Mol Genet Genomics (2016) 291:1595–1605
DOI 10.1007/s00438-016-1202-3

ORIGINAL ARTICLE

iTRAQ-based analysis of developmental dynamics in the soybean leaf proteome reveals pathways associated with leaf photosynthetic rate

Jun Qin[1,7] · Jianan Zhang[3] · Duan Liu[4] · Changcheng Yin[5] · Fengmin Wang[1] · Pengyin Chen[6] · Hao Chen[5] · Jinbing Ma[7] · Bo Zhang[8] · Jin Xu[2] · Mengchen Zhang[1]

Received: 5 December 2015 / Accepted: 15 March 2016 / Published online: 5 April 2016

Abstract Photosynthetic rate which acts as a vital limiting factor largely affects the potential of soybean production, especially during the senescence phase. However, the physiological and molecular mechanisms that underlying the change of photosynthetic rate during the developmental process of soybean leaves remain unclear. In this study, we compared the protein dynamics during the developmental process of leaves between the soybean cultivar *Hobbit* and the high-photosynthetic rate cultivar *JD 17* using the iTRAQ (isobaric tags for relative and absolute quantification) method. A total number of 1269 proteins were detected in the leaves of these two cultivars at three different developmental stages. These proteins were classified into nine expression patterns depending on the expression levels at different developmental stages, and the proteins in each pattern were also further classified into three large groups and 20 small groups depending on the protein functions. Only 3.05–6.53 % of the detected proteins presented a differential expression pattern between these two cultivars. Enrichment factor analysis indicated that proteins involved in photosynthesis composed an important category. The expressions of photosynthesis-related proteins were also further confirmed by western blotting. Together, our results suggested that the reduction in photosynthetic rate as well as chloroplast activity and composition during the developmental process was a highly regulated and complex process which involved a serial of proteins that function as potential candidates to be targeted by

Communicated by K. Chong.

Electronic supplementary material The online version of this article (doi:10.1007/s00438-016-1202-3) contains supplementary material, which is available to authorized users.

✉ Jin Xu
xujin@xtbg.ac.cn

✉ Mengchen Zhang
mengchenzhang@hotmail.com

[1] National Soybean Improvement Center Shijiazhuang Sub-Center, North China Key Laboratory of Biology and Genetic Improvement of Soybean Ministry of Agriculture, Cereal and Oil Crop Institute, Hebei Academy of Agricultural and Forestry Sciences, Shijiazhuang 050031, People's Republic of China

[2] Key Laboratory of Tropical Plant Resource and Sustainable Use, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Menglun, Mengla, Yunnan 666303, People's Republic of China

[3] National Foxtail Millet Improvement Center, Minor Cereal Crops Laboratory of Hebei Province Institute of Millet Crops, Hebei Academy of Agriculture and Forestry Sciences, Shijiazhuang 050035, People's Republic of China

[4] Geochemical Environmental Research Group, Texas A&M University, 833 Graham Road, College Station, TX 77845, USA

[5] Beijing Protein Innovation, B-8, Beijing Airport Industrial Zone, Beijing 101318, People's Republic of China

[6] Department of Crop, Soil and Environmental Sciences, University of Arkansas, Fayetteville, AR 72701, USA

[7] Department of Horticulture, University of Arkansas, Fayetteville, AR 72701, USA

[8] Department of Crop and Soil Environmental Sciences, Virginia Tech, Blacksburg, VA 24061, USA

biotechnological approaches for the improvement of photosynthetic rate and production.

Keywords Soybean · Proteome · iTRAQ · Photosynthesis

Introduction

Cultivated soybean [*Glycine max* (L.) Merr.] is a major global oil crop. In addition to providing vegetable oil and vegetable protein, it also contains many special nutrients, including isoflavone, saponin, and phytosterol. Soybean leaves are the main organs for photosynthesis and transpiration, and their developmental status has a great influence on soybean growth, resistance, yield and quality. Soon after attaining full expansion, the soybean leaves enter a senescence phase marked by a decline in photosynthetic rate and progressive loss of chloroplast activity and composition (Ford and Shibles 1988). Therefore, maintaining a stable photosynthetic rate and chloroplast activity as well as composition in soybean leaves during the senescence phase after full expansion are important factors that affect the potential of soybean production. Le et al. (2012) performed a genome-wide comparative transcriptome analysis of soybean leaf tissues at the late developmental stages under drought stress, and this study provided the basic foundation for further analysis of the functions of drought-responsive candidate genes that may ultimately lead to the development of drought-tolerant soybean cultivars. However, the relevant research on the dynamic development of soybean leaves has been sparse.

The transcriptional dynamics of C4 maize leaf have been well documented in previous studies. Wang et al. (2013) performed a genome-wide comparative analysis of developmental trajectories in Kranz (foliar leaf blade) and non-Kranz (husk leaf sheath) leaves of the C4 maize and the results identified cohorts of genes associated with the procambium initiation and vascular patterning, and provided a broader insight into the regulation of Kranz differentiation. Liu et al. (2013) conducted an anatomical and transcriptional dynamics analysis of maize embryonic leaves during seed germination and provided a foundation for forming hypotheses on the timing of regulatory events, particularly in the context of the Kranz structure development. Moreover, Li et al. (2010) analyzed the maize leaf transcriptome using mRNAseq technology and detected a dynamic transcriptome which regulating the transcripts for primary cell wall and basic cellular metabolism at the leaf base transitioning to transcripts for secondary cell wall biosynthesis and C4 photosynthetic development toward the tip. In addition, Manandhar-Shrestha et al. (2013) performed a comparative proteomics analysis of the chloroplast envelopes from the bundle sheath and mesophyll chloroplasts in Zea mays and described novel membrane proteins involved in C4-related metabolite fluxes and development.

Proteomics technologies are powerful tools for the analysis of protein function. Two-dimensional electrophoresis is one of the main methods used to identify differentially expressed proteins. Samples can be labeled and used for the difference gel electrophoresis (DIGE) methodology. The gel spots identified with this method can be cut out for mass spectrometry analysis, which increases the sensitivity and accuracy of protein identifications (Chen and Tang 2007; Salavati et al. 2012; Roy et al. 2011). Two-dimensional electrophoresis has been successfully applied for protein analysis of different organs and tissues of soybean, including seed (Mooney et al. 2004; Hajduch et al. 2005), leaf (Le et al. 2012), anther (Zeng et al. 2007) and nodule cytosol (Oehrle et al. 2008). A commonly used research strategy is to look for differentially expressed proteins among samples induced by certain factors and then identify and verify molecular mechanisms that are key factors to the process of development, physiological or pathological changes. Examples include differential protein expression analysis under stress, such as drought (Roy et al. 2011), flooding (Komatsu et al. 2010), solar UV-B radiation (Xu et al. 2008), high temperature and humidity (Ren et al. 2009; Wang et al. 2012), contaminated environment (Danchenko et al. 2009), photoperiod treatments (Hao et al. 2009) and resistance to *phytophthora* root rot (Qiu et al. 2009).

