DOI 10.1002/pmic.201200389 *Proteomics* 2013, *13*, 1784–1800

REVIEW

Use of proteomics to understand seed development in rice

Zhu Yun Deng, Chun Yan Gong and Tai Wang

Key Laboratory of Plant Molecular Physiology, Institute of Botany, Chinese Academy of Sciences, Haidianqu, Beijing, China

Received: August 28, 2012
Revised: December 24, 2012
Accepted: January 7, 2013

Rice is an important cereal crop and has become a model monocot for research into crop biology. Rice seeds currently feed more than half of the world's population and the demand for rice seeds is rapidly increasing because of the fast-growing world population. However, the molecular mechanisms underlying rice seed development is incompletely understood. Genetic and molecular studies have developed our understanding of substantial proteins related to rice seed development. Recent advancements in proteomics have revolutionized the research on seed development at the single gene or protein level. Proteomic studies in rice seeds have provided the molecular explanation for cellular and metabolic events as well as environmental stress responses that occur during embryo and endosperm development. They have also led to the new identification of a large number of proteins associated with regulating seed development such as those involved in stress tolerance and RNA metabolism. In the future, proteomics, combined with genetic, cytological, and molecular tools, will help to elucidate the molecular pathways underlying seed development control and help in the development of valuable and potential strategies for improving yield, quality, and stress tolerance in rice and other cereals. Here, we reviewed recent progress in understanding the mechanisms of seed development in rice with the use of proteomics.

Keywords:
Cereal crop / Monocot / Plant proteomics / Rice / Seed development

1 Introduction

Rice seeds represent one of the most important staple foods for humans, currently feeding more than half of the world's population; and the demand for rice seeds is expanding rapidly because of the fast-growing world population [1]. In addition to being consumed as staple food, rice seeds are used to make edible starches, vinegars, syrups, and sweet alcoholic beverages. The bran of rice seeds is a raw material for extracting oil used in the cosmetic and healthcare industries. Besides the aforementioned practical importance for food security and economic production, two other reasons make rice a good model monocot system to study seed biology. First, the genome of rice is relatively smaller but highly conserved as compared with that of other cereal crops such as maize and barley. Second, the complete genome sequence of rice is available in public databases such as the National Center for Biotechnology Information (NCBI; http://www.ncbi.nlm.nih.gov/) and the Rice Genome Annotation Project Database (RGAP; http://rice.plantbiology.msu.edu/). Therefore, rice seed biology has been broadly studied at different levels by use of histological, physiological, cytological, biochemical, genetic, and molecular tools. These studies have greatly expanded our knowledge of rice seed development and clearly demonstrate that (i) seed developmental events are programmed to occur

Correspondence: Professor Tai Wang, Key Laboratory of Plant Molecular Physiology, Institute of Botany, Chinese Academy of Sciences, 20 Nanxincun, Xiangshan, Haidianqu, Beijing 100093, China
E-mail: twang@ibcas.ac.cn
Fax: +0086-10-62594170

Abbreviations: **DAF**, day after fertilization; **DEPs**, differentially expressed proteins; **DHT**, day high temperature; **ER**, endoplasmic reticulum; **GBSS**, granule-bound starch synthase; **hnRNPs**, heterogeneous nuclear ribonucleoprotein; **KH**, K-homology; **LEAs**, late-embryogenesis-abundant proteins; **NCBI**, National Center for Biotechnology Information; **NHT**, night high temperature; **PCD**, programmed cell death; **PPDKB**, pyruvate orthophosphate dikinase B; **RACK**, receptor for activated C-kinase; **Rboh**, respiratory burst oxidase homolog; **RBPs**, RNA-binding proteins; **REG**, rice embryo globulins; **RRM**, RNA recognition motif; **SOD**, superoxide dismutase; **ssDNA**, single-stranded DNA; **SSPs**, seed storage proteins; **TCA**, tricarboxylic acid; **WT**, wild-type

Colour Online: See the article online to view Table 3 in colour.

as a result of expression and activation of different proteins in distinct seed compartments (i.e. embryo, endosperm, and caryopsis coat) and even within specific regions (e.g. apical meristem), at distinct developmental stages [2, 3]; and (ii) seeds regulate the expression and/or activation of specific proteins to respond efficiently to a highly dynamic environment (e.g. temperature, water, and light) during development [4]. However, the number of identified proteins active in these processes is limited, and elucidation of these regulatory networks remains a major challenge.

