

8-6-2011

## Epigenetic regulations in cell wall degradation and regeneration in *Oryza sativa*

Feng Tan

Follow this and additional works at: <https://scholarsjunction.msstate.edu/td>

---

### Recommended Citation

Tan, Feng, "Epigenetic regulations in cell wall degradation and regeneration in *Oryza sativa*" (2011).  
*Theses and Dissertations*. 1908.  
<https://scholarsjunction.msstate.edu/td/1908>

This Dissertation - Open Access is brought to you for free and open access by the Theses and Dissertations at Scholars Junction. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of Scholars Junction. For more information, please contact [scholcomm@msstate.libanswers.com](mailto:scholcomm@msstate.libanswers.com).

EPIGENETIC REGULATIONS IN CELL WALL DEGRADATION AND  
REGENERATION IN *ORYZA SATIVA*

By

Feng Tan

A Thesis  
Submitted to the Faculty of  
Mississippi State University  
in Partial Fulfillment of the Requirements  
for the Degree of Degree of Philosophy  
in Molecular Biology  
in the Department of Biochemistry and Molecular Biology

Mississippi State, Mississippi

August 2011

Copyright 2011

By

Feng Tan

EPIGENETIC REGULATIONS IN CELL WALL DEGRADATION AND  
REGENERATION IN *ORYZA SATIVA*

By

Feng Tan

Approved:

---

Zhaohua Peng  
Associate Professor of Biochemistry and  
Molecular Biology  
(Director of Dissertation)

---

Shane Burgess  
Professor of Basic Sciences  
Associate Dean for Strategic Initiatives  
& Economic Development, CVM  
(Member of Committee)

---

Daniel Peterson  
Associate Professor of Plant and Soil  
Sciences  
(Member of Committee)

---

Susan Bridges  
Professor Emerita of Computer Science  
and Engineering  
(Member of Committee)

---

Jiaxu Li  
Associate Professor of Biochemistry  
and Molecular Biology  
(Member of Committee)

---

Din Pow Ma  
Professor of Biochemistry and  
Molecular Biology  
(Director of Graduate Studies)

---

George Hopper  
Interim Dean of the College of Agriculture  
and Life Sciences

Name: Feng Tan

Date of Degree: August 6, 2011

Institution: Mississippi State University

Major Field: Molecular Biology

Major Professor: Zhaohua Peng

Title of Study: EPIGENETIC REGULATIONS IN CELL WALL DEGRADATION  
AND REGENERATION IN *ORYZA SATIVA*

Pages in Study: 186

Candidate for Degree of Degree of Philosophy

It is well known that chromatin components are key players in establishing and maintaining spatial and temporal gene expression in plants, however, little is known about the epigenetic regulation on cell wall degradation and regeneration. This study aimed to 1) investigate the global proteome and phosphoproteome of rice chromatin, and 2) characterize changes in chromatin components and chromatin structure associated with cell wall degradation and regeneration, and 3) characterize the differentially regulated proteins and eventually explore the mechanism.

In this dissertation, we examine proteins copurified with chromatin using both 2-DE gel and shotgun approaches from rice (*Oryza sativa*) suspension cells. Nine hundred seventy-two distinct protein spots were resolved on 2-DE gels and 509 proteins were identified by MALDI-MS/MS following gel excision, these correspond to 269 unique proteins. When the chromatin copurified proteins are examined using shotgun proteomics, a large number of histone variants in addition to the four common core histones were identified. Furthermore, putative phosphoproteins copurified with chromatin were examined using Pro-Q Diamond phosphoprotein stain and followed by

MALDI-MS/MS. Our studies provided new insights into the chromatin composition in plants.

To study the epigenetic regulation of the cell wall degradation and regeneration, we examined cellular responses to the enzymatic removal of the cell wall in rice suspension cells using proteomic approaches. We found that removal of cell wall stimulates cell wall synthesis from multiple sites in protoplasts instead of from a single site as in cytokinesis. Microscopy examination and chromatin decondensation assay further showed that removal of the cell wall is concomitant with substantial chromatin reorganization. Histone post-translational modification studies using both Western blots and isotope labeling assisted quantitative mass spectrometry analyses revealed substantial histone modification changes, particularly H3K18<sub>AC</sub> and H3K23<sub>AC</sub>, are associated with the degradation and regeneration of the cell wall. Label-free comparative proteome analyses further revealed that chromatin associated proteins undergo dramatic changes upon removal of the cell wall, particularly cytoskeleton, cell wall metabolism, and stress-response proteins. This study demonstrates that cell wall removal is associated with substantial chromatin change and may lead to stimulation of cell wall synthesis using a novel mechanism.

## DEDICATION

To my lovely daughter

Xiaoxiao Lillian Tan

## ACKNOWLEDGEMENTS

. The first person I would like to express my deep and sincere gratitude to is Dr. Zhaohua Peng, my major advisor, who has given me countless suggestions and insightful guidance and during the course of this work. His wide knowledge and fruitful discussions have been of great value to me. I thank him from the depth of my heart for his important inspiration and encouragement for this research. I am extremely grateful to my committee members, Dr. Shane Burgess, Dr. Jiaxu Li, Dr. Susan Bridges, and Dr. Daniel Peterson, for their sustained encouragement, guidance and suggestions throughout the period of this research and preparation of thesis. I also express my gratitude to Dr. Din-Pow Ma for serving as my graduate coordinator and helping me to register for courses and for his kind help throughout my graduate study. I would also like to extend my gratitude to Dr. Scott Willard for his help during the course of my Ph.D.

I want to thank Dr. Tibor Pechan and Ken Pendarvis at Life Sciences and Biotechnology Institute (LSBI) for helping me with the mass spectrometry analyses of my samples. I am thankful to Bill Monroe at Electron Microscope Center for imparting electron and confocal microscopy training and providing technical help. I would also take this opportunity to extend my thanks to Dr. Susan Bridges' students Bryce Magee and Brandon Malone for their collaboration in analyzing the label-free proteomics, microarray and sequencing data. I am also very grateful to Dr. Kanglin Zhang of Loma Linda University for helping me conduct the quantitative histone acetylation study which was an important part of this research.



I would also like to take this opportunity to thank BR Chitteti, who taught me to make cell culture, run gels, analyze my proteomic data. My thanks to my previous and current lab mates Guosheng Li, Aparna, Ramesh, Hana, Hui Wang, Jian Zhang, Guohua Yang for their valuable discussions and our friendship. It was my great pleasure to work with them.

Last, but not least, I would like to thank my family. My parents constantly encourage and inspire me. Without their endless support, I would never achieve my current progress. I am especially grateful to my wife, Lequn Hu, for her encouragement and support, for keeping my life in proper perspective and balance.

## TABLE OF CONTENTS

	Page
DEDICATION .....	ii
ACKNOWLEDGEMENTS .....	iii
LIST OF TABLES .....	viii
LIST OF FIGURES .....	ix
 CHAPTER	
I. INTRODUCTION .....	1
II. LITERATURE REVIEW .....	5
Rice ( <i>Oryza sativa</i> ) .....	5
Plant cell wall.....	6
Cell wall biosynthesis .....	7
Plant protoplasts.....	8
Chromatin .....	9
Histones and histone post-translational modifications .....	11
Chromatin reorganization .....	13
Epigenetics and cell wall regeneration .....	15
Proteomics.....	16
Two-dimensional gel electrophoresis coupled with MALDI-TOF-MS.....	19
2D-LC coupled electrospray ionization mass spectrometry.....	21
Comparative proteomics .....	24
Systems biology .....	25
III. PROTEOME PROFILING OF RICE CHROMATIN.....	27
Abstract.....	27
Introduction.....	28
Materials and methods .....	29
Suspension culture and protoplast isolation.....	29
Nucleus isolation and chromatin isolation.....	29
Electron microscopy .....	30
Protein extraction .....	31

Western blots .....	31
Two-dimensional polyacrylamide gel electrophoresis (2D PAGE) .....	32
In gel digestion and mass spectrometry .....	33
Sample preparation for MudPIT and 2D-LC-MS/MS analysis .....	34
Gene ontology (GO) annotation.....	36
Results.....	37
Chromatin purification from rice ( <i>Oryza sativa</i> ) suspension cells .....	37
Western blots analysis.....	38
2-DE gel, and mass spectrometric analysis of proteins co-purified with chromatin .....	40
Identification of chromatin co-purified proteins using shotgun approach.....	45
Gene ontology analysis .....	58
Discussions .....	61
Chromatin purification.....	61
Protein identification and agriGO annotations .....	62
Proteins co-purified with chromatin .....	68
The efficiency of chromatin protein identification using shotgun approach.....	71
IV. PHOSPHOPROTEOME PROFILING OF RICE CHROMATIN .....	72
Abstract.....	72
Introduction.....	72
Materials and methods .....	74
Identification of phosphoproteins using Pro-Q Diamond gel stain .....	74
Investigating GO annotations of putative phosphoproteins.....	74
Results.....	75
Mapping putative phosphoproteome co-purified with chromatin using Pro-Q Diamond phosphoprotein stain.....	75
GO annotations of putative phosphoproteins.....	76
Discussion.....	83
Putative phosphoproteins co-purified with chromatin revealed by Pro-Q Diamond stain .....	83
V. CHROMATIN REORGANIZATION AND HISTONE POST- TRANSLATIONAL MODIFICATIONS DURING CELL WALL DEGRADATION AND REGENERATION IN <i>ORYZA SATIVA</i> .....	86
Abstract.....	86
Introduction.....	87
Materials and methods .....	89
Cell culture.....	89
Protoplast isolation and culture.....	89
Evaluation of new cell wall formation.....	90

Nuclei isolation and assessment of the integrity of nuclear fractions.....	90
Chromatin decondensation assay .....	90
Immunoblot analysis .....	91
Histone H3 isolation .....	91
Histone post-translational modifications identification and acetylation quantification by mass spectrometry .....	92
Results.....	93
Cell wall regeneration in rice ( <i>Oryza sativa</i> ) protoplasts .....	93
Removal of cell wall stimulates chromatin reorganization.....	94
Characterization of histone H3 post-translational modifications.....	98
Histone H3 modification changes associated with cell wall removal.....	101
Discussion .....	107
Removal of cell wall and chromatin state.....	107
Histone modification and differential histone modification in rice .....	109
The differences between cell wall synthesis in protoplasts and cell wall synthesis during cytokinesis.....	111
 VI.     PROTEOME DYNAMIC CHANGE DURING CELL WALL DEGRADATION AND REGENERATION IN <i>ORYZA</i> <i>SATIVA</i> .....	 112
Abstract.....	112
Introduction.....	112
Materials and methods .....	115
Protein extraction .....	115
Shotgun proteomic analysis .....	115
Protein comparative quantification using the spectral count and $\Sigma$ XCorr methods .....	116
System modeling: <i>Oryza sativa</i> protein interaction network.....	117
Results.....	118
Protoplasts proteome revealed by shotgun proteomics.....	118
Differential expression of proteins in response to cell wall removal revealed by label-free quantification methods.....	119
Systems approach revealed key regulatory nodes of the protein interaction network .....	153
Discussion .....	158
Differentially expressed proteins in response to the removal of the cell wall.....	158
 VII.     SUMMARY .....	 161
REFERENCES .....	168

## LIST OF TABLES

TABLE	Page
1 Proteins co-purified with chromatin in rice ( <i>Oryza sativa</i> ) using MALDI-TOF .....	48
2 Proteins co-purified with chromatin identified using shotgun approach .....	53
3 Proteins identified by searching the reversed sequence database .....	58
4 Unique putative phosphoprotein identified by Pro-Q Diamond and MALDI-TOF .....	82
5 LC-MSMS analysis of histone H3 .....	99
6 Numbers of differentially expressed proteins revealed by label-free quantification.....	122
7 Differentially Up regulated proteins identified using label-free quantification methods .....	125
8 Differentially Down regulated proteins identified using label-free quantification methods .....	134
9 Identified differentially regulated proteins by label-free methods.....	146

## LIST OF FIGURES

FIGURE	Page
1     Microscopy images of isolated rice ( <i>Oryza sativa</i> ) suspension cell protoplasts, nuclei and chromatin.....	38
2     Immunological characterization of proteins in the chromatin fraction, nuclear fraction, and total protein fraction .....	40
3     2-DE gel image of rice chromatin proteome revealed by SYPRO Ruby fluorescence stain .....	43
4     2-DE gel images of three biological replicas of rice chromatin proteome revealed by SYPRO Ruby fluorescence stain .....	44
5     Histone H3 acetylation detected by LC-MS/MS.....	47
6     Functional classification of chromatin proteome .....	60
7     Venn diagram representing the proteins distribution identified using 2- DE and MudPIT .....	64
8     Hierarchical tree graph of total unique proteins in biological process category generated by SEA of agriGO.....	65
9     Hierarchical tree graph of total unique proteins in cellular component category generated by SEA of agriGO.....	66
10    Hierarchical tree graph of total unique proteins in molecular function category generated by SEA of agriGO.....	67
11    2-DE gel images of the putative phosphoproteins associated with rice chromatin.....	78
12    2-DE gel images of the putative phosphoproteins co-purified with rice chromatin.....	79
13    Hierarchical tree graph of overrepresented GO terms in biological process category generated by SEA. ....	80

14	Hierarchical tree graph of overrepresented GO terms in cellular component category generated by SEA. ....	81
15	Fluorescence microscopy images of cultured rice ( <i>Oryza sativa</i> ) protoplasts .....	96
16	Protoplast chromatin decondensation revealed by DAPI stain and MNase assay .....	97
17	Representative MS/MS mass spectra of H3 and H3.2 modifications .....	100
18	Immunological characterization of histone H3 modification in suspension cells, protoplasts, and protoplast-derived cells (48 h) .....	105
19	Quantitative comparisons of acetylations at residues K18 and K23 of histone H3 prepared from suspension cells and fresh protoplasts .....	106
20	Protein distribution among different Gene Ontology categories .....	123
21	Percentage of differentially expressed proteins in gene ontology categories.....	124
22	Interaction network of <i>Oryza Sativa</i> protoplast proteins with significantly altered expression response to absence of cell wall .....	157
23	<i>Oryza Sativa</i> cellular activities dynamic regulations respond to cell wall removing.....	158

## CHAPTER I

### INTRODUCTION

Epigenetics refers to heritable patterns of gene expression which do not depend on alterations of genomic DNA sequence. The two major epigenetic mechanisms are histone modifications and DNA methylation. Dynamic changes in chromatin structure are directly influenced by post-translational modifications of the amino terminal tails of core histone proteins, H2A, H2B, H3, and H4, which build up the basic structural unit of chromatin, the nucleosome. These modifications (e.g., methylation, acetylation, and phosphorylation) recruit other proteins to form complexes that affect chromatin structure and function (Jenuwein and Allis, 2001). Histone modifications are proposed to affect chromosome function through at least two distinct mechanisms. The first mechanism suggests modifications may alter the electrostatic charge of the histone resulting in a structural change in histones or their binding to DNA. The second mechanism proposes that these modifications are binding sites for protein recognition modules, such as the bromodomains or chromodomains used to recognize acetylated lysines or methylated lysine, respectively (Jenuwein and Allis, 2001; Kimmins and Sassone-Corsi, 2005). Lysine acetylation, lysine and arginine mono-, di- and tri-methylation, serine phosphorylation, lysine ubiquitinylation and glutamate poly-ADP ribosylation are the primary reversible histone post-translation modifications (Grunstein, 1997; Struhl, 1998; Ng and Bird, 2000; Strahl and Allis, 2000; Richards and Elgin, 2002; Sun and Allis, 2002; Iizuka and Smith, 2003). The combinatorial nature of histone amino-terminal



modifications reveals a so called "histone code" that considerably extends the information potential of the genetic code (Jenuwein and Allis, 2001). For example, histone acetylation and methylation often have distinct influences on gene expression. The DNA of active genes is preferentially associated with highly acetylated histones while DNA of inactive genes is associated with hypo acetylated histones (Wolffe and Pruss, 1996). In addition, the lysine mono-, di-, and trimethylation states also have a distinct impact on gene expression (Barski et al., 2007). The proteins involved in histone modifications include histone acetyltransferases, histone deacetylases, histone methyltransferases (SET domain proteins), histone demethylases, chromodomain proteins (methylated histone binding), etc. However, we do not know exactly how the genome reprogramming takes place although several remodeling factors and complexes involving these processes have been identified (Cairns, 2005; Hsieh and Gage, 2005; Bultman et al., 2006).

DNA methylation is thought to be a powerful epigenetic mechanism which interferes with the binding of transcription factor proteins and establishes a silent chromatin state. Heterochromatin is characterized by methylation of cytosine nucleotides of the DNA, the methylation of histone H3 at lysine 9, and the specific binding of heterochromatin protein 1 (HP1) to methylated H3 lysine 9. It has been shown that deacetylation of lysine 9 at the amino terminus of H3 is a prerequisite for methylation of this same lysine. Methylation of H3 lysine 9 in turn recruits the binding of HP1 that helps to establish highly compacted and transcriptionally inactive heterochromatin (Rice and Allis, 2001). There is increasing evidence that chromatin components are the key players in establishing and maintaining the spatial and temporal gene expression profile in both plants and animals (Meyerowitz, 2002).

Plant cells can rapidly re-synthesize cell wall after wall is removed (Mishra and Colvin, 1969). The regeneration of wall structure around protoplasts devoid of cell wall has been observed for mosses, tomato, *Haplopappus* suspension culture, *Arabidopsis*, tobacco mesophyll cells, soybean cell suspensions, onion root cells, maize endosperm cells and rice suspension cells (Horine and Ruesink, 1972). Plant protoplasts typically synthesized cell walls containing low amounts of cellulose regardless of the source or developmental stage of the original tissue (Jean H. Gould, 1986). A question raised by these and similar studies involving cell wall regeneration by protoplasts is why the cell walls regenerated by protoplasts are different in composition from the cell walls of the original tissue? Tobacco protoplasts isolated from leaf mesophyll cells containing secondary cell wall, with only 5% cellulose (Blaschek W, 1981), suggests that changes in metabolism occur in protoplasts either during isolation or during culture. On the other hand, there is a possibility that it is a wound response since Blaschek et al. found that the composition of regenerated wall is usually rich with synthesized  $\beta$ -1,3-linked glucans (Blaschek W, 1981). These glucans have been reported as a response to wounding for suspension cultured soybean cells (Brett, 1978).

The plant protoplast system is an excellent example displaying the cell wall degradation and regeneration capability. Although the cell wall re-synthesis activity in protoplasts has been discovered for a long time, the underlying molecular mechanisms are poorly explored. It is very interesting to study if cell wall re-synthesis in protoplasts and new cell wall synthesis during cytokinesis follow the same regulatory mechanism and use the same set of catalytic enzymes. It is also very interesting to investigate whether epigenetic regulations are involved in cell wall degradation and regeneration.

Actin filaments are possibly involved in guiding the phragmoplast to the site of the former preprophase band location at the parent cell. In addition, it has been reported that cellulose synthesis is guided by microtubules underneath the plasma membrane and cell wall and plasma membrane are directly connected via hechtian strands, containing both actins and microtubules (Hecht, 1912; Volgger et al., 2009). On the other hand, a large number of cytoskeleton proteins have been shown to co-purify with rice, human and *C. elegans* chromatin or chromosomes (Uchiyama et al., 2005; Chu et al., 2006; Tan et al., 2007).  $\beta$ -actin has been shown to be located with the entire metaphase chromosome body and it has been reported that  $\beta$ -actin is a component of chromatin-remodeling complex (Zhao et al., 1998; Uchiyama et al., 2005). Despite of this apparent connection between cell wall, chromatin and cytoskeleton proteins, it is still not clear if there is any cross-talk between chromatin and cell wall in mature plant cells at the molecular level.

Proteomics is a major tool in identifying proteins that are present under particular biological conditions. It is most effective when comparing differentially expressed proteins between two samples. This identification will allow for the characterization of biological roles, clarification of biological mechanisms, and identification of biomarkers. In this study, various microscopy techniques and proteomic approaches were employed to investigate dynamic changes in chromatin during cell wall degradation and regeneration. This work provides critical insight into the molecular mechanism of epigenetic regulation and cell wall degradation and regeneration in plant.

## CHAPTER II

### LITERATURE REVIEW

#### **Rice (*Oryza sativa*)**

Rice (*Oryza sativa* L) is an important food crop, providing 23% of the total calories consumed worldwide (Brar and Khush, 2002; Childs, 2004). In the Asian countries, such as Korea, China, and Japan, rice accounts for ~ 60% of daily caloric intake. Although 90 % of world's rice is produced and consumed in Asia, rice is a minor crop in the United States. Nevertheless, the annual per capita consumption of rice in the United States has increased three-fold to 27 pounds (12.3 kg) in the last forty years (Brar and Khush, 2002; Childs, 2004). Cultivated rice (*Oryza sativa* L.) originated from at least two independent domestication events resulting in the *indica* and *japonica* ecotypes (Cheng et al., 2003). The major cultivars planted in US are long-grain types, which are considered tropical *japonicas* (DILDAY, 1990; Mackill, 1995).

Because of its importance as a staple cereal crop, rice was the first fully-sequenced model monocot plant for studying plant biology. Although the completion of the rice genome (*O. sativa Japonica* Group cultivar Nipponbare) sequencing was announced by the International Rice Genome Sequencing Project (IRGSP) in 2002, the assembly was not published until 2005 by the IRGSP (Matsumoto et al., 2005). *O. sativa* has a compact genome compared with the multi-gigabase genomes of maize, wheat and barley. This accomplishment provided the public with 12 fully annotated chromosomes, which are available at the TIGR Rice Genome Annotation browser

(Ouyang et al., 2007). Approximately 57,000 genes have been identified through the TIGR project and, of these, about 25,000 genes have been annotated (Ouyang et al., 2007). The rice genome sequence is serving as a major reference for all cereal genomes. In addition, *Oryza sativa* has the smallest cereal genome consisting of just around 500Mb across 12 chromosomes. It is renowned for being easy to genetically modify and is a model organism for cereal biology. All of these properties led us to choose rice as a material to conduct cell wall degradation and regeneration studies.

### **Plant cell wall**

Plant cells are encased in rigid complex walls consisting of polysaccharides, proteins, and/or lignin, and differing in composition and amount depending on cell types (Zhong and Ye, 2007). The plant cell wall is the major determinant of plant structure and is of fundamental importance in plant growth and development, resistance to pathogen invasion, and the properties of plant fibres and fuels (Farrokhi et al., 2006). Unlike animals in which specialized skeletal system provides physical support, the plant cell shape, and ultimately organ and whole plant morphology highly depend on the cumulative properties of the walls. It is well documented, that the wall exhibits a dynamic nature critical not only for cell division, enlargement and differentiation, but also for intercellular communication, responses to environmental cues including biotic and abiotic stresses, reversible or irreversible modifications (Carpita and McCann, 2000; Kaczkowski, 2003; Kwon et al., 2005; Minic and Jouanin, 2006). The composite wall structure also serves as the main source of cellulose, the most abundant and useful biopolymer on the earth. Cellulose provides major dietary fibers and raw materials that are used to manufacture textiles, lumber and paper (Cosgrove, 2005; Somerville, 2006a;

Zhong and Ye, 2007). Moreover, there is a growing recognition that lignocellulosic biofuel holds tremendous potential as a sustainable energy resource (Somerville, 2006b, 2007). Further improvement of cell wall characteristics may help to optimize lignocellulosic biomass for cost-efficient biofuel production (Farrokhi et al., 2006).

Although plant cell walls are made of common components, they differ in composition and amount depending on cell types and even in different microdomains of the wall of a given cell (Freshour et al., 1996; Zhong and Ye, 2007). In addition, the amount of composition of walls may change during cell growth and in response to biotic and abiotic stresses (Vorwerk et al., 2004; Derbyshire et al., 2007). Therefore, plants have evolved mechanisms to turn on different biosynthetic pathways, synthesize the right amount of wall component, and assemble them in the right place to make a wall to meet specific needs (Zhong and Ye, 2007).

### **Cell wall biosynthesis**

Primary cell wall synthesis is initiated during cytokinesis and further modified during cell expansion. The phragmoplast serves as a scaffold for cell plate assembly and subsequent formation of a new cell wall separating two daughter cells (Raven et al., 2005). The phragmoplast is initially barrel-shaped and derived from the mitotic spindle, which is directly connected with the chromosomes of the two daughter cells while nuclear envelopes reassemble around the nuclei of the daughter cells. The phragmoplast is a complex assembly of microtubules, microfilaments, and endoplasmic reticulum elements which assemble in two opposing sets perpendicular to the plane of the future cell plate during anaphase (Verma and Hong, 2001). The microtubules and actin filaments within the phragmoplast direct vesicles carrying cell wall materials to the

growing cell plate. The involvement of spindle fiber and phragmoplast in cell wall synthesis suggests that cell wall synthesis and chromosome/chromatin state are interconnected although there have been no specific studies to address the question.

The secondary cell wall is a structure found between the primary cell wall and the plasma membrane in specific plant cells. The cell starts producing the secondary cell wall only after the primary cell wall is complete and the cell has stopped growing. The secondary cell wall is composed largely of polysaccharides and lignin with lesser quantities of other compounds. The rice straw cell wall, for example, consists of cellulose (32-37%), hemicellulose (29-37%), lignin (5-15%), and other minor compounds including proteins, aromatic compounds, silicon, etc (Glissmann and Conrad, 2000).

### **Plant protoplasts**

Plant protoplasts have been in constant use for more than three decades, and have become one of the most versatile analytical tools in plant biology. They can be isolated enzymatically in quantities and fused with those of other species, they can regenerate a new cell wall and are capable of nuclear and cell division, and many eventually develop into mature plants (Marx, 1987). These studies on using protoplasts of higher plants may open up new possibilities in plant breeding and in the experimental biology of plants. Efficient plant protoplast culture systems facilitate many research objectives including isolation of valuable biochemical mutants, introduction of foreign DNA into host genomes, transient gene expression experiments, molecular cytogenetics with particular emphasis on *in situ* hybridization for gene localization on metaphase chromosomes, flow-cytometry experiments comprising cell-cycle analysis and chromosome sorting, cloning of large DNA inserts (Negrutiu et al., 1992). Protoplasts have been used to observe

cellular processes and activities such as cell wall synthesis, cell division, embryogenesis, differentiation during regeneration, photosynthesis activity, calcium signaling and regulation, and the modulation of ion channels by light, stress and hormone responses in various plant species (Sheen, 2001).

Viable protoplasts are potentially totipotent. Therefore, when given the correct chemical and physical stimuli, each protoplast is capable, theoretically, of regenerating a new wall and undergoing repeated mitotic division to produce daughter cells from which fertile plants can be regenerated via the tissue culture process (Davey et al., 2005).

Reports on regeneration of plants from rice protoplasts are accumulating, including *japonica* types (Taipei 309 and Nipponbare) and *indica* types (Chinsurah Boro II, IR54 and IR72) (Lee et al., 1989; Ghosh Biswas et al., 1994; Utomo et al., 1995). Overall, dedifferentiated plant protoplast cells may be similar to pluripotent animal stem cells in their ability to continuously perceive extra cellular signals and in maintaining a chromatin organization that allows a fast response to the signals. Understanding the molecular mechanism of chromatin reorganization can provide a new perspective on cell wall degradation and regeneration. In this study, the protoplasts from the model organism *Oryza sativa* were used to explore the cell wall degradation and regeneration.

### **Chromatin**

The eukaryotic chromatin is a highly organized DNA and protein supercomplex that plays a critical role in multiple essential biological processes, including genetic information storage, DNA recombination, DNA replication, gene expression regulation, etc. The fundamental repeating unit of chromatin, the nucleosome, is comprised 147 base pairs of genomic DNA wrapped around a histone octamer, with two copies each of four



histones, H2A, H2B, H3 and H4. This structure gives rise to a “beads on a string” like fiber of ~10 nm in diameter (Kornberg, 1974; Olins and Olins, 1974; Horn and Peterson, 2002). The adjacent nucleosomes are bridged by linker DNA associated with histone 1, establishing a further packaged ~30 nm fiber, termed “solenoid” helical fiber (Finch and Klug, 1976; Robinson and Rhodes, 2006; Woodcock et al., 2006). The *in vivo* chromatin structure beyond the 30 nm fiber remains poorly understood. However, it is well known that it not only stores the genetic information but also controls when, where and how a gene is expressed by working together with its interacting proteins. Chromatin exists in many configurations and undergoes dynamic structural changes. Recent evidence has also shown that folding and unfolding of chromatin have a significant impact on gene activity (Misteli, 2001; Fischle et al., 2003; Espada and Esteller, 2007), and chromatin associated proteins are believed to be essential to the processes of structure modulation (Laemmli, 1978).

The chromatin is packed into domains with different degrees of accessibility to the transcriptional machinery, the more open regions are relatively transcriptional active while the condensed regions are inactive (Tackett et al., 2005). Changes to the chromatin structure, called chromatin remodeling, may facilitate gene transcription by loosening the histone-DNA complex, allowing other proteins such as transcription factors access to the DNA. Alternatively, chromatin remodeling where the histone assumes a more closed conformation blocks transcription factor access to the DNA resulting in loss of gene expression. Studies of gene expression maps and chromosomal localization show that clusters of similar expressed genes constitute uniformly transcribed domains (Cohen et al., 2000). It is believed that the status exchange between silent and active domains is regulated by special proteins/complexes. Further research has revealed that boundary or

insulator elements form sharp boundaries between the open and condensed chromatin domains (Burgess-Beusse et al., 2002; Labrador and Corces, 2002). Insulators may be anchored to nuclear lamina or nuclear pore complexes with emanating chromatin fiber loops of active genes for transcription (Gerasimova et al., 2000; Ishii et al., 2002; Labrador and Corces, 2002; Nemeth and Langst, 2004). Furthermore, chromatin and other nuclear components tend to quickly and transiently interact with each other (Espada and Esteller, 2007). Dynamic changes of chromatin structure occur in many cellular processes, including genome replication, DNA recombination, and spatial and temporal coordination of gene expression during growth and development. Shelby *et al* observed centromere movements *in vivo* by using the DNA binding domain of human centromere protein CENP-B fused to GFP (Shelby et al., 1996). Li and coauthors (Li et al., 1998) discovered chromosome arm movement. Chromosome movements are proposed to be connected with DNA replication occurring at specific sites within the nucleus (Hozak et al., 1993). Relatively little is known about how remodeling factors change nucleosome structure and how different factors work together to promote chromatin remodeling. Chromatin remodeling is typically initiated by post-translational modification of the amino acids that make up the histone proteins, as well as through methylation of neighboring DNA (Turner, 2002).

### **Histones and histone post-translational modifications**

In all eukaryotic cells DNA is wrapped almost two superhelical turns around a histone octamer complex that constitutes the nucleosome. The histone octamer complex consists of a tetramer of two heterodimers of histone protein H3/H4 that interacts with two heterodimers of histone H2A/H2B (Luger et al., 1997; Khorasanizadeh, 2004). Distinct

histone variants have been discovered and it is believed that incorporation of different core histone variants may alter various cellular processes and have specific functions (Kamakaka and Biggins, 2005). Histone H3.3 is believed to be enriched in actively transcribed genes and the centromeric H3 (also termed CENP-A). H2A has numerous variants with diverse functions, some involved in gene activation and others in silencing. Some H2B variants are believed to control compaction of chromatin structure during development. H4 appears to be the only histone protein devoid of any variants.

Apart from the distinct histone variants, various covalent modifications of histones are involving in gene regulation. These histone proteins and particularly their N-terminal tails, are subject to a large number of covalent modifications. Such modifications include acetylation, methylation, phosphorylation, ubiquitylation, and ADP-ribosylation (Biel et al., 2005). Those histone modifications are believed to participate in all processes involving the DNA including gene regulation, replication, apoptosis and repair.

Histone proteins have numerous evolutionarily conserved lysine (K) residues, which are subject to acetylation. Numerous acetylation residues have been identified in different histones. Histone acetylation is highly dynamic through the action of histone acetyltransferases (HATs) that transfer acetyl groups to histones and the histone deacetylases (HDACs) that sequentially remove these groups. Histone acetylation is generally found in euchromatin and has long been correlated with transcribing DNAs. Agalioti *et al.* found that H3K4<sub>AC</sub> and H3K9<sub>AC</sub> participate in the recruitment of transcriptional activators and proteins in the transcriptional machinery (Agalioti et al., 2000). A recent report has also shown that a single acetylation of H4K16 can physically

reorganize the chromatin structure by locally dissociating neighbourhood nucleosomes (Shogren-Knaak and Peterson, 2006).

Different chromatin states correlate with specific methylated lysines that can be either mono-methylated (ME), dimethylated (ME2) or tri-methylated (ME3). Different methylation states at different lysines can have different functional outcomes (Sims et al., 2003). Methylation sites consist of K4, K9, K27, K36, K29 at histone H3 and K20 at histone H4, generally. Arginine (R) is also a target for methylation, which may be mono- or di- methylated. Histone methyl transferases (HMTs) are responsible for adding methyl groups to histone lysine and arginine residues. The lysine methylation of histones has been correlated with both active transcription and negative repression, although arginine methylation is generally associated with activation. Methylation of histones can be reversed by histone demethylases.

Phosphorylation has been shown to occur in all histones at serines and threonines. Phosphorylation of histones is important during gene regulation, DNA repair and apoptosis (Nowak and Corces, 2004) and H1 phosphorylation has been found to participate in the cell cycle during cell division (Gurley et al., 1995). Histone ubiquitination, which unlike the poly-ubiquitin chains serving as degradation mark, are involved in cell epigenetic regulation (Peterson and Laniel, 2004).

### **Chromatin reorganization**

Chromatin has two major forms of structures, heterochromatin and euchromatin. Heterochromatin has densely packed, silenced chromatin as opposed to the permissive, actively transcribed euchromatin. Chromatin must retain the flexibility to make genetic information accessible when needed, and therefore, the degree of compaction has to be

tightly regulated. In plants, all cells need to be able to transcriptionally respond to specific environmental cues, either periodic such as light-dark cycles or random as stress conditions. Plant chromatin must display a very dynamic organization to activate or repress specific sets of genes involved in different developmental responses, allowing the integration of developmental programs with the response to environmental signals (Jarillo et al., 2009).

Epigenetic mechanisms are closely linked to and differ in the above-described chromatin states. Particularly, histone modification marks are binding with chromatin regions that are thought to determine the transcription activity. Histone epigenetic marks and the subsequent binding of proteins to modified residues may induce conformational changes in DNA, alter nucleosome positioning, eventually influencing chromatin structure and recruiting downstream signaling proteins to regulate cellular processes (Minard et al., 2009). For example, methylation of histone H3 at lysine 9 generates a 'code' for the recruitment of heterochromatin protein 1 (HP1), thereby inducing the assembly of restrictive chromatin and, consequently, gene silencing (Bannister et al., 2001). Although those epigenetic markers have served to define heterochromatin and euchromatin, such markers often are more ambiguous than previously thought. For example, even the classical heterochromatin markers Heterochromatin Protein 1 (HP1) and histone 3 lysine 9 trimethylation have now been found in euchromatin (Libault et al., 2005; Vakoc et al., 2005).

It has been widely accepted that chromatin structure plays a critical role in the regulation of eukaryotic gene transcription. More recent studies have led to the striking observation that several protein complexes involved in transcription regulation can function, at least in part, by altering chromatin structure (Kingston and Narlikar, 1999;

Kornberg and Lorch, 1999). The regulation of chromatin structure to expose or occlude a particular DNA segment is controlled by the dynamic interplay between sequence-specific DNA-binding proteins, histone variants, histone-modifying enzymes, chromatin-associated proteins, histone chaperones and ATP-dependent nucleosome remodelers (Li et al., 2007). The most widely characterized chromatin-modifying complexes studied to date are ATP-dependent structure chromatin remodeling complexes, which use the energy of ATP hydrolysis to locally disrupt or alter the nucleosomal conformation. The SWI/SNF complexes and the ISWI-based complexes are two major classes of chromatin remodeling complexes. The mechanisms by which the ATP-dependent chromatin remodeling complexes alter nucleosome and chromatin structure are not yet clear. However, pieces of information that illustrate the similarities and differences of these complexes are becoming available. ATP hydrolysis is the driving force for both complexes. While the SWI/SNF proteins disrupt DNA-histone contacts and release the histones, the NURF family of nucleosome remodelers merely slides the nucleosomes along the DNA (Vignali et al., 2000; Gangaraju and Bartholomew, 2007).

### **Epigenetics and cell wall regeneration**

Little is known about epigenetic regulation of cell wall synthesis. However, gene expression profile studies on the mutant of the *Arabidopsis* homolog of the trithorax gene (ATX1) have revealed some very exciting results (Ndamukong et al., 2009). ATX1 methylates lysine 4 on histone H3. Among the differentially expressed genes in the *atx1* mutant, the genes encoding endomembrane and cell wall metabolic pathway proteins were substantially overrepresented, representing about 20% of the total genes with altered gene expression. Zhong and Ye (Zhong and Ye, 2007) investigated regulation of HD-ZIP

III genes by microRNA 165 and demonstrated that microRNA 165 regulates vascular patterning and consequently cell wall synthesis. A group of small-interfering RNAs from natural antisense transcripts derived from a cellulose synthase gene modulate cell wall biosynthesis in barley (Held et al., 2008).

It has long been assumed that epigenetic mechanisms that take place on chromatin are somehow linked to dedifferentiation and regeneration. Recently, it has been shown that two rounds of chromatin decondensation are required for tobacco cells to undergo dedifferentiation and re enter the cell division (Zhao et al., 2001). The first phase takes place during enzymatic digestion of cell wall in the course of protoplast isolation. The second decondensation occurs only after protoplasts are induced with phytohormones for re-enter of the cell cycle (Grafi, 2004). In the absence of hormonal application, protoplasts undergo cycles of chromatin condensation/decondensation and die (Zhao et al., 2001). Although there are only limited reports on epigenetic regulation of cell wall in plants, we expect that evidence will emerge rapidly.

### **Proteomics**

A proteome can be described as the protein complement to the genome of that particular organism, or put simply, the proteome is the inventory of proteins which are being expressed by an organism under a particular set of conditions at a particular moment in time. Bradshaw and Burlingame defined the proteome as encompassing “all the proteins expressed by a genome, cell or tissue” (Bradshaw et al., 2005). It represents a bridge between genomic analysis and protein expression of regulatory biomolecules. Genomics can provide the data of potential proteins that can present in a cell, but only proteomics can reveal what proteins are really present under particular biological

conditions. Microarrays have been used to study total gene expression profiles and thus the functions of gene products. However, post translational modifications, protein half-lives, and other protein processing events make it difficult to correlate mRNA expression with protein expression. Many studies have shown that there is poor correlation between mRNA and protein expression levels (Anderson and Seilhamer, 1997; Gygi et al., 1999). Unfortunately, unlike DNA, which can be amplified by PCR, proteins cannot be amplified and detection is limited by their *in vivo* abundance, and unlike the genome, which is static, the proteome is dynamic and is dependant on various factors such as environmental stimulation and development stage. The information made available by genomic data can help in the field of proteomics. Experimental data can be compared to theoretical data calculated from an organism's genome to identify hypothetical proteins which have not been previously identified or characterized. Databases searched using bioinformatics programs often incorporate both genomic and proteomic data from previously characterized proteins.

With the proteome as its target, the field of proteomics has among its goals the comprehensive identification, quantitation and functional characterization of a given proteome. Achieving these goals would then allow for the comprehensive comparison of the changes in patterns of protein expression, modification, interactions and activity that occur in response to a given stimulus. However, comprehensive coverage of the proteome remains a distant possibility and there is doubt as to whether it can ever be achieved (Bradshaw et al., 2005). Nevertheless, significant technological and methodological developments in the field have advanced us to a point where we are beginning to see applications for proteomics.



Currently two methods, “top-down” and “bottom-up”, are performed for proteomics analysis. However, only “bottom-up” proteomics is described and used in this dissertation. Both techniques rely on mass spectrometry for the protein analysis. Both have evolved with time and are said to be complimentary techniques rather than one being a better approach than the other. In top-down proteomics, whole proteins (not subjected to enzymatic digestion) are subjected to mass spectrometry analysis (Kelleher, 2004; McLafferty et al., 2007). The information obtained from this type of analysis is often useful in the determination of post-translational modifications of proteins. The intact proteins can then be fragmented and the peptides resulting from this fragmentation can be analyzed to obtain sequence information which can aid in the identification of the protein. (McLafferty et al., 2007).

Bottom-up proteomics on the other hand, involves working with peptides generated from proteolytic digested proteins for mass spectrometry analysis. This approach can be carried out via two basic strategies. The first strategy of bottom-up is known as the peptide mass fingerprinting. In peptide mass fingerprinting, peptide masses obtained from an MS scan are compared to calculated peptide masses generated by "in silico" cleavage of protein in the database. The disadvantages are the requirement for simple and pure mixture of proteins and several peptides to uniquely identify a protein. A second strategy is by peptide fragmentation. The approach starts to convert a protein or protein mixture into peptides by protease digestion. Generated peptides are then resolved by different types of chromatography columns, and each peptide is subjected to MS/MS analysis to produce peptide fragmentations. Through electrospray ionization (ESI), peptides in the HPLC effluent are directly introduced into the mass analyzer, such as ion trap or triple quadrupole, allowing on-line MS/MS analysis.

## **Two-dimensional gel electrophoresis coupled with MALDI-TOF-MS**

The traditional approach for proteomic analyses is separating protein mixtures on Two-dimensional Polyacrylamide Gel Electrophoresis (2D-PAGE) gels followed by characterization of individual proteins/peptides by mass spectrometry. Two-dimensional gel electrophoresis is a combination of two types of gel electrophoresis: isoelectric focusing (IEF) and sodium dodecyl sulfate- polyacrylamide gel electrophoresis (SDS-PAGE) (O'Farrell, 1975). IEF separates proteins by their isoelectric points using a pH gradient across an electric field. The isoelectric point of a protein is a characteristic pH at which its net charge is zero. When a protein is placed in a medium with a pH gradient and subjected to an electric field, it initially moves towards the electrode with the opposite charge. Eventually the protein will arrive at the position of its isoelectric point in the pH gradient and become immobilized (Righetti and Bossi, 1997). Immobilized pH gradient (IPG) gel strips are most commonly used for isoelectric focusing now.

The second dimension in 2-D gel electrophoresis is SDS-PAGE. This method separates proteins according to their molecular weights. Dodecyl sulfate was used to denature proteins and eliminate positive charges on proteins. The net effect is that proteins migrate as ellipsoids with a uniform negative charge-to-mass ratio, with mobility related to mass (Blomberg et al., 1995). As a result, 2-D gel electrophoresis can provide protein mixture as arrays of spots on a polyacrylamide gel. The spots can be visualized by several different gel staining methods. These methods include Coomassie blue staining, silver staining and fluorescent staining, which is used in this study. Proteins will be stained in proportion to the amount of their basic and aromatic amino acids and the amount of sample in the spot. Protein quantitation can be determined based on density of the reagents binding to the proteins. However, low abundance spots are not detected by

Coomassie blue staining due to low sensitivity. Though, silver staining is very sensitive, there is nonspecific binding of the dye to gel and it is difficult to stop the chemical reaction on multiple gels at the same time. The recent development of fluorescent dyes such as SYPRO Ruby overcomes these limitations (Lopez et al., 2000). The binding of SYPRO Ruby to the amount of protein is in linear ratio and quite sensitive. In addition, phosphoprotein specific dye, Pro-Q Diamond, (Schulenberg et al., 2004) specifically binds to phosphoproteins and the Pro-Q Emerald glycoprotein gel stain specifically binds to glycoproteins (Hart et al., 2003).

Proteins of interest are excised and subjected to in-gel digestion. The digests are further analyzed to generate mass spectra by MALDI-TOF-MS. The extracted peptides are first mixed with a matrix ( $\alpha$ -cyanohydroxycinnamic acid), which co-crystallizes with the peptides sample on a plate and when ionization occurs these matrix molecules absorb the energy causing thermal expansion of the matrix and the analyte into gas phase (Egelhofer et al., 2002; Lim et al., 2003). MALDI produces singly charged ions predominantly and these charged ions are analyzed by detectors based on time taken by the ions to reach the detection, which is related to the  $m/z$  ratio of the ion (Egelhofer et al., 2002). The experimental peptide masses are then compared to the theoretical ones from the insilico digestion of the protein database to identify the protein of interest. This process of identifying the proteins by using masses of peptides to search against a pool of known theoretical masses of peptides is called peptide fingerprinting (Cottrell, 1994).

Two-DE is a good tool to differentiate proteins that are isoforms having similar charges and vary in their size or molecular weight. These are also useful for identifying proteins that shift in their positions due to post-translational modifications. 2-DE is a powerful technique that simultaneously separates thousands of proteins and allows

comparative protein profiling between different crude biological samples (Klose and Kobalz, 1995). Zoom gels or narrower focus pH gradients during iso-electric focusing help to detect observing proteins of interest in a particular pH range. Another advantage of 2-DE method is it also can be performed by using two samples labeled with distinct fluorescent dyes on a single gel. This technique is known as Difference In Gel Electrophoresis (DIGE) (Unlu et al., 1997). However, there are some limitations of 2-DE. It is incapable of detecting the majority of protein components, the ones that are detected are mainly the high abundance ones. The limitation in protein loading on the IPG strips as well as the protein abundance within a cell that differ by up to 10 orders of magnitude are the main causes for this discrimination. Besides, buffers that are specified for isoelectric focusing are limited; hence it is difficult to use certain detergents to increase solubility. Few proteins with extreme pI or molecular weight can be resolved on 2-DE gels.

### **2D-LC coupled electrospray ionization mass spectrometry**

Limitations of the 2-DE, such as resolution of proteins with high abundance, co-immigration of proteins, detection of highly acidic or basic proteins have now been overcome by the use of MudPIT (Multidimensional Protein Identification Technology). The MudPIT technique is used for the separation and identification of complex protein and peptide mixtures using liquid chromatography and it is directly connected with the ion source of a mass spectrometer (Link et al., 1999; Washburn et al., 2001). This technique is often called “shotgun” proteomics.

In this dissertation, the 2D-LC-ESI-MS/MS approach was performed and discussed. Profiling of proteins/peptides using 2D-LC can be performed using two step-

chromatography separation. In the first step, ion exchange chromatography (eg, strong cation-exchange column, SCX) is used to separate peptides based on their charge and increasing concentration of salt is used to free peptides from resin after which they bind to a reversed phase resin. In the second step, a gradient of increasing hydrophobicity is used to progressively elute peptides from the reverse phase chromatography resin into Electrospray Ionization Mass Spectrometer (ESI-MS) (Liu et al., 2002). ESI is widely used throughout the biochemical field, since it is easily coupled with high performance liquid chromatography (HPLC), micro-HPLC and capillary electrophoresis. An electrospray is produced by applying a strong electric field which is obtained by the use of a potential difference of 3-6 kV between a capillary and its counter-electrode (Loo et al., 1989; Lin et al., 2003). A unique characteristic of ESI is the production of multiply charged ions for big molecules such as proteins and peptides. These multiply charged ions are either in positive-ion mode, or negative-ion mode, and also cover a range of charge states. Multiple charged ions are produced which lowers the  $m/z$  values for high molecular weight compounds and it allows measurement of  $m/z$  values on mass spectrometers with a weak  $m/z$  ranges limit (Lin et al., 2003). The ionized peptides undergo Collision-Induced Dissociation (CID), and the fragment ions are recorded in a tandem mass spectrometer. The experimental MS/MS spectrum of each peptide is compared with the theoretical MS/MS spectra to elucidate its amino acid sequence using modern protein database search algorithms such as Mascot, Mowse, MS-Fit, Profound (Aebersold and Mann, 2003; Kelleher, 2004; McLafferty et al., 2007). In this dissertation, the spectral information is matched with databases using SEQUEST algorithm and protein identification is done by DTASelect algorithm (Link et al., 1999; Washburn et al., 2001). Lack of quantitative information availability using this technology is probably its

only limitation, however, it can be overcome by adopting ICAT and iTRAQ techniques. In addition, nowadays, several label free quantitation methods have been developed and improved for protein quantification.

The label-free methods are high throughput and completely eliminate the labor intensive 2-DE gel separation or sample labeling steps. In addition, they can separate a very wide range of proteins and overcome the protein solubility problem that is often encountered in the 2-DE gel method. The reported label-free methods include peptide counts, sequence coverage, peak area intensity measurements, spectral counts, and the sum of the TurboSEQUEST cross correlation coefficient (Xcorr) of peptides in a protein (Bantscheff et al., 2007; Wong et al., 2008). It has been shown that spectral counts, the total number of MS/MS spectra taken on peptides from a given protein in a given LC/LC-MS/MS analysis, are linearly correlated with the protein abundance over a dynamic range of 2 orders of magnitude (Liu et al., 2004; Old et al., 2005; He and Zhang, 2006; Schmidt et al., 2007). It has been reported that, compared to peak intensity, the spectral count method provides more sensitivity in detecting proteins undergoing change in abundance and the results match well with 1-D gel staining intensity measurement (Old et al., 2005). The TurboSEQUEST algorithm has been widely used in peptide and protein identification. It calculates a cross correlation to quantitatively measure the relatedness of experimental mass spectra to the *in silico* generated tandem mass spectra based on protein sequence (Eng et al., 1994). The cross correlation coefficient (Xcorr) is determined by factors including the number of fragment ions in the mass spectrum, their relative abundance, continuity of ion series, and presence of immonium ions for certain amino acids in the spectrum, all of which are proportional to the concentration of the precursor ion. Nanduri *et al* (Nanduri et al., 2005) have reported that the sum of

SEQUEST cross correlation coefficient ( $\Sigma X_{corr}$ ) correlates well with the concentrations of a known protein mixture in serial dilutions.

By using this shotgun proteomics coupled with label-free quantification technique, which is sometimes automated, complex mixtures of peptides can be resolved and many proteins identified, including some of the less abundant ones, which have demonstrated great potential in comparative proteomic studies.