However, two-dimensional electrophoresis has its own disadvantages due to the limited numbers of samples and proteins that can be identified within one sample as well as the insensitivity to small proteins and the repeatability. The proteome is distinct in different cells and tissues, and the expression and modifications of proteins are affected by developmental stages and environmental factors. The proteome is also dynamic rather than static. Therefore, it is important to study proteomes during development stage. With progress in proteomics technologies, the scale and focus of proteomics have shifted over time. In the early years, the goal was often to identify a single protein, in contrast to modern high-throughput methods that aim to measure dynamic proteomes, particularly in parallel and in a quantitative manner. The isobaric tags for relative and absolute quantification (iTRAQ) is an advanced and mature technique for protein quantization, and can be used for high-throughput proteomics analysis at different development stages (Ross et al. 2004; Adav et al. 2011; Lan et al. 2011; Qin et al. 2013).

The objective of this study was to compare the differential expressed proteins during the developmental process of leaves between the Chinese elite cultivar *JD 17* and the American elite cultivar *Hobbit* using the iTRAQ method. Our results indicated that photosynthesis-related proteins

exhibited significant differential expression levels at three developmental stages of soybean seedlings. The possible mechanisms for the role of these differential expressed proteins in soybean leaves during the developmental process were further discussed.

Materials and methods

Plant materials and treatment conditions

JD 17 is an elite soybean variety in Huang-Huai Hai. In national regional trials, *JD 17* was the highest yielding soybean variety in the history of the national soybean regional trials which has achieved an average yield greater than 3750 kg per Ha for three consecutive years. During the regional trial in Hebei Province, *JD 17* achieved a 33.11 % yield increase in the comparison with control varieties which was the largest yield increase in the history of the regional trials. In addition, high-yield experts have monitored the yield and found that *JD 17* produces a yield of 4510 kg per Ha which exceeds the national standards for super high-yielding varieties. *Hobbit,* which is imported from the United States, is an elite soybean variety with high yield and high oil content. It is also a dedicated ('Special') variety for the SSS Planting pattern. *JD 17* and *Hobbit* showed similar developing habit, including flowering time and maturity.

Plants were grown in a mixture of 50 % vermiculite and 50 % soil in a greenhouse at Shijiazhuang (114°26″E, 38°03″N). When a fully developed trifoliolate leaf node displayed unrolled leaflets, the first fully developed trifoliolate leaves were collected on 14, 28, and 42 days after sowing, respectively. All the leaves were collected from even growth plants at the same knot. Leaf samples were freezed in liquid nitrogen and the frozen leaves were stored at −80 °C.

Protein extraction

Total protein was extracted with cold acetone procedure. Leaves were grounded in 10 % sample volume of polyvinylpolypyrrolidone and chilled with liquid nitrogen to break cells. Total proteins were precipitated by 10 % (v/v) trichloroacetic acid (TCA) in acetone at −20 °C. Then samples were incubated at −20 °C for 2 h after thorough mixing. Proteins were collected by centrifugation at 30,000*g* at 4 °C for 30 min. The protein pellets were washed 3 times with cold acetone to remove TCA. Each protein pellet was re-suspended in 1 mL protein extraction reagent with sonication which contained 8 M urea, 4 % (w/v) CHAPS, 30 mM HEPES, 1 mM PMSF, 2 mM EDTA and 10 mM DTT. The supernatant was collected after centrifugation at 20,000*g* at 4 °C for 30 min. The protein concentration was determined with a 2-D Quant Kit (General Electric Company, USA). The protein quality and concentration were further determined and verified in SDS-PAGE.

Protein digestion and iTRAQ labeling

For trypsin digestion, 100 μL of total protein was mixed by equal volume of tetraethylammonium bicarbonate (TEAB, pH 8.5; Sigma, St. Louis, MO, USA), and modified trypsin (Promega, Madison, WI, USA) was added to the mixture (3.3 μg trypsin/100 μg protein). The protein was digested at 37 °C for 24 h, and the solvent was removed by speed-vac. The efficiency of protein digestion was verified with MALDI TOF/TOF (Bruker Limited, Coventry, UK) mass spectrometry.

Peptides were labeled with iTRAQ labeling kit (Applied Biosystems, Foster City, USA) following the manufacture's manual. 2-, 4-, and 6-week samples were labeled with reagent 114, 115, and 116, respectively. The peptides were further fractionated offline in the high performance liquid chromatography (HLPC) system (Shimadzu, Japan) with a strong cation exchange column (SCX column, Luna 5 μm column, 4.6 mm I.D. × 250 mm, 5 μm, 100 Å; Phenomenex, Torrence, CA, USA). The retained peptides were eluted with Buffer A (10 mM KH_2PO_4 in 25 % ACN, pH 3.0) and Buffer B (2 M KCl, 10 mM KH_2PO_4 in 25 % ACN, pH 3.0) with flow rate at 1 mL/min. In total, 38 fractions were collected and combined into 17 fractions to reduce peptide complexity, according to protein properties. Subsequently, eluted fractions were lyophilized in a centrifugal speed vacuum concentrator and dissolved with 0.1 % formic acid prior to reversed-phase nanoflow liquid chromatography (nLC) tandem mass spectrometry (nLC-MS/MS).

MS/MS analysis

MS/MS analysis was performed on the Proxeon Easy Nano-LC system connected to a hybrid gradrupole/time-of-flight mass spectrometer (MircoTOF-Q II, Bruker, Billerica,USA) equipped with nano-electrospray ion source. Peptides of each fraction were equalized, so that the same amount of protein from each fraction was mixed and then injected into the Nano-LC system. Peptides were separated on C18 analytical reverse phase column with mixtures of Solution A and Solution B at a flow rate of 300 nL/min (Solution A: 5 % acetonitrile, 0.1 % formic acid; Solution B: 95 % acetonitrile, 0.1 % formic acid). After equilibrating with 5 % Solution B for 10 min, the ratio of Solution B increased to 45 % in 80 min, and further increased to 80 % in 5 min. The ratio of Solution B was maintained at 80 % for 15 min, and decreased to 5 % instantly and

held for 15 min. The MicroTOF-Q II hybrid MS was used to analyze the fractions. The MS/MS survey scan was obtained in the mass-to-charge ratio (m/z) range 50–2000. The ionization tip voltage and interface temperature were run as 1250 V and 150 °C.