Proteomics has provided an efficient large-scale solution complementing traditional and genomic approaches for investigating rice seed development. Compared with traditional molecular and genetic strategies, proteomics involves global comparative studies, allowing for large-scale identification and expression analysis of proteins active in specific seed compartments or regions at certain developmental stages. This capability may make it possible for us to answer many questions that seem unsolvable at the single gene or protein level and to discover novel development-related proteins that may be undetectable by forward or reverse genetic methods (e.g. redundant or lethal proteins involved in rice seed development but without homologues in other species). As compared with genomics and transcriptomics tools such as DNA microarray, proteomics focuses directly on proteins, the final active agents in developing seeds. Although proteins are the translated products of gene transcripts and large-scale analysis of genes or transcripts is relatively easy and efficient, proteomics has irreplaceable superiority. Alternative RNA splicing and various posttranslational modifications such as phosphorylation, methylation, glycosylation, ubiquitylation, and proteolytical processing may change protein localization, activity, stability, and/or function [5], which may play important roles in seed development. Protein isoforms derived from these processes can be identified and analyzed only by proteomic methods, not genomics or transcriptomics. In addition, quantitative proteomics techniques must be used to directly measure changes in protein abundance during seed development and environmental response, because RNA levels are not always correlated with the levels of corresponding proteins. Finally, the localization of gene products, providing important clues of gene function as well as regulatory networks activated in different seed regions, can be determined only at the protein level. Characterization of proteomes expressed in specific cell types, tissues, regions, or organs of seeds (called subproteomics) is a powerful approach for understanding the molecular mechanisms underlying seed development. Due to these advantages, proteomic technologies have been increasingly used to investigate rice seeds over the past two decades. Most proteomic studies of rice seeds have aimed to provide practical or theoretical information helpful for improving the quality and yield of rice and other cereals. Proteomic studies relevant to seed development focused on the important but unanswered issues regarding embryo and endosperm development in rice: (i) metabolic and molecular events that occur at different seed developmental stages, (ii) the effects of environmental factors on seed development and rice quality, and (iii) specific biological processes during seed development such as storage protein biosynthesis and its regulation (Table 1).

In this review, we first briefly describe the structure and development process of rice seeds, providing the anatomical and cellular basis for proteomic studies related to rice seed development. Then, we discuss recent proteomic studies that have contributed to our understanding of the mechanisms underlying seed development in rice, as well as the proteomic technologies used in these studies (Table 1).

2 Rice seed structure and development

2.1 Seed structure

The rice seed is strictly considered a caryopsis. It is a single-seeded fruit with the seed fused with the fruit coat. Therefore, a mature rice seed or caryopsis comprises three main structural elements: a diploid embryo, a triploid endosperm, and a maternal caryopsis coat (Fig. 1A). The embryo represents the new generation, residing at a small corner of the caryopsis. Its main portions include scutellum, coleoptile, plumule, epiblast, and radicle (Fig. 1B). The scutellum is the apical portion of the embryo, acting as a storage organ during seed development and helping to digest and transport endosperm nutrients during seed germination. The coleoptile, the radicle, and the plumule containing the shoot apical meristem and three leaf primordia, are precursor organs of rice seedlings. The coleoptile protects the plumule during germination. The plumule and the radicle eventually develop into the shoot and the seminal root, respectively. The endosperm is the major organ for storage of nutritional reserves, occupying the bulk of a mature rice seed. It consists of the central starchy endosperm filled with lifeless starchy endosperm cells, an aleurone layer surrounding the starchy endosperm, and a basal transfer layer between the embryo and the endosperm. Roughly 90% of the nutrients in the endosperm cells are starches stored in starch granules, and 5–8% of the nutrients are storage proteins held in proteoplasts (protein bodies). In contrast, the aleurone layer mainly accumulates proteins and lipids. The caryopsis coat consists of several layers of the pericarp (fruit coat), one layer of the testa (seed coat) and one layer of the nucellus, coating the seed and serving as a protection for the embryo and the endosperm.

2.2 Seed development

The rice seed initiates its development from double fertilization and matures after approximately 20 days of growth. During double fertilization in the embryo sac, one sperm nucleus fuses with the nucleus of the egg cell to produce the diploid zygote and the other sperm nucleus fuses with

