RESEARCH ARTICLE

Proteome and phosphoproteome analysis of chromatin associated proteins in rice (*Oryza sativa*)

Feng Tan, Guosheng Li, Brahmananda Reddy Chitteti and Zhaohua Peng

Department of Biochemistry and Molecular Biology, Mississippi State University, Mississippi State, MS, USA

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. In this report, we have developed a protocol for chromatin purification from rice suspension cells and examined proteins copurified with chromatin using both 2-DE gel and shotgun approaches. Nine hundred seventy-two distinct protein spots have been resolved on 2-DE gels and 509 proteins have been identified by MALDI-MS/MS following gel excision, which correspond to 269 unique proteins. When the chromatin copurified proteins are examined using shotgun method, 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. Furthermore, putative phosphoproteins copurified with chromatin have been examined using Pro-Q Diamond phosphoprotein stain and followed by MALDI-MS/MS. Our studies have provided valued new insight into chromatin composition in plants.

Received: June 21, 2007
Revised: August 14, 2007
Accepted: August 30, 2007

Keywords:
Chromatin / Chromatin associated proteins / Phosphoproteome / Rice (*Oryza sativa*)

1 Introduction

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 is the nucleosome, which comprises 147 bp of genomic DNA wrapped around a histone octamer, with two copies each of four histones, H2A, H2B, H3, and H4, thus giving rise to a "beads on a string" like fiber of ~10 nm in diameter [1–3]. The adjacent nucleosomes are bridged by linker DNA, associated with histone 1, establishing a further packaged ~30 nm fiber, termed "solenoid" helical fiber [4–6]. The *in vivo* chromatin structure beyond the 30 nm fiber remains poorly understood. Recent evidence has shown that folding and unfolding of chromatin have a significant impact on gene activity [7–9], and it is believed that the chromatin associated proteins are essential to the processes of structure modulation [10].

The chromatin is packed into domains with different degrees of accessibility to the transcriptional machinery: the more opened regions are relatively transcriptional active while the condensed regions are inactive [11]. Studies of gene expression maps and chromosomal localization show that clusters of similar expressed genes constitute uniformly transcribed domains [12]. It is believed that the status

Correspondence: Dr. Zhaohua Peng, Department of Biochemistry and Molecular Biology, Mississippi State University, Mail Box 9650, Mississippi State, MS 39762, USA
E-mail: zp7@BCH.msstate.edu
Fax: +1-662-325-8664

Abbreviations: CIB, chromatin isolation buffer; **GAPDH,** glyceraldehyde-3-phosphate dehydrogenase; **GO,** Gene Ontology; **NCBI,** National Center for Biotechnology Information; **NIB,** nuclei isolation buffer; **RDA,** result dependent analysis

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 opened and condensed chromatin domains [13, 14]. Insulators may be anchored to nuclear lamina or nuclear pore complexes with emanating chromatin fiber loops of active genes for transcription [13, 15–17]. Furthermore, chromatin and other nuclear components tend to quickly and transiently interact with each other [9]. 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.* [18] observed centromere movements *in vivo* by using the DNA binding domain of human centromere protein CENP-B fused to GFP. Meanwhile Li *et al.* [19] discovered chromosome arm movement. It was proposed that the chromosome movements might be connected with DNA replication which occurs at specific sites within the nucleus [20]. Chromatin compartmentalization and interaction in the nucleus are correlated with global regulation of gene expression. The identification of chromatin associated proteins should enhance our understanding of gene regulation, chromatin higher level structure, and conformational dynamics.

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 [21]. One hundred fifty eight proteins were identified in the sucrose gradient fraction and 108 proteins were identified in the percoll gradient fraction. 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.* [22] identified 1099 proteins that were copurified with spermatogenic chromatin in *Caenorhabditis elegans.* Although the copurified protein number is large, the authors are able to reduce the protein number to 132 for further studies on spermatogenesis by subtracting the proteins shared with oogenic chromatin. 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.* [23] studied human B lymphocyte chromatin enriched fractions and identified 64 proteins including 18 putative transcription factors. 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. Five hundred nine proteins, corresponding to 269 unique proteins, have been identified using 2-DE gel method. The proteins identified by shotgun method included histones, histone variants, many known chromatin binding proteins, and function unknown proteins. Our studies have provided new insight into chromatin composition in plants.

2 Materials and methods

2.1 Suspension culture and protoplast isolation

A rice (*Oryza sativa*) suspension culture (line Oc) [24, 25] 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, 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 Yamada *et al.* [26] 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), 0.4 M mannitol, 80 mM $CaCl_2$, 0.125 mM $MgCl_2$, 0.5 mM MES, and B5 organic medium plus 2.0 mg/L 2,4-D (pH 5.6). After 8 h 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 $120 \times g$ for 5 min, and washed three times with protoplast suspension medium (0.4 M mannitol, 80 mM $CaCl_2$, 0.125 mM $MgCl_2$, 0.5 mM MES, and B5 organic medium at pH 5.6). A yield of 2.4×10^6 protoplasts *per* gram suspension cells was obtained.

2.2 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 $MgCl_2$, 1.0 mM $CaCl_2$, 0.1 mM spermidine, 0.5% Ficoll, 0.5% Triton X-100, 1.0 mM DTT, and 1.0 mM PMSF added freshly) with 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 $500 \times g$ for 10 min at 4°C. The pellet of nuclei was resuspended in NIB, layered onto 10 mL 2 M sucrose, and pelleted by a centrifugation at $6000 \times g$ 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, 1.0 mM NaCI, 1.0 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 $750 \times g$ for 10 min at 4°C, resuspended in the CIB buffer, layered onto 10 mL of 2 M sucrose, and pelleted again by spinning at $7600 \times g$ 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.

2.3 Electron microscopy

The freshly isolated chromatin was used for transmission electron microscopy examination using the preparation method described by Vengerov and Popenko [27] with minor modifications. In brief, chromatin was diluted to a final concentration ~5 μg/μL in TE buffer containing 0.25 M ammonium acetate. After 10 min incubation, 2.0 μL of Cytochrome *c* at 0.2 μg/μL was added, mixed gently, and kept at room temperature for 90 s. 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 s, rinsed in 90% ethanol for 2 s, 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).

2.4 Protein extraction

Proteins were extracted as reported [28, 29] 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 β-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 s. 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 $3000 \times g$ for 10 min at 4°C. The proteins were precipitated with five volumes of precipitation buffer (PB: 0.1 M ammonium acetate and 1% β-mercaptoethanol in methanol) at −70°C overnight. The protein pellet was recovered by a centrifugation at $17\,210 \times g$ for 10 min at 4°C and washed three times with prechilled PB and another three washes with prechilled 70% ethanol. The protein pellet was lyophilized to powder in a speed vacuum (LABCONCO, model LYPH-LOCK 6) and stored at −70°C.

2.5 Western blots

Twenty micrograms *per* lane of total protein extracts, nuclear proteins, and chromatin associated proteins were separated by a 12% SDS-PAGE and processed for Western blot using standard procedures. The antibodies against H1, H2B, and acetyl-K14H3 were purchased from UPSTATE. Antibodies against actin and α-tubulin were purchased from Santa Cruz Biotechnology and SIGMA, respectively. Antibodies against COP9 signalosome subunits 3 and 6, and PhyA were kindly provided by Deng [30, 31].

2.6 2-D PAGE

Proteins were dissolved thoroughly in rehydration buffer (7 M urea, 2 M thiourea, 4% CHAPSO, 1% DTT, and 0.2% ampholines) and centrifuged at $12\,000 \times g$ for 10 min to remove undissolved contents. The supernatant was quantified using a BioRad RcDc protein assay kit according to manufacturer's instructions. The quantified proteins were then used for 2-D PAGE. IEF was carried out using a BioRad PROTEAN IEF cell on 17 cm 3–11 pH nonlinear IPG strips (GE). Four hundred micrograms of protein in 400 μL of rehydration buffer was loaded into the IEF tray and active rehydration was carried out at 23°C for 12 h, followed by 250 V for 2 h, a linear increase of voltage to 10 000 V for 4 h, and the IEF was performed at 23°C for a total of 90 000 V·h. After the completion of IEF, the strips were equilibrated in a buffer containing 6 M urea, 0.375 M Tris-HCl (pH 6.8), 20% glycerol, 2% SDS, and 2% DTT for 15 min and followed by equilibration for another 15 min in a buffer containing 6 M urea, 0.375 M 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 mm × 18 mm × 1.5 mm) containing 12% separating gel and 4% stacking gel. Electrophoresis was carried out in a BioRad PROTEAN PLUS horizontal Dodeca cell at 15 mA/gel.

The gels were stained with SYPRO Ruby fluorescence stain (BioRad) according to the protocols provided by the manufacturer and scanned with a VersaDoc4000 image system (BioRad). For each treatment, 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 (BioRad), 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.

