PROTEOMICS GUIDED GENE DISCOVERY FOR PLANT GROWTH AND DEFENSE

A Dissertation

by

YIXIANG ZHANG

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Joshua S. Yuan
Committee Members,	Susie Y. Dai
	Dennis C. Gross
	Herman B. Scholthof
Head of Department,	Leland S. Pierson III

August 2014

Major Subject: Plant Pathology

ABSTRACT

Research carried out in this dissertation aims to discover and validate novel gene functions in plant defense and growth. Plants deploy many mechanisms to respond and overcome biological, physical, and environmental challenges. By taking advantage of proteomics, we can discover dynamic changes of the proteome during the plant response. Nonetheless, the complexity of the proteome and the effect of high abundant proteins hinder the discovery of proteins with vital functions at low abundance.

I developed a new sample preparation method, dubbed Polyethyleneimine Assisted Rubisco Cleanup (PARC) to fractionate and deplete an abundant protein, namely Rubisco, from plant protein samples. The new approach was applied to investigate mechanisms for plant defense against herbivorous insects. My results indicated that PARC can effectively remove Rubisco and almost two times more differentially regulated proteins were identified. Over-expression of jacalin-like and cupin-like genes was carried out to validate their role in insect resistance in rice. The results further highlighted that PARC can serve as an effective strategy for gene discovery.

Furthermore, I integrated a rapid sample preparation method and bioinformatics classification system for comparative analysis of plant responses to two plant hormones, zeatin and brassinosteroid (BR). My data showed the metabolic pathways in sucrose and starch biosynthesis and utilization were significantly changed in zeatin treated plants, yet the lipid biosynthesis remained unchanged. For brassinosteroid treated plants, lipid

biosynthesis and β-oxidation were both down-regulated, yet the changes in sucrose and starch metabolism were minor.

Finally, a prohibitin (PHB8) involved in plant growth regulation was discovered by bioinformatics and proteomics methods. The over-expression of PHB8 in Arabidopsis increased the plant height, stem diameter, branch number and seed yield significantly. Downstream proteomics revealed that due to PHB8 over-expression, an ATPase beta subunit is significantly up-regulated. Further genetic study showed that over-expression of ATPase led to a similar phenotype as PHB8 over-expression lines. Pull-down assays revealed that PHB8 interacts with PHB9 and PHB16 to regulate ATPase level and energy metabolism.

Overall, my research shows that the latest proteomics platform and appropriate sample preparation methods can facilitate the discovery of genes involved in plant defense and growth regulation.

DEDICATION

To my wife, Lin

ACKNOWLEDGEMENTS

I would like to express my deep gratitude to my advisor Dr. Joshua S. Yuan, for supporting me. He guided me to learn useful knowledge, to solve problems, and to think with his constant patience, encouragement and support. He has been patient, caring and helpful during my pursuit of a PhD degree. I would like to thank my committee members, Dr. Dennis Gross, Dr. Herman Scholthof, and Dr. Susie Dai. They were always supportive, helpful and gave me valuable suggestions throughout the course of this research.

Thanks also go to my colleagues, especially Dr. Peng Gao, Dr. Xing Zhuo, Dr. Chao Di, Dr. Yaling Song, Dr. Zhide Chen, Dr. Shumei Jin and all of my lab members, department faculty and staff for making my time at Texas A&M University a great experience.

Finally, thanks to my mother and father for their encouragement and support and to my wife for her love. Without their love and patience, the work would never be done. They are the reason and the strength to make me move forward.

TABLE OF CONTENTS

Page

LIST OF FIGURES

LIST OF TABLES

Page

CHAPTER I

INTRODUCTION

1.1. FROM SINGLE MOLECULE TO SYSTEMS BIOLOGY

With the advancement of technology, the focus of current life science has been switched from single DNA, RNA, protein or metabolite to a systems view of whole sets of biomolecules produced by living organisms or complex systems. The terms, such as genomics, transcriptomics, proteomics and metabolomics, arose to the view of scientists as vital platforms of systems biology. Systems biology is used broadly since a decade ago, and it is still an emerging area which focuses on the structure and dynamics of the whole biological system (1).

Proteomics is considered a powerful approach for systems biology studies regarding protein dynamics and interaction networks (2). However, the nature of the proteome makes proteomics studies insurmountable for scientist with insufficient approaches to address the problem. First of all, the complexity of the whole proteomes, such as variations of chemical and physical properties, the quantity and distribution of expressed protein in a cell, make the protein identification difficult. Secondly, the proteome is dynamically changed in terms of expression and abundance in the cell. Finally, the broad occurrence of post translational modifications (PTMs) results in proteins with different functions and localizations. A new terminology, namely proteoforms, was introduced to describe the different forms of protein molecules arising from a single gene, which includes genetic variations, RNA alternative splicing, and

PTMs (3). The complication of proteomics compared to genomics and transcriptomics makes it a challenging and fascinating research area.

1.2. PROTEOMICS PLATFORMS

Various proteomics platforms have been applied to plant biology studies, and these platforms can be generally classified as gel-based and gel-free platforms (4). The gel-based platforms involve 1D (one dimensional) or 2D (two dimensional) electrophoresis separation of proteins and further protein identification with mass spectrometry-based methods. Even though 2D-PAGE is considered as the golden standard for proteomics analysis, the inherent limitation of 2D-PAGE makes the successful identification of proteins difficult with extreme PI and molecular weights, membrane proteins and low abundant proteins (5).

The recent developments of MudPIT (Multidimensional Protein Identification Technology) techniques have enabled many different gel-free proteomics platforms (6). For MudPIT analysis, the total protein sample is digested into peptides and the mixture of the peptides is separated by the multidimensional HPLC (High Performance Liquid Chromatography). The separated peptides are subjected to the MS/MS (tandem mass spectrometry) analysis and the resultant mass spectra are then searched against protein sequence database with various algorithm programs for protein identification. The MudPIT-based gel-free platforms have generally led to a deeper coverage of proteome, better quantification, and better capacity to analyze membrane proteins as compared to

the gel-based analysis. For these reasons, gel-free platforms have gained popularity to become the next generation proteomics analysis approaches (2, 7).

Even though the MudPIT platform possesses superior performance than gel-based platforms, mapping the whole proteome dynamic is still a challenge for scientists. The protein identification could be improved by many aspects, such as sample preparation, sample processing, instrumentation, and bioinformatics tools. The purpose of developing efficient sample preparation methods is to reduce the complexity of total protein which leads to the improvement of low abundant protein identification. For instance, subproteomics is a term used to describe the study of a subset of proteins. The object of subproteomics could be proteins in specific organelles, or proteins with particular PTMs obtained by enrichment (8).

1.3. PLANT PROTEOMICS

Even though the MudPIT proteomics platform has led to tremendous successes for plant studies, there are still great challenges (9). For example, abundant plant proteins like Rubisco (Ribulose-1, 5-bisphosphate carboxylase oxygenase) impose serious challenges for the proteomics analysis, because they reduce sensitivity, dynamic range and protein identification (10, 11). The signals of low abundant proteins can be easily masked by abundant proteins because of the mode of MS data collection. However, the identification of low abundant proteins is much more important because many of proteins, such as GTPases, kinases and phosphatases (12), are functional and important for biological processes.

Emerging proteomics platforms have charged plant biologists many powerful systems biology tools for plant growth and defense studies. For example, proteomics studies of protein kinases have been performed in plant growth, stress and pathogen defense (13-15). Even thorough transcriptome analysis has revealed systemic elevation of genes and pathways for secondary metabolite biosynthesis and relevant signaling pathway to account for rice defense insect herbivores is caused by abiotic stress (16), few researchers have used shot-gun proteomics approaches to profile the proteome dynamics at the systems level. Nevertheless, post-transcriptional level regulation is crucial for plant responses toward insect herbivores because defense pathways involve the regulation of protein degradation (17). Also, the correlation between transcriptome and proteome is low due to the fact that proteins undergo many post translational modifications (18). The comprehensive proteome profiling thus will allow comparisons of the transcriptome and proteome level regulation to reach a systems level understanding of the mechanisms for plant defense and growth.

1.4. PROTEOMICS STUDIES OF PLANT RESPONSES TO BRASSINOSTEROID

Phytohormones are a group of molecules that regulate plant growth and response to biotic and abiotic stresses. The proteome dynamics of plants in response to growth promotion hormones will shed light onto the mechanisms of how plants react to hormone application. Zeatin and Brassinosteroid (BR) are two plant hormones used in this research.

Brassinosteroids play a very important role in plant physiology, development, and defense. Proteomics studies have been performed to identify proteins involved in BR responses in many plants, such as Arabidopsis, mung bean and rice (19-22). Proteomics combined with genetic studies have revealed the BR signaling pathway by identifying pivotal components in Arabidopsis (20, 23, 24). Wang's lab focused on dissecting the BR pathway by using proteomic and genetic methods (24). They used 2-D DIGE (two-dimensional difference gel electrophoresis) and LC-MS/MS (liquid chromatography-tandem MS) to study BR mutants to investigate early response proteins. Their studies found BR-signaling kinases as an intermediate between plasma membrane receptor kinase BRI1 and downstream regulation components thus built a complete signaling transduction pathway in Arabidopsis (23, 24).

Proteomics was applied not only in studying the signaling pathways but also other gene products under the regulation of BR. Rice seedlings treated with brassinolide (BL), the most active form of BR, showed increased lamina inclination and root elongation. The proteins identified in lamina joints and roots were mainly photosynthesis and stress. With the limitation of MS techniques and bioinformatics tools, the authors were able to identify 21 proteins regulated by BL treatment by 2D-PAGE and MALDI-TOF MS (matrix-assisted laser desorption/ionization time-of-flight mass spectrometry) and only 2 could be identified in both treatments (21). The proteomics of mung bean epicotyl treated with BL under chilling stress was carried out, and 17 proteins down-regulated in a chilling treatment were up-regulated after the application of BL (19). However, few researchers examined the crosstalk between brassinosteroids and other

phytohormones in plants. A study using cDNA microarray and proteomics analyzed rice root after BL and GA treatment. The study successfully identified previously reported and newly found BR and GA pathway related proteins, however, the authors did not make the comparison between proteomics results of the two plant hormones treatments (25). However, finding the differences highlighting the molecular and metabolic mechanisms for response to plant growth promoting hormones can help to design better strategies to promote plant biomass accumulations.

1.5. PROTEOMICS STUDIES OF PLANT-INSECT INTERACTIONS

Besides that, abiotic stress causes yield loss, biotic stress, (i.e. , bacteria, fungus, virus and phytophagous insects), also causes huge losses in crop production (26). Insect herbivores are one of the most severe causes of crop yield loss among biotic stresses. During the long term of evolution, plants developed many mechanisms to cope with insect herbivores. Constitutive defense mechanisms will protect plants from insects without insect attack, such as preexisting physical barriers, toxins, and phenolic compounds. However, the inducible defense mechanism in plants, which are induced after insect attack, will significantly contribute to plant protection against insect feeding by direct and indirect means. For example, plants produce many secondary metabolites, such as terpenoids, glucosinolates, and alkaloids to guard themself against insect herbivores (27). Protease inhibitors are another major group of defense components plants make to inhibit insect nutrient absorption (28).

Understanding how plant proteins are regulated during insect herbivore attack will help find pivotal genes that control defense responses. On the other hand, how insects adapt to the plant defense responses will help reveal the relationship between insects and plants from an evolutionary perspective. How do plants develop different defense mechanisms to protect them against insects attack? How do insects in turn adapt to the responses of plants? By what mechanism do plants perceive the difference between wounding and insect feeding and thus trigger the downstream signaling pathway? Many studies have investigated defense genes against insect herbivores by variant biochemistry and molecular biology methods (29-31). With the development of large scale experimental approaches, scientists performed many 'Omics' studies to decipher the steps involved in plant insect interactions. The development of transcriptomics study tools has allowed large scale exploration of differentially expressed genes during plant insect interaction (32). However, the transcriptome profile change of attacked plants cannot be correlated well with proteome changes. It may be due to that the post-transcriptional and post-translational modifications that determine the different fates of RNAs and proteins. Thus it is inappropriate to predict proteome changes only based on RNA level changes. (33)

As an offensive part of plant-insect interactions, insect herbivores need to break down plant cell walls by degradation enzymes to release nutrition out of the plant cell. Meanwhile, for effective nutrition absorption, insects have to develop a set of mechanisms to degrade or tolerate defense compounds produced by host plants. Therefore, generalist (33, 34) and specialist (35) insect herbivores might develop

different tolerance mechanisms against plant defense. Generalist herbivores, especially for whose host range crosses different families, require more diverse adaptation mechanisms to cope with a broad range of plant defense compounds (34). On the other hand, specialist insect herbivores may possess much more targeted mechanisms to counteract host defense compounds (35-37).

Up to date, many studies have been done in the relationship between insect gut content and plant defense proteins. Many proteins involved in the defense function reduce insect adaptation to plants by interfering with nutrient acquisition and digestion. For example, Howe's lab (38) focused on jasmonic acid (JA) induced plant defense responses against insect herbivores. They identified arginase and threonine deaminase (TD) to be important for JA inducible proteins in the midgut content of *Manduca sexta* (tobacco hornworm) larvae by LC-MS/MS. They further studied the TD isoform stability in the midgut of tobacco hornworm by shot-gun proteomics. TD2, which undergoes C-terminal removal after herbivore attack, was found to be the most abundant proteins in the insect midgut and frass by shotgun proteomics (37). The proteolysis of TD2 occurred in the gut of generalist lepidopteran instead of coleopteran by chymotrypsin-like protease produced by insects (39). Their researches demonstrated a way to study plant sourced defense proteins in the insect guts and how plant defense adapted to insect herbivores by targeting the insect digestion processes. Another type of plant defense compound that targets insect postingestive processes is a group of protease inhibitors (PIs) (40-43). The proteome changes in the *Callosobruchus maculatus* larvae

intestinal tract after ingestion of a cysteine protease inhibitor revealed how this seed-feeding insect responded to counteract the antinutrient effect of PIs (44).

Many studies of aphid proteome changes have been carried out among different host and endosymbionts. A study of proteome changes in *Myzus persicae* aphid was performed on different hosts switched from *Brassicaceae* to *Solanaceae* families. Due to the limitation of the platform and bioinformatics tools, only 14 proteins were identified that were differentially expressed during host switch, which were involved in primary metabolism and cytoskeleton formation (34). The complex interactions among *Macrosiphum euphorbiae* aphids, host and endosymbionts were revealed by different groups of scientists. The study of *M. euphorbiae* proteome responses to different hosts and stresses showed that the aphid proteome underwent specific changes against each stress. The proteome changes mainly included primary metabolism related proteins and endosymbionts derived proteins, which render aphids the adaptation to different host resistance and stresses (45, 46).

As a defensive part of plant insect interactions, plants deployed many approaches to defend against insects. Most of the previous research to characterize proteins was carried out using gel-based platforms, resulting in the identification of limited numbers of differential proteins (33, 47-49). For instance, proteins separated by 2D-PAGE then followed by MALDI-TOF-MS or LC-MS/MS. These methods were broadly used to study pull-down assays (50) or proteome profiling (51). Other efforts have been put into Sitka spruce (*Picea sitchensis*) and Norway spruce (*Picea abies*) by Canadian scientists using SRM, MudPIT and 2D-LC-MS/MS to qualitatively or quantitatively study the

proteome change (47, 52-54). Plants respond to insect attacks with systemic up-regulation of regulatory, metabolic and defense-related proteins, and the coverage of proteome is crucial for revealing in-depth mechanisms (55). My lab previously performed transcriptomic analysis using rice as a model plant. The result revealed molecular and genomic mechanisms for plant defenses, including the regulation of enzymes involved in secondary metabolisms and volatile production (16). Despite the progress, only a few studies have been performed to comprehensively study plant defense against herbivorous insects using a shot-gun proteomics approach. Therefore, a comprehensive proteomics study of plant-insect interaction will give deeper understanding of how plants interact with their enemies.

1.6. STUDIES OF PHB GENE FAMILY

Prohibitins (PHBs) are ubiquitous and evolutionarily conserved proteins with diverse functions, such as signaling, cell cycle regulation, mitochondrial respiration, cell death and aging (56-63). The exact molecular functions of PHBs are still under debating. PHBs were originally indicated to be important in lipid raft formation in mitochondrial and other organelle membranes (58, 64). Then PHBs were found to present in the nucleus to interact with transcriptional factors (65). A recent study indicated that PHB proteins can form a complex with membrane bound ubiquitin ligase to regulate ERAD (endoplasmic reticulum (ER)-associated degradation) (64, 66).

Prohibitins have been discovered in plants for more than 15 years, but studies of PHB genes in plants are still very limited (67). An extensive bioinformatics analysis

regarding motif distribution, intron/exon structure, and digital expression pattern of PHBs in Arabidopsis was carried out previously. The family of PHB was broadened to 5 classes and the members were expanded from 7 to 17 in Arabidopsis (68). The studies of plant PHBs were mainly focused on PHB1, 2, 3, 4 in Arabidopsis and their homologs in tobacco (69-72).

The functional studies of prohibitins in Arabidopsis and tobacco indicated that prohibitins play important roles in ROS responses and mitochondria morphology (69, 73). PHB3 was found to be involved in NO homeostasis and abiotic stress response (72). AtHIR1 (Hypersensitive Induced Reaction proteins) and AtHIR2, which also named PHB8 and PHB9, were showed can form a complex with RPS2 to function in effector triggered immunity (ETI) during plant disease defense (70). Recently, PHB2 was identified as an interacting protein with endoplasmic reticulum BAX INHIBITOR-1 (BI-1) to regulate cell death and plant powdery mildew interaction (74).

Bioinformatics analysis of PHB gene family from prokaryotes to eukaryotes indicated that PHB gene family is a ubiquitous existed gene family (68). As an evolutionally conserved but functionally diverse gene family, the studies of prohibitins in model species could be a good guidance for other species. Prohibitins were related to aging, senescence and obesity in animal which are relevant to energy metabolism. However, not like their homologs in animals, few studies reported that plant prohibitins involved in energy metabolism. The current study progresses revealed the importance of the function study of prohibitins in plant growth regulation and energy metabolism.

1.7. DISSERTATION FOCUS

The hypothesis of the study is the novel mechanisms for protein level regulation of plant growth and defense can be revealed by reducing protein sample complexity. The reduced sample complexity can be achieved by sample fractionation and preparation, and thus to improve protein identification. The study aims to discover important proteins and pathways involved in plant growth regulation and defense responses (Fig. 1.1).

In Chapter II, simple separation steps by enrichment of mitochondria and chloroplast was developed to avoid multiple purification steps. By taking advantage of proteomics and bioinformatics approaches, differences and crosstalk was investigated between different plant hormones triggered molecular and metabolic responses. The study showed that the rapid sample preparation method assisted by bioinformatics classification could provide effective proteomics analysis of plant hormone responses through pathway characterization. In Chapter III, a new protein preparation method was developed to enhance the detection of low abundance proteins and increased the performance of proteomics in terms of both peptide quantification and proteome coverage. Meanwhile, the key genes and pathways involved in plant response to herbivore insects were identified. In Chapter IV, the bioinformatics and proteomics approach led to the discovery of the PHB8 gene involved in the regulation of plant growth. Molecular studies indicated that PHB8 may function as a molecular chaperone to regulate protein stability in different membrane systems. Overall, the latest proteomics platform will be used to identify the genes involved in plant defense and growth

regulation. These genes enable us to a deeper understanding of plant-insect defense and growth regulation at both the molecular and systems levels.

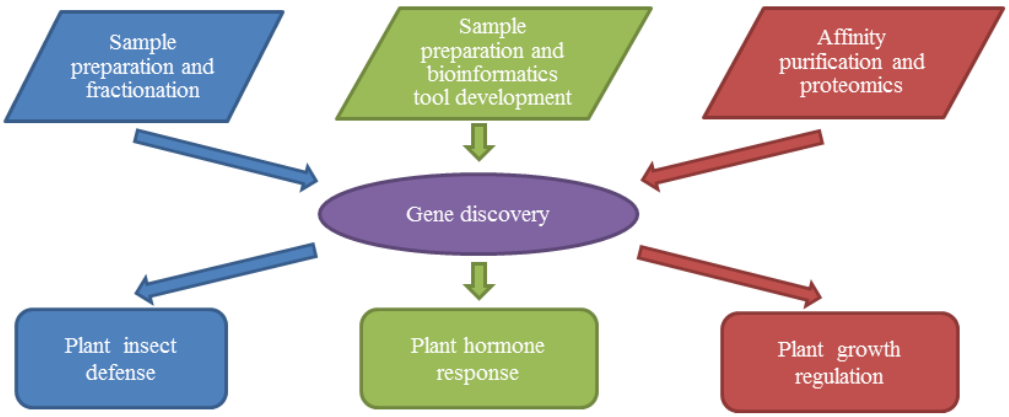

Fig. 1. 1. Outline of the research.

CHAPTER II

APPLICATION OF AN IMPROVED PROTEOMICS METHOD FOR ABUNDANT PROTEIN CLEANUP: MOLECULAR AND GENOMIC MECHANISMS STUDY IN PLANT DEFENSE*

2.1. SUMMARY

High abundance proteins like Ribulose-1,5-bisphosphate carboxylase oxygenase (Rubisco) impose a consistent challenge for the whole proteome characterization using shot-gun proteomics. In order to address this challenge, I developed and evaluated Polyethyleneimine (PEI) Assisted Rubisco Cleanup (PARC) as a new method by combining both abundant protein removal and fractionation. The new approach was applied to a plant insect interaction study to validate the platform and investigate mechanisms for plant defense against herbivorous insects. My results indicated that PARC can effectively remove Rubisco, improve the protein identification, and discover almost three times more differentially regulated proteins. The significantly enhanced shot-gun proteomics performance was translated into in-depth proteomic and molecular mechanisms for plant-insect interactions, where carbon re-distribution was used to play an essential role. Moreover, the transcriptomic validation also confirmed the reliability

*This research was originally published in Molecular & Cellular Proteomics. Yixiang Zhang, Peng Gao, Zhuo Xing, Shumei Jin, Zhide Chen, Lantao Liu, Nasie Constantino, Xinwang Wang, Weibing Shi, Joshua S Yuan and Susie Y Dai. Application of an improved proteomics method for abundant protein cleanup: molecular and genomic mechanisms study in plant defense. *Molecular & Cellular Proteomics*. 2013; 12:3431-3442. © the American Society for Biochemistry and Molecular Biology.

of PARC analysis. Finally, functional studies were carried out for two differentially regulated genes as revealed by PARC analysis. Insect resistance was induced by over-expressing either jacalin-like or cupin-like genes in rice. The results further highlighted that PARC can serve as an effective strategy for proteomics analysis and gene discovery.

2.2. INTRODUCTION

One of the constant challenges for proteomics is inadequate protein identification due to the interference of high abundance proteins (75). The challenge is particularly critical for plant proteomics analysis due to the prevalence of Rubisco (Ribulose-1,5-bisphosphate carboxylase oxygenase) in green tissue. As a major enzyme involved in carbon fixation, Rubisco consists of 30 to 50% of total plant protein from green tissues and causes less sensitivity, dynamic range, and protein identification of plant proteomics (76-78). Influences of high abundance proteins like Rubisco affect both gel-based and shot-gun proteomics analysis. In one of the most popular shot-gun proteomics platforms with the data-dependent MS/MS acquisition, the peptides derived from the abundant proteins have more chance to be sampled by the MS instrument than the peptides from other functional proteins. Thus, the dynamic range and detection sensitivity will be sacrificed due to the prevalence of high abundance proteins (76). In order to address this challenge, I developed and evaluated a new method by combining PEI (Polyethyleneimine) precipitation and protein sample fractionation to improve the performance of Multidimensional Protein Identification Technology (MudPIT)-based proteomics analysis.

PEI is a positively charged polymer broadly used for removing nucleic acids from proteins (79). The compound can also be employed to remove acidic proteins like Rubisco from the total protein (80). PEI precipitation can be considered as a fractionation process to separate acidic proteins from the total protein, and thus can be used for both Rubisco removal as well as fractionation of plant proteins from green tissues. Despite the potential to be used for sample preparation, very few studies optimized the PEI precipitation for plant proteomics analysis and evaluated the effectiveness of the approach for enhancing proteomics performance. In this study, I demonstrated that PEI-Assisted Rubisco Cleanup (PARC) and fractionation could significantly improve protein identification. The method was also used to investigate the molecular mechanisms for plant insect interaction and discover two important regulators for rice defense against herbivorous insects.

Plant-insect interactions represent an important traditional research field due to the relevance for crop growth and yield. Proteomics has emerged as a major approach to study plant physiological, pathological and developmental processes at the systems level (81-83). Despite progress, only a few studies have been performed to comprehensively study plant defense against herbivorous insects using a proteomics approach. Most of the previous research was carried out using gel-based platforms, resulting in the identification of limited numbers of differential proteins (33, 47-49). Plants respond to insect attacks with systemic up-regulation of regulatory, metabolic and defense-related proteins, and the coverage of proteome is crucial for revealing in-depth mechanisms (55). Our previous transcriptomic analysis using rice as a model plant revealed

molecular and genomic mechanisms for plant defenses, including the regulation of enzymes involved in secondary metabolisms and volatile production (16). Despite the transcriptomic analysis, the protein level regulation of rice defense against herbivorous insects is largely unknown. The comprehensive proteome profiling will thus help to reveal novel mechanisms for protein-level defense regulation and to identify the new regulators in plant defense which will guide the crop improvement essential for food security.

In this study, I developed and evaluated a novel plant protein isolation and fractionation method to improve the performance of MudPIT-based shot-gun proteomics. I further applied the method for a plant-insect interactions study. The method increased protein identification as compared to the protein isolated by a commercial kit. The novel method also resulted in discovering almost three times more differential proteins. The significantly improved differential protein identification helped to reveal new proteome-level mechanisms for plant defense against herbivorous insects and to identify key genes involved in defense regulation. In particular, the proteomics results revealed dynamic re-allocation of carbon resources as a defense mechanism against herbivorous insects in plant. Furthermore, I carried out two levels of validation of proteomics data. The first was real-time PCR analysis to verify the gene expression of the up-regulated proteins. The second was the functional validation of two key regulatory proteins for plant defense. The first gene was a cupin-like protein specifically identified in PARC. The cupin-like protein was previously reported to be important for reactive oxygen species (ROS) production during the defense response. For example, in

a study researched by Carrillo et al. (84), a cupin-like protein Os03g48470 was demonstrated to be an oxalate oxidase to produce H_2O_2. However, it was not clear which one of more than 100 cupin-like proteins in the protein family was actually involved in plant defense against herbivorous insects (85-87). The second gene is a jacalin-like protein with unknown molecular function. Transgene over-expression experiments validated both genes to be important regulators for rice defense against herbivorous insects. Overall, the results highlighted that the PARC method could significantly improve the performance of shot-gun proteomics to enable the in-depth analysis of plant defense mechanisms and the discovery of new regulators for plant defense against herbivorous insects.

2.3. MATERIALS AND METHODS

2.3.1. Plant and Insect Growth

Rice (*Oryza sativa* ssp. *Japonica* cv. Nipponbare) seeds were germinated at 30 °C in the dark for five days. Then the seedlings were grown at 28 °C with 14 hours of light for two weeks. Fall armyworm (FAW) (*Spodoptera frugiperda*) eggs were obtained from Benzon Research (Carlisle, PA, USA) and hatched at 28 °C. FAW larvae were raised on an artificial diet and second-instar FAW were used for herbivore treatment. Two larvae were placed on the leaves of a single two-week old rice seedling around nightfall at 10pm. After 12 hours, approximately 5-10% of the leaf area was consumed. Insects were then removed and the rice plants were harvested to snap-freeze in liquid nitrogen. The samples were stored in ultralow freezer at -80 °C until further analysis.

2.3.2. Protein Sample Preparation Methods

The protein sample preparation integrates both PEI precipitation optimization and sample fractionation. For the reference, the total plant protein was extracted with Plant Total Protein Extraction Kit (Sigma-Aldrich, MO, USA) according to the manufacture's instruction with minor modifications. The protein sample is referred to as TP (Total Protein) fraction, which represented plant total protein extracted by a conventional method. Briefly, 100-250 mg of leaf tissue was ground into a fine powder in liquid nitrogen. The powder was washed with methanol and acetone, then pelleted and dried with a SpeedVac. The plant tissue pellet was dissolved in Reagent Type 4 Working Solution supplied by the kit. The protein extraction solution provided by the commercial Sigma kit contains 7 M urea, 2 M thiourea, 40 mM Triszma base, and 1% 3-(4-Heptyl)phenyl-3-hydroxypropyl dimethylammoniopropanesulfonate (C7BzO) as the detergent. The solution was adjusted to pH 10.4. The TP sample preparation using this commercial kit involves chaotropic reagents, particularly, with a zwitterionic detergent to dissolve most proteins including hydrophobic membrane proteins. After the extraction and removal of plant tissue debris, the TP samples were ready for proteomic analysis. However, the TP fraction cannot be used for PEI precipitation to remove Rubisco potentially due to the changes in protein charges by extraction buffer for the high concentration of the zwitterionic detergent.

Besides the TP sample prepared by the commercial kit, a modified SDS (sodium dodecyl sulfate)-based total protein extraction method was used to validate the effectiveness of the commercial kit (88). The SDS solubilized total protein was subject

to SDS-PAGE (sodium dodecyl sulfate polyacrylamide gel electrophoresis) analysis. The gel fragment containing proteins were washed, distained, dehydrated, and further processed for in-gel digestion according to previous studies (89).

In order to carry out PEI precipitation, a soluble protein extraction method was developed with a non-ionic detergent. Approximately 200 mg rice leaves were ground into a fine powder in liquid nitrogen using mortar and pestle. 1 ml of extraction buffer (50 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 150 mM NaCl, 0.1% Nonidet P-40, 1 mM PMSF) was then added and incubated with the sample. The extract was then centrifuged at 13,000 rpm for 10 minutes at 4 °C, and the supernatant was then transferred to a new micro-centrifuge tube. The supernatant contained most of the soluble protein in the cell and was thus referred to as TS (total soluble protein). The pellet was used to continue to extract protein using the same Plant Total Protein Extraction Kit (Sigma-Aldrich, MO, USA) according to the manufacturer's instruction. The sample contained mostly the insoluble protein and was thus referred to as IS. In addition to PEI precipitation, I also evaluated ammonium sulfate for its capacity to selectively precipitate Rubisco. 40% ammonium sulfate was used to precipitate Rubisco according to previous publication (90). The ammonium sulfate precipitated pellets were then re-suspended and analyzed by SDS-PAGE gel in the same way as PEI precipitated proteins.

PEI precipitation was carried out to remove the Rubisco contamination in the TS fraction. A titration was performed to determine the effective concentration for PEI precipitation. For LC-MS/MS experiments, the TS samples were first mixed with 100 mg/g of PEI, vortexed vigorously for 10 sec, and then precipitated on ice for 5 min.

Samples were then centrifuged for 15 min at 13,000 rpm. The supernatant was removed by pipetting and collected as SS (Supernatant Soluble) fraction. The SS fraction is the Rubisco removed soluble protein. The PEI precipitated pellets were then resuspended in resuspension buffer and centrifuged for 10 min at 13,000 rpm to re-pellet the debris. The supernatant was collected as PS (precipitated soluble protein). The fraction should be Rubisco enriched fraction containing many acidic proteins. The protein concentration for each fraction was measured by Bradford Protein Assay Kit (Thermo Scientific Pierce, IL, USA).

2.3.3. MudPIT and Shot-gun Proteomics

MudPIT-based shot-gun proteomics were carried out to analyze each aforementioned fraction as described by Washburn et al (91). Approximately 100 µg protein was digested by Mass Spectrometry Grade Trypsin Gold (Promega, WI, USA) with 1:40 w/w at 37 °C for 24 hr. The digested peptides were desalted using a Sep-Pak C18 plus column (Waters Corporation, Milford, Massachusetts, USA) and then loaded onto a biphasic (strong cation exchange/reversed phase) capillary column using a high pressure tank. The two-dimensional liquid chromatography separation and tandem mass spectrometry were carried out as previously described (92).

The separated peptides were analyzed using a linear ion trap mass spectrometer, Finnigan LTQ (Thermo Finnegan, San Jose, CA, USA). The mass spectrometry was set to the data-dependent acquisition mode. Then the full mass spectra were recorded on the peptides over a 300-1700 m/z range, followed by five tandem mass (MS/MS) events for

the most abundant ions from the first MS analysis. The Xcalibur data system (Thermo Fisher Scientific, San Jose, CA, USA) was used to control the LC-LTQ system and collect the data. The experiments were carried out with duplicate biological samples for all fractions except the TP fraction.

2.3.4. Data Analysis

A data preprocessing pipeline based on Xcalibur Development Kit was used to generate DTA files in the same way as the ThermoFinnigan Bioworks (2.0) software. Tandem mass spectra were extracted from the raw files and converted into the MS2 file. The MS2 file was searched against the rice protein database containing 66,338 protein sequences from the MSU Rice Genome Annotation (Release 7), the same number of reverse sequence, and common contaminant proteins. The reverse sequences of the original dataset were included to calculate confidence levels and false-positive rates. The common containment proteins were included for data quality control. The ProLuCID (version 1.0) algorithm was used to assign peptide sequence to peptide fragmentation spectra using the Texas A&M Supercomputing Facility (93). The ProLuCID parameters were set as follows: the precursor mass accuracy was set at 100 ppm; fragment ion mass accuracy was set at 600 ppm. No fixed or variable modifications were considered. At least two distinct peptides (semi-tryptic) were required to identify a protein with no sequence coverage assigned.

Protein quantification was achieved by spectral counting. The validity of peptide/spectrum matches were assessed in DTASelect v.2.0 using a 0.05 false discovery

cutoff, with a cross-correlation score (XCorr) that's larger than 1, and normalized difference in cross-correlation scores (DeltaCN) larger than 0.08. Semi-tryptic peptide was included in the final calculation. DTASelect software listed the accession numbers together for those proteins combined into one group. Proteins identified with more than two peptides were used for quantification. Protein identification was based on both specific and non-specific peptides. Quantification is based on duplicate biological samples.