Protein identification and quantification

All of the MS data were collected using the software micrOTOF control (Bruker, Billerica, USA) and analyzed using DataAnalysis (Bruker Daltonics, Billerica, USA). The Mascot v2.3.01 (Matrix Science, Boston, USA) was used to identify proteins. The Universal Protein Resource (UniProt) database was accessed as the reference protein database. Data collected from three biological repeats and their fractions were combined for protein search. The parameters were set as follows: specifying trypsin as the digestion enzyme; cysteine carbamidomethylation as fixed modification; iTRAQ8-Plex on N-terminal residue, iTRAQ-8Plex on tyrosine (Y), iTRAQ-8Plex on lysine (K), glutamine-pyroglutamic acid and oxidation on methionine (M) as the variable modification. The peptide tolerance was set to 0.05 Da and MS/MS tolerance was set to 0.05 Da. At last, an additional filter before exporting the data was set as following: significance threshold $P < 0.05$ (with 95 % confidence) and ion score or expected cut-off less than 0.05 (with 95 % confidence). The protein search results were exported by Mascot then normalized and quantified using Scaffold v3.0. The 28- and 42-day protein profiles were quantified base on 14-day results.

Protein annotation

Based on the results of proteins search against UniProt database, DAVID Bioinformatics Resources v6.7 (Huang et al. 2009a, b) was employed to obtain the functional classification of the proteins based on Gene Ontology (GO) terms (Harris et al. 2004) and Clusters of Orthologous Groups (COG) terms (Natale et al. 2000), to find out the protein distribution in each functional group. The Kyoto Encyclopedia of Genes and Genomes (KEGG) database (http://www.genome.jp/kegg/) (Kanehisa and Goto 2009; Ogata et al. 1999) was used to annotate the pathway of these proteins. The enrichment test of functional category proteins was performed by Chi-square test with defining cut-off as 0.01 under Arabidopsis thaliana background. False discovery rate (FDR) significance threshold of 0.05 was used as false-positive control.

Cluster analysis

Proteins with significant expression level changes during three developmental stages of soybean seeding were identified with previously described methods (Cox and Mann 2008; Lan et al. 2011). The 90 % confidence (Z score = 1.645) of log 2 ratios was used to select the proteins whose expression levels were two fold higher or 50 % lower. The mean and SD of the protein expression levels were calculated for the biological repeats. The profiles from different time points were calculated independently and a broader threshold was selected. The results showed that the cut-off value for the up-regulated protein was 1.75-fold (mean + Z score × SD) and the cut-off value for the down-regulated protein was 0.69-fold (mean − Z score × SD). Protein ratios outside this range were defined as being significantly different at $P = 0.1$. The expression profiles of the differentially expressed proteins were determined by cluster analysis based on the k-means method with Pearson's correlation distance in Genesis v1.7.6 (Sturn et al. 2002; Zenoni et al. 2010). Number of clusters was determined by the particular algorithm named Figures of Merit (FOM) (Yeung et al. 2001).

Western blot analysis

The same samples used in the iTRAQ analysis were also used for the western blot analysis which included five differentially expressed proteins. Equal amounts of protein samples from *JD 17* and *Hobbit* were separated by SDS-PAGE and electro-transferred to a PVDF membrane (Millipore Corporation, Bedford, MA, USA) at 100 V for 60 min. The membrane was then immersed in 5 % non-fat milk in TTBS solution (0.2 M Tris–HCl (pH 7.6), 1.37 M NaCl, and 0.1 % Tween-20) for 1 h at room temperature. The proteins were incubated with the corresponding polyclonal antibodies in 5 % non-fat milk in TTBS solution for 3 h at room temperature and then subjected to three 5 min rinses in TTBS solution. The membrane was further incubated with horseradish peroxidase-conjugated goat anti-rabbit antibodies (Beijing Protein Innovation, Beijing, China) for 1 h at room temperature and subjected to three 5 min rinses in TTBS solution. The blot was then developed with the SuperECL Plus kit (Applygen, Beijing, China) and exposed on X-ray film.

Results

Photosynthesis efficiency analysis in leaf-dynamic development

The photosynthesis efficiency was first tested in *JD 17* and *Hobbit*. As shown in Fig. 1, *JD 17* had the higher photosynthetic rate (P_n) than *Hobbit*. The stomata are important limitation factors for photosynthesis efficiency. To investigate whether the difference of photosynthesis

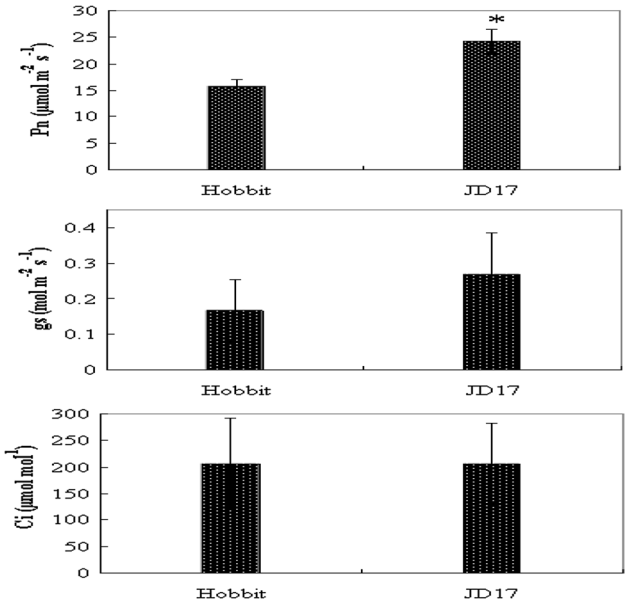

Fig. 1 Photosynthetic performance in the two soybean cultivars *JD 17* and *Hobbit* at three contiguous developmental stages

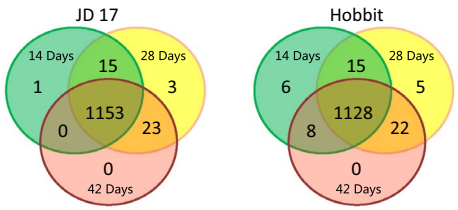

Fig. 2 Venn diagram showing the expressed proteins identified in *JD 17* and *Hobbit* at three contiguous developmental stages

efficiency between these two cultivars was due to stomatal constraint, we measured intercellular CO_2 concentration (C_i) and stomatal conductance (g_s). In contrast to P_n, the two cultivars showed the similar levels of C_i and g_s. These data suggested that the differential level of photosynthesis efficiency between the two cultivars did not result from stomatal conductance and intercellular CO_2 concentration.

Proteome analysis in leaf-dynamic development

To obtain a better understanding of the molecular mechanisms of developmental stages of soybean seedings, the proteome changes in the *Hobbit* and *JD 17* leaves were examined at three contiguous developmental stages and the number of proteins in different development stages and their functional and pathway distributions were analyzed as well. The Venn diagram in Fig. 2 showed the number of leaf proteins from *Hobbit* and *JD 17* in the three developmental stages. A total of 1269 protein were detected in 6 samples in which 87.25 % of protein identities shared in three biological repeats. There were 1184 proteins detected in the *Hobbit* leaves at the 14-, 28-, and 42-day developmental stages and 96.49 % of the protein species detected in all three stages. In *JD 17*, there were 1195 proteins detected at the three stages, with 95.27 % of the proteins commonly detected in all three stages. These results indicated that the detected number of proteins and their identities were unchanged across three contiguous developmental stages.