2.7 In gel digestion and MS

After PDQuest analysis, the spots of interest were robotically excised from 2-DE gels by a Proteome Works Spot Cutter (BioRad). 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 and iodoacetamide, respectively. The resulting peptide mix was desalted with C18 ZipTips (Millipore) and spotted on a MALDI plate in a solution containing 70% ACN, 0.1% TFA, and 5 mg/mL matrix (CHCA).

Mass spectra were collected on an ABI 4700 Proteomics Analyzer (Applied Biosystems) MALDI TOF-TOF mass spectrometer, and protein identification (ID) 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 subdatabase of "nr" (nonredundant) database of National Center for Biotechnology Information (NCBI).

During the initial MS scan, data were analyzed as 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 CI%>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 CI%>95% were considered as a positive ID.

2.8 Identification of phosphoproteins using Pro-Q Diamond phosphoprotein in 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 (BioRad). A spot constantly stained with Pro-Q Diamond dye in all three biological replicas was considered as a putative phosphoprotein spot. The ratios of SYPRO Ruby *versus* 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.

2.9 Protein digestion and LC/LC-MS/MS analysis

The protein pellet was dissolved in 6 M urea with 100 mM Tris-Cl (pH 7.8) and centrifuged at 16 000 × *g* for 10 min. The supernatant was collected and then quantified using a Rc Dc kit (BioRad). Proteins (100 μg in 50 μL) were reduced by mixing with 20 μL of 200 mM DTT in 100 mM Tris-Cl (pH 7.8) for 1 h at room temperature, alkylated with 20 μL of 200 mM iodoacetamide in 100 mM Tris-Cl (pH 7.8) in darkness for 1 h, and diluted to a final urea concentration of 0.6 M, 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 10 μL of 10 mM lysine after 15 h. 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 macrotrap (Michrom Bioresources, Auburn, CA) following the protocol provided by the manufacturer and eluted with 0.1% TFA in 95% ACN. The eluted peptides were vacuum dried to pellet and redissolved in 20 μL of 0.1% formic acid with 5% ACN.

The peptide mixtures were subjected to 2-D LC comprising a separation on a strong cation exchange column (SCX BioBasic 0.32 mm × 100 mm) followed by a RP column (BioBasic C18, 0.18 mm × 100 mm, Thermo Hypersil-Keystone, Bellefonte, PA) coupled directly in-line with ESI IT mass spectrometer (ProteomeX workstation, ThermoFinnigan). A flow rate of 3 μL/min was used for both SCX and RP 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% ACN and 0.1% formic acid were applied. The eluted peptides were loaded directly on the RP column, equilibrated with 0.1% formic acid and 5.0% ACN. The peptides were eluted from RP column by an ACN gradient (in 0.1% formic acid) as follows: 5–30% for 30 min, 30–65% for 9 min, 95% for 5 min, 5% for 15 min, a total of 59 min of elution.

The LCQ Deca XP IT mass spectrometer was configured to optimize the duty cycle length with the quality of data acquire, by alternating between a single full MS scan followed by three 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 m/z range of 200–2000. The mass spectra and tandem mass spectra produced were used to search the rice (*O. sativa*) nonredundant protein database (NCBInrPDB) downloaded from the NCBI by using TurboSEQUEST, BIOWORKS BROWSER 3.1 SR1 (ThermoFinnigan). 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 two. 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 or more peptides from a protein to meet the following criteria: X-correlation >1.9 (+1 charge), >2.5 (+2 charge), >3.75 (+3 charge); delta correlation value ≥0.1; and probability <0.01. The false positive rates were estimated through searching reversed sequence database.

2.10 Gene Ontology (GO) annotation

Proteins were classified according to the GO rules [32]. There are three independent categories of ontologies that are used to describe the molecular function of a gene product. Cellular component, which describes where the gene product

can be found, was used in this study. GO annotations were obtained from GORetriever, a program available at AgBase [33] (http://www.agbase.msstate.edu/). Proteins without annotations at AgBase were searched in other databases, including NCBI, UniProt, and Gramene. GOSlim for GO pie charts were then generated using the program GOSlim-Viewer provided by AgBase (http://www.agbase.msstate.edu/).

3 Results

3.1 Chromatin purification from rice (*O. 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 [21, 22]. To obtain high quality chromatin, rice suspension culture cells [24, 25] were used for chromatin isolation and purification in our studies. The protoplasts were released from suspension cells by enzymatic digestion of the cell wall as reported [26] for 8 h. One modification was that B5 organic medium was included in the cell wall digestion solution to maintain regular cellular activities during the removal of cell wall. As shown in Fig. 1A, pure and intact protoplasts were obtained after filtering through a 25 μm metal mesh and followed by several rounds of washing. Nuclei were released from the protoplasts in a nucleus isolation buffer and purified using a 2 M sucrose density centrifugation. Highly enriched nuclei were obtained on a large scale as revealed by fluorescence microscopy after DAPI staining (Fig. 1B). Chromatin was released from the enriched nuclei in a chromatin isolation buffer with homogenization, purified using a 2 M sucrose density centrifugation, and followed by at least three washes with the chromatin isolation buffer. The chromatin preparation was homogeneous in appearance and no visible contamination of other organelles as revealed by DAPI stain (Fig. 1C) and examination under phase contrast microscopy (data not shown). 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 Fig. 1D, large clusters of thick chromatin fiber were observed. It was evident that the above 30 nm structure (Fig. 1D) of the chromatin was, at least partially, maintained in our chromatin preparation.

3.2 Western blot, 2-DE gel, and mass spectrometric analysis of proteins copurified with chromatin

The establishment of the chromatin isolation protocol made it possible for us to identify proteins copurified with chromatin. We extracted the chromatin associated proteins using the phenol extraction method [28, 29]. The phenol extracted chromatin proteins were free of DNA contamination as revealed by an agarose gel analysis (data not shown). To test the quality and quantity of the isolated chromatin proteins, we used antibodies against histone H2B, H3K14acetyl, and H1 for Western blots, respectively. We found that H2B, H3, and H1 were highly enriched in the chromatin fraction compared to the total protein extract and total nuclear protein extract as shown in Fig. 2. We noticed that two thick

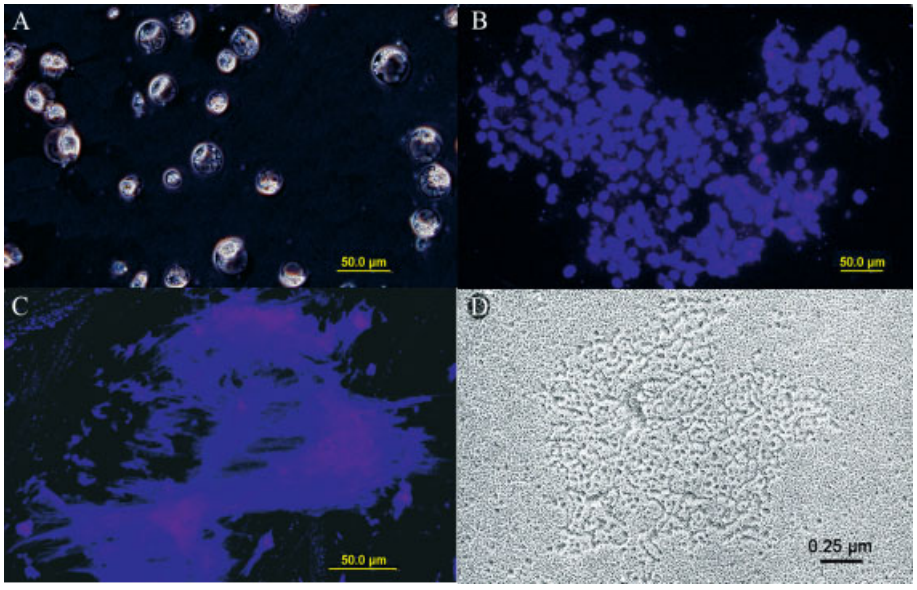

Figure 1. Microscopy images of isolated rice (*O. sativa)* suspension cell protoplasts, nuclei and chromatin mass. (A) Image of rice protoplasts. Isolated rice protoplasts were diluted in protoplast suspension medium, plated on a microscope slide, and visualized using a phase-contrast microscopy. (B) Image of purified rice nuclei after DAPI stain. Purified rice nuclei resuspended 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.