Table 1. Recent proteomic studies related to rice seed development

Focused issues	Studied proteomes	Developmental stages	Extracted protein fractions	Separation methods (p*I* range)	Identification methods	No. of identified spots (identities, *unique proteins*)	Ref.
Whole seed development							
Molecular, cellular, and metabolic events during seed development	Proteome changes in whole seeds during the complete developmental process	2, 4, 6, 8, 10, 12,14, 16, 18, and 20 DAF[a)]	Total low salt-soluble fraction from developing seeds	2DE (4–7)	MALDI-TOF	309 (345, *227*)	[32]
Endosperm development							
Molecular, cellular, and metabolic events during starch accumulation	Proteome changes in endosperms during the period of rapid starch accumulation	12, 15, and 18 DAF[a)]	Total low salt-soluble fraction from developing endosperms	2D-DIGE (4–7)	MALDI-TOF	298 (317, *187*)	[16]
Molecular, cellular, and metabolic regulation by light-dark cycle	Diurnal changes in endosperm proteomes at the middle phase of endosperm development	10 DAF[a)] (0, 4, 8 HAO[b)] at light phase, 12 HAO[b)] at light-dark transition, 16 and 20 HAO[b)] at dark phase)	Total low salt-soluble fraction from developing endosperms	2D-DIGE (4–7)	MALDI-TOF	81 (91, *62*)	[33]
Embryo development or embryogenesis							
Molecular and metabolic events during embryogenesis	Proteomic comparison among embryos at different developmental stages, as well as mature and dry embryos	5, 7, 14, 21, 30 DAF[a)]	Total urea-soluble fraction from developing, mature, and dry embryos	2DE (3–10)	LC/MS/MS	242 (275, *192*)	[28]
Molecular and metabolic events during embryogenesis; globulins and their roles in embryogenesis	Proteome changes in embryos during the middle and late developmental phases	6, 12, 18 DAF[a)]	Total urea-soluble fraction from developing embryos	2DE (3–10)	MALDI-TOF/TOF	53 (55, *23*)	[29]
Stress response and tolerance							
Effects of high temperature on rice quality and seed development	Proteome changes in seeds by high temperature treatment during seed development	3, 6, 9, 12, 15, and 30 DAF[a)]	Total urea and triton-soluble fraction from developing seeds	2DE (3–10)	LC/MS/MS	54 (54, *48*)	[26]
Difference in effects by day and night high temperature during seed development	Proteome changes in seeds by day or night high temperature treatment during seed development	5, 10, 15, and 20 DAF[a)]	Total NP-40 and low salt-soluble fraction from developing seeds	2DE (4–7)	MALDI-TOF; MALDI-TOF/TOF; LTQ-ESI-MS/MS	61 (61, *46*)	[25]

Table 1. Continued

Focused issues	Studied proteomes	Developmental stages	Extracted protein fractions	Separation methods (p*I* range)	Identification methods	No. of identified spots (identities, *unique proteins*)	Ref.
Molecular mechanisms underlying abiotic stress tolerance	Comparison of embryo proteomes among stress-tolerant and -sensitive rice varieties	Mature stage	Total urea-soluble fraction from mature and dry embryos	2DE (3–11)	MALDI-TOF; LC-MS/MS; De novo sequencing	28 (28, *22*)	[30]
Specific biological processes during seed development							
G protein-mediated transduction pathway in embryo development	Proteins regulated by the α subunit of a heterotrimeric G protein during embryogenesis	14, 21, 28, 35 DAF[a)] and germination	Total urea and NP-40-soluble fraction from developing and germinating embryos	2DE (3.5–7 and 6–10); Cleveland peptide mapping	Edman sequencing	7 (7, *3*)	[27]
Nuclear proteins involved in seed development	Nuclear subproteome of seeds at the middle developmental phase	9 DAF[a)]	Phenol-extracted fraction from nuclei of developing seeds	1DE; 2DE (3–11); gel free	MALDI-TOF/TOF; LC/LC-MS/MS	468 (468, *468*)	[35]
RNA-binding proteins serving protective functions during seed maturation and dormancy	RNA-binding proteins in mature dry seeds	Mature stage	Low salt-soluble and ssDNA-binding fraction from dry seeds	2DE (3–10)	MALDI-TOF	18 (18, *9*)	[36]
RNA-binding proteins implicated in seed development	Cytoskeleton-associated RNA-binding proteins in developing seeds	12–14 DAF[a)]	Low salt-soluble, cytoskeletal-enriched, and ploy(U)-binding fraction from developing seeds	2DE (6–11 and 3–10)	Q-TOF	162 (*148* RNA-binding proteins)	[37]
Prolamine mRNA-binding proteins related to mRNA transportation and localization to ER during storage protein synthesis	Cytoskeleton-associated prolamine mRNA-binding proteins in developing seeds	10–15 DAF[a)]	Low salt-soluble, cytoskeletal-enriched, and prolamine zipcode-binding fraction from developing seeds	1DE	LC–MS/MS	Prolamine mRNA-binding proteins captured with (*18*) or without (*132*) heparin	[38, 39]

a) DAF, day after fertilization; b) HAO, hours after the onset of 12-h light and 12-h dark cycle.