### **Comparative proteomics**

Comparative proteomics analyzes relative protein expression in two or more samples. The samples are taken from organisms in the control and treatments.

Traditionally, protein differential expression is determined by contrasting the position and density of spots in 2-D electrophoresis gels. Recently, shotgun proteomics coupled with label-free quantification have demonstrated great potential in comparative proteomic studies. The label-free methods are high throughput and completely eliminate the labor intensive 2-DE gel separation or sample labeling steps. In this research, 2D-LC is used in combination with mass spectrometry to perform quantitative and comparative proteomics analyses together with bioinformatics algorithm for computing the ion intensities. The ion intensity-based quantitative approaches have progressively gained more popularity as mass spectrometry performance and bioinformatics have improved significantly.

The shotgun proteomics and the corresponding protein identification have utility for analysis of cellular protein expression changes in plants. In rice, comparative studies on leaf before and after wounding have led to identifications of proteins that differentially regulate due to wound stress (Shen et al., 2003). Other studies on rice such as green versus etiolated rice shoots (Komatsu et al., 1999) and proteome of rice after treatment

with jasmonic acid (Rakwal and Komatsu, 2000) and brassinolide (Konishi and Komatsu, 2003) has led to identification of many proteins that play a role in plant normal physiological versus the treated conditions. In *Arabidopsis thaliana*, comparative proteomic investigation of cell wall and extra-cellular matrix when treated with an elicitor suspension and the non-treated tissue was performed to identify proteins induced during the stress (Ndimba et al., 2003). In this study, we evaluated protein expression differences to identify markers of cell wall degradation and regeneration progression of *Oryza sativa* to gain further insight into potential mechanisms underlying these changes. 2D-LC fractionation techniques before mass detection simplify the complex proteome. Differentially expressed proteins were identified by a label-free quantitative proteomic approach.

### **Systems biology**

It is estimated that more than 1500 genes are involved in cell wall synthesis in *Arabidopsis thaliana* (Initiative, 2000; Somerville et al., 2004). Due to the high complexity of cell wall synthesis, the cross-talk and integration of different pathways are still poorly understood. Using a systems biology approach may provide a more completed picture of cell wall synthesis and regulation. Kitano (Kitano, 2002) proposed systems biology, a systematic way to visualize multiple related biological processes in a network. This concept is generally applied to interpreting interactions of genes or gene products (Somerville et al., 2004). Molecular systems approach with the interaction networks not only can identify direct and indirect global responses of genes to the objective network, but also will allow us to identify key regulatory nodes in networks (He and Zhang, 2006). Recently, several computational tools and databases have been implemented that can be



directly applied to existing public information and map genes/proteins into networks and pathways. Pathway Studio (Ariadne Inc., Rockville, MD, USA), one of the most widely used and commercially available software, enables researchers to navigate and analyze biological pathways, gene regulation networks and protein interaction maps (Nikitin et al., 2003). Pathway Studio database, MedScan (Ariadne Inc., Rockville, MD, USA), works in many species (Buza and Burgess, 2008) because it is equipped with an automated text mining engine. Multiple aspects of protein function, including cellular location, protein-protein interactions, protein modifications, gene expression regulation, and regulation of various cellular processes are also included for many species (Novichkova et al., 2003). However, the database of plant genes and proteins is still limited. Nevertheless, this available systems biology tool makes it possible to examine cell wall re-synthesis in protoplasts using a molecular systems biology approach.

## CHAPTER III

### PROTEOME PROFILING OF RICE CHROMATIN

#### **Abstract**

The eukaryotic chromatin/chromosome stores genomic information, controls genetic material distribution, and plays an essential role in the establishment and maintenance of spatial and temporal gene expression profile. Despite over a century of research, the protein composition and higher level structure of chromatin still remain obscure, particularly in plants. Here, we have developed a protocol for chromatin purification from rice suspension cells and examined proteins associated with chromatin using 2-Dimensional Gel Electrophoresis (2-DE) and Multidimensional Protein Identification Technology (MudPIT). Total proteins were isolated from freshly prepared rice chromatin and separated on 2-DE gels. Gel images were analyzed by PD Quest software (BioRad). 972 distinct protein spots have been resolved on SYPRO Ruby stained 2-DE gels and 509 proteins have been identified by MALDI-TOF analyses with Confidence Intervals (C.I.%) over 95%. When the chromatin co-purified proteins are examined using shotgun method (2D-LC-MS/MS), 292 unique proteins have been identified with at least two peptides were hit for an individual protein. Interestingly, a large number of histone variants in addition to the four common core histones have been identified. Other proteins identified include nucleosome assembly proteins, high mobility group proteins, histone modification proteins, transcription factors, and a large number of hypothetical and function-unknown proteins. The gene corresponding to all these proteins

were characterized using the Gene Ontology tools. They were grouped into different levels and presented in pie charts. Gene Ontology studies reveal that most chromatin associated proteins are involved in cellular organization, biosynthesis and metabolic pathways. This study has provided new insight into chromatin composition in plant.

### **Introduction**

Recent advances in proteomic technologies have significantly facilitated studies on chromatin and chromosome associated proteins. Human metaphase chromosomes have been purified and followed with comparative proteomics studies (Uchiyama et al., 2005). Among the identified proteins, the main components are mitochondrion proteins (38.6%), nuclear proteins (29.8%), ribosomal proteins (12.7%), cytoplasmic proteins (11.4%), Cytoskeleton proteins (4.4%), and unknown proteins (3.2%). Chu *et al* reported the proteome of spermatogenic chromatin in *C. elegans* (Chu et al., 2006). In addition, they found that 17 out of the 32 knockouts of mouse genes, which were homologous to the *C. elegans* genes encoding chromatin associated proteins, result in male sterility in mouse, demonstrating that targeting the chromatin associated proteins has high potential to identify regulatory proteins critical to cellular processes. Shiio *et al* studied human B lymphocyte chromatin enriched fractions and identified 64 proteins including 18 putative transcription factors (Shiio et al., 2003). Proteome studies on chromatin associated proteins in plants have not been reported.

In this report, we purified nuclei and chromatin from rice suspension culture and examined the chromatin associated proteins using 2-DE gel and shotgun approaches. 509 proteins have been identified using 2-DE gel method and 292 unique proteins identified by shotgun approach including histones, histone variants, many chromatin binding

proteins and function-unknown proteins. Our studies have provided new insight into chromatin composition in plants.

## **Materials and methods**

### **Suspension culture and protoplast isolation**

A rice (*Oryza sativa*) suspension culture was used for protoplast isolation. The suspension cells were grown at 24°C with constant shaking on a gyratory shaker at 150 rpm in B5 organic medium (pH 5.7) supplemented with 20 g/L sucrose, 0.5 g/L MES 2-(*N*-morpholino)ethane-sulfonic acid, and 2.0 mg/L 2,4-Dichlorophenoxyacetic acid (2,4-D) and subcultured weekly. For protoplast isolation, cells were harvested 4 days after subculture. The protoplasts were generated using a method as reported by Yasuyuki Yamada *et al* (Yamada et al., 1986) with modifications. In brief, suspension cultured cells were added to filter-sterilized enzyme solution containing 2.5% Cellulose RS [Onozuka RS], 1% Macerozyme R10 (Research Products International Corp.), 0.4 M mannitol, 80 mM CaCl<sub>2</sub>, 0.125 mM MgCl<sub>2</sub>, 0.5 mM MES, and B5 organic medium plus 2.0 mg/l 2,4-D (pH 5.6). After 9 hours of incubation at 25 °C in the darkness, the released protoplasts were filtered through a 25 µm stainless steel filter, collected by a centrifugation at 120g for 5 min, and washed 3 times with protoplast suspension medium (0.4 M mannitol, 80 mM CaCl<sub>2</sub>, 0.125 mM MgCl<sub>2</sub>, 0.5 mM MES, and B5 organic medium at pH 5.6). A yield of  $24 \times 10^6$  protoplasts per gram suspension cells was obtained.

### **Nucleus isolation and chromatin isolation**

The collected protoplasts were resuspended in Nuclei Isolation Buffer (NIB: 0.25 M sucrose, 10 mM Tris-HCl (pH 8.0), 2.0 mM MgCl<sub>2</sub>, 1.0 mM CaCl<sub>2</sub>, 0.1 mM

spermidine, 0.5% Ficoll, 0.5% Triton X-100, 1.0 mM DTT and 1.0 mM PMSF added freshly) in a concentration of no more than  $10^6$  protoplasts/ml and were ruptured by constant shaking at 4 °C for 15 min. Raw nuclei were collected by centrifugation at 500g for 10 min at 4 °C. The pellet of nuclei was resuspended in NIB, layered onto 10 ml 2M sucrose, and pelleted by a centrifugation at 6000g for 15 min at 4 °C. Pure nuclei were obtained after three washes with NIB. The purified nuclei were suspended in a Chromatin Isolation Buffer (CIB: 10 mM Tris-HCl pH 7.5, 0.1 mM EDTA, 1.0 mM DTT, 1.0 mM PMSF) followed with constant shaking at 4 °C for 25 min to break the nuclei. The raw chromatin was collected after centrifugation at 750g for 10 min at 4 °C, resuspended in the CIB buffer, layered onto 10 ml of 2M sucrose, and pelleted again by spinning at 7600g for 15 min at 4 °C. The chromatin pellet was washed three times with CIB and used for chromatin protein extraction or directly used for electron microscopy.

### **Electron microscopy**

The freshly isolated chromatin was used for transmission electron microscopy examination using the preparation method described by Vengerov *et al* (Vengerov and Popenko, 1977) with minor modifications. In brief, chromatin was diluted to a final concentration  $\sim 5\mu\text{g}/\mu\text{l}$  in TE buffer containing 0.25 M ammonium acetate. After 10 min incubation, 2.0  $\mu\text{l}$  of cytochrome C at  $0.2\mu\text{g}/\mu\text{l}$  was added, mixed gently, and kept at room temperature for 90 seconds. A drop of chromatin solution was carefully placed onto a prepared Carbon-Coated Grid. After removal of the excess liquid by filter paper, the sample was dehydrated by dipping into 75% ethanol for 45 seconds, rinsed in 90% ethanol for 2 seconds and air-dried. For contrast enhancement, the grids were rotary

shadowed with Pt-Pd. Electron microscopy was performed using a JEM-100CX II electron microscope (JEOL USA, Inc.).

### **Protein extraction**

Proteins were extracted as reported (Hurkman and Tanaka, 1986; Saravanan and Rose, 2004) with minor modifications. Briefly, the chromatin pellet was resuspended in phenol extraction buffer (PEB: 0.9 M sucrose, 0.5 M Tris-HCl, 50 mM EDTA, 0.1 M KCl, 1% Triton X-100, and 2% freshly added  $\beta$ -mercaptoethanol, final pH 8.7) and sonicated with a microtip probe (Misonix XL 2020) on an ice bath for 5 min with an intermittent cooling every 30 sec. The sample was then mixed with an equal volume of saturated phenol (PH 8.0) and then vortexed for 1 min. The phenol phase was collected after a centrifugation at 3,000g for 10 min at 4 °C. The proteins were precipitated with five volumes of precipitation buffer (PB: 0.1 M ammonium acetate and 1%  $\beta$ -mercaptoethanol in methanol) at -70 °C overnight. The protein pellet was recovered by a centrifugation at 17,000g for 10 min at 4°C and washed three times with pre-chilled PB and another three washes with pre-chilled 70% ethanol. The protein pellet was lyophilized to powder in a speed vacuum (LABCONCO, model LYPH-LOCK 6) and stored at -70 °C.

### **Western blots**

20  $\mu$ g/lane of total proteins, nuclear proteins, and chromatin associated proteins were separated by a 15% SDS PAGE and processed for Western blotting using standard procedures. The antibodies against H1, H2B and H3K14<sub>AC</sub> were purchased from UPSTATE. Actin and  $\alpha$ -tubulin antibodies were bought from Santa Cruz Biotechnology

and SIGMA, respectively. Antibodies against COP9 signalosome subunit 3 and 6, and PhyA were kindly provided by Xingwang Deng's Lab.

### **Two-dimensional polyacrylamide gel electrophoresis (2D PAGE)**

Proteins were dissolved thoroughly in rehydration buffer (7M urea, 2M thiourea, 4% CHAPSO, 1% DTT, and 0.2% Ampholines) and centrifuged at 12,000g for 10 min to remove un-dissolved contents. The supernatant was quantified using a Bio-Rad Rc Dc protein assay kit according to manufacturer's instructions. The quantified proteins were then used for 2D PAGE. Isoelectric focusing (IEF) was carried out using a Bio-Rad PROTEAN IEF cell on 17cm 3-11 pH non linear IPG strips (GE). Four hundred microgram of protein in 400  $\mu$ l of rehydration buffer was loaded into the IEF tray and active rehydration was carried out at 23°C for 12 hours, followed by 250 V for 2 hours, a linear increase of voltage to 10,000V for 4 hours, and the isoelectric focusing was performed at 23 °C for a total of 90,000 VH. After the completion of IEF, the strips were equilibrated in a buffer containing 6M urea, 0.375M Tris-HCl (pH 6.8), 20% glycerol, 2% SDS, and 2% dithiothreitol for 15min and followed by equilibration for another 15 min in a buffer containing 6M urea, 0.375M Tris-HCl (pH 6.8), 20% glycerol, 2% SDS, 0.1% bromophenol blue, and 2.5% iodoacetamide. The equilibrated IPG strips were then loaded on horizontal slab gels (19 × 18 × 1.5mm) containing 12% (w/v) separating gel and 4% stacking gel (w/v). Electrophoresis was carried out in a Bio-Rad PROTEAN PLUS horizontal Dodeca cell at 15 mA/gel.

The gels were stained with SYPRO Ruby fluorescence stain (Bio-Rad) according to the protocols provided by the manufacturer and scanned with a VersaDoc4000 image system (Bio-Rad). At least three 2-DE gels representing three biological repeats were

used for data analyses. The images were analyzed with PDQUEST 7.4.0. software (Bio-Rad, Hercules, CA), including gel cropping, anchor spots selection, and alignment. The spots with consistent size and shape within replicate groups were considered as a protein spot. The protein spots were also checked manually to ensure that all analyzed spots were true protein spots and the gel alignment was appropriate.

### **In gel digestion and mass spectrometry**

After PDQUEST analysis, the spots of interest were robotically excised from 2-DE gels by a Proteome Works Spot Cutter (Bio-Rad). In-gel trypsin digestion was performed using the ProPrep (Genomic Solutions) robotic digester/spotter. The samples were subjected to disulfide bond reduction and alkylation with DTT (dithiotreitol) and iodoacetamide, respectively. The resulting peptide mix was desalted with C18 ZipTips (Millipore) and spotted on a MALDI plate in a solution containing 70% acetonitrile, 0.1% trifluoroacetic acid, and 5mg/ml matrix (alpha-cyano-4-hydroxycinnamic acid).

Mass spectra were collected on an ABI 4700 Proteomics Analyzer (Applied Biosystems) MALDI TOF-TOF mass spectrometer, and protein identification was performed using the Result Dependent Analysis (RDA) of ABI GPS Explorer software, version 3.5 (Applied Biosystems). Some of the crucial parameters were set as follows: Digestion enzyme: trypsin with one miss cleavage; MS (precursor-ion) peak filtering: 800 - 4000 m/z interval, monoisotopic, minimum S/N=10, mass tolerance = 150 ppm. MSMS (fragment-ion) peak filtering: monoisotopic,  $M+H^+$ , minimum S/N=3, MSMS fragment tolerance = 0.2 Da; Database used: *Oryza* taxonomic sub-database of “nr” (non redundant) database of National Center for Biotechnology Information (NCBI)



During the initial MS scan, data were analyzed as peptide mass fingerprinting (PMF) and preliminary protein ID was done by searching against the database using the MASCOT (Matrix Science) algorithm. Proteins with high confidence ID (Cross Confidence Interval C.I. % > 95%) were automatically selected for “*in silico*” digestion and their three most prevalent corresponding peptides-precursor ions present in the MS spectra were selected for MSMS analysis: RDA\_1 (top protein confirmation). The sample spots not yielding high confidence ID after preliminary PMF ID and/or after RDA\_1 ID, were subjected to RDA\_2 by selecting the first 20 most intense precursor ions in the MS spectra for MSMS analysis. The spectral data from the PMF (initial MS scan), RDA\_1 and RDA\_2 MSMS were together subjected to combined MASCOT search. Only proteins with total Protein Score C.I. % > 95 % were considered as a positive ID.

#### **Sample preparation for MudPIT and 2D-LC-MS/MS analysis**

The protein pellet was dissolved in 6M urea with 100mM Tris-Cl (pH7.8) and centrifuged at 16,000g for 10 min. The supernatant was collected and then quantified using a Rc Dc kit (Bio-Rad). Proteins (100µg in 50 µl) were reduced by mixing with 20 µl of 200mM DTT in 100mM Tris-Cl (pH7.8) for 1hr at room temperature, alkylated with 20 µl of 200mM iodoacetamide in 100mM Tris-Cl (pH7.8) in darkness for 1 hr, and diluted to a final urea concentration of 0.6M, a concentration at which trypsin retains its activity. Trypsin solution was added to a final ratio of enzyme to substrate of 1/50. The digestion was carried out at 37°C and stopped by adding 10µl of 10mM lysine after 15 hrs. The pH of the reaction mixture was then adjusted to below 6.0 and vacuum dried to a final volume of 25 µl. The peptides were desalted using a peptide macro trap (Michrom Bioresources, Inc., Auburn, CA) following the protocol provided by the manufacturer

and eluted with 0.1% trifluoroacetic acid in 95% acetonitrile. The eluted peptides were vacuum dried to pellet and redissolved in 20  $\mu$ l of 0.1% formic acid with 5% acetonitrile.

The peptide mixtures were subjected to two-dimensional liquid chromatography (2-D LC) comprising a separation on a strong cation exchange column (SCX BioBasic 0.32  $\times$  100 mm) followed by a reverse phase (RP) column (BioBasic C18, 0.18  $\times$  100-mm Thermo Hypersil-Keystone, Bellefonte, PA) coupled directly in-line with electro spray ionization ion trap mass spectrometer (ProteomeX workstation ThermoFinnigan). A flow rate of 3  $\mu$ l/min was used for both SCX and reverse phase columns. For SCX, a salt step-gradient of 0, 10, 15, 25, 30, 35, 40, 45, 50, 57, 64, 90, and 700 mM ammonium acetate in 5% acetonitrile and 0.1% formic acid was applied. The eluted peptides were loaded directly on the reverse phase column, equilibrated with 0.1% formic acid and 5.0% acetonitrile. The peptides were eluted from the reverse phase column by an acetonitrile gradient (in 0.1% formic acid) as follows: 5%-30% for 30 min, 30%-65% for 9 min, 95% for 5 min, and 5% for 15 min - a total of 59 min of elution.

The LCQ Deca XP ion trap mass spectrometer was configured to optimize the duty cycle length with the quality of data acquired, by alternating between a single full MS scan followed by three tandem MS/MS scans on the three most intense precursor masses (as determined by XCALIBUR mass spectrometer software in real time) from the full scan. The collision energy was normalized to 35%. Dynamic mass exclusion windows were 2 min long. In addition, MS spectra for all samples were measured with an overall mass/charge ( $m/z$ ) range of 200 to 2,000. The mass spectra and tandem mass spectra produced were used to search the *Oryza sativa* no redundancy protein database (TIGR, V5.0) from TIGR Rice Genome Annotation (<http://rice.plantbiology.msu.edu>) by using TurboSEQUEST, Bioworks Browser 3.2 (Thermo Electron Corp.).

TurboSEQUEST cross-correlates experimentally acquired mass spectra with theoretical mass spectra generated *in silico*. The idealized spectra were weighted with b and y fragment ions. Trypsin digestion was applied to generate the "precursor ions" and the database included mass changes due to cysteine carbamidomethylation and methionine oxidation as differential (variable) modification. The allowance for missed cleavages is one. The peptide (precursor) ion mass tolerance was 1.0 Da, and the fragment ion (MS2) tolerance was 0.5 Da. The general requirement for protein identification was two peptides from a protein to meet the following criteria: X-correlation >1.9 (+1 charge), >2.2 (+2 charge), >3.75 (+3 charge); delta correlation value  $\geq 0.08$ ; probability < 0.01. We used the reverse database functionality in Bioworks 3.2 and searched tandem MS (MS2) data against a reverse *Oryza sativa* database using the same search criteria as described above. The peptide false positive rates were estimated.

### **Gene ontology (GO) annotation**

Functional categorization of proteins was carried out according to the gene ontology (GO) rules using the GO browser at <http://www.geneontology.org/> (Ashburner et al., 2000). The three independent GO ontologies used to describe the gene products are: (i) the biological process (BP) in which the gene product participates; (ii) the molecular function (MF) that describes the gene product activities, such as catalytic or binding activities, at the molecular level; and (iii) the cellular component (CC) where the gene product can be found. GO annotations were obtained from GORetriever, a program available at AgBase (McCarthy et al., 2007) (<http://www.agbase.msstate.edu/>). GO pie charts were then generated by using GOSlim-Viewer provided by AgBase (<http://www.agbase.msstate.edu/>). AgriGO (Du et al., 2010) was also performed to

identify GO annotations for which each of the groups of genes was significantly enriched compared to all genes in the genome.

## Results

### Chromatin purification from rice (*Oryza sativa*) suspension cells

For successful proteome analysis of chromatin associated proteins, a highly purified and structurally intact chromatin preparation is ideal. However, chromatin is composed of a folded string of DNA and protein, which is "floating" in the nuclear solution that contains over thousands of proteins. Therefore, it is almost impossible to obtain 100% pure chromatin without pulling down other proteins. Both studies in human chromosome and *C. elegans* chromatin had revealed contamination by other proteins (Uchiyama et al., 2005; Chu et al., 2006). Rice suspension culture cells were used for chromatin isolation and purification in our studies. As shown in Figure 3.1A, pure and intact protoplasts were obtained from suspension cells by enzymatic treatment of the cell wall for 9 hours. Highly enriched nuclei were obtained on a large scale as revealed by fluorescence microscopy after DAPI staining (Figure 3.1B). Chromatin, which was released from the enriched nuclei, was homogeneous in appearance and no visible contamination of other organelles as confirmed by DAPI stain (Figure 3.1C) and examination under phase contrast microscopy. To further validate the quality of our chromatin preparation, the purified chromatin was examined using a transmission electron microscope after being shadowed with Pt-Pd to increase contrast. As shown in Figure 3.1D, large clusters of thick chromatin fiber were observed. Obviously, the above 30 nm structure (Figure 3.1D) of the chromatin was, at least partially, maintained in our chromatin preparation.

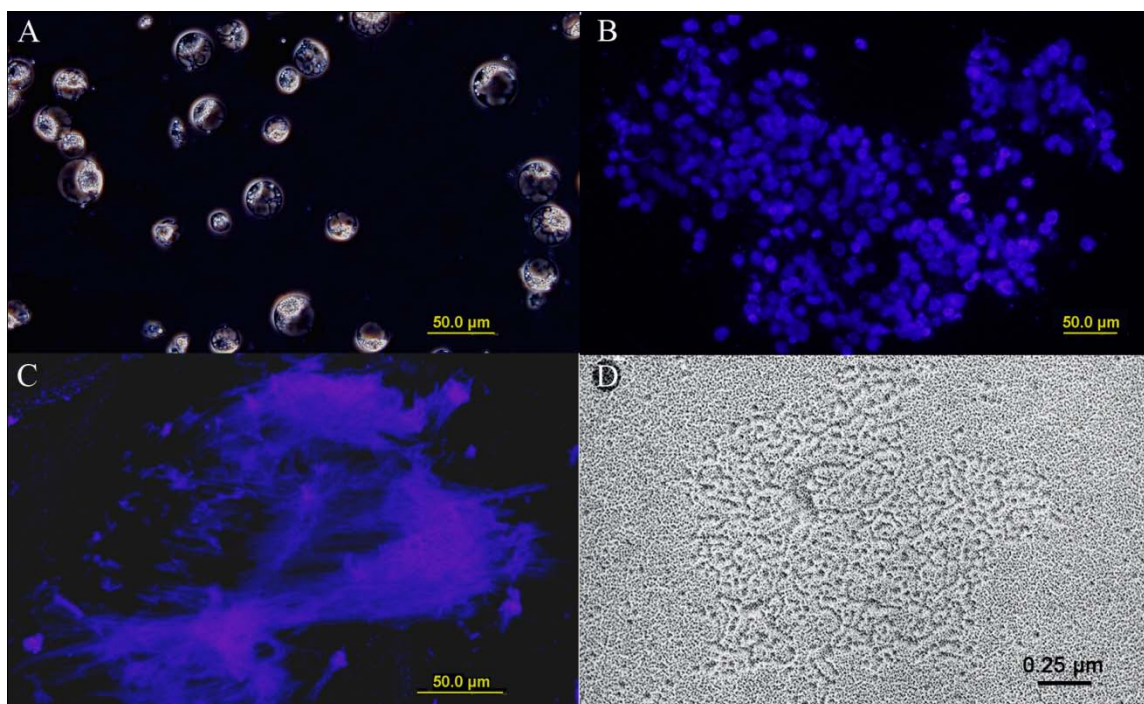


Figure 1 Microscopy images of isolated rice (*Oryza sativa*) suspension cell protoplasts, nuclei and chromatin

A) Image of rice protoplasts. Isolated rice protoplasts were diluted in protoplast suspension medium, plated on a microscope slide, and visualized using a phase-contrast microscopy. B) Image of purified rice nuclei after DAPI stain. Purified rice nuclei re-suspended in NIB buffer were examined using a fluorescence microscopy. C) Image of purified chromatin mass after DAPI stain. D) Electron microscopy image of purified chromatin. Transmission electron microscopy technique was used to examine the quality and detailed structure of the purified chromatin sample. The magnification is revealed by the scale bar.

### Western blots analysis

The establishment of the chromatin isolation protocol made it possible for us to investigate chromatin associated proteins. To test the quality and quantity of the isolated chromatin proteins, we used antibodies against histone H2B, H3K14AC, and H1 for Western immunoblotting, respectively. We found that H2B, H3, and H1 were highly enriched in the chromatin fraction compared to the total protein extract and nuclear protein extract as shown in Figure 3.2. Interestingly, two thick bands were detected when

antibodies against H3K14AC were performed. It was not clear whether the two bands were due to differential modification of H3 at other sites or cross reaction with a H3 variant(s) having the same K14 acetylation. We also examined COP9 complex in total protein, nuclear protein, and chromatin protein fractions respectively with antibodies against subunit 3 and subunit 6. COP9 signalosome is a nuclear enriched and high abundance protein complex involved in the regulation of protein degradation and plant development (Chamovitz et al., 1996; Peng et al., 2001b, a). Both subunit 3 and 6 were high abundant in the nuclear fraction when compared with the total protein fraction. However, they were absent in the chromatin fraction (Figure 3.2). The two bands detected with antibodies for COP9 subunit 6 (CSN6) were probably due to differential modification or alternative splicing because these two bands were missing simultaneously in CSN6 mutants (Peng et al., 2001b). We also examined  $\alpha$ -tubulin and actin in these three protein fractions. Both actin and  $\alpha$ -tubulin could be detected in all three fractions, however, the quantities in chromatin fraction were the lowest and in nucleus were the highest. Our rice suspension cells were grown in darkness except a brief exposure to light while adding enzymes for the removal of cell wall. It was known that Phytochrome A was mainly in cytoplasm in darkness. We investigated PhyA level in these three protein fractions. Our results indicated that PhyA was mainly detected in the total protein fraction (Figure 3.2). Although a trace amount was shown in nuclear fraction, none was detected in the chromatin fraction. Above Western immunoblotting results indicated that histone proteins were significantly enriched in chromatin fraction, and meanwhile, contaminations by other nuclear proteins such as COP9 complex subunit and cytosolic protein such as PhyA were not detectable in Western blots.

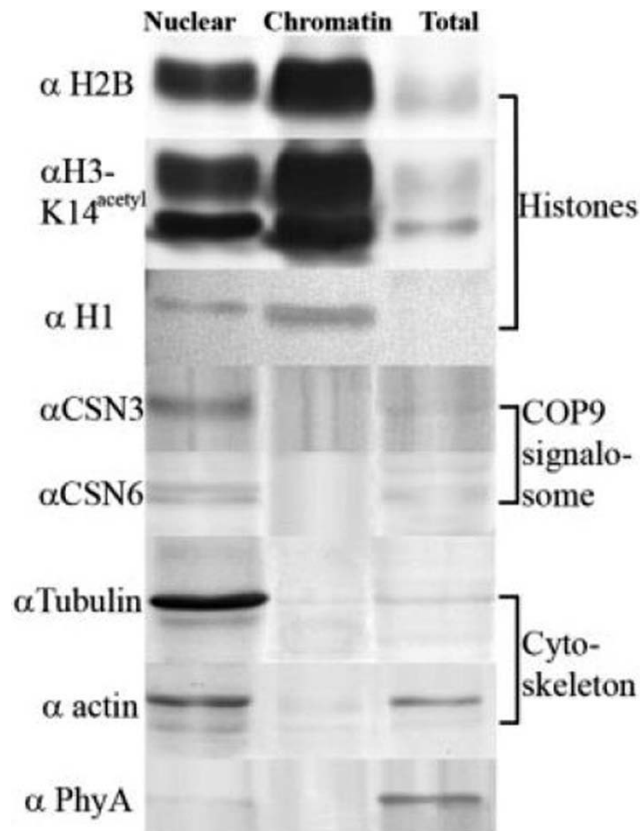


Figure 2 Immunological characterization of proteins in the chromatin fraction, nuclear fraction, and total protein fraction

20  $\mu$ g proteins were loaded on each lane. Western blots were carried out with standard procedures. The second antibodies were conjugated with alkaline phosphatase. Protein sources are indicated on the top. Antibodies are indicated on the left and the characteristics of the corresponding protein are indicated on the right.

## 2-DE gel, and mass spectrometric analysis of proteins co-purified with chromatin

To further investigate proteins associated with chromatin, the proteins were resolved on pH 3 to 11 2-DE gels. A representative gel image was shown in Figure 3.3 and the three biological replicas were shown in Figure 3.4. The separated proteins had a wide range of pIs and molecular weights. About 972 distinct protein spots were consistently resolved on 2-DE gels in all three biological replicas with SYPRO Ruby stain using PDQUEST automatic analyses and manual verification (Figure 3.4).

Among the 972 resolved protein spots, 607 prominent protein spots were excised using a Proteome Works Spot Cutter (Bio-Rad), in gel digested with a ProPrep (Genomic Solutions) robotic digester/spotter, and MS/MS analyzed with a MALDI TOF-TOF mass spectrometer (ABI 4700 Proteomics Analyzer, Applied Biosystems). Among the 607 excised protein spots, 509 proteins were identified with high Confidence Intervals (C.I.%>95%) as marked on the 2-DE gel picture (Figure 3.3) and listed in Table 3.1. We found that many distinct proteins spots had the same protein identities perhaps due to post translational modifications, alternative splicing or other reasons. For example, H2A was identified in 6 distinct protein spots, H2B in 4 distinct protein spots, H3 in 4 distinct protein spots, and H4 in 6 distinct protein spots. In addition, different variants of histones were identified within a same protein spot probably due to the overlapping. For example, spot 104 contained H3.3, H3.2, and H3-maize. Overall, the four core histone proteins have been identified over 60 times, which was in agreement with that histones were presented in many different modification forms and have multiple variants in the cell. Other representative chromatin associated proteins include e.g. putative DNA-directed RNA polymerase, transposon protein, retrotransposon protein, putative WRKY DNA-binding protein, putative transcriptional factor APF, RAN GTPase, etc. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was initially discovered as a glycolytic enzyme in cytoplasm. Later, it was found that GAPDH was both cytoplasmic and nuclear located (Mazzola and Sirover, 2002; Zheng et al., 2003). It acts as an essential component of a transcriptional activator complex regulating histone H2B expression. This protein was identified in three distinct protein spots in our chromatin preparation (spots 3105, 3501, and 4416). A large number of skeleton proteins were identified, including tubulin beta chain, tubulin 5 chain, tubulin R1623, alpha-tubulin, putative actin, and kinesin motor



domain containing protein. Most of these skeleton proteins had also been reported to be associated with human metaphase chromosome and *C. elegans* chromatin (Uchiyama et al., 2005; Chu et al., 2006). In addition, various proteins unrelated to chromatin were detected in the chromatin fraction and most of these proteins were also co-purified with human chromosome and *C. elegans* chromatin, respectively (Uchiyama et al., 2005; Chu et al., 2006), such as Tu translational elongation factor, heat shock proteins, chaperonins, DNA J homologue, RNA binding proteins, several ribosome subunits, putative U3 snoRNP protein IMP4, prohibitin, 26 S proteome regulatory subunits, etc. Many hypothetical and function-unknown proteins were discovered as shown in Table 3.1.

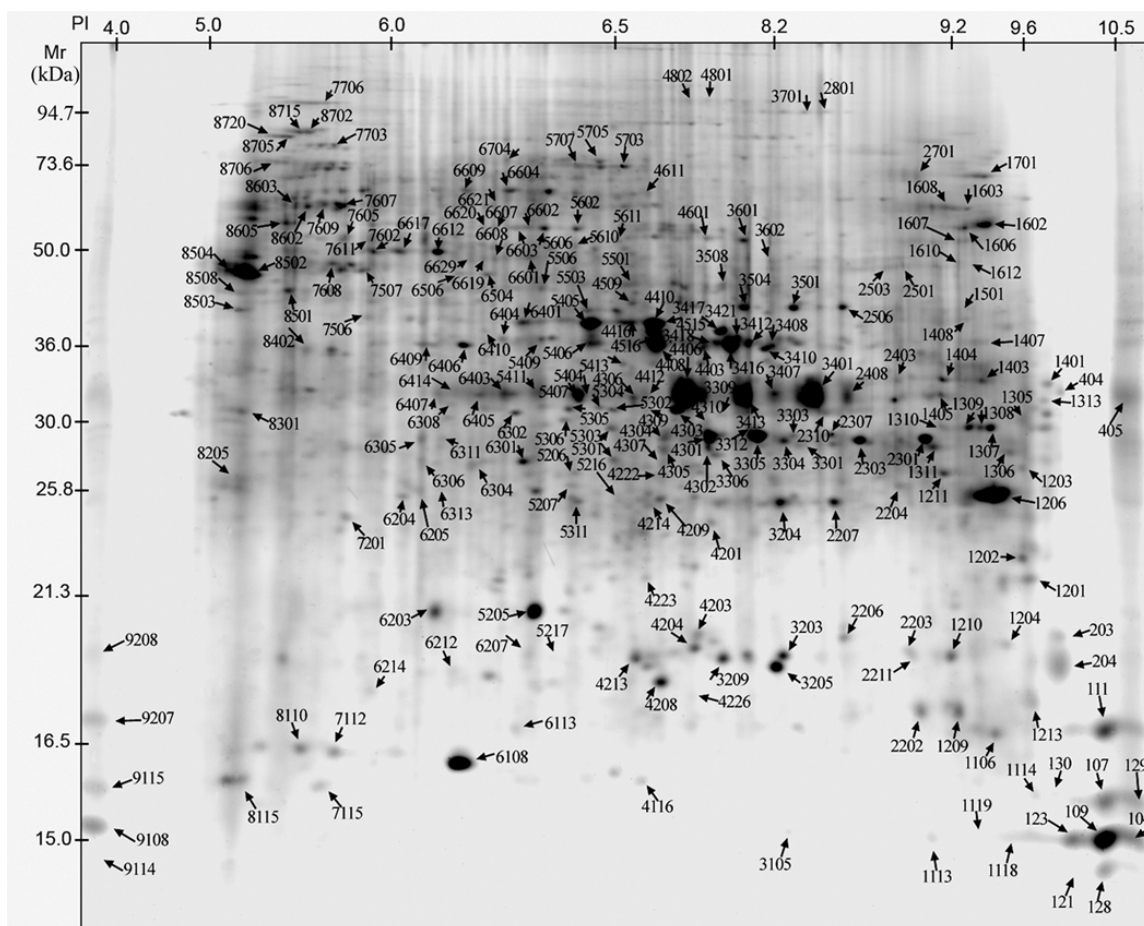


Figure 3 2-DE gel image of rice chromatin proteome revealed by SYPRO Ruby fluorescence stain

Proteins were extracted from purified chromatin of rice suspension cells, separated on 2-DE gels, and stained with SYPRO Ruby. Proteins identified with high confidence (C.I.%>95%) are marked with arrows. Molecular mass markers are on the left and the pH gradient of the first dimension is indicated on the top. The second dimension SDS PAGE was 12%.



### **Identification of chromatin co-purified proteins using shotgun approach**

Recent development in 2D-LC-MS/MS technology has made shotgun proteomics a powerful tool in protein identification using a protein mixture. We performed shotgun approach to study proteins co-purified with chromatin. After multiple runs, we obtained 292 unique proteins with 2 or more peptides as listed in Table 3.2 and another 367 annotations with a single peptide which are not presented in this dissertation due to the page limitation. Although only one peptide was identified in these proteins, the probability of random match to the corresponding protein was low. A striking feature was that a large number of histone variants in addition to the four common core histones were identified, including 11 H2A variants, 10 H2B variants, 2 H3 variants, histone H1 and a histone like protein. Other proteins being identified include high mobility group proteins (a single peptide identified), nucleosome assembly proteins, histone deacetylase HD2, transcription factors, DNA binding proteins, retrotransposon proteins, 26S proteasome regulatory subunits, heat shock proteins, RNA binding proteins, etc (Table 3.2). The reason that we could only detect limited number of nonhistone proteins was probably due to the high abundance of histones, which prevented the identification of relative low abundance proteins in the protein mixture because peptides were selected for MS/MS based on abundance and randomness. Many proteins, including H1, which were not detected using the 2-DE gel based method, were identified using shotgun method. Meanwhile, a large number of proteins identified using 2-DE gel based approach, including the abundant cationic peroxidase, were not detected using shotgun approach. A large number hypothetical proteins and function-unknown proteins have been discovered, which takes almost fifty percents of total proteins, suggesting that the chromatin associated proteins are poorly annotated and few studies have been investigated to explore

the chromatin. To estimate the error rate, the reversed sequence data base was searched. Seven proteins were found based on a single peptide match and no proteins were found based on two peptide matches (Table 3.3). Since the total peptides identified by shotgun approach were 1830, the error rate for peptide miss identification was 7/1830.

Histone post-translational modifications, such as methylation and acetylation, play a critical role in gene regulation. We carried out an *in silico* search for histone peptide with phosphorylation, acetylation or methylation using BioWorks 3.3. Probably because the mass spectrometer was optimized for protein identification instead of choosing peptides with putative modifications for MS/MS analysis, only about two dozens of peptides with putative modifications were identified. Most of the spectra of these peptides had noise probably due to the complexity of the protein mixture. After manual examination, we found that K@QLATK@AAR of H3.2 was probably acetylated at both Lysines (Figure 3.5A); YR^PGTVAL of H3.3 was probably acetylated at Arginine (Figure 3.5B). The corresponding mass spectra of the two peptides were shown in Figure 3.5. Further investigation of histone post-translational modifications was performed and discussed in chapter 5.

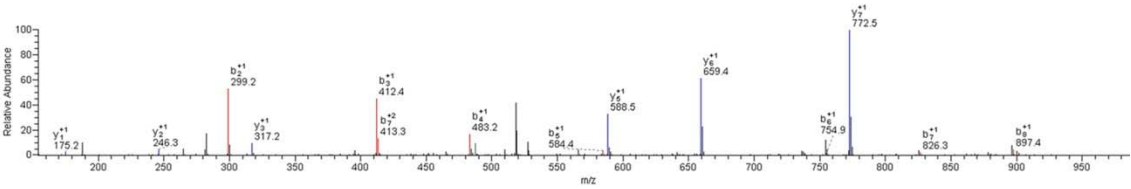
A. Histone H3.2 (Q2RAD9)

Peptide: K<sup>Q</sup>QLATK<sup>Q</sup>AAR

Modification : Acetylation (K<sup>Q</sup>+42.03670) (R<sup>Q</sup>+42.03670)

Precursor ion : 536.18 charge :2

#	AA	B	Y	#
1	K <sup>Q</sup>	171.22	-	9
2	Q	299.35	901.05	8
3	L	412.51	772.92	7
4	A	483.58	659.76	6
5	T	584.69	588.68	5
6	K <sup>Q</sup>	754.90	487.58	4
7	A	825.98	317.37	3
8	A	897.05	246.29	2
9	R	-	175.21	1



B. Histone H3.3 (Q0JCT1)

Peptide: YR<sup>Q</sup>PGTV<sup>Q</sup>ALR

Modification : Acetylation (K<sup>Q</sup>+42.03670) (R<sup>Q</sup>+42.03670)

Precursor ion : 538.60 charge :2

#	AA	B	Y	#
1	Y	164.18	-	9
2	R <sup>Q</sup>	362.40	912.07	8
3	P	459.52	713.85	7
4	G	516.57	616.73	6
5	T	617.68	559.68	5
6	V	716.81	458.58	4
7	A	787.89	359.45	3
8	L	901.04	288.37	2
9	R	-	175.21	1

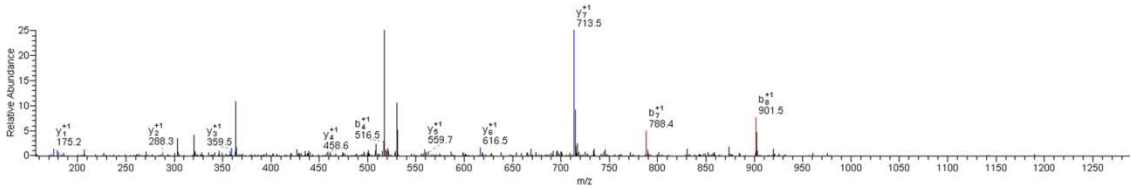


Figure 5 Histone H3 acetylation detected by LC-MS/MS

(A) Histone H3.2 peptide KQLATKAAR was acetylated at both lysine. (B) Histone H3.3 peptide YRPGTVAL was acetylated at arginine.

Table 1 Proteins co-purified with chromatin in rice (*Oryza sativa*) using MALDI-TOF

Location	Spot Number	Accession Number	Protein Name	%.I.	MW (kDa)	PI	Identified Peptides
nucleus	104	AAX92719	histone H3 - maize	100.0	15.3	11.3	7
	107	Q8S857	Putative histone H2A	100.0	14.6	10.4	4
	109	AAC78105	Histone H3.3	99.8	15.4	11.2	5
	111	Q9LGH8	Putative histone H2B	100.0	16.5	10.0	12
	121	EAZ40221	Histone H4	100.0	11.4	11.5	4
	123	BAC75621	putative histone H2A	100.0	13.9	10.2	4
	128	BAC57734	histone H4	100.0	11.4	11.5	8
	129	Q8S857	Putative histone H2A	100.0	14.6	10.4	3
	130	EAY85846	hypothetical protein Osl_007079	99.0	23.0	7.8	7
	204	Q9LGH6	Putative histone H2B	100.0	16.5	10.0	10
	404	Q7XUR6*	OSJNBa0084K11.15 protein	97.7	41.5	11.4	12
	1113	BAC57734	histone H4	100.0	11.4	11.5	7
	1114	BAC75621	putative histone H2A	98.7	13.9	10.2	3
	1118	BAC57734	histone H4	100.0	11.4	11.5	10
	1119	BAC57734	histone H4	100.0	11.4	11.5	5
	1203	BAD68174*	putative DNA-directed RNA polymerase II 23K chain	100.0	24.9	9.5	11
	1210	Q40674*	Cyclophilin 2 (EC 5.2.1.8) (Peptidyl-prolyl cis-trans isomerase) (PPlase) (Rotamase)	99.5	18.3	8.6	6
	1213	ABB48012*	AT-hook protein 1, putative, expressed	100.0	41.4	9.5	7
	1403	EAZ38957*	Glycine-rich RNA-binding protein-like	100.0	33.4	9.3	12
	1404	EAZ38957*	Glycine-rich RNA-binding protein-like	100.0	33.4	9.3	13
	1408	Q948S9*	DNAJ homologue	100.0	46.9	9.5	12
	1501	Q948S9*	DNAJ homologue	100.0	46.9	9.5	12
	1612	Q7XVC4*	OSJNBa0072D21.14 protein	100.0	56.1	9.3	16
	1701	EAY83860*	Os12g0611200	100.0	65.7	9.6	27
	2203	Q93WG4	Putative retroelement (Putative gag-pol polyprotein)	98.6	194.2	8.8	23
	2206	EAY80866	hypothetical protein Osl_034825	96.1	14.2	9.2	8
	2211	Q94JJ4	Putative histone H2B	100.0	16.5	10.0	7
	2310	DAA00397	TPA_exp: unknown	100.0	39.3	6.4	14
	2403	DAA00397	TPA_exp: unknown	100.0	39.3	6.4	12
	2701	EAZ02751	hypothetical protein Osl_023983	99.8	62.4	9.2	10
	3305	Q8W3N9	26S proteasome regulatory particle triple-A ATPase subunit3 (Fragment)	100.0	41.4	5.9	13
	3306	Q9XJ45	Ran (Small GTP-binding protein) (Ran2)	100.0	25.0	6.7	8
	3309	DAA00397	TPA_exp: unknown	100.0	39.3	6.4	12
	3401	Q94CF9*	RSSG8	99.7	129.6	8.6	20
	3508	Q94CQ1	Putative RNA-binding like protein	100.0	57.7	9.0	10
	3601	BAD09992*	hypothetical protein	99.7	36.6	11.7	11
	3602	Q9FXT8	26S proteasome regulatory particle triple-A ATPase subunit4	100.0	44.6	7.0	18
	4222	Q9XJ44	RAN (Putative GTP-binding protein) (Small GTP-binding protein) (Ran1)	100.0	25.0	6.4	9
	4226	Q7XPU6*	OSJNBa0088H09.10 protein	100.0	16.8	6.1	5
	4302	Q9XJ45	Ran (Small GTP-binding protein) (Ran2)	100.0	25.0	6.7	6
	4307	Q9XJ45	Ran (Small GTP-binding protein) (Ran2)	100.0	25.0	6.7	7
	4310	BAC16488	putative transcription factor APF1	100.0	29.6	6.4	5
	4611	DAA00397	TPA_exp: unknown	98.8	39.3	6.4	4
	5301	AAP52582*	transposon protein, putative, CACTA, En/Spm sub-class	99.8	88.6	8.4	19
	5406	BAC92643*	putative TGF-beta receptor-interacting protein	100.0	36.4	5.9	15
	5409	P49027*	Guanine nucleotide-binding protein beta subunit-like protein (GPB-LR) (RWD)	100.0	36.2	6.0	11
	5501	EAZ32941	hypothetical protein Osl_016424	98.0	39.3	6.4	6
	5506	EAY76891*	Cysteine endopeptidase	96.6	40.7	6.2	12
	5602	NP_001054148	OSJNBa0015K02.19 protein	100.0	42.8	5.2	10
	5606	EAY95913*	hypothetical protein Osl_017146	100.0	95.9	6.7	16
	5610	Q9FWK8*	Putative RNA binding protein	100.0	45.6	6.2	7
	5611	Q9FWK8*	Putative RNA binding protein	100.0	45.6	6.2	12
	5703	BAD29050*	ribosomal protein-like	100.0	51.5	6.2	18
	5705	DAA00398	TPA_exp: putative transposase	100.0	55.3	6.3	18
	6108	O22384*	Glycine-rich protein	100.0	15.9	7.8	8
	6207	Q7X8Q5*	OSJNBa0079A21.2 protein (OSJNBb0038F03.16 protein)	96.4	217.9	9.1	23
	6212	Q84Q77	17.4 kDa class I heat shock protein 3	100.0	17.9	5.8	7
	6404	CAD39822*	OSJNBa0079F16.13	94.7	55.2	9.1	14
	6405	Q8W403	Sec13p	99.9	33.3	5.6	5
	6410	ABA96737*	Retrotransposon protein, putative, unclassified	99.2	27.2	9.5	9
	6602	BAD81520	putative Y1 protein	98.4	50.1	5.9	5
	6617	EAY80530*	Cysteine endopeptidase	100.0	40.9	5.9	14
	6620	NP_001054148	OSJNBa0015K02.19 protein	100.0	42.8	5.2	9
	6621	Q8W3N9	26S proteasome regulatory particle triple-A ATPase subunit3 (Fragment)	100.0	41.4	5.9	18
	7115	O22384*	Glycine-rich protein	100.0	15.9	7.8	5

Table 1 (continued)

