthesis from hexosephosphate via sucrose phosphate synthase may thus represent a further Pi-yielding alteration of carbon flux in Pi-deficient roots. The production of Pi, via PPSPASE1, a novel type of inorganic pyrophosphatases that catalyze the cleavage of pyrophosphate (44), was also highly induced in Pi-deficient plants. PPSPASE1 has not been linked to a particular reaction. Other suggested bypass reactions, such as the production of 3-PGA via nonphosphorylating NADP-G3P dehydrogenase or the PPI-dependent generation of pyruvate via PPDF were expressed at a very low level and remained unchanged upon Pi deficiency.
**CONCLUSIONS**

Integration of different omics approaches is a promising route to obtain a multilayered picture of regulatory processes in a perturbed system. The goals of the current approach were threefold. First, we attempted to evaluate the importance of post-transcriptional processes by elucidating how tightly the transcribed mRNA is linked to protein abundance. Second, we aimed to uncover novel aspects of acclimatory processes induced by P deficiency by a thorough parallel analysis of changes in protein and mRNA abundance. As a further goal, we intended to provide a high resolution compendium of changes in gene expression of Arabidopsis root upon exposure of P starvation, an important abiotic stress. The technology used here allowed for an optimal and nearly complete proteome and transcriptome coverage, excluding only a relatively small subset of low-abundance proteins and proteins that are tightly linked to membranes. Whereas up-regulated proteins corresponded well to changes in the abundance of their cognate transcripts, for most of the strongly down-regulated proteins no significant change in the corresponding mRNA was observed. The lack of changes in protein abundance in these genes might be because of post-translational (down) regulation of the protein activity to avoid a de novo cycle of synthesis after the stress is relieved. Another possibility is that the time-course of the decline differs between mRNAs and proteins and does not allow for monitoring changes at the protein level at the observed time point. Third, transcript regulation may be "unwanted" but unavoidable for a subset of mRNAs, because of the broad action of transacting factors, whereas protein abundance remains stable, being controlled by post-translational mechanisms.

We observed the strongest expression changes upon P deficiency for genes involved in the remodeling of membrane lipids, in the acquisition of P, and in carbohydrate flux. Although the importance of these responses has been described previously, our data add several novel aspects to the numerous P deficiency responses of Arabidopsis roots. Protein expression landscapes revealed new players of potential importance for cellular P homeostasis such as members of the acyl carrier proteins family or the binding partner of the TOR kinase RAPTOR1B that were solely regulated at the protein level. Another example is the iron effluxer IREG1 which was strongly up-regulated at the protein level, but the cognate mRNA was unresponsive to P deficiency. Iron levels are increased in P-deficient plants (5), but the reason for this is not obvious. Increased uptake via the major ferrous transporter IRT1 is unlikely because expression of IRT1 and FRO2, encoding the oxidoreductase FRO2 that mediates the obligatory reduction of ferric iron prior to IRT1-mediated uptake of iron, are markedly down-regulated upon P starvation. In the light of the strong effects of P deficiency on lipid remodeling, it is reasonable to speculate that the permeability of root membranes for metals is altered by the substitution of phospholipids by galactolipids. This assumption is supported by the down-regulation of several other transcripts coding for metal transporters in P-deficient plants.

In conclusion, our data demonstrate that a complete picture of gene activity is only observed when both the transcriptome and proteome levels are surveyed and that transcriptional regulation can be, at least partly, uncoupled from protein regulation. The integrated measurement and comparison of mRNA and protein abundances revealed extensive translational and post-translational regulation. We argue that the technology applied here is approaching completeness for both levels which allows for a correct, bias-free assessment of gene activity changes caused by environmental perturbations. It can further be stated that the highly complex relationship between the two levels is rather because of biological regulation than caused by methodical caveats. The technology used here should prove useful in further exploring complex biological systems.