To study the process of soybean seeding development, we examined the functional enrichment of the detected proteins. There were 1083 annotated proteins detected, which were classified into 36 small groups according to biological process, cellular component and molecular function (Fig. 3). The results suggested that the proteins were distributed among all function groups, with enrichment in some specific groups. The large group of biological process included 17 small groups. A total of 32.73 % of the proteins belonged to this biological process group and the identified proteins were further enriched in 4 small groups: cellular component biogenesis, cellular process, metabolic process, and response to stimulus. Another large group, cellular component, which was broken down into nine small groups, included 48.78 % of the identified proteins. Among those protein, the protein species were significantly enriched in six of the small groups: envelope, extracellular region, macromolecular complex, membrane-enclosed lumen, organelle, and organelle part. Furthermore, 18.50 % of the proteins belonged to another large group of molecular functions, which was consisted of ten small groups. The proteins in molecular functions group were enriched in 4 of the small groups: antioxidant activity, catalytic activity, electron carrier activity, and structural molecule activities. In addition, 1083 proteins were found distributed among 103 pathways in which 19.10 % belonged to the metabolic pathways, 12.12 % to the biosynthesis of secondary metabolites, and 6.95 % to ribosomes.

Protein expression pattern analysis in soybean leaf development

In the previous section, we discussed the number of identified proteins and their classification and functional distribution during developmental stages of soybean seeding. In this section, we discuss the magnitude of protein expression and protein expression patterns. The cluster analysis was performed with the software Genesis v1.7.6 (Sturn et al. 2002; Zenoni et al. 2010) and the results indicated that the protein expression patterns were different among the developmental stages (Supplemental Fig. 1A). Based on the protein levels, the protein expression profiles were

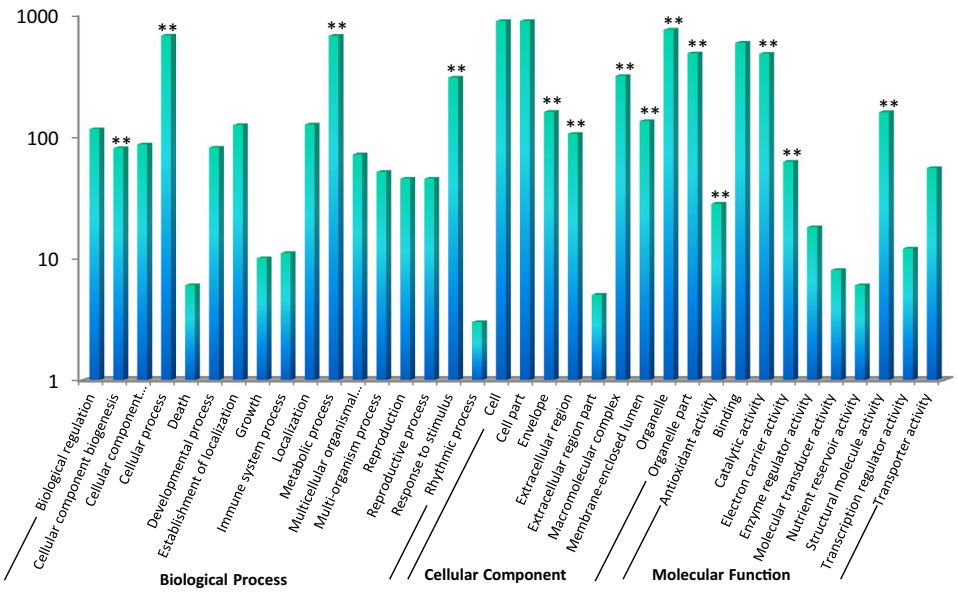

Fig. 3 Functional categorizations and enrich distribution of proteins that were detected in *JD 17* and *Hobbit* at three contiguous developmental stages. The *stars* indicate the enrich functional categorizations

grouped into nine patterns (Supplemental Fig. 1B). which were observed in both *Hobbit* and *JD 17* in all three developmental stages.

Cluster 2 and Cluster 7 belonged to a continuous growing pattern, in which protein abundance increased continuously between 14 and 42 days. Cluster 9 was a continuous decreasing pattern, in which protein abundance decreased continuously during this period. Cluster 3 was a stable pattern, in which the protein abundances were stable between 14 and 42 days. Cluster 1 and Cluster 8 showed to a valley pattern, in which 28 days had the lowest protein abundance. Clusters 4–6 had a hill pattern, in which 28 days had the highest protein abundance. *Hobbit* possessed 2–3 times more proteins in Clusters 1, 5, and 7 than *JD 17*. The 5.23 % of detected proteins of *Hobbit* were grouped in Cluster 1, while 15.96 % of proteins for *JD 17* showed the expression pattern. *Hobbit* had only 13.03 % of protein species in Cluster 5, while *JD 17* exhibited as great as 31.92 %. The percentages of detected protein species in Cluster 7 were 20.04 and 6.18 % for *Hobbit* and JD 17, respectively. The numbers of proteins that belonged to the Cluster 1, 5 and 7 for *Hobbit* and *JD 17* were 125, 125, 113 and 264, 264, 224, respectively, and the main differences presented in three small groups within the cellular component category: cell, cell part, and organelle. In addition to the number of proteins, the protein species were different between *Hobbit* and *JD 17*. Only 45 proteins from the 460 identified proteins were in common among the three small groups for *Hobbit* and *JD 17*. In Cluster 7, *Hobbit* had 182, 182, and 148 proteins in those three small groups, respectively, while *JD 17* showed 57, 57, and 45 proteins; only 7 proteins were in common among 286 detected proteins. In Clusters 8 and 2, *Hobbit* and *JD 17* had a similar number of identified proteins, but the number of common proteins was low.

To follow the discussion in the last section of the nine protein expression patterns observed in *Hobbit* and *JD 17* leaves and the protein identities and number in each of the nine expression patterns, the functions of identified proteins will be discussed next (Supplemental Table 1). As discussed above, the proteins were classified into three large categories and 20 small functional groups. The protein distributions were stable in the above functional groups among different developmental stages of soybean seeding. However, the distributions were very different between *Hobbit* and *JD 17*. In the category of biological processes, the proteins were enriched in the small functional groups of metabolic and cellular processes (Supplemental Fig. 1C). In the category of cellular components, the proteins were enriched in the small functional groups of organelles and organelle parts. In the category of molecular functions, protein species were enriched in the small functional groups of binding and catalytic activities.