	7507	AAP52582*	transposon protein, putative, CACTA, En/Spm sub-class	98.7	88.6	8.4	16
	8110	O22385*	Glycine-rich protein	99.6	16.0	7.8	6
	8205	BAD67781	Putative WRKY DNA-binding protein	98.0	28.5	10.3	10
	8501	EAZ29051*	hypothetical protein Osl_012534	96.9	92.3	9.7	17
	8702	EAZ19310*	DnaK-type molecular chaperone hsp70-rice	100.0	69.5	5.3	15
	8705	Q40693*	Heat shock protein 70	100.0	71.0	5.2	19
	8720	Q40693*	Heat shock protein 70	100.0	71.0	5.2	13
	9108	AAC78105	Histone H3.3	100.0	15.4	11.2	5
	9114	BAC57734	histone H4	100.0	11.4	11.5	10
	9115	Q8S857	Putative histone H2A	100.0	14.6	10.4	4
	9207	AAX92719	histone H3 - maize	100.0	15.3	11.3	6
	9208	Q9LG12	Putative histone H2B	100.0	16.5	10.0	10
ribosome	130	BAF08760	Putative ribosomal protein S15	100.0	14.8	9.9	7
	203	EAZ37303	hypothetical protein Osl_020786	100.0	35.7	10.5	13
	1201	BAF28263	40S ribosomal protein S5, putative, expressed	100.0	22.2	9.7	10
	1202	AAP92747	ribosomal L9-like protein	100.0	21.3	9.6	12
	1204	Q7XNU2	OSJNBa0093F12.16 protein	100.0	17.7	9.2	6
	1407	Q9LIT4	EST AU069389(C61144) corresponds to a region of the predicted gene	100.0	34.1	9.6	13
	4223	EAZ00939	hypothetical protein Osl_022171	99.0	113.2	5.9	18
	5411	Q7XNU2	OSJNBa0093F12.16 protein	97.5	17.7	9.2	4
	6404	EAZ00939	hypothetical protein Osl_022171	95.7	113.2	5.9	18
	6602	Q7XNU2	OSJNBa0093F12.16 protein	100.0	17.7	9.2	6
cytoskeleton	1106	EAY88762	hypothetical protein Osl_009995	100.0	17.4	8.8	6
	1310	BAD37694	putative myosin heavy chain PCR43	99.5	173.7	8.8	22
	1405	EAZ09005	hypothetical protein Osl_030237	100.0	33.7	9.0	16
	6407	Q40665	Beta-tubulin	99.7	48.7	4.7	11
	6410	P37832	Tubulin beta chain	100.0	49.8	4.8	15
	6414	Q40665	Beta-tubulin	99.7	48.7	4.7	10
	6603	Q94DL4	Putative actin	98.7	41.7	5.2	6
	6617	EAY86520	hypothetical protein Osl_007753	100.0	63.6	5.8	14
	7506	EAY83721	Kinesin motor domain containing protein, expressed	99.6	313.6	5.0	36
	7602	EAY86520	hypothetical protein Osl_007753	100.0	63.6	5.8	15
	7608	EAY96234	hypothetical protein Osl_017467	100.0	57.7	4.9	10
	7611	JC2510	beta-tubulin R1623 - rice	100.0	50.1	4.8	13
	8605	BAF21989	Alpha-tubulin	100.0	49.6	4.9	12
plastid	1209	EAY86111*	hypothetical protein Osl_007344	99.8	59.8	9.1	7
	1608	NP_001051013*	Beta-glucosidase	100.0	56.8	9.1	18
	2202	NP_001051013*	Beta-glucosidase	100.0	56.8	9.1	7
	2506	EAY86569*	hypothetical protein Osl_007802	100.0	39.0	6.5	10
	3408	O22510*	Cationic peroxidase	95.3	38.3	8.9	3
	3410	O22510*	Cationic peroxidase	100.0	38.3	8.9	7
	3416	O22510*	Cationic peroxidase	100.0	38.3	8.9	13
	3417	O22510*	Cationic peroxidase	100.0	38.3	8.9	14
	3418	O22510*	Cationic peroxidase	100.0	38.3	8.9	12
	3421	O22510*	Cationic peroxidase	100.0	38.3	8.9	7
	3508	Q9ASH1*	P0686E09.9 protein	100.0	40.2	7.2	7
	3601	Q9ZWM0*	Plastidic ATP sulfurylase	100.0	52.3	9.0	23
	4306	O22510*	Cationic peroxidase	100.0	38.3	8.9	11
	4408	O22510*	Cationic peroxidase	100.0	38.3	8.9	10
	4410	O22510*	Cationic peroxidase	100.0	38.3	8.9	16
	4412	O22510*	Cationic peroxidase	100.0	38.3	8.9	8
	4416	O22510*	Cationic peroxidase	100.0	38.3	8.9	10
	4502	AAP06869	chloroplast import-associated channel protein homolog	100.0	136.8	8.8	19
	4516	O22510*	Cationic peroxidase	100.0	38.3	8.9	12
	4601	Q9ZWM0*	Plastidic ATP sulfurylase	100.0	52.3	9.0	27
	4801	AAP06869	chloroplast import-associated channel protein homolog	100.0	136.8	8.8	32
	5202	O22510*	Cationic peroxidase	100.0	38.3	8.9	4
	5305	O22510*	Cationic peroxidase	100.0	38.3	8.9	10
	5306	O22510*	Cationic peroxidase	100.0	38.3	8.9	5
	5404	O22510*	Cationic peroxidase	100.0	38.3	8.9	14
	5405	O22510*	Cationic peroxidase	100.0	38.3	8.9	12
	5406	O22510*	Cationic peroxidase	100.0	38.3	8.9	14
	5407	O22510*	Cationic peroxidase	100.0	38.3	8.9	13
	5503	O22510*	Cationic peroxidase	100.0	38.3	8.9	7
	6108	Q8VXC4*	Glycine rich RNA binding protein	100.0	19.5	6.6	10
	6204	Q949D1*	Hypothetical protein	100.0	26.4	7.6	6
	6207	Q7XSS2*	OSJNBa0041A02.10 protein	99.5	93.2	11.4	14
	6214	Q7X6C3*	OSJNBb0108J11.2 protein (OSJNBa0033G16.15 protein)	100.0	27.5	9.1	5
	6401	O22510*	cationic peroxidase	100.0	38.3	8.9	12
	6403	O22510*	Cationic peroxidase	100.0	38.3	8.9	9
	7115	Q8VXC4*	Glycine rich RNA binding protein	96.8	19.5	6.6	7
	8603	Q9ZWM0*	Plastidic ATP sulfurylase	100.0	52.3	9.0	13



Table 1 (continued)

mitochondrion	1307	EAZ25961	Porin-like protein	100.0	33.6	9.3	20
	1308	EAZ25961	Porin-like protein	100.0	33.6	9.3	15
	1309	EAY88938	Porin-like protein	100.0	33.6	9.3	11
	2301	Q8VXC7	Voltage-dependent anion channel	100.0	29.6	8.6	12
	2303	EAY98830	Voltage-dependent anion channel	100.0	27.8	7.9	10
	2307	BAB89921	putative porin	100.0	29.9	7.3	12
	3303	EAZ12509	Putative porin	100.0	29.9	7.3	9
	3304	NP_001056162	Voltage-dependent anion-selective channel (VDAC) protein	100.0	29.6	8.6	10
	3312	Q6K548	Outer mitochondrial membrane protein porin (Voltage-dependent anion-selective channel protein) (VDAC)	100.0	29.2	7.1	10
	4301	Q6K548	Outer mitochondrial membrane protein porin (Voltage-dependent anion-selective channel protein) (VDAC)	100.0	29.2	7.1	6
	4304	EAZ08792	Outer mitochondrial membrane protein porin	100.0	29.2	6.7	12
	4305	BAF24897	Os09g0361400	100.0	29.2	7.1	11
	5304	EAZ13326	Putative 36kDa porin II	100.0	41.5	5.7	11
	6604	P15998	ATP synthase alpha chain, mitochondrial (EC 3.6.3.14)	100.0	55.2	5.9	16
	6609	P15998	ATP synthase alpha chain, mitochondrial (EC 3.6.3.14)	100.0	55.2	5.9	18
	7201	EAZ26040*	hypothetical protein OsJ_009523	97.9	53.7	5.2	9
	7607	Q01859	ATP synthase beta chain, mitochondrial precursor (EC 3.6.3.14)	100.0	59.0	6.3	18
	7609	Q01859	ATP synthase beta chain, mitochondrial precursor (EC 3.6.3.14)	100.0	59.0	6.3	16
	8301	Q01859	ATP synthase beta chain, mitochondrial precursor (EC 3.6.3.14)	100.0	59.0	6.3	5
	8602	Q01859	ATP synthase beta chain, mitochondrial precursor (EC 3.6.3.14)	100.0	59.0	6.3	21
	8603	Q01859	ATP synthase beta chain, mitochondrial precursor (EC 3.6.3.14)	99.8	59.0	6.3	7
membrane	404	EAZ39341	putative prohibitin	100.0	39.0	9.8	15
	1307	ABF94514	Eukaryotic porin family protein, expressed	100.0	26.7	8.9	14
	1308	ABF94514	Eukaryotic porin family protein, expressed	100.0	26.7	8.9	11
	1313	EAZ39341	putative prohibitin	100.0	39.0	9.8	15
	1603	EAY98386	Hypothetical protein P0015C02.2	100.0	56.5	5.7	15
	3407	BAB85263*	putative H <sup>+</sup> -exporting ATPase	100.0	26.6	6.9	9
	3421	Q40648	Potassium channel beta subunit protein	100.0	36.1	8.8	6
	4213	EAZ25964	hypothetical protein OsJ_009447	99.7	66.4	7.8	16
	4303	EAY94429	B0812A04.3 protein	100.0	30.6	6.6	16
	5302	EAY94429	B0812A04.3 protein	100.0	30.6	6.6	11
	5303	EAY94429	B0812A04.3 protein	100.0	30.6	6.6	15
	5303	CAE76006	B1358B12.15	100.0	30.8	7.0	14
	5411	Q7X863	OSJNBa0016N04.2 protein (OSJNBa0049H08.21 protein)	99.4	16.4	11.2	9
	6302	EAZ23571	Putative prohibitin	100.0	30.2	5.9	11
	6403	BAB89823	embryonic abundant protein-like	100.0	29.1	5.7	10
	7607	BAF05814	Os01g0685800	100.0	59.7	5.9	17
	7609	BAF05814	Os01g0685800	100.0	59.7	5.9	15
	7703	BAD45853	Putative vacuolar proton-ATPase	100.0	68.4	5.2	16
	8301	BAF05814	Os01g0685800	100.0	59.7	5.9	7
	8602	BAF05814	Os01g0685800	100.0	59.7	5.9	21
	8603	BAF05814	Os01g0685800	99.7	59.7	5.9	7
	8706	Q7XQN5	OSJNBa0089K21.1 protein	97.4	27.4	8.9	10
cytoplasm	1106	EAZ25787	Elongation factor 1-alpha	100.0	47.9	9.2	9
	1210	O04985*	Non-symbiotic hemoglobin 2 (rHb2) (ORYsa GLB1b)	100.0	18.6	9.0	7
	1211	BAC81178	unknown protein	100.0	58.7	8.7	13
	1311	O64937	Elongation factor 1-alpha (EF-1-alpha)	97.9	49.2	9.1	7
	1314	O64937	Elongation factor 1-alpha (EF-1-alpha)	100.0	49.2	9.1	4
	1405	BAD26449*	putative annexin	100.0	35.6	8.9	21
	1602	O64937	Elongation factor 1-alpha (EF-1-alpha)	100.0	49.2	9.1	12
	1606	EAZ25788	Elongation factor 2-alpha	100.0	49.3	9.1	13
	1607	O64937	Elongation factor 1-alpha (EF-1-alpha)	100.0	49.2	9.1	9
	1610	O64937	Elongation factor 1-alpha (EF-1-alpha)	100.0	49.2	9.1	8
	2203	O04985*	Non-symbiotic hemoglobin 2 (rHb2) (ORYsa GLB1b)	100.0	18.6	9.0	8
	2204	BAC81178	unknown protein	100.0	58.7	8.7	13
	2501	O64937	Elongation factor 1-alpha (EF-1-alpha)	94.9	49.2	9.1	9
	2506	A1YR13*	glyceralde-3-phosphate dehydrogenase	100.0	36.5	7.7	12
	2801	Q8W315*	Putative GTP-binding protein	99.9	68.0	8.4	16
	3105	Q42977*	Glyceraldehyde 3-phosphate dehydrogenase, cytosolic (EC 1.2.1.12)	95.5	36.5	6.6	2
	3203	EAZ06044	hypothetical protein Osl_027276	100.0	38.5	5.9	11
	3205	EAZ06044	hypothetical protein Osl_027276	100.0	38.5	5.9	12
	3209	O04986*	Non-symbiotic hemoglobin 1 (rHb1) (ORYsa GLB1a)	100.0	18.4	6.9	7
	3501	A1YR13*	glyceralde-3-phosphate dehydrogenase	100.0	36.5	7.7	17
	3504	AAP54418	N-acetyl-gamma-glutamyl-phosphate reductase, chloroplast precursor, putative, expressed	100.0	45.4	8.5	5
	3701	EAZ13100*	hypothetical protein OsJ_002925	100.0	72.6	8.2	7
	4203	P31673*	17.4 kDa class I heat shock protein 1	100.0	17.4	6.2	5

Table 1 (continued)

	4204	P31673*	17.4 kDa class I heat shock protein 1	100.0	17.4	6.2	7
	4208	EAY96193	hypothetical protein Osl_017426	100.0	101.0	5.7	8
	4209	EAY77115	hypothetical protein Osl_004962	98.8	25.3	8.8	5
	4416	Q42977*	Glyceraldehyde 3-phosphate dehydrogenase, cytosolic (EC 1.2.1.12)	100.0	36.5	6.6	3
	5205	EAZ06044	hypothetical protein Osl_027276	100.0	38.5	5.9	13
	5217	P31673*	17.4 kDa class I heat shock protein 1	100.0	17.4	6.2	7
	6203	EAZ06044	hypothetical protein Osl_027276	100.0	38.5	5.9	13
	6205	Q7XUS2*	OSJNBa0060P14.1 protein	100.0	30.7	6.4	7
	6212	P31673*	17.4 kDa class I heat shock protein 1	100.0	17.4	6.2	7
	6305	BAF13000*	Actin-1, putative, expressed	96.6	41.8	5.3	5
	6410	BAD81440*	Purine rich element binding protein B-like	100.0	33.3	5.7	13
	6504	Q93Y73	Putative dehydrogenase	100.0	40.2	6.7	13
	6506	Q93Y73	Putative dehydrogenase	100.0	40.2	6.7	12
	8115	BAD32133*	putative receptor-like protein kinase 4	96.6	72.7	6.3	16
endoplasmic reticulum	7706	O24182	Endosperm luminal binding protein	100.0	73.5	5.3	10
	8702	O24182	endosperm luminal binding protein	100.0	73.5	5.3	21
	8705	O24182	Endosperm luminal binding protein	100.0	73.5	5.3	22
	8715	O24182	Endosperm luminal binding protein	100.0	73.5	5.3	18
	8720	O24182	Endosperm luminal binding protein	100.0	73.5	5.3	18
intracellular	1309	BAF14566	OSJNBa0072K14.5 protein	100.0	48.3	8.7	14
	1310	BAF14566	OSJNBa0072K14.5 protein	99.8	48.3	8.7	9
	1401	BAC99389	putative U3 snoRNP protein IMP4	100.0	28.9	10.1	18
	1401	BAC99390	putative U3 snoRNP protein IMP4	100.0	33.9	9.6	20
	6506	Q8W2C4	Translational elongation factor Tu	100.0	48.4	6.0	12
	6617	AAF15312	Chloroplast translational elongation factor Tu	100.0	50.3	6.1	12
	6619	AAF15312	Chloroplast translational elongation factor Tu	98.1	50.3	6.1	2
	6619	Q8W2C4	Translational elongation factor Tu	100.0	48.4	6.0	21
	6619	Q8W2C3	Translational elongation factor Tu	98.1	50.4	6.2	1
	6629	Q9AY71	Putative GTP-binding protein	100.0	46.7	7.0	12
	6629	Q8W2C4	Translational elongation factor Tu	100.0	48.4	6.0	12
	7602	AAF15312	chloroplast translational elongation factor Tu	100.0	50.3	6.1	12
	7602	EAZ23662	Translational elongation factor Tu	100.0	50.4	6.2	12
	7605	BAF14566	OSJNBa0072K14.5 protein	100.0	48.3	8.7	12
cytosol	4509	Q8W424	26S proteasome regulatory particle non-ATPase subunit8	100.0	34.9	6.3	17
	5413	Q9SDD1	ESTs D39011(R0609) (26S proteasome regulatory particle non-ATPase subunit11)	100.0	34.3	6.1	16
extracellular region	2207	P93442	Expansin-A4 precursor (OsEXPA4) (Alpha-expansin-4) (OsEXP4) (OsaEXP4.22)	96.6	25.9	8.1	5
unknown	405	EAZ05518	hypothetical protein Osl_026750	100.0	29.8	8.2	9
	1118	BAF06011	Os01g0721800	98.6	58.4	8.8	14
	1206	EAZ29098	hypothetical protein OsJ_012581	100.0	36.0	9.3	8
	1209	ABF98424	Glycosyl hydrolase family 1 protein, expressed	99.8	46.0	9.0	6
	1210	BAF05781	Os01g0679600	100.0	21.3	6.8	4
	1305	BAF14824	Peptidyl-prolyl cis-trans isomerase	100.0	93.7	5.8	18
	1306	EAZ03635	hypothetical protein Osl_024867	100.0	31.8	9.5	11
	1608	ABF98424	Glycosyl hydrolase family 1 protein, expressed	100.0	46.0	9.0	15
	2408	EAZ41442	hypothetical protein OsJ_024925	100.0	29.7	8.2	3
	2503	Q8W2F8	3-ketoacyl-CoA thiolase (Fragment)	100.0	18.6	8.6	7
	3204	JC7138	alpha-amylase (EC 3.2.1.1) isozyme III - rice	100.0	48.7	5.3	6
	3301	BAB67891	Putative thaumatin-like cytokinin-binding protein	100.0	26.2	7.9	9
	3303	BAB67891	Putative thaumatin-like cytokinin-binding protein	100.0	26.2	7.9	7
	3304	BAB67891	Putative thaumatin-like cytokinin-binding protein	100.0	26.2	7.9	11
	3401	EAZ41445	hypothetical protein OsJ_024928	100.0	29.8	8.2	13
	3408	CAJ86336	H0814G11.3	99.0	33.5	5.2	4
	3410	CAJ86336	H0814G11.3	100.0	33.5	5.2	9
	3412	EAZ05175	hypothetical protein Osl_026407	100.0	38.6	7.3	7
	3413	EAZ05518	hypothetical protein Osl_026750	99.9	29.8	8.2	8
	3416	CAJ86336	H0814G11.3	100.0	33.5	5.2	15
	3417	CAJ86336	H0814G11.3	100.0	33.5	5.2	17
	3418	CAJ86336	H0814G11.3	100.0	33.5	5.2	13
	3421	CAJ86336	H0814G11.3	100.0	33.5	5.2	8
	3501	AAN59792	cytosolic glyceraldehyde-3-phosphate dehydrogenase GAPDH	100.0	23.4	7.9	9
	4201	EAZ30021	hypothetical protein OsJ_013504	100.0	32.2	7.8	8
	4208	AAG03091	unknown protein	100.0	14.7	6.8	6
	4209	BAC15855	putative dimethylaniline monooxygenase	95.4	53.0	5.5	12
	4214	BAB93247	putative glutathione transferase III(b)	100.0	24.1	6.2	4
	4306	CAJ86336	H0814G11.3	100.0	33.5	5.2	12
	4309	CAJ86336	H0814G11.3	98.6	33.5	5.2	5
	4403	CAJ86336	H0814G11.3	100.0	33.5	5.2	14
	4406	EAZ41442	hypothetical protein OsJ_024925	95.6	30.0	8.2	5
	4408	CAH69301	TPA: class III peroxidase 59 precursor	100.0	37.0	5.6	17
	4410	CAJ86336	H0814G11.3	100.0	33.5	5.2	18

Table 1 (continued)

4412	Q7XTJ4	OSJNBa0020P07.10 protein	100.0	31.9	9.6	12
4416	CAJ86336	H0814G11.3	100.0	33.5	5.2	10
4515	EAY89848	hypothetical protein Osl_011081	100.0	41.7	6.3	22
4516	CAJ86336	H0814G11.3	100.0	33.5	5.2	13
4601	Q7XQC6	OSJNBb0060M15.5 protein	99.8	39.4	6.1	4
5202	CAJ86336	H0814G11.3	100.0	33.5	5.2	5
5206	Q7XKE6	OSJNBb0017I01.8 protein	100.0	27.5	6.3	9
5207	BAD61316	putative glutathione transferase F4	100.0	25.3	5.7	9
5216	BAB93247	putative glutathione transferase III(b)	100.0	24.1	6.2	8
5305	CAJ86336	H0814G11.3	100.0	33.5	5.2	11
5306	CAJ86336	H0814G11.3	100.0	33.5	5.2	6
5404	CAJ86336	H0814G11.3	100.0	33.5	5.2	14
5405	CAJ86336	H0814G11.3	100.0	33.5	5.2	13
5406	CAJ86336	H0814G11.3	100.0	33.5	5.2	16
5406	BAF16251	Peroxidase	100.0	37.0	5.6	16
5407	CAJ86336	H0814G11.3	100.0	33.5	5.2	13
5503	CAJ86336	H0814G11.3	100.0	33.5	5.2	8
5707	Q7XMP6	OSJNBb0059K02.15 protein	100.0	63.9	6.8	15
6113	BAC22221	unknown protein	100.0	18.2	9.9	7
6212	Q8SA79	Heat shock-like protein (Fragment)	100.0	6.9	9.8	3
6301	BAD61316	putative glutathione transferase F4	100.0	25.3	5.7	12
6304	BAD61316	putative glutathione transferase F4	100.0	25.3	5.7	10
6306	Q94GR4	Putative LN1 protein	99.7	31.7	5.7	10
6308	Q7XKT9	OSJNBa0022H21.18 protein	100.0	32.1	5.5	6
6311	BAC15855	putative dimethylaniline monooxygenase	97.4	53.0	5.5	12
6313	AAP54754	glutathione S-transferase GSTU6, putative, expressed	96.3	26.2	5.5	3
6401	NP_001054337	Peroxidase	100.0	37.0	5.6	15
6403	CAJ86336	H0814G11.3	100.0	33.5	5.2	11
6405	CAJ86336	H0814G11.3	100.0	33.5	5.2	3
6406	Q8VYH7	Isoflavone reductase-like protein	100.0	33.5	5.7	16
6407	CAJ86336	H0814G11.3	98.8	33.5	5.2	6
6409	Q7XKT9	OSJNBa0022H21.18 protein	100.0	32.1	5.5	13
6410	CAJ86336	H0814G11.3	100.0	33.5	5.2	9
6601	Q7XK22	OSJNBa0044K18.23 protein	100.0	40.1	7.6	12
6607	BAB92156	putative S-adenosyl-L-methionine synthetase	100.0	17.8	4.8	3
6607	P93438	S-adenosylmethionine synthetase 2 (EC 2.5.1.6) (Methionine adenosyltransferase 2) (AdoMet synthetase 2)	100.0	42.9	5.7	12
6609	BAB92682	putative selenium binding protein	100.0	53.1	5.7	10
6612	BAC15855	Putative dimethylaniline monooxygenase	95.6	53.0	5.5	13
6704	Q7XMP6	OSJNBb0059K02.15 protein	100.0	63.9	6.8	16
7112	BAC15855	Putative dimethylaniline monooxygenase	99.6	53.0	5.5	14
7507	BAC99688	putative disease resistance protein	95.4	109.7	6.2	18
8402	Q7G765	Probable NAD(P)H-dependent oxidoreductase 2	100.0	35.8	5.4	11
8502	JC7138	alpha-amylase (EC 3.2.1.1) isozyme III - rice	100.0	48.7	5.3	9
8503	Q8S7Z5	Putative 10-deacetylbaecatin III-10-O-acetyl transferase	99.1	47.0	5.7	12
8504	JC7138	alpha-amylase (EC 3.2.1.1) isozyme III - rice	100.0	48.7	5.3	9
8508	JC7138	alpha-amylase (EC 3.2.1.1) isozyme III - rice	98.8	48.7	5.3	7

a) Proteins were separated by 2-DE gels and followed with MS and MS/MS analysis with an ABI 4700 Proteomics Analyzer (Applied Biosystems) MALDI TOF-TOF mass spectrometer.

b) Spot Number: The spot number was given by computer based on spot excision order.

c) Accession Number: Protein accession number

d) MW/PI: predicted molecular mass and pI.

e) Identified Peptides: Number of peptides that matched with the identified protein in mass analyses.

f) C.I. %: Cross Confidence Interval %. Over 95% represents high confidence identification.

g) \*: Proteins without GO annotation. The cellular localization was predicted by LOCTree or PSORT.

Table 2 Proteins co-purified with chromatin identified using shotgun approach

Protein Accession	Protein Name	Protein Probability	pI	MW (KDa)	Identified Peptides No.
Q84P96	3-ketoacyl-CoA thiolase-like protein	6.40E-09	8.30	46.83	11
A2XIT5	60S ribosomal protein L13	1.45E-05	10.50	24.01	3
Q10HT7	60S ribosomal protein L13-2, putative, expressed	1.45E-05	10.50	22.54	3
A1XFD1	AF-4 domain containing protein-like protein	3.02E-05	6.00	165.60	5
JC7138	alpha-amylase isozyme III	5.29E-07	4.25	48.74	8
P27934	Alpha-amylase isozyme 3E precursor	5.29E-07	4.25	48.71	8
Q42975	Beta-glucosidase	3.77E-15	10.07	56.90	4
Q75I92	Beta-glucosidase	3.77E-15	9.60	15.48	2
Q75I93	Beta-glucosidase	3.77E-15	10.07	56.87	5
Q5Z5T9	BKRF1 encodes EBNA-1 protein-like	1.25E-03	9.00	23.08	2
Q762A2	BRI1-KD interacting protein 112	2.89E-06	4.30	30.40	2
Q6K624	BRI1-KD interacting protein 135	6.56E-09	4.30	81.36	4
Q761Y0	BRI1-KD interacting protein 135	6.56E-09	10.50	52.10	4
Q9ST80	CAA303717.1 protein	5.53E-05	6.00	36.77	2
O22510	Cationic peroxidase	5.53E-05	8.30	38.32	2
Q5U1N4	Class III peroxidase 59 precursor	5.53E-05	5.10	36.99	2
Q7F3A8	Cysteine endopeptidase	1.72E-09	6.00	40.72	12
Q9SXM1	Cysteine endopeptidase	3.63E-07	6.00	40.72	4
Q9SYT5	Cysteine endopeptidase precursor	3.63E-07	6.00	40.20	3
O24190	Cysteine proteinase precursor	3.63E-07	6.00	40.18	3
Q0ILZ4	DEAD-box ATP-dependent RNA helicase 9	7.98E-07	10.10	65.59	6
O22523	DNA-binding protein GBP16	2.88E-06	6.00	43.15	3
A3AEP8	Elongation factor 1-alpha	3.36E-09	10.10	47.90	17
O64937	Elongation factor 1-alpha	3.36E-09	10.10	49.29	19
Q10QZ6	Elongation factor 1-alpha, putative, expressed	3.36E-09	10.10	49.29	19
A2X6R0	Elongation factor Tu	2.55E-05	6.00	63.68	2
A2XP46	Elongation factor Tu	2.78E-06	6.00	48.42	3
Q10AI6	Elongation factor Tu, mitochondrial, putative, expressed	2.55E-05	6.00	33.88	2
A2Y5R6	Expansin-A4 precursor	8.96E-05	8.33	25.88	3
Q0DHB7	Expansin-A4 precursor	8.96E-05	8.33	25.88	3
Q852A1	Expansin-A7 precursor	3.78E-09	10.07	28.22	5
Q94LR4	Expansin-B4 precursor	2.06E-04	6.00	31.36	4
Q10H93	Expressed protein	9.99E-07	4.30	77.89	3
A1YR13	Glyceralde-3-phosphate dehydrogenase	3.80E-06	8.30	36.57	3
A2YQT7	Glyceraldehyde-3-phosphate dehydrogenase, cytosolic	3.80E-06	6.00	36.41	2
Q0J8A4	Glyceraldehyde-3-phosphate dehydrogenase, cytosolic	3.80E-06	6.00	36.41	2
ABG22105	Glycine-rich RNA-binding protein GRP1A	5.68E-05	4.25	11.98	3
Q2QLR2	Glycine-rich RNA-binding protein GRP1A,	5.68E-05	6.29	16.08	3
Q6ASX7	Glycine-rich RNA-binding protein, putative	1.06E-03	6.29	15.90	3
Q10EB0	Glycosyl hydrolase family 1 protein, expressed	3.77E-15	10.07	33.00	4
Q10EB1	Glycosyl hydrolase family 1 protein, expressed	3.77E-15	10.07	43.29	4
Q10EB2	Glycosyl hydrolase family 1 protein, expressed	8.95E-07	10.07	46.05	3
Q01JL0	H0112G12.10 protein	6.95E-08	5.10	40.26	17
CAD79700	H0302E05.3	9.61E-04	6.00	36.77	2
Q01I88	H0311C03.6 protein	4.09E-09	6.00	92.50	2
Q259L9	H0701F11.10 protein	5.53E-05	4.30	33.48	2
Q01LB1	H0718E12.4 protein	8.35E-03	10.07	48.28	2
A2YWQ1	Heat shock protein 81-1	3.15E-09	4.30	80.19	4
Q0J4P2	Heat shock protein 81-1	3.15E-09	4.30	80.19	4
Q69QQ6	Heat shock protein 81-2	3.15E-09	4.30	80.20	4
Q07078	Heat shock protein 81-3	3.15E-09	4.30	80.18	4
A2XCU2	Histone H2A	2.40E-05	10.50	14.56	6
A2XLW3	Histone H2A	7.86E-05	4.30	36.36	3
A2YNE9	Histone H2A	1.58E-07	10.50	19.99	5
A2Z7B5	Histone H2A	2.40E-05	10.50	14.61	8
A3AMK7	Histone H2A	7.86E-05	4.30	42.97	3

Table 2 (continued)

A3C4S7	Histone H2A	2.40E-05	10.50	14.61	8
A3CGY8	Histone H2A	2.33E-05	10.50	16.70	2
A3CI01	Histone H2A	5.31E-07	12.50	12.13	10
Q10D30	Histone H2A.Z, putative, expressed	2.40E-05	10.50	14.46	6
Q10RE1	Histone H2A.Z, putative, expressed	2.40E-05	10.50	14.56	6
Q0JMH6	Histone H2B	4.45E-06	10.50	11.34	8
A3AGM4	Histone H2B.1	4.45E-06	10.50	16.51	8
Q9LGI2	Histone H2B.10	4.45E-06	10.50	16.55	8
Q943L2	Histone H2B.11	4.45E-06	10.50	15.37	8
Q6ZBP3	Histone H2B.2	4.45E-06	10.50	16.33	8
Q94JJ7	Histone H2B.3	4.45E-06	10.50	16.53	8
Q94JJ4	Histone H2B.4	4.45E-06	10.50	16.50	8
Q94JE1	Histone H2B.5	4.45E-06	10.50	16.76	8
Q9LGH4	Histone H2B.6	4.45E-06	10.50	16.47	8
Q9LGH8	Histone H2B.8	4.45E-06	10.50	16.49	8
Q6F362	Histone H2B.9	4.45E-06	10.50	16.26	8
A2XU26	Histone H3	7.55E-15	10.53	22.89	5
A2Y8W0	Histone H3	7.30E-12	10.50	41.50	6
A2Y9L8	Histone H3	4.47E-08	11.50	16.76	3
A3AIP0	Histone H3	7.30E-12	10.50	17.10	15
A3AUC6	Histone H3	7.30E-12	12.50	15.41	15
A3AUC6	Histone H3	7.30E-12	12.50	15.41	15
A3B4F3	Histone H3	2.60E-03	12.50	19.63	6
A3C8S4	Histone H3	4.47E-08	10.50	16.53	8
Q2QSX0	Histone H3, putative	1.50E-04	10.53	15.48	5
Q2QSW7	Histone H3, putative, expressed	1.50E-04	11.51	15.43	5
A2Y533	Histone H3.2	4.47E-08	12.50	15.27	5
P69247	Histone H3.2	4.47E-08	12.50	15.27	12
Q2RAD9	Histone H3.2	4.47E-08	12.50	15.27	7
A2XJH3	Histone H3.3	7.30E-12	12.50	15.41	6
Q0JCT1	Histone H3.3	7.30E-12	12.50	15.41	6
Q71U98	Histone H3.3	7.30E-12	12.50	15.41	15
A2WWR4	Histone H4	1.79E-08	12.50	11.41	7
A2Y5H8	Histone H4	2.26E-04	10.50	9.90	2
A3B4U7	Histone H4	2.26E-04	11.50	11.50	2
A3B4W3	Histone H4	2.32E-05	12.50	9.74	15
A3BZ47	Histone H4	1.79E-08	12.50	30.93	22
Q851P9	Histone-like protein	7.18E-05	10.50	29.63	6
Q94D20	Nucleoid DNA-binding protein cnd41-like	2.50E-05	8.33	51.14	5
Q70Z21	Nucleosome assembly protein 1-like protein 1	4.38E-06	4.30	42.61	2
Q7FAH2	OJ000223_09.15 protein	3.80E-06	6.00	36.77	4
Q0JIM0	Os01g0791600 protein	2.77E-06	8.30	22.74	3
Q0E0X4	Os02g0514700 protein	8.35E-03	8.33	54.41	2
Q0E0V5	Os02g0519300 protein	2.54E-04	4.30	29.50	6
Q0DZV9	Os02g0595700 protein	2.55E-05	4.30	24.12	2
Q0DXF4	Os02g0756800 protein	5.73E-06	8.33	34.18	2
Q0DWC1	Os02g0821800 protein	3.25E-09	8.70	9.53	5
Q0DM93	Os03g0822000 protein	3.78E-09	10.07	28.22	5
Q0JD78	Os04g0423200 protein	6.58E-03	5.10	14.50	2
Q0JBZ3	Os04g0501600 protein	1.82E-09	6.00	92.51	4
Q0J9B4	Os04g0662800 protein	8.21E-04	8.30	58.41	2
Q0DL24	Os05g0128000 protein	2.86E-04	6.00	36.95	2
Q0DFD6	Os05g0597100 protein	1.02E-06	4.30	32.55	5
Q0DB64	Os06g0598500 protein	2.77E-06	6.00	47.99	3
Q0D5C7	Os07g0568700 protein	7.91E-07	7.17	35.46	2
Q0D3F6	Os07g0688700 protein	4.62E-10	11.51	28.33	9
Q0J528	Os08g0473600 protein	5.29E-07	4.25	48.71	8
Q0J2B8	Os09g0361400 protein	4.40E-04	8.04	29.22	2

Table 2 (continued)

Q0J0V1	Os09g0482100 protein	3.15E-09	4.30	80.20	4
Q0J0U8	Os09g0482400 protein	3.15E-09	4.30	80.18	4
Q0J0U7	Os09g0482600 protein	3.15E-09	4.30	64.50	4
Q0IY79	Os10g0356000 protein	2.77E-06	6.00	38.48	3
Q0IVG1	Os10g0572900 protein	3.60E-05	10.10	26.16	3
Q0IPF7	Os12g0207600 protein	2.77E-06	10.10	59.03	3
Q01M91	OSIGBa0109M01.2 protein	1.55E-04	6.00	28.14	2
Q7XQZ5	OSJNBa0015K02.19 protein	6.95E-08	5.10	40.26	17
Q7XU87	OSJNBa0029H02.25 protein	1.82E-09	6.00	92.51	4
Q7XSV2	OSJNBa0039K24.4 protein	5.53E-05	5.10	36.99	2
Q7FAE2	OSJNBa0052P16.8 protein	1.55E-04	6.00	28.10	2
Q7XVM2	OSJNBa0072K14.5 protein	8.35E-03	10.07	48.28	2
Q7XM19	OSJNBa0084K01.12 protein	8.21E-04	8.30	55.55	2
Q7XLZ6	OSJNBa0086O06.13 protein	3.67E-08	10.10	38.24	5
Q7XUC9	OSJNBa0088A01.17 protein	1.79E-08	12.50	11.41	22
Q6K548	Outer mitochondrial membrane protein porin	4.40E-04	8.04	29.22	2
O24523	Peroxidase	2.38E-04	10.57	13.13	3
Q8GT95	Polygalacturonase inhibitor 1 precursor	7.91E-07	7.17	35.46	2
Q8H7Y8	Probable histone H2A variant 1	2.40E-05	10.50	14.56	9
Q8S857	Probable histone H2A variant 2	2.40E-05	10.50	14.61	8
Q84MP7	Probable histone H2A variant 3	2.40E-05	10.50	14.46	9
A2YMC5	Probable histone H2A.1	2.33E-05	10.50	32.29	2
Q6ZL43	Probable histone H2A.1	2.33E-05	10.50	14.04	9
A2YMC6	Probable histone H2A.2	6.39E-07	10.50	13.98	2
Q6ZL42	Probable histone H2A.2	6.39E-07	10.50	13.98	9
Q84NJ4	Probable histone H2A.3	7.45E-03	10.50	13.92	8
A2Y5G8	Probable histone H2A.4	7.33E-10	10.50	16.97	9
Q6L500	Probable histone H2A.4	7.33E-10	10.50	16.97	12
A2WQG7	Probable histone H2A.5	7.33E-10	10.50	16.40	9
Q94E96	Probable histone H2A.5	7.33E-10	10.50	16.40	12
A2XZN0	Probable histone H2A.6	7.33E-10	10.50	18.04	9
Q75L11	Probable histone H2A.6	7.33E-10	10.50	16.50	12
A2ZK29	Probable histone H2A.7	2.33E-05	10.50	14.05	2
Q2QS71	Probable histone H2A.7	2.33E-05	10.50	14.05	9
A2ZK26	Probable histone H2A.8	2.33E-05	10.50	14.04	2
Q8LLP5	Probable histone H2AXa	6.46E-05	10.50	14.27	8
A2ZL69	Probable histone H2AXb	5.31E-07	10.50	14.34	3
Q2QPG9	Probable histone H2AXb	5.31E-07	10.50	14.34	10
Q6K9D8	Putative 2-oxoglutarate dehydrogenase E2 subunit	8.35E-03	6.00	49.38	2
Q6ZJX3	Putative 33 kDa secretory protein	1.24E-03	8.30	30.05	2
Q9LWS2	Putative 60S ribosomal protein L13E	1.45E-05	10.50	24.05	2
Q6ZLB8	Putative 60S ribosomal protein L4/L1	2.12E-03	10.50	44.74	2
Q9AYM0	Putative AT-Hook DNA-binding protein	3.60E-05	10.10	41.40	3
Q5JKH1	Putative BRI1-KD interacting protein 112	2.89E-06	6.00	56.27	4
Q5W6H1	Putative DNA-binding protein GBP16	2.88E-06	6.00	43.20	3
Q6K701	Putative fibrillarin	3.25E-09	10.50	32.40	8
Q6AT27	Putative fibrillarin protein	3.25E-09	10.50	37.25	4
Q6K5G8	Putative glyceraldehyde-3-phosphate dehydrogenase	3.80E-06	8.30	36.57	3
Q8GTL0	Putative glycine-rich cell wall protein	4.62E-10	11.51	28.51	9
Q6Z142	Putative glycine-rich protein	1.35E-04	10.53	32.95	12
Q5KQG2	Putative histone deacetylase HD2	1.02E-06	4.30	29.92	5
Q688F7	Putative histone deacetylase HD2	1.02E-06	4.30	32.53	5
Q9M4T5	Putative histone deacetylase HD2	1.02E-06	4.30	32.50	5
Q8H4Z0	Putative histone H1	3.00E-10	10.50	28.53	13
Q6K8A9	Putative Neurofilament triplet M protein	8.59E-11	6.00	61.52	5
Q5VND6	Putative nucleosome assembly protein 1	4.38E-06	4.30	42.64	2
Q6ESK5	Putative peptidyl-prolyl cis-trans isomerase NIMA-interacting 4	6.92E-04	10.30	15.37	2
Q6Z676	Putative phi-1	5.73E-06	8.33	32.79	2

Table 2 (continued)

Q6Z7W3	Putative phi-1	4.55E-04	8.33	32.40	2
Q8H047	Putative serine protease	1.68E-07	6.00	79.37	3
Q5VMT5	Putative Spo76 protein	3.02E-05	6.00	174.77	8
Q943L0	Putative thaumatin-like cytokinin-binding protein	2.01E-04	8.33	26.19	2
A2WS70	Putative uncharacterized protein	2.50E-05	8.33	51.14	5
A2WUB6	Putative uncharacterized protein	4.24E-11	6.00	55.90	6
A2WWU4	Putative uncharacterized protein	2.01E-04	8.33	26.19	2
A2WY45	Putative uncharacterized protein	3.63E-07	6.00	40.72	4
A2WYJ4	Putative uncharacterized protein	2.89E-06	6.00	110.51	4
A2X3K1	Putative uncharacterized protein	8.59E-11	9.90	57.99	5
A2X5E8	Putative uncharacterized protein	2.54E-04	4.30	39.24	6
A2X5K1	Putative uncharacterized protein	2.32E-06	10.07	59.85	2
A2X6V9	Putative uncharacterized protein	3.80E-06	6.00	39.03	3
A2X752	Putative uncharacterized protein	6.56E-09	4.30	81.95	4
A2X9S5	Putative uncharacterized protein	5.73E-06	8.33	32.76	2
A2X9T0	Putative uncharacterized protein	4.55E-04	8.33	32.40	2
A2XB19	Putative uncharacterized protein	6.40E-09	8.30	46.83	11
A2XB50	Putative uncharacterized protein	3.25E-09	10.50	32.40	8
A2XBU1	Putative uncharacterized protein	1.68E-07	7.17	56.27	3
A2XD51	Putative uncharacterized protein	4.12E-07	8.30	19.03	11
A2XD52	Putative uncharacterized protein	4.12E-07	8.30	17.44	11
A2XD55	Putative uncharacterized protein	4.12E-07	10.10	27.77	11
A2XD56	Putative uncharacterized protein	4.12E-07	8.30	17.56	11
A2XJ75	Putative uncharacterized protein	9.99E-07	4.30	74.17	3
A2XKH4	Putative uncharacterized protein	5.68E-05	6.29	15.96	4
A2XMY6	Putative uncharacterized protein	7.18E-05	10.50	29.66	6
A2XNH0	Putative uncharacterized protein	3.78E-09	10.07	28.22	5
A2XUU7	Putative uncharacterized protein	3.80E-06	6.00	42.05	4
A2XV86	Putative uncharacterized protein	4.09E-09	6.00	92.50	2
A2XWY9	Putative uncharacterized protein	3.67E-08	10.10	38.27	5
A2XYK3	Putative uncharacterized protein	6.95E-08	6.00	95.97	15
A2XYL5	Putative uncharacterized protein	8.21E-04	8.30	55.52	2
A2XZ78	Putative uncharacterized protein	5.53E-05	4.30	33.48	2
A2XZZ1	Putative uncharacterized protein	2.86E-04	10.10	33.14	2
A2Y0X1	Putative uncharacterized protein	3.25E-09	6.00	16.34	3
A2Y3I6	Putative uncharacterized protein	2.88E-06	6.00	43.18	3
A2Y4I5	Putative uncharacterized protein	2.21E-07	8.33	44.25	3
A2Y4S4	Putative uncharacterized protein	8.61E-06	8.33	45.77	2
A2Y876	Putative uncharacterized protein	1.02E-06	4.30	11.75	5
A2Y8K1	Putative uncharacterized protein	1.45E-05	10.50	24.05	2
A2YBU6	Putative uncharacterized protein	3.02E-05	6.00	165.56	5
A2YIS2	Putative uncharacterized protein	2.12E-03	10.50	44.71	2
A2YIV4	Putative uncharacterized protein	3.00E-10	10.50	28.38	8
A2YMS8	Putative uncharacterized protein	7.91E-07	7.17	35.46	2
A2YQ51	Putative uncharacterized protein	1.24E-12	11.51	27.89	2
A2YR07	Putative uncharacterized protein	1.09E-05	8.30	29.83	5
A2YR08	Putative uncharacterized protein	1.24E-03	8.30	30.05	2
A2YW60	Putative uncharacterized protein	5.29E-07	4.25	48.71	8
A2Z0D1	Putative uncharacterized protein	4.40E-04	6.00	29.17	2
A2Z1W4	Putative uncharacterized protein	1.55E-04	6.00	28.14	2
A2Z2G1	Putative uncharacterized protein	3.15E-09	4.30	80.21	4
A2Z6T0	Putative uncharacterized protein	2.55E-05	4.30	31.58	2
A2Z852	Putative uncharacterized protein	1.67E-03	8.30	59.17	3
A2ZA46	Putative uncharacterized protein	2.06E-04	6.00	31.36	4
A2ZAG9	Putative uncharacterized protein	3.60E-05	10.10	41.40	3
A2ZLU2	Putative uncharacterized protein	6.47E-06	8.33	23.46	4
A2ZMM6	Putative uncharacterized protein	7.98E-07	10.10	65.77	6
A2ZN20	Putative uncharacterized protein	5.68E-05	6.29	16.07	3

Table 2 (continued)

A2ZV28	Putative uncharacterized protein	1.07E-04	10.07	51.54	2
A2ZX38	Putative uncharacterized protein	4.24E-11	6.00	55.90	6
A2ZZB2	Putative uncharacterized protein	4.61E-05	10.50	34.35	2
A3A133	Putative uncharacterized protein	2.89E-06	6.00	89.40	3
A3A5N7	Putative uncharacterized protein	8.59E-11	10.10	60.19	5
A3A7E8	Putative uncharacterized protein	2.54E-04	4.30	39.20	6
A3A8R9	Putative uncharacterized protein	3.80E-06	6.00	39.03	3
A3A904	Putative uncharacterized protein	6.56E-09	4.30	81.88	4
A3ACP9	Putative uncharacterized protein	6.40E-09	8.30	62.80	14
A3AD86	Putative uncharacterized protein	5.81E-04	4.30	42.20	2
A3AK49	Putative uncharacterized protein	9.99E-07	4.30	74.15	3
A3AL37	Putative uncharacterized protein	5.81E-04	4.30	51.46	2
A3ALU8	Putative uncharacterized protein	3.77E-15	10.07	47.86	3
A3ANA0	Putative uncharacterized protein	5.81E-04	4.30	49.77	2
A3AP59	Putative uncharacterized protein	3.78E-09	9.87	36.03	5
A3AV14	Putative uncharacterized protein	3.80E-06	6.00	42.05	4
A3AVB2	Putative uncharacterized protein	1.82E-09	6.00	89.36	4
A3AWW7	Putative uncharacterized protein	3.67E-08	10.10	38.25	5
A3AYA3	Putative uncharacterized protein	6.95E-08	6.00	95.97	15
A3AYV7	Putative uncharacterized protein	5.53E-05	4.30	33.48	2
A3B2Z6	Putative uncharacterized protein	2.88E-06	6.00	40.33	2
A3B3W3	Putative uncharacterized protein	2.21E-07	8.33	47.67	3
A3B412	Putative uncharacterized protein	5.81E-04	6.00	69.71	2
A3B449	Putative uncharacterized protein	8.61E-06	8.33	45.77	2
A3B7G9	Putative uncharacterized protein	1.02E-06	4.30	11.75	5
A3B8F1	Putative uncharacterized protein	4.38E-06	4.30	49.90	2
A3BAS3	Putative uncharacterized protein	3.02E-05	6.00	165.65	6
A3BEJ5	Putative uncharacterized protein	5.81E-04	4.30	50.12	2
A3BH97	Putative uncharacterized protein	3.00E-10	10.50	65.51	6
A3BY14	Putative uncharacterized protein	4.40E-04	10.07	29.29	2
A3BYR8	Putative uncharacterized protein	6.92E-04	10.10	16.87	2
A3C008	Putative uncharacterized protein	3.15E-09	4.30	47.45	2
A3C777	Putative uncharacterized protein	2.06E-04	6.00	31.36	4
A3C7J3	Putative uncharacterized protein	3.60E-05	8.30	28.43	3
A3CJC5	Putative uncharacterized protein	7.98E-07	10.10	62.77	6
Q2QTC8	Putative uncharacterized protein	1.25E-03	12.50	17.39	2
Q7PC92	Putative uncharacterized protein	2.77E-05	6.00	39.32	2
Q84TZ6	Putative uncharacterized protein OSJNBa0087M10.17	9.99E-07	4.30	76.96	3
Q6AUL0	Putative uncharacterized protein OSJNBb0006J12.11	2.21E-07	8.33	47.71	3
Q60DV9	Putative uncharacterized protein P0426G01.13	8.61E-06	8.33	42.23	2
Q5NAI9	Putative Y1 protein	1.48E-11	6.00	50.11	8
Q10FE5	Retrotransposon protein, putative, Ty1-copia subclass, expressed	5.68E-05	4.25	15.18	4
Q10FE7	Retrotransposon protein, putative, Ty1-copia subclass, expressed	5.68E-05	7.17	20.38	4
Q7XJB4	Ribosomal protein large subunit 13	1.45E-05	10.50	24.01	3
Q6EQL9	Ribosomal protein-like	2.77E-05	10.10	48.16	2
P0C510	Ribulose biphosphate carboxylase large chain precursor	2.77E-06	6.00	52.88	3
P0C511	Ribulose biphosphate carboxylase large chain precursor	2.77E-06	6.00	53.70	3
P0C512	Ribulose biphosphate carboxylase large chain precursor	2.77E-06	6.00	52.88	3
Q6ENG6	Ribulose biphosphate carboxylase large chain precursor	2.77E-06	6.00	52.88	3
Q2QND8	Thaumatococcus-like cytokinin-binding protein, putative, expressed	6.47E-06	8.33	23.43	4
Q6ZL53	Translational elongation factor Tu	2.55E-05	6.00	50.41	2
Q851Y8	Translational elongation factor Tu	2.55E-05	6.00	48.42	2
Q43594	Tubulin beta-1 chain	5.81E-04	4.30	50.29	2
Q8H7U1	Tubulin beta-2 chain	5.81E-04	4.30	50.18	2
P45960	Tubulin beta-4 chain	5.81E-04	4.30	50.29	2
Q76FS3	Tubulin beta-6 chain	5.81E-04	4.30	49.89	2
P37832	Tubulin beta-7 chain	5.81E-04	4.30	49.82	2
Q76FS2	Tubulin beta-8 chain	5.81E-04	4.30	49.63	2



Table 2 (continued)

CAJ46982	unnamed protein product	3.67E-08	10.10	38.24	5
CAJ46983	unnamed protein product	3.67E-08	10.30	37.51	5

- a) Protein accession: Protein accession number  
b) MW: predicted molecular mass  
c) PI: predicted protein pI  
d) Identified Peptides: Number of peptides that matched with the identified protein in MS/MS analyses.