2.3.5. Ontology and Pathway Analysis

PatternLab (94) software (version 2.1.1.5) was used for data analysis to discover differentially expressed proteins. Isoforms of protein were combined in groups by PatternLab. The Row Sigma normalization was carried out to adjust systemic errors. The TFold test was applied to derive differentially expressed proteins. The cutoff of p value (BH q-value) and FDR were 0.05 for both. F-stringency was optimized by software automatically. Gene ontology analysis was performed using agriGO (95). Each protein was classified with respect to its biological process using GO annotation. The pathway analysis of differentially regulated proteins was analyzed using KEGG (96).

2.3.6. Alignments and Phylogenetic Analysis

A phylogenetic tree was generated using sequences download from NCBI reference proteins and UniProtKB/Swiss-Prot database. The threshold for blasted sequences of *Os03g48770* was 1e-54. The amino acid sequences were initially aligned

using ClustalX (97). The trees were created using the neighbor-joining method of MEGA version 5.05 with the Poisson model having uniform rates for all sites. The phylogenetic trees were tested by 1000 bootstrapping (98). In addition to a phylogenetic tree, multiple sequence alignments were carried out for wheat oxalate oxidase GF-3.8 (P26759), rice germin-like protein 3-6 (Os03g48780) and the cupin domain containing protein Os03g48770 by Clustal Omega (http://www.ebi.ac.uk/Tools/msa/clustalo/).

2.3.7. Real-time Quantitative PCR

For the qRT-PCR analysis, wild type rice seedlings were infested with FAW larvae and the gene expression was analyzed at series of time points of 0 hr, 12 hr, 24 hr and 48 hr. RNA isolation using TRI Reagent was performed following manufacture's instruction (99).

Reverse transcription of total RNA was performed using SuperScript III First-Strand Synthesis kit (Invitrogen, CA, USA). Triplicate quantitative assays were performed using the SYBR Green Master Mix (Applied Biosystems, CA, USA) with an ABI 7900HT sequence detection system. All data were normalized with 18S rRNA as an internal reference gene using the following primers: forward 5'-CGGCTACCACATCCAAGGAA-3'; reverse 5'-TGTCACTACCTCCCCGTGTCA-3'. The relative quantification method (modified ΔΔCT) was used to evaluate the differential gene expression. The ΔCTs were normalized against the lowest expressed samples within the group to derive ΔΔCT, where the smallest ΔΔCT for each group will be 0. The $2^{\Delta\Delta CT}$ will represent the ratio of gene expression over the lowest expressed

sample in the group. Both the ratio of gene expression and 95% confidence intervals derived from triplicate assays were presented.

2.3.8. Rice Transformation

For overexpression of *Os03g48770* and *Os12g14440,* full-length cDNA were sub-cloned into a binary vector *pCXUN* with ubiquitin promoter (100) and introduced into Nipponbare rice embryonic calli by *Agrobacterium tumefaciens*-mediated transformation (101). Transgenic plants were selected on selection plates with 50 mg/L hygromycin. T1 generation plants were used to perform insect feeding assays.

2.3.9. Leaf Area Damage Measurement

Transgenic and wild type rice leaves from T1 plants were collected and cut into small pieces. The pieces were inserted into 10 ml 0.7% agarose in a 30 ml plastic cup. One 3rd instar larva was put on the leaf fragment and allowed it to feed for 48 hr. The percentage of leaf area damage was calculated by ImageJ (102). ANOVA analysis was carried out to compare control groups vs. transgenic lines with 95% confidence intervals. Three validated transgenic lines from each target genes were used. For each transgenic line, six leaf segments were used for insect treatment bioassay and leaf area measurement.

2.4. RESULTS

The study integrated method development with functional proteomics analysis. First, a new strategy for plant shot-gun proteomics, PARC (PEI-Assisted Rubisco Cleanup**)** was developed to improve the protein identification and differential expression analysis. Second, the method was used to investigate the proteomic and molecular mechanisms for plant defense against herbivorous insects, which further validated PARC as an effective method to reveal in-depth mechanisms and novel regulators for plant insect interaction.

2.4.1. PARC as an Efficient Approach to Remove Rubisco

The method development started with optimization of PEI precipitation for removing Rubisco from extracted protein samples (Fig. 2.1). PEI precipitation was carried out using two types of input protein, the total soluble protein (TS) extracted by soluble protein extraction buffer and the total protein (TP) extracted by the urea/thiourea based buffer in the Plant Total Protein Extraction Kit (See Experimental Procedures for details). As shown in Fig. 2.1*A*, PEI precipitation could effectively remove Rubisco at the concentration of 50 to 100 mg/g from the total soluble proteins TS. However, PEI precipitation could not selectively precipitate Rubisco from the total protein (TP) extracted by the commercial Sigma Aldrich kit (Fig. 2.1*B*). As shown in Fig. 2.1*B*, when 25 and 50 mg/g concentrations of PEI were used, both Rubisco and other proteins were precipitated from the total protein sample TP. The difference in selectivity might result

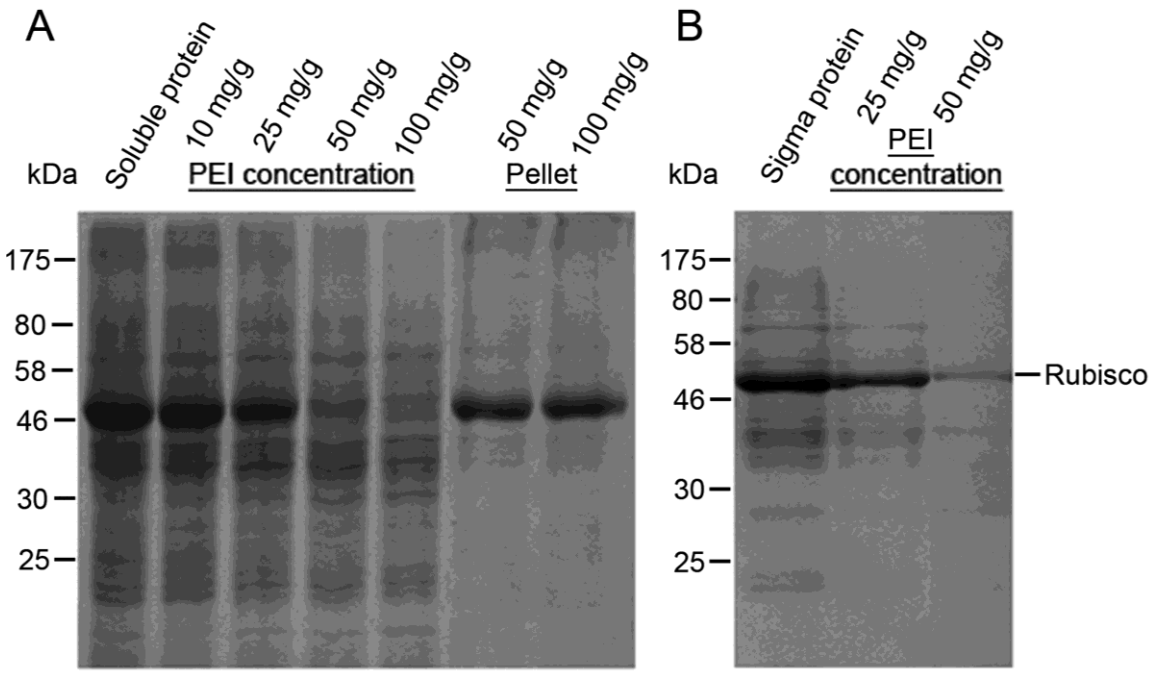

Fig. 2. 1. Efficiency of Rubisco removal by PEI (Polyethyleneimine) precipitation. A, PEI precipitation of total soluble protein isolated by soluble protein extraction buffer. From the left to the right, lane 1 shows the total soluble protein extracted by soluble protein extraction buffer. Lane 2 to 5 shows the supernatant of soluble protein after precipitation with 10, 25, 50, and 100 mg/g PEI (weight/total protein), respectively. Lane 6 and 7 are dissolved protein pellets after PEI precipitation. B, Incompatibility between PEI precipitation and total protein extracted by Sigma Aldrich Plant Total Protein Extraction kit. From the left to the right, lane 1 shows the total protein extracted by the kit. Lane 2 and 3 shows total protein supernatant after PEI precipitation at 25 and 50 mg/g, respectively. PEI concentration (mg/g) is as indicated at the top. Molecular weight markers in kDa are shown on the left.

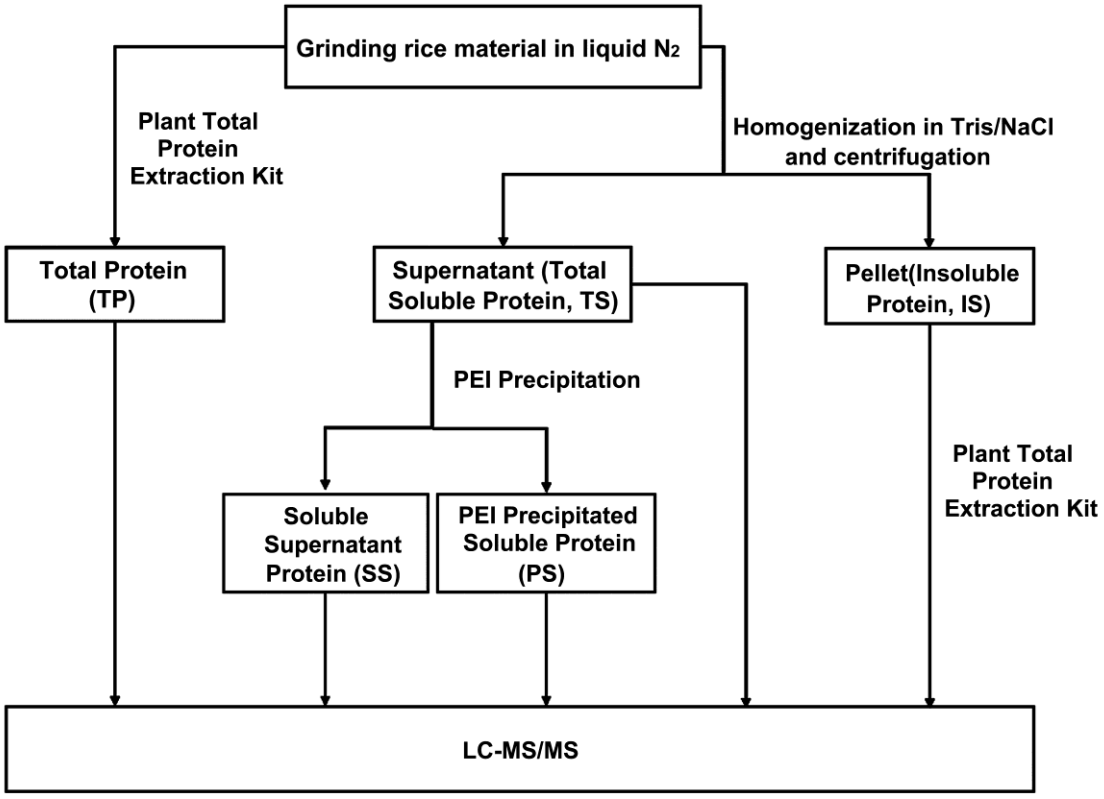

Fig. 2. 2. Flowchart of protein extraction, Rubisco removal and fractionation method. Rice protein samples were extracted and fractionized into five fractions. Aerial part of rice seedlings were ground into fine powder in liquid nitrogen. Total protein was extracted by Sigma Plant Total Protein Extraction Kit (TP). Total soluble protein TS and Insoluble protein (IS) were obtained by centrifugation after the powder was homogenized in extraction buffer. PEI precipitation was used to fractionize the TS fraction into Soluble Supernatant protein (SS) and PEI Precipitated Soluble protein (PS).

from the change of protein charges in the TP sample caused by the urea and thiourea buffer.

Based on the results, I developed the strategy as described in materials and supplies to remove Rubisco and to fractionate protein (Fig. 2.2). In this strategy, the total soluble protein TS was first extracted and then fractionated into supernatant Rubisco reduced fraction (SS) and pelletized Rubisco enriched fraction (PS) using 100 mg/g PEI. In the meantime, the insoluble protein (IS) can be extracted by the aforementioned the Sigma Aldrich Plant Total Protein Extraction Kit, and the IS, PS and SS constituted the three fractions of the total protein, which can be compared to TP. Shot-gun proteomics analysis was then carried out for each fraction (SS, PS, IS, TS, and TP).

Two additional control experiments were also carried out. First, the effectiveness and suitability of the commercial kit as reference was further evaluated by comparison of the protein identification using SDS-based protein sample preparation with that of TP. As shown in Fig. 2.3, more non-redundant proteins were identified in the TP fraction than those from the SDS-based method. Second, different chemical compounds including PEI and ammonium sulfate can be used to 'selectively' precipitate Rubisco. I briefly compared the effectiveness of PEI versus ammonium sulfate. As shown in Fig. 2.4, the addition of ammonium sulfate resulted in the precipitation more non-Rubisco proteins as compared to PEI. The PARC method development thus is majorly based on PEI due to its higher specificity. In particular, the efficacy of PARC for protein identification and differential protein expression can be evaluated comparing 1) SS and PS vs. TS for soluble proteins as well as 2) SS, PS and IS vs. TP for total proteins.

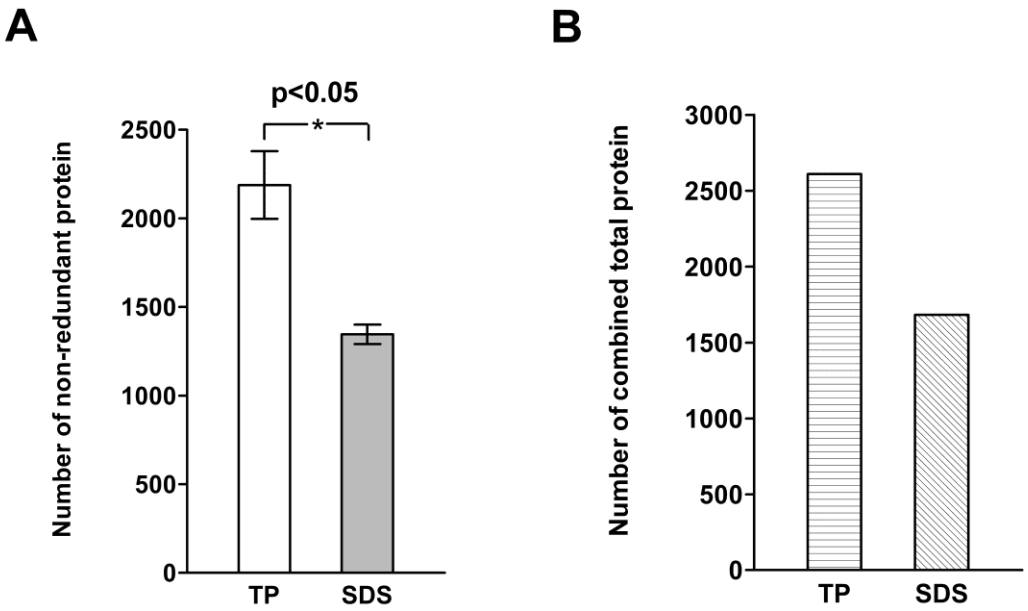

Fig. 2. 3. Comparison of total protein identified from SDS-based protein extraction (SDS) and total protein (TP) extracted by commercial kit. The number of non-redundant protein identified by SDS-based protein extraction and total protein extraction using the commercial Sigma kit was presented (*A*). The standard error was calculated based on duplicate biological samples. In addition, the combined number of proteins (sum of duplicated biological samples) identified by each method was presented (*B*).

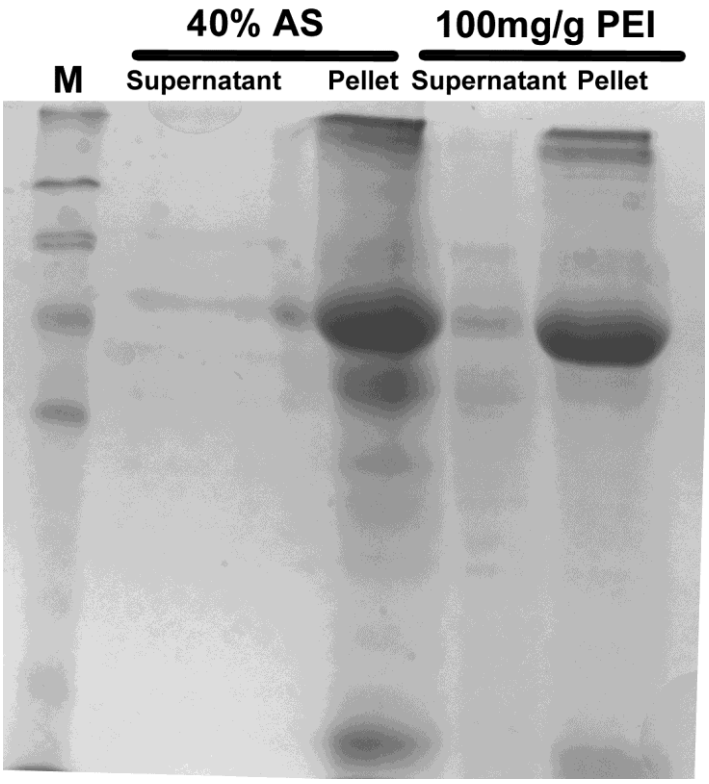

Fig. 2. 4. Comparison of PEI and ammonium sulfate-precipitated proteins using SDS PAGE gel. The optimal precipitation conditions for both AS (ammonium sulfate, 40%) and PEI (100 mg/g) were used to remove Rubisco from total proteins. The SDS PAGE analysis of both the supernatant and the pellets was shown in the figure.

2.4.2. Effective Removal of Rubisco, Improved Protein Identification and Enhanced Differential Analysis of Shot-gun Proteomics

The aforementioned protein fractions were subject to the MudPIT-based proteomics analysis. The effectiveness of PARC was evaluated from three aspects. First, I analyzed the percentage of spectral counts from Rubisco as part of the total spectral count for each fraction (TS, SS, PS, IS, TP). As shown in Fig. 2.5, PARC could efficiently remove Rubisco from soluble proteins. The Rubisco spectra counts were 21.33% of the total protein spectral counts TP. Similar to TP sample, the total soluble protein TS fraction contained 23.22% of Rubisco spectral counts. The insoluble fraction (IS) contained a surprisingly low percentage of Rubisco peptide counts at 9.81% of total spectral counts. SS and PS fractions were derived from TS when the sample was subject to PEI precipitation. As expected, the supernatant Rubisco-reduced fraction (SS) had very little Rubisco contamination with the Rubisco at 0.14% of total spectral counts. Meanwhile, the precipitated Rubisco-enriched fraction (PS) had a significantly higher percentage of Rubisco spectral counts at 36.12%.

Second, the protein identification was improved by fractionation and Rubisco removal as shown in Fig. 2.6. Total soluble protein TS was fractionated into SS and PS fractions through Rubisco removal and PEI precipitation. The fractionation process led to the identification of 523 more proteins and a total of 24.2% increase in the number of proteins identified (Fig. 2.6*A*). Furthermore, the Rubisco removed SS, Rubisco enriched PS, and the total insoluble proteins (IS) should represent the total proteins, which can be compared to the total protein (TP) extracted by the commercial kit. The three fractions

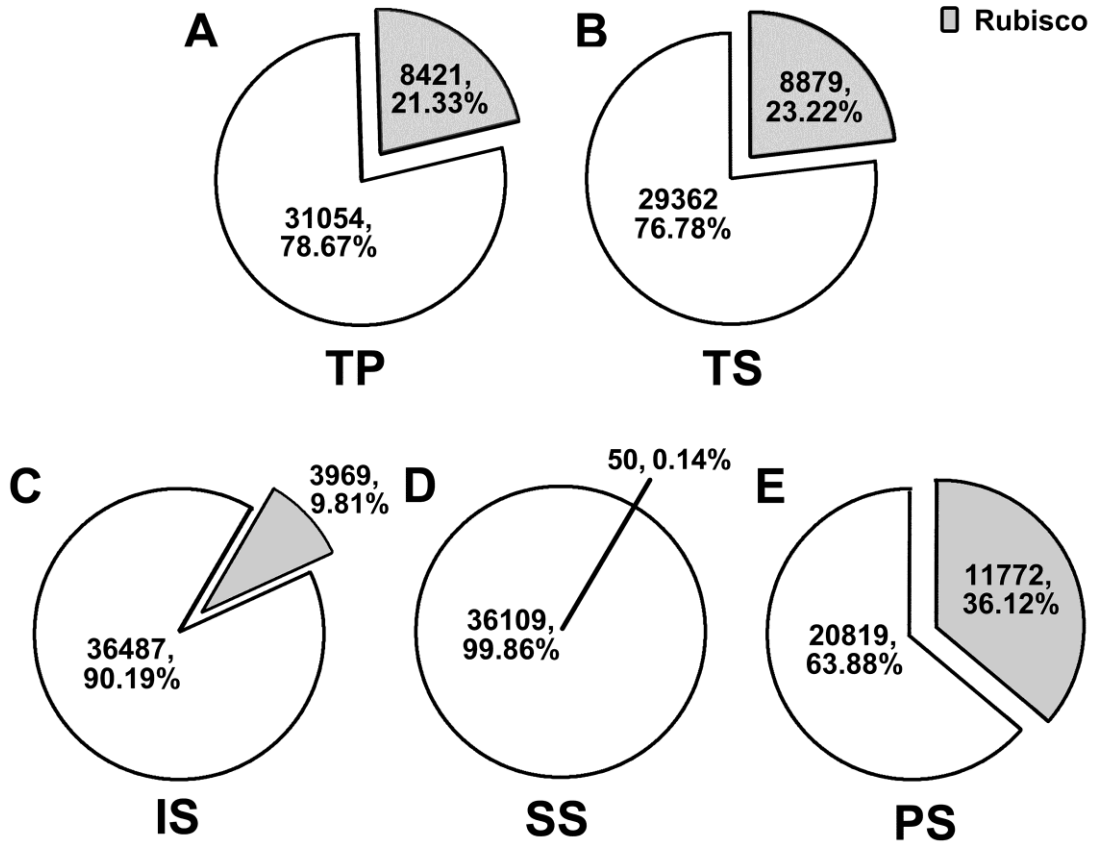

Fig. 2. 5. Rubisco removal efficiency evaluated by proteomics. The pie charts showed the percentage of spectral counts for Rubisco in total spectral counts identified in each fraction. *A*, TP. *B*, TS. *C*, IS. *D*, SS. *E*, PS.

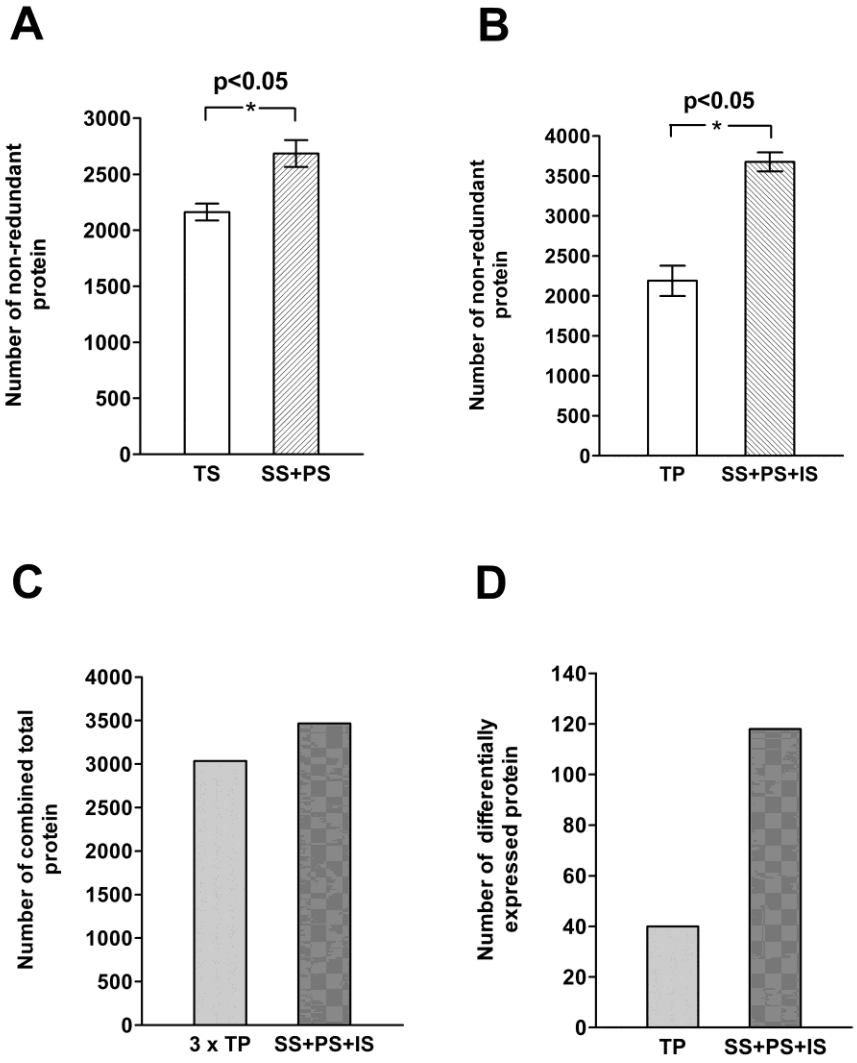

Fig. 2. 6. Improvement of protein identification and differential protein expression analysis. *A*, Comparison of the number of protein identified by TS and the combination of SS and PS. *B*, Comparison of the number of protein identification for TP and the combination of homogenizations, SS, PS, and IS. *C*, Comparison of total number of combined proteins between 3 independent TP biological samples and 1 biological samples of SS, PS, and IS. *D*, Comparison of the number of differentially expressed proteins identified in FAW larvae treated and untreated rice samples with the same fractionations of *B*. The standard errors were calculated based on duplicates of the biological samples from different factions.

together identified 1488 more proteins than TP. In other words, PARC increased the protein identification by 68.0% (Fig. 2.6*B*). Further validation was carried out by comparison of the number of proteins identified by three biological TP samples vs. the one set of SS, PS and IS fractions (Fig. 2.6*C*). This comparison is justified by the same number of LC gradient to remove the bias caused by more experimental LC-MS/MS runs for combination of the three fractions. As shown in Fig. 2.6*C*, the combined SS, PS and IS fractions still identify 431 more proteins than the three times of TP sampling. The result highlighted that the fractionation can increase protein identification, even though the exact level of improvement is hard to define.

Third and most importantly, the improved protein identification led to the better evaluation of differential protein expression. In order to further validate the effectiveness of PARC and to investigate mechanisms for plant insect interaction at the systems level, differential protein expression analysis was carried out between insect treated rice and untreated reference. The study revealed superior performance of the PARC, as shown in Fig. 2.6*D*. The comparison of TP samples between insect treated and control only identified 40 differentially expressed proteins. PARC enabled significantly improved differential protein discovery. One hundred and eighteen differentially expressed proteins were identified when SS, PS and IS fractions were analyzed. PARC therefore increased the identification of differentially expressed proteins by almost three times.

2.4.3. Overview of Differential Protein Expression for Plant Insect Interaction as Revealed by Proteomics

The systems-level overview of proteomics results included both Gene Ontology (GO) and pathway analysis. Considering that PARC significantly improved the performance of proteomics analysis, the functional proteomics analysis of rice-FAW (Fall Armyworm) interaction was focused on the PARC-based study (i.e. SS, PS and IS fractions) rather than the traditional approach (i.e. TP fraction). In fact, the PARC analysis identified much more relevant pathways than those of the traditional methods.

Gene Ontology analyses revealed that proteomics analysis based on PARC can identify much more GO terms than the reference sample (TP), assumingly due to the greater amount of proteins identified in the PARC-based analysis (Fig. 2.7). Even though both methods identified metabolic processes and responses to stimuli as major GO groups, the PARC-based analyses enabled the identification of new GO terms such as signaling and cellular component organization.

A group of differentially expressed metabolic proteins were shown in Table 2.1 with detailed functional classification of all differentially expressed proteins was shown in Table A-1. The analyses revealed that the plant responded to herbivore insect damage by up-regulating proteins involved in a broad range of processes including primary metabolism, secondary metabolism, upstream jasmonic acid biosynthesis like lipoxygenase, defense relevant proteins and others. An important aspect of metabolism changes was the carbohydrate biosynthesis and degradation. As shown in Fig. 2.8, both the starch and carbohydrate degradation pathway and sucrose biosynthesis were up-

regulated, leading to the net carbon fluxes toward both sucrose and secondary metabolite biosynthesis.

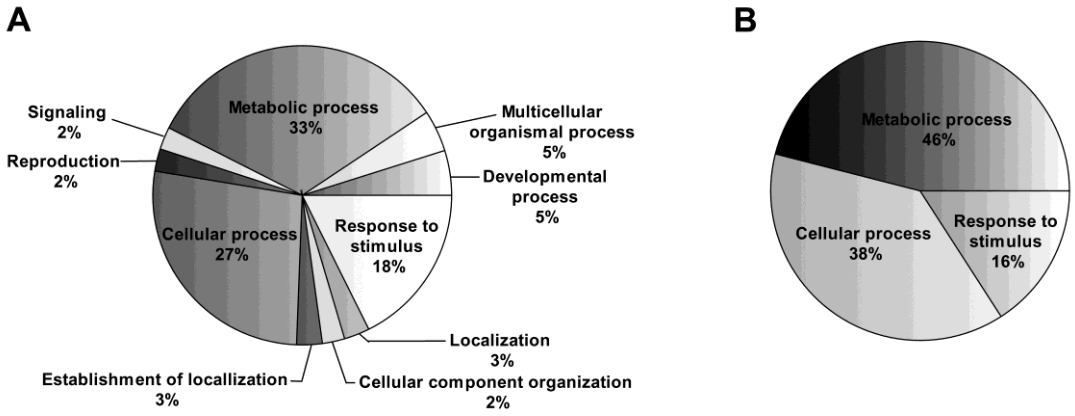

Fig. 2. 7. Gene Ontology analysis of differentially expressed protein. Pie charts shows GO distribution of (*A*) combination of SS, PS, and IS as well as (*B*) TP according to their biological processes.

Table 2. 1. Selected significantly up-regulated metabolic proteins after FAW herbivory as identified by PARC.

Gene function	Gene ID	pValue	Ratio
Carbohydrates biosynthesis and degradation			
Sucrose synthase	Os03g28330.1	0.000177	22.3
Sucrose synthase	Os06g09450.1	0.004775	8.20
Glucose-6-phosphate isomerase	Os06g14510.1	0.046986	3.94
Triosephosphate isomerase	Os09g36450.1	0.029399	2.92
2,3-bisphosphoglycerate-independent phosphoglycerate mutase	Os05g40420.1	0.032074	4.22
Enolase	Os06g04510.1	0.022889	3.57
Pyruvate kinase	Os11g05110.1	0.006019	1.98
Glyceraldehyde-3-phosphate dehydrogenase	Os02g38920.1	0.015524	2.47
Glyceraldehyde-3-phosphate dehydrogenase	Os08g03290.1	0.003714	1.79
Ribulose bisphosphate carboxylase small chain	Os12g17600.1	0.006673	5.30
Triosephosphate isomerase	Os09g36540.1	0.029399	2.92
UDP-glucose 6-dehydrogenase	Os12g25690.1	0.036053	4.38
Alpha-glucan phosphorylast isozyme	Os01g63270.1	0.003241	2.56
4-alpha-glucanotransferase	Os07g46790.1	0.031757	2.82
Amino acids biosynthesis			
5-methyltetrahydropteroyltriglutamate-homocysteine methyltransferase	Os12g42876.1	0.008758	7.21
5-methyltetrahydropteroyltriglutamate-homocysteine methyltransferase	Os12g42884.1	0.011412	7.47
Aminotransferase	Os02g55420.1	0.035815	3.94
glutamine synthetase	Os03g12290.1	0.036183	3.36
Secondary metabolites biosynthesis			
Caffeoyl-coa O-methyltransferase	Os08g38900.1	0.011654	3.71
Isoflavone reductase	Os01g01660.1	0.039708	3.20
Lipoxygenase 2.1	Os12g37260.1	0.000453	3.36

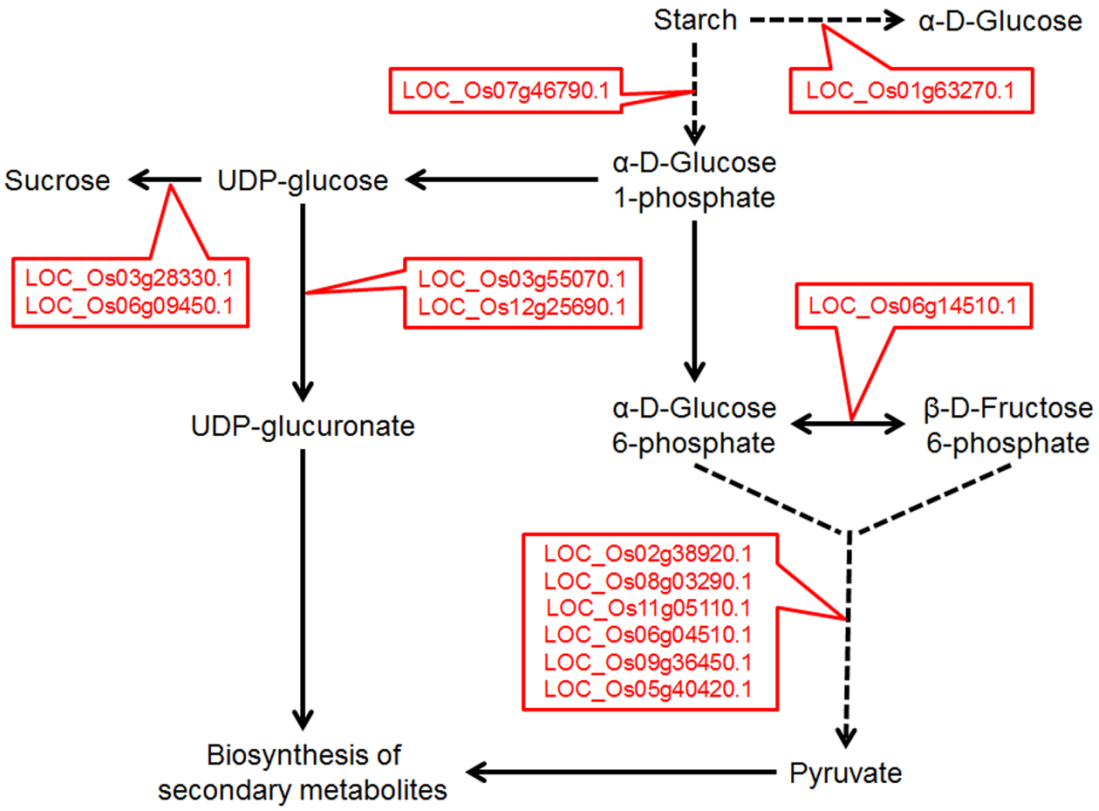

Fig. 2. 8. Pathway analysis of up-regulated proteins by FAW herbivory as revealed by PARC. The figure focused on the carbohydrate related metabolisms.