Acknowledgments — We thank Yuki Nakamura, Wan-Hsing Cheng, and Tom Buchkout for helpful comments on the manuscript, Cole Lu (Schmidt lab) for figure artwork and the bioinformatics and proteomic core facilities at the Institute of Plant and Microbial Biology for proficient support.

* This work was supported by an Academia Sinica Thematic grant to W.S. (grant No. 02234g). Proteomic mass spectrometry analyses were performed by the core facilities for Proteomic Research located at the Institute of Biological Chemistry, Academia Sinica, supported by a National Science Council grant (NSC 98-3112-B-001-023, NSC 99-3112-B-001-025) and Academia Sinica.

This article contains supplemental Figs. S1 and S2 and Data Sets S1 to S5.

Public data base accession number for RNA-seq data (NCBI): SRA050356.1.

To whom correspondence should be addressed: Institute of Plant and Microbial Biology, Academia Sinica, Taipei, 11529 Taiwan. Tel.: 886-0912034404; E-mail: wosh@gate.sinica.edu.tw.

**REFERENCES**

1. Ong, S. E., and Mann, M. (2005) Mass spectrometry-based proteomics turns quantitative. Nat. Chem. Biol. 1, 252–262.
2. Liberman, L. M., Sozzani, R., and Benfey, P. N. (2012) Integrative systems biology: an attempt to describe a simple weed. Curr. Opin. Plant Biol. 15, 162–167.
3. Wu, P., Ma, L., Hou, X., Wang, M., Wu, Y., Liu, F., and Deng, X. W. (2003) Phosphate starvation triggers distinct alterations of genome expression in Arabidopsis roots and leaves. Plant Physiol. 132, 1260–1271.
4. Hammond, J. P., Bennett, M. J., Bowen, H. C., Broadley, M. R., Eastwood, D. C., May, S. T., Rahn, C., Swarup, R., Woolaway, K. E., and White, P. J. (2003) Changes in gene expression in Arabidopsis shoots during phosphate starvation and the potential for developing smart plants. Plant Physiol. 132, 578–596.
5. Misson, J., Raghothama, K. G., Jain, A., Jouhet, J., Block, M. A., Bligny, R., Ortet, P., Creff, A., Somerville, S., Rolland, N., Doumas, P., Nacey, P., Herrera-Estrella, L., Nussaume, L., and Thibaud, M. C. (2005) A genome-wide transcriptional analysis using Arabidopsis thaliana Affymetrix gene chips determined plant responses to phosphate deprivation. Proc. Natl. Acad. Sci. U.S.A. 102, 11934–11939.
6. Bari, R., Datt, Pant, B., Stitt, M., and Scheible, W. R. (2006) PHO2, microRNA399, and PHR1 define a phosphate-signaling pathway in plants. Plant Physiol. 141, 986–999.
7. Morcuende, R., Bari, R., Gibon, Y., Zheng, W., Pant, B. D., Blázing, O.,
Proteome and Transcriptome Profiling in Arabidopsis Roots

Usadel, B., Czeckowski, T., Udvardi, M. K., Stitt, M., and Scheible, W. R. (2007) Genome-wide reprogramming of metabolism and regulatory networks of Arabidopsis in response to phosphorus. Plant Cell Environ. 30, 85–112

Müller, R., Morant, M., Jarmer, H., Nilsson, L., and Nielsen, T. H. (2007) Genome-wide analysis of the Arabidopsis leaf transcriptome reveals interaction of phosphate and sugar metabolism. Plant Physiol. 143, 156–171

Marchive, C., Yehudai-Resheff, S., Germain, A., Fei, Z., Jiang, X., Juddins, J., Wu, H., Fernie, A. R., Fait, A., and Stern, D. B. (2009) Abnormal physiological and molecular mutant phenotypes link chloroplast polynucleotide phosphorylase to the phosphate deprivation response in Arabidopsis. Plant Physiol. 151, 905–921

Lin, W. D., Liao, Y. Y., Yang, T. J., Pan, C. Y., Buckhout, T. J., and Schmidt, W. (2011) Coexpression-based clustering of Arabidopsis root genes predicts functional modules in early phosphate deficiency signaling. Plant Physiol. 155, 1383–1402