Table 3 Proteins identified by searching the reversed sequence database

UniProt ID No.	Protein Name	pI	MW (Da)	Identified Peptides No.
A2WWI2	Hypothetical protein	6.0	101589	1
A2Z4H7	Hypothetical protein	6.0	118054.7	1
A2ZHK0	Hypothetical protein	10.07	50300.9	1
A3AB78	Hypothetical protein	10.07	51652.89	1
Q2QYM3	CBL-interacting serine/threonine-protein kinase 15, putative, expresse	10.07	50323.96	1
Q6Z7S3	Putative Altered Response to Gravity (Os02g0741100 protein)	10.07	49325.11	1
Q7G768	Putative receptor-like protein kinase (Hypothetical protein)	6.0	118109.7	1

- a) Protein accession: Protein accession number  
b) MW: predicted molecular mass  
c) PI: predicted protein pI  
d) Identified Peptides: Number of peptides that matched with the identified protein in MS/MS analyses.

### Gene ontology analysis

To help understand the distribution and function of the chromatin associated proteins, we obtained GO annotations of the identified proteins from the AgBase

(McCarthy et al., 2007) (<http://www.agbase.msstate.edu/>). As shown in Figure 3.6, the GO term distributions of proteins from both 2-DE MALDI-TOF and shotgun approaches are pretty similar. Regarding the protein locations, distributions of identified proteins from 2-DE MALDI method were intracellular (28%), nuclear proteins (15%), cytoskeleton (14%), cytoplasm proteins (9%), membrane protein (6%), cytosol (4%), mitochondrion (3%), plastid (2%), ribosome (3%), etc., while the shotgun proteomics identified proteins distributed as intracellular (26%), nuclear proteins (19%), cytoskeleton (12%), cytoplasm proteins (10%), membrane protein (7%), mitochondrion (4%), ribosome (3%). In comparison, the human metaphase chromosome proteome is much different from our results, which contains 38% mitochondrion proteins, 29.8% nuclear proteins, 12.7% ribosome proteins, 11.4% cytoplasmic proteins, 4.4% cytoskeleton proteins, and 3.2% unknown proteins (Uchiyama et al., 2005). Gene Ontology analyses from both approaches based on biological processes indicated that about 30% of the genes were involved in several metabolic pathways and 25% genes were involved in biosynthesis and cell organization (Figure 3.6). In addition, the analyses based on molecular function showed that the predominant protein functions are nucleotide binding, binding, catalytic activity, structural molecular activity, which was consistent with that many histone and histone variants, nuclear proteins were presented in our protein lists.

In the current GO classification system, one protein may be grouped into more than one GO category. Therefore, sometimes the accumulative percentage is over 100%. On the other hand, the information for proteins with multiple functions is not completely compiled. For example, glyceraldehyde-3-phosphate dehydrogenase (GAPDH) is classified as cytoplasm instead of nuclear protein although it has been confirmed that GAPDH is involved in transcriptional regulation. The situation is the same for many

proteins involved in translation and RNA binding. Thus, the GO analysis results presented should be taken with the information described above in mind.

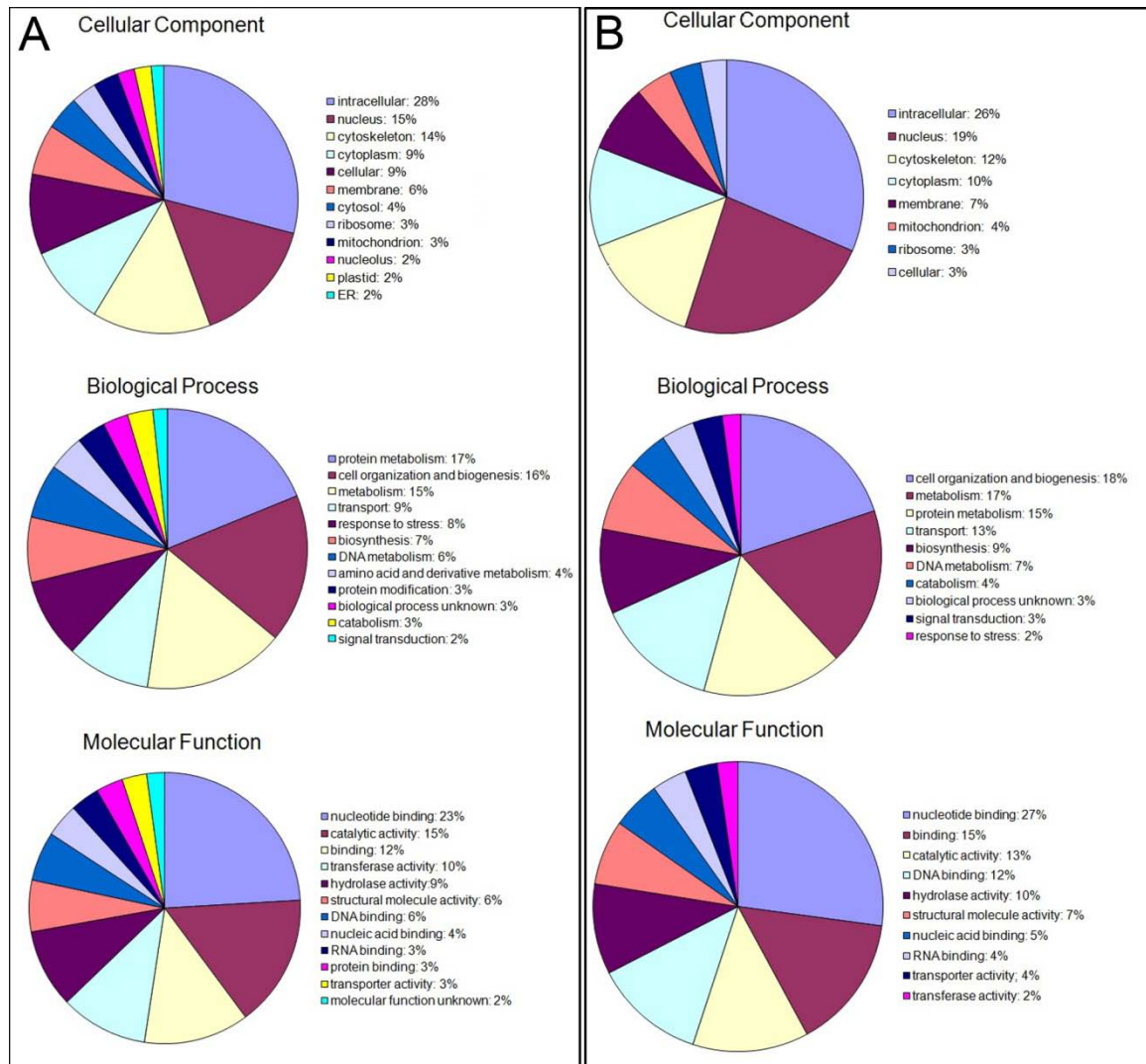


Figure 6 Functional classification of chromatin proteome

Distribution of the proteins copurified with chromatin in different functional classifications. Percentage distribution of the unique proteins was used to make the pie chart based on the GO of cellular localizations. The pie chart was generated using the analysis results of the “GOSlimViewer” tool at AgBase. (A) Pie chart was made from the proteins identified by 2-DE MALDI-TOF analyses. (B) Pie chart was made from the proteins discovered by shotgun approach.

## Discussions

### Chromatin purification

Like all organelle sub-proteome studies, our goal is to obtain high quality chromatin permitted by current technologies and identify all possible proteins associated. We are aware that it is impossible to isolate the chromatin without pulling down some other non-chromatin proteins. The chromatin proteome study is intended to identify chromatin associated candidate proteins instead of a final proof of their association with chromatin. We have developed a protocol for purification of chromatin from rice suspension cells. The procedure includes protoplast isolation, nucleus purification, and chromatin purification. Several lines of evidences indicate that chromatin is highly enriched and of high quality in our preparation. 1) Electron microscopy and optical microscopy reveal that the chromatin preparation maintains, at least partially, high level structure and is free of visible organelle contamination. 2) Western blots have shown that abundant protein complex COP9 signalosome, which is nuclear enriched, and cytoplasmic protein PhyA did not co-purify with chromatin. 3) Among the 509 identified proteins using 2-DE gel approach, nuclear proteins represent 15%, which rise to 19% when using shotgun proteomics. Proteins from cytoskeleton, mitochondrion, plastid, and ribosome are 14%, 3%, 2%, and 3%, respectively. Because cytoskeleton and ribosome proteins can be both nuclear and cytoplasmic, the ratio of nuclear protein should be higher than 15%. The primary contaminants of human chromosome are mitochondrion (38.6%), cytoskeleton (12.7%) and ribosome (4.4%), respectively (Uchiyama et al., 2005). Compared with the human chromosome, our sample had less mitochondria contamination. Using our chromatin preparation, we have resolved 972 protein spots on 2-DE gels in all three biological replicas. 607 prominent protein spots were excised for

mass analyses. 509 proteins were identified with high confidence (C.I.%>95%). As shown in Figure 3.3, over half of the protein spots, particularly the low abundance proteins, have not been identified probably due to the inefficiency of in-gel digestion and peptide recovery as well as the limitation of mass spectrometer sensitivity. Further identification of these proteins will provide a more complete picture of the chromatin proteome, particularly the low abundance proteins. Since there is no cross-linking treatment prior to chromatin isolation and the chromatin purification is a procedure involved multiple steps of washing, we can not exclude that weakly associated proteins might have dissociated from chromatin. It will be interesting to examine the differences after treating the cell with cross-linking reagent prior to chromatin isolation. Nevertheless, the established protocols of chromatin purification, chromatin protein isolation, and mass analysis will be very useful tools for further studies on chromatin proteome in responses to environmental and biological stimuli.

### **Protein identification and agriGO annotations**

In this study 509 *Oryza sativa* chromatin associated protein spots have been successfully identified, corresponding to 269 unique proteins, with high confidence using 2-DE followed by MALDI-TOF analysis. While performing MudPIT approach, using ESI MS/MS, 507 proteins have been identified, corresponding to 292 unique proteins. By combining the data from both methods, the proteome map of chromatin was constructed with 519 proteins. There are only 41 proteins commonly found both in 2-DE and MudPIT, which corresponded about 15% indicating that the two methods have their own biases to identify proteins and they are complementary with each other. This protein distribution is represented in Figure 3.7 using a Venn diagram. The genes corresponding

to all 519 unique proteins were characterized using the agriGO Gene Ontology tool. AgriGO (Du et al., 2010) is aim to identify GO annotations for which each of the groups of genes was significantly enriched compared to all genes in the genome. SEA (Singular enrichment analysis) of agriGO was carried out and generated a tree structure graph. As shown in Figure 3.8 (biological process), Figure 3.9 (cellular component) and Figure 3.10 (molecular function), GO terms are represented as boxes containing detailed description with statistical information, organized and connected based on their relationship. The graphical result is a GO hieratical image containing all statistically significant terms. These nodes in the image are classified into ten levels which are associated with corresponding specific colors. The smaller of the term's adjusted p-value, the more significant statistically, and the node's color is darker and redder. Compared with whole genome GO annotation, the rice chromatin proteome seems to particularly enrich in certain organism processes, such as nucleosome and chromatin assembly, cellular protein complex assembly, protein folding and biosynthesis, transport, and response to stress (Figure 3.8). Most of those proteins were located in nucleosome, chromatin, cytoskeleton, and mitochondrion membrane etc as shown in the cellular component analysis by agriGO in Figure 3.9. These results strongly substantiate that the primary role of chromatin associated proteins is to control and support nucleosome and chromatin structure. The molecular function annotation study (Figure 3.10) revealed that the most predominant protein functions are nucleotide binding, hydrolase activity, transferase activity and channel activity, which are consistent with the results from Agbase annotation (Figure 3.6).

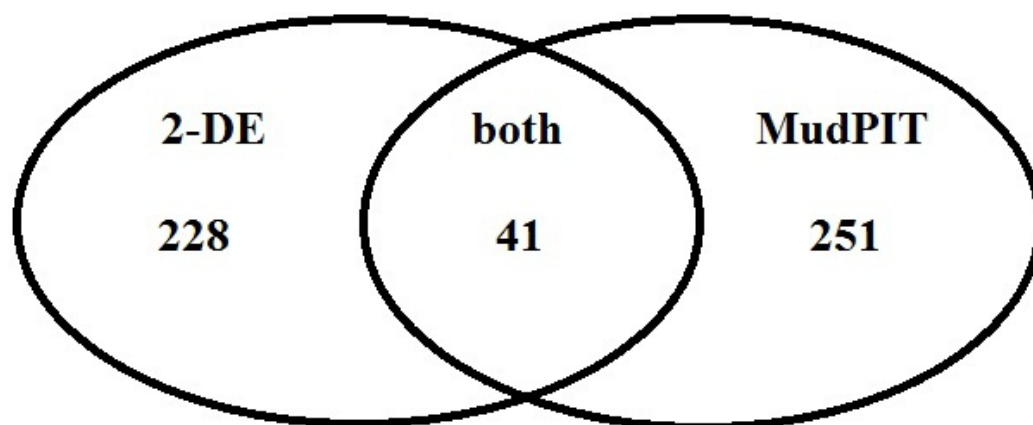


Figure 7 Venn diagram representing the proteins distribution identified using 2-DE and MudPIT

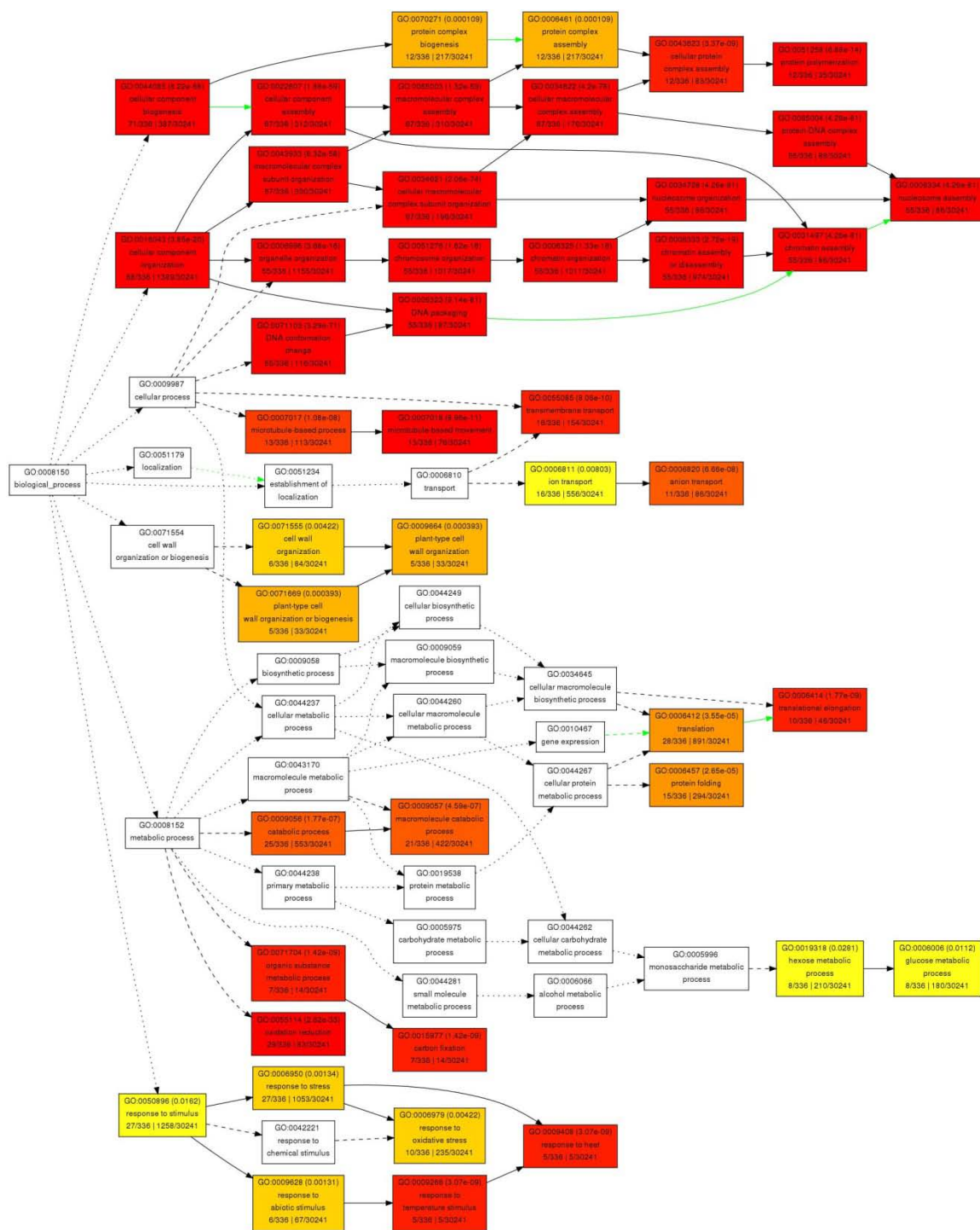
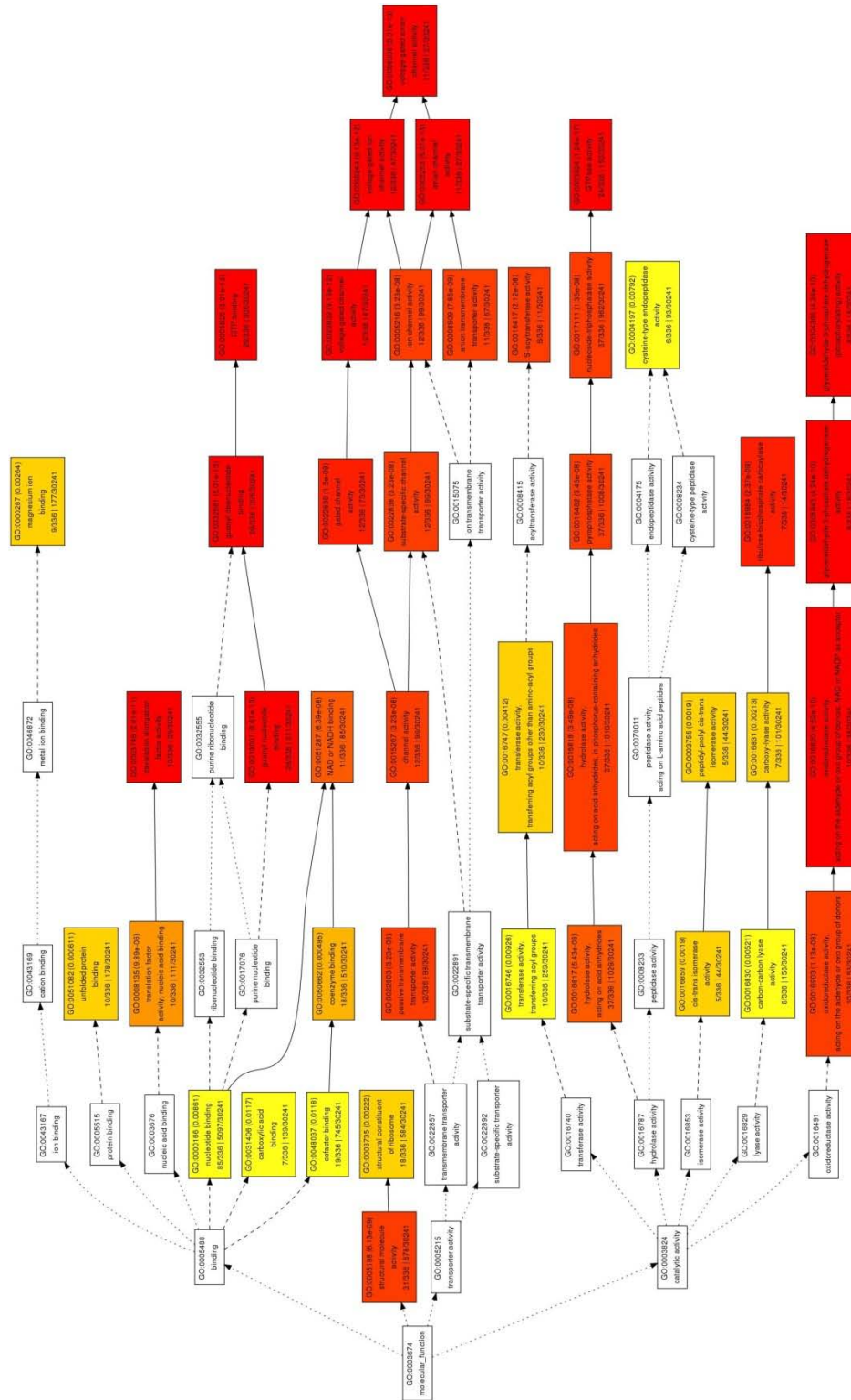


Figure 8 Hierarchical tree graph of total unique proteins in biological process category generated by SEA of agriGO







### **Proteins co-purified with chromatin**

One of the most interesting observations in this study is the existence of a large number of histone variants in rice. We have observed 11 possible H2A variants and 10 possible H2B variants in addition to the 4 common core histone proteins. In mammals, only 3 H2B variants and 6 H2A variants have been identified thus far (Bernstein and Hake, 2006). While 5 histone H3 variants have been reported in mammals, we only identified two histone H3 variants. These observations suggest that there is a significant difference between the mammalian chromatin and plant chromatin. As in mammals, on the other hand, no H4 variant has been detected. Because H4 plays a unique structural role in the histone core, it indicates that the overall structure of the nucleosome is still conserved. It is believed that histone variants provide different sequence modules or cassettes that can be post-translationally modified and subsequently recognized by specific effector proteins to bring about downstream effects. Therefore, the differences in protein composition among histone variants contribute to distinct, variant-specific biological functions. It has been proposed that specific histone variants in the nucleosome generate distinct chromosomal domains, called the nucleosome code, for the regulation of gene expression (Bernstein and Hake, 2006; Hake and Allis, 2006). MS/MS was carried out to investigate histone post-translational modifications. Probably because our mass spectrometer was optimized for protein identification not for post translational modification during the mass analyses, only about two dozens of peptides with putative posttranslational modification were identified. After manual examination of the mass spectra, only a few peptides were identified with confidence (Figure 3.5). The others had high noise peaks probably because we were analyzing a very complicated protein mixture

and without specifically selecting the modified peptide for MS/MS. Further investigation of histone post-translational modifications was performed and discussed in chapter 5 by an alternative approach.

In addition to histones, high mobility group proteins, including HMG1, HMGB1 and multiple At-hook proteins, have been identified. Another major group of proteins are DNA binding proteins, including several BRI1-KD interacting proteins, several WRKY transcription factors, BHLH transcription factor, BKRF1 encodes EBNA-1 protein, Putative DNA-Binding protein GBP16, Putative nucleoid DNA-binding proteins, nucleoid DNA-binding protein cnd41-like, DNA-directed RNA polymerase II 23K chain, and auxin responsive protein CsPK3. Proteins involved in nucleosome assembly include nucleosome assembly protein 1-like protein, nucleosome assembly protein 1, and Putative nucleosome/chromatin assembly factor A. Other known chromatin proteins are putative histone deacetylase, putative Spo76 protein, retrotransposon proteins, etc.

In addition to these well known chromatin associated proteins, such as histones and DNA binding proteins, we also have identified a large number of proteins whose relationships with chromatin are not clear. Since chromatin is a large protein-DNA supercomplex, it is not a surprise if some proteins are trapped within the chromatin during purification. However, it is very interesting to notice that many of the proteins also co-purify with *C. elegans* chromatin and metaphase chromosome of human cell lines even though different methods have been used for purification. These proteins include Tu translational elongation factor, heat shock and chaperonin proteins, RNA binding proteins, several ribosome subunit proteins, putative U3 snoRNP protein IMP4, prohibitin, several 26S proteome regulatory subunits, anti-oxidative stress proteins, etc. The co-purification of these proteins with chromatin in different organisms suggests that

these proteins, at least, can bind chromatin tightly during purification although it is unknown whether they are indeed associated with chromatin *in vivo*. In addition, we identified a plastid cationic peroxidase at multiple distinct protein spots while other high abundance plastid proteins such as RUBISCO were not detected. Mitochondrial proteins N-terminal N-acetyl-gamma-glutamyl phosphate reductase (Arg6) and C-terminal acetylglutamate kinase (Arg5) in yeast and cytoplasmic protein GAPDH in mammals have been reported to act as transcriptional regulator although their roles in metabolism are well established (Mazzola and Sirover, 2002; Zheng et al., 2003; Hall et al., 2004). The GAPDH proteins were co-purified with chromatin in our studies as well as in the human and *C. elegans* studies. These observations suggest that it will be interesting and necessary to further test if proteins co-purified with chromatin are indeed associated with chromatin *in vivo* in plants.

A large number of cytoskeleton proteins co-purified with chromatin in our studies and also co-purified with human metaphase chromosome and *C. elegans* chromatin (Uchiyama et al., 2005; Chu et al., 2006).  $\beta$ -actin has been shown to be located with the entire metaphase chromosome body and it has been reported that it is a component of chromatin-remodeling complex (Zhao et al., 1998; Uchiyama et al., 2005). In addition, injection of anti-actin antibodies into *Xenopus* oocytes blocks chromosome condensation (Scheer et al., 1984). The association of cytoskeleton proteins to chromatin may be related to chromatin spatial organization and dynamic movements (Uchiyama et al., 2005).

### **The efficiency of chromatin protein identification using shotgun approach**

The previous studies in our lab using total protein extracts in Arabidopsis have shown that shotgun approach is more efficient than 2-DE gel based approach in protein identifications (Chitteti et al., 2008). However, the number of proteins identified using shotgun method is much less than our expectation when chromatin proteins are examined in this study. Since the same protocol was used for the analysis of Arabidopsis total protein mixture and we have repeated the experiments multiple times, we believe that the failure is not due to a technique mistake in our experiments, although we can not exclude the possibility. Several factors related to the characteristics of chromatin associated proteins may have contributed to the failure in nonhistone protein identification using the shotgun method. One possible reason is that histones, such as H2A and H2B, are highly abundant compared with other nonhistone proteins in the chromatin protein mixture. Meanwhile, many chromatin associated proteins, such as transcriptional factors, are expressed in very low quantity. Therefore, those proteins can not be easily detected. Another possible reason is that many DNA binding proteins are basic proteins rich in lysine and arginine, trypsin is not a good enzyme for these proteins in mass analysis. The third possibility is that 2D-LC-MS/MS process bias against some proteins (Mallick et al., 2007). On the other hand, shotgun method identified H1 protein and many other known chromatin binding factors that were not identified using 2-DE method either due to the corresponding spot(s) was not selected for mass analysis or failure in detecting the proteins during mass analysis.

## CHAPTER IV

### PHOSPHOPROTEOME PROFILING OF RICE CHROMATIN

#### **Abstract**

Chromatin phosphoproteome has been investigated in rice using suspension cell as a starting material. Among the 205 putative phosphoprotein spots, 154 proteins have been identified with a Confidence Interval (C.I. %) over 95% using MALDI-TOF/TOF following Pro-Q Diamond Phosphoprotein 2-DE Gel stain. Among them, many proteins previously have been reported have the potential to be phosphorylated, suggesting that Pro-Q Diamond Phosphoprotein Gel Stain is a useful tool in exploring phosphoproteome. Further GO annotation study indicated that phosphoproteins involved in diversified cellular processes and cellular components, especially the nucleosome and chromatin assembly and conformational organization.

#### **Introduction**

Protein phosphorylation plays a critical role in gene transcription, DNA replication, and chromatin remodeling. Identification of chromatin associated phosphoproteins should provide new insight into chromatin structure and potentially the regulation of chromatin structure and function. Since many protein spots with similar molecular weight but different pIs had been observed on 2-DE gel, we stained the 2-DE gel with Pro-Q Diamond Phosphoprotein in Gel Stain (Molecular Probe) after protein separation. The Pro-Q Diamond fluorescence dye has been widely used in the identification of phosphoproteins (Steinberg et al., 2003; Schulenberg et al., 2004;

Agrawal and Thelen, 2005; Stasyk et al., 2005; Agrawal and Thelen, 2006; Chitteti and Peng, 2007b, a). These reports substantiate that the Pro-Q Diamond dye is a useful tool in the identification of candidate phosphoproteins. Besides, using Pro-Q Diamond Phosphoprotein gel stain, one can follow quantitative changes of particular protein spots on 2-DE gels in a time course under a treatment or at different developmental stages. It is a useful tool in revealing differential regulation of phosphoproteins.

Direct visualization of the putative phosphoprotein spots using Pro-Q Diamond dye enables us to select the potentially interesting proteins for mass analyses. However, the limited protein quantity, from an excised protein spot of 2-DE gel, makes mapping the phosphorylated residue(s) extremely challenging. It is also found that using an appropriate concentration of Pro-Q Diamond dye is critical for reducing non specific background. Recently, several phosphoproteome analysis methods have been developed, including titanium dioxide microcolumn (Larsen et al., 2005), immobilized metal-affinity chromatography method (Ficarro et al., 2002), phosphopeptide enrichment by IEF (Maccarrone et al., 2006), and the phosphoprotein extraction kit of QIAGEN (Jones et al., 2006). These methods in combination with mass spectrometry analysis may provide a solution for the identification of phosphoproteins associated with chromatin.

In this study, the proteins of chromatin extract in rice have been investigated using the Pro-Q Diamond Phosphoprotein Gel Stain. 205 prominent protein spots were constantly stained by the Pro-Q Diamond, and 154 confident putative phosphoproteins have been identified by MALDI-TOF/TOF analyses.



## **Materials and methods**

Rice nuclei isolation, chromatin isolation, protein extraction, 2D PAGE analysis were carried out as explained in materials and methods section of Chapter 3.

### **Identification of phosphoproteins using Pro-Q Diamond gel stain**

Detection of phosphoproteins after separation on 2-DE gels was conducted by following the instructions from the manufacturer (Molecular Probes). In brief, 2-DE gels were fixed in solution containing 50% methanol and 10% acetic acid, washed with several changes of water to remove SDS, and stained with the Pro-Q Diamond dye. After destain, the gel images were recorded using a VersaDoc4000 (Bio-RAD). A spot constantly stained with Pro-Q Diamond dye in all three biological replicas was considered as a putative phosphoprotein spot. Protein gels of the same protein samples were also stained with SYPRO Ruby to correlate the protein spots revealed by these two different dyes. The ratios of SYPRO Ruby vs Pro-Q diamond stain in each protein spot were calculated after gel stain intensity normalization using PDQuest 7.4.0 software. The average of three biological replicas was used for the calculation of ratio.

In gel digestion and Mass Spectrometry were carried out as explained in materials and methods section of Chapter 3.

### **Investigating GO annotations of putative phosphoproteins**

Functional categorization of genes was carried out according to the GO rules (Ashburner et al., 2000). AgriGO (Du et al., 2010) was performed to identify GO annotations for which each of the groups of genes was significantly enriched compared to all genes in the genome. In all cases, the query list for agriGO consisted of the genes which met the specified conditions, while the background was the TIGR gene model.

Statistical significance was determined using a hypergeometric test using the Yekutieli multi-test adjustment (Benjamini). We used SEA in agriGO to do the analysis, and generated a tree structure graph. GO terms are represented as boxes containing detailed description, organized and connected based on their relationship (Du et al., 2010).

## **Results**

### **Mapping putative phosphoproteome co-purified with chromatin using Pro-Q Diamond phosphoprotein stain**

Gel image analyses were carried out by PDQUEST 7.4.0. software (Bio-Rad, Hercules, CA), which identified 390 putative protein spots. Further manual examination confirmed that 205 prominent protein spots were constantly stained by the Pro-Q Diamond dye (Figure 4.1 and Figure 4.2). Compared with SYPRO Ruby stain (Figure 4.1C), whose stain intensity is proportional to the protein quantity, the Pro-Q Diamond phosphoprotein stain displayed high specificity to some specific proteins (Figure 4.1B). For example, spots 1309, 1307, and 2306 were heavily stained by SYPRO Ruby but weakly stained by Pro-Q Diamond. Meanwhile, spots 2408, 1404, and 2304 were weakly stained by SYPRO Ruby but intensively stained by the Pro-Q Diamond dye. To reveal the Pro-Q stain specificity, the ratios of Pro-Q Diamond vs. SYPRO Ruby staining were calculated. The intensities of the gels among the biological replicas and between SYPRO Ruby and Pro-Q Diamond stain were normalized using PDQuest 7. 4.0. before data collection and calculating the ratios.

Proteins that were intensively stained by Pro-Q stain were excised and processed for MS/MS analysis using a MALDI-TOF mass spectrometer. Totally, 154 protein annotations that corresponding to 102 unique phosphoproteins were discovered and were

shown in Table 4.1. These proteins include H3-maize, H3.3, H2A, H2B, putative WRKY DNA-binding protein, putative retrotransposon protein, putative transposon protein, etc. In some protein spots, more than two proteins were identified, making it impossible to pinpoint which one was phosphosphorylated. On the other hand, some proteins were presented in multiple protein spots that shared the similar molecular weight but different pIs. These proteins are putative phosphorylated proteins, including spots 1403 and 1404 for glycine-rich RNA binding protein-like protein; spots 4306, 5407, and 6403 for cationic peroxidase or H0814G11.3; spots 2506 and 3501 for glyceralde-3-phosphodehydrogenase; spots 3401, 3413 and 4406 for hypothetical protein (OsJ\_024928); 4303 and 5302 for B0812A04.3 protein or B1358B12.5; 8602 and 8603 for Os01g0685800, etc.

### **GO annotations of putative phosphoproteins**

To understand the biological processes and cellular components involved in chromatin associated proteins in perspective of identified phosphoproteome in rice, Gene Ontology analyses were performed using the all identified phosphoproteins detected by Pro-Q diamond fluorescent staining and further validated by MALDI-TOF mass spectrometry. SEA analysis computes GO term enrichment in one set of protein by comparing it to another set, named the reference lists. In this dissertation, we used rice TIGR whole genome as the reference.

As shown in Figure 4.3 (biological process) and Figure 4.4 (cellular component), boxes in the graph represent GO terms labeled by their GO ID, term definition and statistical information. The significant term (adjusted  $P \leq 0.05$ ) are marked with color, while non-significant terms are shown as white boxes. The diagram, the degree of color

saturation of a box is positively correlated to the enrichment level of the term. Solid, dashed, and dotted lines represent two, one and zero enriched terms at both ends connected by the line, respectively. Phosphoproteins seem to be particularly enriched in certain organism processes and development, including nucleosome assembly, nucleosome organization, DNA packaging, DNA conformation change, and chromatin assembly and disassembly as shown in Figure 4.3. Furthermore, substantially more phosphoproteins located in nucleosome, cytoplasmic membrane-bounded vesicle, and mitochondrial membrane (Figure 4.4). No significant enrichment was found in the molecular function category. The detailed distribution of enriched biological process and cellular component for these putative phosphoproteins are clearly depicted in Figure 4.3 and Figure 4.4. These results indicated that the phosphoproteins involved in diversified cellular processes and cellular components, especially the nucleosome and chromatin assembly and conformational organization.

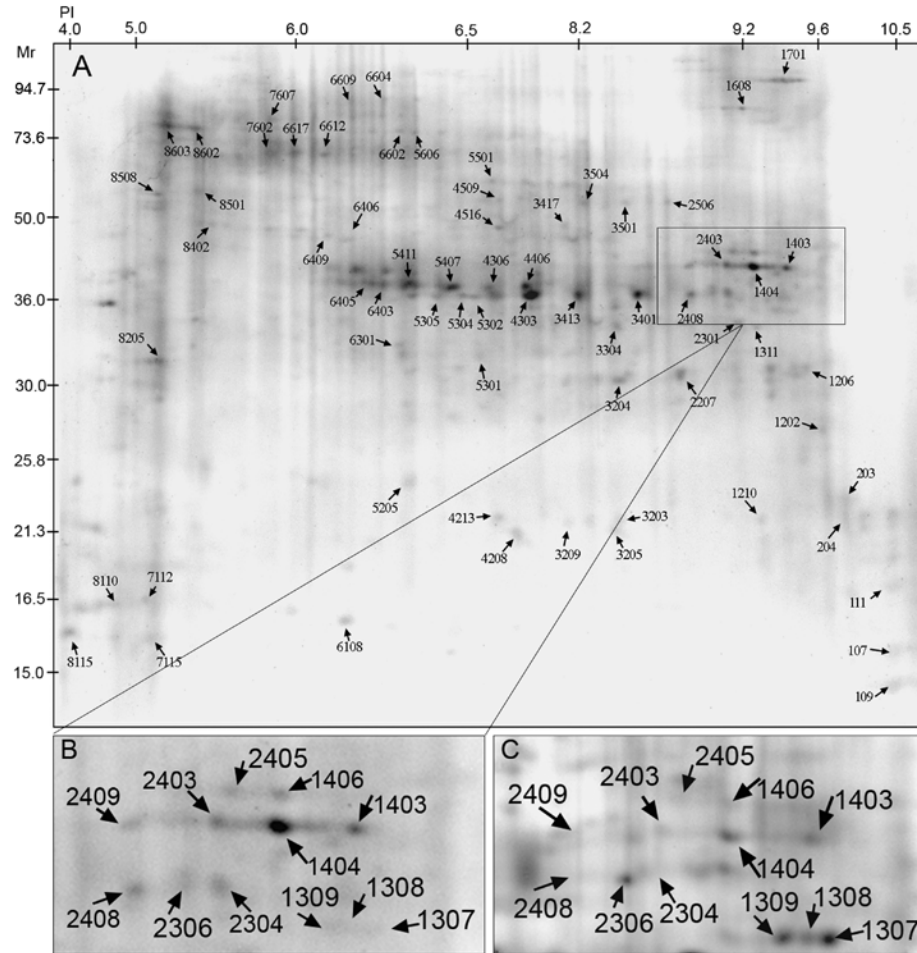


Figure 11 2-DE gel images of the putative phosphoproteins associated with rice chromatin.

Proteins were extracted from purified chromatin of rice suspension cells, separated on 2-DE gel, and stained with Pro-Q Dimond phosphoprotein stain and SYPRO Ruby, respectively. A). 2-DE gel image of chromatin associated phosphoproteins revealed by Pro-Q Diamond stain. Proteins identified with high confidence (C.I.%>95%) are marked with arrows. Molecular mass markers are on the left and the pH gradient of the first dimension is indicated on the top. The second dimension SDS PAGE was 12%. B) An enlarged section of the phosphoproteome image stained by Pro-Q Diamond dye. C) SYPRO Ruby stain image corresponding to the region shown in B.

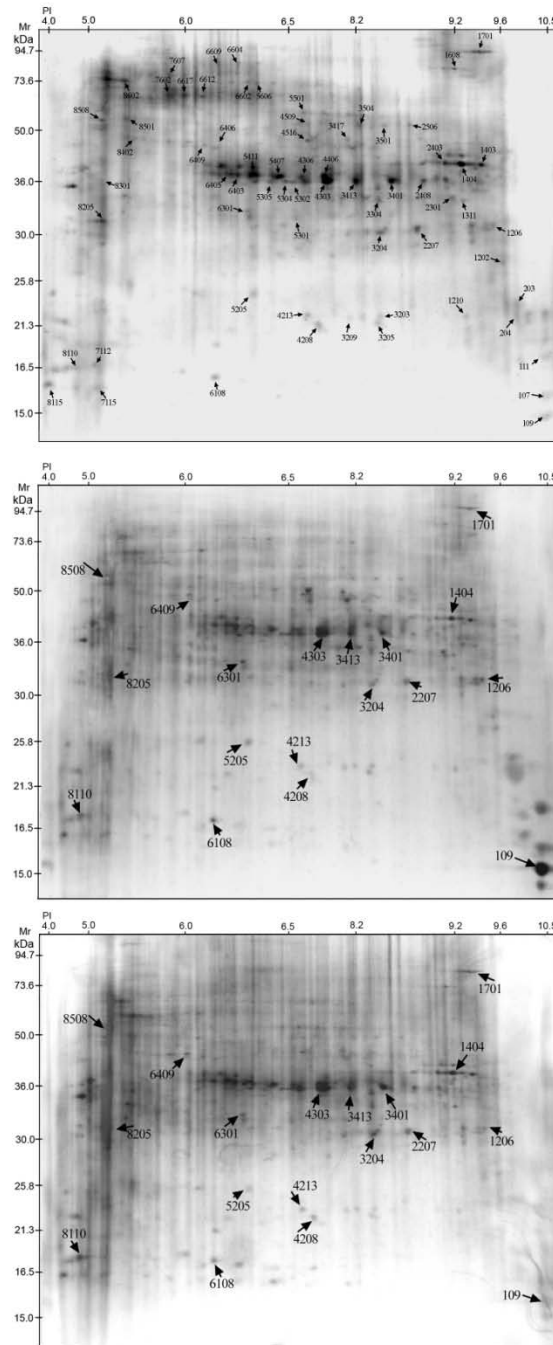


Figure 12 2-DE gel images of the putative phosphoproteins co-purified with rice chromatin.

The images of the three gels represent three biological replicates.

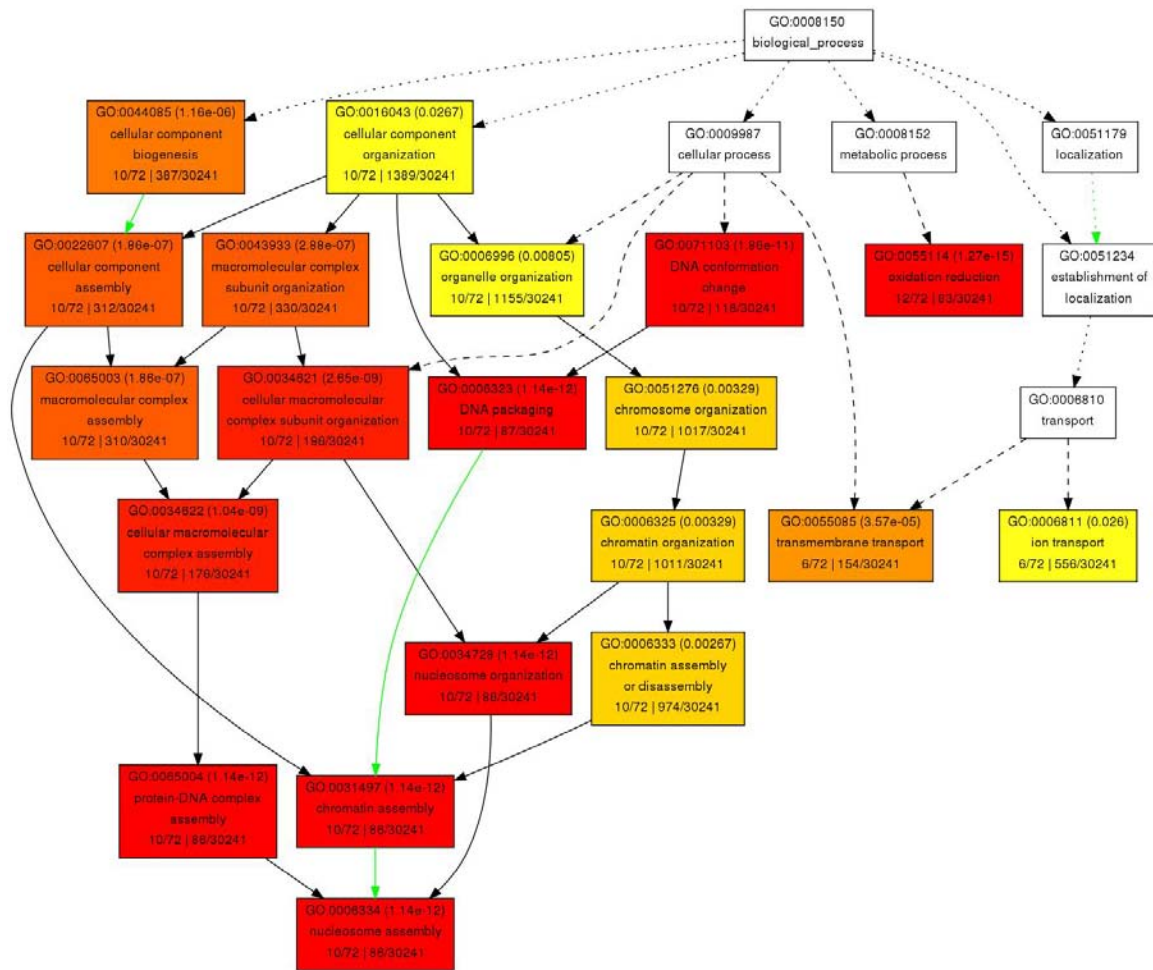


Figure 13 Hierarchical tree graph of overrepresented GO terms in biological process category generated by SEA.

This figure shows the significant biological process GO annotations for chromatin associated putative phosphoproteins. Boxes in the graph represent GO terms labeled by their GO ID, term definition and statistical information. The significant term (adjusted  $P \leq 0.05$ ) are marked with color, while non-significant terms are shown as white boxes. The diagram, the degree of color saturation of a box is positively correlated to the enrichment level of the term. Solid, dashed, and dotted lines represent two, one and zero enriched terms at both ends connected by the line, respectively.

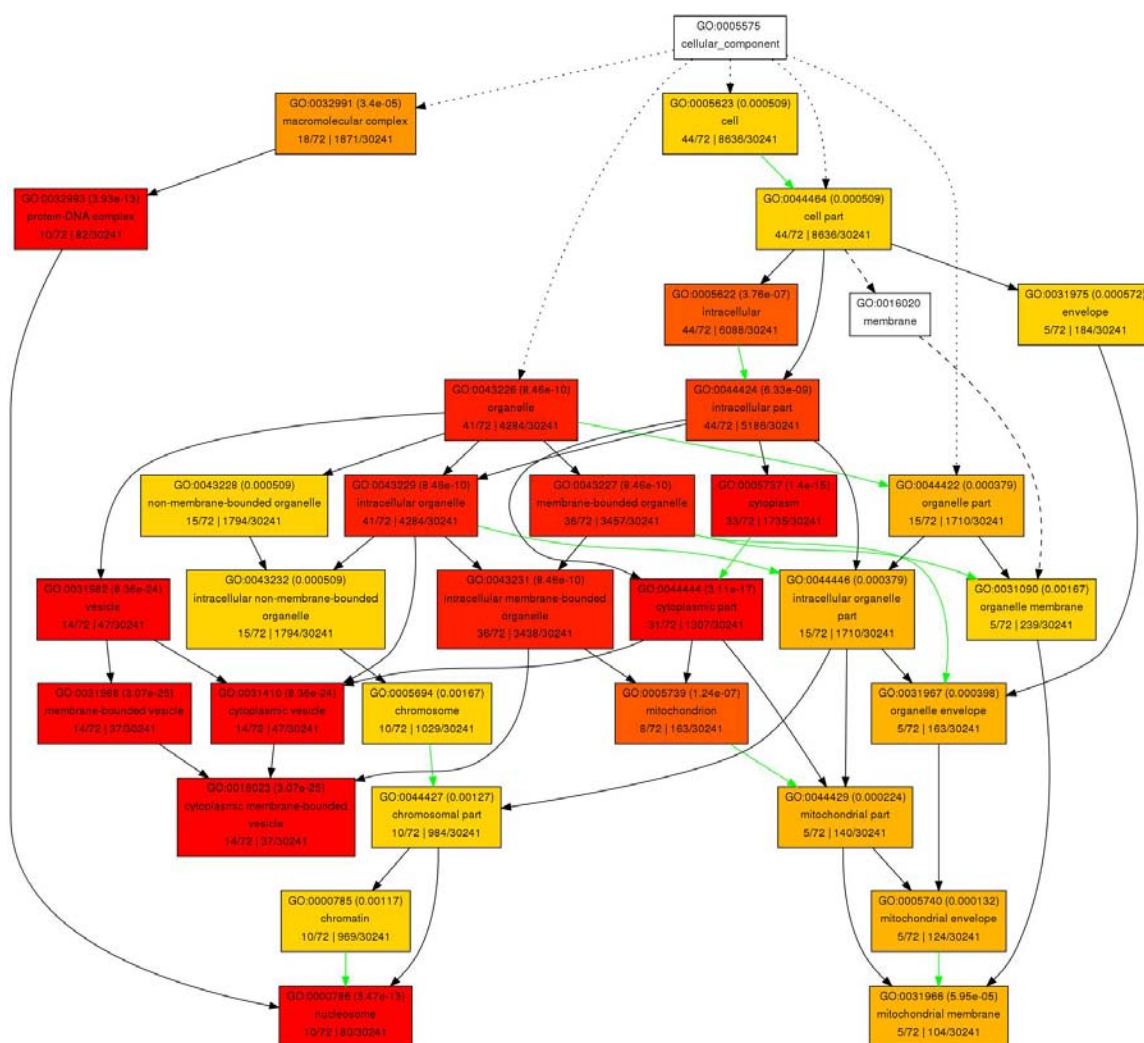


Figure 14 Hierarchical tree graph of overrepresented GO terms in cellular component category generated by SEA.

This figure shows the significant cellular component GO annotations for chromatin associated putative phosphoproteins.