The correlation of protein and mRNA changes induced by insect damage was evaluated at both global and gene-specific levels. The comparison of this proteomics study with previous transcriptomics analysis of FAW treated rice revealed correlated mRNA and protein expression for some key secondary metabolism genes. These genes include Os10g28200, an NAD dependent epimerase/dehydratase family protein, and Os12g37260, a lipoxygenase chloroplast precursor (16). In addition, Os12g14440 (jacalin-like lectin domain containing protein) was also shown to be up-regulated by the microarray study. This gene was chosen for the downstream functional analysis. It should be noted that a previous study utilized an early version of long-oligo microarray which contains only 22,000 features and a relatively high error rate (16). A more comprehensive comparison of protein and mRNA expression can be achieved by comparing the proteomics data with the MPSS data for insect treatment (103). As shown in Table A-2 among the 88 up-regulated proteins caused by FAW treatment, 32 were up-regulated at mRNA level when treated with beet armyworm (BAW).

2.4.4. Validation of Proteomics Results by qRT-PCR

The gene-specific validation was focused on the up-regulated proteins as identified by the PARC-based proteomics analysis. Eleven up-regulated proteins were chosen for gene expression validation using real-time PCR. Most of these proteins were chosen because they are related to insect defense according to previous studies. As shown in Fig. 2.9 and Fig. 2.10 eight out of eleven genes showed correlated mRNA and protein expression levels. In other words, the triplicate qRT-PCR assays revealed that

eight genes have significantly increased mRNA expression between insect treatment and control plants at 12 hr after treatments.

Among the 8 genes verified by qRT-PCR, *Os03g48770* (cupin domain containing protein) and *Os12g14440* (jacalin-like lectin domain containing protein, JRL), were chosen for functional study using transgenic analysis. The cupin protein was exclusively identified as an up-regulated protein by the PARC approach. Both genes were up-regulated at 12 hr after insect treatment according to the qRT-PCR results (Fig. 2.9).

2.4.5. Biological Validation of Proteomics Results by Transgenic Analysis

In order to verify the gene function of *Os03g48770* and *Os12g14440*, transgenic lines have been generated to over-express the two target genes for insect treatment bioassay. The transgene overexpression was validated for T1 rice (Fig 2.11). Six leaf segments from each transgenic line and the wild-type rice were subject to insect treatment bioassay by feeding the leaves with the FAW larvae. After two days of larvae feeding, the leaf area damage was measured and calculated. The average percentages of leaf area damage were 1.73%, 2.71% and 2.48% for the *Os03g48770* over-expressing lines and 2.36%, 4.01% and 5.71% for the *Os12g14440* over-expressing lines, respectively. In contrast, the wild type rice underwent 12.19% leaf damage.

Statistical analysis of leaf area damage after FAW larvae feeding showed significant difference between the six transgenic lines and wild type rice with $p<0.01$ (Fig. 2.12). The statistical analysis were also carried out for each individual line, which

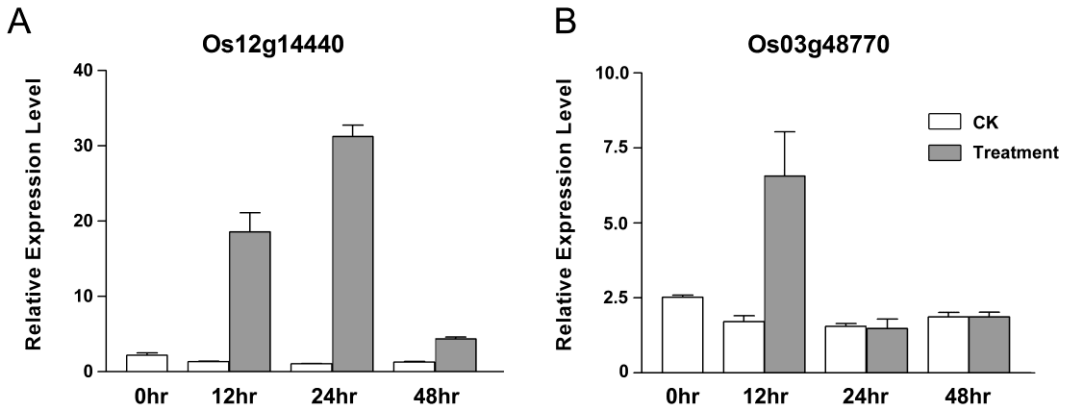

Fig. 2. 9. qRT-PCR validation of mRNA expression for selected differentially expressed proteins. qRT-PCR validation of (*A*) *Os12g14440* and (*B*) *Os03g48770* gene expression after 0hr, 12 hr, 24 hr and 48 hr treatments. The relative expression levels as compared to the lowest expressing sample in the group were derived from modified ΔΔCT method. The average ratio and 95% confidence interval were derived from triplicate assays.

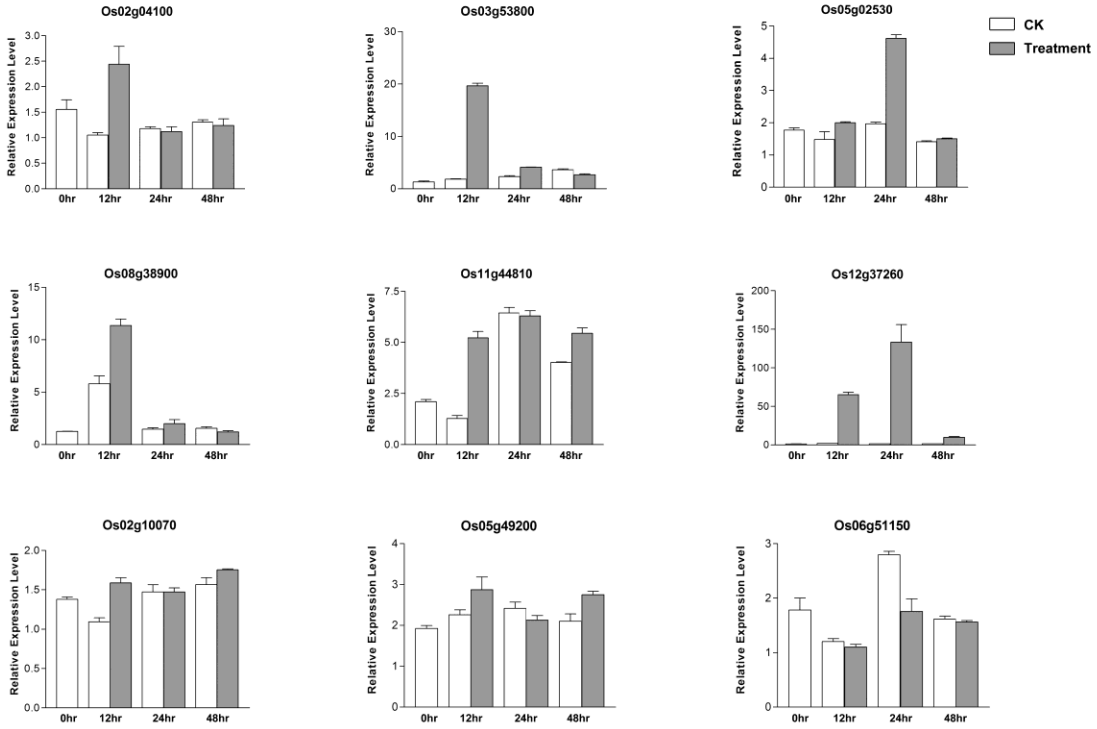

Fig. 2. 10. Quantitative RT-PCR analysis of mRNA expression level for the insect-induced response genes revealed by PARC-based proteomics. Nine up-regulated genes were selected from the proteomics result for qRT-PCR analysis of mRNA expression. This data was used to validate correlation between protein and gene expression levels. The relative expression level of mRNA for each gene at 0 hr, 12 hr, 24 hr and 48 hr was studied. 18S rRNA was used as internal control. The relative ratio and 95% confidence interval were derived from triplicate assays.

separated the samples into two groups (a and b). The samples in group b (all transgenic lines) have significantly decreased leaf area damage as compared to group a (wild type). The result indicated that the overexpression of the *Os03g48770* and *Os12g14440* in rice contributed to the resistance to the FAW larvae. More reliable functional studies need to be carried out in T3 transgenic lines in the future and is beyond the scope of PARC method development. The functional validation suggested that the differentially regulated proteins identified by PARC played important roles in insect defense.

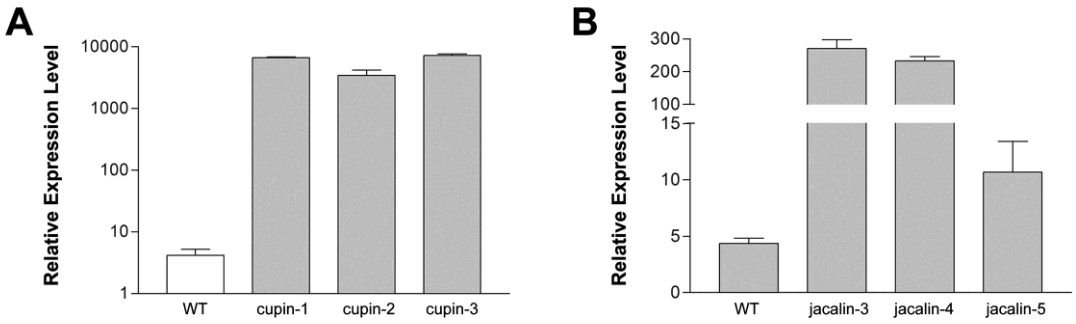

Fig. 2. 11. Quantitative RT-PCR verification of mRNA over-expression in the transgenic lines. Three individual transgenic lines of (A) the Os03g48770 (cupin-1, cupin-2 and cupin-3) and (B) the Os12g14440 (jacalin-3, jacalin-4 and jacalin-5) were verified by qRT-PCR for transgene mRNA expression levels. The relative expression levels were obtained by comparing the transgenic lines with wild type. The relative ratio and 95% confidence interval were derived from triplicate assays.

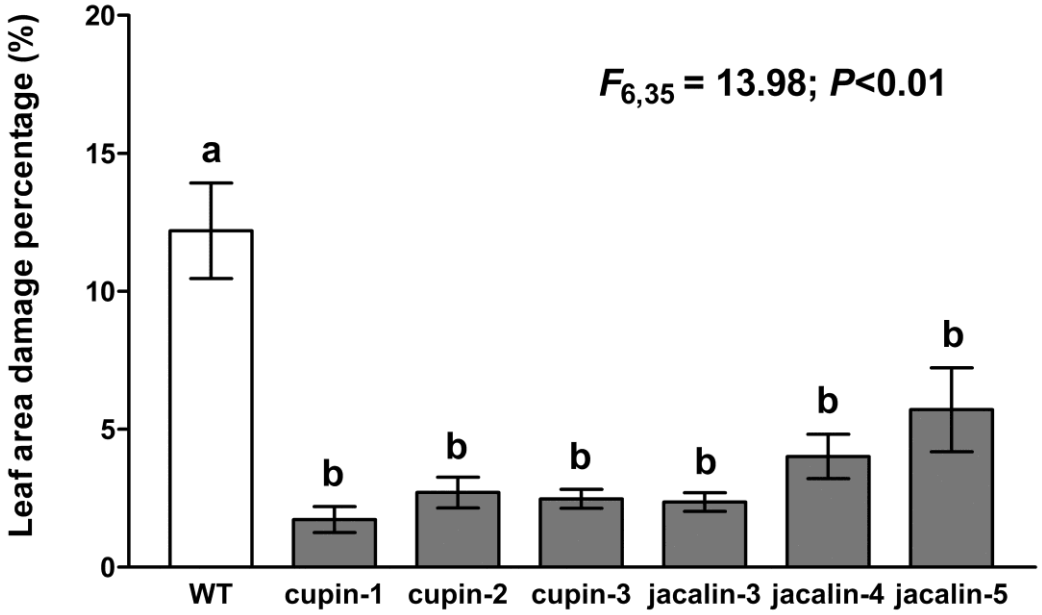

Fig. 2. 12. Comparison of leaf area damage caused by FAW larvae feeding on wild type and transgenic rice. The figure shows ANOVA analysis of three transgenic lines for both *Os03g48770* (Cupin) and *Os12g14440* (Jacalin) after FAW consumption. Wild type rice (group a) showed significant difference with transgenic lines (group b) ($F_{6,35}$=13.98; $p<0.01$). The measurement is based on bioassay of 6 leaf segments of each independent line. The average percentage of leaf area damage and 95% confidence intervals were presented.

2.5. DISCUSSION

2.5.1. PARC as an Effective Platform for Shot-gun Proteomics

Even though proteomics has emerged as a powerful platform to study systems and molecular mechanisms for biological processes like plant defense, the application of the technology is still limited by some inherent challenges. One example is that high abundance proteins will reduce sensitivity and hinder protein identification. Different strategies have been developed to address these issues caused by high abundance proteins and to improve the detection limit and performance of shot-gun proteomics. These strategies included target proteomics, fractionation, antibody-based protein removal, improved algorithm for protein identification, and others. Recently, I developed an organelle-enriched method to enhance the identification and differential expression analysis of mitochondrial proteins (described in Chapter III) (92). The method involved both organelle enrichment for sample preparation and bioinformatics classification. Even though the method improved the identification of mitochondrial proteins, the impact of Rubisco was still not properly mitigated. I hereby presented another approach to improve the sensitivity, detection, protein identification and differential expression analysis for MudPIT-based shot-gun proteomics. Moreover, the method was used to dissect the proteomic and molecular mechanisms for plant defense against herbivorous insects, leading to the identification of two key regulators in plant defense.

Several sample preparation methods were reported to deplete Rubisco from plant protein samples. These methods include sucrose density gradient centrifugation,

Ca^{2+}/phytate precipitation, and Rubisco antibody column (78, 104-107). Despite progress, these procedures were either labor intensive, time consuming, costly, or leading to limited improvement in protein identification. I therefore developed PARC as a new strategy to combine the Rubisco removal and protein fractionation. In PARC, the total protein was first separated into two fractions with a nonionic detergent. The two fractions were a total soluble protein (TS) and an insoluble protein (IS). The TS fraction was then treated with PEI to derive two other fractions as Rubisco-removed soluble protein (SS) and PEI precipitated fraction (PS). PEI is a polymer with repeating units that can contain primary, secondary and tertiary amino groups. The negative charge of the molecule can selectively precipitate positively charged proteins like Rubisco. The results indicated that PEI could be more selective for Rubisco removal than ammonium sulfate, which is another common compound for Rubisco precipitation.

In the PS fraction, proteins other than Rubisco were also precipitated. In Fig. 2.5*E*, the spectra count from Rubisco composes 36.12% of the total spectra count in PS fraction, which suggested co-precipitation of other positively charged proteins. However, since the PEI precipitated fraction will also be subjected to the shotgun proteomics analysis, other co-precipitated proteins will still be included in the identification process, so that no extra information will be lost due to Rubisco removal. These studies eventually led us to the finding that the integration of fractionation and Rubisco removal by PEI led to the detection of more than 3,500 proteins in the rice proteome. This translated into 68% increase of the number of proteins identified using a commercial kit without any fractionation. The number of identified proteins was also

significantly higher than those previously published (108, 109). More importantly, the PARC method essentially tripled the number of differentially regulated proteins identified in the plant insect interaction study.

The identification of more plant proteins may also benefit from other novel approaches such as combining proteomic and genomic forces together (110). Briggs' lab has identified novel proteins by using the unique peptide sequence information from a comprehensive proteomic survey of the *Arabidopsis* proteome. The new proteo-genomics and bioinformatics approach has led to discovery of more than 10,000 novel peptides. This suggests the current state-of-the-art *Arabidopsis* genome is still incomplete. The informatics-intensive proteo-genomics research can be further pursued in future research.

Nevertheless, considering the limitation of current proteomics platforms, fractionation is still one of the viable alternatives to expand the dynamic range of proteomics detection. The disadvantage of the fractionation approach is prolonged analysis time and increased reagent cost. However, the gain of more protein identification and in-depth differential protein expression analysis in this study could compensate the tripled instrumentation time and the increased data analysis time, because the proteomic setup and data analysis are heavily based on computerized and instrumentation-based platform. The principle of abundant protein precipitation and fractionation can be adapted and applied to studies beyond the plant species as PEI can be used to precipitate acidic proteins in general. More importantly, the improved protein identification led to significantly improved differential protein expression analysis,

enabling much more in-depth systems and molecular mechanism analysis in the rice-FAW interaction study. The increased number of differential expressed proteins led to many more GO terms being discovered, more comprehensive pathway annotation, and therefore a more in-depth understanding of mechanisms for plant defense.

2.5.2. Systems and Molecular Mechanisms of Plant Defense against Herbivorous Insects as Revealed by PARC

As aforementioned, the PARC-based analysis enabled a greater comprehensive proteome profiling than the traditional approach, delivering in-depth mechanisms from the systems levels to the molecular levels. The GO and protein classification indicated the up-regulation of pathways involved in secondary metabolism, primary metabolism, defense, and other processes. The data correlates with the fact that plants respond to herbivore attack through many dimensions, such as direct and indirect, physical and chemical, constitutive and inducible defenses (27, 29, 55, 111, 112). The proteomics data suggested a complex, yet surprising protein-level regulation of carbon utilization. As shown in Fig. 2.8 the starch degradation enzymes were up-regulated, which could potentially lead to increased production of α-D-Glucose-1P. Furthermore, the α-D-Glucose-1P could be used for either glycolysis and secondary metabolite production or sucrose biosynthesis. Some enzymes leading to the channeling of carbon flux to secondary metabolites were up-regulated. In addition, enzymes involved in sucrose biosynthesis were also up-regulated. Previous studies suggested that sucrose biosynthesis was up-regulated during plant defense against pathogen and insects in a

wide range of crop and model plant species, including rice, cotton, tomato, and *Arabidopsis* (30, 113-119). The study correlated with previous studies indicating the important roles for sucrose in plant defense against herbivorous insects. In addition, as compared to our previous transcriptomic analysis, the proteomic analysis successfully identified new up-regulated pathways in carbohydrate metabolism during defense (16). Overall, active changes in carbon metabolism seemed to be a major feature of plant defense against herbivorous insects. Further validation of the hypothesis will rely on metabolite analysis, which is a major focus in plant defense field.

The study also revealed the correlations of transcriptomic and proteomic regulations. Even though the comparison of previous microarray data and this proteomics analysis only revealed a few over-lapping genes, the comparison of the more comprehensive MPSS analysis with this study revealed almost 36% correlation between transcriptomic and proteomic up-regulated genes (Table A-2). The correlation is significant as several factors contributed to the differences between transcriptomics and proteomics dynamics. First, the insect species, treatment conditions, and plant growth stages were all different between the previous transcriptomics and current proteomics studies. Second, as shown in the qRT-PCR experiments (Fig. 2.10), the mRNA level changes are often transient and dynamic, which might not correlate with protein level changes at a given point. Third, several biological processes including RNA degradation, protein translation, and protein stabilization could lead to the differences between protein and gene expression level. In spite of these key issues, I still found that 32 out of 88 up-regulated proteins were also up-regulated at mRNA levels. The study indicated

that transcriptional regulation of plant defense is crucial in regulating the defense-related protein level. Despite the significant correlation of up-regulated genes and proteins, the down-regulated proteins are not well correlated with mRNA level down regulation. This could be due to the dynamic protein degradation during defense process.

2.5.3. Functional Validation of Plant Defense Genes as Revealed by PARC

Besides the systems and pathway level analysis, further functional verification was carried out by transgenic functional study of two genes identified from the proteomics study. The *Os03g48770* (cupin-like protein) and *Os12g14440* (jacalin-like lectin protein) over-expressing lines showed improved resistance against rice herbivore FAW. These two genes represented new regulators for rice defense. Jacalin-related lectins (JRLs) can bind specifically with different carbohydrates to regulate biological processes (120). Previous studies showed that several JRLs were up-regulated significantly during insect defense and it was suggested that binding with foreign glycans might be important for the defense process (121). Other research has indicated that JRLs might regulate the signal transduction by modulating plant protein-carbohydrate interactions (122). For example, JLRs can regulate the size of *Arabidopsis* beta-glucosidase PYK10 complex antagonistically to retain the activity against pathogens instead of losing the enzyme activity (123). For the phytophagous insects, JRLs can bind with sugars on the gut epithelial cells and therefore inhibit nutrient uptake of insects (124). I hereby showed that a specific JRL induced by insect treatment can

promote the rice defense against FAW larvae. However, the detailed mechanisms for this JRL to regulate plant defense need further investigation.

Os03g48770 belonged to the cupin super gene family and was identified by the PARC method only. The members of this superfamily have very diverse functions and their roles in plant disease defense have been studied in several plants species (87, 125, 126). In particular, one of the cupin superfamily members was proven to be an oxalate oxidase which produces H_2O_2 and can serve as an important signal in plant disease defense (127). In this study, I showed that insect treatment led to the over-expression of a member of rice cupin superfamily genes and the over-expression of this gene resulted in significantly improved defense against FAW. The further phylogenic analysis indicated that this particular cupin gene is a homology of a wheat oxalate oxidase, GER3 (128, 129) (Fig. 2.13). *Os03g48770* also shares 97% similarity with *Os03g48780* (Fig. 2.14), which encodes an oxalate oxidase involved in ROS signaling (16). It will be interesting to further study if this gene is involved in rice defense against both insect and pathogen, and how the up-stream and down-stream regulation is involved. The *Os03g48770* gene has the potential to be directly used in crop improvement. Even though the F1 transgenic lines have shown promising insect resistance for both genes, extensive research in homozygous T3 transgenic lines need to be carried out to validate if these genes can be used in engineering crop resistance to herbivorous insects.

Overall, PARC represented an effective method to improve the sensitivity, protein identification and differential expression characterization in shot-gun proteomics studies. The method was implemented to identify both novel mechanisms and important

regulators for plant defense against herbivorous insects. The technique can be broadly applied to proteomics studies for new gene discovery and mechanism elucidation.

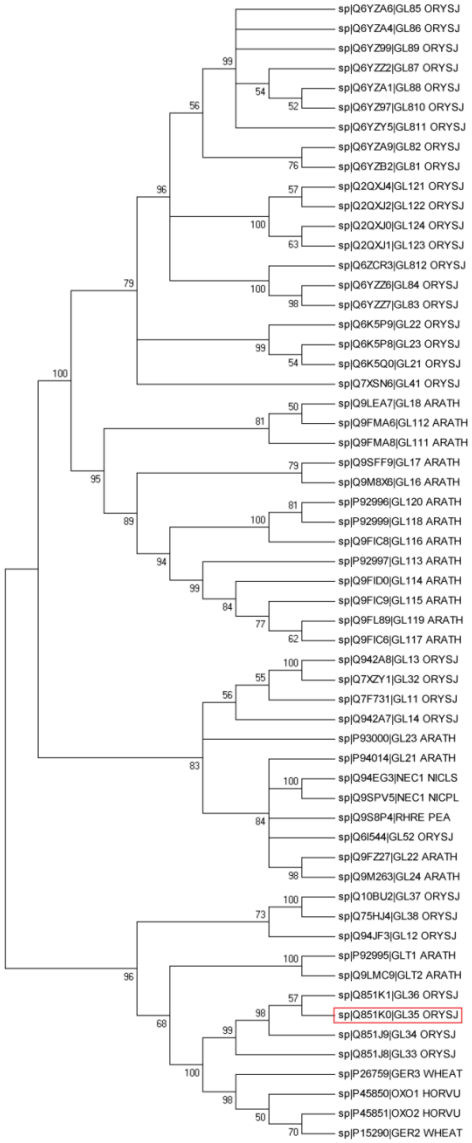

Fig. 2. 13. Phylogenetic analysis of *Os03g48770* homologs. The phylogenetic tree was built for *Os03g48770* homologs across multiple monocot species. Protein in red box was the query protein.

Fig. 2. 14. Sequence alignment of *Os03g48770* homologs. The sequence alignment of wheat oxalate oxidase GF-3.8 (P26759), rice germin-like protein 3-6 (Os03g48780, Q851K1) and the cupin domain containing protein (Os03g48770, Q851K0) were also presented.

CHAPTER III

INTEGRATION OF SHOT-GUN PROTEOMICS AND BIOINFORMATICS ANALYSIS TO EXPLORE PLANT HORMONE RESPONSES*

3.1. SUMMARY

Multidimensional protein identification technology (MudPIT)-based shot-gun proteomics has been proven to be an effective platform for functional proteomics. In particular, the various sample preparation methods and bioinformatics tools can be integrated to improve the proteomics platform for applications like target organelle proteomics. I have integrated a rapid sample preparation method and bioinformatics classification system for comparative analysis of plant responses to two plant hormones, zeatin and brassinosteroid (BR). These hormones belong to two distinct classes of plant growth regulators, yet both can promote cell elongation and growth. An understanding of the differences and the cross-talk between the two types of hormone responses will allow us to better understand the molecular mechanisms and to identify new candidate genes for plant engineering.

As compared to traditional organelle proteomics, the organelle-enrichment method both simplifies the sample preparation and increases the number of proteins identified in the targeted organelle as well as the entire sample. Both zeatin and BR

induce dramatic changes in signaling and metabolism. Their shared-regulated protein components indicate that both hormones may down-regulate some key components in auxin responses. However, they have shown distinct induction and suppression of metabolic pathways in mitochondria and chloroplast. For zeatin, the metabolic pathways in sucrose and starch biosynthesis and utilization were significantly changed, yet the lipid biosynthesis remained unchanged. For BR, lipid biosynthesis and β-oxidation were both down-regulated, yet the changes in sucrose and starch metabolism were minor. In this chapter, a rapid sample preparation method and bioinformatics classification for effective proteomics analysis of plant hormone responses was presented. The study highlighted the largely differing response to zeatin and brassinosteroid by the metabolic pathways in chloroplast and mitochondria.

3.2. INTRODUCTION

Proteomics can directly address many biological questions by revealing the abundance of certain proteins within organisms. Traditionally, two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) was the golden standard for proteomics analysis, yet the platform is limited by both protein identification and quantification capacities. The recent advances in mass spectrometry instrumentation, separation methods, data acquisition and analysis tools have enabled use of the so-called ‘shot-gun’ proteomics. It uses tandem mass spectrometry and the multidimensional protein identification technology (MudPIT) (2). In the MudPIT platform, the whole proteome is directly digested with protease, and the resulting peptides are subjected to

multidimensional chromatography separation. The separated peptides are then analyzed online by mass spectrometry. The so called MudPIT platform eliminates the tedious gel separation and has been broadly applied in plant biology studies (4, 6). Even though the platform has superior performance as compared to 2-D gel platforms, limitations still exist for several reasons. First, profiling the whole proteome is complicated by the complexity of the protein sample, the number of proteins expressed, the differing molecular weights, and other variations in chemical and physical characteristics(10, 11). Also, many functional proteins such as GTPases, kinases and phosphatases exist in low abundance. Their signals can be easily masked by highly abundant proteins such as ribulose 1, 5-bisphosphate carboxylase/oxygenase (Rubisco) (12). These challenges can be addressed by improving sample preparation methods, bioinformatics analysis, sample processing, and mass spectrometry instrumentation. I hereby present the integration of a rapid sample preparation method with bioinformatics analysis to achieve better peptide identification and focused study of chloroplast and mitochondrial proteins.

I am particularly interested in chloroplast and mitochondria because the two organelles are important for energy metabolism and plant growth, among many other functions. In particular, the proteome dynamics of these two organelles in response to growth relevant hormones like auxin, cytokine, and brassinosteroid will shed light onto the mechanisms for plant hormone responses. It will also identify candidate genes for improving crop seed and biomass yield for food, fiber and energy usages.

Traditionally, in order to identify proteins in a particular type of organelle, the organelle is separated by gradient density centrifugation and ultra-centrifuged from a

large quantity of initial samples (12, 130). Proteomics studies toward specific organelles have been performed on nuclei, mitochondria, chloroplasts, Golgi apparatuses, and endoplasmic reticulum, etc. (11, 131). For example, Dunkley and colleagues used localization of organelle proteins by isotope tagging (LOPIT) to simultaneously localize 527 proteins out of 689 proteins identified in several organelles of Arabidopsis (132). Most of the traditional organelle purification involves time-consuming and tedious separation steps, which could introduce extra errors (133). I hereby simplified the traditional method of organelle separation by implementing a rapid centrifugation step. The rapid sample preparation method integrated with bioinformatics classification was evaluated as an alternative to study mitochondria and chloroplast proteomics in plant's responses to growth hormones.

The proteomics analysis of hormone responses is part of long-term efforts in the lab to identify important genes involved in plant biomass increases for bioenergy purposes. Several plant hormones such as auxin and gibberellic acid can promote plant growth and are known to be able to increase plant biomass accumulation through different mechanisms. Among these plant hormones, zeatin and BR are of particular interest to us. Zeatin is a plant hormone belonging to cytokinins and regulates plant development and growth. Zeatin has been widely applied in agriculture to increase fruit or seed size and is well known to promote cell elongation and root development (134). Interestingly, BR is also known to be able to promote plant growth through cell elongation (135). Even though both plant hormones can promote cell elongation and growth, the underlying mechanisms are believed to be widely different; the hormone

signaling pathways for the two are unique to one another. However, very few studies have focused on studying the differences and cross-talk between the responses in the two hormones at the proteome level. I hereby utilized the aforementioned platform to explore the proteome responses of Arabidopsis in response to treatment by the two plant hormones. The metabolic pathways in chloroplasts and mitochondria are of particular interest.

Overall, in this chapter, I have integrated a simple sample preparation method with bioinformatics classification to analyze plant responses to zeatin and BR. The new method has been shown to improve protein identification, in particular in mitochondria. Using this platform, I have revealed that both zeatin and BR induce significant changes in signaling and metabolism. The shared regulated protein components indicated that both hormones may down-regulate some components in auxin responses. However, the two plant hormones have shown distinctive induction and suppression of metabolic pathways. For zeatin, the metabolic pathways in sucrose and starch biosynthesis and utilization were significantly up-regulated, yet the lipid biosynthesis remained unchanged. For BR, the lipid biosynthesis and β-oxidation were both down-regulated, yet the changes in sucrose and starch metabolism are minor. These differences highlight the different molecular and metabolic mechanisms for response to zeatin and BR. The data can help us to design better strategies to promote plant biomass accumulations.

3.3. MATERIALS AND METHODS

3.3.1 Plant Material and Growth Conditions

Arabidopsis thaliana ecotype Col-0 was used. Seeds were stratified at 4℃ to synchronize germination for 2 days and then grown at 23℃/19℃ under a 12 h/12 h light/dark cycle for 4 weeks.

3.3.2. Hormone Treatment

Zeatin (Sigma-Aldrich, St. Louis, MO) and 24-epibrassinolide (PhytoTechnology Laboratories, Shawnee Mission, KS) were sprayed at 100 μm and 0.5 mg/L, respectively. 0.8% methanol solution was sprayed as mock. The aerial parts of plant were collected at 24 hours after the spray.

3.3.3. Plant Total Protein Isolation

For the total protein isolation, a plant total protein extraction kit (Sigma-Aldrich, St. Louis, MO) was used, and the entire procedure followed the manufacturer's manual. 150 mg of aerial tissue from Arabidopsis leave was collected and ground in liquid nitrogen to a fine powder. Pre-cooled methanol solution with protease inhibitor was added to the powder and vortexed for 30 seconds. The mixture was incubated at -20℃ then centrifuged at 16,000×g for 5 minutes at 4℃. Supernatant was removed and the pellet was washed by methanol solution for two more times. The resulting pellet was washed by pre-cooled acetone and centrifuged at 16,000×g for 5 minutes at 4℃. SpeedVac was used to remove residual acetone and Reagent Type 4 Working Solution

provided by the kit was used to incubate the pellet for 15 minutes at room temperature. The pellet was then centrifuged at 16,000×g for 30 minutes, and supernatant was collected and stored at -80℃ for future proteomics use (See Fig. 3.1).

3.3.4. Organelle Enrichment and Protein Isolation

The organelle enrichment procedure was developed based on the method from Santoni (136) with some modification. Five gram of fresh aerial tissue of Arabidopsis was collected and washed by ice-cold water to remove the soil. A blender was used to disrupt the tissue after adding a 2:1 (mL medium/g fresh weight) homogenization buffer (50 mM TRIZMA base, 500 mM Sucrose, 10% Glycerol, 20 mM EDTA-Na_2, 20 mM EGTA, 50 mM NaF, 5 mM beta-glycerophosphate, 1 mM phenantroline, 0.6% PVP40, 10 mM ascorbic acid, 1 mM leupeptin, 5 mM DTT, 1 mM Na-orthovanadate, pH 8.0 adjusted by ES). The homogenate was then filtered through Miracloth to remove plant debris. Centrifugation of filtered homogenate was conducted at 1,000×g for 5 minutes to remove the nuclei. The supernatant was then centrifuged at 26,000×g for 25 minutes to pellet organelles.