Leaf proteome comparison between *Hobbit* and *JD17* in the same developmental stage

Previous studies shown that the protein groups and numbers were stable across three developmental stages for both cultivars. In this section, the differences between the cultivars will be compared. The differences in protein species and the magnitude of expression will be analyzed. In this study, a differential protein is defined as a protein with an expression level higher than double the control or lower than half the control. Differential proteins in *Hobbit* were

defined against *JD 17* as a control. A total of 1240 proteins were detected in 14-day leaf proteomes in which 1159 proteins (93.47 %) were in common for both *Hobbit* and *JD 17* (Fig. 4). There were 141 differential proteins (6.53 %) which included 84 up-regulated proteins and 57 down-regulated proteins and the numbers of proteins assigned to biological process category, cell component and molecular function were 40, 89, and 12, respectively. These 141 proteins were distributed in 34 pathways, with the Ribosomal assembly pathway including most differential proteins (25), followed by 4 photosynthesis-related proteins in the 'carbon fixation in photosynthetic organisms' pathway. For 28-day leaf proteomes, 1244 proteins were detected, with 1206 proteins (96.95 %) in common between *Hobbit* and *JD 17* (Fig. 4). There were 540 differential proteins (3.05 %), which included 221 up-regulated and 319 down-regulated proteins and the numbers of proteins belonged to biological process category, cellular component and molecular function were 158, 324 and 58, respectively. The differential proteins were distributed among 15 pathways. There were 51 differential proteins in the Ribosomal assembly pathway, which contained the most differential proteins of any other pathway. The Photosynthetic pathway was second most enriched, with 20 differential proteins. For 42-day leaf proteomes, 1240 proteins were detected, with 1182 proteins (95.32 %) in common between *Hobbit* and *JD 17* (Fig. 4). There were 126 differential proteins (4.68 %) which included 98 up-regulated and 28 down-regulated proteins. Among 126 differential proteins, 39 proteins were assigned to the biological process category and the rest proteins were grouped to cellular component category (78) and molecular function category (9), respectively.

The differential proteins were distributed among 23 pathways. The ribosomal assembly pathway was most populated with differential proteins (13), followed by the photosynthetic pathway, with two differential proteins. The remaining 21 pathways included one differential protein for each pathway.

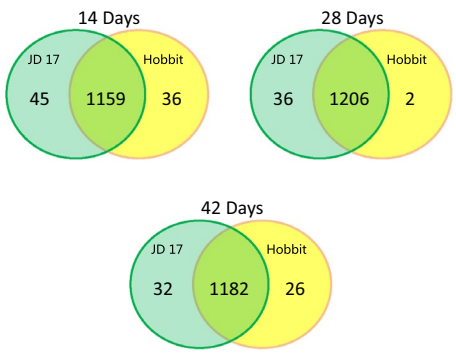

Fig. 4 Leaf proteome comparison between *Hobbit* and *JD 17* in the same developmental stage

Proteins detected in both *Hobbit* and *JD 17*

During developmental stages of soybean seeding, 584 shared proteins for *Hobbit* and *JD 17* were detected in which 290 proteins were expressed consistently higher in Hobbit or consistently lower than *JD 17* during the 14- to 42-day period. Among these 290 proteins, 239 had functional annotations in 132 pathways and there were 176 up-regulated proteins and 114 down-regulated proteins. There were 92 out of the 290 proteins in *Hobbit* showed differential expression compared to *JD 17 and* the functional analysis revealed that these 92 proteins were annotated in 23 small groups according to their biological processes, cellular components and molecular functions (Fig. 5). These 92 proteins were enriched in 13 out of 23 groups and distributed among 33 pathways, with 9 enriched in the photosynthesis pathway. The expression levels of these proteins will be discussed further in the following sections.

Photosynthetic protein expression patterns and their immunological verifications

Photosynthesis is critical for soybean yield; hence, we investigated the abundance of enzymes involved in photosynthesis. From 14 to 42 days, the amount of CAB2 increased by 32 % in *JD 17* and 15 % in *Hobbit*. In addition, three homologues of OEE1 (oxygen-evolving enhancer protein 1) were detected including C6T7N2, C6TC92, and C6TKL8. In *Hobbit*, the amount of C6T7N2 and C6TC92 increased by 23 and 15 %, respectively. From 14 to 42 days, the amount of C6TKL8 increased by 15 % for *JD 17* and by 7 % for *Hobbit*. Furthermore, ATPsB (ATP synthase beta subunit, C6T308) which is involved in photophosphorylation has increased by 23 % between 14 and 42 days in *JD 17* but decreased by 24 % in *Hobbit* (Fig. 6). These results demonstrated that from 14 to 42 days, two proteins (CAB2 and OEE1) with an important role in the photosynthesis primary reactions increased in abundance more in *JD 17* than in *Hobbit*. During the same period of time, the abundance of the photophosphorylation protein ATPsB increased in *JD 17* but decreased in *Hobbit* (Fig. 6). These results corroborated our previous physiological study that showed *JD 17* exhibiting higher photosynthetic efficiency. In addition, RCA (ribulose bisphosphate carboxylase/oxygenase activase), a CO2 fixation related protein, and two RCA homologues (D4N5G0 and D4N5G2) in the Calvin cycle were also detected in this study. From 14 to 42 days, the abundance of D4N5G0 dropped by 38 and 7 % in *JD 17* and *Hobbit*, respectively. During the same period of time, the abundance of D4N5G2 dropped by 19 % in *JD 17* but increased by 7 % in *Hobbit*. Furthermore, two GAPDH (glyceraldehyde-3-phosphate dehydrogenase) homologues, Q38IX0 and C6T9R8, were also detected during the 14- to