Table 4 Unique putative phosphoprotein identified by Pro-Q Diamond and MALDI-TOF

Location	Spot Number	Accession Number	Protein Name	C.I %	MW (kDa)	PI	Identified Peptides	Ratio
nucleus	107	AAN06860	Putative histone H2A	100.0	14.55	10.3	4	2.36
	109	AAC78105	Histone H3.3	99.8	15.4	11.2	5	1.57
	111	Q9LGH6	Putative histone H2B	100.0	16.46	10.0	12	0.81
	204	NP_001042044	Os01g0152300	100.0	19.29	9.9	12	0.93
	1403	ABA96443	retrotransposon protein, putative, Ty3-gypsy subclass	98.9	162.14	9.1	23	1.67
	1404	NP_001059075*	Glycine-rich RNA-binding protein-like	100.0	33.23	9.3	14	3.29
	1701	EAY83860*	Os12g0611200	100.0	65.73	9.6	27	1.71
	2403	DAA00397	TPA_exp: unknown	100.0	39.29	6.4	12	3.84
	5301	AAP52582*	transposon protein, putative, CACTA, En/Spm sub-	99.8	88.56	8.4	19	1.75
	5606	EAY95913*	hypothetical protein OsI_017146	100.0	95.91	6.7	16	1.09
	6617	ABA92414*	Thiol protease SEN102 precursor, putative,	100.0	40.91	5.7	15	1.88
	8110	O22385*	Glycine-rich protein	99.6	16.02	7.8	6	0.63
	8205	BAD67781	Putative WRKY DNA-binding protein	98.0	28.48	10.3	10	2.58
	8501	EAZ29051*	hypothetical protein OsJ_012534	96.9	92.32	9.7	17	0.77
ribosome	203	EAZ37303	hypothetical protein OsJ_020786	100.0	35.73	10.5	13	1.71
	1202	AAP92747	ribosomal L9-like protein	100.0	21.34	9.6	12	1.50
mitochondrion	2301	Q8VXC7	Voltage-dependent anion channel	100.0	29.58	8.6	12	0.54
	5304	EAZ13326	Putative 36kDa porin II	100.0	41.49	5.7	11	1.52
	6604	P15998	ATP synthase alpha chain, mitochondrial (EC	100.0	55.25	5.9	16	2.38
	6609	P15998	ATP synthase alpha chain, mitochondrial (EC	100.0	55.25	5.9	18	2.31
membrane	4213	EAZ25964	hypothetical protein OsJ_009447	99.7	66.37	7.8	16	0.98
	4303	EAY94429	B0812A04.3 protein	100.0	30.62	6.6	16	2.80
	5302	EAY94429	B0812A04.3 protein	100.0	30.62	6.6	11	1.80
	6403	BAB89823	embryonic abundant protein-like	100.0	29.08	5.7	10	0.86
cytoplasm	1311	O64937	Elongation factor 1-alpha (EF-1-alpha)	97.9	49.25	9.1	7	0.26
	3203	EAZ06044	hypothetical protein OsI_027276	100.0	38.48	5.9	11	0.67
	3205	EAZ06044	hypothetical protein OsI_027276	100.0	38.48	5.9	12	0.68
	3209	O04986*	Non-symbiotic hemoglobin 1 (rHb1) (ORYsa GLB1a)	100.0	18.43	6.9	7	0.56
	3501	A1YR13*	glyceralde-3-phosphate dehydrogenase	100.0	36.54	7.7	17	1.02
	3504	AAP54418	N-acetyl-gamma-glutamyl-phosphate reductase,	100.0	45.45	8.5	5	0.84
	5205	EAZ06044	hypothetical protein OsI_027276	100.0	38.48	5.9	13	0.70
	8115	BAD32133*	putative receptor-like protein kinase 4	96.6	72.75	6.3	16	1.21
cytosol	4509	Q8W424	26S proteasome regulatory particle non-ATPase	100.0	34.87	6.3	17	1.08
extracellular region	2207	P93442	Expansin-A4 precursor (OsEXPA4) (Alpha-expansin-	96.6	25.87	8.1	5	5.78
unknown	1206	EAZ29098	hypothetical protein OsJ_012581	100.0	36.01	9.3	8	0.40
	3204	JC7138	alpha-amylase (EC 3.2.1.1) isozyme III - rice	100.0	48.71	5.3	6	0.54
	3413	EAZ05518	hypothetical protein OsI_026750	99.9	29.82	8.2	8	1.03
	4406	EAZ41445	hypothetical protein OsJ_024928	99.9	29.81	8.2	12	0.61
	6301	BAD61316	putative glutathione transferase F4	100.0	25.29	5.7	12	1.02
	6406	Q9FTN5	Putative isoflavone reductase homolog IRL	100.0	33.48	5.7	18	0.75
	6409	Q7XKT9	OSJNBa0022H21.18 protein	100.0	32.08	5.5	13	0.67
	6612	EAZ38461	hypothetical protein OsJ_021944	98.1	45.73	5.5	13	1.14
	7112	BAC15855	Putative dimethylaniline monooxygenase	99.6	53.01	5.5	14	0.42
	8402	Q7G765	Probable NAD(P)H-dependent oxidoreductase 2	100.0	35.76	5.4	11	0.54
	8508	JC7138	alpha-amylase (EC 3.2.1.1) isozyme III - rice	98.8	48.71	5.3	7	0.33

Spot Number: The spot number was given by computer based on spot excision order.

Accession Number: Protein accession number

MW/PI: Predicted molecular mass and pI.

Identified Peptides: Number of peptides that matched with the identified protein in mass analyses.

C.I. %: Cross Confidence Interval %. Over 95% represents high confidence identification.

\*: Proteins without GO annotation. The cellular localization was predicted by LOCtree or PSORT, two sub-cellular prediction programs.

Ratio: The normalized intensity of Pro-Q Diamond divided by SYPRO Ruby intensity.

## **Discussion**

### **Putative phosphoproteins co-purified with chromatin revealed by Pro-Q Diamond stain**

Protein phosphorylation plays an essential role in multitude biological processes in plants, including developmental regulation and hormone responses. To investigate the phosphoproteomic map of plant chromatin, the rice chromatin proteins have been examined using Pro-Q Diamond Phosphoprotein in Gel Stain and performed MS/MS analysis of the stained proteins using a MALDI-TOF/TOF mass spectrometer. One significant advantage of this method is that one can follow the quantitative and qualitative changes of a particular putative phosphoprotein spot in a time course and under different treatments. Special attention can be applied to particular protein spots that are interesting. Obviously, comparing SYPRO Ruby gel images and Pro-Q Diamond gel images, many protein spots that were heavily stained with SYPRO Ruby were not stained or weakly stained by Pro-Q Diamond dye. Conversely, many spots that were weakly stained by SYPRO Ruby were heavily stained by Pro-Q Diamond dye (Figure 4.1). Since SYPRO Ruby stain intensity has a linear correlation to protein concentration in a broad range (Berggren et al., 1999), the results suggest that Pro-Q Diamond dye stain intensity is not proportional to protein concentration. One possibility is that the stain intensity of

the Pro-Q Diamond dye is proportional to the level of phosphorylation although further tests are required to confirm this.

It is well known that many chromatin associated proteins may have extensive posttranslational modifications or alternative splicing thus generating multiple protein spots on 2-DE gels. Indeed, mass spectrometric analysis reveals that many distinct protein spots share the same protein identity. For example, glycine-rich RNA binding protein-like protein was identified in spots 1403 and 1404; cationic peroxidase was found in spots 4306, 5407, and 6403; glyceralde-3-phosphodehydrogenase was found in spots 2506 and 3501; and a hypothetical protein (OsJ\_024928) was found in spots 3401, 3413 and 4406. Interestingly, these proteins are all stained by Pro-Q Diamond stain, suggesting that they can be modified by phosphorylation. Indeed, among 154 putative phosphoproteins identified with high confidence, most of them have been reported that either the proteins themselves or their close family members are regulated by phosphorylation in published literatures, including, for example, H3-maize, H3.3, H2A, H2B, 26S proteasome regulatory subunit 8 (Rivett et al., 2001), embryonic abundant protein (Chitteti and Peng, 2007a), ribosome subunit L9 (Kruiswijk et al., 1978), Glycine rich protein (Vilardell et al., 1990), WRKY DNA-binding protein (Yang et al., 1999), etc. These reports substantiates that the Pro-Q Diamond dye is a useful tool in the identification of candidate phosphoproteins on 2-DE gel effectively and it provides new insight into the role of chromatin associated phosphoproteins in plants.

Additionally, GO analysis revealed that many of the putative chromatin associated phosphoproteins were located in membrane-bounded organs and nucleosome, chromatin. Interestingly, those phosphoproteins were highly involved in oxidation reduction, nucleosome assembly, nucleosome organization, DNA packaging, DNA

conformation change, and chromatin assembly and disassembly. Thoses results suggest that phosphoprotein not only plays a key role in regulating development, signal transduction, and response to external and endogenous stimuli, but also probably is a key regulator of nucleosome assembly and chromatin conformation organization.

Although Pro-Q diamond is believed to be one of the best methods to investigate the phosphoproteome and quantitation, it is very difficult for us to pinpoint which protein is phosphorylated in specific spot because many spots have more than one protein being identified. Many chromatin binding proteins, including histones, are from a gene family with closely related members, which lead to overlay of proteins in the same spots and result in difficulty in applying Pro-Q Diamond stain for phosphoprotein identification. Specific phosphorylated peptides enrichment, which could collect and enrich the phosphorylated peptides for mass spectrometry analysis and eventually obtain precise modified residues, might be a solution to overcome this issue.

CHAPTER V

CHROMATIN REORGANIZATION AND HISTONE POST-TRANSLATIONAL  
MODIFICATIONS DURING CELL WALL DEGRADATION AND  
REGENERATION IN *ORYZA SATIVA*

**Abstract**

The plant cell wall is rigid structure deposited outside the cell membrane. The highly dynamic structure not only acts as the first barrier of defense against biotic, abiotic stresses, but also controls cell growth and provides mechanical and structural support to plants. In an attempt to study the correlations between cell wall degradation & regeneration and chromatin organization, we applied nucleus microscopy examination and chromatin decondensation assay to observe cell development and chromatin conformational changes. We find that removal of cell wall stimulates cell wall synthesis from multiple sites in protoplasts instead of from a single location as in cytokinesis. And most interestingly, the removal of the cell wall leads to substantial chromatin reorganization. Cell wall regeneration was started and at the meantime chromatin lost compaction in protoplasts and recovered the condensed heterochromatin in terms of reforming chromocenters. Histone post-translational modification studies using both Western blots and isotope labeling assisted quantitative mass spectrometry analyses reveal that the chromatin decondensation is associated with hyperacetylation at H3K18 and H3K23, suggesting a possible role of these two modifications in chromatin organization. The cell wall removal and regeneration device is an excellent system to

study the molecular mechanisms underlying cell wall synthesis and has provided novel insight into intensive cross-talks between cell wall and epigenetic regulations in plant cells.

## Introduction

The mature plant cell can be separated from their original tissue by cell wall degrading enzymes, resulting in the formation of a large population of protoplast cells. During the cell wall removal process, the cells undergo remarkable changes in their pattern of gene expression and lose their differentiated state (Zhao et al., 2001). Following nutritive culturing, the protoplasts can re-enter the cell cycle, recover cell walls and eventually form new plantlets (Damm and Willmitzer, 1988). The plant protoplast is an attractive experimental system to investigate the biochemical and molecular basis for the cell wall removal induced cell dedifferentiation and cell wall regeneration. The protoplasts were characterized by the acquisition of totipotency, with a new balance between the euchromatin that is permitted to transcribe, and heterochromatin which is largely repressive. In mature plant, it's easy to recognize the highly condensed chromatin by microscopy after DAPI-staining nuclei. Zhao *et al.* (Zhao et al., 2001) has discovered that the DAPI-positive domains (chromocenter) became decondensed during tobacco protoplasts preparation. This phenomena has also been found in *Arabidopsis thaliana* (Tessadori et al., 2007) and *Cucumis sativus* (Ondrej et al., 2009). Tessadori *et al.* (Tessadori et al., 2007) discovered that in *Arabidopsis* protoplasts the heterochromatin decondensation was accompanied with all major repeats relaxation including centromeric, pericentromeric and 5S rDNA repeats. It has also been reported that the subtelomeric repeats (DNA type I repeats) were involved in heterochromatin

disassembly (Ondrej et al., 2009). All these studies above ascribed the chromatin change to cell dedifferentiation, it is not yet tested whether removal of cell wall itself could lead to chromatin reorganization without the involvement of cell dedifferentiation. Using dedifferentiated cells, such as well established suspension culture cells, to isolate protoplasts and examining the chromatin change may answer this question.

It is well known that methylation, acetylation, and other post-translational modifications of histone proteins at different amino acid residues comprise the histone codes that switch on and off genes within cell development (Martin and Zhang, 2005; Chen and Tian, 2007; Cheng and Zhang, 2007). Arney and Fisher (Arney and Fisher, 2004) stated that epigenetic modifications play a critical role in chromatin organization and gene expression during cell differentiation in yeast, *Drosophila* and mammalian cells. Chromatin and histones have been studied in plants. We have purified chromatin from rice in a large scale and examined the chromatin associated proteome and phosphoproteome using both 2-DE gel and shotgun proteomics approaches (Tan et al., 2007). Most of the histones and histone variants were identified. We also studied the rice nuclear proteome in endosperm (Li et al., 2008). Johnson *et al.* (Johnson et al., 2004) observed that H3K27<sub>ME2</sub> and H3K36<sub>ME</sub> were frequently found together in *Arabidopsis*. Zhang *et al.* (Zhang et al., 2007) did an extensive study of core histone modifications in *Arabidopsis*. Several unique plant post translational modifications were identified. In addition, Bergmüller *et al.* (Bergmuller et al., 2007) examined the post translational modifications in *Arabidopsis* H2B-variants. Smith *et al.* (Smith et al., 2003) developed a highly efficient isotope labeling method to quantitatively analyze histone H4 N-terminal acetylation, in which deuterated acetic anhydride was used to acetylate free lysine

residues. The histone acetylation level change between two samples was revealed by comparing the deuterated and protiated acetyl groups, respectively.

Here, we investigated the *Oryza Sativa* protoplasts development when devoid of cell walls by microscopic examinations. Strikingly, we found that the protoplasts were induced to start cell wall regeneration from multiple sites of a protoplast simultaneously and at the meantime the chromatin underwent dramatically conformational changes. This structural shifting was also confirmed by chromatin decondensation assay. In addition, differential histone modifications are detected and validated by immunoblot assay and mass spectrometry with isotope labeling.

## **Materials and methods**

### **Cell culture**

Rice (*Oryza sativa*) cell culture have been described previously (Lee et al., 2004). The suspension cells were grown at 24°C with constant shaking on a gyratory shaker at 150 rpm in liquid B5 organic medium (pH 5.7) supplemented with 20 g/L sucrose, 0.5 g/L MES, 2.0 mg/L 2,4-dichlorophenoxyacetic acid (2,4-D), 2 g/L casein enzymatic hydrolysate and 0.005% pectinase as reported (Lee et al., 2004). Synchronized growth of cultures was achieved.

### **Protoplast isolation and culture**

The protoplasts were generated using a method mentioned in chapter 3. After enzyme digestion, protoplasts were obtained and adjusted to a final concentration of  $5 \times 10^5$  / ml, and cultured in the darkness using protoplast medium (PTM) (0.4 M mannitol, 20 mM CaCl<sub>2</sub>, 0.125 mM MgCl<sub>2</sub>, 0.5 mM MES, and 2 g/L N-Z-Amine A in B5 organic medium plus 2.0 mg/L 2,4-D at pH 5.6) (Yamada et al., 1986).



### **Evaluation of new cell wall formation**

New cell wall formation was evaluated by monitoring the fluorescence of Fluorescent Brightener 28 (Calcofluor White M2R, Fluostain I, Sigma Aldrich, St. Louis, MO) using a confocal laser scanning microscope (CLSM) Zeiss Axiovert 200 M (Zeiss, Germany). Protoplasts were pelleted at  $120 \times g$  for 5 min, washed once with PTM, and stained for 10 min with 0.001% (W/V) Fluorescent Brightener 28 in PTM medium. After removing excess dye, the stained protoplasts were washed once with PTM, and then observed at an excitation wavelength of 492 nm and an emission wavelength of 520 nm (Yamamoto et al., 2004; Suzuki et al., 2007).

### **Nuclei isolation and assessment of the integrity of nuclear fractions**

For the microscopic examination of suspension cell nuclei and protoplast nuclei, both samples were fixed using 1% formaldehyde. Suspension cell (treated with PTM medium for 9 hours) nuclei were released using a prechilled blender at low speed in Nuclei Isolation Buffer (NIB) buffer. The homogenized slurry was filtered through two layers of Miracloth and the nuclei were collected by centrifugation at 500 g for 10 min at 4°C. Protoplast nuclei were isolated from rice protoplasts as described in chapter 3. The integrity of freshly obtained nuclei was assessed by staining with DAPI. For each sample, thirty nuclei were randomly selected for the statistical analysis of chromocenters.

### **Chromatin decondensation assay**

The nuclei preparations from both suspension cells and protoplasts were resuspended in 1.2 ml digestion buffer (10 mM Tris-HCl, pH 7.5, 15 mM NaCl, 60 mM KCl, 4 mM CaCl<sub>2</sub>, 0.1 mM PMSF, 0.5 mM Spermidine, 0.15 mM Spermine, 1 mM  $\beta$ -mercaptoethanol, Protease Inhibitor Cocktail). Micrococcal nuclease (40 units/ml, Roche

Molecular) was then added for various time periods. The reaction was terminated by shifting the reaction to 4 °C and mixing with an equal amount of stop solution (20 mM EDTA, 400 mM NaCl, 2% Triton, 0.2% sodium deoxycholate, 0.2% SDS) for 15 min on ice. After 10 min centrifugation at  $15,000 \times g$ , the supernatant was collected and extracted with phenol/chloroform/isoamyl alcohol (25:24:1). The DNA was precipitated with ethanol, dissolved in 50  $\mu$ l TE buffer, treated with RNase A, and resolved on a 2% agarose gel.

### **Immunoblot analysis**

Proteins extracted from suspension cells, protoplast cells, and protoplast-derived cells (48h) were separated by 15% SDS-PAGE gel and processed for quantitative western blots. Anti-histone H3 antibodies (Rb pAb, lot 31949, Upstate) were used to estimate H3 quantity by western blots and consequently to adjust the protein loadings to achieve equal histone H3 loading in all subsequent Western blots. Immunoblots against anti- H3K18<sub>AC</sub> (Rb pAb, lot 676244, Abcam), anti-H3K23<sub>AC</sub> (Rb pAb, lot 746169, Abcam), anti-H3K14<sub>AC</sub> (Rb pAb lot 30020, Upstate), anti-H3K27<sub>AC</sub> (Rb pAb lot 26817, Millipore), anti-H3K4<sub>ME3</sub> (Ms mAb lot 772308, Abcam), anti-H3K9<sub>ME3</sub> (Rb pAb lot 27759, Millipore), anti-H3K27<sub>ME3</sub> (Rb pAb lot DAM1703508, Millipore) and anti-H3K36<sub>ME3</sub> (Rb pAb lot 716801, Abcam) were carried out using a standard Western blot procedure (Tan et al., 2007).

### **Histone H3 isolation**

Acid extraction of histones was performed as previously described (Shechter et al., 2007) with slight modifications. Briefly, the nuclei pellet was re-suspended in 0.4N sulphuric acid and incubated on a rotator for two hours. After centrifugation at  $16,000 \times g$

for 10 min at 4°C, the supernatant fraction that contained acid-soluble proteins was retained. The acid-soluble proteins were precipitated by adding TCA at a final concentration of 33%. The pellet was washed twice with ice-cold acetone and air-dried. The acid-soluble proteins were dissolved in ddH<sub>2</sub>O and quantified using the Bradford method. The proteins were separated by a 15% SDS-PAGE gel and stained with Coomassie Blue dye. The locations of core histone H3 was confirmed by immunoblotting with anti-histone H3 (Rb pAb, lot 31949, Upstate). The histone H3 bands were manually chopped and collected.

#### **Histone post-translational modifications identification and acetylation quantification by mass spectrometry**

The gel bands containing histones were dried with reduced vacuum and in-gel derivatization of unmodified lysines using d<sub>6</sub>-acetyl anhydride in d<sub>3</sub>-acetyl acid solution was performed as previously described with a minor modification (Smith et al., 2003). The gel bands were sequentially destained with a mixture of methanol and water (50:50), washed by water, punched into fine powder, neutralized to pH 7.8 by 100 mM ammonium bicarbonate, and then digested with trypsin (Roche Applied Science) for eight hours (enzyme : protein is ~1:100). After digestion, the peptides were extracted with acetonitrile, dried with reduced vacuum, and then redissolved in 0.1% formic acid for LC/MS/MS analysis. LC/MS/MS was carried out by nano-electrospray on Waters QTOF Ultima instrument which was coupled with a Waters capillary HPLC. A 140 minute gradient (mobile phase A: 0.1% formic acid; B: 0.1% formic acid in acetonitrile) with a flow rate of 5 µl/min was run through a splitter. After splitting the flow (~300 nl/min) was run through a Waters nano C18 column (150 mm x 75 µm) to the nano-ESI source. 2.6 KV capillary voltage and 45 V cone voltage were optimized for nano-electrospray.

Modifications of histones were manually determined with the aid of the Prospector program ([prospector.ucsf.edu/product](http://prospector.ucsf.edu/product)). Quantitative LC/MS/MS analyses of histone acetylation were performed by derivatization of unmodified lysines using d6-acetoanhydride as previously described (Smith et al., 2003) and the approaches utilized consisted of a full-mass scan mode in the mass range from 300 to 1000 and an MS/MS experiment with selected ion at  $m/z$   $537 \pm \sim 4$  Da bracketing the tryptic K18/K23 peptides with natural and chemical acetylation through a widely opened quadrupole (LM & HM = 2.3) for collision activated dissociation (collision gas: Ar and collision energy: 26 V) analysis. The QTOF instrument was calibrated with the MS/MS fragmentation ions of the peptide-Glu-Fib resulting in less than 20 ppm mass accuracy for precursor ions when a lock-mass was applied and  $\pm 0.02$  Da for product ions. The acetylation proportions of K18 and K23 were calculated by the intensity relative to the total peak value of diacetylated peptide (K18/K23 di-Ac), K18 or K23 monoacetylated peptide (mono-Ac), and K18 and K23 un-acetylated peptide (no-Ac).

## Results

### Cell wall regeneration in rice (*Oryza sativa*) protoplasts

Protoplast preparation from rice suspension cells was achieved after a 9 h digestion with 2.5% Cellulase RS and 1% Macroenzyme R10. Due to the unique cell wall structure of plants in grass family, multiple hours of enzyme digestion is required to completely remove the cell wall (Yamada et al., 1986; Tan et al., 2007). Protoplasts released after 9 h of enzyme treatment lacked any detectable cell wall structure (Figure 5.1A). Within 2 h, distinct fluorescent spots appeared on the surface of protoplasts. While cell wall synthesis during cytokinesis occurs in only one spot - the center of the

phragmoplast, the cell wall re-synthesis in protoplasts occurred at multiple positions simultaneously as shown in single layer Figure 5.1B and Figure 5.1C, and 3-D projection images of multiple layers in Figure 5.1G and the images of different layers (Figure 5.1H to L), suggesting that novel mechanisms might be involved in the cell wall re-synthesis process. With an increase in cell culture time, the fluorescence became stronger and spread over the entire cell surface (Figure 5.1). The 48h protoplast-derived cells (Figure 5.1F) were differed in shape from the freshly isolated protoplasts which stayed relatively spherical and smooth on the surface (Figure 5.1A), indicating that the cell wall regeneration progressed rapidly and cell wall synthesis at multiple sites could somehow integrate to form the wall cage appropriately.

### **Removal of cell wall stimulates chromatin reorganization**

It was reported that isolation and culture of protoplasts led to chromatin decondensation/reorganization and the chromatin change was ascribed to cell dedifferentiation (Zhao et al., 2001; Tessadori et al., 2007; Ondrej et al., 2009). With numerous experimental replicates, we found that removal of the cell wall from rice suspension culture cells, which had undergone cell dedifferentiation, was also associated with chromatin decondensation/reorganization (Figure 5.2). We randomly selected 30 nuclei from each sample to perform a statistical analysis of chromocenters. In suspension cells, most of the nuclei (over 90%) have multiple distinct chromocenters when stained with DAPI (Figure 5.2A). The average chromocenter number was  $8.2 \pm 2.9$ . This large deviation was because some cells (about 5%) had no chromocenters at all probably because these cell cultures were not synchronized and cells were in different cell division stages. In freshly isolated protoplast nuclei, the distinct speckles of chromocenters

disappeared in most protoplasts. The average chromocenter number was  $0.7 \pm 0.6$ . For the few chromocenter like structures still observed, they were blurry instead of being sharp (Figure 5.2B). After a 48 h culture, chromocenters recovered in almost all surviving cells (over 90% of the cells re-synthesized a cell wall). Interestingly, the average number of chromocenters was  $14.3 \pm 3.7$ , higher than the original suspension cells. Meanwhile, the speckles appeared to be smaller when compared with the original suspension cells (Figure 5.2C). These results indicated that when the cell wall was removed, complete or partial chromatin disassembling was concomitant. When the cell wall was recovered, chromocenters reassembled. It appeared that the reassembled chromocenters of protoplast-derived cells might not be exactly identical to the chromocenters of suspension cells because the number had increased and the speckle size had decreased.

To further examine the chromatin change in protoplasts, the nuclei from both suspension cells and protoplasts were treated with MNase. MNase cut DNA in regions without the binding of core histones, therefore, generating DNA fragments with the size intervals of a nucleosome when separated on an agarose gel, which appeared as ladders of DNA bands. In highly packaged chromatin, the DNA digestion would be slower than decompacted chromatin because the enzymes had more difficulty to gain access to the highly packaged chromatin. Consequently, the chromatin structural status can be revealed by MNase digestion. As shown in Figure 5.2E, MNase digestion of suspension cell chromatin was clearly slower than the digestion of protoplast chromatin, suggesting a decondensation of chromatin in response to the removal of cell wall.

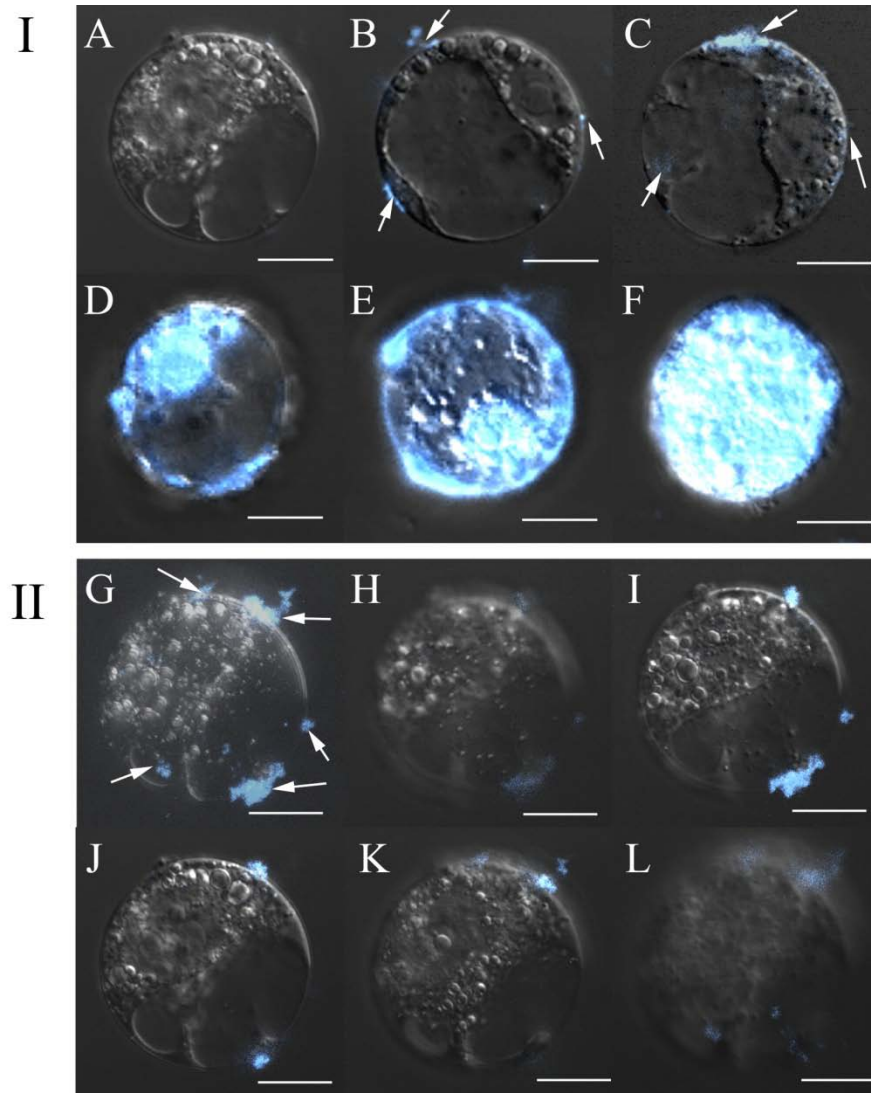


Figure 15 Fluorescence microscopy images of cultured rice (*Oryza sativa*) protoplasts

CLSM was used to observe the polysaccharide specific fluorescent dye, Fluorescent Brightener 28, using an excitation at 492 nm and emission at 520 nm. Part I, Fresh (0h) (A) along with 2h (B), 4h (C), 6h (D), 12h (E), and 48h (F) cultured samples of protoplasts were examined. The arrows point at the positions of cell wall syntheses. Part II, (G) confocal 3-D projection image of a protoplast cultured for 2hrs; (H~L) multiple scanning layers of the protoplast shown in G. The magnification is revealed by the scale bar, 10 $\mu$ m.

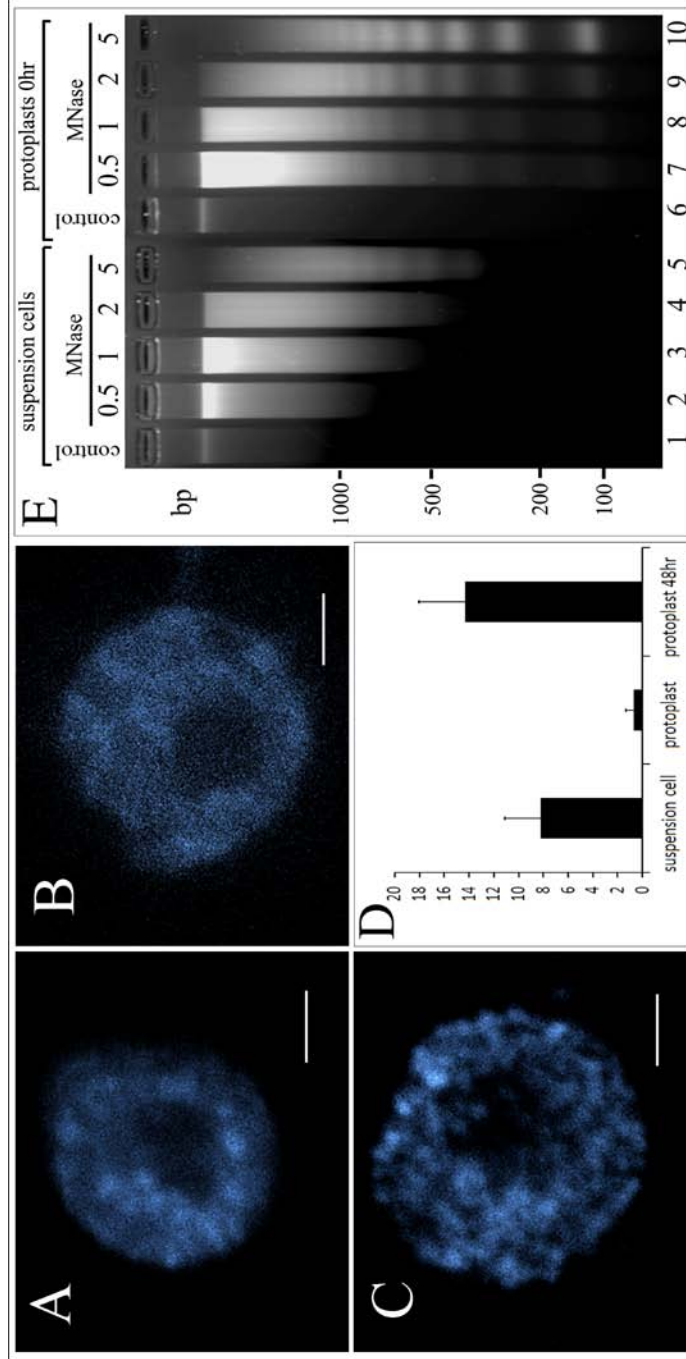


Figure 16 Protoplast chromatin decondensation revealed by DAPI stain and MNase assay

Chromocenters were revealed by DAPI stain as bright sharp speckles in nuclei. The scale bar represents 2 $\mu$ m. (A), DAPI-stained nuclei in suspension cell. (B), DAPI-stained nuclei in protoplast-derived cells (48 h). (D), Statistical data of chromocenter numbers in suspension cells, fresh protoplasts and protoplast-derived cells (48 h). Thirty nuclei were randomly selected for the analysis. (E), Chromatin from freshly prepared protoplasts showed increased sensitivity to MNase digestion, MNase (40units/ml) was applied to nuclei isolated from suspension cells (lanes 1-5) and nuclei isolated from freshly obtained protoplasts (lanes 6-10) for the indicated time periods (min). DNA was extracted and resolved by a 2% agarose gel. Control (lanes 1 and 6) is DNA prepared from nuclei without MNase treatment.



### Characterization of histone H3 post-translational modifications

While performing LC/LC-MS/MS analysis in protein identification, we observed many H3 post-translational modifications. A total of 11 modified histone H3 and H3.2 peptides, including 5 acetylation sites (H3K14, H3K18, H3K23, H3K27, H3.2K36) and 3 methylation sites (H3K27, H3K36, H3.2K36), were mapped (Table 5.1).

Monomethylation was identified on residue H3K36. Acetylation, mono-, and dimethylation were identified on residue H3K27. Acetylation was identified on residues H3K14, H3K18, and H3K23. In addition, histone H3.2 peptides were modified at Lys 36 (acetylation, di-, and trimethylation). A few selected examples are presented in Figure 3. A peptide with the sequence STGG<sup>14</sup>K<sub>AC</sub>APR from the N-terminal of H3 was obtained in which lysine 14 was determined to be acetylated (Figure 3A), and lysine 23 acetylation at peptide KQLAT<sup>23</sup>K<sub>AC</sub>AAR (Figure 3B) and di-acetylation on lysine 18 and lysine 23 at peptide <sup>18</sup>K<sub>Ac</sub>QLAT<sup>23</sup>K<sub>AC</sub>AAR (Figure 3C), were also identified. The product ion spectra presented a complete y-ion series for an unambiguous sequence assignment. These acetylation modifications assignments are based on the identification of a unique ion ( $m/z$  126.1), which is a further fragment ion induced by the loss of NH<sub>3</sub> from the acetylated lysine immonium ions at  $m/z$  143.1, typical of an acetylated lysine residue, along with a 42 mass unit added to lysine 14, lysine 18 or lysine 23. Moreover, the detection of a series of y ions corresponded to the acetylated peptide fragmentation rather than the tri-methyl group neutral loss of y<sub>i</sub>-59 ions that corresponded to the tri-methylated peptide fragmentation (Zhang et al., 2007). A dimethylated peptide (<sup>27</sup>K<sub>ME2</sub>SAPATGGVK) with a double-charged precursor ion at  $m/z$  472.31 in which K27 was illustrated to be di-methylated was also identified (Figure 3D). The di-methylated lysine residue generated a marker ion at  $m/z$  129 and it was also observed that a 28 mass

unit was added to lysine 27. The ion spectrum of a double-charged precursor ion at  $m/z$  689.45 was shown here to reveal lysine 36 tri-methylation at histone H3.2 peptide SAPTTGGV<sup>36</sup>K<sub>ME3</sub>KPHR (Figure 3E). The detection of neutral loss of C-terminal ions,  $y_6-59$ ,  $y_8-59$ ,  $y_9-59$ ,  $y_{10}-59$  and  $y_{11}-59$ , which corresponded to the fragmentation containing a tri-methylated lysine residue, as well as a 42 mass unit added to lysine 36 confirmed a tri-methylation at lysine 36.

Table 5 LC-MSMS analysis of histone H3

Peptide	Detected PTM <sup>a</sup>	MW <sup>b</sup>	$m/z$	Mascot score
<sup>10</sup> STGGKAPR <sup>17</sup>	Lys <sup>14</sup> AC	814.43	408.28 (2+)	59
<sup>27</sup> KSAPATGGVK <sup>36</sup>	Lys <sup>27</sup> ME	928.53	465.33 (2+)	34
<sup>27</sup> KSAPATGGVK <sup>36</sup>	Lys <sup>27</sup> ME2	942.55	472.34 (2+)	54
<sup>27</sup> KSAPATGGVK <sup>36</sup>	Lys <sup>27</sup> AC	956.53	479.32 (2+)	45
<sup>18</sup> KQLATKAAR <sup>26</sup>	Lys <sup>23</sup> AC	1027.61	514.83 (2+)	39
<sup>18</sup> KQLATKAAR <sup>26</sup>	Lys <sup>18</sup> AC, Lys <sup>23</sup> AC	1069.62	535.88 (2+)	35
<sup>28</sup> SAPATGGVKKPHR <sup>40</sup> (H3.2)	Lys <sup>36</sup> AC	1376.75	459.96 (3+)	24
<sup>28</sup> SAPATGGVKKPHR <sup>40</sup> (H3.2)	Lys <sup>36</sup> ME3	1376.88	689.45 (2+)	69
<sup>28</sup> SAPATGGVKKPHR <sup>40</sup> (H3.2)	Lys <sup>36</sup> ME2	1362.77	455.29 (3+)	18
<sup>27</sup> KSAPATGGVKKPHR <sup>40</sup>	Lys <sup>27</sup> AC, Lys <sup>36</sup> ME	1488.85	745.50 (2+)	16
<sup>27</sup> KSAPATGGVKKPHR <sup>40</sup>	Lys <sup>27</sup> ME2, Lys <sup>36</sup> ME	1474.87	738.49 (2+)	61

a) Superscript letters following the position numbers represent the following:  
AC, acetylation; ME, monomethylation; ME2, di-methylation; ME3, tri-methylation.  
b) Calculated molecular weight.

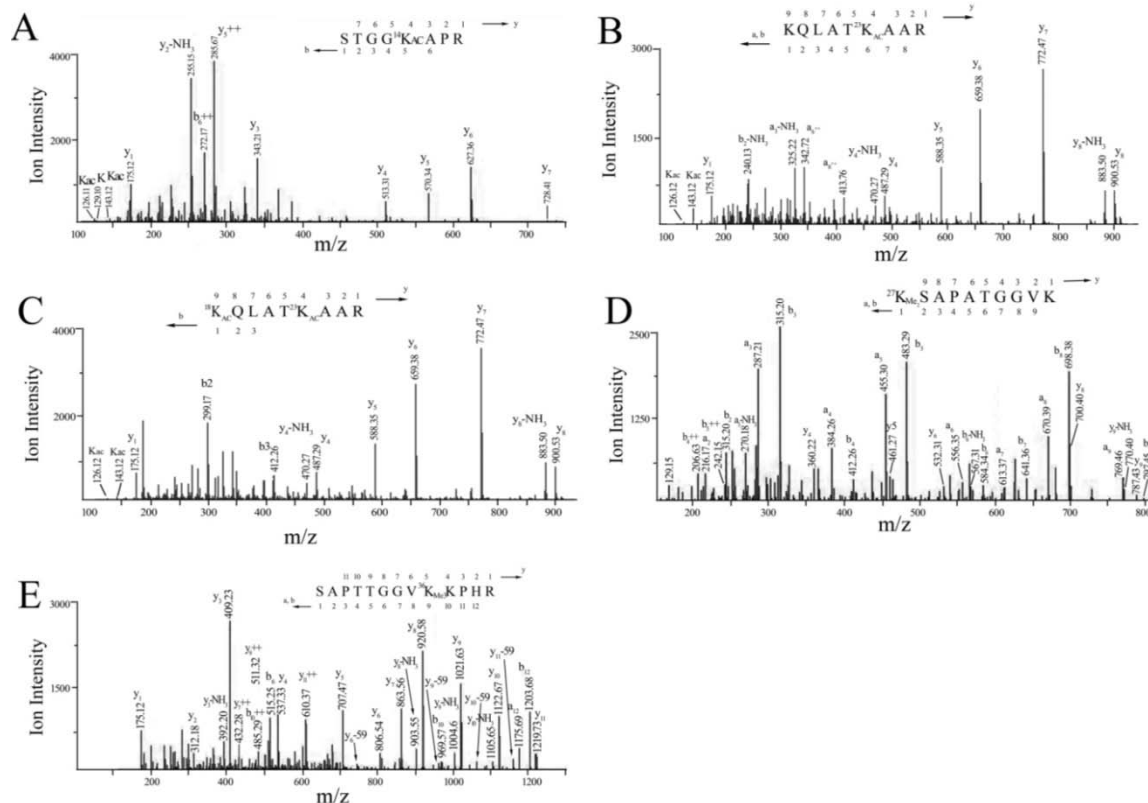


Figure 17 Representative MS/MS mass spectra of H3 and H3.2 modifications

Peaks were labeled according to ion number for b and y ions, respectively. (A) MS/MS spectrum of acetylated peptide STGG<sup>14</sup>K<sub>ac</sub>APR isolated from rice (*Oryza sativa*) histone H3. The immonium-NH<sub>3</sub> ion at *m/z* 126.11 and immonium ion itself at *m/z* 143.10, typical of an acetylated lysine residue, was observed. These observations along with a 42 mass unit added to lysine 14 indicated an acetylation at lysine 14. (B) MS/MS spectrum of acetylated H3 lysine 23 peptide KQLAT<sup>23</sup>K<sub>ac</sub>AAR. (C) MS/MS spectrum of di-acetylated H3 lysine 18 and lysine 23 peptide <sup>18</sup>K<sub>ac</sub>QLAT<sup>23</sup>K<sub>ac</sub>AAR. (D) MS/MS spectrum of di-methylated peptide <sup>27</sup>K<sub>me2</sub>SAPATGGVK. An observation of di-methylation specific ion at *m/z* 126 and a 28 mass unit added to lysine 27 indicated a di-methylation at lysine 27. (E) MS/MS spectrum of tri-methylated peptide SAPTTGGV<sup>36</sup>K<sub>me3</sub>KPHR (H3.2). The neutral loss of a series of C-terminal ions, y<sub>i</sub>-59, corresponding to the fragmentation contains a tri-methylated lysine residue, was observed. These observations as well as a 42 mass unit added to lysine 36 indicated a tri-methylation at lysine 36. Peaks were labeled according to ion number for the b and y ions shown above

### **Histone H3 modification changes associated with cell wall removal**

The substantial chromatin reorganization in response to cell wall removal and regeneration led us to the hypothesis that the histone modification state may also change with cell wall. Proteins extracted from suspension cells, protoplast cells and protoplast-derived cells (48 h) were immunoblotted with anti-H3K18<sub>AC</sub>, anti-H3K23<sub>AC</sub>, anti-H3K14<sub>AC</sub>, anti-H3K27<sub>AC</sub>, anti-H3K4<sub>ME3</sub>, anti-H3K9<sub>ME3</sub>, anti-H3K27<sub>ME3</sub> and anti-H3K36<sub>ME3</sub>. The signal from a western blot of anti-H3 confirmed identical histone H3 loads for each sample (Figure 5.4). As shown in Figure 4, the anti-H3K18 signal almost doubled in protoplasts and protoplasts derived cells compared with suspension cells. The anti-H3K23 signal was barely detected in suspension cells. It had a substantial increase in fresh protoplasts but declined in protoplast derived cells. H3K14 acetylation increased in protoplasts and continued to incline in protoplast derived cells. In contrast, H3K27 acetylation signal was strongest in suspension cells compared with protoplasts and protoplast-derived cells. Anti-H3K4<sub>ME3</sub>, H3K9<sub>ME3</sub>, and H3K36<sub>ME3</sub> had stronger signal in protoplasts and protoplast-derived cells than suspension cells. There was no distinct difference in protoplasts and protoplasts derived cells for these three modifications. Finally, there was no or very little change in antiH3K27<sub>ME3</sub> signal in these three samples. Since we maintained the same hormone concentration in the entire cell wall removal and regeneration process as in suspension culture and the suspension cells were treated with protoplast isolation buffer sans cell wall digestion enzyme, hormone response and osmotic stress response should not contribute to the histone modification change. Thus, the differential histone modifications were due to the removal of the cell wall either directly or indirectly.

Smith and co-authors (Smith et al., 2003) developed an isotope labeling method to quantitatively reveal histone acetylation changes among different samples. We decided to examine differential histone acetylation with this method. In this experiment, deuterated acetic anhydride was used to acetylate all unacetylated lysine residues *in vitro*. In contrast, the *in vivo* acetylated lysine is in protiated forms. Thus, there is a 3 Da difference between protiated and deuterated acetyl groups (42 Da versus 45 Da, respectively). The mass difference is used to determine the endogenous level of acetylation. We adopted this method to quantify the acetylation at H3K18 and H3K23. Histones extracted from purified nuclei were separated by SDS-PAGE (15%) and excised manually. In-gel derivatization and digestion was performed using d<sub>6</sub>-acetyl anhydride and trypsin, respectively. The produced tryptic digests of H3 were analyzed by LC-ESI-MS and MS/MS. The averaged mass spectra for the period in which all types of acetylated H3 (18-26) eluted from the column are presented in Figure 5.5A & 5.5B for protoplasts and suspension cells, respectively. Peak *m/z* 535.8 corresponds to K18 and K23 both naturally acetylated peptide <sup>18</sup>K<sub>AC(H)</sub>QLAT<sup>23</sup>K<sub>AC(H)</sub>AAR; Peak *m/z* 537.3 corresponds to either K18 or K23 naturally acetylated peptide <sup>18</sup>K<sub>AC(H)</sub>QLAT<sup>23</sup>K<sub>AC(D)</sub>AAR or <sup>18</sup>K<sub>AC(D)</sub>QLAT<sup>23</sup>K<sub>AC(H)</sub>AAR; Peak *m/z* 538.8 corresponds to K18 and K23 both chemically acetylated peptide <sup>18</sup>K<sub>AC(D)</sub>QLAT<sup>23</sup>K<sub>AC(D)</sub>AAR. AC(H) is natural acetylation while AC(D) is chemical acetylation by d<sub>6</sub>-acetyl anhydride. Applying the quantitative analyses below (Smith et al., 2003), the relative intensities of di-acetylated ion and mono-acetylated ion were found to be 0.212 and 0.383 respectively in protoplast H3, and 0.067 and 0.309 in suspension cells H3.

$$K18/K23_{Mono\_AC} = I_{537.33} / (I_{535.84} + I_{537.33} + I_{538.85}) \quad (Eq\ 1)$$

$$K18/K23_{Di\_AC} = I_{535.84} / (I_{535.84} + I_{537.33} + I_{538.85}) \quad (Eq\ 2)$$

These results indicated that 21.2% of histone H3 molecules were acetylated at both lysine 18 and 23 and 38.3% were acetylated at either one of the two lysine sites in protoplasts *in vivo*. In contrast, only 6.7% of histone H3 molecules were di-acetylated at lysine 18 and 23 and 30.9% of them were mono-acetylated at either one of the two sites in suspension cells *in vivo*. Weighing these intensities for the number of *in vivo* acetylated lysines indicated the average level of acetylation in protoplast H3(18-26) was 0.807 acetylation per molecule and 0.423 acetylation per molecules for suspension cell H3(18-26). It is obvious that the acetylation level of H3 (H3K18 and H3K23) substantially increased upon removal of the cell wall. Particularly, the di-acetylation level at H3K18 and H3K23 in protoplasts has a three folds increment when compared to that in suspension cells.

While the MS study described above provided useful information about the global distribution of endogenous acetylation level of H3(18-26) peptide in both samples, it was not sufficient to confirm which lysine was modified and provide quantitative information for each modified lysine. Therefore, we carried out MS/MS analyses, in which fragment ions were generated to obtain detailed information of the acetylated peptide. Endogenous acetylation at each of the two lysine residues within the H3 tail was confirmed by measuring the relative ion intensities of protiated versus deuterated fragment ions of the isotopically tagged histone H3(18-26) peptide. Tandem mass spectra revealing the relative ion intensities of fragment ions of fully acetylated H3 (18-26) peptides are shown in Figure 5.5C & 5.5D for protoplast and suspension cells, respectively. Zoomed images of interesting mass peaks are presented on top of the spectra, including  $b_2$ ,  $y_5$ ,  $y_6$ , and  $y_7$ . Quantitative analyses were conducted using the following equations (Smith et al., 2003).

$$K18_{AC} = I_{b_2 (AC\_H)} / [I_{b_2 (AC\_H)} + I_{b_2 (AC\_D)}] \quad (\text{Eq 3})$$

$$K23_{AC} = 1/3 (I_{y5(AC\_H)}/[I_{y5(AC\_H)}+I_{y5(AC\_D)}]+I_{y6(AC\_H)}/[I_{y6(AC\_H)}+I_{y6(AC\_D)}]+I_{y7(AC\_H)}/[I_{y7(AC\_H)}+I_{y7(AC\_D)}]) \quad (\text{Eq 4})$$

Based on the intensities of four expanded ions in Figure 5C, the relative intensity for H3K18<sub>AC</sub> and H3K23<sub>AC</sub> were 0.522 and 0.274 in protoplasts, indicating that 52.2% of H3K18 molecules were acetylated and 27.4% of H3K23 were acetylated. Similar quantitative analyses were applied to suspension cells H3 peptide tandem mass spectrum in Figure 5.5D. We found that 27.5% lysine 18 and 10.5% lysine 23 were modified with acetylation in suspension cells. Obviously, the acetylation levels at both H3K18 and H3K23 had a significant increment in protoplast H3. Specifically, the acetylation level at H3 lysine 18 in protoplasts was twice more than that in suspension cells, and the acetylation at H3 lysine 23 was almost three times more in protoplasts than in suspension cells.