For protein isolation of enriched organelles, pre-cooled methanol with protease inhibitor was added to the organelle-enriched pellet, which was collected after the centrifugation described in the Organelle enrichment section. The sample was then vortexed for 30 seconds. The mixture was incubated at -20℃ then centrifuged at 16,000×g for 5 minutes at 4℃. The supernatant was removed and the pellet was washed twice by methanol solution. The resulting pellet was again washed by pre-cooled acetone

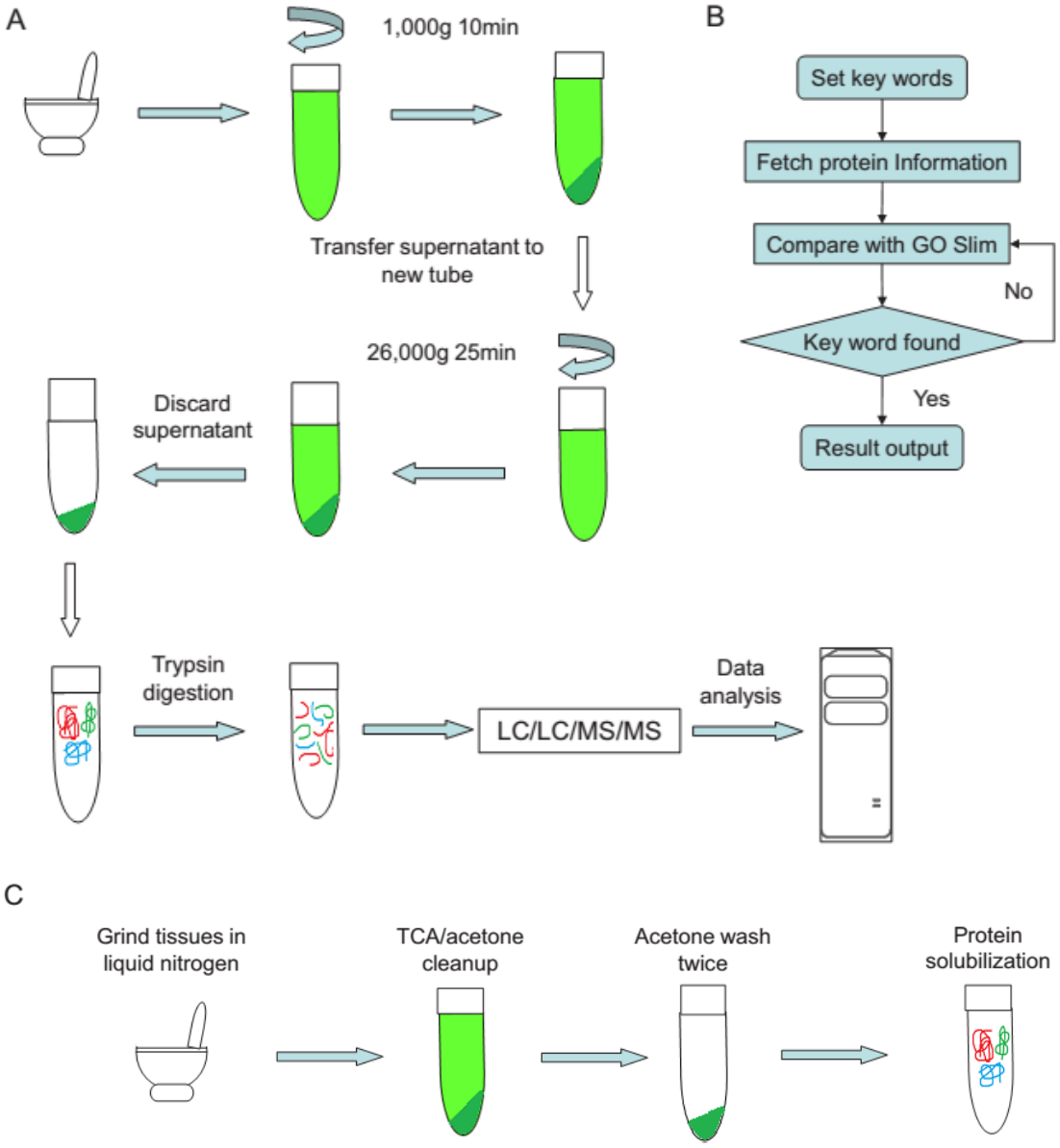

Fig. 3. 1. Workflow of sample preparation and bioinformatics analysis.
Additional file 1A shows the workflow of organelle enrichment, protein isolation and 2D LC/MS/MS. Additional file 1B illustrates the flow of protein classification package. Additional file1C shows workflow of traditional plant total protein isolation by TCA/acetone.

and centrifuged at 16,000×g for 5 minutes at 4℃. Residual acetone was removed by SpeedVac, and Reagent Type 4 Working Solution was used to incubate the pellet for 15 minutes at room temperature. The pellet was then centrifuged at 16,000×g for 30 minutes, and supernatant was collected and stored at -80℃ for future proteomics use (See Fig. 3.1*A*).

3.3.5. MudPIT

MudPIT-based shot-gun proteomics was carried out to analyze each sample. Approximately 100 µg of protein was digested by Trypsin Gold, Mass Spectrometry Grade (Promega, WI, USA) with 1:40 w/w at 37 ℃ for 24 h. The digested peptides were desalted using a Sep-Pak plus C18 column (Waters Limited, ON, Canada) and then loaded onto a biphasic (strong cation exchange/reversed phase) capillary column using a pressure tank. The 2D back column was composed of 5 cm of C18 reverse phase resin and 3 cm of strong cation exchange (SCX) resin. The back column was then connected to a 15-cm-long 100 um-ID C18 column (packed in house with the same C18 reverse phase in the back column) and sprayed through a SilicaTip (New objective, Inc, Woburn, MA). The two-dimensional liquid chromatography separation and tandem mass spectrometry conditions followed the protocols previously described by Washburn et al. (91). Before SCX separation, a 1 h RP gradient from 100% Solvent A (95% H_2O, 5% ACN, and 0.1% formic acid) to 100% Solvent B (30% H_2O, 70% ACN, and 0.1% formic acid) was configured to move peptides from C18 resin to SCX resin in the back column. The SCX LC separation was performed with eleven salt pulses containing

increasing concentrations of ammonium acetate. Each salt pulse was followed by a 2 h reverse phase gradient from 100% Solvent A to 60% Solvent B. The LC eluent was directly nanosprayed into a linear ion trap mass spectrometer, Finnigan LTQ (Thermo Fisher Scientific, San Jose, CA). The mass spectrometer was set to the data-dependent data acquisition mode, and full mass spectra were recorded on the peptides over a 300-1700 m/z range, followed by five tandem mass (MS/MS) events for the most abundant ions from the first MS analysis. The Xcalibur data system (Thermo Fisher Scientific, San Jose, CA) was used to control the LC-LTQ system and collect the data.

3.3.6. Data Analysis

Tandem mass spectra were extracted from the raw files and converted into the MS2 file. The MS2 file was searched against the Arabidopsis protein database downloaded from The Arabidopsis Information Resource (TAIR); it contains reverse sequence and common contaminant proteins. A ProLuCID algorithm was used to search for data using the Texas A&M Supercomputing Facility. The validity of peptide/spectrum matches was assessed in DTASelect2.0 using a 0.05 false discovery cutoff, with a cross-correlation score (XCorr) that's larger than 1, and normalized difference in cross-correlation scores (DeltaCN) larger than 0.08. Proteins with more than two peptides were identified as detected and were recorded.

3.3.7. Ontology and Pathway Analysis

PatternLab (94) software is used for data analysis to discover differentially expressed proteins. The cutoff of p-value and Fold-change is 0.05 and 2.0 respectively. Gene ontology annotations for proteins were performed by VirtualPlant (137). The pathway analysis of proteins differentially expressed was analyzed by Aracyc (http://www.arabidopsis.org/biocyc/). Cluster analysis was carried out by MeV(138).

3.3.8. Protein Classification Software

A python package was developed to parse proteins based on their GO keywords (See Fig. 3.1*B*). The report containing differentially expressed proteins searched each protein ID against GO Slim, which can be downloaded from ftp://ftp.arabidopsis.org/home/tair/Ontologies/Gene_Ontology/. If the annotation of the protein matches the keyword set, the ID will be output to a text file and the number of matched protein ID will be displayed.

3.4. RESULTS

3.4.1. The Organelle Enrichment Method Improved Total and Mitochondrial Protein Identification

The protein identification and mass spectra were compared between samples prepared by the organelle enrichment and traditional methods. The average number of proteins identified from the organelle enrichment samples and those from total protein isolation was 3099 and 2897, respectively (shown in Table 3.1). The average identified peptide increased from 20128 to 21547. The average spectra count increased 23.44%, from 55565 to 68588. The pair-wise student's t-test showed a significant difference for the number of peptides and spectra count between the organelle enriched sample and the traditional protein sample (Table 3.1). However, there was not a significant difference in the number of identified proteins. This is probably due to the dynamic range of the gel-free shotgun proteomics platform, which determines the detection up-limit of the described platform. The protocol used here presents a digestion of about 100 μg of the total protein, tryptic digestion, and chromatography separation. Prior research with a similar platform also reported similar protein identification numbers of the global proteome profiling and/or similar peptide counts and spectra counts (139, 140).

I further processed the proteomics data with protein classification software. The analysis indicated the organelle enrichment method has identified over 30% more mitochondrial proteins, even though the chloroplast protein identification didn't change significantly. As compared to the traditional method, the organelle enrichment method led to a greater percentage of mitochondrial protein identified (Table 3.1). This suggests

Table 3. 1. Improved protein identification using the organelle enrichment method (OEM) as compared traditional method (TM)

	Protein Identified*	**Peptide IDs***	**Spectra Count***	**Mitochondrial proteins**	**Percentage of mitochondrial proteins (%)**
OEM 1	2956	21283	73181	228	7.71
OEM 2	3121	21848	68127	236	7.56
OEM 3	3221	21511	64458	201	6.24
TM sample 1	2880	20386	64733	179	6.22
TM sample 2	2732	19939	52752	170	6.22
TM sample 3	3081	20061	49211	181	5.87
The pair-wise student's t-test	0.0845	0.0295	0.04	<0.01	

*All of the data are filtered and forward matches

the protocol has effectively enriched the proteins in mitochondria for the research purpose. Overall, with this simplified sample preparation method, I successfully enriched mitochondrial protein and identified more proteins that are involved in energy metabolism. The integration of bioinformatics classification allowed us to focus more on mitochondrial pathways. I therefore used this method to explore the proteome dynamics during plant hormone responses.

3.4.2. Overview of Zeatin and BR-Regulated Proteins

As aforementioned, I focused on comparing zeatin and BR treated *Arabidopsis thaliana* Col-0 plants with wild-type plants. A total of 267 proteins were up-regulated and 88 were down-regulated in the zeatin-treated sample. A total of 60 up-regulated and 228 down-regulated proteins were identified in BR treated samples. These proteins could be classified into several groups based on the biological process category of GO. As shown in Fig. 3.2*A*, most zeatin-triggered proteins are involved in cellular processes (36%), compared with metabolic processes (25%), response to stimuli (15%), developmental processes (6%), cellular component organization or biogenesis (6%), biological regulation (5%), etc. Fig. 3.2*B* shows the category percentage of up-regulated proteins in BR treated plants: cellular processes (31%), metabolic processes (26%), developmental processes (9%), cellular component organization or biogenesis (7%), response to stimuli (6%), biological regulation (5%), etc.

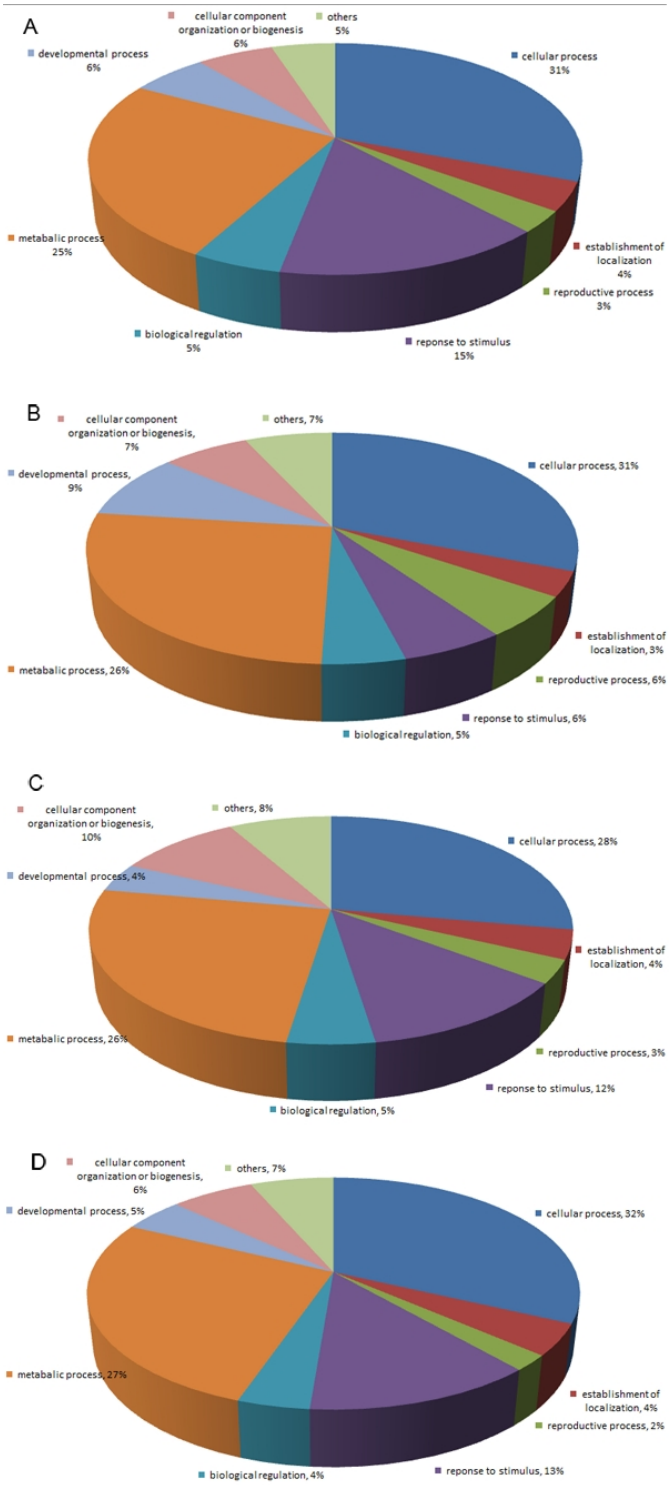

Fig. 3. 2. Pie charts of GO distribution of up-regulated proteins in zeatin (A) and BR (B) as well as down-regulated proteins in zeatin (C) and BR (D) treated Arabidopsis according to their biological process.

Among the differentially regulated proteins between the two hormone treatments, I particularly focused on the shared genes as shown in Table 3.2. Among all of the differentially expressed proteins present in both of the zeatin and BR treated samples, a total of 12 proteins were up-regulated in both treatments. These proteins include DEAD/DEAH box helicase (AT3G18600), 5'-adenylylsulfate reductase 2 (APR2, AT1G62180), protoporphyrinogen oxidase (putative, AT5G14220), co-chaperone grpE family protein (AT5G17710), and others. A second group contains the 35 genes that are down-regulated in both samples. This group contains sulfate adenylyltransferase 4 (APS4, AT5G43780), auxin-binding protein 1 (ABP1, AT4G02980), jacalin lectin family protein (AT2G33070), plastid-lipid associated protein (PAP, AT2G46910) and oligosaccharyl transferase STT3 subunit family protein (STT3A, AT5G19690) and others. Only two genes were found with opposite regulation. AT2G40360 was up-regulated in the zeatin-treated sample but down-regulated in BR treated samples. Kiba and colleagues found this gene up-regulated after cytokinin treatment, confirming this result (141).

3.4.3. Cluster Analysis of Zeatin and BR Treated Sample

Besides the differentially regulated genes, I further carried out two types of global analysis, the cluster analysis of protein abundance based on normalized mass spectra counts and the pathway analysis of differentially regulated proteins. Fig. 3.3 shows the overview of cluster analysis; it revealed a dynamic proteome profile among the wild type, the zeatin treated sample, and the BR treated sample. It also revealed that

Table 3. 2. Shared differentially expressed proteins between zeatin and BR treated samples

Gene Locus	Fold change (zeatin)	Fold change (BR)	Description
AT3G53520	15.44	9.79	NAD-dependent epimerase/dehydratase family protein
AT5G17710	9.07	6.72	co-chaperone grpE family protein
AT2G33430	8.45	5.69	plastid developmental protein DAG, putative
AT3G18600	8.36	3.52	DEAD/DEAH box helicase, putative
AT1G26340	7.74	4.54	cytochrome b5, putative
AT5G44320	7.70	3.52	eukaryotic translation initiation factor 3 subunit 7
AT5G08260	6.49	4.09	serine carboxypeptidase S10 family protein
AT1G62180	5.83	3.52	5'-adenylylsulfate reductase 2
AT5G18280	5.83	3.07	apyrase (APY2)
AT5G38990	4.49	3.07	protein kinase family protein
AT5G28050	2.12	2.49	cytidine/deoxycytidylate deaminase family protein
AT5G14220	2.10	2.01	protoporphyrinogen oxidase, putative
AT2G40360	2.09	-3.33	transducin family protein/WD-40 repeat family protein
AT1G52410	-2.10	-2.15	caldesmon-related
AT1G60420	-2.33	2.05	DC1 domain-containing protein
AT4G00620	-2.33	-2.93	tetrahydrofolate dehydrogenase/cyclohydrolase, putative
AT1G49820	-2.33	-2.93	5-methylthioribose kinase family
AT1G55690	-2.33	-2.93	SEC14 cytosolic factor family protein/phosphoglyceride transfer family protein
AT4G02980	-2.65	-3.33	auxin-binding protein 1 (ABP1)
AT1G66070	-2.65	-3.33	translation initiation factor-related
AT4G16580	-2.65	-3.33	expressed protein
AT5G33320	-2.65	-3.33	triose phosphate/phosphate translocator, putative
AT1G06650	-2.65	-3.33	2-oxoglutarate-dependent dioxygenase, putative
AT4G30840	-2.65	-3.33	WD-40 repeat protein
AT1G05560	-2.65	-3.33	UDP-glucose transferase (UGT75B2)
AT3G14010	-2.65	-3.33	hydroxyproline-rich glycoprotein family protein
AT2G43160	-2.65	-3.33	epsin N-terminal homology (ENTH) domain-containing protein
AT2G32810	-2.65	-3.33	beta-galactosidase, putative/lactase, putative

Table 3.2. Continued

Gene Locus	Fold change (zeatin)	Fold change (BR)	Description
AT2G38000	-2.97	-3.73	chaperone protein dnaJ-related
AT5G40170	-2.97	-3.73	disease resistance family protein
AT2G34680	-2.97	-3.73	AIR9
AT4G24090	-3.10	-3.91	expressed protein
AT1G16860	-3.10	-3.91	merozoite surface protein-related
AT5G23210	-3.10	-3.91	SCPL34, similar to serine carboxypeptidase S10 family
AT3G56130	-3.29	-4.13	biotin/lipoyl attachment domain-containing protein
AT1G53590	-3.29	-4.13	C2 domain-containing protein
AT2G33070	-3.88	-4.89	jacalin lectin family protein
AT2G28760	-4.06	-5.11	NAD-dependent epimerase/dehydratase family protein
AT4G26555	-4.06	-5.11	immunophilin/FKBP-type peptidyl-prolyl cis-trans isomerase family protein
AT3G63150	-4.06	-5.11	GTP-binding protein-related
AT2G46910	-4.20	-5.29	plastid-lipid associated protein PAP/fibrillin family protein
AT5G19690	-4.84	-6.09	oligosaccharyl transferase STT3 subunit family protein
AT2G33830	-5.61	-7.07	dormancy/auxin associated family protein
AT5G05740	-5.75	-2.41	S2P-like putative metalloprotease
AT5G43780	-6.25	-3.41	sulfate adenylyltransferase 4/ATP-sulfurylase 4 (APS4)
AT1G33360	-7.80	-9.82	ATP-dependent Clp protease ATP-binding subunit ClpX, putative
AT3G53520	-8.90	-11.20	NAD-dependent epimerase/dehydratase family protein
AT4G36530	-10.08	-4.63	hydrolase, alpha/beta fold family protein
AT5G22880	-26.74	-4.23	histone H2B, putative
AT2G38000	-2.97	-3.73	chaperone protein dnaJ-related
AT5G40170	-2.97	-3.73	disease resistance family protein
AT2G34680	-2.97	-3.73	AIR9
AT4G24090	-3.10	-3.91	expressed protein

*IDs in bold are genes with opposite regulation

many proteins with similar function showed similar expression patterns. I focused particularly on some mitochondria and chloroplast-located proteins. One group of the zeatin treated, specific up-regulated proteins contains: AT2G34460 (NAD(P)-binding Rossmann-fold superfamily protein), AT1G72640 (NAD(P)-binding Rossmann-fold superfamily protein], AT1G54010 (GDSL-like Lipase/Acylhydrolase superfamily protein), ATCG00680(subunit of the photosystem II reaction center), and others. The first three genes are involved in lipid metabolism, especially lipid oxidation and catabolism.

3.4.4. Pathway Analysis Revealed Distinctive Responses of Two Hormones

Both zeatin and BR are relevant to plant growth regulation, cell elongation, and energy metabolism. I therefore carried out pathway analysis using AraCyc to investigate if both hormone treatments promote the plant and cellular growth with the same metabolic pathway or not. The pathway analysis revealed distinctive patterns.

The most impressive pathway-level differences between zeatin and BR triggered responses are the regulation within the fatty acid biosynthesis pathway. Three proteins (AT1G24360, AT2G05990 and AT2G04540) were found down-regulated in the BR treated sample. AT1G24360 is an NAD(P)-binding Rossmann-fold superfamily protein, and AT2G05990 is an enoyl-ACP reductase, a component of the fatty acid synthase complex. AT2G04540 is a beta-ketoacyl synthase. All of these enzymes are involved in fatty acid biosynthesis and elongation. In addition, two other proteins relevant to very long-chain fatty acid biosynthesis were down-regulated. These two proteins are

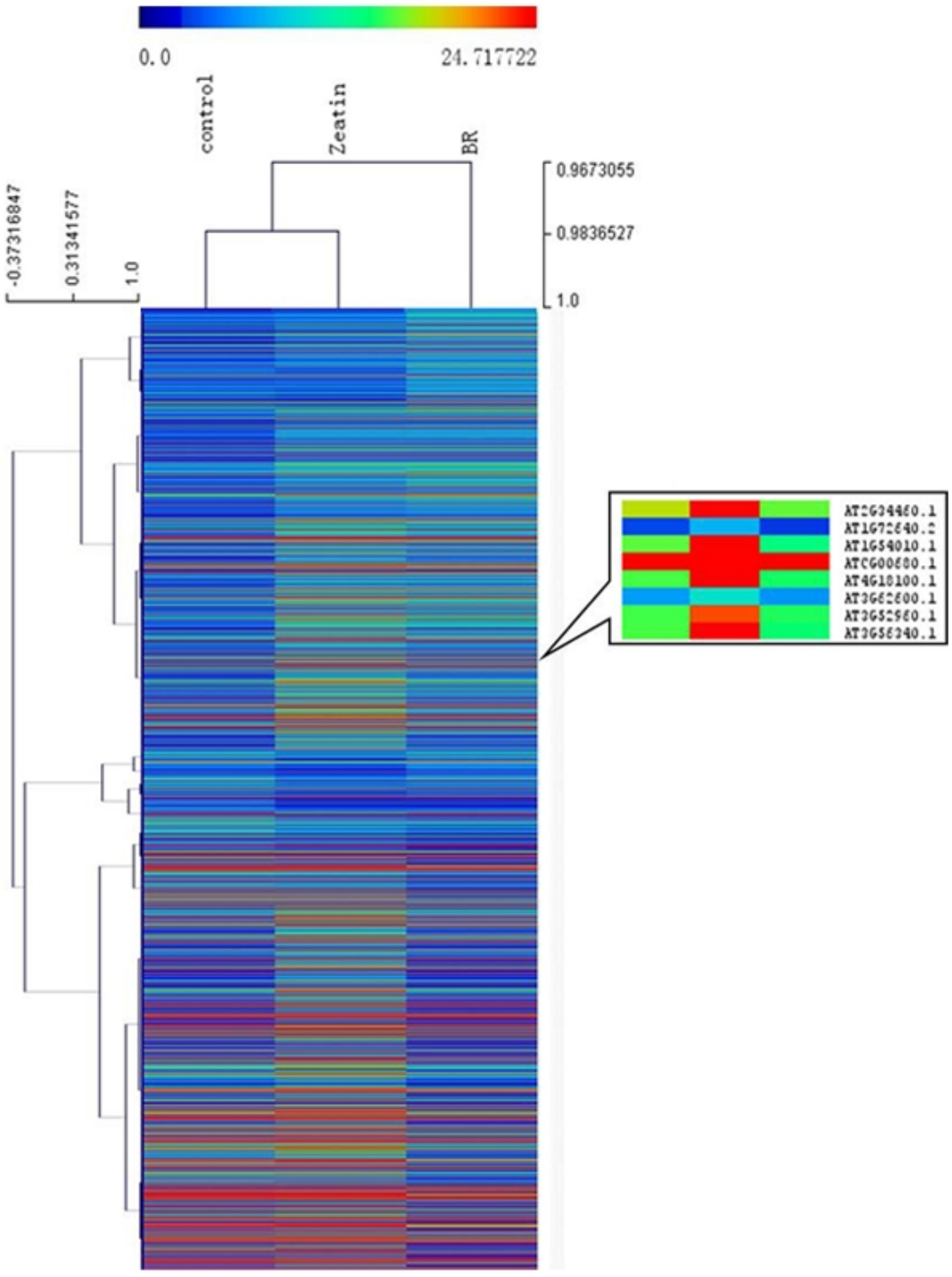

Fig. 3. 3. Overview of cluster analysis of zeatin and BR treated samples and a snapshot of a group of zeatin proteins.

AT1G76150 encoding an enoyl-CoA hydratase and AT5G27600 encoding a peroxisomal long-chain acyl-CoA synthetase. Despite the many down-regulated proteins in BR responses, few proteins can be found differentially expressed in the zeatin treated sample for lipid biosynthesis (Fig. 3.4).

Besides the down-regulation of fatty acid biosynthesis, the pathways for utilization and oxidation of fatty acid were also down-regulated in BR treated samples. Four down-regulated gene products (AT1G76150, AT3G06860, AT5G65110 and AT5G27600) play distinct roles in fatty acid β-oxidation pathway. AT1G76150 degrades even *cis*-unsaturated fatty acids. AT5G65110 encodes an acyl-CoA oxidase for fatty acid oxidation. AT5G27600 involves in oxidation of very long chain fatty acid in peroxisomes.

Even though lipid metabolisms were significantly changed in response to BR treatment, sucrose and starch metabolisms seem to be changed more by zeatin treatments. Two gene products in the starch biosynthesis pathway were up-regulated in the zeatin-treated sample. These two proteins are AT5G48300, a small subunit of ADP-glucose pyrophosphorylase and AT5G03650, starch branching enzyme. Meanwhile, proteins were found up-regulated in both the starch degradation pathway I and pathway II. Among these proteins are AT2G40840, a crucial enzyme for starch to sucrose conversion; AT1G10760, α-glucan dikinase; and AT5G26570, chloroplastidic phosphoglucan water dikinase. However, in the BR treated sample, only ApL4, a large subunit of ADP-glucose pyrophosphorylase catalyzing the first rate limiting step in starch biosynthesis was found down-regulated in the starch biosynthesis pathway.

In addition to starch metabolism, zeatin treatment has induced three enzymes in sucrose biosynthesis and catabolism. These enzymes are AT5G20830 (sucrose synthase, Sus1), AT2G22480 (phosphofructokinase) and AT5G52920 (pyruvate kinase beta subunit). No proteins in the sucrose degradation pathway showed significant changes in the BR treated sample.

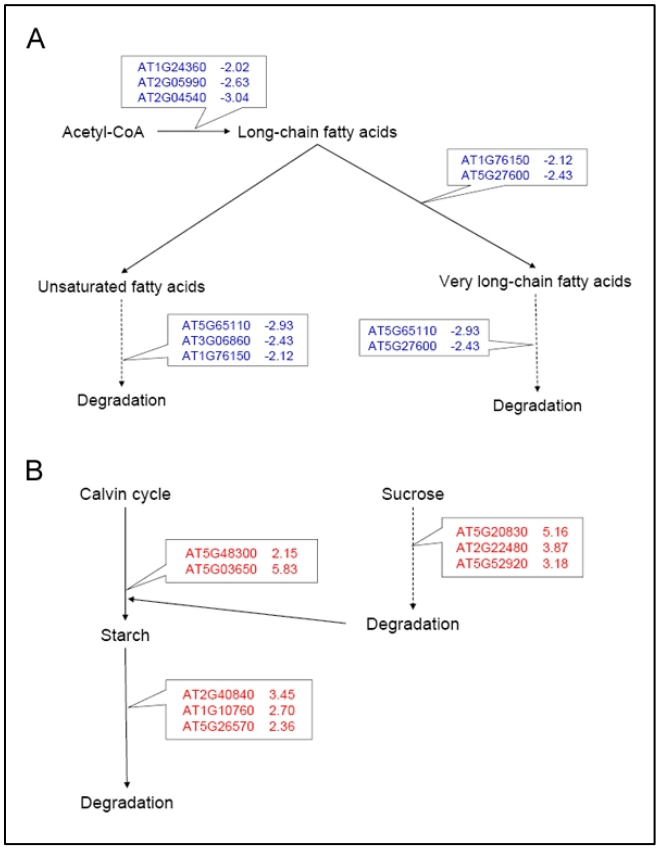

Fig. 3. 4. Pathways analysis of differentially regulated proteins in zeatin and BR responses. A. BR down-regulates many lipid biosynthesis sand utilization proteins. B. Zeatin up-regulates many sucrose and starch metabolism proteins.

3.5. DISCUSSION

3.5.1. Plant Organelle Proteomics

Different strategies can be used for proteomics analysis in organelles like mitochondria and chloroplast (142). The traditional approach is to isolate these organelles with multiple steps of gradient centrifugation and ultra-centrifugation. Proteins were further isolated from the organelles for proteomics analysis. The limitation of the strategy lies in the requirement of a large quantity of initial sample and the potential errors that could be introduced during the multiple step purification (143, 144). I hereby adopted another strategy to combine a simple and rapid sample preparation method with bioinformatics classification. One simple centrifugation step was used to separate the mitochondria and chloroplast from other plant organelles. The separated mitochondria and chloroplast protein was then used for shot-gun proteomics and bioinformatics classification. The method has led to the enrichment of mitochondrial protein identification by 30%, and reduction of initial sample amount by more than 10-fold. I went ahead and utilized the method to study an important biological question in plant hormone responses.

3.5.2. Improved Protein Identification for Hormone Response Proteomics Analysis

Previous research, mainly utilizing 2D DIGE (Two-dimensional differential gel electrophoresis), was carried out to study the BR-treated Arabidopsis (20). The study has led to the discovery of 103 of differentially expressed proteins. As compared to the previous studies, more differentially proteins were identified in the present study,

demonstrating the effectiveness of the shot-gun proteomics platform and the sampling strategy. A total of 355 proteins have been identified in the zeatin treated samples and a total of 288 proteins were found to be differentially expressed in the BR treated samples. The deep coverage of differentially regulated proteins and focused study of energy-related pathways in mitochondria and chloroplasts allow us to have a global comparison of metabolic pathway regulations at the proteome level between the two types of hormone responses.

3.5.3. The Distinct and Shared Pathways Induced by Zeatin and BR Treatment

The study has revealed significantly differential regulation of metabolic pathways in zeatin and BR, in particular for pathways located in mitochondria and chloroplasts. Even though both zeatin and BR can promote cell elongation, the mechanisms are expected to be different. The results highlighted that BR down-regulates key proteins in both fatty acid biosynthesis and oxidation. Fatty acid β-oxidation eventually breaks down the long-chain fatty acids and produces acetyl-CoA to enter TCA cycle (145). The fact that both fatty acid biosynthesis and catabolism are down-regulated indicates that BR may promote the cell elongation and growth through shutting down the energy storage through lipid biosynthesis. Interestingly, zeatin treated plants showed essentially no changes in these two pathways, indicating a completely different metabolic regulatory mechanism.

For zeatin treatment, some of the sucrose and starch biosynthesis proteins were up-regulated. Additionally, proteins involved in sucrose and starch degradation were

also up-regulated. The use of sucrose is one way that plants transport energy; synthesized sucrose from photosynthetic tissues can be transported to other tissues and cells for utilization (146). The fact that both biosynthesis and degradation were up-regulated indicates the rapid metabolism of these energy source compounds. Interestingly, BR treatment only induces the down-regulation of one gene involved in starch biosynthesis.

The comparison of the two hormone responses indicated that the two types of plant hormones regulate cell elongation and growth through distinctive pathways. BR down-regulates key proteins in lipid metabolisms and energy storage, while zeatin up-regulates key proteins in sucrose and starch metabolisms for energy utilization. The future work can be developed to coordinate the expression of genes involved in the responses to two plant hormones to develop new ways for manipulating plant growth and development.

CHAPTER IV

ENGINEERING PLANT PROHIBITIN TO PROMOTE PLANT GROWTH THROUGH IMPROVING ENERGY EFFICIENCY

4.1. SUMMARY

New gene discovery for plant defense and growth mechanism is challenging. The bottle neck is the high-throughput screening method and confirmation and characterization process. Compared to the current next generation sequence techniques, and gene array techniques, proteomics technology is an important part of systems biology. PHB8 (AT3G01290) is a Class III PHB gene which may regulate plant yield and defense against pathogen and insect through a totally different mechanism from the traditional studies. The aim of the study is to characterize how PHB8 regulates plant growth and defense using Arabidopsis as the model plant. The PHB8 gene may form hetero oligomers with other Class III PHB gene products to function as the molecular chaperone to stabilize mitochondrial proteins. Shot-gun proteomics analysis can be used to characterize the PHB protein complex that localizes on the membrane as the molecular chaperone. By using this powerful tool, several plant defense and growth mechanisms were studied, which involved plant hormone treatment and insect treatment. The findings of the present studies indicate that direct protein level study of the dynamically changing biological system is the most straight forward path to characterize novel genes in various pathways. In summary, methodology development integrated with bioinformatics can help us study protein function more efficiently. The research

presented here may lead to clarify novel mechanisms to regulate plant growth and defense by stabilizing membrane proteins as molecular chaperones.