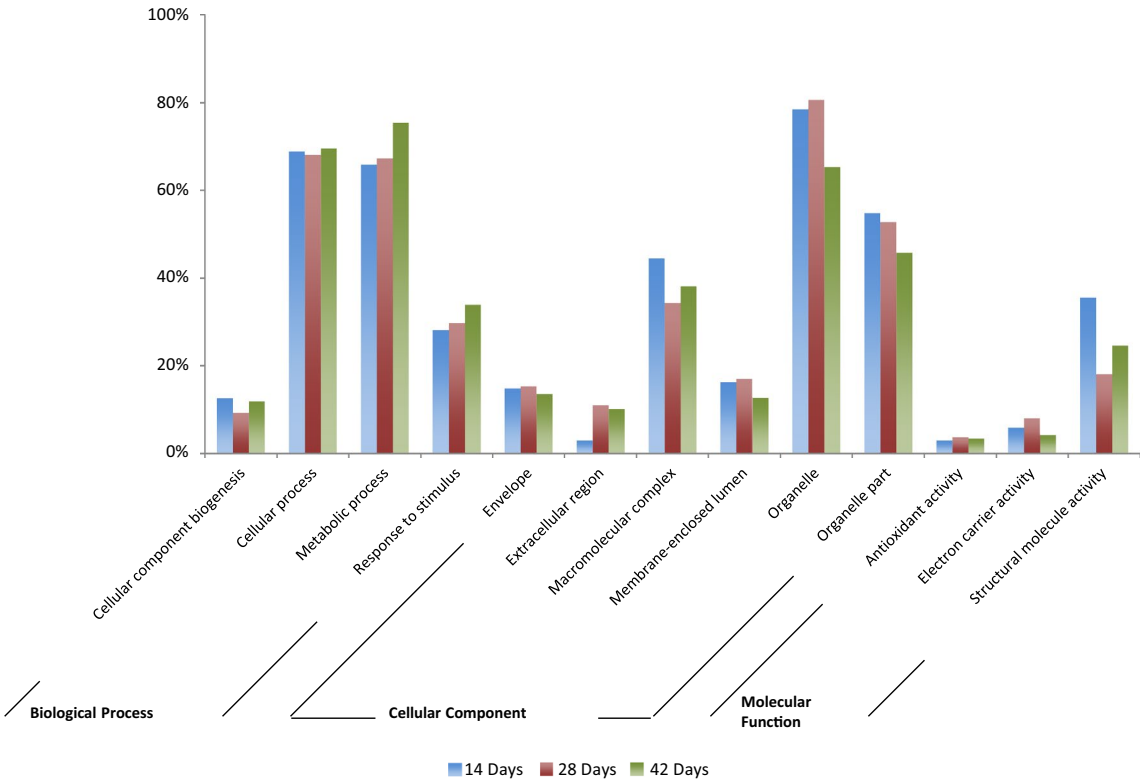

Fig. 5 Functional categorisation of the proteins that were differentially expressed between Jidou17 and *Hobbit* at three contiguous developmental stages

42-day period. The Q38IX0 abundance dropped by 13 and 38 % in *JD 17* and *Hobbit*, respectively. Similar decreasing was found for C6T9R8 in *JD 17* and *Hobbit* which were 7 and 34 %, respectively. In addition, the abundance of SBPase (sedoheptulose-1,7-bisphosphatase, C6TK14) dropped by 19 and 29 % in *JD 17* and *Hobbit*, respectively (Fig. 6). For the enzyme 2-Cysp (2-cysteine peroxiredoxins, C6TGM9), which is involved in the removal of reactive oxygen species (ROS), its abundance did not change from 14 to 42 days in *JD 17*, but it decreased by 19 % in *Hobbit*. Thioredoxin Trx (C6TLX7) has many functions, including being involved in the scavenging of ROS. The abundance of C6TLX7 increased by 24 % in *JD 17* but decreased by 34 % in *Hobbit* from 14 to 42 days (Fig. 6). These five proteins are important enzymes in the Calvin cycle and in scavenging ROS. Our results showed that the abundances of these five proteins decreased from 14 to 42 days (Fig. 6) but the decreasing trend was more pronounced in *Hobbit* than in *JD 17*.

For further confirmation, the protein expression patterns of CAB-2, RCA, SBPase, and Trx were identified by western blotting. As shown in Fig. 7, this pattern of expression was consistent with our iTRAQ results and thereby verified the reliability of the iTRAQ system and the data analysis methods.

Discussion

JD 17 has a higher photosynthetic rate (Pn) than *Hobbit*. In this study, we first demonstrated that the differential level of photosynthesis efficiency between *JD 17* and *Hobbit* did not result from stomatal conductance and intercellular CO_2 concentration. Then, we examined the proteomes during soybean leaf development in two cultivars using the high-throughput iTRAQ technique. We analyzed the proteomes in different leaf developmental stages and identified the protein species, numbers and the levels of expression. We performed a comparative analysis of the dynamic patterns of protein expression during development and some protein expression patterns were found to be similar in different cultivars. We also analyzed proteomes at the same developmental stages but from different cultivars and identified the proteins commonly detected in both *Hobbit* and *JD 17*. We found that the protein groups and number of proteins were stable across three developmental stages as well as in both cultivars. Our study indicated that photosynthesis-related proteins had significant differential expression levels at different stages of leaf development, and the reduction in photosynthetic rate and chloroplast activity as well as composition during the developmental process were highly regulated and complex processes that involves potential candidate proteins that will

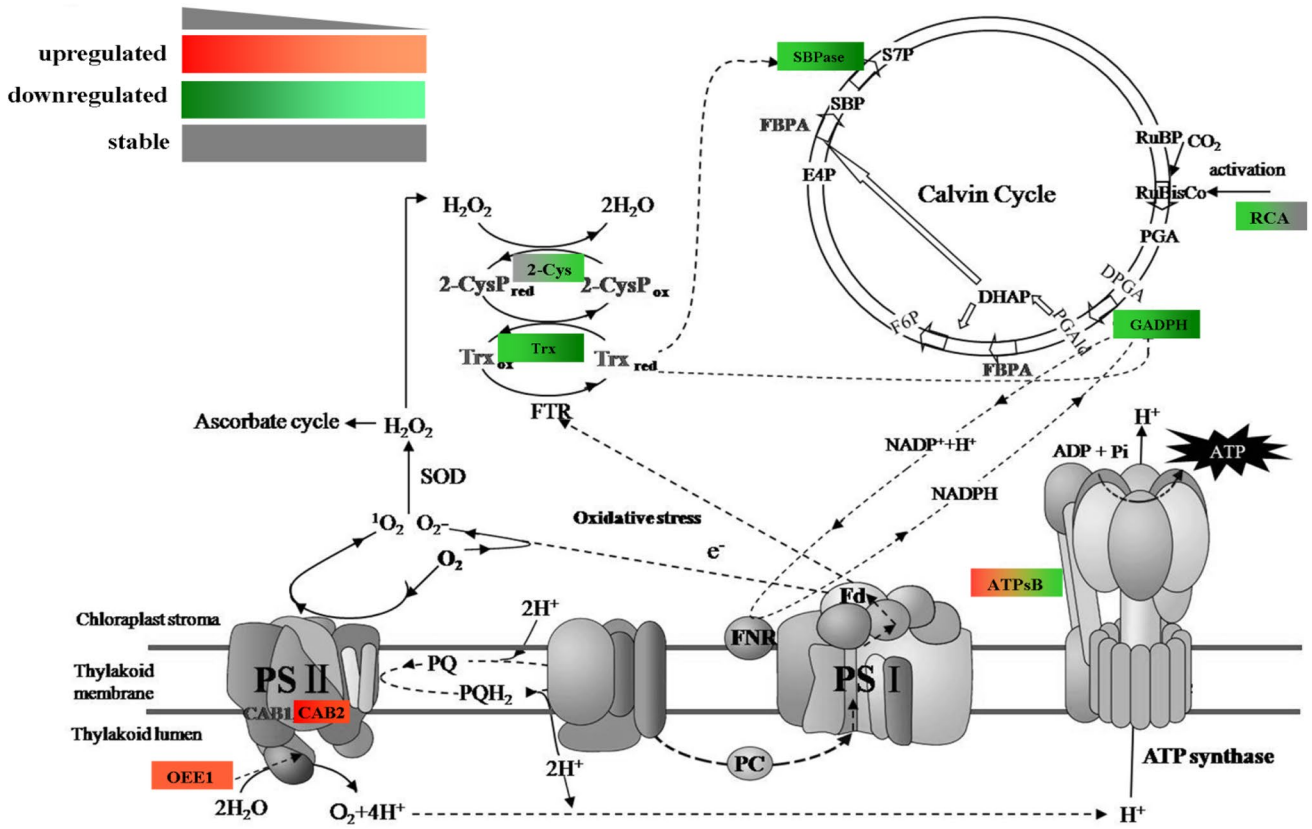

Fig. 6 The changes of the abundance of enzymes involved in photosynthesis efficiency in soybean leaf during the developmental stages from 14 to 42 days. The abundance of enzymes are indicated in *colorful rectangular boxes*, with *red* for up-regulated, *green* for down-regulated, and *gray* for stable expression of proteins. The *left* part of the *box* represents the changes of the abundance of proteins in *JD 17* from 14 to 42 days, and the *right* part of the *box* represents the changes of the abundance of proteins in *Hobbit* form 14 to 42 days (color figure online)

be targeted by biotechnological approaches to augment the aforementioned reduction. Such a study will be important for revealing the molecular mechanism of soybean leaf development and provide a better understanding of plant growth and development. In addition, this study may provide new clues for enhancing photosynthetic efficiency and developing new soybean breeding cultivars.