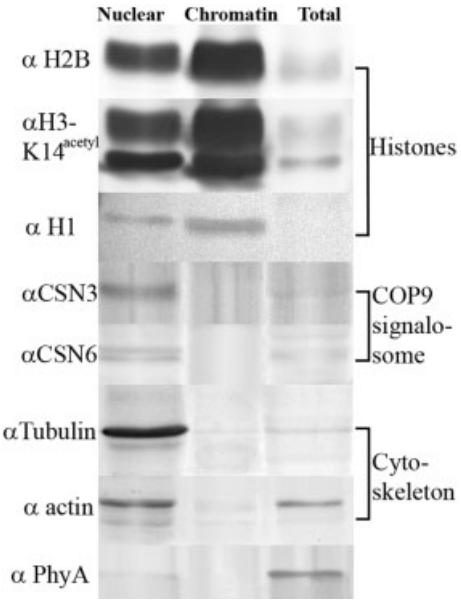

Figure 2. Immunological characterization of proteins in the chromatin fraction, nuclear fraction, and total protein fraction. 20 μg proteins were loaded on each lane. Western blots were carried out with standard procedures. The second antibodies were conjugated with alkaline phosphotase. Protein sources are indicated on the top. Antibodies are indicated on the left and the characteristics of the corresponding protein are indicated on the right.

bands were detected when antibodies against H3K14acetyl were used for Western blots. 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. COP9 signalosome is a nuclear enriched and high abundance protein complex involved in the regulation of protein degradation and plant development [30, 31, 34]. We examined COP9 complex in total protein, nuclear protein, and chromatin protein fractions, respectively, with antibodies against subunits 3 and 6. Both subunits 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 (Fig. 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 [30]. We also examined α-tubulin and actin in these three protein fractions. Western results indicated both actin and α-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 examined PhyA level in these three protein fractions. Our results indicated that PhyA was mainly detected in the total protein fraction (Fig. 2). Although a trace amount was detected in nuclear fraction, no PhyA protein was detected in the chromatin fraction. Above Western blot results indicated that histone proteins were significantly enriched in chromatin fraction. Meanwhile, contaminations by other nuclear proteins such as COP9 complex subunit and cytosolic protein such as PhyA were not detectable in Western blots.

To further analyze proteins copurified with chromatin, the proteins were resolved on pH 3.0–11.0 2-DE gels. A representative gel image was shown Fig. 3 and the three biological replicas were shown in Fig. 1 of Supporting Information. The separated proteins had a wide range of p*I*s and molecular weights. PDQuest software 7.4.0 analyses, followed by manual verification, found that about 972 distinct protein spots could be consistently resolved on 2-DE gels in all three biological replicas after SYPRO Ruby stain (Fig. 1 of Supporting Information).

Among the 972 resolved protein spots, 607 prominent protein spots were excised using a Proteome Works Spot Cutter (BioRad), 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 (CI%>95%) as marked on the 2-DE gel picture (Fig. 3) and listed in Table 1 of Supporting Information. These 509 identified proteins correspond to 269 unique proteins. We found that many distinct proteins spots had the same protein identities perhaps due to PTMs, alternative splicing, or other reasons. For example, H2A was identified in six distinct protein spots, H2B in four distinct protein spots, H3 in four distinct protein spots, and H4 in six 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 [35, 36]. 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

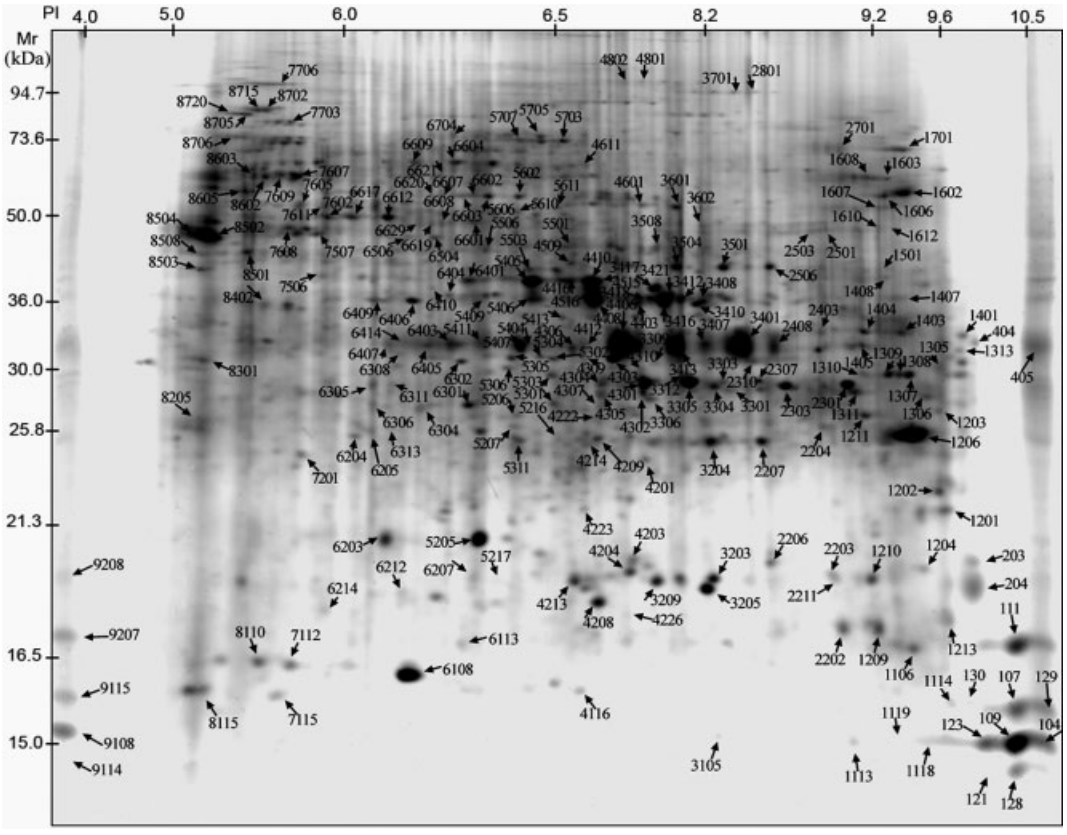

Figure 3. 2-DE gel image of rice chromatin proteome revealed by SYPRO Ruby fluorescence stain. Proteins were extracted from purified chromatin of rice suspension cells, separated on 2-DE gels, and stained with SYPRO Ruby. Proteins identified with high confidence (CI%>95%) are marked with arrows. Molecular 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%.

[21, 22]. In addition, many proteins with known functions unrelated to chromatin were detected in the chromatin fraction and most of these proteins were also copurified with human chromosome and *C. elegans* chromatin, respectively [21, 22]. These proteins included 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.* A plastid protein, cationic peroxidase, was detected in multiple protein spots. RUBISCO and other major plastid proteins, however, were not detected in our chromatin extracts, suggesting that the high abundance of the peroxidase in the chromatin preparation was not due to a large scale plastid contamination. Many hypothetical and function unknown proteins were also identified as shown in Table 1 of Supporting Information.

3.3 GO analysis of the proteins identified using 2-DE gel approach

To help understand the distribution and function of the chromatin associated proteins, we obtained GO annotations of the identified proteins from the AgBase [33] (http://www.agbase.msstate.edu/) and other databases such as NCBI, UniProt, Gramene and conducted GO analyses. As shown in Fig. 4, the cellular distributions of identified unique proteins were nuclear proteins (39.2%), function unknown proteins (20.4%), cytoplasm proteins (15.8%), membrane protein (8.8%), cytoskeleton (8.4%), mitochondrion (8.4%), plastid (7.0%), ribosome (6.0%), intracellular (3.7%), cytosol (0.9%), *etc.* In comparison, the human metaphase chromosome proteome 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 [21].

In the current GO classification system, one protein may be grouped into more than one GO category. Therefore, the accumulative percentage is over 100%. On the other hand, the information for proteins with multiple functions is not completely compiled. For example, 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.

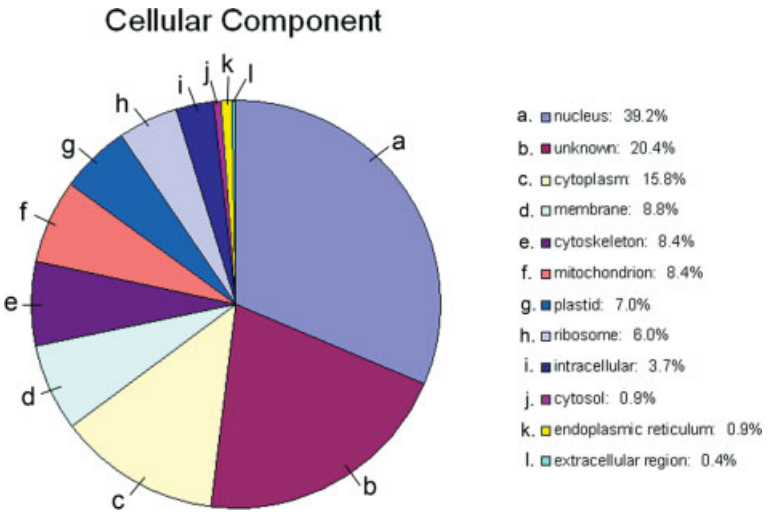

Figure 4. Distribution of the proteins copurified with chromatin in different subcellular compartments. Percentage distribution of the unique proteins was used to make the pie chart based on the GO of cellular localizations. The pie chart was generated using the analysis results of the "GOSlimViewer" tool at AgBase.