4.2. INTRODUCTION

Higher yield and rapid growth are major challenges for crop improvements (147, 148). The traditional pathways studied for plant yield improvement include photosynthesis, nutrition usage, abiotic stress, delayed flowering, cellulose biosynthesis, GA (gibberellic acid) biosynthesis and signaling (147-153). Despite progress, a bottleneck for further increases in crop yield and plant biomass has been witnessed in recent years (154). Novel approaches and mechanisms to increase crop yield are necessary to break the bottleneck (152, 154-156). My recent discovery of a prohibitin (PHB)-based protein regulatory network may provide a novel approach for crop improvement for improved growth and yield.

Prohibitins (PHBs) are ubiquitous, pleotropic and evolutionarily conserved proteins involved in signaling, cell cycle regulation, mitochondrial respiration, cell death and others (56-63, 157). PHBs were originally found to be associated tumor-suppressor genes in mammalian cells (58, 59, 61, 71, 158-160). Recent studies indicated that prohibitin genes are associated with a broader range of disease phenotypes including Parkinson's disease, Type 2 diabetes, Hepatitis C, obesity and others (61, 161-167). The exact molecular functions of PHBs are still under debate. PHBs were originally indicated to be important in raft formation in mitochondrial and other organelle membranes, then found to be present in the nucleus to interact with transcriptional factors, and later

implied for the broader function of stabilizing and degrading membrane proteins (64, 66, 168, 169). Regardless, significant evidence indicated that PHB proteins are involved in energy metabolism through its regulation of mitochondria function. Classic studies in yeast indicated that PHB1 and PHB2 form a ring-structure complex to act as a molecular chaperone to stabilize proteins in mitochondria (58, 170). In the nematode *Caenorhabditis elegans*, prohibitin was found to negatively regulate energy metabolism and aging by controlling mitochondrial biogenesis (63, 171). The role of PHB genes in energy metabolism makes the plant homologs potential candidates to manipulate plant growth and yield.

A recent comprehensive genome analysis of PHB gene family from prokaryotes to eukaryotes indicated that the PHB gene family is an evolutionally conserved but functionally diverse gene family (68). The studies of plant PHBs mainly focused on PHB1, 2, 3, 4 in Arabidopsis and their homologs in tobacco (69-72). Studies indicated that PHB proteins play important roles in ROS responses and mitochondria morphology (69, 73). PHB3 was found to be involved in NO homeostasis and abiotic stress response (72). AtHIR1 (Hypersensitive Induced Reaction proteins) and AtHIR2, which also named PHB8 and PHB9, were shown to form a complex with RPS2 to act as an important role in effector triggered immunity (ETI) in plant disease defense (70). Recently, PHB2 was identified as an interacting protein with endoplasmic reticulum BAX INHIBITOR-1 (BI-1) to regulate cell death in the plant powdery mildew interaction (74). The existence of m-AAA protease-PHB2 complex suggested prohibitins regulate protein metabolism (172). In addition, previous studies showed that the down

regulation of PHB3 and PHB4 in Class IV-A can lead to dramatic changes in plant morphology and smaller plant size (69, 71, 73, 159). Despite the extensive study in plant defense, few studies focused on how PHB proteins may regulate energy metabolism and plant growth, even though evidence indicated that the PHB genes are important in energy metabolism in human and animal studies.

In this study, I discovered a PHB8-involved protein network to regulate ATPase level and energy metabolism. PHB8 gene was found to be significantly up-regulated in several growth-promoting hormone treatments. The over-expression of PHB8 in Arabidopsis increased plant growth rate, shortened flowering time, and increased the plant height, stem diameter and seed yield significantly. Downstream proteomics revealed that PHB8 over-expression impacts the energy metabolism. In particular, an ATPase beta subunit is up-regulated due to the PHB8 over-expression. The results correlated well with the increase of ATP production in the PHB8 over-expression lines. Further genetic study showed that over-expression (OE) of ATPase led to a similar phenotype as PHB8 OE lines. The results indicated that PHB8 might regulate ATPase protein level to regulate energy metabolism. Further analysis of a protein interaction network revealed that PHB8, 9, 16 interacts with one another. The results supported that class III PHB proteins form a complex to selectively stabilize membrane proteins. In particular, in mitochondria, this machinery selectively assists the folding or stabilizes the key energy producing proteins like ATPase to regulate energy metabolism. PHB8 proteins and the protein metabolism cascade regulated by PHB8 can be exploited to improve plant growth, biomass and seed yield. The finding will not only help to develop

new strategies for plant improvement, but also shed light into human disease studies on how PHB proteins may regulate protein and energy metabolism.

4.3. MATERIALS AND METHODS

4.3.1. Plant Materials and Growth Conditions

Arabidopsis thaliana Col-0 seeds mutagenized by T-DNA insertion were obtained from ABRC. Seeds were stratified at 4ºC for 2 days and then were transferred into growth chamber at 22 ºC/19 ºC with a circadian cycle of 12-h-light/12-h-dark.

4.3.2. Identification of the T-DNA Insertion Mutant

The *PHB8* mutant (Salk_092306) contains a T-DNA insertion in the third intron. The T-DNA insertion mutant plants were confirmed by PCR using a combination of a T-DNA border primer LBb1.3 (5'-ATTTTGCCGATTTCGGAAC -3') and gene specific primers (LP:5'-GGGCAACTACTGATCTTTCCC-3', RP:5'-CGAGCAGAAGAATGAAATTGC-3')

4.3.3. Construction of Transgenic Arabidopsis Lines

The *PHB8* and *ATPase* were amplified from total RNA of *Arabidopsis thaliana* Col-0 and cloned into the pENTR vector (Invitrogen). To create transgenic plant line overexpressing the *PHB8* gene and *ATPase*, the ORF of *PHB8* and *ATPase* were inserted into the pEarleyGate100 vector respectively (173). All Gateway LR reactions were performed by Gateway® LR Clonase® II enzyme mix (Invitrogen) following the

instructor's manual. For intracellular localization, the full-length *PHB8* gene coding sequence was cloned into a rebuilt vector pX-DG driven by the CaMV *35S* promoter (100).The construct was verified by sequencing and introduced into *Agrobacterium tumerfaciens* strain GV3101. The construction was transformed by floral infiltration method (174). Transgenic plants were selected by Basta resistance and confirmed by PCR. The homozygous T2 seeds of the transgenic plants were used for further analysis.

4.3.4. Real-time Quantitative PCR

For the qRT-PCR analysis, total RNA isolation using TRI Reagent was performed following manufacture's instruction (99). Then, 1.0 μg of total RNA was incubated with RNase-free DNase I following the protocol provided by the manufacturer to remove possible genomic DNA. Reverse transcription of total RNA was performed using SuperScript III First-Strand Synthesis kit (Invitrogen, CA, USA).

Triplicate quantitative assays were performed using the SYBR Green Master Mix (Applied Biosystems, CA, USA) with an ABI 7900HT sequence detection system. All data were normalized with 18S rRNA or UBQ10 as internal reference. Primers used for quantitative qRT-PCR analysis were listed in supplemental Table 3. The relative quantification method (modified ΔΔCT) was used to evaluate the differential gene expression.

4.3.5. Intracellular Localization

The GFP fluoreescene signal was observed by using OLYMPUS FV1000

confocal microscope. Protoplast of transgenic plants was prepared by following previous protocol (175).

4.3.6. Pull-down and Protein Digestion

To better pull down the PHB8 membrane protein complex, the 80 amino acid sequence (BCCD) at the C-terminus of Arabidopsis MCCA (At1g03090) was cloned as the biotin affinity tag. The purification of BCCD fusion protein was modified method of Qi (176). The BCCD tag was cloned into PUC19, fused with FLAG tag. The FB tag was fused at C-terminal of PHB8 gene then cloned into pEarleyGate 100 by LR reaction (Fig. 4.1). The cross-linker, i.e. dithiobis (succinimidyl propionate) (DSP), was used to stabilize protein complexes with weak or transient interactions. The captured protein complexes were digested on Dynabeads magnetic beads by Trypsin digestion solution for 12 h.

4.3.7. Plant Total Protein Isolation

For the total protein isolation, a plant total protein extraction kit (Sigma-Aldrich, St. Louis, MO) was used, and the entire procedure followed the manufacturer's manual. 150 mg of aerial tissue from Arabidopsis was collected and ground in liquid nitrogen to a fine powder. Pre-cooled methanol solution with protease inhibitor was added to the powder and vortexed for 30 seconds. The mixture was incubated at -20℃ then centrifuged at 16,000×g for 5 minutes at 4℃. Supernatant was removed and the pellet

was washed by methanol solution for two more times. The resulting pellet was washed by pre-cooled acetone and centrifuged at 16,000×g for 5 minutes at 4℃. SpeedVac was used to remove residual acetone and Reagent Type 4 Working Solution provided by the kit was used to incubate the pellet for 15 minutes at room temperature. The pellet was then centrifuged at 16,000×g for 30 minutes, and supernatant was collected and stored at -80℃ for future proteomics use.

Fig. 4.1. Diagrams of the PHB8-FB expression cassettes after LR reaction in pEarleyGate100 vector. FLAG and BCCD were fussed with PHB8 in one ORF. LB and RB, left and right border sequence of T-DNA transfer; BAR, the Basta resistance gene; 35S, the CaMV 35S promoter and its upstream enhancer; *attB1* and *attB2*, Gateway® cloning cassettes generated after LR reaction; FLAG, tag used for affinity purification; BCCD, biotin carboxyl carrier protein domain; OCS, the 3' sequences of the octopine synthase gene with transcriptional terminator.

4.3.8. LC-MS/MS

For pulldown sample analysis, a 120 min LC liner gradient separation from buffer A (0.1% FA) to buffer B (80% acetonitrile, 0.1% FA) was used to separate digested peptides.

For MudPIT-based shot-gun proteomics were carried out to analyze each aforementioned fraction as described by Washburn et al. (91). Approximately 100 μg protein was digested by Mass Spectrometry Grade Trypsin Gold (Promega, WI, USA) with 1:40 w/w at 37 °C for 24 hr. The digested peptides were desalted using a Sep-Pak C18 plus column (Waters Corporation, Milford, Massachusetts, USA) and then loaded onto a biphasic (strong cation exchange/reversed phase) capillary column using a high pressure tank. The two-dimensional liquid chromatography separation and tandem mass spectrometry were carried out as previously described (177).

The separated peptides were analyzed using a linear ion trap mass spectrometer, Finnigan LTQ (Thermo Finnegan, San Jose, CA, USA). The mass spectrometry was set to the data-dependent acquisition mode. Then the full mass spectra were recorded on the peptides over a 300-1700 m/z range, followed by five tandem mass (MS/MS) events for the most abundant ions from the first MS analysis. The Xcalibur data system (Thermo Fisher Scientific, San Jose, CA, USA) was used to control the LC-LTQ system and collect the data.

4.3.9. Proteomics Data Analysis

A data preprocessing pipeline based on the Xcalibur Development Kit was used to generate DTA files in the same way as the ThermoFinnigan Bioworks (2.0) software. Tandem mass spectra were extracted from the raw files and converted into the MS2 file. The MS2 file was searched against the Arabidopsis protein database version 10 containing 27,416 protein sequences from the TAIR, the same number of reverse sequence, and common contaminant proteins. The reverse sequences of the original dataset were included the calculate confidence levels and false-positive rates. The common containment proteins were included for data quality control. The ProLuCID (version 1.0) algorithm was used to assign peptide sequence to peptide fragmentation spectra using the Texas A&M Supercomputing Facility. The ProLuCID parameters were set as the follows: the precursor mass accuracy was set at 100 ppm; fragment ion mass accuracy was set at 600 ppm. No fixed or variable modifications were considered. At least two distinct peptides (semi-tryptic) were required to identify a protein with no sequence coverage assigned.

Protein quantification was achieved by spectral counting. The validity of peptide/spectrum matches were assessed in DTASelect v.2.0 using a 0.05 false discovery cutoff, with a cross-correlation score (XCorr) that's larger than 1, and normalized difference in cross-correlation scores (DeltaCN) larger than 0.08. Semi-tryptic peptide was included in the final calculation. DTASelect software listed the accession numbers together for those proteins combined into one group. Each individual accession numbers were then listed in the supplementary data. Proteins identified with more than two

peptides were used for quantification. Protein identification was based on both specific and non-specific peptides.

4.3.10. ATP Measurement

The ATP production was analyzed with two-week-old Arabidopsis seedlings, which were cultured under dark conditions to avoid plastid ATP production. The seedlings were collected and blotted on filter paper to remove water on surface then weighted. Samples were ground in liquid nitrogen and re-suspended in 2.5% TCA solutions to extract ATP. The samples were centrifuged at 16,000 g to precipitate plant tissues and the supernatant were collected for further analysis. ATP production was measured using ATP ENLITEN® ATP Assay System Bioluminescence Detection Kit (Promega) according to user instructions.

4.4. RESULTS

4.4.1. Discovering PHB8 As A Candidate Gene for Improving Plant Yield

Previous genome analysis discovered a diverse expression pattern for prohibitin genes at different developmental stages, tissues, and responses to stimulus (68). Considering that prohibitins are involved in energy metabolisms, detailed analysis of gene expression in response to plant growth hormones were carried out to identify which prohibitin gene respond to the growth promoting hormones. As shown in Fig. 4.2, the visualization of an AtGenExpression dataset for hormone treatment indicated that PHB8 is significantly up-regulated by brassinosteroid and IAA. The gene is also up-regulated

by GA application in the GA biosynthesis mutant ga1-3 and ga1-5 plants. The ga1-3 mutant possesses a phenotype of very late-flowering under short light growth condition (178, 179). Compared to gal1-3 mutant, ga1-5 is weak GA biosynthesis mutant which has low level of endogenous GA resulted in the expression of anthocyanin pathway related genes (180, 181).The results indicated that PHB8 gene might be involved in plant growth regulation.

4.4.2. PHB8 Over-expression Increases Biomass and Seed Yield

In order to explore if PHB8 gene can promote plant growth or not, both phb8 T-DNA knock out mutants and over-expression lines were generated. Two PHB8 OE lines were confirmed with significantly increased PHB8 gene expression (Fig. 4.3). Both of the PHB8 OE lines have shown clear phenotype of rapid growth (Fig. 4.4), early flowering, increased height, stem diameter and seed yield (Fig. 4.5 and Fig. 4.6).

Even though the phb8-3 mutants did not have a phenotype in plant growth, growth phenotype can be observed for the fully grown Arabidopsis plants after 70 days. As compared to the wild-type, the stem height of OE 4-14 and OE11-7 line were increased by 19.6% and 20.9%, respectively (Fig. 4.5*b*). The stem diameter of OE 4-14 and OE11-7 lines were increased by 24.4% and 24.2% as compared with WT (Fig. 4.5*c*). The seed yield was calculated indirectly by counting total silliques number per plant and the seeds in each sillique. As shown in Fig. 4.6, the seed number per sillique is the same. The total number of sillique per plant for OE4-14 and OE11-7 lines increased by 22.1%

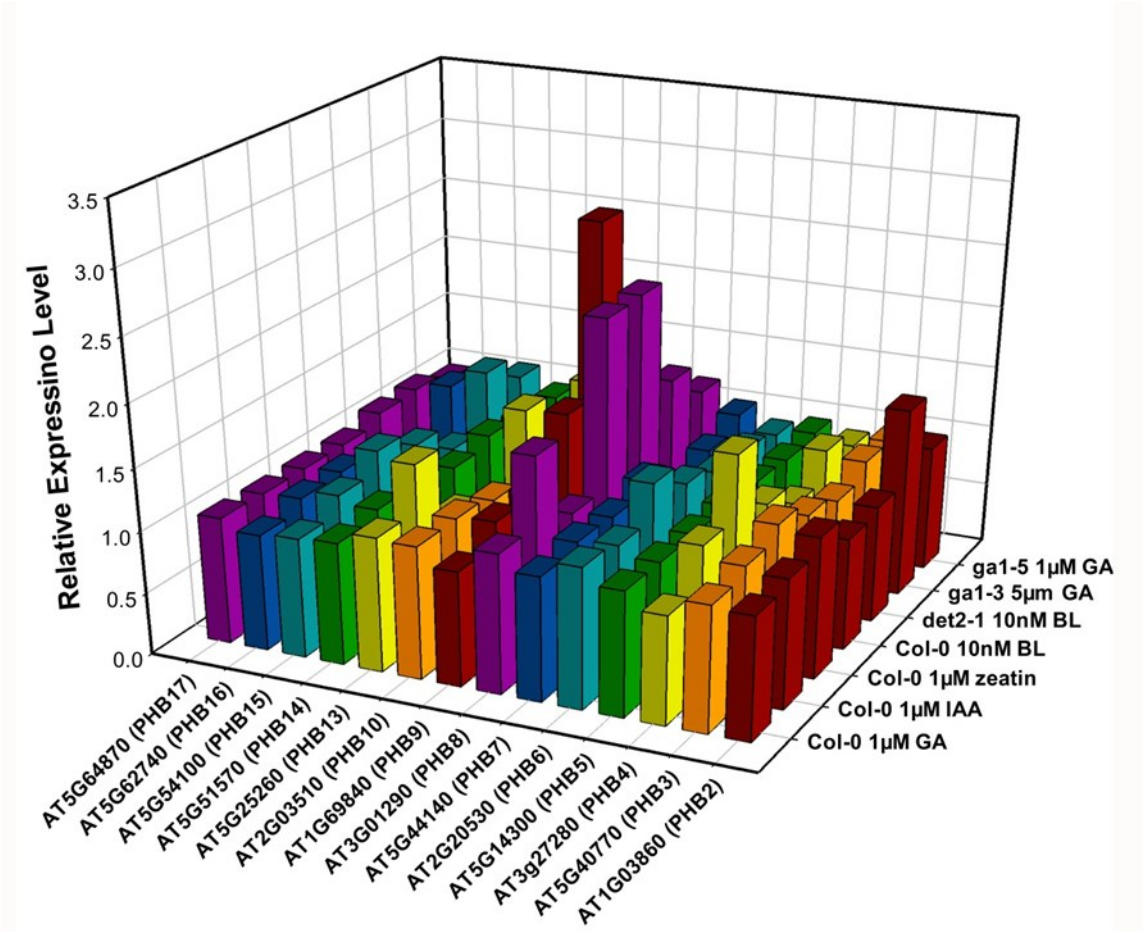

Fig. 4. 2. Expression level of PHB gene family in Arabidopsis under plant growth promoting hormone treatment with different genetic background. BL, brassinolide; IAA, indole acetic acid; GA, gibberellic acid; ga1-3 and ga1-5, GA biosynthesis mutant; Col-0, *Arabidopsis thaliana* ecotype Columbia wild type. The data showed the ratio of mean-normalized expression level of PHB gene in treatment sample versus control for 3 h in each tissue cluster. To better visualize the data, one of the data point of AT5G25260 (PHB13) treated with 10 nM BL in Col-0 was removed from the figure due to the high value of ratio.

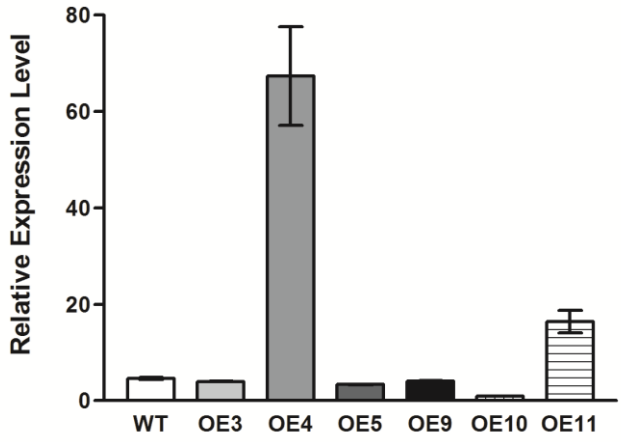

Fig. 4. 3. qRT-PCR result of PHB8 expression in overexpression lines. The relative expression levels were obtained by comparing the transgenic lines with wild type. The relative ratio and 95% confidence interval were derived from triplicate assays.

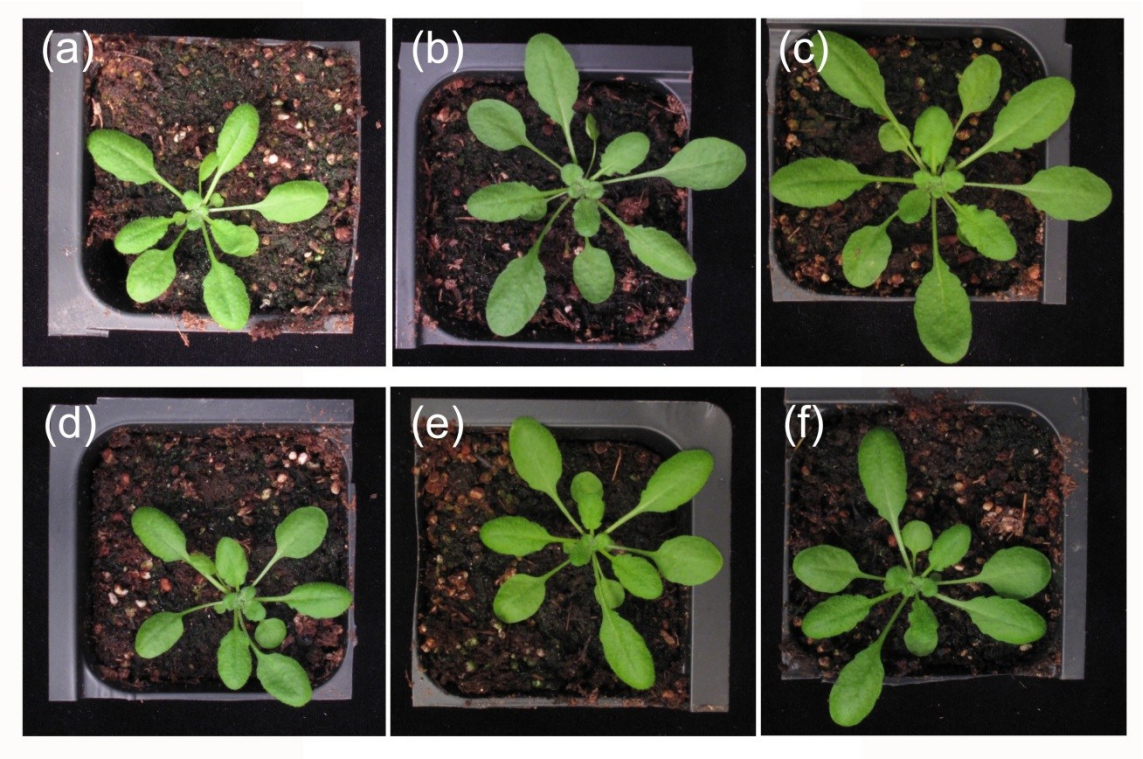

Fig. 4. 4. Leaf size of wild-type *Arabidopsis thaliana*, PHB8 T-DNA insertion mutant, PHB8 and ATPase overexpression transgenic plants at day 28. (a) wild type *Arabidopsis thaliana* ecotype Columbia; (b) OE4-14 and (c) OE11-7, two PHB8 OE lines; (d) *phb8-3*, T-DNA insertion mutant line (Salk_092306); (e) ATPase OE19-16 and (f) OE24-14, two transgenic lines with ATPase overexpression construct.

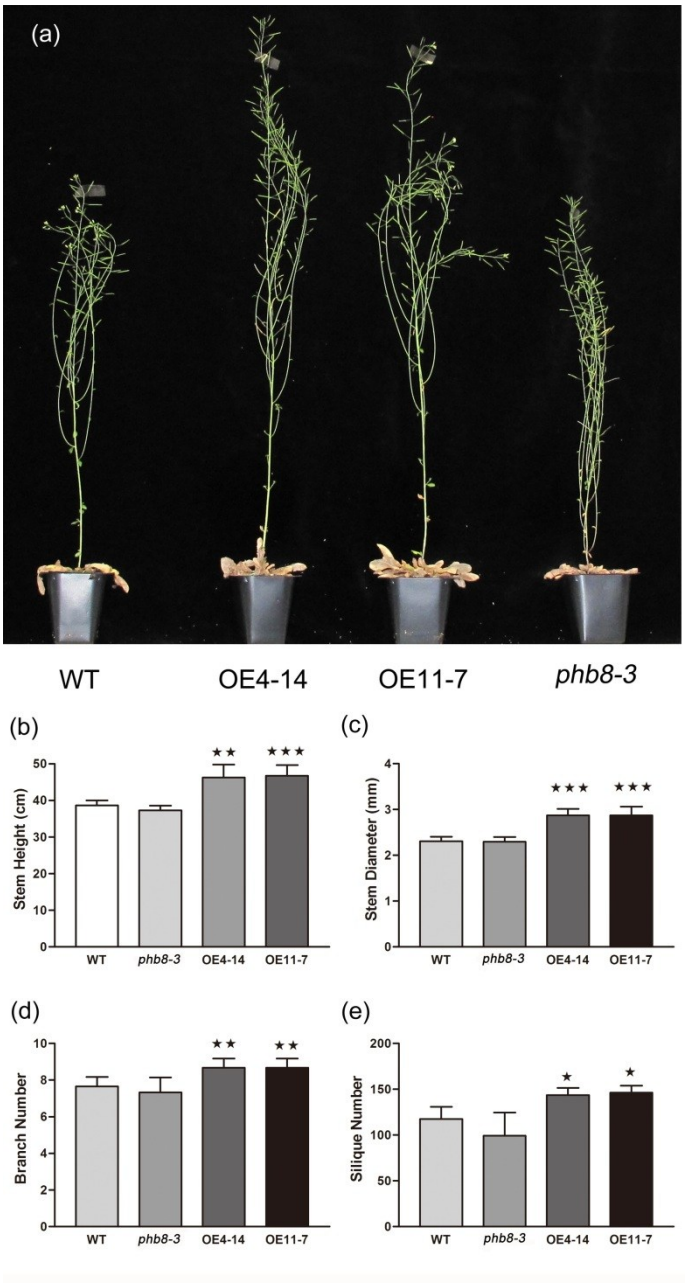

Fig. 4. 5. Phenotype of wild-type *Arabidopsis thaliana*, *PHB8* transgenic lines and T-DNA insertion mutant line. WT, wild type *Arabidopsis thaliana* ecotype Columbia; OE4-14 and OE11-7, two lines with single copy of *PHB8* overexpression construct; *phb8-3*, T-DNA insertion mutant line (Salk_092306). (a) Phenotypes of WT, PHB8 OE4-14, OE11-7 and *phb8-3*. (b) Stem height. (c) Stem diameter. (d) Branch number. (e) Silique number. Data are shown as means with standard deviation (n=6). Asterisks indicate difference level between transgenic and WT plants with **P*< 0.01, ***P*<0.05, and ****P*<0.001 using one-way ANOVA.

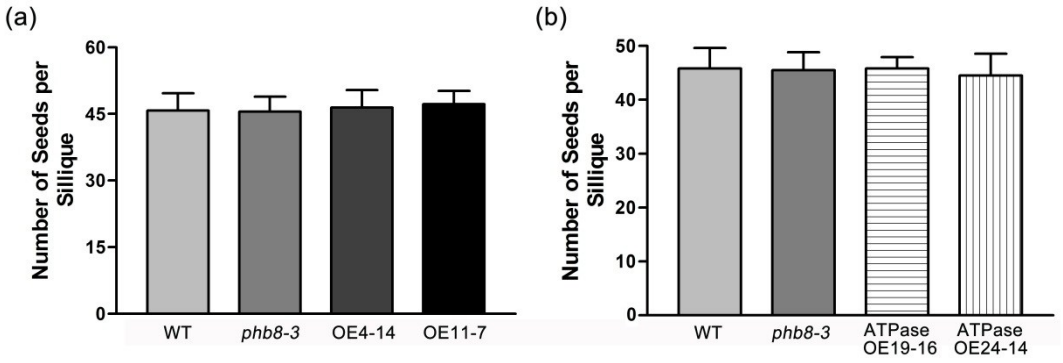

Fig. 4. 6. The number of seeds per sillique for all transgenic lines and WT. WT, wild type *Arabidopsis thaliana* ecotype Columbia; *phb8-3*, T-DNA insertion mutant line (Salk_092306), OE4-14 and OE11-7, two lines with single copy of *PHB8* overexpression construct; ATPase OE19-16 and OE24-14, two transgenic lines with *ATPase* overexpression construct. (a) Result of WT, *phb8-3,* PHB8 OE4-14, OE11-7. (b) Result of WT, *phb8-3,* ATPase OE19-16, OE24-14. Data are shown as means with standard deviation (n=10).

and 24.5%, indicating the increased seed yield for phb8 OE lines. In consistency with the increased seed yield, the branching number for the over-expression lines increased too.

4.4.3. Protein Network Regulated by PHB8

In order to further investigate the downstream network regulated by PHB8, shot-gun proteomics were carried out to compare the wild-type and OE4-14 lines. Gene Ontology analysis indicated that ATP processing and energy metabolism-related genes were up-regulated. Further computational analysis revealed that a network of ATP production and energy generating proteins were synergistically enriched in the PHB8 over-expression lines. These proteins included ATPase, NAD (P) H dehydrogenase, and other proteins in the electron transport chain. As shown in Table 4.1, the most significantly up-regulated protein in the OE line is an F0/F1 type ATPase. AT5G08680 is the beta subunit of an F0/F1 type ATPase, which produces ATP through proton gradient in mitochondria. The result indicated that PHB8 OE line might achieve rapid growth as well as increased seed and biomass yield through enhancing energy metabolism and ATP production. There is no other prohibitin family proteins were identified differentially expressed in this assay. An internal NAD(P)H dehydrogenase (NDA1) was found up-regulated more than 7 times. Recent research showed that the reduced expression of NDA1 in Arabidopsis displays a smaller leaf diameter compared to wild type as well as a slower growth rate. The NDA1 suppression dramatically affects the redox balance in leaves, which increases NADH/NAD^+ ratio in mitochondria. However,

Table 4. 1. Top 10 proteins of up-regulated and down-regulated proteins for the *PHB8* OE line.

Locus	Fold Change	p-Value	Description
AT5G08680.1	274.75464	0.000772	ATP synthase beta chain, mitochondrial, putative
AT4G20890.1	77.8959036	0.000047	Tubulin beta-9 chain
AT3G11250.1	10.0134945	0.00934	60S acidic ribosomal protein P0
AT1G07180.1	7.4147608	0.011903	Internal NAD(P)H dehydrogenase in mitochondria
AT3G01290.1	5.16227093	0.000117	Band 7 family protein
AT2G40290.1	5.11839231	0.022288	Eukaryotic translation initiation factor 2 subunit 1
AT4G19120.1	4.9420945	0.018542	Early-responsive to dehydration stress protein
AT3G09740.1	4.72954721	0.026761	Syntaxin 71
AT5G61970.1	4.43540645	0.004326	Signal recognition particle-related / SRP-related
AT2G30620.1	4.2107136	0.015241	Histone H1.2
AT1G29965.1	-4.04506236	0.006277	60S ribosomal protein L18A
AT4G16150.1	-4.057472075	0.008749	Calmodulin-binding protein,
AT3G16530.1	-4.291189539	0.0305	Legume lectin family protein
AT5G42220.1	-4.382416661	0.002328	Ubiquitin family protein
AT4G30490.1	-4.467202202	0.005992	AFG1-like ATPase family protein,
AT1G56190.1	-4.654514345	0.01118	Phosphoglycerate kinase, putative
AT5G39830.1	-4.735807281	0.001504	DegP protease, putative
AT3G17170.1	-5.00023264	0.014827	Ribosomal protein S6 family
AT5G47930.1	-5.526821499	0.000126	40S ribosomal protein S27
AT1G02930.1	-6.286924789	0.001052	Glutathione S-transferase, putative
AT5G26860.1	-6.352723332	0.010671	Similar to Lon protease, putative

there is no evidence to support NDA1 protein involved in photosynthesis in high light conditions (182-184).

4.4.4. PHB8 OE Increases Pool ATP Amount

The ATP production for PHB8 OE lines was further verified to confirm that PHB8 over-expression leads to increased ATP production. Considering that the chloroplasts also produce ATP under light conditions, I used using etiolated plants in dark to compare the phb8 mutant, wild-type, and over-expression lines for ATP production. As shown in Fig. 4.7, the ATP levels in OE-4 and OE-11 lines are 10.57% and 9.72% higher than that of wild-type Even though phb8-3 mutant does not have a strong growth phenotype, the mutant thus has a slightly decreased ATP production by 6.61%.