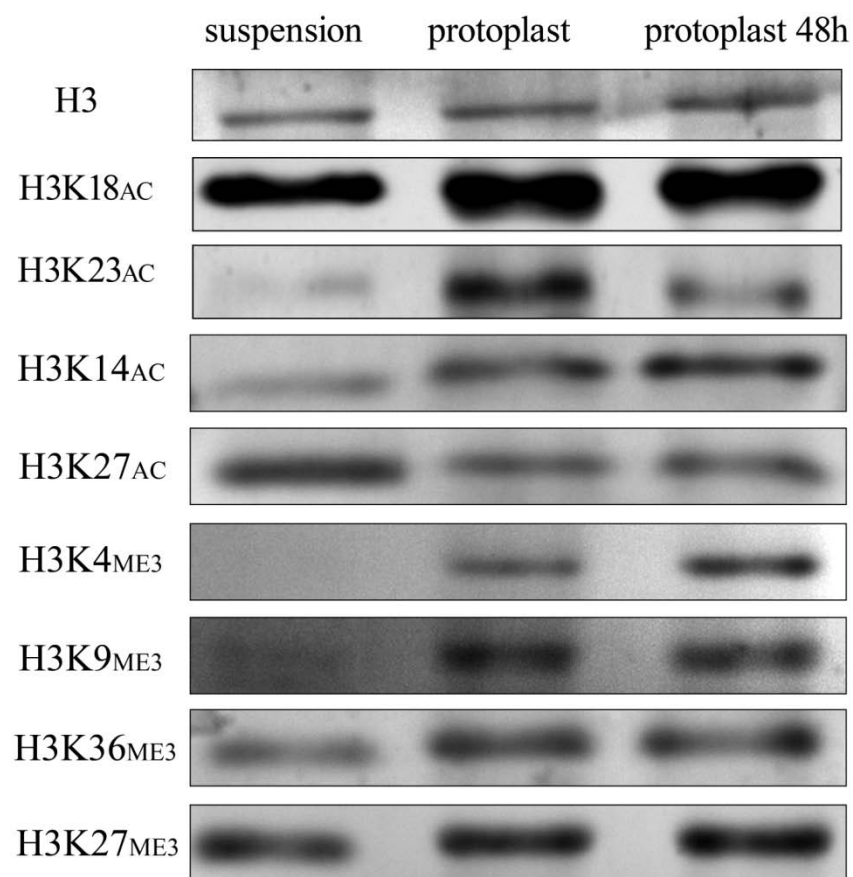


Figure 18 Immunological characterization of histone H3 modification in suspension cells, protoplasts, and protoplast-derived cells (48 h)

Western blots were carried out with standard procedures. The secondary antibodies were conjugated with alkaline phosphatase. Protein sources are indicated on the top and antibodies are indicated on the left.



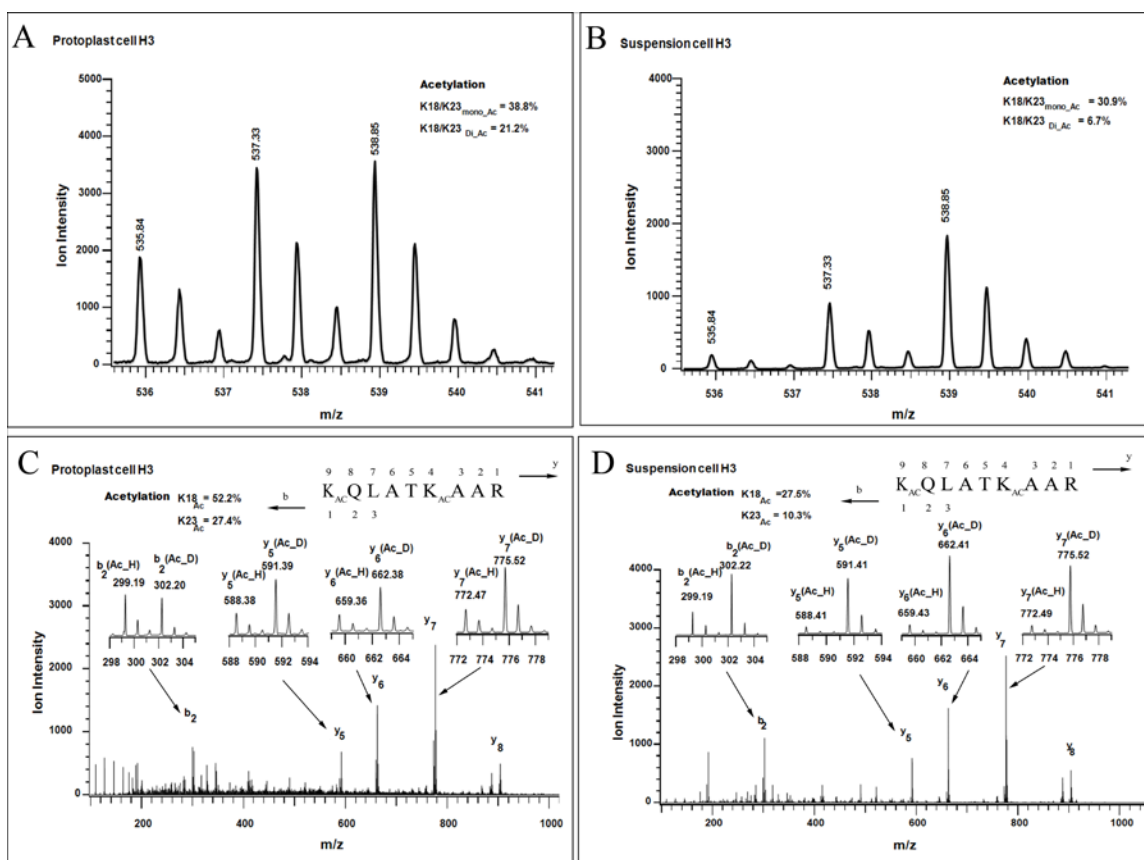


Figure 19 Quantitative comparisons of acetylations at residues K18 and K23 of histone H3 prepared from suspension cells and fresh protoplasts

Endogenous acetylation at each of the two lysine residues within the tail of histone H3 was determined by measuring the relative abundance of protiated versus deuterated fragmentation ions of the isotopically tagged histone H3 peptide  $^{18}KQLAT^{23}KAAR$ . (A & B), mass spectrum showing the mass-to-charge ratio and relative ion intensities of all types of acetylated H3(18-26) fragments. Peak  $m/z$  535.8 corresponds to K18 and K23 both naturally acetylated peptide  $^{18}K_{AC(H)}QLAT^{23}K_{AC(H)}AAR$ ; Peak  $m/z$  537.3 corresponds to K18 or K23 naturally acetylated peptide  $^{18}K_{AC(H)}QLAT^{23}K_{AC(D)}AAR$  or  $^{18}K_{AC(D)}QLAT^{23}K_{AC(H)}AAR$ ; Peak  $m/z$  538.8 corresponds to K18 and K23 both chemically acetylated peptide  $^{18}K_{AC(D)}QLAT^{23}K_{AC(D)}AAR$ . Ac(H) is natural acetylation; AC(D) is chemical acetylation by  $d_6$ -acetyl anhydride. (C & D), tandem mass spectrum showing the relative ion intensities of fragment ions of  $^{18}KQLAT^{23}KAAR$ . Expanded above the spectrum were the set of peaks that correspond to the protiated and deuterated forms of respective ion fragments from the acetylated isotopically labeled H3(18-26) peptide.

## Discussion

### Removal of cell wall and chromatin state

Several recent studies have shown that protoplast culture in *Arabidopsis* leads to chromatin decondensation or reorganization (Zhao et al., 2001; Tessadori et al., 2007; Ondrej et al., 2009). Since protoplast culture is mainly used for cell dedifferentiation and plantlet regeneration studies and genome reprogramming is expected to be required for cell dedifferentiation, the observed chromatin change in protoplasts was ascribed to cell dedifferentiation. Since our protoplasts were isolated from dedifferentiated suspension culture cells with active cell division activities and we maintained the same hormone concentration as in the suspension culture in the entire protoplast isolation and culture process, the substantial chromatin decondensation/reorganization observed should not be due to either cell dedifferentiation or re-entrance into the cell cycle. Moreover, it has been reported that the removal of the cell wall alone may not automatically result in cell dedifferentiation without hormone treatment (Birnbaum et al., 2003; Birnbaum et al., 2005; Lee et al., 2006; Brady et al., 2007). Therefore, cell dedifferentiation is not a factor that affects chromatin state in our experimental system. We treated the suspension cells with protoplast isolation buffer and did not observe any chromatin change, suggesting that the factor that caused chromatin change was enzymatic removal of the cell wall instead of the effect of the buffer. Several lines of evidence support this hypothesis. First, most chromocenters disassembled when the cell wall was removed and reassembled when the cell wall was recovered, indicating that cell wall is critical to chromatin state. Second, many histone modifications displayed substantial change upon removal of cell wall and some of the histone modifications, such as H3K23 acetylation, immediately declined with the recovery of the cell wall. Third, histone variants such as H3.3 and H2A

variant 2 declined upon removal of cell wall and inclined upon recovery of cell wall (in chapter 6). All these observations suggested that the chromatin change was associated with cell wall.

The question remains why removal of the cell wall results in chromatin decondensation/reorganization. One possibility is that re-synthesis of the cell wall requires activation of a large number of genes simultaneously. Our analyses with our own ChIP-Seq results and published data sets (He et al., 2010) indicated that cell wall metabolic genes had substantially higher ratio of subjection to the regulation of histone modifications (Feng and Peng, unpublished results). In such a scenario, a signaling cascade which senses the cell wall state and activates the cell wall synthesis pathway genes must be involved. Another possibility is that cell wall and chromatin have some uncovered connections. For example, the cell wall synthesis and chromosome decondensation are concomitant processes during cytokinesis. The wall synthesis is initially guided by spindle fibers, which is directly linked with the chromosomes of daughter cells and then reorganized into phragmoplast. Thus, there is a possibility that cell wall synthesis and the chromatin state have an intrinsic connection, which remains to be explored.

Although our results suggest that removal of the cell wall treatment results in chromatin reorganization, directly or indirectly. Our results do not conflict with the concept that cell dedifferentiation causes chromatin decondensation/reorganization. Zhao *et al.* (Zhao et al., 2001) observed two rounds of chromatin decondensation in protoplast isolated from leaf tissues, we observed only one round of chromocenter decondensation in protoplasts isolated from dedifferentiated suspension cells, suggesting that dedifferentiation also contributed to chromatin change observed during leaf protoplast

culture. The previous proteome studies in our lab on cell dedifferentiation using cotyledons found that a substantial histone differential expression was associated with cell dedifferentiation (Chitteti et al., 2008). Since no cell wall removal was involved in that study, the histone differential expression observed should be due to cell dedifferentiation, supporting the concept that cell dedifferentiation is associated with chromatin structure change to some extent. It appears that both cell wall removal and cell dedifferentiation are cellular processes with a tight connection to chromatin state.

### **Histone modification and differential histone modification in rice**

Histone modifications have been documented in *Arabidopsis* using a mass spectrometry approach (Johnson et al., 2004; Bergmuller et al., 2007; Zhang et al., 2007). However, mass spectrometry analysis in rice and other plants is still poorly explored. We have identified acetylation, mono-, di-, and tri-methylation on residue H3K36; acetylation, mono- and di-methylation on residue H3K27; and acetylation on residues H3K14, H3K18, and H3K23 using mass spectrometry analyses of peptides derived from rice histone H3. Multisite modifications in individual tryptic peptides at lysine residues occurred due to closely spaced lysine residues. During MSMS analysis, all the acetylated peptides have the  $m/z$  126 marker ion. All the modifications described above and trimethylation on H3K9 and H3K27 have been verified using Western blots. Western blots further revealed that many of the tested histone modifications displayed differential expression between suspension cells, protoplast cells and protoplast-derived cells. In addition, the quantity change of H3K18 and H3K23 acetylation was estimated using an isotope labeling mediated method. Our results clearly demonstrated that in addition to

cellular histone and histone variant change, substantial histone modification change is associated with enzymatic removal of the cell wall in rice.

Histone H3 acetylation at K18 and K23 has been observed in *Arabidopsis*, *S. cerevisiae*, and *H. sapiens* (Zhang et al., 2007; Horwitz et al., 2008; Liu et al., 2009). However, the precise functions of these two modifications are still unclear and even controversial. Early studies in *S. cerevisiae* suggest a potential correlation between H3K18 acetylation and gene expression activity (Kurdistani et al., 2004), and recent investigations in *H. sapiens* revealed H3K18 hypoacetylation leads to gene inactivation (Horwitz et al., 2008). Qin and coauthors (Qin et al., 2009) found that simultaneous deletion of SEM1 and UBP6 in *S. cerevisiae* induced a dramatic silencing defect in the telomere. Interestingly, the silencing defect was accompanied by significantly reduced levels of acetylated H3K14 and H3K23 at the telomeres, suggesting that H3K14 and H3K23 acetylation associates either with telomere structure formation or telomere gene silencing. Offermann *et al.* (Offermann et al., 2008) treated *S. cerevisiae* with Trichostatin A, a histone deacetylase inhibitor. They found that acetylation remained limited to the gene regions even after Trichostatin A treatment for most of the lysine residues. However, H3K14 and H3K18 acetylation were induced in sub-telomeric regions which contain repetitive stretches of DNA, indicating H3K14 and H3K18 residues in the sub-telomeric region are the targets of histone acetyltransferases and suggesting a possible role of H3K14 and H3K18 acetylation in sub-telomeric structure. In this study, we found that removal of cell wall led to disassembly of chromocenters, a structure rich in heterochromatin. Our preliminary histone post-translational modification analyses suggested that acetylation of H3K18 and H3K23 is the most obvious modification change associated with cell wall removal. Further quantitative MS/MS studies using isotope

labeling approach confirmed that acetylation of H3K18 and H3K23 increased up to three folds. Meanwhile, our Western blot analysis revealed H3K18 and H3K23 acetylation level increased while the cell wall is removed and decreased while the cell wall is recovered, suggesting a possible role of H3K18 and H3K23 in regulation of chromatin structure. It is highly possible that, however, the chromatin structure change associated with hyperacetylation at H3K18 and H3K23 will lead to activation of some specific genes.

### **The differences between cell wall synthesis in protoplasts and cell wall synthesis during cytokinesis**

The phragmoplast plays an essential role in cell wall synthesis during late cytokinesis. It services as a scaffold for building the cell plate and formation of a new cell wall (Verma and Hong, 2001), which separates the two daughter cells. Only one phragmoplast and one new cell wall are produced for each dividing cell. When the cell wall was removed enzymatically, we found that cell wall re-synthesis started from multiple locations simultaneously (Figure 5.1). This observation raised many interesting questions, such as how the starting sites are determined, whether phragmoplast and cell plate like structures are involved in cell wall re-synthesis, and more importantly whether the same set of enzymes and apparatus are used in cell wall re-synthesis as during cytokinesis.

## CHAPTER VI

### PROTEOME DYNAMIC CHANGE DURING CELL WALL DEGRADATION AND REGENERATION IN *ORYZA SATIVA*

#### **Abstract**

Cell wall is a critical extracellular structure that provides protection and structural support in plant cells. Cell wall is built-de novo during cytokinesis and is guided by phragmaplast, which is derived from spindle fibers that are directly connected to the chromosomes of daughter cells. Plant cells can also re-synthesize the cell wall rapidly if the cell wall is removed. To study the biological function of cell wall and the regulation of cell wall synthesis, we examined cellular responses to enzymatic removal of cell wall in rice (*Oryza sativa*) suspension cells using comparative proteomics approach with label-free quantification methods, revealing 136 up-regulated proteins and 94 down-regulated proteins. Furthermore, these differentially expressed proteins were further studied in the context of an *Oryza Sativa* protein interaction network by Pathway Studio software. Our results show that several cellular processes have tight connections in response to cell wall removal and regrowth, like energy metabolism, cell growth and division, protein synthesis and transport, and chromosome reorganization.

#### **Introduction**

The composition of the cell wall was well studied for many species for years, until recently, only a few discoveries have revealed the mechanism of plant cells synthesize wall polysaccharides, assemble them into a rigid fibrous network and regulate

wall expansion during cell growth (Kaczkowski, 2003; Cosgrove, 2005; Zhong and Ye, 2007). However, special attention should be given to the fact that it is estimated that more than 1500 putative cell wall related genes are occurred in the genome of *Arabidopsis Thaliana* (Initiative, 2000; Somerville et al., 2004). The combinations of molecular biological, functional genomics and proteomics approaches, as well as public accessibility of the genomics resources will significantly facilitate the clarification of the complete picture. Despite cell wall proteins (CWPs) of plants and/or suspension cultured cells have been determined (Chivasa et al., 2002; Watson et al., 2004; Bayer et al., 2006), there is little about how cell wall affect cellular activities.

The mature plant cell can be separated from their original tissue by cell wall degrading enzymes, resulting in the formation of a large population of protoplast cells. During the cell wall removal process, the cells undergo remarkable changes in their pattern of gene expression and lose their differentiated state (Zhao et al., 2001). Following nutritive culturing, the protoplasts can re-enter the cell cycle, recover cell walls and eventually form new plantlets (Damm and Willmitzer, 1988). The plant protoplast is an attractive experimental system to investigate the biochemical and molecular basis for the cell wall removal induced cell dedifferentiation and cell wall regeneration. The protoplasts were characterized by the acquisition of totipotency, with a new balance between the euchromatin that is permitted to transcribe, and heterochromatin which is largely repressive.

Differential comparative proteomics is important for identifying molecular processes involved in different physiological conditions, is crucial to discover and unravel the molecular functions and biological roles of gene products. Recently, shotgun proteomics coupled with label-free quantification have demonstrated great potential in



comparative proteomic studies. The label-free methods are high throughput and completely eliminate the labor intensive 2-DE gel separation or sample labeling steps. In addition, they can separate a very wide range of proteins and overcome the protein solubility problem that is often encountered in the 2-DE gel method. The reported label-free methods include peptide counts, sequence coverage, peak area intensity measurements, spectral counts, and the sum of the TurboSEQUEST cross correlation coefficient ( $\Sigma$ XCorr) of peptides in a protein (Bantscheff et al., 2007; Wong et al., 2008). In this dissertation, we used both spectral counts and  $\Sigma$ XCorr methods to perform comparative analyses.

It is estimated that more than 1500 genes are involved in cell wall synthesis in *Arabidopsis Thaliana* (Initiative, 2000; Somerville et al., 2004). Due to the high complexity of cell wall synthesis, the cross-talk and integration of different pathways are still poorly understood. Using a systems biology approach may provide a more completed picture of cell wall synthesis and regulation. Kitano (Kitano, 2002) proposed systems biology, a systematic way to visualize multiple related biological processes in a network. This concept is generally applied to interpreting interactions of genes or gene products (Somerville et al., 2004). Molecular systems approach with the interaction networks not only can identify direct and indirect global responses of genes to the objective network, but also will allow us to identify key regulatory nodes in networks (He and Zhang, 2006). Recently, several computational tools and databases have been implemented that can be directly applied to existing public information and map genes/proteins into networks and pathways. Pathway Studio (Ariadne Inc., Rockville, MD, USA), one of the most widely used and commercially available software, enables researchers to navigate and analyze biological pathways, gene regulation networks and protein interaction maps (Nikitin et

al., 2003). Pathway Studio database, MedScan (Ariadne Inc., Rockville, MD, USA), works in many species (Buza and Burgess, 2008) because it is equipped with an automated text mining engine. Multiple aspects of protein function, including cellular location, protein-protein interactions, protein modifications, gene expression regulation, and regulation of various cellular processes are also included for many species (Novichkova et al., 2003). However, the database of plant genes and proteins is still limited. Nevertheless, this available systems biology tool makes it possible to examine cell wall re-synthesis in protoplasts using a molecular systems biology approach.

### **Materials and methods**

Rice suspension culture, protoplasts preparation and culture, Gene Ontology analysis were carried out as explained in materials and methods section of Chapter 5.

#### **Protein extraction**

Protoplasts were harvested at the start of culture (0hr) and after 2, 6, 12, 24, 48 and 72 hours of culture. Suspension cells treated with PTM medium for 9 hours were used as a control. The collected samples were ground in liquid nitrogen with pre-chilled mortar and pestle into fine powder. The procedures of phenol extraction method were described in chapter 3. Three biological replicas were extracted for each treatment including the control.

#### **Shotgun proteomic analysis**

Shotgun proteomics analyses were performed as we described before at the Life Sciences and Biotechnology Institute of Mississippi State University (McCarthy et al., 2005; Tan et al., 2007; van den Berg et al., 2007; Chitteti et al., 2008). Briefly, protein pellets were dissolved in 6 M urea and reduced with 10 mM DTT and alkylated with 50

mM iodoacetamide. The proteins were digested overnight at 37 °C with trypsin (Promega). Tryptic peptides were concentrated and desalted with a peptide macro trap (Michrom Bioresources, Auburn, CA).

The peptide mixtures were subjected to two-dimensional liquid chromatography (2-D LC) comprising a separation on a strong cation exchange column (SCX BioBasic 0.32 × 100 mm) followed by a reverse phase column (BioBasic C18, 0.18×100-mm Thermo Hypersil-Keystone, Bellefonte, PA) coupled directly in-line with electro spray ionization ion trap mass spectrometer (ProteomeX workstation, Thermo Finnigan). All the parameters and steps are described in chapter 3.

#### **Protein comparative quantification using the spectral count and $\Sigma$ XCorr methods**

The spectral count quantification method was used essentially as reported (Liu et al., 2004). The  $\Sigma$ XCorr quantification method was as reported by Nanduri and Bridges (Nanduri et al., 2005). Both spectral count and XCorr were generated by TurboSEQUENT (Bioworkers Brower 3.2, Thermo Electron Corp.), a commercial software widely used in mass spectral data analysis. The ProtQuant software (Bridges et al., 2007), a java-based tool for label-free quantification, which was used for spectral count and  $\Sigma$ XCorr quantification comparison was downloaded from AgBase (McCarthy et al., 2007) database tool box (<http://www.agbase.msstate.edu/>). The criteria and procedures of quantitative analysis of protein differential expression were as reported previously (Chitteti et al., 2008). One-way ANOVA analysis was conducted to determine the statistical significance of differential expression (*p*-value). The proteins with a *p*-value less than 0.05 in one or more time points were defined as differentially regulated.

Only the peptides having more than a 1.5 fold quantity change between control and treatments were considered for further analyses.

### **System modeling: *Oryza sativa* protein interaction network**

The differentially expressed proteins obtained from the label-free quantification were analyzed to visualize the cellular pathways and protein interactions applying Pathway Studio (Nikitin et al., 2003) (Ariadne Inc., Rockville, MD, USA) using ResNet-Plant database, which was updated on 07/08/2008. We first imported the plants 1.0 database from Ariadne Genomics downloads center into Pathway Studio 5.0. Then all the plants database updates available on the Ariadne Genomics downloads center were updated. The newly updated database contains functional protein information including GO (Gene Ontology) and various pathways and networks of protein-protein interactions (Buza and Burgess, 2008). The differentially regulated proteins were imported into Pathway Studio program. We built interaction networks with proteins, up and down respectively, including the upstream regulators and downstream targets. The blue marked ones were the identified proteins in our samples, while the yellow ones indicated the cellular processes involved. To further study the direct proteins and cellular processes related to responses to cell wall removal, we used “Find groups” tool in Pathway Studio to identify Gene Ontology (GO) groups that were significant ( $p\text{-value} \leq 0.05$ ). The statistical significance level of overlap between the protein list and the GO group was calculated by Pathway Studio using Fisher exact test (Nanduri et al., 2008). Pathway Studio provides tools to pick up the proteins with significance and cellular processes of interest, and establish the interaction network. Here, different colors were used to indicate

whether the protein was significantly increase (red), or decrease (grey) in rice protoplast culture in response to cell wall removal.

## **Results**

### **Protoplasts proteome revealed by shotgun proteomics**

As a first step in understanding the effects of cell wall removal, we analyzed the total protoplast proteome utilizing a shotgun proteomics approach. Protein mixtures were isolated via phenol extraction, digested by trypsin, and examined using LC/LC-MS/MS approach with a LCQ Deca XP ion trap mass spectrometer. A total of 345 proteins were identified based on the criterion of two or more peptide matches and another 453 proteins were identified based on a single peptide match. Due to the length limitation the protein information is not presented in this dissertation. Identified representative proteins related to cell wall metabolism included alpha-1,4-glucan-protein synthase, GDP-mannose 3,5-epimerase 1, GDP-mannose 3,5-epimerase 2, glycosyltransferase 6, glycosyl hydrolases family 16 protein, indole-3-acetate beta-glucosyltransferase, NAD dependent epimerase/dehydratase family protein and etc.

To verify our results, a reverse database of *Oryza sativa* was searched. The reversed database was generated by the functionality in Bioworks 3.2. Although we obtained 22 single peptide hits in the reverse database, no protein had two or more peptide matches. Therefore, proteins identified with two or more matched peptides were considered highly confident identifications. In the forward database, a total of 1,430 peptides were identified. Therefore, the peptide false discovery rate was estimated at 1.5% for the entire data set. The protein false discovery rate was estimated at 2.6%.

## **Differential expression of proteins in response to cell wall removal revealed by label-free quantification methods**

To gain insight into cellular responses to cell wall removal, we studied protein differential expression following the time course of cell wall removal and regeneration using spectral count and  $\Sigma$ XCorr methods (Liu et al., 2004; Old et al., 2005; He and Zhang, 2006; Schmidt et al., 2007). Proteins extracted from protoplasts cultured for 0, 2, 6, 12, 24, 48, and 72 h and extracted from suspension cells were subjected to shotgun proteomic analysis with three biological repeats for every treatment. The ProtQuant software, available at <http://www.agbase.msstate.edu/>, was used to carry out quantitative comparison among different treatments using the spectral count and  $\Sigma$ XCorr method as reported previously (Bridges et al., 2007; Chitteti et al., 2008). We converted the protein ID from TIGR (Rice Genome Annotation Project) to UniProt (Universal Protein Resource) entries, which contain high-quality annotation and are non-redundant and cross-referenced to many other databases. Proteins with 1.5 fold-changes ( $p\text{-value} \leq 0.05$ ) at one or more time points compared with suspension cells revealed by either one of the two label-free quantification methods were chosen for further analyses. As a result, a total of 230 proteins, of which 136 were up-regulated and 94 were down-regulated, were shown to be differentially expressed (Tables 6.2 and 6.3). To better understand the overall protein differential expression trend, the number of up and down expressed proteins, revealed by spectral count and  $\Sigma$ XCorr method at each time point are listed in Table 6.1. No substantial differences were observed between the two label-free quantification methods, suggesting that the quantification results were supported by both methods. There were 136 unique up-regulated proteins accumulatively and 94 unique down-regulated proteins accumulatively. In the protoplast culture process from 0 h to 72 h, the number of differentially up-regulated proteins at each time point decreased with

time. In contrast, the number of differentially down-regulated proteins increased with time, suggesting that the cell responded to the removal of the cell wall with a very rapid action with accumulation of new proteins to start to recover the cell wall.

The protein quantity change was calculated and the fold-change is presented in Tables 6.2 & 6.3. In addition, the value of control  $\Sigma$ XCorr versus treatment  $\Sigma$ XCorr and the total number of spectral counts of control versus treatment samples are also presented in parentheses for every protein at each time point. Many proteins were detected only in the treatment or the control. For these proteins, the fold-change is not presented but the  $\Sigma$ XCorr and the cumulative peptide count numbers are presented in parentheses.

Consistent with our observation with DAPI stain and MNase digestion (chapter 5), H1, H2A, three H2A variants, several H2B variants, H3.2, H4, and H3 as well as some histone modification enzymes were differentially expressed upon removal of cell wall and during the cell wall regeneration process. Together with our Western blots and isotope labeling assisted quantification of H3K18 and H3K23 acetylation described in chapter 5, our results indicated that there were dramatic changes in chromatin in response to the removal of the cell wall.

To display the relative protein expression level changes for each of the differentially expressed proteins, we generated a graph utilizing the sum of the XCorr value, produced by the ProtQuant program. Some examples are presented in Table 6.4. Glycosyltransferase family (GTs) proteins including glycosyltransferase 6 and indole-3-acetate beta-glucosyltransferase were up regulated while glycosyl hydrolase family (GHs) proteins such as glycosyl hydrolases family 16 protein and basic endochitinase 1 were down regulated. Several nuclear proteins, such as histone, histone deacetylase 2b, HMG-Y-related protein A and Putative Y1 protein, were significantly up regulated at an early

stage but declined over time. Meanwhile, some other histones, histone variants and nucleosome assembly protein SET showed down regulation at an early stage, followed by a slight recovery at a later stage, indicating that the chromatin structure was undergoing complicated dynamic changes.

We further classified these differentially expressed proteins following the Gene Ontology rules as described in the Materials and Methods section. The number of differentially expressed proteins within each gene ontology category is shown in Figure 6.1. As revealed by the pie chart, differentially expressed proteins were identified in most cellular components defined by GO (Figure 6.1A). The cellular components with the highest number of proteins displaying differential expression were intracellular, nucleus, ribosome, cytoplasm, and membrane. The sum of these cellular components occupied about 80% of the total differentially expressed proteins. To identify the ontology categories with the highest percentage of proteins undergoing differential expression in response to cell wall removal, the number of differentially expressed proteins was divided by the total number of proteins identified in the same category and is presented in Figure 6.2. Interestingly, ER, membrane, and cytoskeleton proteins were mainly subjected to up regulation (Figure 6.2A). Based on molecular function, DNA binding, protein binding, nucleotide and nucleic acid binding, structural molecule activities and translational factor activities had more proteins subjected to up regulation (Figure 6.2B). According to biological process, protein metabolic process, cellular component organization and biogenesis, and transport were mainly subjected to up regulation. In contrast, response to biotic stimulus, carbohydrate metabolic process, and electron transport were mainly subjected to down regulation (Figure 6.2C).



Table 6      Numbers of differentially expressed proteins revealed by label-free quantification

Regulation	Method	0hr	2hr	6hr	12hr	24hr	48hr	72hr	total
No of Up-Regulated	SC	66	72	63	46	33	38	36	136
	Xcorr	66	72	63	47	34	39	37	
No of Down-Regulated	SC	53	45	67	62	67	72	77	94
	Xcorr	54	46	66	62	67	71	78	

SC: The spectral count quantification method. Xcorr: The  $\Sigma$ XCorr quantification method.

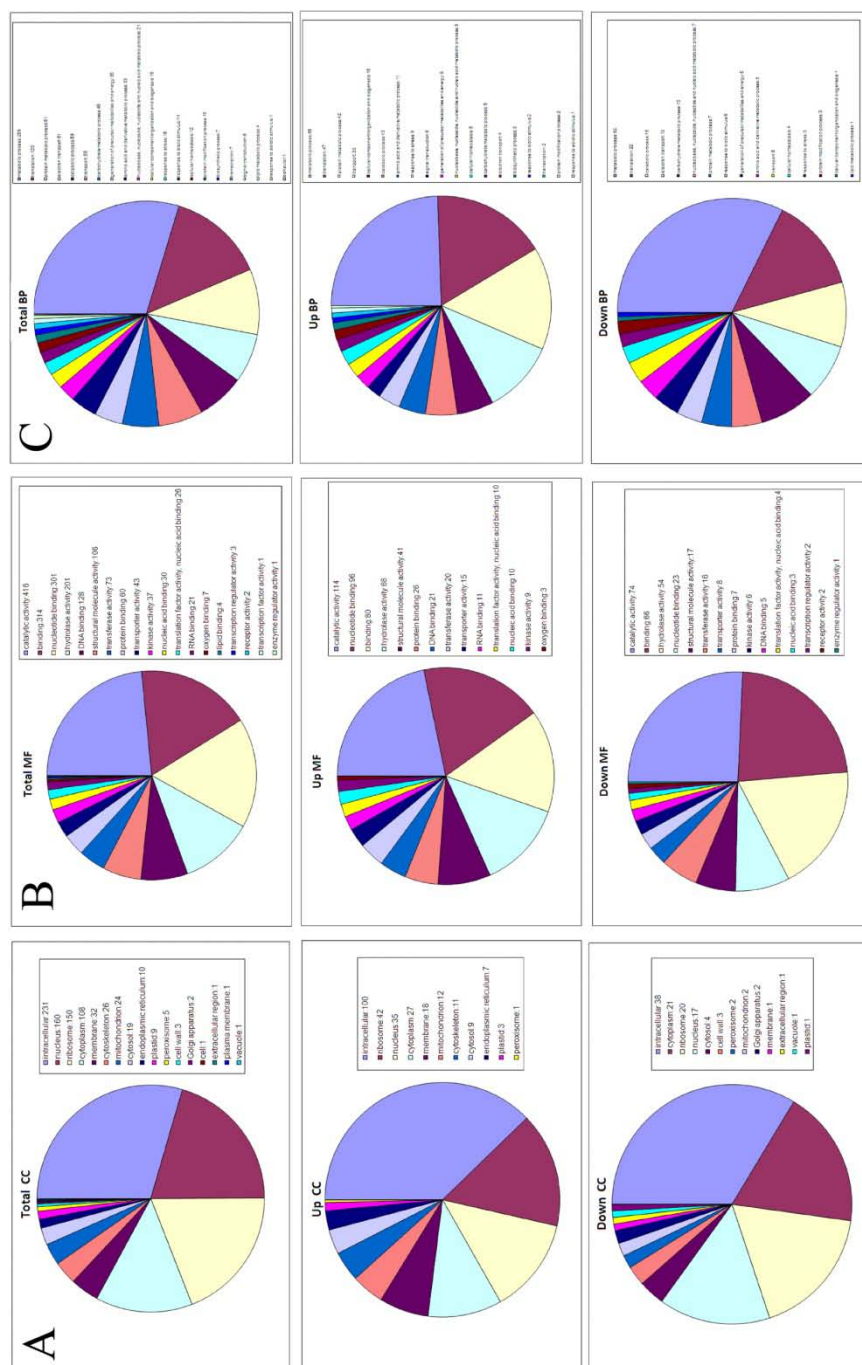


Figure 20 Protein distribution among different Gene Ontology categories

The total identified proteins, the up regulated proteins and the down regulated proteins are presented, respectively. The pie charts were generated using the program, GOSlimViewer, provided by AgBase, Mississippi State University. Gene distribution was grouped on the basis of Cellular Components (CC), Molecular Functions (MF), and Biological Processes (BP).

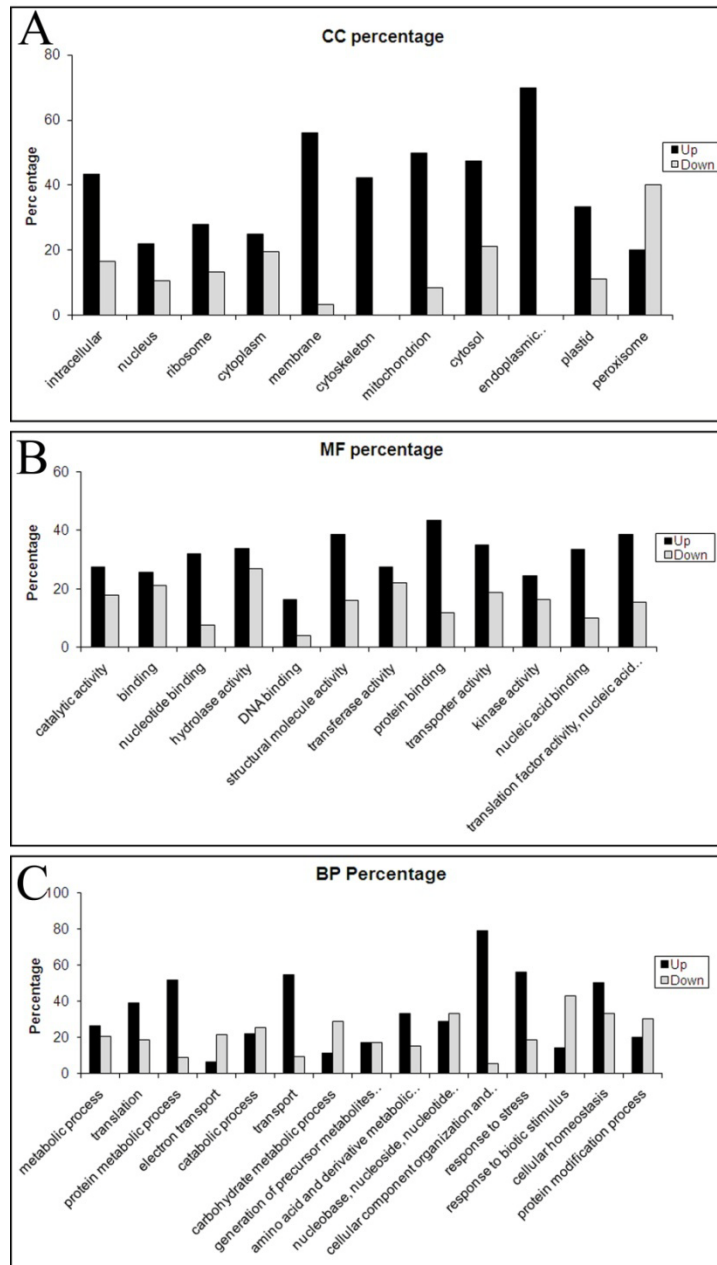


Figure 21 Percentage of differentially expressed proteins in gene ontology categories

The percentage ratios were calculated as follows: the number of differentially expressed proteins divided by total number of proteins identified in the same GO category. The GO annotations were retrieved from AgBase.

Table 7 Differentially Up regulated proteins identified using label-free quantification methods

ID	Description	Ratio of Expression (Control/0hr)		Ratio of Expression (Control/2hr)		Ratio of Expression (Control/6hr)		Ratio of Expression (Control/12hr)		Ratio of Expression (Control/24hr)		Ratio of Expression (Control/48hr)		Ratio of Expression (Control/72hr)	
		Xcorr	SC	Xcorr	SC	Xcorr	SC	Xcorr	SC	Xcorr	SC	Xcorr	SC	Xcorr	SC
A3C4S4	GDP-mannose 3,5-epimerase			-0.9(0.03)	~(0.4)										
1	GDP-mannose 3,5-epimerase														
Q2R1V8															
2	GDP-mannose 3,5-epimerase														
Q2R2W8	glycosyltransferase 6			~(0.7)52	~(0.3)					~(0.10.37)	~(0.4)	~(0.7)82	~(0.3)	~(0.7)47	~(0.3)
	indole-3-acetate beta-														
Q7XWK2	glucosyltransferase			3.5(4.01/14.19)	4.0(1.4)					3.3(2.4/20.01)	8.0(1.8)	5.1(2.4/12.33)	5.0(1.5)	4.1(2.4/9.83)	4.0(1.4)
Q00FD6	histone deacetylase 2b	5.1(2.4/12.19)	5.0(1.5)	2.8(2.1/6.78)	3.0(1.3)	5.7(2.4/13.61)	6.0(1.6)	12.8(2.4/30.61)	11.0(1/11)						
	core histone H2A/H2B/H3/H4														
Q7GBK0	family protein														
Q8H4Z0	histone H1							5.4(2.59/13.89)	5.0(1.5)						
Q7SL11	histone H2A					2.3(14.22/32.59)	2.0(5/10)	3.2(14.22/46.16)	2.8(5/14)	2.6(14.22/36.49)	2.2(5/11)			2.5(26.97/10.77)	2.0(6/3)
Q6L500	histone H2A							1.3(54.52/70.92)	1.7(12/20)						
Q8LLP5	histone H2A							4.7(2.33/10.93)	4.0(1.4)						
Q94E96	histone H2A							1.3(54.52/70.92)	1.7(12/20)						
Q2QST1	histone H2A.2							5.4(2.59/13.89)	5.0(1.5)						
Q6LGI2	Histone H2B.10							5.4(2.59/13.89)	5.0(1.5)						
Q943L2	Histone H2B.11							5.4(2.59/13.89)	5.0(1.5)						
Q6ZBP3	Histone H2B.2							5.4(2.59/13.89)	5.0(1.5)						
Q94JJ7	Histone H2B.3							5.4(2.59/13.89)	5.0(1.5)						
Q94JJ4	Histone H2B.4							5.4(2.59/13.89)	5.0(1.5)						
Q94JE1	Histone H2B.5							5.4(2.59/13.89)	5.0(1.5)						
Q6LGH4	Histone H2B.6							5.4(2.59/13.89)	5.0(1.5)						
Q6LGH8	Histone H2B.8							5.4(2.59/13.89)	5.0(1.5)						
Q6F362	Histone H2B.9							5.4(2.59/13.89)	5.0(1.5)						
Q2RAD9	Histone H3.2							5.4(2.59/13.89)	5.0(1.5)						
Q005M3	histone H4			1.7(15.05/25.19)	1.7(6/10)			5.2(4.6/23.71)	5.0(1.5)						
Q6JHY5	histone H4			1.7(15.05/25.19)	1.7(6/10)									5.2(15.04/2.89)	6.0(6/1)
Q7G823	Histone H4			2.4(4.84/11.79)	2.5(2/5)									5.2(15.04/2.89)	6.0(6/1)
	histone H4			3.6(2.85/10.22)	3.0(1/3)										
Q69MW7	HMG-Y-related protein A	3.8(2.85/10.9)	4.0(1/4)	3.4(8.95/30.52)	2.7(3/8)	4.1(8.95/36.79)	3.3(3/10)								
Q43594	tubulin beta-1 chain	5.3(6.95/47.54)	4.3(3/13)	4(0.6)8	~(0.3)										

Table 7 (continued)

[illegible]

Table 7 (continued)

[illegible]

Table 7 (continued)

[illegible]

Table 7 (continued)

[illegible]



Table 7 (continued)

Q6L590	luminal-binding protein 4 precursor	3.0(18.6/155.99)	3.3(4/13)	~(0/8.16)	~(0/3)	3.9(18.6/173.5)	4.5(4/18)	2.0(18.6/137.59)	2.3(4/9)			~(0/10.34)	~(0/3)
Q7XDC8	malate dehydrogenase		3.3(4/13)	3.2(18.6/159.36)	3.3(4/13)								
Q6ZG90	mitochondrial ATP synthase precursor			~(0/10.88)	~(0/4)	~(0/24.74)	~(0/9)	~(0/8.37)	~(0/3)	~(0/10.02)	~(0/3)		
Q10A19	mitochondrial prohibitin			5.3(3.23/17.25)	5.0(1/5)								
Q7EYR6	complex protein 2			2.7(8.13/21.91)	3.0(2/6)	2.9(8.13/23.91)	3.0(2/6)	~(0/15.14)	~(0/3)	~(0/13.23)	~(0/3)		
Q7XIR7	complex protein 2												
Q7XIR7	myosin-like protein												
Q6AV34	N-acetyl-gamma-glutamyl-phosphate reductase											~(0/13.15)	~(0/3)
Q6AV34	NADH ubiquinone												
Q8GS72	oxidoreductase B14 subunit	5.0(2.85/14.14)	4.0(1/4)										
O04985	non-symbiotic hemoglobin 2			4.1(3.98/16.31)	4.0(1/4)								
O04986	non-symbiotic hemoglobin 2												
Q0JB08	Os04g056390.0 protein					~(0/6.96)	~(0/3)					~(0/8.47)	~(0/3)
P55142	OsGrx_C2.2 - glutaredoxin subgroup I			~(0/16.53)	~(0/5)								
Q53LQ0	OsPDIL1.1 - Oryza saliva protein disulfide isomerase	1.6(69.71/112.32)	1.7(18/30)	2.1(69.71/147.08)	2.2(18/40)	2.2(69.71/150.96)	2.8(18/41)	1.6(69.71/110.94)	1.7(18/30)				
Q75M08	OsPDIL2.1 - Oryza saliva protein disulfide isomerase	~(0/15.35)	~(0/4)	~(0/11.65)	~(0/4)	~(0/20.44)	~(0/6)					~(0/13.71)	~(0/4)
Q942L2	OsPDIL2.2 - Oryza saliva protein disulfide isomerase			~(0/12.51)	~(0/4)								
Q67UF5	OsPDIL2.3 - Oryza saliva protein disulfide isomerase			~(0/17.71)	~(0/6)	~(0/24.37)	~(0/7)	~(0/26.30)	~(0/7)	~(0/22.30)	~(0/5)	~(0/15.73)	~(0/4)
Q6L515	outer mitochondrial membrane protein porin	1.7(37.89/63.69)	1.9(7/13)	2.1(37.89/79.15)	2.4(7/17)	2.0(37.89/76.12)	2.4(7/17)						
Q84P97	outer mitochondrial membrane protein porin									~(0/8.8)	~(0/3)		
Q7F4F8	outer plastidial membrane protein porin	~(0/15.48)	~(0/5)	~(0/11.55)	~(0/3)								
Q0DR93	PD-like protein	5.8(3.85/22.35)	6.0(1/6)	5.9(3.85/22.76)	6.0(1/6)								
Q0DAB3	peptidase C22G7.01c											3.0(6.81/17.67)	4.0(1/4)
Q10L32	peptide methionine sulfoxide reductase matB							~(0/14.96)	~(0/3)				~(0/3)
Q8FR35	peroxiredoxin-5	3.9(4.96/19.37)	4.0(1/4)	2.0(25.79/51.35)	2.1(8/17)	3.8(4.96/18.89)	5.0(1/5)	1.7(25.78/44.29)	1.9(8/15)	1.9(25.78/48.81)	1.9(8/15)	2.2(25.78/55.62)	2.4(8/19)
Q6H6C7	phosphoglycerate kinase												

Table 7 (continued)

[illegible]

Table 7 (continued)

[illegible]

Table 7 (continued)

	succinate dehydrogenase	1.6(49.177.24)	1.5(12/18)	1.8(49.190.73)	1.8(12/21)	2.2(49.1/105.62)	2.2(12/26)					
Q6ZDY8	flavo-protein subunit											
Q00TX5	superoxide dismutase 2					~(0/12.52)	~(0/3)					
Q43008	superoxide dismutase			~(0/7.99)	~(0/3)	~(0/23.94)	~(0/9)	~(0/10.82)	~(0/4)			
P48404	triosephosphate isomerase	2.9(38.44/112.79)	2.5(11/28)	3.0(38.44/116.11)	2.5(11/28)	3.0(38.44/116.9)	2.9(11/32)	3.0(38.44/114.33)	2.6(11/29)	3.3(38.44/124.98)	2.5(38.44/95.23)	2.6(38.44/99.33)
Q0J8H7	tyrosyl-tRNA synthetase	~(0/10.13)	~(0/3)									2.5(11/27)
Q84PX0	tyrosyl-tRNA synthetase	~(0/10.13)	~(0/3)									
Q5ZAV7	USP family protein							2.5(8.73/21.44)	3.5(2/7)		3.0(8.73/25.87)	4.0(2/8)
Q7XXS5	USP family protein							5.3(4.89/26)	5.0(1/5)	6.0(1/6)	5.2(4.89/25.34)	5.0(1/5)
Q0ISB6	xaa-Pro aminopeptidase 1	3.4(4.89/46.19)	9.0(1/9)	3.8(4.89/33.04)	6.0(1/6)	4.4(4.89/21.49)	4.0(1/4)			5.4(4.89/31.09)	3.0(6.81/17.67)	5.0(4.89/24.64)

a) ID: UniProt Accession number

b) Protein Identification: Protein name

c) Hours: The hours indicate the time of protoplasts cultured

d) Control: Suspension cell treated with PM medium and without multienzymes digestion was used as control.

e)  $\Sigma X_{corr}$ : Result based on  $\Sigma X_{corr}$  quantification method

f) SC: Results based on Spectral Count quantification method

g) Ratio of Expression: Protein quantity in protoplasts cultured sample / control sample

h) ( ): Parentheses includes the quantification results in control sample / protoplasts cultured sample

i) "~": represents the ratio is not available because the protein is not detected either in the control or in the treatment

j) Note: the blank cells indicate either the data didn't meet the P-value significance ( $P \leq 0.05$ ) criteria or less than three peptides were identified in protoplast sample

Table 8 Differentially Down regulated proteins identified using label-free quantification methods

ID	Descriptio n	Ratio of Expression (Control/0hr)		Ratio of Expression (Control/2hr)		Ratio of Expression (Control/6hr)		Ratio of Expression (Control/12hr)		Ratio of Expression (Control/24hr)		Ratio of Expression (Control/48hr)		Ratio of Expression (Control/72hr)	
		Xcorr	SC	Xcorr	SC	Xcorr	SC	Xcorr	SC	Xcorr	SC	Xcorr	SC	Xcorr	SC
	alpha-1,4-														
Q6Z4G3	glucan-protein synthase 1	~(10.38/0)	~(3/0)	~(10.38/0)	~(3/0)	~(10.38/0)	~(3/0)	~(10.38/0)	~(3/0)	~(10.38/0)	~(3/0)	~(10.38/0)	~(3/0)	~(10.38/0)	~(3/0)
	alpha-1,4-														
Q8H8T0	glucan-protein synthase	3.3(35.62/4.27)	7.0(7/1)	3.3(35.62/4.27)	7.0(7/1)	3.3(35.62/4.27)	7.0(7/1)	3.3(35.62/4.27)	7.0(7/1)	3.3(35.62/4.27)	7.0(7/1)	3.3(35.62/4.27)	7.0(7/1)	3.3(35.62/4.27)	7.0(7/1)
	basic														
P246Z6	endochitinase 1 precursor	~(24.76/0)	~(4/0)	~(24.76/0)	~(4/0)	~(24.76/0)	~(4/0)	~(24.76/0)	~(4/0)	~(24.76/0)	~(4/0)	~(24.76/0)	~(4/0)	~(24.76/0)	~(4/0)
	basic														
Q0D9D2	endochitinase 1 precursor	~(13.23/0)	~(3/0)	~(13.23/0)	~(3/0)	~(13.23/0)	~(3/0)	~(13.23/0)	~(3/0)	~(13.23/0)	~(3/0)	~(13.23/0)	~(3/0)	~(13.23/0)	~(3/0)
	glycine-rich														
	RNA-binding														
Q10FE7	protein 2														
	glycine-rich														
	RNA-binding														
Q2QLR2	protein 2														
Q851P9	histone H1	~(6.90/0)	~(3/0)	~(6.90/0)	~(3/0)	~(6.90/0)	~(3/0)	~(6.90/0)	~(3/0)	~(6.90/0)	~(3/0)	~(6.90/0)	~(3/0)	~(6.90/0)	~(3/0)
	histone H2A														
Q8S857	variant 2														
Q6ZL42	histone H2A.3	1.9(20.20/10.49)	2.3(7/3)	1.9(20.20/10.49)	2.3(7/3)	1.9(20.20/10.49)	2.3(7/3)	1.9(20.20/10.49)	2.3(7/3)	1.9(20.20/10.49)	2.3(7/3)	1.9(20.20/10.49)	2.3(7/3)	1.9(20.20/10.49)	2.3(7/3)
Q84NJ4	histone H2A.4	7.0(23.73/3.39)	7.0(7/1)	7.0(23.73/3.39)	7.0(7/1)	7.0(23.73/3.39)	7.0(7/1)	7.0(23.73/3.39)	7.0(7/1)	7.0(23.73/3.39)	7.0(7/1)	7.0(23.73/3.39)	7.0(7/1)	7.0(23.73/3.39)	7.0(7/1)
Q0JCT1	histone H3	2.3(48.57/17.66)	2.7(8/3)	2.3(48.57/17.66)	2.7(8/3)	2.3(48.57/17.66)	2.7(8/3)	2.3(48.57/17.66)	2.7(8/3)	2.3(48.57/17.66)	2.7(8/3)	2.3(48.57/17.66)	2.7(8/3)	2.3(48.57/17.66)	2.7(8/3)
	nuclear transport														
Q9XJ54	factor 2	4.3(21.28/4.96)	4.0(4/1)	4.3(21.28/4.96)	4.0(4/1)	4.3(21.28/4.96)	4.0(4/1)	4.3(21.28/4.96)	4.0(4/1)	4.3(21.28/4.96)	4.0(4/1)	4.3(21.28/4.96)	4.0(4/1)	4.3(21.28/4.96)	4.0(4/1)
	nuclear WD														
Q69IK7	protein														
Q69JW2	protein SET														