3.4 Mapping putative phosphoproteome copurified with chromatin using Pro-Q Diamond phosphoprotein stain

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 p*I*s 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 [37–43]. Analysis using PDQuest 7.4.0. software (BioRad) identified 390 putative protein spots. Further manual examination confirmed that 205 prominent protein spots were constantly stained by the Pro-Q Diamond dye (Fig. 5, and Fig. 2 of Supporting Information). Compared with SYPRO Ruby stain (Fig. 5C), whose stain intensity is proportional to the protein quantity, the Pro-Q Diamond phosphoprotein stain displayed high specificity to some specific proteins (Fig. 5B). For example, spots 1309, 1307, and 2306 were heavily stained by SYPRO Ruby but not or 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 ratio of Pro-Q Diamond *versus* SYPRO Ruby staining were calculated and listed in Table 2 of Supporting Information. 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. software 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/TOF spectrometer. One hundred fifty-four protein annotations were obtained as shown in Table 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 phosphophosphorylated. On the other hand, some proteins were presented in multiple protein spots that shared the similar molecular weight but different p*I*s. 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 glyceraldehyde-3-phosphodehydrogenase; spots 3401, 3413, and 4406 for hypothetic protein (OsJ_024928); 8602 and 8603 for Os01g0685800, *etc.* The cellular distributions of the putative phosphoproteins were shown in Fig. 3 of Supporting Information. The distribution was nuclear protein 36.4%, unknown protein 23.4%, cytoplasm 11.2%, ribosome 6.5%, *etc.*

3.5 Identification of chromatin copurified proteins using shotgun approach

Recent development in 2-D-LC-MS/MS technology has made shotgun proteomics a powerful tool in protein identification using a protein mixture. To examine proteins copurified with chromatin, we analyzed the protein mixture using shotgun approach. Since our previous studies using total protein extracts had shown that shotgun proteomics was more efficient in protein identification than 2-DE gel based approach in *Arabidopsis* (Chitteti and Peng, unpublished results), we had high expectations to the shotgun methods. Unfortunately, the proteins identified were predominantly histones and histone variants with only a few nonhistone proteins identified in each run. However, different new proteins could be identified in each new run. After multiple examinations, we obtained 70 annotations with two or more peptides identified as listed in Table 2 and another 58 annotations with a

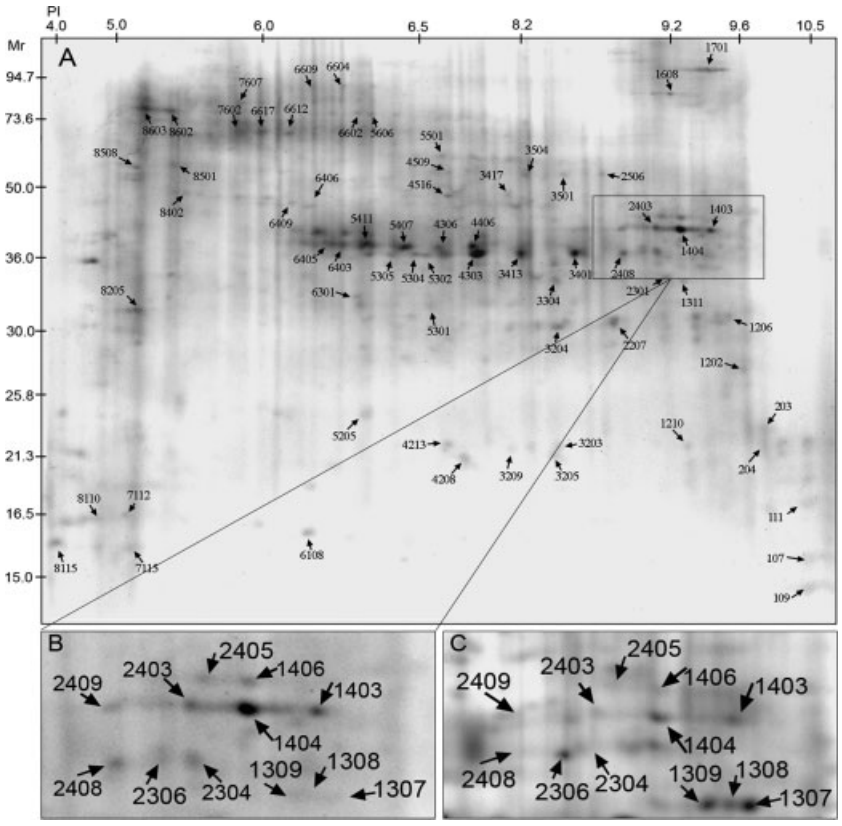

Figure 5. 2-DE gel images of the putative phosphoproteins associated with rice chromatin. Proteins were extracted from purified chromatin of rice suspension cells, separated on 2-DE gel, and stained with Pro-Q Diamond phosphoprotein stain and SYPRO Ruby, respectively. (A) 2-DE gel image of chromatin associated phosphoproteins revealed by Pro-Q Diamond stain. Proteins identified with high confidence (CI%>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).

single peptide identified as shown in Table 3 of Supporting Information. Although only one peptide was identified in these proteins, the probability of random match to the corresponding protein was low (Table 3 of Supporting Information). 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. 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, 5 H2B variants, 2 H3 variants, 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 protesome regulatory subunits, heat shock proteins, RNA binding proteins, *etc.* (Table 2, and Table 3 of Supporting Information). 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 catonic peroxidase, were not detected using shotgun approach. To estimate the error rate, the reversed sequence database was searched. Nine proteins were found based on a single peptide match and no proteins were found based on two peptide matches (Table 4 of Supporting Information). For proteins identified with two peptides match, the error rate should be very low. Since the total peptides identified in Table 2 and Table 3 of Supporting Information were 445, the error rate for peptide miss identification was 9/454.

4 Discussions

4.1 Chromatin purification

Chromatin is a supercomplex of a long string of DNA and numerous proteins "floating" in a solution containing over thousands of other proteins. Like all organelle subproteome 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. By investigating the mouse homologous of *C. elegans* proteins copurified with chromatin, Chu *et al.* [22] found that 37% of the mouse knockouts cause sterility, indicating that identifying the proteins copurified with chromatin can be very useful in various biological studies. We have developed a protocol for purification of chromatin from rice suspension

Table 1. Putative phosphoproteins copurified with chromatin in rice suspension cells (*O. sativa*)