4.4.5. Genetic Evidence Indicating PHB8 Regulates Plant Growth Partially Through ATPase

Considering that F0/F1-type ATPase beta-subunit (AT5G08680) is the most enriched protein in the PHB8 OE lines, the gene was over-expressed to elucidate the mechanisms for PHB8 OE phenotypes (Fig. 4.8). As shown in Fig. 4.9, the OE lines of F0/F1 type ATPase beta subunit lead to more rapid growth and an increase of plant height by about 8%. No significant differences were observed for the stem diameter, branch number, and seed yield. The similar rapid growth and height increase between the PHB8 and ATPase beta subunit over-expression lines indicate that PHB8 regulated the ATP production and plant growth at least partially due to the regulation of ATPase

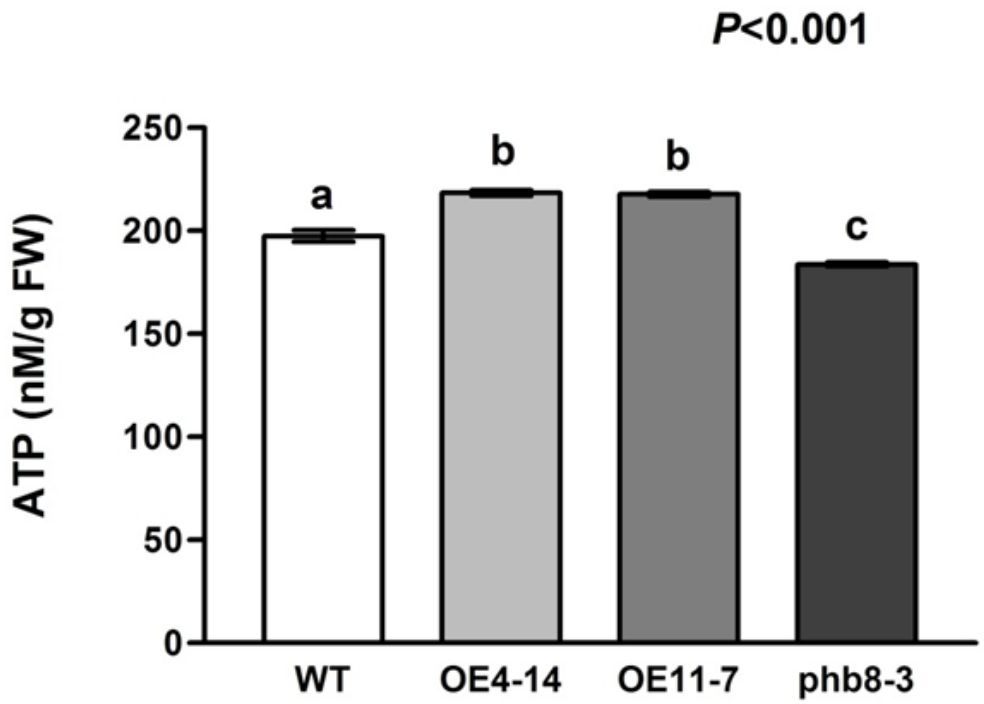

Fig. 4. 7. The level of pool ATP in wild-type *Arabidopsis thaliana*, *PHB8* transgenic lines and T-DNA insertion mutant line. WT, wild type *Arabidopsis thaliana* ecotype Columbia; OE4-14 and OE11-7, two lines with single copy of *PHB8* overexpression construct; *phb8-3*, T-DNA insertion mutant line (Salk_092306).

levels. In the meantime, PHB8 OE lines showed much more significant enhancement of plant growth as compared to over-expressing ATPase beta subunit alone, indicating that other proteins in the regulatory network also played an important role in PHB8's regulation of plant growth.

4.4.6. Exploration of Molecular Mechanisms for PHB8's Regulation of Growth

In order to further explore the molecular mechanisms for PHB-regulated plant growth, pull down assays were carried out to verify the proteins interacting with PHB8. PHB8 protein was fused with Flag and BCCD tags to be transformed in Arabidopsis.

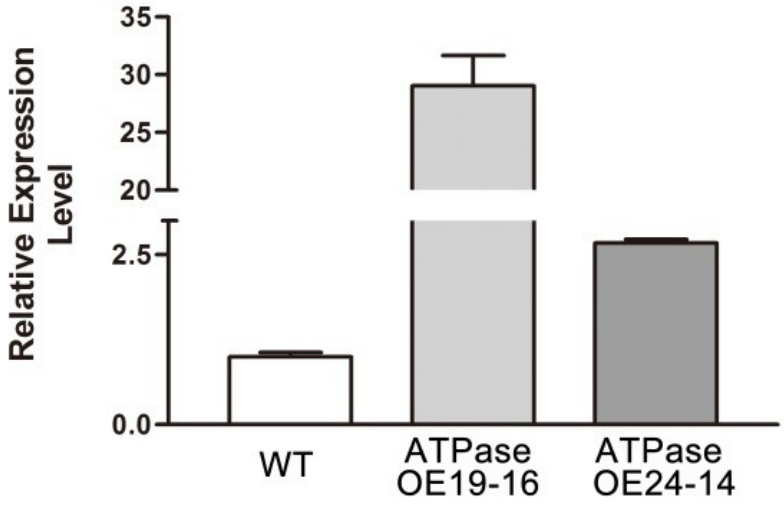

Fig. 4. 8. qRT-PCR result of ATPase OE line. The relative expression levels were obtained by comparing the transgenic lines with wild type. The relative ratio and 95% confidence interval were derived from triplicate assays.

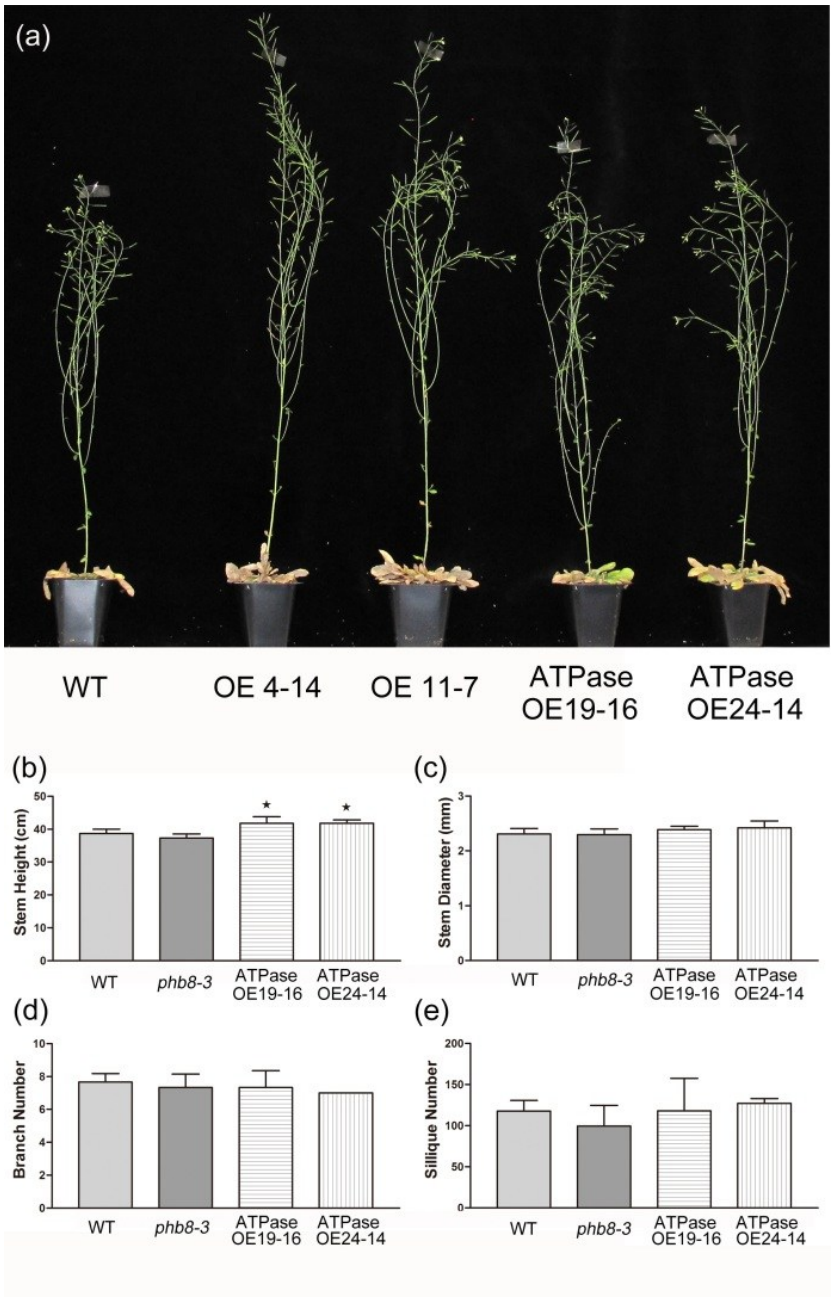

Fig. 4. 9. Phenotype of wild-type *Arabidopsis thaliana*, PHB8 and ATPase transgenic lines. WT, wild type *Arabidopsis thaliana* ecotype Columbia; OE4-14 and OE11-7, two lines with single copy of PHB8 overexpression construct; ATPase OE19-16 and OE24-14, two transgenic lines with ATPase overexpression construct. (a) Phenotypes of WT, PHB8 OE4-14, OE11-7 and ATPase OE19-16, OE24-14. (b) Stem height. (c) Stem diameter. (d) Branch number. (e) Silique number. Data are shown as means with standard deviation (n=6). Asterisks indicate difference level between transgenic and WT plants with *P< 0.01 using one-way ANOVA.

BCCD is a biotin carboxyl carrier protein domain of Arabidopsis 3-methylcrotonal CoA carboxylase, can be biotinylated in vivo with a high affinity and be captured by streptavidin-coupled Dynabeads magnetic beads. The BCCD tag was previously used to purify membrane bound protein complex (176). The potential candidate interacting with PHB8 protein was identified by liquid chromatograph coupled with tandem mass spectrometry. After comparison with the reference samples and removing contaminant proteins (Table A-3), the pull down assay consistently showed that PHB9 and PHB16 showed up in all three replicates. PHB14 were identified in one out of three replicates (Table 4.2). PHB8, 9, 14, and 16 are all among Class III prohibitins according to previous analysis. The result is consistent with previous studies showing that prohibitins may form homo-oligomers or hetero-oligomers to serve as molecular chaperons. Further localization analysis indicated that PHB8 can be located in both plasma membrane and mitochondrial membrane, indicating the protein may function in a broad membrane system (Fig. 4.10).

Table 4. 2. Co-IP result of *PHB8*-BCCD.

Protein name	Accession number	Times identified	Description
PHB8	AT3G01290.1	3	SPFH/Band 7/PHB domain-containing membrane-associated protein family
PHB16	AT5G62740.1	3	SPFH/Band 7/PHB domain-containing membrane-associated protein family
PHB9	AT1G69840.1 AT1G69840.2 AT1G69840.3 AT1G69840.4 AT1G69840.5 AT1G69840.6 AT1G69840.7	3	SPFH/Band 7/PHB domain-containing membrane-associated protein family
PSAB	ATCG00340.1	3	Photosystem I, PsaA/PsaB protein
PIP1B	AT2G45960.1	2	Plasma membrane intrinsic protein 1B
LHCA3	AT1G61520.1 AT1G61520.3	2	Photosystem I light harvesting complex gene 3
AHA2 **HA2**	AT4G30190.1 AT4G30190.2	2	H(+)-ATPase 2
AHA1	AT2G18960.1	2	H(+)-ATPase 1
PEN3	AT1G59870.1	2	ABC-2 and Plant PDR ABC-type transporter family protein
ATSPS1F	AT5G20280.1	2	Sucrose phosphate synthase 1F
AAC1	AT3G08580.1 AT3G08580.2	1	ADP/ATP carrier 1
ACA10	AT4G29900.1	1	Autoinhibited Ca(2+)-ATPase 10
ACC2	AT1G36180.1	1	Acetyl-CoA carboxylase 2
APX1	AT1G07890.1 AT1G07890.2 AT1G07890.3 AT1G07890.4 AT1G07890.5 AT1G07890.6 AT1G07890.7 AT1G07890.8	1	Ascorbate peroxidase 1
ATP5	AT5G13450.1	1	Delta subunit of Mt ATP synthase
ATPH	ATCG00140.1	1	ATP synthase subunit C family protein
AVP1	AT1G15690.1	1	Inorganic H pyrophosphatase family protein
BGL1U18	AT1G52400.1 AT1G52400.2 AT1G52400.3	1	Beta glucosidase 18

Table 4.2. Continued

Protein name	Accession number	Times identified	Description
CA2	AT5G14740.1 AT5G14740.2 AT5G14740.3 AT5G14740.4 AT5G14740.5	1	Carbonic anhydrase 2
CAB3 **CAB1** **CAB2**	AT1G29910.1 AT1G29930.1 AT1G29920.1	1	Chlorophyll A/B binding protein 3
CAT3	AT1G20620.1 AT1G20620.2 AT1G20620.4 AT1G20620.5	1	Catalase 3
Clathrin	AT3G08530.1	1	Clathrin, heavy chain
DEAD/DEAH box helicase	AT5G11170.1 AT5G11170.2 AT5G11200.1 AT5G11200.2 AT5G11200.3	1	DEAD/DEAH box RNA helicase family protein
EDA9	AT4G34200.1	1	D-3-phosphoglycerate dehydrogenase
HSC70-1	AT5G02500.1	1	Heat shock cognate protein 70-1
Hsp 70	AT3G09440.1	1	Heat shock protein 70 (Hsp 70) family protein
LHCA2	AT3G61470.1	1	Photosystem I light harvesting complex gene 2
LHCA4	AT3G47470.1	1	Light-harvesting chlorophyll-protein complex I subunit A4
LHCB5	AT4G10340.1	1	Light harvesting complex of photosystem II 5
MSS1	AT5G26340.1	1	Major facilitator superfamily protein
MTHSC70-2	AT5G09590.1	1	Mitochondrial HSP 70 2
PGM2	AT1G70730.1 AT1G70730.2 AT1G70730.3	1	Phosphoglucomutase/phosphomanno mutase family protein
PHB14	AT5G51570.1	1	SPFH/Band 7/PHB domain-containing membrane-associated protein family
PIP2A	AT3G53420.1 AT3G53420.2	1	Plasma membrane intrinsic protein 2A
PIP2A	AT3G53420.3	1	Plasma membrane intrinsic protein 2A
PSAA	ATCG00350.1	1	Photosystem I, PsaA/PsaB protein
PSAD-2 **PSAD-1**	AT1G03130.1 AT4G02770.1	1	Photosystem I subunit D

Table 4.2. Continued

Protein name	Accession number	Times identified	Description
PSAE-1	AT4G28750.1	1	Photosystem I reaction centre subunit IV / PsaE protein
PSAF	AT1G31330.1	1	Photosystem I subunit F
PSAG	AT1G55670.1	1	Photosystem I subunit G
PSAL	AT4G12800.1	1	Photosystem I subunit l
PSBB	ATCG00680.1	1	Photosystem II reaction center protein B
SYP122	AT3G52400.1	1	Syntaxin of plants 122
Pyruvate decarboxylase	AT5G17380.1	1	Thiamine pyrophosphate dependent pyruvate decarboxylase family protein
VHA-A	AT1G78900.1	1	Vacuolar ATP synthase subunit A

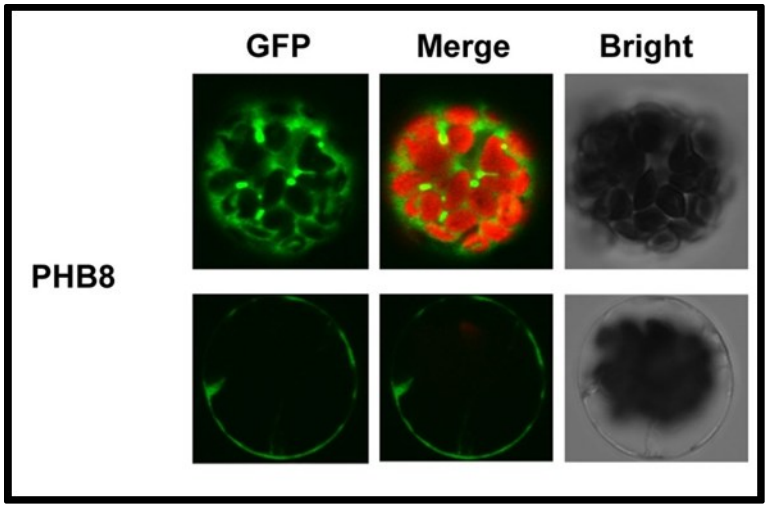

Fig. 4. 10. Subcellular localization of PHB8. Upper panel showed mitochondria localization and lower panel showed plasma membrane localization of PHB8.

4.5. DISCUSSION

4.5.1. An Overall Mechanism for PHB8

The PHB gene family is a conserved gene family which also is called SPFH (stomatins, prohibitins, flotillins and HflK/C) or band 7 domain containing proteins (68, 185). In mammals, the PHB1 was originally considered to inhibit cell proliferation but further studies assigned the function to the 3'-UTR of Phb1 mRNA (171, 186, 187). Several molecular mechanisms were proposed for PHB function. Firstly, prohibitins can regulate Rb/E2F and p53 pathways and thus affect tumorgenesis in breast cancer cells (188). Secondly, yeast prohibitins can function as a holdase/unfoldase to stabilize nascent mitochondrial proteins (58). Thirdly, the yeast PHB1/2 complex associates with m-AAA protease to regulate mitochondria proteins by negative control of m-AAA protease activity (59). It is most likely that the function of prohibitin is relevant to the subcellular localization and cell type (189).

In this study, I found that members of subclass III of the PHB family may form a molecular chaperone to stabilize mitochondrial proteins. Previous study showed that the prohibitins of subclass IV will form hetero-oligomers to regulate protein synthesis in yeast. It seems that the oligomerization of prohibitins requires their relatively high sequence similarity. The functions of the chaperone will not detriment substantially in the PHB8 single mutant to cause significant phenotype. By comprehensive analysis of all the data I obtained, a model was established for a PHB8-based membrane protein regulatory network is shown in Fig. 4.11. In this model, several PHB proteins, mainly PHB8, PHB9 and PHB16, may form a protein complex to stabilize the mitochondrial

membrane proteins such as F0/F1-type ATPase, which leads to higher ATP production and more efficient carbon utilization. Newly synthesized peptides may be stabilized by the PHB8 protein complex, but the identity of peptides needs to be further studied. Taking the various membrane localization of PHB8 into consideration, the complex may have distinct functions on different membrane systems. Moreover, other complex components may be recruited upon different stimuli.

4.5.2. Broad Perspectives

Study of the PHB gene family revealed functional diversity but conservative gene structure (68). In this study, I showed that Class III prohibitins could form protein complex to stabilize ATPase proteins, thus increasing ATP production and promoting plant growth and seed yield. It is not clear if PHB14 is a component of the complex or not based on the data. The discovery of function of prohibitin in the model plant Arabidopsis will be a hint for exploring prohibitin function in human. The involvement of prohibitins in regulating plant vegetative growth and plant pathogen defense may have relation with the subcellular localization. In this study, prohibitins are involved in energy metabolism by upregulating ATPase expression. Prohibitins were thought to be involved in diseases with increased oxidative stress and mitochondria dysfunction in humans, such as type 2 diabetes and obesity (189). The finding that prohibitins control ATP level may lead to the insight of how to deal this type of diseases and development of new therapeutic method. Furthermore, extensive studies are needed to determine if

PHB genes in other Class of PHB gene family will form protein complex that function as molecular chaperones or possess distinct functions.

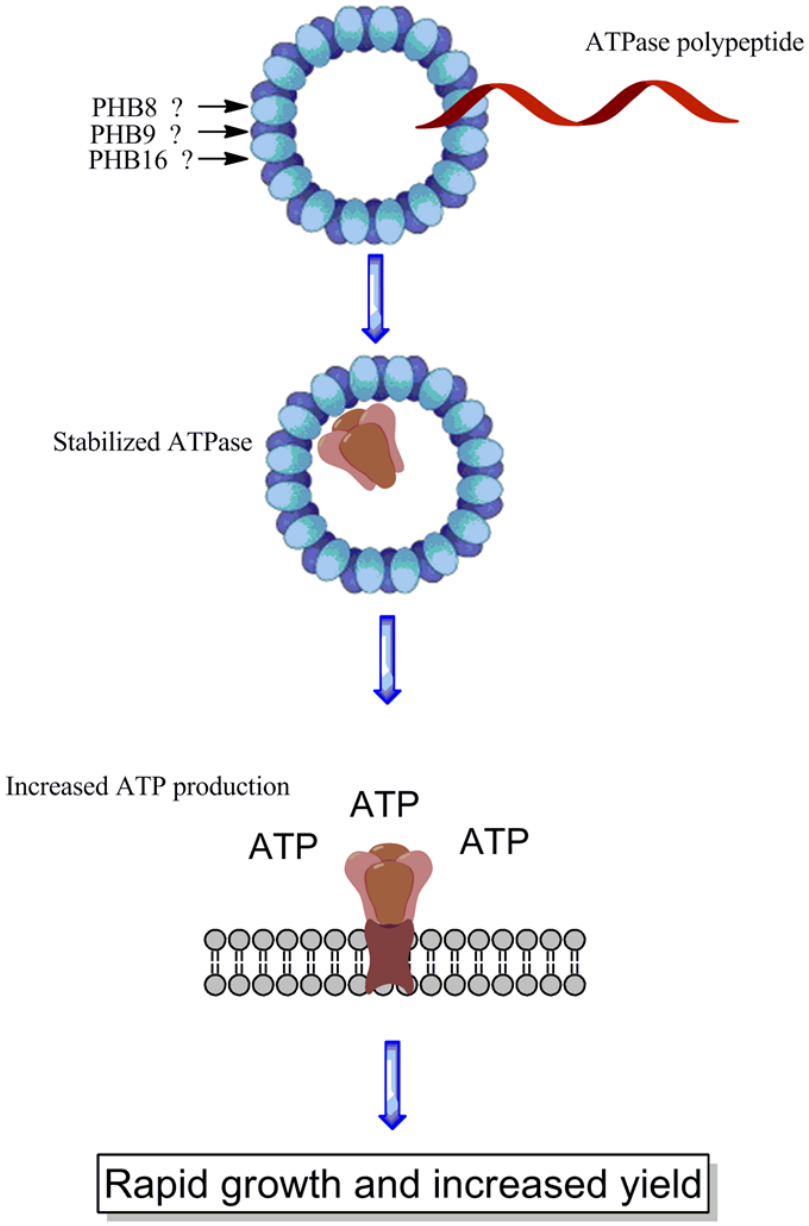

Fig. 4. 11. Model of *PHB8* regulates energy metabolism and plant growth

CHAPTER V

CONCLUSIONS AND FUTURE DIRECTIONS

5.1. CONCLUSIONS

The purpose of this work was to explore methods to improve protein identification in shot-gun proteomics study, and thus to apply new methods to biotechnology applications to improve plant defense and growth. A polymer, namely Polyethyleneimine, was used to remove Rubisco by means of fractionation to reduce the sample complexity. This method was carried out in rice and fall armyworm relationship studies. I also used a subproteomics method and bioinformatics tools to study plant responses to plant hormones. The organelle enriched method served as an efficient and time-saving approach to address plant growth related questions. The data given in the research showed that by reducing the sample complexity, the number of proteins identified increased dramatically. Lastly, I used bioinformatics, molecular biology and proteomics to identify PHB8 as a potential molecular chaperone to promote plant growth.

To dig out low abundant functional proteins, a Polyethyleneimine (PEI) assisted Rubisco cleanup protocol was developed as a new method combining both abundant protein removal and fractionation to reduce sample complexity. The new approach was applied to a plant-insect interaction study to validate the platform and investigate mechanisms for plant defense against herbivorous insects. Only 0.14% of Rubisco spectral counts were identified in the supernatant after PEI precipitation, on the other

hand, the Rubisco accounted for 36.12% in the pellet of PEI precipitation. The PARC method led to 1488 more proteins and a total of 68% increase in the number of proteins identified. As compared to differentially expressed protein identification, the PARC method enabled almost 2 times more than traditional methods. The results indicated that PEI can effectively remove Rubisco, improve the protein identification, and discover more differentially regulated proteins. Carbon re-distribution was indicated to play an essential role for plant-insect interactions by pathway analysis. Moreover, the transcriptomic validation and functional studies were carried out for two differentially regulated genes as revealed by PARC analysis. Insect resistance was induced by over-expressing either jacalin-like or cupin-like genes in rice. The leaf area damage assay showed the transgenic rice plants more resistant to fall armyworm larvae feeding.

Similarly, another rapid sample preparation method and bioinformatics classification systems have been integrated for comparative analysis of plant responses to plant hormones. As compared to traditional proteomics, the organelle-enrichment method both simplifies the sample preparation and increases the number of proteins identified in the targeted organelle. The organelle enriched method identified over 30% more mitochondrial proteins, but did not change significantly for chloroplast protein identification. A total of 355 and 288 differentially expressed proteins were identified in zeatin and BR treated samples respectively. Pathway analysis revealed both zeatin and brassinosteroid may down-regulate some key components in auxin responses. However, they have shown distinct induction and suppression of metabolic pathways in mitochondria and chloroplasts. For zeatin, the metabolic pathways in sucrose and starch

biosynthesis and utilization were significantly changed, yet the lipid biosynthesis remained unchanged. For BR, lipid biosynthesis and β-oxidation were both down-regulated, yet the changes in sucrose and starch metabolism were minor.

I conducted bioinformatics research which led to the discovery of PHB8 involved in the regulation of plant growth. The over-expression of PHB8 in Arabidopsis increased biomass and seed yield by increased stem height, diameter, branch and sillique number. A proteomics survey of PHB8 OE lines revealed ATPase beta subunit upregulated about 275 fold. The pool ATP level in PHB8 OE lines was about 10% higher than wild type, on the contrary, phb8 T-DNA mutant plant contained 6.61% less ATP compared to wild type. The subsequent pull-down assay of PHB8 implied PHB8/9/16 may form a protein complex, just like the PHB1/2 functions as a molecular chaperone in yeast. The mechanism of PHB8 regulated plant growth is proposed as a molecular chaperone to stabilize nascent polypeptides to facilitate protein folding, resulting in increased ATPase amount and ATP production.

5.2. FUTURE DIRECTIONS

For the improvement of protein identification in plant proteomics, not only the sample preparation and separation methods should be improved, but also the dynamic range and sensitivity of instruments, and the bioinformatics algorisms as well as the separation methods. The methods presented in this research could be an indication of how to reduce the complexity of protein samples and thus improves identification. To further study the mechanisms of plant growth regulation and defense responses, PTMs

(post translational modifications) and protein-protein interaction network studies should be performed. The studies carried out in this research focused on global proteome changes and pathway dynamics. With the demand of deep understanding of how proteins regulate plant response, PTMs will be the next aspect for particular protein to be studied. Different types of PTMs, such as phosphorylation, acetylation, methylation, ubiquitination and glycosylation, determine protein turnover and function. Meanwhile, the protein-protein interaction network analysis will help us find out the 'hub gene' in the network. Either acting as a connector between different regulation networks or controller within a regulatory network, the hub gene will be a pivotal for the plant functional protein study.

I proposed that PHB8/9/16 may form a protein complex to facilitate protein folding/unfolding on mitochondria membrane based on previous research and the results I obtained in this study. However, the formation and function of the protein complex should be studied by biochemistry and genetics approaches. The phenotype of the single PHB8 mutation was not remarkable compare to PHB8 over-expression lines. That may be due to the complementary effects of PHB9/16. So the double mutants among each pairs of PHB8/9/16 or triple mutant will be ideal genetics materials to study the function of the proposed protein complex. How many different polypeptides interact with PHB8 complex? Is PHB8 the center component of the complex? All these questions need to be addressed by further research.

REFERENCES

1. Chuang, H. Y., Hofree, M., and Ideker, T. (2010) A decade of systems biology. *Annu Rev Cell Dev Bi* **26**, 721-744.

2. Yates, J. R., Ruse, C. I., and Nakorchevsky, A. (2009) Proteomics by mass spectrometry: approaches, advances, and applications. *Annu Rev Biomed Eng* **11**, 49-79.

3. Smith, L. M., Kelleher, N. L., and Proteomics, C. T. D. (2013) Proteoform: a single term describing protein complexity. *Nat Methods* **10**, 186-187.

4. Lee, J., Garrett, W. M., and Cooper, B. (2007) Shotgun proteomic analysis of Arabidopsis thaliana leaves. *J Sep Sci* **30**, 2225-2230.

5. Arruda, S. C. C., Barbosa, H. D., Azevedo, R. A., and Arruda, M. A. Z. (2011) Two-dimensional difference gel electrophoresis applied for analytical proteomics: fundamentals and applications to the study of plant proteomics. *Analyst* **136**, 4119-4126.

6. Maor, R., Jones, A., Nuhse, T. S., Studholme, D. J., Peck, S. C., and Shirasu, K. (2007) Multidimensional protein identification technology (MudPIT) analysis of ubiquitinated proteins in plants. *Mol Cell Proteomics* **6**, 601-610.

7. Wolters, D. A., Washburn, M. P., and Yates, J. R. (2001) An automated multidimensional protein identification technology for shotgun proteomics. *Anal Chem* **73**, 5683-5690.

8. Yang, X. Y., Lu, J., Sun, X. S., and He, Q. Y. (2012) Application of subproteomics in the characterization of Gram-positive bacteria. *J Proteomics* **75**, 2803-2810.

9. Garbis, S., Lubec, G., and Fountoulakis, M. (2005) Limitations of current proteomics technologies. *J Chromatogr A* **1077**, 1-18.

10. Jung, E., Heller, M., Sanchez, J. C., and Hochstrasser, D. F. (2000) Proteomics meets cell biology: The establishment of subcellular proteomes. *Electrophoresis* **21**, 3369-3377.

11. Brunet, S., Thibault, P., Gagnon, E., Kearney, P., Bergeron, J. J. M., and Desjardins, M. (2003) Organelle proteomics: looking at less to see more. *Trends Cell Biol* **13**, 629-638.

12. Huber, L. A., Pfaller, K., and Vietor, I. (2003) Organelle proteomics - Implications for subcellular fractionation in proteomics. *Circ Res* **92**, 962-968.

13. Xing, T., Higgins, V. J., and Blumwald, E. (1996) Regulation of plant defense response to fungal pathogens: Two types of protein kinases in the reversible phosphorylation of the host plasma membrane H+-ATPase. *Plant Cell* **8**, 555-564.

14. Deprost, D., Yao, L., Sormani, R., Moreau, M., Leterreux, G., Nicolai, M., Bedu, M., Robaglia, C., and Meyer, C. (2007) The Arabidopsis TOR kinase links plant growth, yield, stress resistance and mRNA translation. *Embo Rep* **8**, 864-870.

15. Zheng, Z. F., Xu, X. P., Crosley, R. A., Greenwalt, S. A., Sun, Y. J., Blakeslee, B., Wang, L. Z., Ni, W. T., Sopko, M. S., Yao, C. L., Yau, K., Burton, S., Zhuang, M. B., McCaskill, D. G., Gachotte, D., Thompson, M., and Greene, T. W. (2010) The protein kinase snrk2.6 mediates the regulation of sucrose metabolism and plant growth in Arabidopsis. *Plant Physiol* **153**, 99-113.

16. Yuan, J. S., Kollner, T. G., Wiggins, G., Grant, J., Degenhardt, J., and Chen, F. (2008) Molecular and genomic basis of volatile-mediated indirect defense against insects in rice. *Plant J* **55**, 491-503.

17. Ahearn, I. M., Haigis, K., Bar-Sagi, D., and Philips, M. R. (2012) Regulating the regulator: post-translational modification of RAS. *Nat Rev Mol Cell Bio* **13**, 39-51.

18. Waters, K. M., Pounds, J. G., and Thrall, B. D. (2006) Data merging for integrated microarray and proteomic analysis. *Briefings in functional genomics & proteomics* **5**, 261-272.

19. Huang, B., Chu, C. H., Chen, S. L., Juan, H. F., and Chen, Y. M. (2006) A proteomics study of the mung bean epicotyl regulated by brassinosteroids under conditions of chilling stress. *Cell Mol Biol Lett* **11**, 264-278.

20. Deng, Z. P., Zhang, X., Tang, W. Q., Oses-Prieto, J. A., Suzuki, N., Gendron, J. M., Chen, H. J., Guan, S. H., Chalkley, R. J., Peterman, T. K., Burlingame, A. L., and Wang, Z. Y. (2007) A proteomics study of brassinosteroid response in Arabidopsis. *Mol Cell Proteomics* **6**, 2058-2071.

21. Konishi, H., and Komatsu, S. (2003) A proteomics approach to investigating promotive effects of brassinolide on lamina inclination and root growth in rice seedlings. *Biological & pharmaceutical bulletin* **26**, 401-408.

22. Wang, F., Bai, M. Y., Deng, Z., Oses-Prieto, J. A., Burlingame, A. L., Lu, T., Chong, K., and Wang, Z. Y. (2010) Proteomic study identifies proteins involved in brassinosteroid regulation of rice growth. *J Integr Plant Biol* **52**, 1075-1085.

23. Tang, W. Q., Deng, Z. P., and Wang, Z. Y. (2010) Proteomics shed light on the brassinosteroid signaling mechanisms. *Curr Opin Plant Biol* **13**, 27-33.

24. Tang, W. Q., Deng, Z. P., Oses-Prieto, J. A., Suzuki, N., Zhu, S. W., Zhang, X., Burlingame, A. L., and Wang, Z. Y. (2008) Proteomics studies of brassinosteroid signal transduction using prefractionation and two-dimensional DIGE. *Mol Cell Proteomics* **7**, 728-738.

25. Yang, G., and Komatsu, S. (2004) Microarray and proteomic analysis of brassinosteroid- and gibberellin-regulated gene and protein expression in rice. *Genomics, proteomics & bioinformatics* **2**, 77-83.

26. Chiarappa, L., Chiang, H. C., and Smith, R. F. (1972) Plant pests and diseases: assessment of crop losses. *Science* **176**, 769-773.

27. Mithofer, A., and Boland, W. (2012) Plant defense against herbivores: chemical aspects. *Annu Rev Plant Biol* **63**, 431-450.

28. Furstenberg-Hagg, J., Zagrobelny, M., and Bak, S. (2013) Plant defense against insect herbivores. *Int J Mol Sci* **14**, 10242-10297.

29. Singh, V., Louis, J., Ayre, B. G., Reese, J. C., and Shah, J. (2011) TREHALOSE PHOSPHATE SYNTHASE11-dependent trehalose metabolism promotes Arabidopsis thaliana defense against the phloem-feeding insect Myzus persicae. *Plant J* **67**, 94-104.

30. Hao, P. Y., Liu, C. X., Wang, Y. Y., Chen, R. Z., Tang, M., Du, B., Zhu, L. L., and He, G. (2008) Herbivore-induced callose deposition on the sieve plates of rice: An important mechanism for host resistance. *Plant Physiol* **146**, 1810-1820.

31. Singh, V., and Shah, J. (2012) Tomato responds to green peach aphid infestation with the activation of trehalose metabolism and starch accumulation. *Plant signaling & behavior* **7**, 605-607.

32. Thompson, G. A., and Goggin, F. L. (2006) Transcriptomics and functional genomics of plant defence induction by phloem-feeding insects. *J Exp Bot* **57**, 755-766.