Hammond, J. P., Broadley, M. R., Craigon, D. J., Higgins, J., Emmerson, Z. F., Townsend, H. J., White, P. J., and May, S. T. (2005) Using genomic DNA-based probe-selection to improve the sensitivity of high-density oligonucleotide arrays when applied to heterologous species. Plant Methods 1, 10

Wasaki, J., Yonetani, R., Kuroda, S., Shinano, T., Yazaki, J., Fujii, F., Shimbo, K., Yamamoto, K., Sakata, S., Katsuki, N., Kikuchi, S., Yamagishi, M., and Osaki, M. (2003) Transcriptomic analysis of metabolic changes by phosphate stress in rice plant roots. Plant Cell Physiol. 44, 1515–1523

Calderon-Vazquez, C., Ibarra-Laclette, E., Caballero-Perez, J., and Herrera-Estrella, L. (2008) Transcript profiling of Zea mays roots reveals gene responses to phosphate deficiency at the plant- and species-specific levels. J. Exp. Bot. 59, 2479–2497

Wang, Y. H., Garvin, D. F., and Kochian, L. V. (2002) Rapid induction of regulatory and transporter genes in response to phosphorus, potassium, and iron deficiencies in tomato roots. Evidence for cross talk and root/ rhizosphere-mediated signals. Plant Physiol. 130, 1361–1370

Hernández, G., Ramírez, M., Valdes-Lopez, O., Tesfaye, M., Graham, M. A., Czechowski, T., Schlereth, A., Wndray, M., Erban, A., Cheung, F., Wu, H. C., Lara, M., Town, C. D., Kopka, J., Udvardi, M. K., and Vance, C. P. (2007) Phosphate stress in common bean: root transcript and metabolic responses. Plant Physiol. 144, 752–767

Ideker, T., Thorsson, V., Flanagan, J. A., Christmas, R., Buhrer, J., Eng, J. K., Bumgarner, R., Goodlett, D. R., Aebersold, R., and Hood, L. (2002) Integrated genomic and proteomic analyses of a systematically perturbed metabolic network. Science 292, 929–934

Rogers, S., Girolami, M., Koich, W., Waters, K. M., Liu, T., Thrall, B., and Wiley, H. S. (2008) Investigating the correspondence between transcriptomic and protein expression profiles using coupled cluster models. Bioinformatics 41, 2894–2900

Griffin, T. J., Gygi, S. P., Ideker, T., Rist, B., Eng, J., Hood, L., and Aebersold, R. (2002) Complementary profiling of gene expression at the transcriptome and proteome levels in Saccharomyces cerevisiae. Mol. Cell. Proteomics 1, 323–333

Chen, G., Gharib, T. G., Huang, C. C., Taylor, J. M., Misek, D. E., Kardia, S. L., Ghodsi, M. T., Siminovitch, L., Conde, V. M., Laird, P. W., Fleming, A. R., et al. (2005) Utility of the cancer genome atlas project to develop new cancer therapies. Science 310, 1944–1949

Baginsky, S., Kleffmann, T., von Zychlinski, A., and Gruissem, W. (2005) Analysis of short oligonucleotide probes and RNA profiling data from Arabidopsis thaliana chloroplasts. J. Proteome Res. 4, 637–640

Li, K., Xu, C., Zhang, K., Yang, A., and Zhang, J. (2007) Proteome analysis of roots growth and metabolic changes under phosphate deficit in maize (Zea mays L.) plants. Proteomics 7, 1501–1512

Tian, Q., Wissuwa, M., Heidari, M., Nghi, N. M., Detz, K., Cahoon, D. A., and Bowers, J. K. (2009) Aramid (Aramid) resistant in response to abiotic stress. Plant Physiol. 149, 1490–1499

Chen, W., and Selbach, M. (2011) Global quantification of mammalian gene expression control. Nature 473, 337–342