Location	Spot number	Accession number	Protein name	CI%	MW (kDa)	p*I*	Identified peptides
Nucleus	107	AAN06860	Putative histone H2A	100.0	14.55	10.3	4
	107	Q8S857	Putative histone H2A	100.0	14.6	10.4	4
	107	BAC75621	Putative histone H2A	99.6	13.91	10.2	1
	107	Q94E96	Putative histone H2A	99.6	16.39	10.7	1
	109	AAX92952	Histone H3 – maize	99.5	16.04	10.0	4
	109	AAC78105	Histone H3.3	99.8	15.4	11.2	5
	109	AAN06860	Putative histone H2A	99.5	14.55	10.3	4
	109	Q8S857	Putative histone H2A	98.7	14.6	10.4	3
	111	Q94E96	Putative histone H2A	100.0	16.39	10.7	4
	111	Q943L2	Putative histone H2B	100.0	15.36	10.0	9
	111	Q94JE1	Putative histone H2B	100.0	16.75	10.0	10
	111	Q9LGH4	Putative histone H2B	100.0	16.46	10.0	10
	111	Q9LGH6	Putative histone H2B	100.0	16.46	10.0	12
	111	Q9LGH8	Putative histone H2B	100.0	16.48	10.0	12
	204	NP_001042044	Os01g0152300	100.0	19.29	9.9	12
	204	EAY72569	Putative histone H2B	100.0	16.52	10.0	11
	204	Q9LGH6	Putative histone H2B	100.0	16.46	10.0	10
	1210	Q40674*	Cyclophilin 2 (EC 5.2.1.8) (peptidyl-prolyl *cis-trans* isomerase) (PPIase) (rotamase)	99.5	18.32	8.6	6
	1210	Q40673*	Cyclophilin 2 (EC 5.2.1.8) (peptidyl-prolyl *cis-trans* isomerase) (PPIase) (rotamase)	99.3	18.35	8.6	6
	1403	EAZ38957*	Glycine-rich RNA-binding protein-like	100.0	33.36	9.3	12
	1403	NP_001059075*	Glycine-rich RNA-binding protein-like	100.0	33.23	9.3	13
	1403	ABA96443	Retrotransposon protein, putative, Ty3-gypsy subclass	98.9	162.14	9.1	23
	1404	EAZ38957*	Glycine-rich RNA-binding protein-like	100.0	33.36	9.3	13
	1404	NP_001059075*	Glycine-rich RNA-binding protein-like	100.0	33.23	9.3	14
	1701	EAY83860*	Os12g0611200	100.0	65.73	9.6	27
	1701	EAZ21188*	Os12g0611200	100.0	62.73	9.4	26
	1701	NP_001067252*	Os12g0611200	100.0	65.55	9.5	26
	2403	NP_001067743	Ribosomal protein-like	95.8	48.13	8.6	11
	2403	DAA00397	TPA_exp: unknown	100.0	39.29	6.4	12
	3401	Q94CF9*	RSSG8	99.7	129.58	8.6	20
	5301	AAP52582*	Transposon protein, putative, CACTA, En/Spm subclass	99.8	88.56	8.4	19
	5501	EAZ32941	Hypothetical protein OsJ_016424	98.0	39.25	6.4	6
	5606	EAY95913*	Hypothetical protein Osl_017146	100.0	95.91	6.7	16
	5606	NP_001054148	OSJNBa0015K02.19 protein	100.0	40.23	5.9	13
	6108	O22384*	Glycine-rich protein	100.0	15.86	7.8	8
	6108	O22385*	Glycine-rich protein	100.0	16.02	7.8	9
	6108	O22390*	Glycine-rich protein	100.0	16.53	9.0	7
	6108	O24187*	OsGRP1	100.0	15.8	5.5	9
	6405	Q8W403	Sec13p	99.9	33.35	5.6	5
	6602	BAD81520	Putative Y1 protein	98.4	50.05	5.9	5
	6617	EAY80530*	Cysteine endopeptidase	100.0	40.95	5.9	14
	6617	ABA92414*	Thiol protease SEN102 precursor, putative, expressed	100.0	40.91	5.7	15
	7115	O22384*	Glycine-rich protein	100.0	15.86	7.8	5
	8110	O22385*	Glycine-rich protein	99.6	16.02	7.8	6
	8110	O24184*	Glycine-rich RNA-binding protein	99.5	16.32	7.8	6
	8110	O24187*	OsGRP1	97.0	15.8	5.5	5
	8205	BAD67781	Putative WRKY DNA-binding protein	98.0	28.48	10.3	10
	8501	EAZ29051*	Hypothetical protein OsJ_012534	96.9	92.32	9.7	17
Ribosome	203	EAY96929	Hypothetical protein Osl_018162	100.0	13.23	9.8	8
	203	EAZ37303	Hypothetical protein OsJ_020786	100.0	35.73	10.5	13
	203	BAF16813	Os05g0207300	100.0	20.82	9.9	10

Table 1. Continued

Location	Spot number	Accession number	Protein name	CI%	MW (kDa)	p*I*	Identified peptides
	1202	P49210	60S ribosomal protein L9	100.0	21.35	9.6	11
	1202	Q8S9N3	ARE1 (Fragment)	97.0	8.1	11.7	3
	1202	AAP92747	Ribosomal L9-like protein	100.0	21.34	9.6	12
	5411	Q7XNU2	OSJNBa0093F12.16 protein	97.5	17.74	9.2	4
	6602	Q7XNU2	OSJNBa0093F12.16 protein	100.0	17.74	9.2	6
Plastid	1608	NP_001051013*	Beta-glucosidase	100.0	56.84	9.1	18
	1608	AAA84906*	Beta-glucosidase	100.0	56.86	9.1	17
	2506	EAY86569*	Hypothetical protein OsI_007802	100.0	39.01	6.5	10
	3417	O22510*	Cationic peroxidase	100.0	38.3	8.9	14
	4306	O22510*	Cationic peroxidase	100.0	38.3	8.9	11
	4516	O22510*	Cationic peroxidase	100.0	38.3	8.9	12
	5305	O22510*	Cationic peroxidase	100.0	38.3	8.9	10
	5407	O22510*	Cationic peroxidase	100.0	38.3	8.9	13
	6108	Q8VXC4*	Glycine rich RNA binding protein	100.0	19.46	6.6	10
	6403	O22510*	Cationic peroxidase	100.0	38.3	8.9	9
	7115	Q8VXC4*	Glycine rich RNA binding protein	96.8	19.46	6.6	7
Mitochondrion	2301	Q8VXC7	Voltage-dependent anion channel	100.0	29.58	8.6	12
	3304	NP_001056162	Voltage-dependent anion-selective channel (VDAC) protein	100.0	29.58	8.6	10
	5304	NP_001044063	Os01g0715500	100.0	29.74	5.5	10
	5304	EAZ13326	Putative 36kDa porin II	100.0	41.49	5.7	11
	5304	EAY75604	Putative 36kDa porin II	99.9	44.68	5.2	11
	6604	P15998	ATP synthase alpha chain, mitochondrial (EC 3.6.3.14)	100.0	55.25	5.9	16
	6609	P15998	ATP synthase alpha chain, mitochondrial (EC 3.6.3.14)	100.0	55.25	5.9	18
	7607	Q01859	ATP synthase beta chain, mitochondrial precursor (EC 3.6.3.14)	100.0	59.02	6.3	18
	8301	Q01859	ATP synthase beta chain, mitochondrial precursor (EC 3.6.3.14)	100.0	59.02	6.3	5
	8602	Q01859	ATP synthase beta chain, mitochondrial precursor (EC 3.6.3.14)	100.0	59.02	6.3	21
Membrane	4213	EAZ25964	Hypothetical protein OsJ_009447	99.7	66.37	7.8	16
	4303	EAY94429	B0812A04.3 protein	100.0	30.62	6.6	16
	4303	CAE76006	B1358B12.15	100.0	30.78	7.0	15
	4303	NP_001053006	Os04g0462900	100.0	31.45	6.0	15
	5302	EAY94429	B0812A04.3 protein	100.0	30.62	6.6	11
	5302	CAE76006	B1358B12.15	100.0	30.78	7.0	10
	5302	NP_001053006	Os04g0462900	100.0	31.45	6.0	10
	5411	Q7X863	OSJNBa0016N04.2 protein (OSJNBa0049H08.21 protein)	99.4	16.38	11.2	9
	6403	BAB89823	Embryonic abundant protein-like	100.0	29.08	5.7	10
	7607	BAF05814	Os01g0685800	100.0	59.72	5.9	17
	8301	BAF05814	Os01g0685800	100.0	59.72	5.9	7
	8602	BAF05814	Os01g0685800	100.0	59.72	5.9	21
Cytoplasm	1311	O64937	Elongation factor 1-alpha (EF-1-alpha)	97.9	49.25	9.1	7
	1210	O04985*	Nonsymbiotic hemoglobin 2 (rHb2) (ORYsa GLB1b)	100.0	18.6	9.0	7
	2506	A1YR13*	Glyceralde-3-phosphate dehydrogenase	100.0	36.54	7.7	12
	3203	EAZ06044	Hypothetical protein OsI_027276	100.0	38.48	5.9	11
	3203	EAZ41920	Hypothetical protein OsJ_025403	100.0	31.4	5.1	8
	3203	NP_001061280	Os08g0224900	100.0	24.26	5.0	6
	3205	EAZ06044	Hypothetical protein OsI_027276	100.0	38.48	5.9	12
	3205	EAZ41920	Hypothetical protein OsJ_025403	100.0	31.4	5.1	9
	3205	NP_001061281	Os08g0225000	100.0	16.36	9.4	6