Table 8 (continued)

[illegible]

Table 8 (continued)

4,5-DOPA											
QQUH82	dioxygenase extracellular	~(3.0)	~(3.0)	~(3.0)	~(3.0)	~(3.0)	~(3.0)	~(3.0)	~(3.0)	~(3.0)	~(3.0)
QQLP8	40S ribosomal protein S12	2.3(29.5/12.56)	2.3(7/3)					5.2(29.5/4.79)	7.0(7/1)		
Q8H2J8	40S ribosomal protein S12	2.3(29.5/12.56)	2.3(7/3)					5.2(29.5/4.79)	7.0(7/1)		
Q0DBK8	40S ribosomal protein S24					5.8(13.2/4.27)	4.0(4/1)	~(13.2/4.0)	~(4.0)	~(13.2/4.0)	~(4.0)
Q8H541	40S ribosomal protein S24					5.8(13.2/4.27)	4.0(4/1)	~(13.2/4.0)	~(4.0)	~(13.2/4.0)	~(4.0)
4	hydroxyphenylpyruvate dioxygenase										
Q0E3L4	60S ribosomal protein L13-2	~(3.0)	~(3.0)	~(16.1/1.0)	~(3.0)	~(16.1/1.0)	~(3.0)	~(16.1/1.0)	~(3.0)	~(16.1/1.0)	~(3.0)
Q7XJB4	60S ribosomal protein L13a	3.3(3/1.0)	3.3(3/1.0)	2.5(8.19/20.64)	2.7(3/8)			~(8.19/0)	~(3.0)	~(8.19/0)	~(3.0)
Q75J18	60S ribosomal protein L13a-2										9.37(25.77/2.75)
Q0DNG2	60S ribosomal protein L13a-2										9.37(25.77/2.75)
Q7F225	60S ribosomal protein L2										10.2(28.06/2.75)
Q2QNF3	60S ribosomal protein L21										1.8(63.66/29.05)
Q0DVA0	60S ribosomal protein L21	1.5(11/16)	1.5(11/16)	3.7(32.23/8.60)	3.7(11/3)			~(32.23/0)	2.8(11/4)	~(32.23/0)	~(11/0)
Q8AV87	60S ribosomal protein L21	1.5(11/16)	1.5(11/16)	3.7(32.23/5.60)	3.7(11/7)			~(32.23/0)	2.8(11/4)	~(32.23/0)	~(11/0)
Q8F331	60S ribosomal protein L30	3.0(6/2)	3.0(6/2)			5.8(27.44/3.95)	6.0(6/1)	5.8(27.44/4.70)	6.0(6/1)		
Q0J0Q8	60S ribosomal protein L32					5.6(27.68/4.91)	5.0(5/1)	5.8(27.68/4.07)	5.0(5/1)	~(27.68/0)	~(5/0)
Q0J0R0	60S ribosomal protein L32					5.6(27.68/4.91)	5.0(5/1)	5.8(27.68/4.07)	5.0(5/1)	~(27.68/0)	~(5/0)
Q84QW3	60S ribosomal protein L32					5.6(27.68/4.91)	5.0(5/1)	5.8(27.68/4.07)	5.0(5/1)	~(27.68/0)	~(5/0)

Table 8 (continued)

60S ribosomal protein L9	~(3/0)	~(10.56/0)	~(3/0)	~(10.56/0)	~(3/0)	~(10.56/0)	~(3/0)
60S ribosomal protein L9	~(3/0)	~(10.56/0)	~(3/0)	~(10.56/0)	~(3/0)	~(10.56/0)	~(3/0)
adenylate kinase A	~(3/0)	~(10.50/0)			~(3/0)	~(10.50/0)	
adenyllyl cyclase-associated protein 1, putative, expressed			~(19.05/0)	~(4/0)	~(19.05/0)	~(4/0)	~(4/0)
alcohol dehydrogenase 1 aldehyde			~(31.61/0)	~(8/0)	5.6(31.61/5.63)	4.0(8/2)	~(8/0)
Q9FRX7 dehydrogenase	4.0(8/2)	4.1(31.61/7.69)	2.7(8/3)				
Q7X7N2 arginase		~(26.00/0)	~(5/0)				
aspartyl aminopeptidase		~(25.25/0)	~(5/0)	~(5/0)	~(25.25/0)	~(5/0)	~(5/0)
Q20VB4 basic 7S globulin precursor	~(7/0)	~(19.15/0)	~(7/0)	~(7/0)	~(19.15/0)	~(7/0)	~(7/0)
Q6AUL0 bifunctional 3-phosphoadenosine 5-phosphosulfate synthetase bifunctional polymyxin resistance araA protein			~(11.61/0)	~(3/0)	~(11.61/0)	~(3/0)	~(3/0)
Q859Z2 protein carrier/ steroid binding protein	4.0(4/1)	3.8(12.87/3.41)	~(12.87/0)	~(4/0)	~(12.87/0)	~(4/0)	~(4/0)
Q7XXQ9 catalase isozyme A					~(9.77/0)	~(3/0)	~(3/0)
Q0E4K1 catalase isozyme B						~(8.18/0)	~(3/0)
Q0D9C4						2.5(77.96/31.72)	2.1(19/9)



Table 8 (continued)

[illegible]

Table 8 (continued)

[illegible]

Table 8 (continued)

Q0D8X9	expressed protein	~(15.180)	~(50)	5.3(15.18/2.42)	5.0(5/1)	5.1(15.18/2.49)	5.0(5/1)	~(15.180)	~(50)	~(15.180)	~(50)	~(15.180)	~(50)
Q0DLI3	expressed protein	~(19.510)	~(40)	~(19.510)	~(40)	~(19.510)	~(40)	~(19.510)	~(40)	~(19.510)	~(40)	~(19.510)	~(40)
Q2QSF1	expressed protein	~(19.510)	~(40)	~(19.510)	~(40)	~(19.510)	~(40)	~(19.510)	~(40)	~(19.510)	~(40)	~(19.510)	~(40)
Q6AUK5	expressed protein	~(9.630)	~(30)	3.6(16.44/4.52)	4.0(4/1)	3.9(16.44/4.22)	4.0(4/1)	4.3(16.44/3.85)	4.0(4/1)	~(16.440)	~(40)	~(16.440)	~(40)
Q7XIE9	expressed protein	~(12.300)	~(40)	~(12.300)	~(40)	~(12.300)	~(40)	~(12.300)	~(40)	~(9.630)	~(30)	~(9.630)	~(30)
Q8LHX6	expressed protein	~(12.300)	~(40)	~(12.300)	~(40)	~(12.300)	~(40)	~(12.300)	~(40)	~(12.300)	~(40)	~(12.300)	~(40)
Q8LHY4	expressed protein	~(8.520)	~(30)	~(8.520)	~(30)	~(8.520)	~(30)	~(8.520)	~(30)	~(8.520)	~(30)	~(8.520)	~(30)
Q40684	expressed protein	~(10.110)	~(30)	~(10.110)	~(30)	~(10.110)	~(30)	~(10.110)	~(30)	~(10.110)	~(30)	~(10.110)	~(30)
Q40684	ferrodoxin-6	~(10.110)	~(30)	~(10.110)	~(30)	~(10.110)	~(30)	~(10.110)	~(30)	~(10.110)	~(30)	~(10.110)	~(30)
Q40684	fructose-bisphosphate	~(10.110)	~(30)	~(10.110)	~(30)	~(10.110)	~(30)	~(10.110)	~(30)	~(10.110)	~(30)	~(10.110)	~(30)
Q94JU0	aldolase	~(10.110)	~(30)	~(10.110)	~(30)	~(10.110)	~(30)	~(10.110)	~(30)	~(10.110)	~(30)	~(10.110)	~(30)
Q7XU92	geranylgeranyl pyrophosphate synthetase 1	~(30.840)	~(50)	~(30.840)	~(50)	~(30.840)	~(50)	~(30.840)	~(50)	~(30.840)	~(50)	~(30.840)	~(50)
Q6YTH5	gibberellin receptor	~(15.340)	~(30)	~(15.340)	~(30)	~(15.340)	~(30)	~(15.340)	~(30)	~(15.340)	~(30)	~(15.340)	~(30)
Q6H3Y7	GID1L2 glutamate dehydrogenase A	~(22.820)	~(50)	~(22.820)	~(50)	~(22.820)	~(50)	~(22.820)	~(50)	~(22.820)	~(50)	~(22.820)	~(50)
P14656	glutamine synthetase	~(22.820)	~(50)	~(22.820)	~(50)	~(22.820)	~(50)	~(22.820)	~(50)	~(22.820)	~(50)	~(22.820)	~(50)
Q0J9K2	root isozyme 3	~(10.4(59.08/5.71)	8.5(17/2)	~(10.4(59.08/5.71)	8.5(17/2)	~(10.4(59.08/5.71)	8.5(17/2)	~(10.4(59.08/5.71)	8.5(17/2)	~(10.4(59.08/5.71)	8.5(17/2)	~(10.4(59.08/5.71)	8.5(17/2)
Q0J9K2	glycerol kinase	~(19.650)	~(60)	~(19.650)	~(60)	~(19.650)	~(60)	~(19.650)	~(60)	~(19.650)	~(60)	~(19.650)	~(60)
Q0J9U2	glycosyl hydrolases	~(44.210)	~(90)	~(44.210)	~(90)	~(44.210)	~(90)	~(44.210)	~(90)	~(44.210)	~(90)	~(44.210)	~(90)
Q7XXR4	family 16 protein	~(44.210)	~(90)	~(44.210)	~(90)	~(44.210)	~(90)	~(44.210)	~(90)	~(44.210)	~(90)	~(44.210)	~(90)
Q6H660	haloacid dehalogenase-like hydrolase domain- containing protei heat shock protein	~(2.5(43.10/7.46)	1.8(9/5)	~(2.5(43.10/7.46)	1.8(9/5)	~(2.5(43.10/7.46)	1.8(9/5)	~(2.5(43.10/7.46)	1.8(9/5)	~(2.5(43.10/7.46)	1.8(9/5)	~(2.5(43.10/7.46)	1.8(9/5)
Q10LW8	hydroxyacylglutathion e hydrolase	~(20.610)	~(50)	~(20.610)	~(50)	~(20.610)	~(50)	~(20.610)	~(50)	~(20.610)	~(50)	~(20.610)	~(50)
Q6Y2Z7	hypothetical protein	~(19.090)	~(80)	~(19.090)	~(80)	~(19.090)	~(80)	~(19.090)	~(80)	~(19.090)	~(80)	~(19.090)	~(80)
Q0D7G4	IAA-amino acid hydrolase ILR1-like 6 precursor	~(19.880)	~(30)	~(19.880)	~(30)	~(19.880)	~(30)	~(19.880)	~(30)	~(19.880)	~(30)	~(19.880)	~(30)

Table 8 (continued)

Q6ETD9	IQ calmodulin-binding motif family protein isoflavone reductase	~(17/0)	~(69.63/0)	~(17/0)	~(69.63/0)	~(3/0)	~(13.60/0)	~(3/0)	~(13.60/0)	~(3/0)	~(13.60/0)	~(3/0)
Q8FTN5	homolog IRL isoflavone reductase	~(17/0)	~(69.63/0)	~(17/0)	~(69.63/0)	~(17/0)	~(69.63/0)	~(17/0)	31.3(69.63/2.23)	17(17/1)	~(69.63/0)	~(17/0)
Q8FTN6	homolog IRL reductase	~(3/0)	~(13.06/0)	~(3/0)	~(13.06/0)	~(3/0)	~(13.06/0)	~(3/0)	~(13.06/0)	~(3/0)	~(13.06/0)	~(3/0)
Q0DJE6	lactoyglutathione lyase late embryogenesis abundant protein L-lactate dehydrogenase A major pollen allergen Bet v 1-1	~(4/0)	~(15.72/0)	~(4/0)	~(15.72/0)	~(4/0)	~(15.72/0)	~(4/0)	~(15.72/0)	~(4/0)	~(15.72/0)	~(4/0)
Q94JF2	embryogenesis abundant protein L-lactate dehydrogenase A major pollen allergen Bet v 1-1	~(7/0)	~(24.96/0)	~(7/0)	~(24.96/0)	~(7/0)	~(24.96/0)	~(7/0)	~(24.96/0)	~(7/0)	~(24.96/0)	~(7/0)
Q0E4Q5	dehydrogenase A major pollen allergen Bet v 1-1	~(3/0)	~(13.35/0)	~(3/0)	~(13.35/0)	~(3/0)	~(13.35/0)	~(3/0)	~(13.35/0)	~(3/0)	~(13.35/0)	~(3/0)
Q2QNS7	D/H major pollen allergen Bet v 1-1	~(103/0)	~(349.70/0)	22.1(349.70/15.85)	20.6(103/5)	51.5(103/2)	42.6(349.70/8.20)	34.3(103/3)	43.3(349.70/8.09)	34.3(103/3)	~(349.70/0)	~(103/0)
Q2QNT0	D/H malate dehydrogenase mitochondrial-processing peptidase beta subunit NADH-ubiquinone oxidoreductase 13 kDa-B subunit	2.6(18/7)	~(49.74/5.45)	9.0(18/2)	~(49.74/0)	~(18/0)	~(49.74/0)	~(18/0)	~(49.74/0)	~(18/0)	~(49.74/0)	~(18/0)
Q6F361	malate dehydrogenase mitochondrial-processing peptidase beta subunit NADH-ubiquinone oxidoreductase 13 kDa-B subunit	~(4/0)	~(12.94/0)	~(4/0)	~(12.94/0)	~(4/0)	~(12.94/0)	~(4/0)	~(12.94/0)	~(4/0)	~(12.94/0)	~(4/0)
Q10Q21	peptidase beta subunit NADH-ubiquinone oxidoreductase 13 kDa-B subunit	~(4/0)	~(11.13/0)	~(4/0)	~(11.13/0)	~(4/0)	~(11.13/0)	~(4/0)	~(11.13/0)	~(4/0)	~(11.13/0)	~(4/0)
Q10MD5	subunit NADH-ubiquinone oxidoreductase 13 kDa-B subunit	~(4/0)	~(11.13/0)	~(4/0)	~(11.13/0)	~(4/0)	~(11.13/0)	~(4/0)	~(11.13/0)	~(4/0)	~(11.13/0)	~(4/0)

Table 8 (continued)

nascent polypeptide-associated complex alpha subunit-like protein	~(11.48/0)	~(3/0)	~(11.48/0)	~(3/0)	~(11.48/0)	~(3/0)	~(11.48/0)	~(3/0)
Q0DIP0								
nascent polypeptide-associated complex alpha subunit-like protein	~(13.91/0)	~(4/0)	~(13.91/0)	~(4/0)	~(13.91/0)	~(4/0)	~(13.91/0)	~(4/0)
Q8RUI4								
nucleoside diphosphate kinase 1	1.8(91.13/50.72)	2.0(18/9)	1.3(91.13/67.86)	1.5(18/12)	1.5(91.13/61.79)	1.6(18/11)	1.6(91.13/56.62)	1.8(18/10)
Q07661								
ornithine aminotransferase	2.6(31.35/12.08)	3.0(9/3)	2.6(31.35/12.21)	3.0(9/3)	2.8(31.35/11.21)	3.0(9/3)	2.2(31.35/14.19)	2.3(9/4)
Q10G56								
protein	2.5(31.35/12.63)	3.0(9/3)	5.6(27.68/4.91)	5.0(5/1)	5.5(27.68/4.28)	5.0(5/1)	5.8(27.68/4.07)	5.0(5/1)
Q0J052								
OsAPX1 - Cytosolic Ascorbate Peroxidase encoding gene	2.2(44.69/20.56)	1.8(11/6)	4.0(44.69/11.07)	3.7(11/3)	3.9(44.69/11.33)	2.8(11/4)	2.3(44.69/19.28)	1.8(11/6)
Q10N21								
OsAPX7 - Stromal Ascorbate Peroxidase encoding gene	~(12.88/0)	~(4/0)	5.5(12.88/2.32)	4.0(4/1)			5.3(12.88/2.04)	4.0(4/1)
Q7XJ02								
osmolin-like protein precursor	10.9(61.74/5.65)	14.0(14/1)	4.8(61.74/12.86)	4.7(14/3)	~(61.74/0)	~(14/0)	~(61.74/0)	~(14/0)
Q2CND8								
osmolin-like protein precursor	~(32.92/0)	~(8/0)	~(32.92/0)	~(8/0)	~(32.92/0)	~(8/0)	~(32.92/0)	~(8/0)
Q943L0								
pathogenesis-related protein 10	~(741.06/0)	~(197/0)	55.8(741.06/13.29)	49.3(197/4)	264.0(741.06/2.81)	197.0(197/1)	116.8(741.06/6.34)	98.5(197/2)
Q75T45								

Table 8 (continued)

[illegible]

Q6Z1G7	pyruvate dehydrogenase E1 component subunit beta	2.4(32.36/13.50)	2.3(7/3)	3.5(32.36/9.33)	3.5(7/2)	~(32.36/0)	~(7/0)
Q0D3V9	sex determination protein	~(16.42/0)	~(5/0)	~(16.42/0)	~(5/0)	~(16.42/0)	~(5/0)
Q7FAE2	sex determination protein	~(16.42/0)	~(3/0)	~(8.58/0)	~(3/0)	~(8.58/0)	~(3/0)
Q9TTY4	stress responsive protein	~(8.58/0)	~(3/0)	~(18.54/0)	~(4/0)	~(18.54/0)	~(4/0)
Q6ZL94	succinyl-CoA ligase alpha-chain 2	1.8(52.07/28.92)	1.8(14/8)	1.7(52.07/30.08)	1.6(14/9)	1.5(52.07/34.12)	1.5(14/9)
Q5JK78	ubiquitin carboxyl-terminal hydrolase 6	2.7(34.93/13.16)	2.7(6/3)	3.3(34.93/3.77)	8.0(8/1)	4.2(34.93/8.32)	4.0(6/2)
Q8H6G9	ubiquitin-conjugating enzyme E2 I1	~(15.2/0)	~(3/0)	~(15.2/0)	~(3/0)	~(15.2/0)	~(3/0)
Q7XHV1	USP family protein	3.7(32.97/8.83)	3.0(6/2)	8.7(32.97/3.77)	6.0(6/1)	~(32.97/0)	~(6/0)
Q10RH2	XS domain containing protein	~(13.28/0)	~(3/0)	~(13.28/0)	~(3/0)	~(13.28/0)	~(3/0)

b) Protein Identification: Protein name

c) Hours: The hours indicate the time of protoplasts cultured

- d) Control: Suspension cell treated with PM medium and without multienzymes digestion was used as control.
- e)  $\Sigma X_{corr}$ : Result based on  $\Sigma X_{corr}$  quantification method
- f) SC: Results based on Spectral Count quantification method
- g) Ratio of Expression: Protein quantity in control sample / protoplasts cultured sample
- h) ( ): Parentheses includes the quantification results in control sample / protoplasts cultured sample
- i) "~": represents the ratio is not available because the protein is not detected either in the control or in the treatment
- j) Note: the blank cells indicate the data didn't meet the P-value significance ( $P \leq 0.05$ ) criteria or less than three peptides were identified in protoplast sample



Table 9 Identified differentially regulated proteins by label-free methods

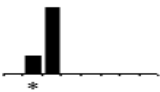
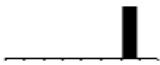


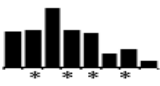

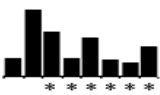
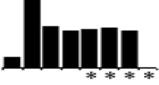
Uniprot ID	Description	Relative Expression	P (pro)	Coverage	MW (KDa)
A3C4S4	GDP-mannose 3,5-epimerase 1		1.31E-003	7.41	42.8
P27933	alpha-amylase isozyme 3D precursor		1.22E-008	5.04	52.4
P48494	triosephosphate isomerase		1.73E-010	48.62	27.1
Q01859	ATP synthase beta chain		5.15E-010	41.67	58.9
Q0D5M3	histone H4		1.06E-004	33.98	11.4
Q0DFD6	histone deacetylase 2b		1.02E-005	4.04	32.5
Q0DIY4	ran-binding protein 1 homolog c		6.89E-008	15.45	24.1
Q0DWC1	fibrillarin-2		1.69E-010	6.61	23.7

Table 9 (continued)



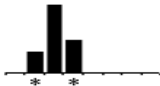

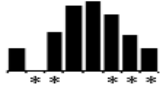
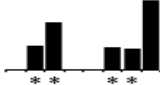
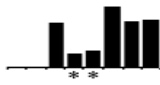


Q0J8A4	glyceraldehyde-3-phosphate dehydrogenase		1.94E-007	31.16	36.4
Q0JKB4	ATP synthase beta chain		5.34E-011	36.09	59.5
Q10D68	serine hydroxymethyltransferase		4.76E-008	9.55	56.4
Q10RW9	chaperonin CPN60-1		2.38E-009	23.74	61
Q2QVC1	argininosuccinate synthase		3.31E-005	7.48	52.2
Q2R1V8	GDP-mannose 3,5-epimerase 2		1.89E-003	7.01	42.1
Q2R2W8	glycosyltransferase 6		2.52E-005	1.86	52.9
Q42971	enolase		6.53E-010	34.17	51.6
Q53NM9	heat shock cognate 70 kDa protein 2		4.15E-008	14.78	57.1

Table 9 (continued)



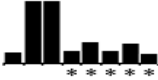







Q5N726	fructose-bisphosphate aldolase cytoplasmic isozyme		2.62E-009	38.83	38.8
Q5NAI9	Putative Y1 protein		2.82E-007	14.29	51.1
Q65XK0	ketol-acid reductoisomerase		2.66E-011	10.73	62.4
Q67UF5	OsPDIL2-3 - Oryza sativa protein disulfide isomerase		3.35E-007	7.26	47.3
Q69MW7	HMG-Y-related protein A		6.83E-005	19.72	21.8
Q6AT27	fibrillarin-2		1.24E-005	13.92	32.6
Q7XDC8	malate dehydrogenase		5.43E-013	28.31	35.6
Q7XKB5	pyruvate kinase		1.07E-006	7.24	55.2
Q7XWK2	indole-3-acetate beta- glucosyltransferase		4.88E-007	3.27	53.4
Q7XXS5	USP family protein		1.93E-009	10.05	20.9

Table 9 (continued)




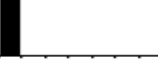

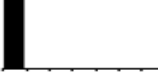

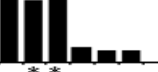

Q84UR8	glycine-rich protein 2b		4.37E-005	31.98	18.7
Q8H3I7	chaperonin		7.69E-006	40.82	10.6
Q8H903	chaperonin CPN60-1		3.92E-006	17.82	67.3
P24626	basic endochitinase 1 precursor		3.86E-007	18.44	33.7
Q06967	14-3-3-like protein S94		1.21E-007	28.85	29.2
Q0D9D2	basic endochitinase 1 precursor		4.12E-008	16.91	36.7
Q0IXF3	cell division control protein 48 homolog E		1.01E-005	8.66	89.8
Q0J0S2	Os09g0501100 protein		1.69E-004	23.31	15.7
Q0J9U2	glycosyl hydrolases family 16 protein		6.19E-006	18.18	34.7

Table 9 (continued)










Q0JC04	eukaryotic initiation factor iso-4F subunit p82-34		9.50E-008	2.14	87
Q0JCT1	Histone H3.3		1.77E-006	13.79	30.7
Q10DR8	adenylyl cyclase-associated protein 1		3.00E-006	7.72	52.1
Q10G56	ornithine aminotransferase		4.63E-010	9.73	51.4
Q10RP0	cell division control protein 48 homolog E		1.01E-005	8.65	89.7
Q2QND8	osmotin-like protein precursor		7.30E-007	26.61	23.4
Q2QNS7	major pollen allergen Bet v 1-D/H		2.77E-010	76.58	16.7
Q2QNT0	major pollen allergen Bet v 1-D/H		1.00E-006	39.24	17
Q69JW2	protein SET		1.27E-009	7.14	28.6

Table 9 (continued)


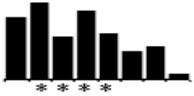

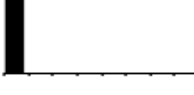

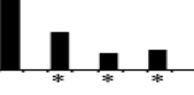
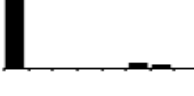


Q6AUK5	expressed protein		5.64E-006	7.86	24.7
Q6EUP4	14-3-3-like protein B		2.36E-006	17.56	29.7
Q6H660	heat shock protein STI		3.06E-009	11.76	64.9
Q6I5R5	Putative cold regulated protein		2.04E-006	27.54	18.2
Q6YYB3	protein kinase		1.18E-008	9.29	30.4
Q6Z4G3	alpha-1,4-glucan-protein synthase 1		8.82E-006	11.48	41.3
Q6ZJX3	protein kinase		1.09E-006	30.43	30
Q6ZJX4	protein kinase		1.09E-006	38.32	29.8
Q6ZJX8	protein kinase		2.47E-006	20.73	29.7

Table 9 (continued)

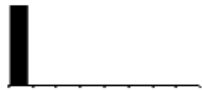

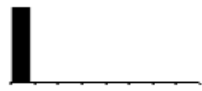
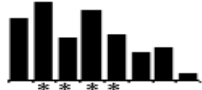
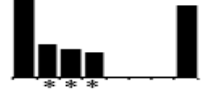
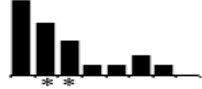
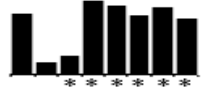
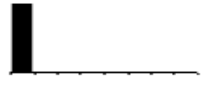


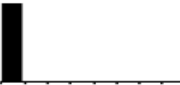



Q75GR1	protein P21		3.83E-006	27.16	24.1
Q75T45	pathogenesis-related protein 10		1.23E-007	68.13	16.9
Q7XI92	geranylgeranyl pyrophosphate synthetase 1		8.95E-009	12.84	38.7
Q7XTE8	14-3-3-like protein GF14-6		2.36E-006	17.56	29.9
Q8H3Q7	xylose isomerase		1.34E-013	4.8	53.5
Q8H8T0	alpha-1,4-glucan-protein synthase		2.78E-014	9.02	30.2
Q8LMR0	phosphoserine aminotransferase		1.22E-008	9.62	44.9
Q8S1V0	xylanase inhibitor		8.78E-009	26.73	43.8
Q8S857	histone H2A variant 2		6.16E-005	21.58	14.6

Table 9 (continued)

Q8SA35	vacuolar ATP synthase subunit E		7.93E-006	23.04	26.6
Q943L0	osmotin-like protein precursor		7.13E-005	36.84	26.2
Q9ASP4	dihydrolipoyl dehydrogenase		7.55E-014	16.5	52.6
Q9XJ54	nuclear transport factor 2		4.88E-009	58.2	13.4
P24626	basic endochitinase 1 precursor		3.86E-007	18.44	33.7

a) ID: UniProt Accession number; b) Description: protein name; c) P (pro): the probability of random match for this protein in *Oryza Sativa* database; d) Coverage: the percent of the AA sequence that have been verified with MS/MS; e) MW: predicted molecular mass; f) Peptide: number of peptides that matched with the identified protein in mass analyses; g) Relative Expression: graphical representation of relative expression is given for control and each protoplasts cultured time point (from left to right: control, 0, 2, 6, 12, 24, 48, and 72hr) based on  $\Sigma$ Xcorr generated by ProtQuant; h) the blank of the columns indicate no peptide was identified; the star indicates at this time point the difference does not meet the P-value significance ( $P \leq 0.05$ ) or less than three peptides were identified. However, in order to make the bar pictures better looking these ones were included.

### Systems approach revealed key regulatory nodes of the protein interaction network

Systems biology approach is a powerful tool to investigate the global trends and the interaction networks of genes, proteins, and pathways. To analyze the differentially expressed proteins using a systems biology approach, we uploaded these proteins into the



Pathway Studio (Nikitin et al., 2003) program, 136 up-regulated and 94 down-regulated proteins. Since all these imported proteins were covered by the plant database of Pathway Studio, the database was adequate for further analysis. We utilized the “build pathway” functionality to make the interaction maps and to identify cellular processes involved. The known upstream regulators and downstream targets were also included to facilitate construction of protein-protein interaction network.

The constructed networks (Figure 6.3) contained hundreds of proteins in which the nodes (proteins) indicated functional annotation, the edges (links) between proteins described binding, regulation, expression, or chemical reactions, and the rectangle (cellular processes) displayed the biological event involved. In summary, 887 proteins were included in the up regulation network (Figure 6.3A) and 581 proteins were included in the down regulation map (Figure 6.3B). The blue marked ones were the differentially expressed proteins identified in our proteome studies. As revealed by the map, a large number of proteins had no interactive proteins identified thus far. We found that more up regulated proteins had no edges compared with the down regulated proteins, suggesting that the up-regulated proteins and their related pathways were still poorly studied. In the established interaction map, many of the down-regulated proteins acted as key nodes, which were central elements connected with several other proteins. Further examination of these down regulated proteins in key node positions found that most of them were cytoplasmic and mitochondrial proteins.

Using the “find group” tool to identify potential GO groups ( $p\text{-value} \leq 0.05$  as a criterion), we identified and selected some cellular processes and the corresponding proteins for further examination, including energy metabolism, cytoskeleton related, protein synthesis and transportation, DNA proliferation, and chromatin assembly. For the

selected groups, the objective processes and proteins were pulled together, regardless of regulation, and a network was established (Figure 6.4). The red ovals represented the proteins that were up regulated while the grey ovals indicated those that were down regulated. The yellow rectangles described the biological processes involved and the edges between each other were relations, such as binding, regulation, expression, *etc.* Overall fifty-five differentially expressed proteins were displayed on the interaction network, 34 up and 21 down regulated respectively, which encompassed 35 various cellular processes (Figure 6.4). Many cell growth and development related processes were identified including cell expansion and regulation of cell size and cell shape, which were directly connected with cell wall synthesis.

Profilin (F6F22.20) was generally believed to bind with actin (T21B14.7), a major component of cytoskeleton, and affects the skeleton structure. With a high concentration of profilin, the polymerization of actin is prevented. On the other hand, profilin also catalyzes the exchange of actin-bound ADP to ATP thereby converting poorly polymerizing ADP-actin monomers into readily polymerizing ATP-actin monomers. Thus in a mixture of actin, profilin and nucleotides (ADP and ATP), actin will polymerize to a certain extent (Korenbaum et al., 1998). Cyclase-associated protein (CAP) (ATCAP1) has recently been shown to be an important regulator of actin dynamics. When functional CAP1 is present, actin was subjected to depolymerization and ADP-actin monomers are funneled to profilin and subsequently to growing actin filament ends (Baum et al., 2000; Benlali et al., 2000). The plant CAP, profilin, and other actin binding factors act cooperatively to regulate actin assembly and depolymerization dynamically, in which CAP1 is a key element in regulating the pool of unpolymerized actin (Chaudhry et al., 2007). Other cytoskeleton proteins such as tubulins (TUB1) and prohibitin

(AT1G03860) were up regulated (Figure 4). Ran and RanBP were previously reported to be involved in nuclear transport during interphase (Dasso and Pu, 1998) and mitotic spindle assembly (Gruss and Vernos, 2004). We also found Ran proteins and Ran binding protein (RanBP), both were differentially up-regulated and linked with cell expansion and cell cycle related processes. Catalase (AT4G35090), a common enzyme involved in oxidative stress, was dramatically increased under the stimulus. This protein relates to several plant defense and stress response processes in the network.

To our surprise, GTs and GHs were not apparent in gene ontology groups shown in Figure 7. Given that Pathway Studio drew biological information from published articles to establish the interaction map, the absence of these proteins suggested that the cell wall metabolic pathways still remain poorly explored. Thus, Pathway Studio can only provide limited information in the current version. More research is required to achieve comprehensive understanding of cellular processes in response to the removal of cell wall.

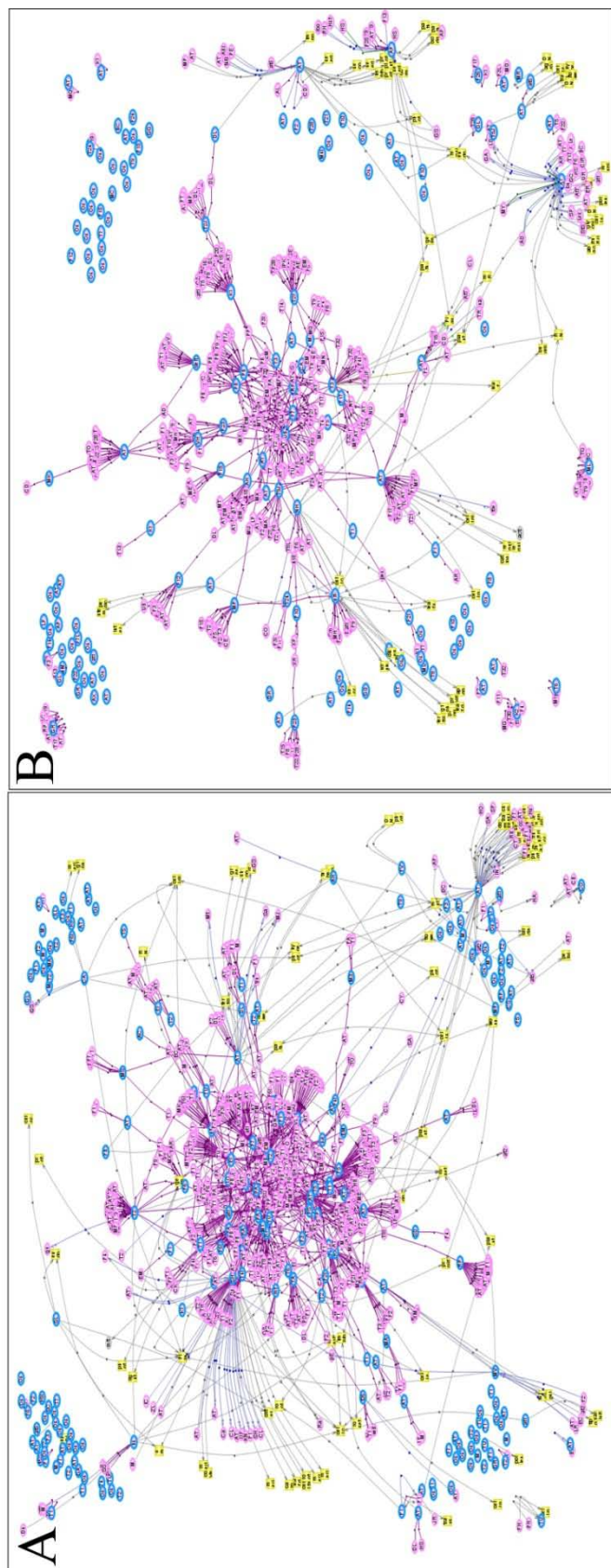


Figure 22 Interaction network of *Oryza Sativa* protoplast proteins with significantly altered expression response to absence of cell wall

All 223 up regulation proteins (A) and 150 down regulation proteins (B) from rice protoplasts with significant changes in expression level were constructed into networks in Pathway Studio. Both upstream regulators and downstream targets proteins were included. The identified differential proteins were marked with blue cycles.



involvement in cell wall re-synthesis. Meanwhile, we also noted that several cell wall and cell wall synthesis related genes are down regulated. Further examination revealed that these down regulated proteins are mainly extracellular proteins. Removal of the cell wall may have removed these proteins, thus reducing the contents of these proteins in the protoplast samples. These proteins may also be critical for cell wall re-synthesis. It is also worthy to note that suspension cells carry out rapid cell division and cell wall synthesis. Therefore, the expression level of cell wall synthesis proteins should be high in suspension cells already. The cell wall synthesis pathway proteins can be actively involved in cell wall re-synthesis without being further induced by the removal of the cell wall. Digital transcription profiling and microarray analysis can be a good method to help estimate the genes involved in cell wall re-synthesis due to the high sensitivity of the methods.

We further examined these proteins using gene ontology tools. We found that ER and membrane proteins were mainly subjected to up regulation, which was consistent with active cell wall synthesis activity upon removal of the cell wall. We also observed that the cytoskeleton proteins were only subjected to up-regulation, suggesting that the skeleton system was induced in response to the removal of the cell wall probably due to the need to support the protoplasts and facilitate cell wall synthesis. Another possible scenario is that devoid of cell wall induced chromatin reorganization, which requires a large number of cytoskeleton proteins to assistant the movements. In addition, we found that cellular component organization and biogenesis was the biological process with the highest number of proteins displaying differential up-regulation. A large number of up-regulated proteins are involved in protein metabolic process and translation, suggesting active protein synthesis upon removal of the cell wall. While the number of proteins

response to biotic stress decreased, the total number of proteins response to all stresses increased, indicating that removal of the cell wall causes severe stress to the cell. We used the protoplast isolation buffer, which contained 0.4 M mannitol, treated suspension cells as control. Therefore, the effect of osmotic stress should not contribute to the difference between suspension cells and protoplasts, suggesting that removal of the cell wall may result in other stresses.

Strikingly, various histones and histone variants, such as H1, H2A, three H2A variants, several H2B variants, H3.2, H4, and H3, as well as some histone modification enzymes were differentially expressed upon removal of cell wall and during the cell wall regeneration process, which were consistent with our observation with DAPI stain and MNase digestion (chapter 5). Together with our Western blots and isotope labeling assisted quantification of H3K18 and H3K23 acetylation described in chapter 5, all these abover discoveries indicated that there were dramatic changes associated with chromatin in response to the removal of the cell wall. Our results suggested that there might be intensive cross-talks between chromatin and cell wall, directly or indirectly. Our results also indicated that cell wall re-synthesis in protoplasts may also involve novel mechanisms when compared with cell wall synthesis during cytokinesis because the cell wall re-synthesis is initiated at multiple sites. Our results suggest that the cell wall removal and regeneration is an excellent system to study cell wall function and answer many fundamental questions related to cell wall synthesis and regulation of biomass characteristics.

## CHAPTER VII

### SUMMARY

The plant cell wall is a dynamic structure that plays a critical role not only in determining cell shape and formation of the plant body, but also in interactions with environmental factors including those required for nutrition, response to abiotic stress and biological attack by other organisms. The regeneration of a cell wall or wall-like structure around isolated higher plant protoplasts has been demonstrated for decades and in many species. Recent sequencing technology has annotated thousands of cell-wall related genes which are estimated to occur in the plant genome database. These genes are considered to encode proteins in categories with differing biological functions. Despite these genomic approaches, little is known about the mechanism of concerted action of cell wall re-synthesis by these proteins which are involved in various cellular processes. It is believed that the remarkable regeneration capacity of plant cells is based on their capability to dedifferentiate. It has long been assumed that the plant cells underwent cell dedifferentiation during protoplast preparation. Several studies reported that chromatin conformational changes have been detected during plant protoplast cell preparation and it is thought that epigenetic mechanisms that took place on chromatin were somehow linked to dedifferentiation and differentiation. However, some studies as well as this dissertation observed dramatic conformation changes in chromatin when cell walls were enzymatically removed using *Oryza sativa* suspension cell, which had dedifferentiated already. Therefore, the chromatin reorganization may be not due to dedifferentiation and



it is likely that a novel mechanism exists to regulate cell wall degradation and regeneration. We propose that epigenetic regulation probably involves cell wall degradation and regeneration and histone post-translational modifications may play a critical role. Hence, the main goal of this research is to improve the understanding of the molecular mechanisms of epigenetic regulation and cell wall degradation and regeneration in plants. Using *Oryza sativa* as the model species, research was carried out to achieve the following objectives:

1. To study the global proteome of rice chromatin using Two-Dimensional Gel Electrophoresis (2-DE) and Multidimensional Protein Identification technique (MudPIT) approaches.
2. To investigate the phosphoproteome map of rice chromatin using specific fluorescent staining and 2-DE coupled MALDI-TOF mass analysis.
3. To examine and validate the chromatin dynamic reorganization and study histone modifications and their quantitative changes during cell wall degradation and regeneration.
4. To investigate the total proteome dynamic changes involved in cell wall degradation and regeneration in rice and corresponding genes will be characterized according to Gene Ontology and pathway studies.

Rice has a small and completely sequenced genome. This small genome size and short life cycle make rice an excellent choice for studying epigenetic regulation of cell wall. For example, the rice genome is less than 1/6 of the maize genome; the small genome size substantially reduces the cost of establishing the database of histone modification binding sites within the cell wall genes. Meanwhile, the available mutants and the efficient transformation system have set a foundation for isolating knockdown

and knockout mutants. More importantly, results obtained in rice can be applied to other biomass crops in the grass family due to the conservation of epigenetic regulatory mechanisms.

Chromatin and its components are the key targets in this study since epigenetic mechanisms are intimately linked to chromatin and are defined as heritable changes in gene expression that do not include the DNA sequence itself. Chromatin basic structure and its major components have been well documented. However, a comprehensive chromatin proteome map has yet to be studied. The identification of chromatin associated proteins should enhance our understanding of gene regulation, chromatin higher level structure, and conformational dynamics. One of the techniques is using modern tools such as proteomics to characterize the proteins and histone modifications that associated with chromatin. But before that, the first step is to establish a highly efficient method to prepare chromatin. In this study, large scale and high efficient nuclei and chromatin purification methods have been developed (Figure 3.1). Various microscopic examinations and Western blot assay were conducted to characterize chromatin structure and purification quality. As shown in Figure 3.1C and Figure 3.1D, large clusters of thick chromatin fiber were observed. It was evident that chromatin structure above the 30 nm fiber is preserved, at least partially, in our rice chromatin preparation. Western immunoblots (Figure 3.2) have shown that the purified chromatin is highly enriched and free of visible organelle contamination.

In this dissertation, the traditional Two-Dimensional Gel Electrophoresis coupled MALDI-TOF and shotgun proteomics approaches have been used to establish the chromatin proteome map of *Oryza sativa*. Among the 972 excised protein spots, 607 were annotated after MS/MS mass analyses. Of these annotated spots, 509 proteins were

identified with confidence intervals (C.I. %) over 95% corresponding to 269 unique proteins. When using shotgun proteomics approach, 507 proteins corresponding to 292 unique proteins were identified. Functional categorization of proteins was carried out after getting the corresponding protein accession number by Agbase and agriGO base on Gene Ontology rules. The distribution was grouped on the basis of cellular localization, biological process and molecular functions.

Chromatin phosphorproteome was examined using Pro Q Diamond phosphoprotein gel stain followed by mass spectrometry analysis. Among the 390 putative phosphoprotein spots, 154 annotations which correspond to 101 unique proteins were identified by mass analysis with a C.I. % over 95. Further Gene Ontology analysis demonstrated that the phosphoproteins were highly involved in diversified cellular processes and cellular components, especially the nucleosome and chromatin assembly and conformational organization (Figure 4.3 and Figure 4.4).

In plant protoplast culture, it has been well documented that plant cells re-synthesize the cell wall rapidly after removal of the cell wall, indicating that the cell wall synthesis pathways are highly activated in protoplasts. Since all living protoplasts synthesize the cell wall simultaneously, protoplast culture is an excellent system to study cell wall regeneration and regulation. Cell wall regeneration was evaluated by microscopic monitoring. We found that the cell wall regeneration progressed rapidly and cell wall synthesis at multiple sites could integrate to form the wall cage appropriately (Figure 5.1). In addition to rapid cell wall regeneration, we discovered that removal of cell wall resulted in chromatin decondensation (Figure 5.2B). Since we used dedifferentiated cells for protoplast isolation, the chromatin should not be due to cell dedifferentiation caused by the removal of cell wall but the cell wall itself. In fresh

protoplasts (without cell wall) the chromocenters largely disappear. When cell wall regenerated, the chromocenters reformed (Figure 5.2C). Given that primary cell synthesis during cytokinesis is also concomitant with chromosome decondensation and guided by phragmoplast derived spindle fiber connected with chromosomes, the cell wall synthesis and chromatin state connect with each other. To further examine the chromatin change in protoplasts, we performed chromatin decondensation assay. As shown in Figure 5.2E, MNase digestion of suspension cell chromatin was clearly slower than the digestion of protoplast chromatin, suggesting a decondensation of chromatin in response to the enzymatic removal of cell wall.

To further investigate the chromatin changes associated with chromatin decondensation, we conducted histone post-translational modification characterization using mass spectrometry. Mass spectrometry alone possesses the sensitivity and speed necessary to provide accurate identification of modified species and simultaneous pinpointing of the modification site. While performing LC/LC-MS/MS analysis in protein identification, we observed many H3 post-translational modifications. Totally, 11 modified histone H3 and H3.2 peptides, including 5 acetylation sites and 3 methylation sites were mapped (Table 5.1). Monomethylation was identified on residue H3K36. Acetylation, mono-, and dimethylation were identified on residue H3K27. Acetylation was identified on residues H3K14, H3K18, and H3K23. In addition, histone H3.2 peptides were modified at Lys 36 (acetylation, di-, and trimethylation). Despite these findings, it is clear that the identification and characterization of epigenetic markers remains a challenging area of our research.

The substantial chromatin reorganization in response to cell wall removal and regeneration led us to the hypothesis that the histone modification state may also change

with cell wall. The traditional Western blot and immunochemistry was conducted to detect the relatively quantitative changes of modifications. Immunodetection of histone modifications using specific antibodies is a highly reliable method, particularly when a Western blot with antibodies for the histone protein (regardless histone modifications) is used as a control for quantitative comparison (Figure 5.4). Among these histone modifications we found that H3K18 and H3 K23 acetylation was the most obvious change. Highly specific antibodies for a particular modification are required, which are unfortunately still not available for many plant specific modifications. Another limitation is that only one modification can be detected each time usually. To validate our findings, LC-MS/MS with isotop labeling was performed to compare H3K18 and H3K23 acetylation levels between protoplasts and suspension cells. The acetylation level at H3K18 and H3K23 were 27.5% and 10.3% in suspension cells, respectively, and 52.2% and 27.4% in protoplasts, respectively (Figure 5.5). Combining those results from Western blots and LC-MS/MS, it is obvious that substantial histone modification changes are associated with enzymatic removal of the cell wall in rice.

Protein differential expression profile at 0, 2, 6, 12, 24, 48, and 72 hr time points after induction of protoplast cells from *Oryza sativa* suspension cells has been explored. A 1.5 fold difference in expression ( $p\text{-value} \leq 0.05$ ) compared to control in all three biological replicas was used as cut off for the identification of differentially regulated proteins by label-free quantitation method, namely ProtQuant, which can quantify the MudPIT data based on the sum of Cross Correlation values or Spectral Count method. Totally, 230 proteins, of which 136 were upregulated and 94 were down-regulated, were shown to be differentially expressed. Several cell wall synthesis related proteins were indentified and have been differentially regulated such as Glycosyltransferase family

(GTs) proteins including glycosyltransferase 6 and indole-3-acetate beta-glucosyltransferase, glycosyl hydrolase family (GHs) proteins including glycosyl hydrolases family 16 protein and basic endochitinase 1. Many of these have a complicated dynamic expression profile. In addition, various nuclear proteins, such as histone, histone variants, nucleosome assembly protein, histone deacetylase 2b, HMG-Y-related protein A and Putative Y1 protein, also appear to have differential expression, indicating that the chromatin structure was undergoing complicated dynamic changes. This provides additional evidence to show the connections between epigenetic regulations and cell wall degradation and regeneration. Moreover, the differentially expressed proteins were further studied in the context of an *Oryza Sativa* protein interaction network by PathwayStudio software. Our results show that several cellular processes have tight connections in response to cell wall removal and regrowth including energy metabolism, cell growth and division, protein synthesis and transport, and chromosome reorganization.

In conclusion, the work presented in this dissertation demonstrates the power of proteomic methods in the identification and characterization of chromatin associated proteins and histone post-translational modifications, and it further provides significant new insight into epigenetic regulations during cell wall degradation and regeneration in plants.