33. Zhang, J. H., Sun, L. W., Liu, L. L., Lian, J., An, S. L., Wang, X., Zhang, J., Jin, J. L., Li, S. Y., and Xi, J. H. (2010) Proteomic analysis of interactions between the generalist herbivore *Spodoptera exigua* (Lepidoptera: Noctuidae) and *Arabidopsis thaliana*. *Plant Mol Biol Rep* **28**, 324-333.

34. Francis, F., Gerkens, P., Harmel, N., Mazzucchelli, G., De Pauw, E., and Haubruge, E. (2006) Proteomics in Myzus persicae: Effect of aphid host plant switch. *Insect Biochem Molec* **36**, 219-227.

35. Wittstock, U., Agerbirk, N., Stauber, E. J., Olsen, C. E., Hippler, M., Mitchell-Olds, T., Gershenson, J., and Vogel, H. (2004) Successful herbivore attack due to metabolic diversion of a plant chemical defense. *P Natl Acad Sci USA* **101**, 4859-4864.

36. Giri, A. P., Wunsche, H., Mitra, S., Zavala, J. A., Muck, A., Svatos, A., and Baldwin, I. T. (2006) Molecular interactions between the specialist herbivore Manduca sexta (Lepidoptera, Sphingidae) and its natural host Nicotiana attenuata. VII. Changes in the plant's proteome. *Plant Physiol* **142**, 1621-1641.

37. Chen, H., Gonzales-Vigil, E., Wilkerson, C. G., and Howe, G. A. (2007) Stability of plant defense proteins in the gut of insect herbivores. *Plant Physiol* **143**, 1954-1967.

38. Chen, H., Wilkerson, C. G., Kuchar, J. A., Phinney, B. S., and Howe, G. A. (2005) Jasmonate-inducible plant enzymes degrade essential amino acids in the herbivore midgut. *P Natl Acad Sci USA* **102**, 19237-19242.

39. Gonzales-Vigil, E., Bianchetti, C. M., Phillips, G. N., and Howe, G. A. (2011) Adaptive evolution of threonine deaminase in plant defense against insect herbivores. *P Natl Acad Sci USA* **108**, 5897-5902.

40. Kirsch, R., Wielsch, N., Vogel, H., Svatos, A., Heckel, D. G., and Pauchet, Y. (2012) Combining proteomics and transcriptome sequencing to identify active plant-cell-wall-degrading enzymes in a leaf beetle. *Bmc Genomics* **13**, 587.

41. Philippe, R. N., Ralph, S. G., Kulheim, C., Jancsik, S. I., and Bohlmann, J. (2009) Poplar defense against insects: genome analysis, full-length cDNA cloning, and transcriptome and protein analysis of the poplar Kunitz-type protease inhibitor family. *New Phytol* **184**, 865-884.

42. Mishra, M., Tamhane, V. A., Khandelwal, N., Kulkarni, M. J., Gupta, V. S., and Giri, A. P. (2010) Interaction of recombinant CanPIs with Helicoverpa armigera gut proteases reveals their processing patterns, stability and efficiency. *Proteomics* **10**, 2845-2857.

43. De Leo, F., Volpicella, M., Licciulli, F., Liuni, S., Gallerani, R., and Ceci, L. R. (2002) PLANT-PIs: a database for plant protease inhibitors and their genes. *Nucleic Acids Res* **30**, 347-348.

44. Nogueira, F. C. S., Silva, C. P., Alexandre, D., Samuels, R. I., Soares, E. L., Aragao, F. J. L., Palmisano, G., Domont, G. B., Roepstorff, P., and Campos, F. A. P.

(2012) Global proteome changes in larvae of Callosobruchus maculatus Coleoptera:Chrysomelidae:Bruchinae) following ingestion of a cysteine proteinase inhibitor. *Proteomics* **12**, 2704-2715.

45. An Nguyen, T. T., Michaud, D., and Cloutier, C. (2007) Proteomic profiling of aphid Macrosiphum euphorbiae responses to host-plant-mediated stress induced by defoliation and water deficit. *J Insect Physiol* **53**, 601-611.

46. Francis, F., Guillonneau, F., Leprince, P., De Pauw, E., Haubruge, E., Jia, L., and Goggin, F. L. (2010) Tritrophic interactions among Macrosiphum euphorbiae aphids, their host plants and endosymbionts: investigation by a proteomic approach. *J Insect Physiol* **56**, 575-585.

47. Lippert, D., Chowrira, S., Ralph, S. G., Zhuang, J., Aeschliman, D., Ritland, C., Ritland, K., and Bohlmann, J. (2007) Conifer defense against insects: Proteome analysis of Sitka spruce (Picea sitchensis) bark induced by mechanical wounding or feeding by white pine weevils (Pissodes strobi). *Proteomics* **7**, 248-270.

48. George, D., Babalola, O. O., and Gatehouse, A. M. R. (2011) Differential protein expression in maize (Zea mays) in response to insect attack. *Afr J Biotechnol* **10**, 7700-7709.

49. Ferry, N., Stavroulakis, S., Guan, W. Z., Davison, G. M., Bell, H. A., Weaver, R. J., Down, R. E., Gatehouse, J. A., and Gatehouse, A. M. R. (2011) Molecular interactions between wheat and cereal aphid (Sitobion avenae): Analysis of changes to the wheat proteome. *Proteomics* **11**, 1985-2002.

50. Schouppe, D., Ghesquiere, B., Menschaert, G., De Vos, W. H., Bourque, S., Trooskens, G., Proost, P., Gevaert, K., and Van Damme, E. J. M. (2011) Interaction of the tobacco lectin with histone proteins. *Plant Physiol* **155**, 1091-1102.

51. Sangha, J. S., Yolanda, H. C., Kaur, J., Khan, W., Abduljaleel, Z., Alanazi, M. S., Mills, A., Adalla, C. B., Bennett, J., Prithiviraj, B., Jahn, G. C., and Leung, H. (2013) Proteome analysis of rice (Oryza sativa l.) mutants reveals differentially induced proteins during brown planthopper (Nilaparvata lugens) infestation. *Int J Mol Sci* **14**, 3921-3945.

52. Lippert, D. N., Ralph, S. G., Phillips, M., White, R., Smith, D., Hardie, D., Gershenzon, J., Ritland, K., Borchers, C. H., and Bohlmann, J. (2009) Quantitative iTRAQ proteome and comparative transcriptome analysis of elicitor-induced Norway spruce (Picea abies) cells reveals elements of calcium signaling in the early conifer defense response. *Proteomics* **9**, 350-367.

53. Zulak, K. G., Lippert, D. N., Kuzyk, M. A., Domanski, D., Chou, T., Borchers, C. H., and Bohlmann, J. (2009) Targeted proteomics using selected reaction monitoring reveals the induction of specific terpene synthases in a multi-level study of methyl jasmonate-treated Norway spruce (Picea abies). *Plant J* **60**, 1015-1030.

54. Hall, D. E., Robert, J. A., Keeling, C. I., Domanski, D., Quesada, A. L., Jancsik, S., Kuzyk, M. A., Hamberger, B., Borchers, C. H., and Bohlmann, J. (2011) An integrated genomic, proteomic and biochemical analysis of (+)-3-carene biosynthesis in Sitka spruce (Picea sitchensis) genotypes that are resistant or susceptible to white pine weevil. *Plant J* **65**, 936-948.

55. Howe, G. A., and Jander, G. (2008) Plant immunity to insect herbivores. *Annu Rev Plant Biol* **59**, 41-66.

56. Welburn, S. C., and Murphy, N. B. (1998) Prohibitin and RACK homologues are up-regulated in trypanosomes induced to undergo apoptosis and in naturally occurring terminally differentiated forms. *Cell Death Differ* **5**, 615-622.

57. Kirchman, P. A., Miceli, M. V., West, R. L., Jiang, J. C., Kim, S., and Jazwinski, S. M. (2003) Prohibitins and Ras2 protein cooperate in the maintenance of mitochondrial function during yeast aging. *Acta Biochim Pol* **50**, 1039-1056.

58. Nijtmans, L. G. J., de Jong, L., Sanz, M. A., Coates, P. J., Berden, J. A., Back, J. W., Muijsers, A. O., van der Spek, H., and Grivell, L. A. (2000) Prohibitins act as a membrane-bound chaperone for the stabilization of mitochondrial proteins. *Embo J* **19**, 2444-2451.

59. Steglich, G., Neupert, W., and Langer, T. (1999) Prohibitins regulate membrane protein degradation by the m-AAA protease in mitochondria. *Mol Cell Biol* **19**, 3435-3442.

60. Browman, D. T., Resek, M. E., Zajchowski, L. D., and Robbins, S. M. (2006) Erlin-1 and erlin-2 are novel members of the prohibitin family of proteins that define lipid-raft-like domains of the ER. *J Cell Sci* **119**, 3149-3160.

61. Sievers, C., Billig, G., Gottschalk, K., and Rudel, T. (2010) Prohibitins Are Required for Cancer Cell Proliferation and Adhesion. *Plos One* **5**, e12735.

62. Coates, P. J., Jamieson, D. J., Smart, K., Prescott, A. R., and Hall, P. A. (1997) The prohibitin family of mitochondrial proteins regulate replicative lifespan. *Curr Biol* **7**, 607-610.

63. Artal-Sanz, M., and Tavernarakis, N. (2009) Prohibitin couples diapause signalling to mitochondrial metabolism during ageing in C-elegans. *Nature* **461**, 793-797.

64. Jo, Y., Sguigna, P. V., and Debose-Boyd, R. A. (2011) Membrane-associated ubiquitin ligase complex containing gp78 mediates sterol-accelerated degradation of 3-hydroxy-3-methylglutaryl-coenzyme a reductase. *J Biol Chem* **286**, 15022-15031.

65. Kuramori, C., Azuma, M., Kume, K., Kaneko, Y., Inoue, A., Yamaguchi, Y., Kabe, Y., Hosoya, T., Kizaki, M., Suematsu, M., and Handa, H. (2009) Capsaicin binds to prohibitin 2 and displaces it from the mitochondria to the nucleus. *Biochem Bioph Res Co* **379**, 519-525.

66. Chander, H., Halpern, M., Resnick-Silverman, L., Manfredi, J. J., and Germain, D. (2011) Skp2B overexpression alters a prohibitin-p53 axis and the transcription of papp-a, the protease of insulin-like growth factor binding protein 4. *Plos One* **6**, e22456.

67. Snedden, W. A., and Fromm, H. (1997) Characterization of the plant homologue of prohibitin, a gene associated with antiproliferative activity in mammalian cells. *Plant Mol Biol* **33**, 753-756.

68. Di, C., Xu, W. Y., Su, Z., and Yuan, J. S. (2010) Comparative genome analysis of PHB gene family reveals deep evolutionary origins and diverse gene function. *Bmc Bioinformatics* **11**:S22.

69. Chen, J. C., Jiang, C. Z., and Reid, M. S. (2005) Silencing a prohibitin alters plant development and senescence. *Plant J* **44**, 16-24.

70. Qi, Y. P., Tsuda, K., Nguyen, L. V., Wang, X., Lin, J. S., Murphy, A. S., Glazebrook, J., Thordal-Christensen, H., and Katagiri, F. (2011) Physical association of arabidopsis hypersensitive induced reaction proteins (HIRs) with the immune receptor RPS2. *J Biol Chem* **286**, 31297-31307.

71. Van Aken, O., Pecenkova, T., van de Cotte, B., De Rycke, R., Eeckhout, D., Fromm, H., De Jaeger, G., Witters, E., Beemster, G. T. S., Inze, D., and Van Breusegem, F. (2007) Mitochondrial type-I prohibitins of Arabidopsis thaliana are required for supporting proficient meristem development. *Plant J* **52**, 850-864.

72. Wang, Y., Ries, A., Wu, K. T., Yang, A., and Crawford, N. M. (2010) The arabidopsis prohibitin gene phb3 functions in nitric oxide-mediated responses and in hydrogen peroxide-induced nitric oxide accumulation. *Plant Cell* **22**, 249-259.

73. Ahn, C. S., Lee, J. H., Hwang, A. R., Kim, W. T., and Pai, H. S. (2006) Prohibitin is involved in mitochondrial biogenesis in plants. *Plant J* **46**, 658-667.

74. Weis, C., Pfeilmeier, S., Glawischnig, E., Isono, E., Pachl, F., Hahne, H., Kuster, B., Eichmann, R., and Huckelhoven, R. (2013) Co-immunoprecipitation-based identification of putative BAX INHIBITOR-1-interacting proteins involved in cell death regulation and plant-powdery mildew interactions. *Mol Plant Pathol* **14**, 791-802.

75. Wagner, K., Miliotis, T., Marko-Varga, G., Bischoff, R., and Unger, K. K. (2002) An automated on-line multidimensional HPLC system for protein and peptide mapping with integrated sample preparation. *Anal Chem* **74**, 809-820.

76. Rose, J. K. C., Bashir, S., Giovannoni, J. J., Jahn, M. M., and Saravanan, R. S. (2004) Tackling the plant proteome: practical approaches, hurdles and experimental tools. *Plant J* **39**, 715-733.

77. Widjaja, I., Naumann, K., Roth, U., Wolf, N., Mackey, D., Dangl, J. L., Scheel, D., and Lee, J. (2009) Combining subproteome enrichment and Rubisco depletion enables identification of low abundance proteins differentially regulated during plant defense. *Proteomics* **9**, 138-147.

78. Krishnan, H. B., and Natarajan, S. S. (2009) A rapid method for depletion of Rubisco from soybean (Glycine max) leaf for proteomic analysis of lower abundance proteins. *Phytochemistry* **70**, 1958-1964.

79. Atkinson, A., and Jack, G. W. (1973) Precipitation of nucleic acids with polyethyleneimine and the chromatography of nucleic acids and proteins on immobilised polyethyleneimine. *Biochimica et biophysica acta* **308**, 41-52.

80. Holler, C., and Zhang, C. M. (2008) Purification of an acidic recombinant protein from transgenic tobacco. *Biotechnol Bioeng* **99**, 902-909.

81. Jorrin-Novo, J. V., Maldonado, A. M., Echevarria-Zomeno, S., Valledor, L., Castillejo, M. A., Curto, M., Valero, J., Sghaier, B., Donoso, G., and Redondo, I. (2009) Plant proteomics update (2007-2008): Second-generation proteomic techniques, an appropriate experimental design, and data analysis to fulfill MIAPE standards, increase plant proteome coverage and expand biological knowledge. *J Proteomics* **72**, 285-314.

82. Yuan, J. S., Galbraith, D. W., Dai, S. Y., Griffin, P., and Stewart, C. N. (2008) Plant systems biology comes of age. *Trends Plant Sci* **13**, 165-171.

83. Zhang, Y. X., Gao, P., and Yuan, J. S. (2010) Plant protein-protein interaction network and interactome. *Curr Genomics* **11**, 40-46.

84. Carrillo, M. G. C., Goodwin, P. H., Leach, J. E., Leung, H., and Cruz, C. M. V. (2009) Phylogenomic relationships of rice oxalate oxidases to the cupin superfamily and their association with disease resistance QTL. *Rice* **2**, 67-79.

85. Banerjee, J., and Maiti, M. K. (2010) Functional role of rice germin-like protein1 in regulation of plant height and disease resistance. *Biochem Bioph Res Co* **394**, 178-183.

86. Davidson, R. M., Reeves, P. A., Manosalva, P. M., and Leach, J. E. (2009) Germins: A diverse protein family important for crop improvement. *Plant Sci* **177**, 499-510.

87. Lou, Y. G., and Baldwin, I. T. (2006) Silencing of a germin-like gene in Nicotiana attenuata improves performance of native herbivores. *Plant Physiol* **140**, 1126-1136.

88. Mechin, V., Damerval, C., and Zivy, M. (2007) Total protein extraction with TCA-acetone. *Methods Mol Biol* **355**, 1-8.

89. Shevchenko, A., Tomas, H., Havlis, J., Olsen, J. V., and Mann, M. (2006) In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nature protocols* **1**, 2856-2860.

90. Loewen, P. C., Didychuk, A. L., Switala, J., Perez-Luque, R., Fita, I., and Loewen, M. C. (2013) Structure of Pisum sativum Rubisco with bound ribulose 1,5-bisphosphate. *Acta Crystallogr F* **69**, 10-14.

91. Washburn, M. P., Wolters, D., and Yates, J. R. (2001) Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat Biotechnol* **19**, 242-247.

92. Zhang, Y., Liu, S., Dai, S., and Yuan, J. (2012) Integration of shot-gun proteomics and bioinformatics analysis to explore plant hormone responses. *Bmc Bioinformatics* **13**, S8.

93. Xu, T., Venable, J., Park, S., Cociorva, D., Lu, B., Liao, L., Wohlschlegel, J., Hewel, J., and Yates III, J. (2006) ProLuCID, a fast and sensitive tandem mass spectra-based protein identification program. *Mol Cell Proteomics* **5**, S174.

94. Carvalho, P. C., Fischer, J. S., Chen, E. I., Yates, J. R., and Barbosa, V. C. (2008) PatternLab for proteomics: a tool for differential shotgun proteomics. *Bmc Bioinformatics* **9**:316.

95. Du, Z., Zhou, X., Ling, Y., Zhang, Z. H., and Su, Z. (2010) agriGO: a GO analysis toolkit for the agricultural community. *Nucleic Acids Res* **38**, W64-W70.

96. Kanehisa, M. (2002) The KEGG database. *Novart Fdn Symp* **247**, 91-103.

97. Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., and Higgins, D. G. (2007) Clustal W and clustal X version 2.0. *Bioinformatics* **23**, 2947-2948.

98. Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* **28**, 2731-2739.

99. Portillo, M., Fenoll, C., and Escobar, C. (2006) Evaluation of different RNA extraction methods for small quantities of plant tissue: Combined effects of reagent type and homogenization procedure on RNA quality-integrity and yield. *Physiol Plantarum* **128**, 1-7.

100. Chen, S. B., Songkumarn, P., Liu, J. L., and Wang, G. L. (2009) A versatile zero background t-vector system for gene cloning and functional genomics. *Plant Physiol* **150**, 1111-1121.

101. Toki, S., Hara, N., Ono, K., Onodera, H., Tagiri, A., Oka, S., and Tanaka, H. (2006) Early infection of scutellum tissue with Agrobacterium allows high-speed transformation of rice. *Plant J* **47**, 969-976.

102. Baviskar, S. N. (2011) A quick & automated method for measuring cell area using ImageJ. *Am Biol Teach* **73**, 554-556.

103. Venu, R., Madhav, M. S., Sreerekha, M., Nobuta, K., Zhang, Y., Carswell, P., Boehm, M. J., Meyers, B. C., Korth, K. L., and Wang, G.-L. (2010) Deep and comparative transcriptome analysis of rice plants infested by the beet armyworm (Spodoptera exigua) and water weevil (Lissorhoptrus oryzophilus). *Rice* **3**, 22-35.

104. Xi, J. H., Wang, X., Li, S. Y., Zhou, X., Yue, L., Fan, J., and Hao, D. Y. (2006) Polyethylene glycol fractionation improved detection of low-abundant proteins by two-dimensional electrophoresis analysis of plant proteome. *Phytochemistry* **67**, 2341-2348.

105. Peltier, J. B., Ripoll, D. R., Friso, G., Rudella, A., Cai, Y., Ytterberg, J., Giacomelli, L., Pillardy, J., and van Wijk, K. J. (2004) Clp protease complexes from photosynthetic and non-photosynthetic plastids and mitochondria of plants, their

predicted three-dimensional structures, and functional implications. *J Biol Chem* **279**, 4768-4781.

106. Hashimoto, M., and Komatsu, S. (2007) Proteomic analysis of rice seedlings during cold stress. *Proteomics* **7**, 1293-1302.

107. Alam, I., Sharmin, S., Kim, K. H., Kim, Y. G., Lee, J., and Lee, B. H. (2012) An improved plant leaf protein extraction method for high resolution two-dimensional polyacrylamide gel electrophoresis and comparative proteomics. *Biotechnic & histochemistry : official publication of the Biological Stain Commission* **88**, 61-75.

108. Gammulla, C. G., Pascovici, D., Atwell, B. J., and Haynes, P. A. (2011) Differential proteomic response of rice (Oryza sativa) leaves exposed to high- and low-temperature stress. *Proteomics* **11**, 2839-2850.

109. Mirzaei, M., Soltani, N., Sarhadi, E., Pascovici, D., Keighley, T., Salekdeh, G. H., Haynes, P. A., and Atwell, B. J. (2012) Shotgun proteomic analysis of long-distance drought signaling in rice roots. *J Proteome Res* **11**, 348-358.

110. Castellana, N. E., Payne, S. H., Shen, Z. X., Stanke, M., Bafna, V., and Briggs, S. P. (2008) Discovery and revision of Arabidopsis genes by proteogenomics. *P Natl Acad Sci USA* **105**, 21034-21038.

111. Steinbrenner, A. D., Gomez, S., Osorio, S., Fernie, A. R., and Orians, C. M. (2011) Herbivore-induced changes in tomato (Solanum lycopersicum) primary metabolism: a whole plant perspective. *J Chem Ecol* **37**, 1294-1303.

112. Uefune, M., Kugimiya, S., Sano, K., and Takabayashi, J. (2012) Herbivore-induced plant volatiles enhance the ability of parasitic wasps to find hosts on a plant. *J Appl Entomol* **136**, 133-138.

113. Clancy, K. M. (1992) The role of sugars in western spruce budworm nutritional ecology. *Ecol Entomol* **17**, 189-197.

114. Awmack, C. S., and Leather, S. R. (2002) Host plant quality and fecundity in herbivorous insects. *Annu Rev Entomol* **47**, 817-844.

115. Moghaddam, M. R. B., and Van den Ende, W. (2012) Sugars and plant innate immunity. *J Exp Bot* **63**, 3989-3998.

116. Gomez-Ariza, J., Campo, S., Rufat, M., Estopa, M., Messeguer, J., San Segundo, B., and Coca, M. (2007) Sucrose-mediated priming of plant defense responses and broad-spectrum disease resistance by overexpression of the maize pathogenesis-related PRms protein in rice plants. *Mol Plant Microbe In* **20**, 832-842.

117. Singh, V., Louis, J., Ayre, B. G., Reese, J. C., Pegadaraju, V., and Shah, J. (2011) TREHALOSE PHOSPHATE SYNTHASE11-dependent trehalose metabolism promotes Arabidopsis thaliana defense against the phloem-feeding insect Myzus persicae (vol 67, pg 94, 2011). *Plant J* **68**, 938-938.

118. Schmidt, L., Schurr, U., and Rose, U. S. R. (2009) Local and systemic effects of two herbivores with different feeding mechanisms on primary metabolism of cotton leaves. *Plant Cell Environ* **32**, 893-903.

119. Heil, M., Ibarra-Laclette, E., Adame-Alvarez, R. M., Martinez, O., Ramirez-Chavez, E., Molina-Torres, J., and Herrera-Estrella, L. (2012) How plants sense wounds:

damaged-self recognition is based on plant-derived elicitors and induces octadecanoid signaling. *Plos One* **7**,e30537.

120. Sharon, N., and Lis, H. (2004) History of lectins: from hemagglutinins to biological recognition molecules. *Glycobiology* **14**, 53r-62r.

121. Jiang, S. Y., Ma, Z. G., and Ramachandran, S. (2010) Evolutionary history and stress regulation of the lectin superfamily in higher plants. *Bmc Evol Biol* **10**:79.

122. Van Damme, E. J. M., Barre, A., Rouge, P., and Peumans, W. J. (2004) Cytoplasmic/nuclear plant lectins: a new story. *Trends Plant Sci* **9**, 484-489.

123. Nagano, A. J., Fukao, Y., Fujiwara, M., Nishimura, M., and Hara-Nishimura, I. (2008) Antagonistic jacalin-related lectins regulate the size of ER body-type beta-glucosidase complexes in Arabidopsis thaliana. *Plant Cell Physiol* **49**, 969-980.

124. Vasconcelos, I. M., and Oliveira, J. T. A. (2004) Antinutritional properties of plant lectins. *Toxicon* **44**, 385-403.

125. Dunwell, J. M., Khuri, S., and Gane, P. J. (2000) Microbial relatives of the seed storage proteins of higher plants: Conservation of structure and diversification of function during evolution of the cupin superfamily. *Microbiol Mol Biol R* **64**, 153-179.

126. Dunwell, J. M., Gibbings, J. G., Mahmood, T., and Naqvi, S. M. S. (2008) Germin and germin-like proteins: Evolution, structure, and function. *Crit Rev Plant Sci* **27**, 342-375.

127. Berna, A., and Bernier, F. (1999) Regulation by biotic and abiotic stress of a wheat germin gene encoding oxalate oxidase, a H2O2-producing enzyme. *Plant Mol Biol* **39**, 539-549.

128. Lane, B. G., Dunwell, J. M., Ray, J. A., Schmitt, M. R., and Cuming, A. C. (1993) Germin, a protein marker of early plant development, is an oxalate oxidase. *J Biol Chem* **268**, 12239-12242.

129. Ramputh, A. I., Arnason, J. T., Cass, L., and Simmonds, J. A. (2002) Reduced herbivory of the European corn borer (Ostrinia nubilalis) on corn transformed with germin, a wheat oxalate oxidase gene. *Plant Sci* **162**, 431-440.

130. Gatto, L., Vizcaino, J. A., Hermjako, H., Huber, W., and Lilley, K. S. (2010) Organelle proteomics experimental designs and analysis. *Proteomics* **10**, 3957-3969.

131. Agrawal, G. K., Bourguignon, J., Rolland, N., Ephritikhine, G., Ferro, M., Jaquinod, M., Alexiou, K. G., Chardot, T., Chakraborty, N., Jolivet, P., Doonan, J. H., and Rakwal, R. (2011) Plant organelle proteomics: collaborating for optimal cell function. *Mass Spectrom Rev* **30**, 772-853.

132. Dunkley, T. P. J., Hester, S., Shadforth, I. P., Runions, J., Weimar, T., Hanton, S. L., Griffin, J. L., Bessant, C., Brandizzi, F., Hawes, C., Watson, R. B., Dupree, P., and Lilley, K. S. (2006) Mapping the Arabidopsis organelle proteome. *P Natl Acad Sci USA* **103**, 6518-6523.

133. Lilley, K. S., and Dupree, P. (2006) Methods of quantitative proteomics and their application to plant organelle characterization. *J Exp Bot* **57**, 1493-1499.

134. Tzou, D. S., Galson, E. C., and Sondheimer, E. (1973) The metabolism of hormones during seed germination and release from dormancy: iii. the effects and metabolism of zeatin in dormant and nondormant ash embryos. *Plant Physiol* **51**, 894-897.

135. Mandava, N. B. (1988) Plant growth-promoting brassinosteroids. *Annual Review of Plant Physiology and Plant Molecular Biology* **39**, 23-52.

136. Santoni, V. (2007) Plant plasma membrane protein extraction and solubilization for proteomic analysis. *Methods Mol Biol* **355**, 93-109.

137. Katari, M. S., Nowicki, S. D., Aceituno, F. F., Nero, D., Kelfer, J., Thompson, L. P., Cabello, J. M., Davidson, R. S., Goldberg, A. P., Shasha, D. E., Coruzzi, G. M., and Gutierrez, R. A. (2010) VirtualPlant: a software platform to support systems biology research. *Plant Physiol* **152**, 500-515.

138. Saeed, A. I., Hagabati, N. K., Braisted, J. C., Liang, W., Sharov, V., Howe, E. A., Li, J. W., Thiagarajan, M., White, J. A., and Quackenbush, J. (2006) TM4 microarray software suite. *Method Enzymol* **411**, 134-193.

139. Li, Z., Adams, R. M., Chourey, K., Hurst, G. B., Hettich, R. L., and Pan, C. Systematic comparison of label-free, metabolic labeling, and isobaric chemical labeling for quantitative proteomics on LTQ Orbitrap Velos. *J Proteome Res* **11**, 1582-1590.

140. Chourey, K., Thompson, M. R., Shah, M., Zhang, B., Verberkmoes, N. C., Thompson, D. K., and Hettich, R. L. (2009) Comparative temporal proteomics of a response regulator (SO2426)-deficient strain and wild-type Shewanella oneidensis MR-1 during chromate transformation. *J Proteome Res* **8**, 59-71.

141. Kiba, T., Naitou, T., Koizumi, N., Yamashino, T., Sakakibara, H., and Mizuno, T. (2005) Combinatorial microarray analysis revealing Arabidopsis genes implicated in cytokinin responses through the His -> Asp phosphorelay circuitry. *Plant Cell Physiol* **46**, 339-355.

142. Friso, G., Giacomelli, L., Ytterberg, A. J., Peltier, J. B., Rudella, A., Sun, Q., and van Wijk, K. J. (2004) In-depth analysis of the thylakoid membrane proteome of Arabidopsis thaliana chloroplasts: New proteins, new functions, and a plastid proteome database. *Plant Cell* **16**, 478-499.

143. Yates, J. R., Gilchrist, A., Howell, K. E., and Bergeron, J. J. M. (2005) Proteomics of organelles and large cellular structures. *Nat Rev Mol Cell Bio* **6**, 702-714.

144. Eubel, H., Heazlewood, J. L., and Millar, A. H. (2007) Isolation and subfractionation of plant mitochondria for proteomic analysis. *Methods Mol Biol* **355**, 49-62.

145. Poirier, Y., Ventre, G., and Caldelari, D. (1999) Increased flow of fatty acids toward beta-oxidation in developing seeds of Arabidopsis deficient in diacylglycerol acyltransferase activity or synthesizing medium-chain-length fatty acids. *Plant Physiol* **121**, 1359-1366.

146. Koch, K. (2004) Sucrose metabolism: regulatory mechanisms and pivotal roles in sugar sensing and plant development. *Curr Opin Plant Biol* **7**, 235-246.

147. Yuan, J. S., Tiller, K. H., Al-Ahmad, H., Stewart, N. R., and Stewart, C. N. (2008) Plants to power: bioenergy to fuel the future. *Trends in Plant Science* **13**, 421-429.

148. Hamblin, M. T., Buckler, E. S., and Jannink, J. L. (2011) Population genetics of genomics-based crop improvement methods. *Trends in Genetics* **27**, 98-106.

149. Xu, Y. L., Li, L., Wu, K., Peeters, A. J., Gage, D. A., and Zeevaart, J. A. (1995) The GA5 locus of Arabidopsis thaliana encodes a multifunctional gibberellin 20-

oxidase: molecular cloning and functional expression. *Proceedings of the National Academy of Sciences* **92**, 6640-6644.

150. Kebeish, R., Niessen, M., Thiruveedhi, K., Bari, R., Hirsch, H.-J., Rosenkranz, R., Stabler, N., Schonfeld, B., Kreuzaler, F., and Peterhansel, C. (2007) Chloroplastic photorespiratory bypass increases photosynthesis and biomass production in Arabidopsis thaliana. *Nat Biotech* **25**, 593-599.

151. Xing, Y. Z., and Zhang, Q. F. (2010) Genetic and molecular bases of rice yield. *Annual Review of Plant Biology, Vol 61*, pp. 421-442.

152. Tester, M., and Langridge, P. (2010) Breeding technologies to increase crop production in a changing world. *Science* **327**, 818-822.

153. Ashraf, M. (2010) Inducing drought tolerance in plants: Recent advances. *Biotechnology Advances* **28**, 169-183.

154. Stamm, P., Ramamoorthy, R., and Kumar, P. P. (2011) Feeding the extra billions: strategies to improve crops and enhance future food security. *Plant Biotechnology Reports* **5**, 107-120.

155. Ferrier, T., Matus, J. T., Jin, J., and Riechmann, J. L. (2011) Arabidopsis paves the way: genomic and network analyses in crops. *Current Opinion in Biotechnology* **22**, 260-270.

156. Chew, Y. H., and Halliday, K. J. (2011) A stress-free walk from Arabidopsis to crops. *Current Opinion in Biotechnology* **22**, 281-286.

157. Zhou, T. B., Qin, Y. H., Lei, F. Y., Huang, W. F., and Drummen, G. P. C. (2013) Prohibitin attenuates oxidative stress and extracellular matrix accumulation in renal interstitial fibrosis disease. *Plos One* **8**, e77187.

158. Zhang, Y., Wang, Y., Xiang, Y., Lee, W., and Zhang, Y. (2012) Prohibitins are involved in protease-activated receptor 1-mediated platelet aggregation. *J Thromb Haemost* **10**, 411-418.

159. Van Aken, O., Whelan, J., and Van Breusegem, F. (2010) Prohibitins: mitochondrial partners in development and stress response. *Trends Plant Sci* **15**, 275-282.

160. Osman, C., Wilmes, C., Tatsuta, T., and Langer, T. (2007) Prohibitins interact genetically with Atp23, a novel processing peptidase and chaperone for the F1FO-ATP synthase. *Mol Biol Cell* **18**, 627-635.

161. Tsutsumi, T., Matsuda, M., Aizaki, H., Moriya, K., Miyoshi, H., Fujie, H., Shintani, Y., Yotsuyanagi, H., Miyamura, T., Suzuki, T., and Koike, K. (2009) Proteomics analysis of mitochondrial proteins reveals overexpression of a mitochondrial protein chaperon, prohibitin, in cells expressing Hepatitis C Virus core protein. *Hepatology* **50**, 378-386.

162. Dang, S. S., Sun, M. Z., Yang, E., Xun, M., Ma, L., Jia, Z. S., Wang, W. J., and Jia, X. L. (2011) Prohibitin is overexpressed in Huh-7-HCV and Huh-7.5-HCV cells harboring in vitro transcribed full-length hepatitis C virus RNA. *Virol J* **8**:424.

163. Supale, S. M., Brun, T., Avril, I., Gjinovci, A., Merkwirth, C., Herrera, P. L., Langer, T., and Maechler, P. (2010) Deletion of the mitochondrial chaperone prohibitin-

2 in beta cells results in beta cell apoptosis and promotes diabetes in transgenic Bet-Phb2(-/-) mice. *Diabetologia* **53**, S75-S76.

164. Jo, Y., and DeBose-Boyd, R. A. (2010) Control of cholesterol synthesis through regulated ER-associated degradation of HMG CoA reductase. *Crit Rev Biochem Mol* **45**, 185-198.