Table 1. Continued

Location	Spot number	Accession number	Protein name	CI%	MW (kDa)	p*I*	Identified peptides
	3209	O04986*	Nonsymbiotic hemoglobin 1 (rHb1) (ORYsa GLB1a)	100.0	18.43	6.9	7
	3501	A1YR13*	Glyceralde-3-phosphate dehydrogenase	100.0	36.54	7.7	17
	3501	EAZ23708*	Putative glyceraldehyde-3-phosphate dehydrogenase	100.0	39.01	6.5	15
	3504	AAP54418	*N*-acetyl-gamma-glutamyl-phosphate reductase, chloroplast precursor, putative, expressed	100.0	45.45	8.5	5
	4208	EAY96193	Hypothetical protein Osl_017426	100.0	101	5.7	8
	5205	EAZ06044	Hypothetical protein Osl_027276	100.0	38.48	5.9	13
	5205	EAZ41920	Hypothetical protein OsJ_025403	100.0	31.4	5.1	9
	5205	NP_001061280	Os08g0224900	100.0	24.26	5.0	5
	8115	BAD32133*	Putative receptor-like protein kinase 4	96.6	72.75	6.3	16
Intracellular	6617	AAF15312	Chloroplast translational elongation factor Tu	100.0	50.32	6.1	12
	7602	AAF15312	Chloroplast translational elongation factor Tu	100.0	50.32	6.1	12
	7602	EAZ23662	Translational elongation factor Tu	100.0	50.38	6.2	12
Cytoskeleton	6617	EAY86520	Hypothetical protein Osl_007753	100.0	63.64	5.8	14
	7602	EAY86520	Hypothetical protein Osl_007753	100.0	63.64	5.8	15
Cytosol	4509	Q8W424	26S proteasome regulatory particle non-ATPase subunit8	100.0	34.87	6.3	17
Extracellular region	2207	P93442	Expansin-A4 precursor (OsEXPA4) (alpha-expansin-4) (OsEXP4) (OsaEXPa1.22)	96.6	25.87	8.1	5
Unknown	1206	EAY92380	Hypothetical protein Osl_013613	100.0	28.2	9.5	7
	1206	EAZ29098	Hypothetical protein OsJ_012581	100.0	36.01	9.3	8
	1210	BAF05781	Os01g0679600	100.0	21.3	6.8	4
	1608	ABF98424	Glycosyl hydrolase family 1 protein, expressed	100.0	46.02	9.0	15
	3204	JC7138	Alpha-amylase (EC 3.2.1.1) isozyme III - rice	100.0	48.71	5.3	6
	3304	BAB67891	Putative thaumatin-like cytokinin-binding protein	100.0	26.17	7.9	11
	3401	EAZ41445	Hypothetical protein OsJ_024928	100.0	29.81	8.2	13
	3413	EAZ05518	Hypothetical protein Osl_026750	99.9	29.82	8.2	8
	3413	EAZ41442	Hypothetical protein OsJ_024925	95.6	30.03	8.2	5
	3413	EAZ41445	Hypothetical protein OsJ_024928	99.9	29.81	8.2	8
	3417	CAJ86336	H0814G11.3	100.0	33.46	5.2	17
	3417	O24523	Peroxidase (Fragment)	100.0	13.12	10.6	6
	3501	AAN59792	Cytosolic glyceraldehyde-3-phosphate dehydrogenase GAPDH	100.0	23.37	7.9	9
	4208	AAG03091	Unknown protein	100.0	14.67	6.8	6
	4306	CAJ86336	H0814G11.3	100.0	33.46	5.2	12
	4306	O24523	Peroxidase (Fragment)	100.0	13.12	10.6	2
	4306	CAH69301	TPA: class III peroxidase 59 precursor	100.0	36.97	5.6	10
	4406	EAZ41442	Hypothetical protein OsJ_024925	95.6	30.03	8.2	5
	4406	EAZ41445	Hypothetical protein OsJ_024928	99.9	29.81	8.2	12
	4516	CAJ86336	H0814G11.3	100.0	33.46	5.2	13
	5305	CAJ86336	H0814G11.3	100.0	33.46	5.2	11
	5407	CAJ86336	H0814G11.3	100.0	33.46	5.2	13
	5407	O24523	Peroxidase (Fragment)	100.0	13.12	10.6	2
	6301	BAD61316	Putative glutathione transferase F4	100.0	25.29	5.7	12
	6403	CAJ86336	H0814G11.3	100.0	33.46	5.2	11
	6403	O24523	Peroxidase (Fragment)	99.4	13.12	10.6	4
	6405	CAJ86336	H0814G11.3	100.0	33.46	5.2	3
	6406	Q8VYH7	Isoflavone reductase-like protein	100.0	33.47	5.7	16
	6406	Q9FTN5	Putative isoflavone reductase homolog IRL	100.0	33.48	5.7	18
	6406	Q9FTN6	Putative isoflavone reductase homolog IRL	100.0	34.28	6.3	5
	6409	Q7XKT9	OSJNBa0022H21.18 protein	100.0	32.08	5.5	13
	6609	BAB92682	Putative selenium binding protein	100.0	53.1	5.7	10

Table 1. Continued

Location	Spot number	Accession number	Protein name	CI%	MW (kDa)	p*I*	Identified peptides
	6609	Q9AVA6	Putative selenium binding protein	100.0	50.93	5.7	10
	6612	EAZ38461	Hypothetical protein OsJ_021944	98.1	45.73	5.5	13
	6612	BAC15855	Putative dimethylaniline monooxygenase	95.6	53.01	5.5	13
	7112	EAZ38461	Hypothetical protein OsJ_021944	97.7	45.73	5.5	13
	7112	BAC15855	Putative dimethylaniline monooxygenase	99.6	53.01	5.5	14
	8402	Q7G764	Probable NAD(P)H-dependent oxidoreductase 1	97.3	35.69	5.3	9
	8402	Q7G765	Probable NAD(P)H-dependent oxidoreductase 2	100.0	35.76	5.4	11
	8508	JC7138	Alpha-amylase (EC 3.2.1.1) isozyme III - rice	98.8	48.71	5.3	7

a) Spot number: the spot number was given by computer based on spot excision order.
b) Accession number: protein accession number.
c) MW/p*I*: predicted molecular mass and p*I*.
d) Identified peptides: number of peptides that matched with the identified protein in mass analyses.
e) CI%: cross confidence interval%. Over 95% represents high confidence identification.
f) *: Proteins without GO annotation. The cellular localization was predicted by LOCtree or PSORT, two subcellular prediction programs.

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. (i) Electron microscopy and optical microscopy reveal that the chromatin preparation maintains, at least partially, high level structure and is free of visible organelle contamination. (ii) Western blots have shown that abundant protein complex COP9 signalosome, which is nuclear enriched, and cytoplasmic protein PhyA did not copurify with chromatin. (iii) Among the 509 identified proteins using 2-DE gel approach, nuclear proteins represent 39.2%. Proteins from cytoskeleton, mitochondrion, plastid, and ribosome are 8.4, 8.4, 7.0, and 6.0%, respectively. Because cytoskeleton and ribosome proteins can be both nuclear and cytoplasmic, the ratio of nuclear protein should be higher than 39.2%. The metaphase chromosome of human cell line, which probably is the best chromatin/chromosome preparation reported thus far, contains 29.8% nuclear proteins. The primary contaminants of human chromosome are mitochondrion (38.6%), cytoskeleton (12.7%), and ribosome (4.4%), respectively [21]. 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. Six hundred seven prominent protein spots were excised for mass analyses. Five hundred nine proteins were identified with high confidence (CI%$>$95%). These proteins correspond to 269 unique proteins. As shown in Fig. 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.

4.2 Proteins copurified 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 five possible H2B variants in addition to the four common core histone proteins. In mammals, only three bona fide H2B variants and six H2A variants have been identified thus far [44]. While five histone H3 variants have been reported in mammals, we only identified three 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 effecter 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 [44, 45].