Nakamura, Y., Arai, K., Masuda, T., Yoshio, Y., Takamaki, K., and Ohta, H. (2005) A novel phosphatidylcholine-activated phospholipase C induced by phosphate starvation in Arabidopsis. J. Biol. Chem. 280, 7469–7476

Plaxton, W. C., and Tran, H. T. (2011) Metabolic adaptations of phosphate-starved plants. Plant Physiol. 156, 1006–1015

Choi, J. H., Von Aung, K., Li, S. W., Xu, C., Chiang, S. F., and Su, C. L. (2006) Regulation of phosphate homeostasis by MicroRNA in Arabidopsis. Plant Cell 18, 412–421

Sivavand, S., Creff, A., Reymond, M., Sigoillot-Claude, C., Ricaud, L., Blanchet, A., Nussaume, L., and Desnos, T. (2007) Root tip contact with low-phosphate media reprograms plant root architecture. Nat. Genet. 39, 792–796

Estelle, M. A., and Somerville, C. (1987) Auxin-resistant mutants of Arabidopsis thaliana with an altered morphology. Mol. Gen. Genet. 206, 299–306

Cox, J., and Mann, M. (2008) MaxQuant enables high-accuracy identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat. Biotechnol. 26, 1367–1372

Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L., and Wold, B. (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat. Methods 5, 621–628

Kent, W. J. (2002) BLAT—the BLAST-like alignment tool. Genome Res. 12, 656–664

Clark, G., Konopka-Postupska, D., Hennig, J., and Roux, S. (2010) Is annexin 1 a multifunctional protein during stress responses? Plant Signal Behav. 5, 303–307

Bayle, V., Arighi, J. F., Creff, A., Nespoulous, C., Vilaret, J., Rossignol, M., Gonzalez, E., Paz-Ares, J., and Nussaume, L. (2011) Arabidopsis thaliana high-affinity phosphate transporters exhibit multiple levels of posttranslational regulation. Plant Cell 23, 1523–1535

Chen, J., Liu, Y., Ni, J., Wang, Y., Bai, Y., Shi, J., Gan, J., Wu, Z., and Wu, P. (2011) OsPHF1 regulates the plasma membrane localization of low- and high-affinity inorganic phosphate transporters and determines inorganic phosphate uptake and translocation in rice. Plant Physiol. 157, 269–276

Mahfouz, M. M., Kim, S., Delauney, A. J., and Verma, D. P. (2006) Arabidopsis TARGET OF RAPAMYCIN interacts with RAPTOR, which regulates the activity of S6 kinase in response to osmotic stress signals. Plant Cell 18, 477–490

Swaraz, A. M., Park, Y. D., and Hur, Y. (2011) Knock-out mutations of Arabidopsis SmD3-b induce pleotropic phenotypes through altered transcript splicing. Plant Sci. 180, 661–671

Shimoyama, M. (2011) Bio-synthesis and functions of the plant sulfolipid. Proc. Lipid Res. 50, 234–239

Nakamura, Y., Koizumi, R., Shui, G., Shimoyama, M., Wenk, M. R., Ito, T., and Ohta, H. (2009) Arabidopsis lipins mediate eukaryotic pathway of lipid metabolism and cope critically with phosphate starvation. Proc. Natl. Acad. Sci. U.S.A. 106, 20978–20983

May, A., Spinka, M., and Kick, K. (2012) Arabidopsis thaliana PECP1 - Enzymatic characterization and structural organization of the first plant phosphoethanolamine/phosphocholine phosphotransferase. Biochim. Biophys. Acta 1824, 319–325

Hammond, J. P., Broadley, M. R., and White, P. J. (2004) Genetic responses to phosphorus deficiency. Annu. Bot. 94, 323–332

Miyazawa, A., Nakanishi, K., and Ohtaki, Y. (2011) The Arabidopsis thaliana phosphate starvation responsive gene AtPPHase1 encodes a novel type of inorganic pyrophosphatase. Biochim. Biophys. Acta 1810, 178–185