165. Ferrer, I., Perez, E., Dalfo, E., and Barrachina, M. (2007) Abnormal levels of prohibitin and ATP synthase in the substantia nigira and frontal cortex in Parkinson's disease. *Neurosci Lett* **415**, 205-209.

166. Schechinger, W., Hojlund, K., Levin, K., Beck-Nielsen, H., and Klein, H. H. (2008) Potential role of prohibitin-1 in type 2-diabetes-associated insulin resistance. *Diabetes* **57**, A372-A372.

167. Supale, S., Thorel, F., Merkwirth, C., Gjinovci, A., Herrera, P. L., Scorrano, L., Meda, P., Langer, T., and Maechler, P. (2013) Loss of prohibitin induces mitochondrial damages altering beta-cell function and survival and is responsible for gradual diabetes development. *Diabetes* **62**, 3488-3499.

168. Chiu, C. F., Ho, M. Y., Peng, J. M., Hung, S. W., Lee, W. H., Liang, C. M., and Liang, S. M. (2013) Raf activation by Ras and promotion of cellular metastasis require phosphorylation of prohibitin in the raft domain of the plasma membrane. *Oncogene* **32**, 777-787.

169. Merkwirth, C., Dargazanli, S., Tatsuta, T., Geimer, S., Lower, B., Wunderlich, F. T., von Kleist-Retzow, J. C., Waisman, A., Westermann, B., and Langer, T. (2008)

Prohibitins control cell proliferation and apoptosis by regulating OPA1-dependent cristae morphogenesis in mitochondria. *Gene Dev* **22**, 476-488.

170. Tatsuta, T., Model, K., and Langer, T. (2005) Formation of membrane-bound ring complexes by prohibitins in mitochondria. *Mol Biol Cell* **16**, 248-259.

171. Artal-Sanz, M., and Tavernarakis, N. (2009) Prohibitin and mitochondrial biology. *Trends Endocrin Met* **20**, 394-401.

172. Piechota, J., Kolodziejczak, M., Juszczak, I., Sakamoto, W., and Janska, H. (2010) Identification and characterization of high molecular weight complexes formed by matrix AAA proteases and prohibitins in mitochondria of Arabidopsis thaliana. *J Biol Chem* **285**, 12512-12521.

173. Earley, K. W., Haag, J. R., Pontes, O., Opper, K., Juehne, T., Song, K. M., and Pikaard, C. S. (2006) Gateway-compatible vectors for plant functional genomics and proteomics. *Plant J* **45**, 616-629.

174. Zhang, X. R., Henriques, R., Lin, S. S., Niu, Q. W., and Chua, N. H. (2006) Agrobacterium-mediated transformation of Arabidopsis thaliana using the floral dip method. *Nat Protoc* **1**, 641-646.

175. Wu, F. H., Shen, S. C., Lee, L. Y., Lee, S. H., Chan, M. T., and Lin, C. S. (2009) Tape-Arabidopsis Sandwich - a simpler Arabidopsis protoplast isolation method. *Plant Methods* **5**: 16-26.

176. Qi, Y. P., and Katagiri, F. (2009) Purification of low-abundance Arabidopsis plasma-membrane protein complexes and identification of candidate components. *Plant J* **57**, 932-944.

177. Zhang, Y. X., Liu, S. M., Dai, S. S. Y., and Yuan, J. S. (2012) Integration of shot-gun proteomics and bioinformatics analysis to explore plant hormone responses. *Bmc Bioinformatics* **13**:S8.

178. Sun, T. P., Goodman, H. M., and Ausubel, F. M. (1992) Cloning the Arabidopsis Ga1 locus by genomic subtraction. *Plant Cell* **4**, 119-128.

179. Sun, T. P., and Kamiya, Y. (1994) The Arabidopsis Ga1 Locus Encodes the Cyclase Ent-Kaurene Synthetase-a of Gibberellin Biosynthesis. *Plant Cell* **6**, 1509-1518.

180. Loreti, E., Povero, G., Novi, G., Solfanelli, C., Alpi, A., and Perata, P. (2008) Gibberellins, jasmonate and abscisic acid modulate the sucrose-induced expression of anthocyanin biosynthetic genes in Arabidopsis. *New Phytol* **179**, 1004-1016.

181. Fridborg, I., Kuusk, S., Moritz, T., and Sundberg, E. (1999) The Arabidopsis dwarf mutant shi exhibits reduced gibberellin responses conferred by overexpression of a new putative zinc finger protein. *Plant Cell* **11**, 1019-1031.

182. Moore, C. S., Cook-Johnson, R. J., Rudhe, C., Whelan, J., Day, D. A., Wiskich, J. T., and Soole, K. L. (2003) Identification of AtNDI1, an internal non-phosphorylating NAD(P)H dehydrogenase in arabidopsis mitochondria. *Plant Physiol* **133**, 1968-1978.

183. Rasmusson, A. G., Soole, K. L., and Elthon, T. E. (2004) Alternative NAD(P)H dehydrogenases of plant mitochondria. *Annu Rev Plant Biol* **55**, 23-39.

184. Wallstrom, S. V., Florez-Sarasa, I., Araujo, W. L., Escobar, M. A., Geisler, D. A., Aidemark, M., Lager, I., Fernie, A. R., Ribas-Carbo, M., and Rasmusson, A. G. (2014) Suppression of NDA-Type alternative mitochondrial NAD(P)H dehydrogenases

in Arabidopsis thaliana modifies growth and metabolism, but not high light stimulation of mitochondrial electron transport. *Plant & cell physiology* 55, 881-896.

185. Browman, D. T., Hoegg, M. B., and Robbins, S. M. (2007) The SPFH domain-containing proteins: more than lipid raft markers. *Trends Cell Biol* **17**, 394-402.

186. Jupe, E. R., Liu, X. T., Kiehlbauch, J. L., McClung, J. K., and DellOrco, R. T. (1996) Prohibitin in breast cancer cell lines: Loss of antiproliferative activity is linked to 3' untranslated region mutations. *Cell Growth Differ* **7**, 871-878.

187. Mcclung, J. K., Danner, D. B., Stewart, D. A., Smith, J. R., Schneider, E. L., Lumpkin, C. K., Dellorco, R. T., and Nuell, M. J. (1989) Isolation of a cDNA that hybrid selects antiproliferative messenger-RNA from rat-liver. *Biochem Bioph Res Co* **164**, 1316-1322.

188. Fusaro, G., Dasgupta, P., Rastogi, S., Joshi, B., and Chellappan, S. (2003) Prohibitin induces the transcriptional activity of p53 and is exported from the nucleus upon apoptotic signaling. *J Biol Chem* **278**, 47853-47861.

189. Theiss, A. L., and Sitaraman, S. V. (2011) The role and therapeutic potential of prohibitin in disease. *Bba-Mol Cell Res* **1813**, 1137-1143.

APPENDIX

Table A- 1. Significantly up-regulated rice proteins after FAW herbivory as identified by PARC-based proteomics analysis.

Gene function	Gene ID	pValue	Ratio
Metabolic Enzymes			
Glycosyl hydrolase, putative	Os05g15770.1	0.007025	37.229
Sucrose synthase, putative	Os03g28330.1	0.023199	22.332
Uroporphyrinogen decarboxylase, putative	Os03g21900.1	0.016138	9.180
Sucrose synthase, putative	Os06g09450.1	0.022999	8.203
5-methyltetrahydropteroyltriglutamate-homocysteine methyltransferase, putative	Os12g42884.1	0.067437	7.470
5-methyltetrahydropteroyltriglutamate-homocysteine methyltransferase, putative	Os12g42876.1	0.074498	7.208
Dehydrogenase, putative	Os01g46610.1	0.013124	6.852
Aminopeptidase, putative	Os08g44860.2	0.022001	5.891
Thiamine biosynthesis protein thic, putative	Os03g47610.1	0.012162	5.769
Glutathione S-transferase, N-terminal domain containing protein	Os05g02530.1	0.034798	5.391
Bifunctional 3-phosphoadenosine 5-phosphosulfate synthetase, putative	Os03g53230.1	0.0112	5.345
Ribulose bisphosphate carboxylase small chain, chloroplast precursor, putative	Os12g17600.1	0.185299	5.299
Alpha-1,4-glucan-protein synthase, putative	Os03g40270.1	0.018062	4.765
UDP-glucose 6-dehydrogenase, putative	Os12g25690.1	0.010637	4.382
2,3-bisphosphoglycerate-independent phosphoglycerate mutase, putative	Os05g40420.1	0.016701	4.216
Glucose-6-phosphate isomerase, putative	Os06g14510.1	0.01995	3.942
Aminotransferase, classes I and II, domain containing protein	Os02g55420.1	0.0171	3.940
ATP-citrate synthase subunit 1, putative	Os01g19450.1	0.013487	3.932
Methylenetetrahydrofolate reductase, putative	Os03g60090.1	0.016138	3.924
Ospop7 - Putative Prolyl Oligopeptidase	Os03g19410.1	0.012725	3.812
Caffeoyl-coa O-methyltransferase, putative	Os08g38900.1	0.01425	3.710
Erythronate-4-phosphate dehydrogenase, putative	Os11g26850.2	0.042187	3.701
M16 domain containing zinc peptidase, putative	Os02g52390.1	0.026576	3.632
IN2-1 protein, putative	Os03g17470.1	0.012725	3.631
Glycine cleavage system H protein, putative	Os10g37180.1	0.126069	3.600
Enolase, putative	Os06g04510.1	0.017463	3.575
Catalase isozyme A, putative	Os02g02400.1	0.06976	3.470
Amine oxidase, flavin-containing, domain	Os03g08570.1	0.02545	3.391

containing protein			
Glutamine synthetase, catalytic domain containing protein	Os03g12290.1	0.029661	3.357
Lipoxygenase 2.1, chloroplast precursor, putative	Os12g37260.1	0.01995	3.356
Isoflavone reductase, putative	Os01g01660.1	0.015375	3.197
26S protease regulatory subunit 6B, putative	Os02g21970.1	0.015775	3.195
UDP-glucose 6-dehydrogenase, putative	Os03g55070.1	0.014813	3.063
NAD dependent epimerase/dehydratase family protein, putative	Os10g28200.1	0.026412	3.015
Glycosyl hydrolase, family 31, putative	Os03g11720.1	0.016337	2.959
Bifunctional 3-phosphoadenosine 5-phosphosulfate synthetase, putative	Os04g02050.1	0.039137	2.957
Soluble inorganic pyrophosphatase, putative	Os02g52940.2	0.049974	2.940
Triosephosphate isomerase, chloroplast precursor, putative	Os09g36450.1	0.051862	2.915
Aspartic proteinase oryzasin-1 precursor, putative	Os05g49200.1	0.029098	2.915
4-alpha-glucanotransferase, putative	Os07g46790.1	0.014414	2.817
4-nitrophenylphosphatase, putative	Os04g41340.1	0.027574	2.670
Enolase-phosphatase E1, putative	Os01g01120.1	0.018225	2.652
Aminotransferase, putative	Os08g39300.1	0.107772	2.621
Alpha-glucan phosphorylase isozyme, putative	Os01g63270.1	0.035525	2.557
Peptide-N4-asparagine amidase A, putative	Os01g10950.1	0.014449	2.542
Glyceraldehyde-3-phosphate dehydrogenase, putative	Os02g38920.1	0.095174	2.471
Aminotransferase, putative	Os08g41990.1	0.043839	2.421
Carbamoyl-phosphate synthase large chain, putative	Os01g38970.1	0.024324	2.383
Pyruvate kinase, putative	Os11g05110.1	0.021838	1.980
Glyceraldehyde-3-phosphate dehydrogenase, putative	Os08g03290.1	0.158437	1.793
Oligopeptidase, putative	Os02g58340.1	0.038175	1.616
Protein synthesis			
Chloroplast 30S ribosomal protein S8, putative	Os04g16832.1	0.010637	4.181
40S ribosomal protein S7, putative	Os03g18570.1	0.026213	4.084
Eukaryotic translation initiation factor 3 subunit K, putative	Os03g08450.1	0.009113	3.913
Ribosomal L9, putative	Os02g57670.1	0.020036	3.756
Ribosomal protein L6, putative	Os09g31180.1	0.03576	3.521
40S ribosomal protein S17, putative	Os10g27190.1	0.030623	3.370
Ribosomal protein, putative	Os02g06700.2	0.027377	3.344
40S ribosomal protein S7, putative	Os05g27940.1	0.016138	3.322

Elongation factor, putative	Os04g02820.1	0.121387	2.313
Elongation factor 1-gamma, putative	Os06g37440.1	0.063492	2.107
Chaperonin			
Chaperonin, putative	Os02g54060.1	0.038793	3.484
Chaperonin, putative	Os06g09679.1	0.021275	3.197
ATP-dependent Clp protease ATP-binding subunit, putative	Os03g14280.1	0.016501	2.940
Dnak family protein, putative	Os01g08560.1	0.0226	2.820
Heat shock protein, putative	Os08g39140.1	0.114326	1.949
Heat shock protein, putative	Os09g30418.1	0.134675	1.935
Heat shock protein, putative	Os09g30412.1	0.13315	1.917
Transcription factors			
Calreticulin precursor protein, putative	Os07g14270.1	0.034762	3.648
RNA recognition motif containing protein, putative	Os08g44290.1	0.038012	2.793
Defense proteins			
Cupin domain containing protein	Os03g48770.1	0.009312	4.846
Jacalin-like lectin domain containing protein, putative	Os12g14440.1	0.033274	4.559
Others			
PPR repeat domain containing protein, putative	Os04g46010.1	0.009875	5.287
Tubulin/ftsz domain containing protein, putative	Os07g38730.1	0.019151	4.954
Transposon protein, putative, unclassified	Os01g03070.1	0.009676	4.341
WD repeat-containing protein, putative	Os01g49290.1	0.036323	3.804
Expressed protein	Os07g31490.1	0.028663	3.580
Coatomer alpha subunit, putative	Os03g50340.1	0.013851	3.526
Chlorophyll A-B binding protein, putative	Os09g26810.1	0.016337	3.493
Ossub19 - Putative Subtilisin homologue	Os02g44520.1	0.011763	3.136
Dynamin, putative	Os06g13820.1	0.012326	3.093
TUDOR protein with multiple snc domains, putative	Os02g32350.1	0.016138	3.092
Chlorophyll A-B binding protein, putative	Os08g33820.1	0.209169	3.018
Chloroplast lumen common family protein, putative	Os05g08930.1	0.013288	2.776
IAP100, putative	Os10g35030.1	0.348678	2.742
Retrotransposon protein, putative, unclassified	Os04g52460.1	0.02545	2.715
Clathrin heavy chain, putative	Os11g01380.1	0.078965	2.288
Expressed protein	Os03g64020.1	0.0682	1.520

Table A- 2. Comprehensive comparison of differentially regulated genes using PARC-based protein expression and MPSS-based mRNA expression.

Gene ID	Protein LOG2 Fold	Protein p Value	BAW Treatment 1	BAW Treatment 2	Wounding Control 1	Wounding Control 2	mRNA p Value	mRNA LOG2 Fold	Correlation
LOC_Os07g22930	-6.0131	0.0474	159	203	250	283	0.0449	-0.5581	Y
LOC_Os10g29470	-5.2525	0.0036	189	172	361	296	0.0239	-0.8639	Y
LOC_Os02g13970	-5.0993	0.0450	232	221	555	479	0.0085	-1.1907	Y
LOC_Os06g12600	-5.0083	0.0144	0	0	0	0	N/A	N/A	
LOC_Os02g14440	-4.6533	0.0451	34	91	0	0	0.0798	N/A	
LOC_Os03g04220	-4.5127	0.0444	0	2	0	0	0.2113	N/A	
LOC_Os02g39920	-4.4903	0.0100	0	0	0	0	N/A	N/A	
LOC_Os09g04790	-4.4753	0.0089	163	206	243	311	0.0741	-0.5863	Y
LOC_Os04g16748	-4.3344	0.0488	0	0	0	0	N/A	N/A	
LOC_Os03g04240	-4.2855	0.0236	0	0	0	0	N/A	N/A	
LOC_Os06g45120	-4.0612	0.0410	120	79	69	81	0.1850	0.4078	
LOC_Os01g39270	-3.9435	0.0177	0	0	0	0	N/A	N/A	
LOC_Os01g66940	-3.9073	0.0167	86	95	49	62	0.0237	0.7054	
LOC_Os10g38700	-3.4396	0.0418	206	210	7	2	0.0001	5.5305	
LOC_Os07g43700	-3.3449	0.0159	0	0	0	0	N/A	N/A	
LOC_Os09g08910	-3.0719	0.0251	0	0	0	0	N/A	N/A	
LOC_Os02g08480	-3.0630	0.0158	0	0	0	0	N/A	N/A	
LOC_Os03g03910	-2.9547	0.0143	615	632	709	854	0.0814	-0.3259	
LOC_Os12g08730	-2.9143	0.0053	1439	1228	598	547	0.0099	1.2199	
LOC_Os01g13210	-2.8686	0.0281	520	448	389	291	0.0707	0.5095	
LOC_Os06g39040	-2.8328	0.0125	0	0	56	96	0.0314	N/A	

LOC_Os11g05110	-2.8215	0.0186	589	463	444	319	0.1225	0.4634	
LOC_Osp1g00650	-2.8121	0.0276	0	0	0	0	N/A	N/A	
LOC_Os05g33280	-2.7846	0.0436	292	274	493	494	0.0009	-0.8022	Y
LOC_Os02g47600	-2.4773	0.0206	99	58	0	0	0.0310	N/A	
LOC_Os04g51270	-2.3416	0.0033	0	0	0	0	N/A	N/A	
LOC_Os07g04840	-2.2154	0.0083	54502	55738	67929	64283	0.0147	-0.2622	Y
LOC_Os06g01390	-2.1728	0.0130	877	556	62	98	0.0294	3.1629	
LOC_Os11g47970	-2.0921	0.0131	163	236	118	142	0.1061	0.6179	
LOC_Os06g02144	-1.7493	0.0000	1573	1879	437	613	0.0105	1.7170	
LOC_Os10g18340	-1.6430	0.0026	15056	14069	17248	17189	0.0165	-0.2417	Y
LOC_Os01g01120	2.6524	0.0114	580	418	160	138	0.0252	1.7437	Y
LOC_Os01g01660	3.1975	0.0397	4	0	139	221	0.0246	-6.4919	
LOC_Os01g03070	4.3408	0.0060	0	2	0	0	0.2113	N/A	
LOC_Os01g08560	2.8203	0.0373	95	48	14	2	0.0601	3.1599	Y
LOC_Os01g10950	2.5419	0.0144	142	163	76	104	0.0351	0.7608	Y
LOC_Os01g19450	3.9320	0.0219	0	2	0	0	0.2113	N/A	
LOC_Os01g38970	2.3834	0.0098	34	58	0	0	0.0309	N/A	Y
LOC_Os01g46610	6.8520	0.0118	120	132	118	111	0.1198	0.1381	
LOC_Os01g49290	3.8042	0.0378	249	212	479	451	0.0048	-1.0125	
LOC_Os01g63270	2.5569	0.0032	77	81	0	2	0.0004	6.3038	Y
LOC_Os02g02400	3.4697	0.0461	163	261	0	2	0.0250	7.7279	Y
LOC_Os02g06700	3.3445	0.0310	60	58	49	81	0.3721	-0.1397	
LOC_Os02g21970	3.1949	0.0292	0	0	0	0	N/A	N/A	
LOC_Os02g32350	3.0925	0.0026	0	0	0	4	0.2113	N/A	
LOC_Os02g38920	2.4712	0.0155	86	60	0	6	0.0172	4.6049	Y
LOC_Os02g44520	3.1364	0.0169	0	2	62	113	0.0386	-6.4512	
LOC_Os02g52390	3.6319	0.0120	1474	1521	1298	1413	0.0748	0.1437	
LOC_Os02g52940	2.9399	0.0014	0	0	174	175	0.0000	N/A	

LOC_Os02g54060	3.4840	0.0147	0	0	0	0	N/A	N/A	
LOC_Os02g55420	3.9397	0.0358	0	0	0	0	N/A	N/A	
LOC_Os02g57670	3.7560	0.0116	4	3	0	0	0.0099	N/A	Y
LOC_Os02g58340	1.6158	0.0023	4030	3193	3513	2736	0.2418	0.2090	
LOC_Os03g08450	3.9130	0.0101	82	71	181	198	0.0040	-1.3087	
LOC_Os03g08570	3.3912	0.0037	60	77	42	98	0.4819	-0.0313	
LOC_Os03g11720	2.9590	0.0061	0	0	0	0	N/A	N/A	
LOC_Os03g12290	3.3566	0.0362	0	0	0	0	N/A	N/A	
LOC_Os03g14280	2.9396	0.0248	653	595	542	639	0.3067	0.0796	
LOC_Os03g17470	3.6305	0.0078	30	58	0	0	0.0440	N/A	Y
LOC_Os03g18570	4.0843	0.0062	82	57	243	187	0.0208	-1.6293	
LOC_Os03g19410	3.8118	0.0004	2088	2250	1236	1546	0.0235	0.6409	Y
LOC_Os03g21900	9.1799	0.0395	95	89	181	123	0.0879	-0.7244	
LOC_Os03g28330	22.3322	0.0002	0	0	0	4	0.2113	N/A	
LOC_Os03g40270	4.7645	0.0170	77	71	7	13	0.0022	2.8875	Y
LOC_Os03g47610	5.7692	0.0185	2547	2197	5680	5143	0.0055	-1.1899	
LOC_Os03g48770	4.8460	0.0242	N/A				N/A	N/A	
LOC_Os03g50340	3.5258	0.0160	0	4	0	0	0.2113	N/A	
LOC_Os03g53230	5.3452	0.0326	0	2	0	0	0.2113	N/A	
LOC_Os03g55070	3.0635	0.0388	163	190	0	2	0.0029	7.4635	Y
LOC_Os03g60090	3.9245	0.0488	211	305	97	94	0.0373	1.4338	Y
LOC_Os03g64020	1.5203	0.0015	2802	2159	1548	1088	0.0494	0.9123	Y
LOC_Os04g02050	2.9565	0.0251	69	60	0	0	0.0024	N/A	Y
LOC_Os04g02820	2.3127	0.0069	56	60	0	2	0.0008	5.8580	Y
LOC_Os04g16832	4.1806	0.0027	0	2	0	0	0.2113	N/A	
LOC_Os04g41340	2.6700	0.0190	301	260	410	345	0.0638	-0.4285	
LOC_Os04g46010	5.2873	0.0060	0	0	0	0	N/A	N/A	
LOC_Os04g52460	2.7149	0.0144	0	0	0	0	N/A	N/A	

LOC_Os05g02530	5.3909	0.0104	3860	3979	319	385	0.0002	3.4770	Y
LOC_Os05g08930	2.7765	0.0105	249	215	90	106	0.0095	1.2433	Y
LOC_Os05g15770	37.2287	0.0330	2080	2761	757	1467	0.0585	1.1221	Y
LOC_Os05g27940	3.3221	0.0057	116	139	153	102	0.5000	0.0000	
LOC_Os05g40420	4.2159	0.0321	39	57	56	100	0.1671	-0.7004	
LOC_Os05g49200	2.9150	0.0403	0	2	0	0	0.2113	N/A	
LOC_Os06g04510	3.5747	0.0229	838	2168	778	1980	0.4513	0.1242	
LOC_Os06g09450	8.2031	0.0048	146	165	0	0	0.0019	N/A	Y
LOC_Os06g09679	3.1974	0.0030	5728	5135	7680	7556	0.0093	-0.4881	
LOC_Os06g13820	3.0934	0.0275	1044	967	986	1007	0.4213	0.0130	
LOC_Os06g14510	3.9420	0.0470	0	0	0	0	N/A	N/A	
LOC_Os06g37440	2.1067	0.0339	52	67	146	104	0.0495	-1.0710	
LOC_Os07g14270	3.6485	0.0429	730	647	132	115	0.0028	2.4789	Y
LOC_Os07g31490	3.5798	0.0043	0	2	0	0	0.2113	N/A	
LOC_Os07g38730	4.9536	0.0114	0	0	0	0	N/A	N/A	
LOC_Os07g46790	2.8168	0.0318	95	38	555	317	0.0472	-2.7129	
LOC_Os08g03290	1.7932	0.0037	812	984	674	701	0.0684	0.3854	Y
LOC_Os08g33820	3.0178	0.0407	14188	15786	12478	15333	0.2883	0.1081	
LOC_Os08g38900	3.7102	0.0117	223	222	0	2	0.0000	7.7977	Y
LOC_Os08g39140	1.9492	0.0033	52	55	132	123	0.0020	-1.2529	
LOC_Os08g39300	2.6209	0.0165	7683	6544	6610	5872	0.1636	0.1888	
LOC_Os08g41990	2.4214	0.0178	0	0	0	0	N/A	N/A	
LOC_Os08g44290	2.7927	0.0048	0	0	0	0	N/A	N/A	
LOC_Os08g44860	5.8914	0.0145	172	261	312	398	0.0773	-0.7135	
LOC_Os09g26810	3.4931	0.0283	116	167	0	0	0.0155	N/A	Y
LOC_Os09g30412	1.9167	0.0046	56	57	0	2	0.0002	5.8202	Y
LOC_Os09g30418	1.9354	0.0024	52	55	132	123	0.0020	-1.2529	
LOC_Os09g31180	3.5213	0.0387	129	163	222	219	0.0243	-0.5948	

LOC_Os09g36450	2.9153	0.0294	343	282	208	143	0.0458	0.8324	Y
LOC_Os10g27190	3.3697	0.0188	86	48	0	6	0.0398	4.4811	Y
LOC_Os10g28200	3.0152	0.0092	0	2	257	166	0.0219	-7.7245	
LOC_Os10g35030	2.7425	0.0395	288	395	973	988	0.0035	-1.5216	
LOC_Os10g37180	3.5999	0.0214	60	41	278	309	0.0028	-2.5390	
LOC_Os11g01380	2.2881	0.0148	120	157	0	0	0.0087	N/A	Y
LOC_Os11g05110	1.9798	0.0060	589	463	444	319	0.1225	0.4634	
LOC_Os11g26850	3.7011	0.0159	4	0	0	2	0.3492	1.0000	
LOC_Os12g14440	4.5590	0.0059	9402	10102	583	656	0.0007	3.9765	Y
LOC_Os12g17600	5.2994	0.0067	54641	44138	48084	38720	0.2422	0.1864	
LOC_Os12g25690	4.3820	0.0361	129	151	0	0	0.0031	N/A	Y
LOC_Os12g37260	3.3561	0.0005	430	427	0	4	0.0000	7.7432	Y
LOC_Os12g42876	7.2076	0.0088	559	446	326	294	0.0409	0.6969	Y
LOC_Os12g42884	7.4701	0.0114	580	456	493	403	0.2287	0.2095	

Table A- 3. Common contaminant proteins from Streptavidin pulldown assays

Locus	Times identified	Description
AT3G56130.4	2	Biotin/lipoyl attachment domain-containing protein
AT1G36160.1	2	Acetyl-coa carboxylase 1
AT1G36160.2	2	Acetyl-coa carboxylase 1
AT1G03090.2	2	Methylcrotonyl-coa carboxylase alpha chain, mitochondrial / 3-methylcrotonyl-coa carboxylase 1 (MCCA)
ATCG00490.1	2	Ribulose-bisphosphate carboxylases
AT1G42970.1	2	Glyceraldehyde-3-phosphate dehydrogenase B subunit
AT3G04120.1	2	Glyceraldehyde-3-phosphate dehydrogenase C subunit 1
AT1G52670.1	2	Single hybrid motif superfamily protein
AT1G12900.1	2	Glyceraldehyde 3-phosphate dehydrogenase A subunit 2
AT1G12900.4	2	Glyceraldehyde 3-phosphate dehydrogenase A subunit 2
AT1G12900.3	2	Glyceraldehyde 3-phosphate dehydrogenase A subunit 2
AT1G13440.1	2	Glyceraldehyde-3-phosphate dehydrogenase C2
AT5G38410.2	2	Ribulose bisphosphate carboxylase (small chain) family protein
AT3G26650.1	2	Glyceraldehyde 3-phosphate dehydrogenase A subunit
AT5G35360.1	2	Acetyl Co-enzyme a carboxylase biotin carboxylase subunit
AT5G38430.1	2	Ribulose bisphosphate carboxylase (small chain) family protein
AT1G67090.1	2	Ribulose bisphosphate carboxylase small chain 1A
AT5G16390.1	2	Chloroplastic acetylcoenzyme A carboxylase 1
AT5G16390.2	2	Chloroplastic acetylcoenzyme A carboxylase 1
AT3G15690.2	2	Single hybrid motif superfamily protein
AT2G39730.1	2	Rubisco activase
AT2G39730.3	2	Rubisco activase
AT2G39730.2	2	Rubisco activase
ATCG00120.1	2	ATP synthase subunit alpha
AT4G20360.1	2	RAB gtpase homolog E1B
AT3G14420.1	2	Aldolase-type TIM barrel family protein
AT3G14420.6	2	Aldolase-type TIM barrel family protein
AT3G14420.5	2	Aldolase-type TIM barrel family protein
AT3G14420.4	2	Aldolase-type TIM barrel family protein
AT3G14420.2	2	Aldolase-type TIM barrel family protein
AT1G33120.1	2	Ribosomal protein L6 family
AT1G33140.1	2	Ribosomal protein L6 family
AT5G15530.1	2	Biotin carboxyl carrier protein 2
AT1G05190.1	2	Ribosomal protein L6 family
AT1G04820.1	2	Tubulin alpha-4 chain
AT1G50010.1	2	Tubulin alpha-2 chain

AT5G17920.1	2	Cobalamin-independent synthase family protein
AT3G45140.1	2	Lipoxygenase 2
AT4G04640.1	2	Atpase, F1 complex, gamma subunit protein
AT1G32060.1	2	Phosphoribulokinase
AT3G46970.1	2	Alpha-glucan phosphorylase 2
AT1G13440.2	1	Glyceraldehyde-3-phosphate dehydrogenase C2
AT4G34030.1	1	3-methylcrotonyl-coa carboxylase
AT1G10630.1	1	ADP-ribosylation factor A1F
AT5G14670.1	1	ADP-ribosylation factor A1B
AT3G62290.3	1	ADP-ribosylation factor A1E
AT3G62290.2	1	ADP-ribosylation factor A1E
AT3G62290.1	1	ADP-ribosylation factor A1E
AT2G47170.1	1	Ras-related small GTP-binding family protein
AT1G70490.3	1	Ras-related small GTP-binding family protein
AT1G70490.2	1	Ras-related small GTP-binding family protein
AT1G70490.1	1	Ras-related small GTP-binding family protein
AT1G23490.1	1	ADP-ribosylation factor 1
AT3G14415.1	1	Aldolase-type TIM barrel family protein
AT3G14415.3	1	Aldolase-type TIM barrel family protein
AT3G14415.2	1	Aldolase-type TIM barrel family protein
AT3G14420.3	1	Aldolase-type TIM barrel family protein
AT1G49240.1	1	Actin 8
AT3G18780.2	1	Actin 2
AT3G18780.1	1	Actin 2
AT5G09810.1	1	Actin 7
AT3G23810.1	1	S-adenosyl-l-homocysteine (SAH) hydrolase 2
AT4G13940.1	1	S-adenosyl-L-homocysteine hydrolase
AT1G07920.1	1	GTP binding Elongation factor Tu family protein
AT5G60390.3	1	GTP binding Elongation factor Tu family protein
AT5G60390.1	1	GTP binding Elongation factor Tu family protein
AT1G07940.2	1	GTP binding Elongation factor Tu family protein
AT1G07940.1	1	GTP binding Elongation factor Tu family protein
AT1G07930.1	1	GTP binding Elongation factor Tu family protein
AT5G42650.1	1	Allene oxide synthase
ATCG00480.1	1	ATP synthase subunit beta
AT3G58610.1	1	Ketol-acid reductoisomerase
AT3G58610.3	1	Ketol-acid reductoisomerase
AT3G58610.2	1	Ketol-acid reductoisomerase
AT4G14960.2	1	Tubulin/ftsz family protein
AT4G14960.1	1	Tubulin/ftsz family protein
AT5G17920.2	1	Cobalamin-independent synthase family protein
AT5G64040.1	1	Photosystem I reaction center subunit PSI-N, chloroplast, putative / PSI-N, putative (PSAN)
AT5G64040.2	1	Photosystem I reaction center subunit PSI-N,

		chloroplast, putative / PSI-N, putative (PSAN)
AtMg01190	1	Atp1 atpase subunit 1
AT5G35630.1	1	Glutamine synthetase 2
AT5G35630.3	1	Glutamine synthetase 2
AT5G35630.2	1	Glutamine synthetase 2
AT2G42100.1	1	Actin-like atpase superfamily protein
AT2G33800.1	1	Ribosomal protein S5 family protein
AT1G23190.1	1	Phosphoglucomutase/phosphomannomutase family protein
AT5G19770.1	1	Tubulin alpha-3
AT5G19780.1	1	Tubulin alpha-5
AT5G49910.1	1	Chloroplast heat shock protein 70-2
AT1G29900.1	1	Carbamoyl phosphate synthetase B
AT1G23310.1	1	Glutamate:glyoxylate aminotransferase
AT1G23310.2	1	Glutamate:glyoxylate aminotransferase
AT5G50920.1	1	CLPC homologue 1
AT1G03090.1	1	Methylcrotonyl-coa carboxylase alpha chain, mitochondrial / 3-methylcrotonyl-coa carboxylase 1 (MCCA)
AT5G38410.1	1	Ribulose bisphosphate carboxylase (small chain) family protein
AT5G38420.1	1	Ribulose bisphosphate carboxylase (small chain) family protein
AT5G38410.3	1	Ribulose bisphosphate carboxylase (small chain) family protein
AT3G01500.1	1	Carbonic anhydrase 1
AT3G01500.3	1	Carbonic anhydrase 1
AT3G01500.2	1	Carbonic anhydrase 1
AT1G58380.1	1	Ribosomal protein S5 family protein
AT1G59359.1	1	Ribosomal protein S5 family protein
AT1G58983.1	1	Ribosomal protein S5 family protein
AT1G58684.1	1	Ribosomal protein S5 family protein