Table 2. Proteins copurified with chromatin identified using shotgun approach

Protein accession	UniProt ID no.	Protein name	Protein probability	p*I*	MW (KDa)	Identified peptides no.
BAD36042	P46465	26S protease regulatory subunit 6A homolog	2.21E-04	4.3	47.8	2
ABC94594	A1XFD1	AF-4 domain containing protein-like protein	3.02E-05	6	165.6	5
BAD11340	Q762A2	BRI1-KD interacting protein 112	2.89E-06	4.3	30.4	2
BAD11362	Q761Y0	BRI1-KD interacting protein 135	6.56E-09	10.5	52.1	4
BAF30271	Q0ILZ4	DEAD-box ATP-dependent RNA helicase 9	2.32E-06	10.1	65.6	8
BAF30363	Q2QLR2	Glycine-rich RNA-binding protein GRP1A, putative, expressed	1.47E-03	6.3	16.1	2
BAC24887	Q8H4Z0	Putative histone H1	3.96E-05	10.5	28.5	3
EAY78499	A2Z7B5	Histone H2A	2.40E-05	10.5	14.6	8
EAY88652	A2XCU2	Histone H2A	2.40E-05	10.5	14.6	6
EAY91823	A2XLW3	Histone H2A	7.86E-05	4.3	36.4	3
EAZ04610	A2YNE9	Histone H2A	1.58E-07	10.5	20	5
EAZ16090	A3C4S7	Histone H2A	2.40E-05	10.5	14.6	8
EAZ20351	A3CGY8	Histone H2A	2.33E-05	10.5	16.7	2
EAZ20714	A3CI01	Histone H2A	5.31E-07	12.5	12.1	10
EAZ28546	A3AMK7	Histone H2A	7.86E-05	4.3	43	3
Q8H7Y8	Q8H7Y8	Probable histone H2A variant 1	4.22E-07	10.5	14.6	2
BAF26525	Q8S857	Probable histone H2A variant 2	2.40E-05	10.5	14.6	8
Q84MP7	Q84MP7	Probable histone H2A variant 3	9.61E-04	10.5	14.5	2
EAZ04236	A2YMC5	Probable histone H2A.1	2.33E-05	10.5	32.3	2
EAZ40190	Q6ZL43	Probable histone H2A.1	2.33E-05	10.5	14	6
EAZ04237	A2YMC6	Probable histone H2A.2	6.39E-07	10.5	14	2
EAZ40191	Q6ZL42	Probable histone H2A.2	6.39E-07	10.5	14	6
EAZ42802	Q84NJ4	Probable histone H2A.3	7.45E-03	10.5	13.9	5
Q94E96	Q94E96	Probable histone H2A.5	7.33E-10	10.5	16.4	12
EAY96290	A2XZN0	Probable histone H2A.6	7.33E-10	10.5	18	9
EAY82963	A2ZK29	Probable histone H2A.7	2.33E-05	10.5	14.1	2
Q2QS71	Q2QS71	Probable histone H2A.7	3.32E-05	10.5	14.1	3
EAY82960	A2ZK26	Probable histone H2A.8	2.33E-05	10.5	14	2
Q8LLP5	Q8LLP5	Probable histone H2AXa	9.86E-03	10.5	14.3	3
EAY83353	A2ZL69	Probable histone H2AXb	5.31E-07	10.5	14.3	3
Q2QPG9	Q2QPG9	Probable histone H2AXb	1.42E-03	10.5	14.3	4
BAF13146	Q10D30	Histone H2A. Z, putative, expressed	2.40E-05	10.5	14.5	6
EAZ12048	Q0JMH6	Histone H2B	4.45E-06	10.5	11.3	8
EAZ10545	Q94JJ7	Histone H2B.3	8.44E-04	10.5	16.5	2
EAZ10581	Q9LGH4	Histone H2B.6	1.97E-05	10.5	16.5	2
EAY94336	A2XU26	Histone H3	7.55E-15	10.5	22.9	5
EAY99520	A2Y8W0	Histone H3	7.30E-12	10.5	41.5	6
EAY99778	A2Y9L8	Histone H3	4.47E-08	11.5	16.8	3
EAZ27179	A3AIP0	Histone H3	7.30E-12	10.5	17.1	15
EAZ34442	A3B4F3	Histone H3	2.60E-03	12.5	19.6	6
EAZ20268	Q2QSW7	Histone H3, putative, expressed	1.50E-04	11.5	15.4	5
EAY99779	A2Y533	Histone H3.2	4.47E-08	12.5	15.3	5
EAZ35903	Q2RAD9	Histone H3.2	4.47E-08	12.5	15.3	12
EAY90303	A2XHJ3	Histone H3.3	7.30E-12	12.5	15.4	6
Q7F8L1	Q0JCT1	Histone H3.3	7.30E-12	12.5	15.4	15
EAY98338	A2Y5H8	Histone H4	2.26E-04	10.5	9.9	2
EAZ10018	A2WWR4	Histone H4	1.79E-08	12.5	11.4	7
EAZ34603	A3B4W3	Histone H4	2.32E-05	12.5	9.7	15
EAZ44836	A3BZ47	Histone H4	1.79E-08	12.5	30.9	22
AAO37519	Q851P9	Histone-like protein	7.18E-05	10.5	29.6	6
BAB63755	Q94D20	Nucleoid DNA-binding protein cnd41-like	2.50E-05	8.3	51.1	5
CAD27458	Q70Z21	Nucleosome assembly protein 1-like protein 1	4.38E-06	4.3	42.6	2
BAF10342	Q0DWP6	Os02g0803700 protein	2.21E-04	4.3	47.8	2
BAD72286	Q5SNC0	Putative 26S proteasome regulatory particle triple-A ATPase subunit5a	2.21E-04	4.3	47.8	2

Table 2. Continued

Protein accession	UniProt ID no.	Protein name	Protein probability	p*I*	MW (KDa)	Identified peptides no.
CAJ46990	Q9AYM0	Putative AT-Hook DNA-binding protein	3.60E-05	10.1	41.4	3
BAD88036	Q5JKH1	Putative BRI1-KD interacting protein 112	2.89E-06	6	56.3	4
BAC66711	Q84UR8	Putative cold shock protein-1	3.09E-07	6	18.7	2
AAV44069	Q5W6H1	Putative DNA-binding protein GBP16	2.88E-06	6	43.2	3
BAD22886	Q6K701	Putative fibrillarin	3.25E-09	10.5	32.4	8
AAT77341	Q6AT27	Putative fibrillarin protein	3.25E-09	10.5	37.2	4
NP_001059520	Q6Z142	Putative glycin-rich protein	1.35E-04	10.5	33	12
AAF70196	Q9M4T5	Putative histone deacetylase HD2	1.02E-06	4.3	32.5	5
BAD21720	Q6K8A9	Putative Neurofilament triplet M protein	8.59E-11	6	61.5	5
BAD68630	Q5VND6	Putative nucleosome assembly protein 1	4.38E-06	4.3	42.6	2
BAD69240	Q5VMT5	Putative Spo76 protein	3.02E-05	6	174.8	8
EAY76961	A2WYB5	Putative uncharacterized protein	3.51E-07	9.6	15.5	3
BAD81520	Q5NAI9	Putative Y1 protein	1.48E-11	6	50.1	8
ABF98117	Q10FE5	Retrotransposon protein, putative, Ty1-copia subclass, expressed	6.75E-04	4.3	15.2	3
BAF12784	Q10FE7	Retrotransposon protein, putative, Ty1-copia subclass, expressed	5.68E-05	7.2	20.4	4
ABA92999	Q6EQL9	Ribosomal protein-like	2.77E-05	10.1	48.2	2

b) UniProt no: UniProt number of the protein.
c) Probability: The probability of random match to the identified protein.
d) MW: predicted molecular mass.
e) p*I*: predicted protein p*I*.
f) Identified peptides: number of peptides that matched with the identified protein in MS/MS analyses.

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 copurify 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, antioxidative stress proteins, *etc*. The copurification 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 [35, 36, 46]. The GAPDH proteins were copurified 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 copurified with chromatin are indeed associated with chromatin *in vivo* in plants.

A large number of cytoskeleton proteins copurified with chromatin in our studies and also copurified with human metaphase chromosome and *C. elegance* chromatin [21, 22]. β-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 [21, 47]. In addition, injection of antiactin antibodies into *Xenopus* oocytes blocks chromosome condensation [48]. The associa-

tion of cytoskeleton proteins to chromatin may be related to chromatin spatial organization and dynamic movements [21].

4.3 The efficiency of chromatin protein identification using shotgun approach

Our previous studies using total protein extracts in *Arabidopsis* have shown that shotgun approach is more efficient than 2-DE gel based approach in protein identifications (data not shown). However, the number of proteins identified using shotgun method is much less than 2-DE gel based method 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. Therefore, other 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 2-D-LC-MS/MS process bias against some proteins [49]. 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.

4.4 Putative phosphoproteins copurified with chromatin revealed by Pro-Q Diamond stain

We have identified 509 proteins and these proteins correspond to 269 unique proteins using 2-DE gel approach. In our previous studies on *Arabidopsis* cotyledon proteins, the 500 proteins identified correspond to 353 unique proteins [42]. The ratio of unique protein in the chromatin proteome is lower. These results suggest a possibility that many chromatin associated proteins may have extensive PTMs 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 hypothetic protein (OsJ_024928) was found in spots 3401, 3413, and 4406. To validate if these proteins are phosphproteins, we stained the gel with Pro-Q Diamond phosphoprotein stain. Interestingly, these proteins are all stained by Pro-Q Diamond stain, suggesting that they can be modified by phosphorylation. Consistent with our results, many of them have been reported to be phosphoproteins, including H3-maize, H3.3, H2A, H2B, 26S proteasome regulatory subunit 8 [50], embryonic abundant protein [43], ribosome subunit L9 [51], glycine rich protein [52], WRKY DNA-binding protein [53], *etc*. These reports substantiates that the Pro-Q Diamond dye is a useful tool in the identification of candidate phosphoproteins. However, because many spots have more than one protein being identified, it is impossible for us to pinpoint which protein is phosphorylated in these protein spots. 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. Recently, several phosphoproteome analysis methods have been developed, including titanium dioxide microcolumn [54], immobilized metal-affinity chromatography method [55], phosphopeptide enrichment by IEF [56], and the phosphoprotein extraction kit of QIAGEN [57]. These methods in combination with MS analysis may provide a solution for the identification of phosphoproteins associated with chromatin.