Light reactions of photosynthesis take place on the grana lamella membrane in the chloroplast. The light reaction system consists of photosystem I (PS I), photosystem II (PS II), cytochrome b6f complex, and ATP synthase (Knaff and Arnon 1971). The dark reactions of photosynthesis take place in the chloroplast stroma. In the dark reactions, CO_2 is fixed in the Calvin cycle, which has the following three phases: carbon fixation, reduction, and regeneration. In the Calvin cycle, 11 enzymes catalyze 13 continuous reactions (Raines 2003). CAB2 (C6T8R7, chlorophyll *a–b* binding protein 2) is involved in the primary reaction of photosynthesis. In the present study, we observed a decrease from 14 to 42 days in the abundance of the enzymes that are involved in the Calvin cycle and the scavenging of ROS, and this reduction could be one of the reasons for diminished photosynthetic efficiency in soybean leaves during this period. On the other hand, the leaves maintained high abundance of the enzymes involved in the primary reactions of photosynthesis and this high abundance could be one of the important factors that maintain steady photosynthetic efficiency.

The application of proteomics methods to global expression analyses and protein identification has been highly efficient in plant research in recent years. In this study, we performed a comparative analysis of the dynamic patterns of protein expression during development. Our results showed that the photosynthetic rate and chloroplast activity and composition during the developmental process are highly regulated and complex processes. In future studies, we will focus more on these proteins with the goal of adjusting their levels using biotechnology methods to enhance photosynthetic efficiency and increase soybean yield.

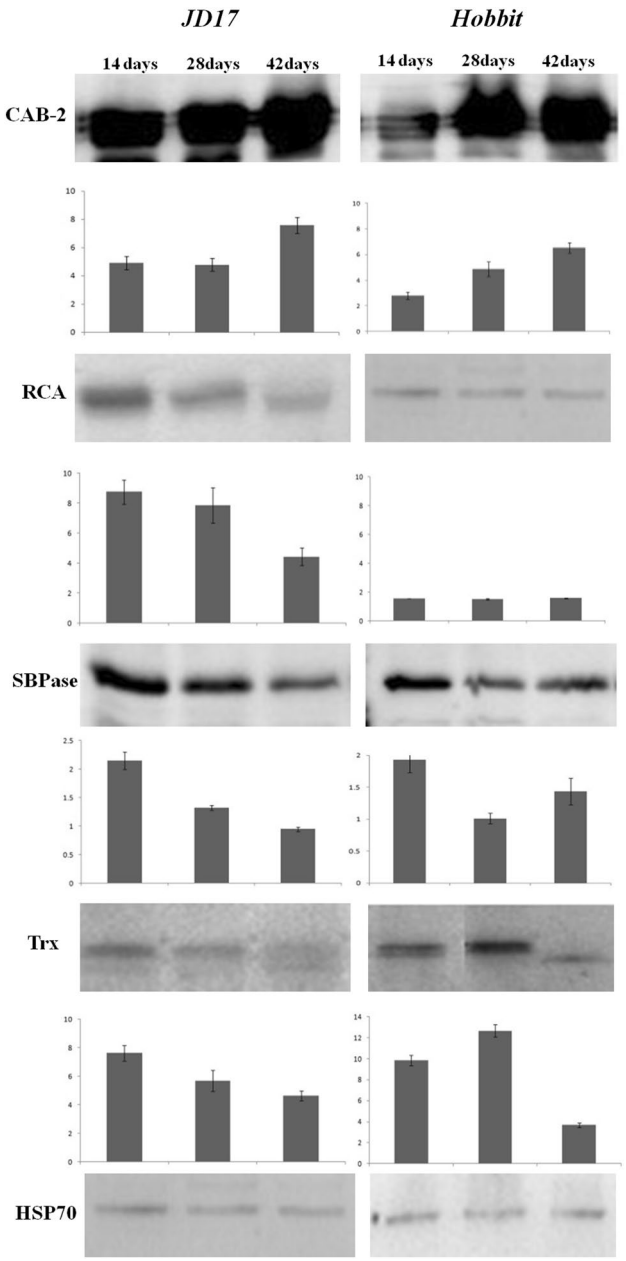

Fig. 7 Western blotting detection of soybean proteins CAB-2 RCA SBPase Trx HSP70

Acknowledgments This study was supported by the Natural science fund for distinguished young scholars of Hebei Province (C2014301035), National Natural Science Foundation of China (31100880), the Key Project of the Natural Science Foundation of Hebei Province (C2012301020), the Key Research Foundation for Excellent Returned Overseas Chinese (C2011006001), and China Scholarship Council Program (201208130236).

Compliance with ethical standards

Conflict of interest The authors declare that they have no competing interests.

Ethical approval This article does not contain any studies with human participants performed by any of the authors.

References

Adav SS, Ng CS, Sze SK (2011) iTRAQ-based quantitative proteomic analysis of *Thermobifida fusca* reveals metabolic pathways of cellulose utilization. J Proteomics 74:2112–2122

Chen XY, Tang ZC (2007) Plant physiology and molecular biology. Higher Education Press, 3rd edn. China, Beijing, pp 50–64

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

Danchenko M, Skultety L, Rashydov NM, Berezhna W, Mátel L, Salaj T, Pret'ová A, Hajduch M (2009) Proteomic analysis of mature soybean seeds from the Chernobyl area suggests plant adaptation to the contaminated environment. J Proteome Res 8:2915–2922

Ford DM, Shibles R (1988) Photosynthesis and other traits in relation to chloroplast number during soybean leaf senescence. Plant Physiol 86:108–111

Hajduch M, Ganapathy A, Stein JW, Thelen JJ (2005) A systematic proteomic study of seed filling in soybean establishment of high-resolution two-dimension al reference maps, expression profiles, and an interactive proteome database. Plant Physiol 137:1397–1419

Hao W, Jiang ZF, Zhao L, Han YP, Li WB (2009) A comparative study on soybean leaf proteomics under different photoperiod treatments. Soybean Science 28:388–393