## REFERENCES

- Aebersold, R., and Mann, M. (2003). Mass spectrometry-based proteomics. *Nature* 422, 198-207.
- Agalioti, T., Lomvardas, S., Parekh, B., Yie, J., Maniatis, T., and Thanos, D. (2000). Ordered recruitment of chromatin modifying and general transcription factors to the IFN-beta promoter. *Cell* 103, 667-678.
- Agrawal, G.K., and Thelen, J.J. (2005). Development of a simplified, economical polyacrylamide gel staining protocol for phosphoproteins. *Proteomics* 5, 4684-4688.
- Agrawal, G.K., and Thelen, J.J. (2006). Large scale identification and quantitative profiling of phosphoproteins expressed during seed filling in oilseed rape. *Molecular & Cellular Proteomics* 5, 2044-2059.
- Anderson, L., and Seilhamer, J. (1997). A comparison of selected mRNA and protein abundances in human liver. *Electrophoresis* 18, 533-537.
- Arney, K.L., and Fisher, A.G. (2004). Epigenetic aspects of differentiation. *Journal of Cell Science* 117, 4355-4363.
- Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., Harris, M.A., Hill, D.P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J.C., Richardson, J.E., Ringwald, M., Rubin, G.M., and Sherlock, G. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nature Genetics* 25, 25-29.
- Bannister, A.J., Zegerman, P., Partridge, J.F., Miska, E.A., Thomas, J.O., Allshire, R.C., and Kouzarides, T. (2001). Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* 410, 120-124.
- Bantscheff, M., Schirle, M., Sweetman, G., Rick, J., and Kuster, B. (2007). Quantitative mass spectrometry in proteomics: a critical review. *Analytical and Bioanalytical Chemistry* 389, 1017-1031.
- Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. *Cell* 129, 823-837.

- Baum, B., Li, W., and Perrimon, N. (2000). A cyclase-associated protein regulates actin and cell polarity during *Drosophila* oogenesis and in yeast. *Current Biology* 10, 964-973.
- Bayer, E.M., Bottrill, A.R., Walshaw, J., Vigouroux, M., Naldrett, M.J., Thomas, C.L., and Maule, A.J. (2006). Arabidopsis cell wall proteome defined using multidimensional protein identification technology. *Proteomics* 6, 301-311.
- Benjamini, Y. Yekutieli D: The control of the false discovery rate in multiple testing under dependency. *Ann Stat* 2001, pp. 29--1165.
- Benlali, A., Draskovic, I., Hazelett, D.J., and Treisman, J.E. (2000). Act up controls actin polymerization to alter cell shape and restrict hedgehog signaling in the *Drosophila* eye disc. *Cell* 101, 271-281.
- Berggren, K., Steinberg, T.H., Lauber, W.M., Carroll, J.A., Lopez, M.F., Chernokalskaya, E., Zieske, L., Diwu, Z., Haugland, R.P., and Patton, W.F. (1999). A luminescent ruthenium complex for ultrasensitive detection of proteins immobilized on membrane supports. *Anal Biochem* 276, 129-143.
- Bergmuller, E., Gehrig, P.M., and Gruissem, W. (2007). Characterization of post-translational modifications of histone H2B-variants isolated from *Arabidopsis thaliana*. *J Proteome Res* 6, 3655-3668.
- Bernstein, E., and Hake, S.B. (2006). The nucleosome: a little variation goes a long way. *Biochem Cell Biol* 84, 505-517.
- Biel, M., Wascholowski, V., and Giannis, A. (2005). Epigenetics--an epicenter of gene regulation: histones and histone-modifying enzymes. *Angew Chem Int Ed Engl* 44, 3186-3216.
- Birnbaum, K., Shasha, D.E., Wang, J.Y., Jung, J.W., Lambert, G.M., Galbraith, D.W., and Benfey, P.N. (2003). A gene expression map of the Arabidopsis root. *Science (New York, N Y)* 302, 1956-1960.
- Birnbaum, K., Jung, J.W., Wang, J.Y., Lambert, G.M., Hirst, J.A., Galbraith, D.W., and Benfey, P.N. (2005). Cell type-specific expression profiling in plants via cell sorting of protoplasts from fluorescent reporter lines. *Nature Methods* 2, 615-619.
- Blaschek W, H.D., Koeler H, Franz G. (1981). Cell wall regeneration by *Nicotiana tabacum* protoplasts: chemical and biochemical aspects. *Plant Sci Lett* 22, 47-57.
- Blomberg, A., Blomberg, L., Norbeck, J., Fey, S.J., Larsen, P.M., Larsen, M., Roepstorff, P., Degand, H., Boutry, M., Posch, A., and et al. (1995). Interlaboratory reproducibility of yeast protein patterns analyzed by immobilized pH gradient two-dimensional gel electrophoresis. *Electrophoresis* 16, 1935-1945.



- Bradshaw, R.A., Burlingame, A.L., Carr, S., and Aebersold, R. (2005). Protein identification: the good, the bad, and the ugly. *Mol Cell Proteomics* 4, 1221-1222.
- Brady, S.M., Orlando, D.A., Lee, J.Y., Wang, J.Y., Koch, J., Dinneny, J.R., Mace, D., Ohler, U., and Benfey, P.N. (2007). A high-resolution root spatiotemporal map reveals dominant expression patterns. *Science (New York, N Y)* 318, 801-806.
- Brar, D.S., and Khush, G.S. (2002). Transferring genes from wild species into rice. In *Quantitative genetics, genomics and plant breeding.*, M.S. Kang, ed (Oxford: CABI).
- Brett, C.T. (1978). Synthesis of beta-(1-->3)-Glucan from Extracellular Uridine Diphosphate Glucose as a Wound Response in Suspension-cultured Soybean Cells. *Plant Physiol* 62, 377-382.
- Bridges, S.M., Magee, G.B., Wang, N., Williams, W.P., Burgess, S.C., and Nanduri, B. (2007). ProtQuant: a tool for the label-free quantification of MudPIT proteomics data. *BMC Bioinformatics* 8 Suppl 7, S24.
- Bultman, S.J., Gebuhr, T.C., Pan, H., Svoboda, P., Schultz, R.M., and Magnuson, T. (2006). Maternal BRG1 regulates zygotic genome activation in the mouse. *Genes & Development* 20, 1744-1754.
- Burgess-Beusse, B., Farrell, C., Gaszner, M., Litt, M., Mutskov, V., Recillas-Targa, F., Simpson, M., West, A., and Felsenfeld, G. (2002). The insulation of genes from external enhancers and silencing chromatin. *Proc Natl Acad Sci U S A* 99 Suppl 4, 16433-16437.
- Buza, J.J., and Burgess, S.C. (2008). Different signaling pathways expressed by chicken naive CD4(+) T cells, CD4(+) lymphocytes activated with staphylococcal enterotoxin B, and those malignantly transformed by Marek's disease virus. *Journal Of Proteome Research* 7, 2380-2387.
- Cairns, B.R. (2005). Chromatin remodeling complexes: strength in diversity, precision through specialization. *Current Opinion in Genetics & Development* 15, 185-190.
- Carpita, N.C., and McCann, M.C. (2000). The cell wall. In *Biochemistry and Molecular Biology of Plants*, B.B. Buchanan, W. Gruissem, and R.L. Jones, eds (Rockville, MD, USA: American Society of Plant Physiologists), pp. 52-109.
- Chamovitz, D.A., Wei, N., Osterlund, M.T., von Arnim, A.G., Staub, J.M., Matsui, M., and Deng, X.W. (1996). The COP9 complex, a novel multisubunit nuclear regulator involved in light control of a plant developmental switch. *Cell* 86, 115-121.

- Chaudhry, F., Guerin, C., von Witsch, M., Blanchoin, L., and Staiger, C.J. (2007). Identification of Arabidopsis cyclase-associated protein 1 as the first nucleotide exchange factor for plant actin. *Molecular Biology of the Cell* 18, 3002-3014.
- Chen, Z.J., and Tian, L. (2007). Roles of dynamic and reversible histone acetylation in plant development and polyploidy. *Biochim Biophys Acta* 1769, 295-307.
- Cheng, C., Motohashi, R., Tsuchimoto, S., Fukuta, Y., Ohtsubo, H., and Ohtsubo, E. (2003). Polyphyletic origin of cultivated rice: based on the interspersed pattern of SINEs. *Mol Biol Evol* 20, 67-75.
- Cheng, X.D., and Zhang, X. (2007). Structural dynamics of protein lysine methylation and demethylation. *Mutation Research-Fundamental and Molecular Mechanisms of Mutagenesis* 618, 102-115.
- Childs, N.W. (2004). Production and Utilization of rice. In: *Rice Chemistry and Technology*. Ed: Champagne, Elaine. T., St Paul, Minnesota. pp1-20.
- Chitteti, B.R., and Peng, Z. (2007a). Proteome and Phosphoproteome Differential Expression under Salinity Stress in Rice (*Oryza sativa*) Roots. *J Proteome Res* 6, 1718-1727.
- Chitteti, B.R., and Peng, Z. (2007b). Proteome and phosphoproteome dynamic change during cell dedifferentiation in Arabidopsis. *Proteomics* 7, 1473-1500.
- Chitteti, B.R., Tan, F., Mujahid, H., Magee, B.G., Bridges, S.M., and Peng, Z. (2008). Comparative analysis of proteome differential regulation during cell dedifferentiation in Arabidopsis. *Proteomics* DOI 10.1002/pmic.200701149.
- Chivasa, S., Ndimba, B.K., Simon, W.J., Robertson, D., Yu, X.L., Yu, X.L., Knox, J.P., Bolwell, P., and Slabas, A.R. (2002). Proteomic analysis of the Arabidopsis thaliana cell wall. *Electrophoresis* 23, 1754-1765.
- Chu, D.S., Liu, H.B., Nix, P., Wu, T.F., Ralston, E.J., Yates, J.R., and Meyer, B.J. (2006). Sperm chromatin proteomics identifies evolutionarily conserved fertility factors. *Nature* 443, 101-105.
- Cohen, B.A., Mitra, R.D., Hughes, J.D., and Church, G.M. (2000). A computational analysis of whole-genome expression data reveals chromosomal domains of gene expression. *Nat Genet* 26, 183-186.
- Cosgrove, D.J. (2005). Growth of the plant cell wall. *Nature Reviews Molecular Cell Biology* 6, 850-861.
- Cottrell, J.S. (1994). Protein identification by peptide mass fingerprinting. *Pept Res* 7, 115-124.

- Damm, B., and Willmitzer, L. (1988). Regeneration of fertile plants from protoplasts of different *Arabidopsis thaliana* genotypes. *Molecular and General Genetics* 213, 15-20.
- Dasso, M., and Pu, R.T. (1998). Nuclear transport: run by Ran? *American Journal Of Human Genetics* 63, 311-316.
- Davey, M.R., Anthony, P., Power, J.B., and Lowe, K.C. (2005). Plant protoplasts: status and biotechnological perspectives. *Biotechnology Advances* 23, 131-171.
- Derbyshire, P., Findlay, K., McCann, M.C., and Roberts, K. (2007). Cell elongation in *Arabidopsis* hypocotyls involves dynamic changes in cell wall thickness. *J Exp Bot* 58, 2079-2089.
- DILDAY, R.H. (1990). Contribution of ancestral lines in the development of new cultivars of rice. *Crop Sci.* 30, 905-911.
- Du, Z., Zhou, X., Ling, Y., Zhang, Z., and Su, Z. (2010). agriGO: a GO analysis toolkit for the agricultural community. *Nucleic Acids Res* 38, 64-70.
- Egelhofer, V., Gobom, J., Seitz, H., Giavalisco, P., Lehrach, H., and Nordhoff, E. (2002). Protein identification by MALDI-TOF-MS peptide mapping: a new strategy. *Anal Chem* 74, 1760-1771.
- Eng, J.K., McCormack, A.L., and Yates, J.R. (1994). An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J Am Soc Mass Spectrom* 5, 14.
- Espada, J., and Esteller, M. (2007). Epogenetics control of nuclear architecture. *Cellular and Molecular Life Sciences* 10.1007/s00018-007-6358-x.
- Farrokhi, N., Burton, R.A., Brownfield, L., Hrmova, M., Wilson, S.M., Bacic, A., and Fincher, G.B. (2006). Plant cell wall biosynthesis: genetic, biochemical and functional genomics approaches to the identification of key genes. *Plant Biotechnology Journal* 4, 145-167.
- Ficarro, S.B., McClelland, M.L., Stukenberg, P.T., Burke, D.J., Ross, M.M., Shabanowitz, J., Hunt, D.F., and White, F.M. (2002). Phosphoproteome analysis by mass spectrometry and its application to *Saccharomyces cerevisiae*. *Nature Biotechnology* 20, 301-305.
- Finch, J.T., and Klug, A. (1976). Solenoidal model for superstructure in chromatin. *Proc Natl Acad Sci U S A* 73, 1897-1901.
- Fischle, W., Wang, Y., and Allis, C.D. (2003). Histone and chromatin cross-talk. *Curr Opin Cell Biol* 15, 172-183.

- Freshour, G., Clay, R.P., Fuller, M.S., Albersheim, P., Darvill, A.G., and Hahn, M.G. (1996). Developmental and Tissue-Specific Structural Alterations of the Cell-Wall Polysaccharides of *Arabidopsis thaliana* Roots. *Plant Physiol* 110, 1413-1429.
- Gangaraju, V.K., and Bartholomew, B. (2007). Mechanisms of ATP dependent chromatin remodeling. *Mutat Res* 618, 3-17.
- Gerasimova, T.I., Byrd, K., and Corces, V.G. (2000). A chromatin insulator determines the nuclear localization of DNA. *Mol Cell* 6, 1025-1035.
- Ghosh Biswas, G.C., Iglesias, V.A., Datta, S.K., and Potrykus, I. (1994). Transgenic *Indica* rice (*Oryza sativa* L.) plants obtained by direct gene transfer to protoplasts. *J Biotechnol* 32, 1-10.
- Glissmann, K., and Conrad, R. (2000). Fermentation pattern of methanogenic degradation of rice straw in anoxic paddy soil. *FEMS Microbiol Ecol* 31, 117-126.
- Grafi, G. (2004). How cells dedifferentiate: a lesson from plants. *Developmental Biology* 268, 1-6.
- Grunstein, M. (1997). Histone acetylation in chromatin structure and transcription. *Nature* 389, 349-352.
- Gruss, O.J., and Vernos, I. (2004). The mechanism of spindle assembly: functions of Ran and its target TPX2. *Journal of Cell Biology* 166, 949-955.
- Gurley, L.R., Valdez, J.G., and Buchanan, J.S. (1995). Characterization of the mitotic specific phosphorylation site of histone H1. Absence of a consensus sequence for the p34cdc2/cyclin B kinase. *J Biol Chem* 270, 27653-27660.
- Gygi, S.P., Rochon, Y., Franza, B.R., and Aebersold, R. (1999). Correlation between protein and mRNA abundance in yeast. *Mol Cell Biol* 19, 1720-1730.
- Hake, S.B., and Allis, C.D. (2006). Histone H3 variants and their potential role in indexing mammalian genomes: The "H3 barcode hypothesis". *Proc Natl Acad Sci U S A* 103, 6428-6435.
- Hall, D.A., Zhu, H., Zhu, X.W., Royce, T., Gerstein, M., and Snyder, M. (2004). Regulation of gene expression by a metabolic enzyme. *Science* 306, 482-484.
- Hart, C., Schulenberg, B., Steinberg, T.H., Leung, W.Y., and Patton, W.F. (2003). Detection of glycoproteins in polyacrylamide gels and on electroblots using Pro-Q Emerald 488 dye, a fluorescent periodate Schiff-base stain. *Electrophoresis* 24, 588-598.

- He, G.M., Zhu, X.P., Elling, A.A., Chen, L.B., Wang, X.F., Guo, L., Liang, M.Z., He, H., Zhang, H.Y., Chen, F.F., Qi, Y.J., Chen, R.S., and Deng, X.W. (2010). Global Epigenetic and Transcriptional Trends among Two Rice Subspecies and Their Reciprocal Hybrids. *Plant Cell* 22, 17-33.
- He, X., and Zhang, J. (2006). Why do hubs tend to be essential in protein networks? *Plos Genetics* 2, e88.
- Hecht, K. (1912). Studien uber den Vorgang der Plasmolyse. *Beitr. Biol. Pflanz* 11, 133-189.
- Held, M.A., Penning, B., Brandt, A.S., Kessans, S.A., Yong, W., Scofield, S.R., and Carpita, N.C. (2008). Small-interfering RNAs from natural antisense transcripts derived from a cellulose synthase gene modulate cell wall biosynthesis in barley. *Proc Natl Acad Sci U S A* 105, 20534-20539.
- Horine, R.K., and Ruesink, A.W. (1972). Cell Wall Regeneration around Protoplasts Isolated from *Convolvulus* Tissue Culture. *Plant Physiol* 50, 438-445.
- Horn, P.J., and Peterson, C.L. (2002). Molecular biology. Chromatin higher order folding--wrapping up transcription. *Science* 297, 1824-1827.
- Horwitz, G.A., Zhang, K., McBrien, M.A., Grunstein, M., Kurdistani, S.K., and Berk, A.J. (2008). Adenovirus Small e1a Alters Global Patterns of Histone Modification. *science* 321, 1084-1085.
- Hozak, P., Hassan, A.B., Jackson, D.A., and Cook, P.R. (1993). Visualization of replication factories attached to nucleoskeleton. *Cell* 73, 361-373.
- Hsieh, J., and Gage, F.H. (2005). Chromatin remodeling in neural development and plasticity. *Current Opinion in Cell Biology* 17, 664-671.
- Hurkman, W.J., and Tanaka, C.K. (1986). Solubilization of Plant Membrane Proteins for Analysis by Two-Dimensional Gel Electrophoresis. *Plant Physiol.* 81, 802-806.
- Iizuka, M., and Smith, M.M. (2003). Functional consequences of histone modifications. *Current opinion in genetics & development* 13, 154-160.
- Initiative, T.A. (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408, 796-815.
- Ishii, K., Arib, G., Lin, C., Van Houwe, G., and Laemmli, U.K. (2002). Chromatin boundaries in budding yeast: the nuclear pore connection. *Cell* 109, 551-562.
- Jarillo, J.A., Pineiro, M., Cubas, P., and Martinez-Zapater, J.M. (2009). Chromatin remodeling in plant development. *Int J Dev Biol* 53, 1581-1596.

- Jean H. Gould, R.L.P.a.W.M.D. (1986). Isolation and culture of anucleate protoplasts from cotton fiber; assessment of viability and analysis of regenerated wall polymers. *PLANT CELL, TISSUE AND ORGAN CULTURE* 6, 61-72.
- Jenuwein, T., and Allis, C.D. (2001). Translating the histone code. *Science* 293, 1074-1080.
- Johnson, L., Mollah, S., Garcia, B.A., Muratore, T.L., Shabanowitz, J., Hunt, D.F., and Jacobsen, S.E. (2004). Mass spectrometry analysis of Arabidopsis histone H3 reveals distinct combinations of post-translational modifications. *Nucleic Acids Research* 32, 6511-6518.
- Jones, A.M., Bennett, M.H., Mansfield, J.W., and Grant, M. (2006). Analysis of the defence phosphoproteome of Arabidopsis thaliana using differential mass tagging. *Proteomics* 6, 4155-4165.
- Kaczkowski, J. (2003). Structure, function and metabolism of plant cell wall. *Acta Physiologiae Plantarum* 25, 287-305.
- Kamakaka, R.T., and Biggins, S. (2005). Histone variants: deviants? *Genes Dev* 19, 295-310.
- Kelleher, N.L. (2004). Top-down proteomics. *Anal Chem* 76, 197A-203A.
- Khorasanizadeh, S. (2004). The nucleosome: from genomic organization to genomic regulation. *Cell* 116, 259-272.
- Kingston, R.E., and Narlikar, G.J. (1999). ATP-dependent remodeling and acetylation as regulators of chromatin fluidity. *Genes Dev* 13, 2339-2352.
- Kitano, H. (2002). Systems biology: a brief overview. *Science* 295, 1662-1664.
- Klose, J., and Kobalz, U. (1995). Two-dimensional electrophoresis of proteins: an updated protocol and implications for a functional analysis of the genome. *Electrophoresis* 16, 1034-1059.
- Komatsu, S., Muhammad, A., and Rakwal, R. (1999). Separation and characterization of proteins from green and etiolated shoots of rice (*Oryza sativa* L.): towards a rice proteome. *Electrophoresis* 20, 630-636.
- Konishi, H., and Komatsu, S. (2003). A proteomics approach to investigating promotive effects of brassinolide on lamina inclination and root growth in rice seedlings. *Biol Pharm Bull* 26, 401-408.
- Korenbaum, E., Nordberg, P., Bjorkegren-Sjogren, C., Schutt, C.E., Lindberg, U., and Karlsson, R. (1998). The role of profilin in actin polymerization and nucleotide exchange. *Biochemistry* 37, 9274-9283.

- Kornberg, R.D. (1974). Chromatin structure: a repeating unit of histones and DNA. *Science* 184, 868-871.
- Kornberg, R.D., and Lorch, Y. (1999). Chromatin-modifying and -remodeling complexes. *Curr Opin Genet Dev* 9, 148-151.
- Kruiswijk, T., de Hey, J.T., and Planta, R.J. (1978). Modification of yeast ribosomal proteins. Phosphorylation. *Biochem J* 175, 213-219.
- Kurdistani, S.K., Tavazoie, S., and Grunstein, M. (2004). Mapping global histone acetylation patterns to gene expression. *Cell* 117, 721-733.
- Kwon, H.K., Yokoyama, R., and Nishitani, K. (2005). A proteomic approach to apoplastic proteins involved in cell wall regeneration in protoplasts of *Arabidopsis* suspension-cultured cells. *Plant & Cell Physiology* 46, 843-857.
- Labrador, M., and Corces, V.G. (2002). Setting the boundaries of chromatin domains and nuclear organization. *Cell* 111, 151-154.
- Laemmli, U.K. (1978). Levels of organization of the DNA in eucaryotic chromosomes. *Pharmacol Rev* 30, 469-476.
- Larsen, M.R., Thingholm, T.E., Jensen, O.N., Roepstorff, P., and Jorgensen, T.J. (2005). Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. *Mol Cell Proteomics* 4, 873-886.
- Lee, J.Y., Colinas, J., Wang, J.Y., Mace, D., Ohler, U., and Benfey, P.N. (2006). Transcriptional and posttranscriptional regulation of transcription factor expression in *Arabidopsis* roots. *Proc Natl Acad Sci U S A* 103, 6055-6060.
- Lee, L., Schroll, R.E., Grimes, H.D., and Hodges, T.K. (1989). Plant regeneration from indica rice (*Oryza sativa* L.) protoplasts. *PLANTA* 178, 325-333.
- Lee, T.J., Shultz, R.W., Hanley-Bowdoin, L., and Thompson, W.F. (2004). Establishment of rapidly proliferating rice cell suspension culture and its characterization by fluorescence-activated cell sorting analysis. *Plant Molecular Biology Reporter* 22, 259-267.
- Li, B., Carey, M., and Workman, J.L. (2007). The role of chromatin during transcription. *Cell* 128, 707-719.
- Li, G., Sudlow, G., and Belmont, A.S. (1998). Interphase cell cycle dynamics of a late-replicating, heterochromatic homogeneously staining region: precise choreography of condensation/decondensation and nuclear positioning. *J Cell Biol* 140, 975-989.

- Li, G., Nallamilli, B.R., Tan, F., and Peng, Z. (2008). Removal of high-abundance proteins for nuclear subproteome studies in rice (*Oryza sativa*) endosperm. *Electrophoresis* 29, 604-617.
- Libault, M., Tessadori, F., Germann, S., Snijder, B., Fransz, P., and Gaudin, V. (2005). The *Arabidopsis* LHP1 protein is a component of euchromatin. *PLANTA* 222, 910-925.
- Lim, H., Eng, J., Yates, J.R., 3rd, Tollaksen, S.L., Giometti, C.S., Holden, J.F., Adams, M.W., Reich, C.I., Olsen, G.J., and Hays, L.G. (2003). Identification of 2D-gel proteins: a comparison of MALDI/TOF peptide mass mapping to  $\mu$ LC-ESI tandem mass spectrometry. *J Am Soc Mass Spectrom* 14, 957-970.
- Lin, D., Tabb, D.L., and Yates, J.R. (2003). Large-scale protein identification using mass spectrometry. *Biochimica Et Biophysica Acta-Proteins and Proteomics* 1646, 1-10.
- Link, A.J., Eng, J., Schieltz, D.M., Carmack, E., Mize, G.J., Morris, D.R., Garvik, B.M., and Yates, J.R., 3rd. (1999). Direct analysis of protein complexes using mass spectrometry. *Nat Biotechnol* 17, 676-682.
- Liu, B., Lin, Y., Darwanto, A., Song, X., Xu, G., and Zhang, K. (2009). Identification and characterization of propionylation at histone H3 lysine 23 in mammalian cells. *J Biol Chem* 284, 32288-32295.
- Liu, H., Lin, D., and Yates, J.R., 3rd. (2002). Multidimensional separations for protein/peptide analysis in the post-genomic era. *Biotechniques* 32, 898, 900, 902 passim.
- Liu, H., Sadygov, R.G., and Yates, J.R., 3rd. (2004). A model for random sampling and estimation of relative protein abundance in shotgun proteomics. *Anal Chem* 76, 4193-4201.
- Loo, J.A., Udseth, H.R., and Smith, R.D. (1989). Peptide and protein analysis by electrospray ionization-mass spectrometry and capillary electrophoresis-mass spectrometry. *Anal Biochem* 179, 404-412.
- Lopez, M.F., Berggren, K., Chernokalskaya, E., Lazarev, A., Robinson, M., and Patton, W.F. (2000). A comparison of silver stain and SYPRO Ruby Protein Gel Stain with respect to protein detection in two-dimensional gels and identification by peptide mass profiling. *Electrophoresis* 21, 3673-3683.
- Luger, K., Rechsteiner, T.J., Flaus, A.J., Waye, M.M., and Richmond, T.J. (1997). Characterization of nucleosome core particles containing histone proteins made in bacteria. *J Mol Biol* 272, 301-311.



- Maccarrone, G., Kolb, N., Teplytska, L., Birg, I., Zollinger, R., Holsboer, F., and Turck, C.W. (2006). Phosphopeptide enrichment by IEF. *Electrophoresis* 27, 4585-4595.
- Mackill, D.J. (1995). Classifiting japonica rice cultivars with RAPD markers. *Corp Sci* 35.
- Mallick, P., Schirle, M., Chen, S.S., Flory, M.R., Lee, H., Martin, D., Ranish, J., Raught, B., Schmitt, R., Werner, T., Kuster, B., and Aebersold, R. (2007). Computational prediction of proteotypic peptides for quantitative proteomics. *Nat Biotechnol* 25, 125-131.
- Martin, C., and Zhang, Y. (2005). The diverse functions of histone lysine methylation. *Nature Reviews Molecular Cell Biology* 6, 838-849.
- Marx, J.L. (1987). Rice Plants Regenerated From Protoplasts: The ability to regenerate rice protoplasts means that for the first time a major cereal will become subject to modern biotechnological methods. *Science* 235, 31-32.

Matsumoto, T., Wu, J.Z., Kanamori, H., Katayose, Y., Fujisawa, M., Namiki, N., Mizuno, H., Yamamoto, K., Antonio, B.A., Baba, T., Sakata, K., Nagamura, Y., Aoki, H., Arikawa, K., Arita, K., Bito, T., Chiden, Y., Fujitsuka, N., Fukunaka, R., Hamada, M., Harada, C., Hayashi, A., Hijishita, S., Honda, M., Hosokawa, S., Ichikawa, Y., Idonuma, A., Iijima, M., Ikeda, M., Ikeno, M., Ito, K., Ito, S., Ito, T., Ito, Y., Ito, Y., Iwabuchi, A., Kamiya, K., Karasawa, W., Kurita, K., Katagiri, S., Kikuta, A., Kobayashi, H., Kobayashi, N., Machita, K., Maehara, T., Masukawa, M., Mizubayashi, T., Mukai, Y., Nagasaki, H., Nagata, Y., Naito, S., Nakashima, M., Nakama, Y., Nakamichi, Y., Nakamura, M., Meguro, A., Negishi, M., Ohta, I., Ohta, T., Okamoto, M., Ono, N., Saji, S., Sakaguchi, M., Sakai, K., Shibata, M., Shimokawa, T., Song, J.Y., Takazaki, Y., Terasawa, K., Tsugane, M., Tsuji, K., Ueda, S., Waki, K., Yamagata, H., Yamamoto, M., Yamamoto, S., Yamane, H., Yoshiki, S., Yoshihara, R., Yukawa, K., Zhong, H.S., Yano, M., Sasaki, T., Yuan, Q.P., Shu, O.T., Liu, J., Jones, K.M., Gansberger, K., Moffat, K., Hill, J., Bera, J., Fadrosch, D., Jin, S.H., Johri, S., Kim, M., Overton, L., Reardon, M., Tsitrin, T., Vuong, H., Weaver, B., Ciecko, A., Tallon, L., Jackson, J., Pai, G., Van Aken, S., Utterback, T., Reidmuller, S., Feldblyum, T., Hsiao, J., Zismann, V., Iobst, S., de Vazeille, A.R., Buell, C.R., Ying, K., Li, Y., Lu, T.T., Huang, Y.C., Zhao, Q., Feng, Q., Zhang, L., Zhu, J.J., Weng, Q.J., Mu, J., Lu, Y.Q., Fan, D.L., Liu, Y.L., Guan, J.P., Zhang, Y.J., Yu, S.L., Liu, X.H., Zhang, Y., Hong, G.F., Han, B., Choisine, N., Demange, N., Orjeda, G., Samain, S., Cattolico, L., Pelletier, E., Couloux, A., Segurens, B., Wincker, P., D'Hont, A., Scarpelli, C., Weissenbach, J., Salanoubat, M., Quetier, F., Yu, Y., Kim, H.R., Rambo, T., Currie, J., Collura, K., Luo, M.Z., Yang, T.J., Ammiraju, J.S.S., Engler, F., Soderlund, C., Wing, R.A., Palmer, L.E., de la Bastide, M., Spiegel, L., Nascimento, L., Zutavern, T., O'Shaughnessy, A., Dike, S., Dedhia, N., Preston, R., Balija, V., McCombie, W.R., Chow, T.Y., Chen, H.H., Chung, M.C., Chen, C.S., Shaw, J.F., Wu, H.P., Hsiao, K.J., Chao, Y.T., Chu, M.K., Cheng, C.H., Hour, A.L., Lee, P.F., Lin, S.J., Lin, Y.C., Liou, J.Y., Liu, S.M., Hsing, Y.I., Raghuvanshi, S., Mohanty, A., Bharti, A.K., Gaur, A., Gupta, V., Kumar, D., Ravi, V., Vij, S., Kapur, A., Khurana, P., Khurana, P., Khurana, J.P., Tyagi, A.K., Gaikwad, K., Singh, A., Dalal, V., Srivastava, S., Dixit, A., Pal, A.K., Ghazi, I.A., Yadav, M., Pandit, A., Bhargava, A., Sureshbabu, K., Batra, K., Sharma, T.R., Mohapatra, T., Singh, N.K., Messing, J., Nelson, A.B., Fuks, G., Kavchok, S., Keizer, G., Llaca, E.L.V., Song, R.T., Tanyolac, B., Young, S., Il, K.H., Hahn, J.H., Sangsakoo, G., Vanavichit, A., de Mattos, L.A.T., Zimmer, P.D., Malone, G., Dellagostin, O., de Oliveira, A.C., Bevan, M., Bancroft, I., Minx, P., Cordum, H., Wilson, R., Cheng, Z.K., Jin, W.W., Jiang, J.M., Leong, S.A., Iwama, H., Gojobori, T., Itoh, T., Niimura, Y., Fujii, Y., Habara, T., Sakai, H., Sato, Y., Wilson, G., Kumar, K., McCouch, S., Juretic, N., Hoen, D., Wright, S., Bruskiewich, R., Bureau, T., Miyao, A., Hirochika, H., Nishikawa, T., Kadowaki, K., Sugiura, M., and Project, I.R.G.S. (2005). The map-based sequence of the rice genome. *Nature* 436, 793-800.

- Mazzola, J.L., and Sirover, M.A. (2002). Alteration of intracellular structure and function of glyceraldehyde-3-phosphate dehydrogenase: A common phenotype of neurodegenerative disorders? *Neurotoxicology* 23, 603-609.
- McCarthy, F.M., Burgess, S.C., van den Berg, B.H., Koter, M.D., and Pharr, G.T. (2005). Differential detergent fractionation for non-electrophoretic eukaryote cell proteomics. *J Proteome Res* 4, 316-324.
- McCarthy, F.M., Bridges, S.M., Wang, N., Magee, G.B., Williams, W.P., Luthe, D.S., and Burgess, S.C. (2007). AgBase: a unified resource for functional analysis in agriculture. *Nucleic Acids Research* 35, D599-D603.
- McLafferty, F.W., Breuker, K., Jin, M., Han, X., Infusini, G., Jiang, H., Kong, X., and Begley, T.P. (2007). Top-down MS, a powerful complement to the high capabilities of proteolysis proteomics. *FEBS J* 274, 6256-6268.
- Meyerowitz, E.M. (2002). Comparative genomics - Plants compared to animals: The broadest comparative study of development. *Science* 295, 1482-1485.
- Minard, M.E., Jain, A.K., and Barton, M.C. (2009). Analysis of epigenetic alterations to chromatin during development. *Genesis* 47, 559-572.
- Minic, Z., and Jouanin, L. (2006). Plant glycoside hydrolases involved in cell wall polysaccharide degradation. *Plant Physiology And Biochemistry: PPB / Societe Francaise De Physiologie Vegetale* 44, 435-449.
- Mishra, A.K., and Colvin, J.R. (1969). The formation of wall-like envelopes by isolated tomato-fruit protoplasts. *Protoplasma* 67, 295-305.
- Misteli, T. (2001). Protein dynamics: implications for nuclear architecture and gene expression. *Science* 291, 843-847.
- Nanduri, B., Lawrence, M.L., Vanguri, S., and Burgess, S.C. (2005). Proteomic analysis using an unfinished bacterial genome: the effects of subminimum inhibitory concentrations of antibiotics on *Mannheimia haemolytica* virulence factor expression. *Proteomics* 5, 4852-4863.
- Nanduri, B., Shah, P., Ramkumar, M., Allen, E.B., Swiatlo, E., Burgess, S.C., and Lawrence, M.L. (2008). Quantitative analysis of *Streptococcus pneumoniae* TIGR4 response to in vitro iron restriction by 2-D LC ESI MS/MS. *Proteomics* 8, 2104-2114.
- Ndamukong, I., Chetram, A., Saleh, A., and Avramova, Z. (2009). Wall-modifying genes regulated by the *Arabidopsis* homolog of trithorax, ATX1: repression of the XTH33 gene as a test case. *Plant J* 58, 541-553.

- Ndimba, B.K., Chivasa, S., Hamilton, J.M., Simon, W.J., and Slabas, A.R. (2003). Proteomic analysis of changes in the extracellular matrix of Arabidopsis cell suspension cultures induced by fungal elicitors. *Proteomics* 3, 1047-1059.
- Negrutiu, I., Hinnisdaels, S., Cammaerts, D., Cherdshewasart, W., Gharti-Chhetri, G., and Jacobs, M. (1992). Plant protoplasts as genetic tool: selectable markers for developmental studies. *Int. J. Dev. Biol.* 36, 73-84.
- Nemeth, A., and Langst, G. (2004). Chromatin higher order structure: opening up chromatin for transcription. *Brief Funct Genomic Proteomic* 2, 334-343.
- Ng, H.H., and Bird, A. (2000). Histone deacetylases: silencers for hire. *Trends in biochemical sciences* 25, 121-126.
- Nikitin, A., Egorov, S., Daraselia, N., and Mazo, I. (2003). Pathway studio - the analysis and navigation of molecular networks. *Bioinformatics* 19, 2155-2157.
- Novichkova, S., Egorov, S., and Daraselia, N. (2003). MedScan, a natural language processing engine for MEDLINE abstracts. *Bioinformatics* 19, 1699-1706.
- Nowak, S.J., and Corces, V.G. (2004). Phosphorylation of histone H3: a balancing act between chromosome condensation and transcriptional activation. *Trends Genet* 20, 214-220.
- O'Farrell, P.H. (1975). High resolution two-dimensional electrophoresis of proteins. *J Biol Chem* 250, 4007-4021.
- Offermann, S., Dreesen, B., Horst, I., Danker, T., Jaskiewicz, M., and Peterhansel, C. (2008). Developmental and environmental signals induce distinct histone acetylation profiles on distal and proximal promoter elements of the C4-Pepc gene in maize. *Genetics* 179, 1891-1901.
- Old, W.M., Meyer-Arendt, K., Aveline-Wolf, L., Pierce, K.G., Mendoza, A., Sevinsky, J.R., Resing, K.A., and Ahn, N.G. (2005). Comparison of label-free methods for quantifying human proteins by shotgun proteomics. *Mol Cell Proteomics* 4, 1487-1502.
- Olins, A.L., and Olins, D.E. (1974). Spheroid chromatin units (v bodies). *Science* 183, 330-332.
- Ondrej, V., Kitner, M., Dolezalova, I., Nadvornik, P., Navratilova, B., and Lebeda, A. (2009). Chromatin structural rearrangement during dedifferentiation of protoplasts of *Cucumis sativus* L. *Molecules and Cells* 27, 443-447.

- Ouyang, S., Zhu, W., Hamilton, J., Lin, H., Campbell, M., Childs, K., Thibaud-Nissen, F., Malek, R.L., Lee, Y., Zheng, L., Orvis, J., Haas, B., Wortman, J., and Buell, C.R. (2007). The TIGR Rice Genome Annotation Resource: improvements and new features. *Nucleic Acids Res* 35, D883-887.
- Peng, Z., Serino, G., and Deng, X.W. (2001a). A role of Arabidopsis COP9 signalosome in multifaceted developmental processes revealed by the characterization of its subunit 3. *Development* 128, 4277-4288.
- Peng, Z., Serino, G., and Deng, X.W. (2001b). Molecular characterization of subunit 6 of the COP9 signalosome and its role in multifaceted developmental processes in Arabidopsis. *Plant Cell* 13, 2393-2407.
- Peterson, C.L., and Laniel, M.A. (2004). Histones and histone modifications. *Curr Biol* 14, R546-551.
- Qin, S., Wang, Q., Ray, A., Wani, G., Zhao, Q., Bhaumik, S.R., and Wani, A.A. (2009). Sem1p and Ubp6p orchestrate telomeric silencing by modulating histone H2B ubiquitination and H3 acetylation. *Nucleic Acids Res* 37, 1843-1853.
- Rakwal, R., and Komatsu, S. (2000). Role of jasmonate in the rice (*Oryza sativa* L.) self-defense mechanism using proteome analysis. *Electrophoresis* 21, 2492-2500.
- Raven, P.H., Evert, R.F., and Eichhorn, S.E. (2005). *Biology of Plants*. (New York: W. H. Freeman Company).
- Rice, J.C., and Allis, C.D. (2001). Gene regulation - Code of silence. *Nature* 414, 258-+.
- Richards, E.J., and Elgin, S.C. (2002). Epigenetic codes for heterochromatin formation and silencing: rounding up the usual suspects. *Cell* 108, 489-500.
- Righetti, P.G., and Bossi, A. (1997). Isoelectric focusing in immobilized pH gradients: an update. *J Chromatogr B Biomed Sci Appl* 699, 77-89.
- Rivett, A.J., Bose, S., Brooks, P., and Broadfoot, K.I. (2001). Regulation of proteasome complexes by gamma-interferon and phosphorylation. *Biochimie* 83, 363-366.
- Robinson, P.J., and Rhodes, D. (2006). Structure of the '30 nm' chromatin fibre: a key role for the linker histone. *Curr Opin Struct Biol* 16, 336-343.
- Saravanan, R.S., and Rose, J.K. (2004). A critical evaluation of sample extraction techniques for enhanced proteomic analysis of recalcitrant plant tissues. *Proteomics* 4, 2522-2532.
- Scheer, U., Hinssen, H., Franke, W.W., and Jockusch, B.M. (1984). Microinjection of actin-binding proteins and actin antibodies demonstrates involvement of nuclear actin in transcription of lampbrush chromosomes. *Cell* 39, 111-122.

- Schmidt, M.W., Houseman, A., Ivanov, A.R., and Wolf, D.A. (2007). Comparative proteomic and transcriptomic profiling of the fission yeast *Schizosaccharomyces pombe*. *Mol Syst Biol* 3, 79.
- Schulenberg, B., Goodman, T.N., Aggeler, R., Capaldi, R.A., and Patton, W.F. (2004). Characterization of dynamic and steady-state protein phosphorylation using a fluorescent phosphoprotein gel stain and mass spectrometry. *Electrophoresis* 25, 2526-2532.
- Shechter, D., Dormann, H.L., Allis, C.D., and Hake, S.B. (2007). Extraction, purification and analysis of histones. *Nat Protoc* 2, 1445-1457.
- Sheen, J. (2001). Signal transduction in maize and *Arabidopsis* mesophyll protoplasts. *Plant Physiology* 127, 1466-1475.
- Shelby, R.D., Hahn, K.M., and Sullivan, K.F. (1996). Dynamic elastic behavior of alpha-satellite DNA domains visualized in situ in living human cells. *J Cell Biol* 135, 545-557.
- Shen, S.H., Jing, Y.X., and Kuang, T.Y. (2003). Proteomics approach to identify wound-response related proteins from rice leaf sheath. *Proteomics* 3, 527-535.
- Shiio, Y., Eisenman, R.N., Yi, E.C., Donohoe, S., Goodlett, D.R., and Aebersold, R. (2003). Quantitative proteomic analysis of chromatin-associated factors. *Journal of the American Society for Mass Spectrometry* 14, 696-703.
- Shogren-Knaak, M., and Peterson, C.L. (2006). Switching on chromatin: mechanistic role of histone H4-K16 acetylation. *Cell Cycle* 5, 1361-1365.
- Sims, R.J., 3rd, Nishioka, K., and Reinberg, D. (2003). Histone lysine methylation: a signature for chromatin function. *Trends Genet* 19, 629-639.
- Smith, C.M., Gafken, P.R., Zhang, Z.L., Gottschling, D.E., Smith, J.B., and Smith, D.L. (2003). Mass spectrometric quantification of acetylation at specific lysines within the amino-terminal tail of histone H4. *Analytical Biochemistry* 316, 23-33.
- Somerville, C. (2006a). Cellulose synthesis in higher plants. *Annual Review Of Cell And Developmental Biology* 22, 53-78.
- Somerville, C. (2006b). The billion-ton biofuels vision. *Science* 312, 1277.
- Somerville, C. (2007). Biofuels. *Current Biology: CB* 17, R115-119.
- Somerville, C., Bauer, S., Brininstool, G., Facette, M., Hamann, T., Milne, J., Osborne, E., Paredes, A., Persson, S., Raab, T., Vorwerk, S., and Youngs, H. (2004). Toward a systems approach to understanding plant cell walls. *Science* 306, 2206-2211.

- Stasyk, T., Morandell, S., Bakry, R., Feuerstein, I., Huck, C.W., Stecher, G., Bonn, G.K., and Huber, L.A. (2005). Quantitative detection of phosphoproteins by combination of two-dimensional difference gel electrophoresis and phosphospecific fluorescent staining. *Electrophoresis* 26, 2850-2854.
- Steinberg, T.H., Agnew, B.J., Gee, K.R., Leung, W.Y., Goodman, T., Schulenberg, B., Hendrickson, J., Beechem, J.M., Haugland, R.P., and Patton, W.F. (2003). Global quantitative phosphoprotein analysis using multiplexed proteomics technology. *Proteomics* 3, 1128-1144.
- Strahl, B.D., and Allis, C.D. (2000). The language of covalent histone modifications. *Nature* 403, 41-45.
- Struhl, K. (1998). Histone acetylation and transcriptional regulatory mechanisms. *Genes & development* 12, 599-606.
- Sun, Z.W., and Allis, C.D. (2002). Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast. *Nature* 418, 104-108.
- Suzuki, M., Sakuda, S., and Nagasawa, H. (2007). Identification of chitin in the prismatic layer of the shell and a chitin synthase gene from the Japanese pearl oyster, *Pinctada fucata*. *Bioscience Biotechnology and Biochemistry* 71, 1735-1744.
- Tackett, A.J., Dilworth, D.J., Davey, M.J., O'Donnell, M., Aitchison, J.D., Rout, M.P., and Chait, B.T. (2005). Proteomic and genomic characterization of chromatin complexes at a boundary. *J Cell Biol* 169, 35-47.
- Tan, F., Li, G.S., Chitteti, B.R., and Peng, Z.H. (2007). Proteome and phosphoproteome analysis of chromatin associated proteins in rice (*Oryza sativa*). *Proteomics* 7, 4511-4527.
- Tessadori, F., Chupeau, M.C., Chupeau, Y., Knip, M., Germann, S., van Driel, R., Fransz, P., and Gaudin, V. (2007). Large-scale dissociation and sequential reassembly of pericentric heterochromatin in dedifferentiated *Arabidopsis* cells. *Journal Of Cell Science* 120, 1200-1208.
- Turner, B.M. (2002). Cellular memory and the histone code. *Cell* 111, 285-291.
- Uchiyama, S., Kobayashi, S., Takata, H., Ishihara, T., Hori, N., Higashi, T., Hayashihara, K., Sone, T., Higo, D., Nirasawa, T., Takao, T., Matsunaga, S., and Fukui, K. (2005). Proteome analysis of human metaphase chromosomes. *Journal of Biological Chemistry* 280, 16994-17004.
- Unlu, M., Morgan, M.E., and Minden, J.S. (1997). Difference gel electrophoresis: a single gel method for detecting changes in protein extracts. *Electrophoresis* 18, 2071-2077.

- Utomo, H.S., Croughan, S.S., and Croughan, T.P. (1995). Suspension and protoplast culture of U.S. rice cultivars. *PLANT CELL REPORTS* 15, 34-37.
- Vakoc, C.R., Mandat, S.A., Olenchock, B.A., and Blobel, G.A. (2005). Histone H3 lysine 9 methylation and HP1gamma are associated with transcription elongation through mammalian chromatin. *Mol Cell* 19, 381-391.
- van den Berg, B.H.J., Harris, T., McCarthy, F.M., Lamont, S.J., and Burgess, S.C. (2007). Non-electrophoretic differential detergent fractionation proteomics using frozen whole organs. *Rapid Communications in Mass Spectrometry* 21, 3905-3909.
- Vengerov, Y.Y., and Popenko, V.I. (1977). Changes in chromatin structure induced by EDTA treatment and partial removal of histone H1. *Nucleic Acids Res* 4, 3017-3027.
- Verma, D.P., and Hong, Z. (2001). Plant callose synthase complexes. *Plant Mol Biol* 47, 693-701.
- Vignali, M., Hassan, A.H., Neely, K.E., and Workman, J.L. (2000). ATP-dependent chromatin-remodeling complexes. *Mol Cell Biol* 20, 1899-1910.
- Vilardell, J., Goday, A., Freire, M., Torrent, M., Martinez, M., Torne, J., and Pages, M. (1990). Gene sequence, developmental expression, and protein phosphorylation of RAB-17 in maize. *Plant Mol Biol* 14, 423-432.
- Volgger, M., Lang, I., Ovecka, M., and Lichtscheidl, I. (2009). Plasmolysis and cell wall deposition in wheat root hairs under osmotic stress. *Protoplasma*.
- Vorwerk, S., Somerville, S., and Somerville, C. (2004). The role of plant cell wall polysaccharide composition in disease resistance. *Trends Plant Sci* 9, 203-209.
- Washburn, M.P., Wolters, D., and Yates, J.R., 3rd. (2001). Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat Biotechnol* 19, 242-247.
- Watson, B.S., Lei, Z.T., Dixon, R.A., and Sumner, L.W. (2004). Proteomics of *Medicago sativa* cell walls. *Phytochemistry* 65, 1709-1720.
- Wolffe, A.P., and Pruss, D. (1996). Targeting chromatin disruption: Transcription regulators that acetylate histones. *Cell* 84, 817-819.
- Wong, J.W.H., Sullivan, M.J., and Cagney, G. (2008). Computational methods for the comparative quantification of proteins in label-free LCn-MS experiments. *Briefings in Bioinformatics* 9, 156-165.



- Woodcock, C.L., Skoultchi, A.I., and Fan, Y. (2006). Role of linker histone in chromatin structure and function: H1 stoichiometry and nucleosome repeat length. *Chromosome Res* 14, 17-25.
- Yamada, Y., Yang, Z.Q., and Tang, D.T. (1986). Plant regeneration from protoplast-derived callus of rice (*Oryza sativa* L.). *Plant Cell Reports* 5, 85-88.
- Yamamoto, M., Fujishita, M., Hirata, A., and Kawano, S. (2004). Regeneration and maturation of daughter cell walls in the autospore-forming green alga *Chlorella vulgaris* (Chlorophyta, Trebouxiophyceae). *Journal of Plant Research* 117, 257-264.
- Yang, P., Chen, C., Wang, Z., Fan, B., and Chen, Z. (1999). A pathogen- and salicylic acid-induced WRKY DNA-binding activity recognizes the elicitor response element of the tobacco class I chitinase gene promoter. *The Plant Journal* 18, 141-149.
- Zhang, K., Sridhar, V.V., Zhu, J., Kapoor, A., and Zhu, J.K. (2007). Distinctive core histone post-translational modification patterns in *Arabidopsis thaliana*. *PLoS One* 2, e1210.
- Zhao, J., Morozova, N., Williams, L., Libs, L., Avivi, Y., and Grafi, G. (2001). Two phases of chromatin decondensation during dedifferentiation of plant cells - Distinction between competence for cell fate switch and a commitment for S phase. *Journal of Biological Chemistry* 276, 22772-22778.
- Zhao, K., Wang, W., Rando, O.J., Xue, Y., Swiderek, K., Kuo, A., and Crabtree, G.R. (1998). Rapid and phosphoinositol-dependent binding of the SWI/SNF-like BAF complex to chromatin after T lymphocyte receptor signaling. *Cell* 95, 625-636.
- Zheng, L., Roeder, R.G., and Luo, Y. (2003). S phase activation of the histone H2B promoter by OCA-S, a coactivator complex that contains GAPDH as a key component. *Cell* 114, 255-266.
- Zhong, R., and Ye, Z.H. (2007). Regulation of cell wall biosynthesis. *Curr Opin Plant Biol* 10, 564-572.