Molecular biology and biochemistry studies have clearly demonstrated that chromatin is not simply a way of DNA packaging in the nucleus. Chromatin represents a highly conserved regulatory entity that provides a means of integrating multiple endogenous and exogenous signals for the establishment and maintenance of cellular gene expression profile during development and in response to environmental stimuli. We have purified chromatin from rice suspension cells, resolved 972 copurified protein spots on 2-DE gels, and identified 509 proteins with high confidence. In addition, 128 chromatin associated proteins are identified using shotgun approach, including histone variants, chromatin assembly factors, histone modification proteins, DNA binding proteins, transcription factors, and a large number of function unknown and hypothetical proteins. These studies should have critical reference value in future studies on chromatin regulation of gene expression as well as chromatin structure and function in plants.

We thank Dr. John Boyle, Dr. Kenneth Willeford, and Dr. Din-Pow Ma, for a critical reading of the manuscript. This project is supported by a USDA functional genomics grant (2005-35604-15357) to Z. Peng. This research was approved for publication as Journal Article J-10944 of the MAFES, Mississippi State University.

5 References

[1] Kornberg, R. D., *Science* 1974, *184,* 868–871.

[2] Olins, A. L., Olins, D. E., *Science* 1974, *183,* 330–332.

[3] Horn, P. J., Peterson, C. L., *Science* 2002, *297,* 1824–1827.

[4] Finch, J. T., Klug, A., *Proc. Natl. Acad. Sci. USA* 1976, *73*, 1897–1901.

[5] Robinson, P. J., Rhodes, D., *Curr. Opin. Struct. Biol.* 2006, *16*, 336–343.

[6] Woodcock, C. L., Skoultchi, A. I., Fan, Y., *Chromosome Res.* 2006, *14*, 17–25.

[7] Misteli, T., *Science* 2001, *291*, 843–847.

[8] Fischle, W., Wang, Y., Allis, C. D., *Curr. Opin. Cell Biol.* 2003, *15*, 172–183.

[9] Espada, J., Esteller, M., *Cell. Mol. Life Sci.* 2007, *64*, 449–457.

[10] Laemmli, U. K., *Pharmacol. Rev.* 1978, *30*, 469–476.

[11] Tackett, A. J., Dilworth, D. J., Davey, M. J., O'Donnell, M. *et al.*, *J. Cell. Biol.* 2005, *169*, 35–47.

[12] Cohen, B. A., Mitra, R. D., Hughes, J. D., Church, G. M., *Nat. Genet.* 2000, *26*, 183–186.

[13] Labrador, M., Corces, V. G., *Cell* 2002, *111*, 151–154.

[14] Burgess-Beusse, B., Farrell, C., Gaszner, M., Litt, M. *et al.*, *Proc. Natl. Acad. Sci. USA* 2002, *99*, 16433–16437.

[15] Nemeth, A., Langst, G., *Brief. Funct. Genomic. Proteomic.* 2004, *2*, 334–343.

[16] Gerasimova, T. I., Byrd, K., Corces, V. G., *Mol. Cell.* 2000, *6*, 1025–1035.

[17] Ishii, K., Arib, G., Lin, C., Van Houwe, G., Laemmli, U. K., *Cell* 2002, *109*, 551–562.

[18] Shelby, R. D., Hahn, K. M., Sullivan, K. F., *J. Cell. Biol.* 1996, *135*, 545–557.

[19] Li, G., Sudlow, G., Belmont, A. S., *J. Cell. Biol.* 1998, *140*, 975–989.

[20] Hozak, P., Hassan, A. B., Jackson, D. A., Cook, P. R., *Cell* 1993, *73*, 361–373.

[21] Uchiyama, S., Kobayashi, S., Takata, H., Ishihara, T. *et al.*, *J. Biol. Chem.* 2005, *280*, 16994–17004.

[22] Chu, D. S., Liu, H. B., Nix, P., Wu, T. F. *et al.*, *Nature* 2006, *443*, 101–105.

[23] Shiio, Y., Eisenman, R. N., Yi, E. C., Donohoe, S. *et al.*, *J. Am. Soc. Mass Spectrom.* 2003, *14*, 696–703.

[24] Baba, A., Hasezawa, S., Syono, K., *Plant Cell Physiol.* 1986, *27*, 463–471.

[25] Kyozuka, J., Shimamoto, K., in: Lindsey (Ed.), *Plant Tissue Culture Manual, B1*, Kluwer Academic Publisher, Dordrecht 1991, pp. 1–16.

[26] Yamada, Y., Zhi-Qi, Y., Ding-Tai, T., *Plant Cell Rep.* 1986, *5*, 85–88.

[27] Vengerov, Y. Y., Popenko, V. I., *Nucleic Acids Res.* 1977, *4*, 3017–3027.

[28] Hurkman, W. J., Tanaka, C. K., *Plant Physiol.* 1986, *81*, 802–806.

[29] Saravanan, R. S., Rose, J. K., *Proteomics* 2004, *4*, 2522–2532.

[30] Peng, Z., Serino, G., Deng, X. W., *Plant Cell* 2001, *13*, 2393–2407.

[31] Peng, Z., Serino, G., Deng, X. W., *Development* 2001, *128*, 4277–4288.

[32] Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D. *et al.*, *Nat. Genet.* 2000, *25*, 25–29.

[33] McCarthy, F. M., Bridges, S. M., Wang, N., Magee, G. B. *et al.*, *Nucleic Acids Res.* 2007, *35*, D599–D603.

[34] Chamovitz, D. A., Wei, N., Osterlund, M. T., von Arnim, A. G. *et al.*, *Cell* 1996, *86*, 115–121.

[35] Mazzola, J. L., Sirover, M. A., *Neurotoxicology* 2002, *23*, 603–609.

[36] Zheng, L., Roeder, R. G., Luo, Y., *Cell* 2003, *114*, 255–266.

[37] Schulenberg, B., Goodman, T. N., Aggeler, R., Capaldi, R. A., Patton, W. F., *Electrophoresis* 2004, *25*, 2526–2532.

[38] Steinberg, T. H., Agnew, B. J., Gee, K. R., Leung, W. Y. *et al.*, *Proteomics* 2003, *3*, 1128–1144.

[39] Agrawal, G. K., Thelen, J. J., *Proteomics* 2005, *5*, 4684–4688.

[40] Stasyk, T., Morandell, S., Bakry, R., Feuerstein, I. *et al.*, *Electrophoresis* 2005, *26*, 2850–2854.

[41] Agrawal, G. K., Thelen, J. J., *Mol. Cell. Proteomics* 2006, *5*, 2044–2059.

[42] Chitteti, B. R., Peng, Z., *Proteomics* 2007, *7*, 1473–1500.

[43] Chitteti, B. R., Peng, Z., *J. Proteome Res.* 2007, *6*, 1718–1727.

[44] Bernstein, E., Hake, S. B., *Biochem. Cell Biol.* 2006, *84*, 505–517.

[45] Hake, S. B., Allis, C. D., *Proc. Natl. Acad Sci. USA* 2006, *103*, 6428–6435.

[46] Hall, D. A., Zhu, H., Zhu, X. W., Royce, T. *et al.*, *Science* 2004, *306*, 482–484.

[47] Zhao, K., Wang, W., Rando, O. J., Xue, Y. *et al.*, *Cell* 1998, *95*, 625–636.

[48] Scheer, U., Hinssen, H., Franke, W. W., Jockusch, B. M., *Cell* 1984, *39*, 111–122.

[49] Mallick, P., Schirle, M., Chen, S. S., Flory, M. R. *et al.*, *Nat. Biotechnol.* 2007, *25*, 125–131.

[50] Rivett, A. J., Bose, S., Brooks, P., Broadfoot, K. I., *Biochimie* 2001, *83*, 363–366.

[51] Kruiswijk, T., de Hey, J. T., Planta, R. J., *Biochem. J.* 1978, *175*, 213–219.

[52] Vilardell, J., Goday, A., Freire, M., Torrent, M. *et al.*, *Plant Mol. Biol.* 1990, *14*, 423–432.

[53] Yang, P., Chen, C., Wang, Z., Fan, B., Chen, Z., *Plant J.* 1999, *18*, 141–149.

[54] Larsen, M. R., Thingholm, T. E., Jensen, O. N., Roepstorff, P., Jorgensen, T. J., *Mol. Cell. Proteomics* 2005, *4*, 873–886.

[55] Ficarro, S. B., McCleland, M. L., Stukenberg, P. T., Burke, D. J. *et al.*, *Nat. Biotechnol.* 2002, *20*, 301–305.

[56] Maccarrone, G., Kolb, N., Teplytska, L., Birg, I. *et al.*, *Electrophoresis* 2006, *27*, 4585–4595.

[57] Jones, A. M., Bennett, M. H., Mansfield, J. W., Grant, M., *Proteomics* 2006, *6*, 4155–4165.