Harris MA, Clark J, Ireland A, Lomax J, Ashburner M, Foulger R, Eilbeck K, Lewis S, Marshall B, Mungall C, Richter J, Rubin GM, Blake JA, Bult C, Dolan M, Drabkin H, Eppig JT, Hill DP, Ni L, Ringwald M, Balakrishnan R, Cherry JM, Christie KR, Costanzo MC, Dwight SS, Engel S, Fisk DG, Hirschman JE, Hong EL, Nash RS, Sethuraman A, Theesfeld CL, Botstein D, Dolinski K, Feierbach B, Berardini T, Mundodi S, Rhee SY, Apweiler R, Barrell D, Camon E, Dimmer E, Lee V, Chisholm R, Gaudet P, Kibbe W, Kishore R, Schwarz EM, Sternberg P, Gwinn M, Hannick L, Wortman J, Berriman M, Wood V, de la Cruz N, Tonellato P, Jaiswal P, Seigfried T, White R (2004) Gene Ontology Consortium The Gene Ontology (GO) database and informatics resource. Nucleic Acids Res 32:258–261

Huang daW, Sherman BT, Lempicki RA (2009a) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res 37:1–13

Huang daW, Sherman BT, Lempicki RA (2009b) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 4:44–57

Kanehisa M, Goto S (2009) KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Res 28:27–30

Knaff DB, Arnon DI (1971) On two photoreactions in system II of plant photosynthesis. Biochim Biophys Acta 2:400–408

Komatsu S, Kobayashi Y, Nishizawa K, Nanjo Y, Furukawa K (2010) Comparative proteomics analysis of differentially expressed proteins in soybean cell wall during flooding stress. Amino Acids 39:1435–1449

Lan P, Li WF, Wen TN, Shiau JY, Wu YC, Li W, Schmidt W (2011) iTRAQ protein profile analysis of arabidopsis roots reveals new aspects critical for iron homeostasis. Plant Physiol 155:821–834

Le DT, Nishiyama R, Watanabe Y, Tanaka M, Seki M, Ham le H, Yamaguchi-Shinozaki K, Shinozaki K, Tran LS (2012) Differential gene expression in soybean leaf tissues at late developmental stages under drought stress revealed by genome-wide transcriptome analysis. PLoS One. doi:10.1371/journalpone0049522

Li PH, Ponnala L, Gandotra N, Wang L, Si YQ, Tausta SL, Kebrom TH, Provart N, Patel R, Myers CR, Reidel EJ, Turgeon R, Liu P, Sun Q, Nelson T, Brutnell TP (2010) The developmental dynamics of the maize leaf transcriptome. Nat Genet 42:1060–1067

Liu WY, Chang YM, Chen SC, Lu CH, Wu YH, Lu MY, Chen DR, Shih AC, Sheue CR, Huang HC, Yu CP, Lin HH, Shiu SH, Ku MS, Li WH (2013) Anatomical and transcriptional dynamics of maize embryonic leaves during seed germination. Proc Natl Acad Sci USA. doi:10.1073/pnas1301009110

Manandhar-Shrestha K, Tamot B, Pratt EP, Saitie S, Bräutigam A, Weber AP, Hoffmann-Benning S (2013) Comparative proteomics of chloroplasts envelopes from bundle sheath and mesophyll chloroplasts reveals novel membrane proteins with a possible role in c4-related metabolite fluxes and development. Front Plant Sci. doi:10.1111/tpj12229

Mooney BP, Krishnan HB, Thelen JJ (2004) High-throughput peptide mass finger printing of soybean seed proteins: automated workflow and utility of UniGene expressed sequence tag databases for protein identification. Phytochemistry 65:1733–1744

Natale DA, Galperin MY, Tatusov RL, Koonin EV (2000) Using the COG database to improve gene recognition in complete genomes. Genetica 108:9–17

Oehrle NW, Sarma AD, Waters JK, Emerich DW (2008) Proteomic analysis of soybean nodule cytosol. Phytochemistry 69:2426–2438

Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M (1999) KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Res 27:29–34

Qin J, Gu F, Liu D, Yin CC, Zhao SJ, Chen H, Zhang JN, Yang CY, Zhan X, Zhang MC (2013) Proteomic analysis of elite soybean Jidou17 and its parents using iTRAQ-based quantitative approaches. Proteome Sci 11:12

Qiu HM, Liu CY, Zhang DJ, Xin XJ, Wang JL, Wang J, Shan CY, Shan DP, Hu GH, Chen QS (2009) Proteome analysis on resistance to phytophora root rot in soybean. Acta Agron Sin 35:418–423

Raines CA (2003) The Calvin cycle revisited. Photosynth Res 75:1–10

Ren C, Bilyeu KD, Beuselinck PR (2009) Composition, vigor, and proteome of mature soybean seeds developed under high temperature. Crop Sci 49:1010–1022

Ross PL, Huang YN, Marchese JN, Williamson B, Parker K, Hattan S, Khainovski N, Pillai S, Dey S, Daniels S, Purkayastha S, Juhasz P, Martin S, Bartlet-Jones M, He F, Jacobson A, Pappin DJ (2004) Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. Mol Cell Proteomics 3:1154–1169

Roy A, Rushton PJ, Rohila JS (2011) The potential of proteomics technologies for crop improvement under drought conditions. Crit Rev Plant Sci 30:471–490

Salavati A, Khatoon A, Nanjo Y, Komatsu S (2012) Analysis of proteomic changes in roots of soybean seedlings during recovery after flooding. J Proteomics 75:878–893

Sturn A, Quackenbush J, Trajanoski Z (2002) Genesis: cluster analysis of microarray data. Bioinformatics 18:207–208

Wang LQ, Ma H, Song LR, Shu YJ, Gu WH (2012) Comparative proteomics analysis reveals the mechanism of pre-harvest seed deterioration of soybean under high temperature and humidity stress. J Proteomics 75:2109–2127

Wang P, Kelly S, Fouracre JP, Langdale JA (2013) Genome-wide transcript analysis of early maize leaf development reveals gene cohorts associated with the differentiation of C4 Kranz anatomy. Plant J 75:656–670

Xu C, Sullivan JH, Garrett WM, Caperna TJ, Natarajan S (2008) Impact of solar ultraviolet-B on the proteome in soybean lines differing in flavonoid contents. Phytochemistry 69:38–48

Yeung KY, Haynor DR, Ruzzo WL (2001) Validating clustering for gene expression data. Bioinformatics 17:309–318

Zeng WY, Yang SP, Yu DY, Gai JY (2007) A comparative study on anther proteomics between cytoplasmic nuclear male-sterile line NJCMS2A and its maintainer of soybean. Acta Agron Sin 33:1637–1643

Zenoni S, Ferrarini A, Giacomelli E, Xumerle L, Fasoli M, Malerba G, Bellin D, Pezzotti M, Delledonne M (2010) Characterization of transcriptional complexity during berry development in *Vitis vinifera* using RNA-Seq. Plant Physiol 152:1787–1795