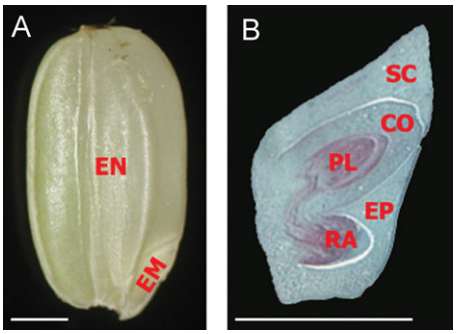

Figure 1. Structure of the rice seed and its embryo. (A) Morphology of a mature seed of rice. EN, endosperm; EM, embryo; scale bar = 1 mm. (B) Longitudinal section of the rice embryo. SC, scutellum; CO, coleoptile; PL, plumule; EP, epiblast; and RA, radicle; scale bar = 1 mm. Adapted from [32].

the two polar nuclei of the central cell to create the triploid primary endosperm nucleus. Afterwards, the zygote and the primary endosperm nucleus develop relatively independently to become the embryo and the endosperm, respectively. Coincidently, the maternal ovary undergoes concordant growth and changes to fit the expanding embryo and endosperm. Finally, the ovary wall and the ovule integument form the caryopsis coat. Here, we concisely describe, respectively, the processes of embryo (Table 2) and endosperm (Table 3) development. Detailed processes of embryo and endosperm development have been documented in previous reviews [2, 6].

2.2.1 Embryo development

The zygote begins its first cell division at about 10 h after double fertilization [7] and continues to undergo rapid cell division in an unfixed direction, thus resulting in a 25-celled early globular embryo at 1 day after fertilization (DAF) [2]. The cell division rate decreases after the 25-celled stage, and the middle globular embryo reaches 150 cells at 2 DAF [2]. Accompanied by cell division, early embryonic events that include determination of organ differentiation, establishment of organ position, and control of embryo size occur before morphogenetic changes are visible [8]. In the late globular stage, at 3 DAF, besides increasing cell number from 150 to 800 cells, the embryo starts organ differentiation and forms the dorsal-ventral axis, making the embryo an oblong shape [2]. At early 4 DAF, the first organogenetic events take place when the coleoptile develops from the ventral surface of the embryo and the scutellum is distinguishable by cells with vacuolated cytoplasm [3, 9]. At late 4 DAF, differentiating shoot apical meristem and radical are morphologically recognized below the coleoptile [9]. At 5 DAF, the size of the scutellum increases obviously and the first leaf primordium emerges [2]. Then the second and third leaf primordia become visible at 7 and 8 DAF, respectively. At the same time, the epiblast stretches out and enlarges [2]. At 10 DAF, morphological differentiation of embryonic organs is accomplished, but volume augmentation of these organs is remarkable [2]. After 10 DAF, the embryo increases its size slightly and undergoes rapid maturation [2]. Eventually, the embryo completes the development process and becomes dormant after 20 DAF [2].

2.2.2 Endosperm development

The endosperm grows at a relatively rapid rate as compared with the embryo. The primary endosperm nucleus starts its first division of nucleus immediately after double fertilization without halting, and keeps on dividing nuclei in the absence of cell-wall formation at a rate faster than the division of the zygote, thus leading to formation of free nuclei with dense cytoplasm [7]. These free nuclei distribute to the periphery of the embryo sac cavity and establish a multinucleate layer at 2 DAF [7]. Accompanied by the division of endosperm nuclei, nuclear DNA content is increased from 3 to 6–12 C by endoreduplication [10]. At 3 DAF, the endosperm begins to cellularize, and cell walls form initially in the region near to the developing embryo and then extend to the entire endosperm layer [7]. Simultaneously, all of the cells undergo an inward division, thus generating an endosperm with two cell layers [7]. Afterwards, the endosperm divides rapidly to significantly increase the number of cell layers and cell numbers of each layer until peak number of cells is achieved at 7 DAF [11]. In addition to cell division, many important developmental events occur during this period. The embryo sac cavity is stuffed with endosperm cells at 4 DAF [11]. DNA endoreduplication continues in the dividing endosperm cells, thus leading to increased nuclear DNA content to 12–24 C at 6 DAF [10]. Endosperm cells begin to accumulate starch and starch granules become visible in endosperm cells at 4 DAF [11]. The outermost endosperm cells differentiate into aleurone cells at 5 DAF. Cell divisions, accompanied by differentiation, leads to the formation of the aleurone layer at 7 DAF [12]. Through increasing cell number, the endosperm grows to its full length at 7 DAF. However, after the cell division stage, further growth of the endosperm continues by enlargement of individual cells. Finally, the endosperm reaches its full width and thickness at 12 and 15 DAF, respectively [13]. Along with cell enlargement, the endosperm continues to deposit starches mainly in endosperm cells and proteins and lipids mainly in aleurone cells. The size of starch granules in endosperm cells increases and reaches the maximum at 15 DAF [14, 15]. Endosperm cells keep on endoreduplicating nuclear DNA, and the DNA content peaks (30 C) at 16 DAF [10]. Ultimately, mature endosperm cells undergo programmed cell death (PCD) at 12–20 DAF [16], leaving only the aleurone layer alive at seed maturity.

Table 2. Embryo development

DAF	Embryo developmental stages		Embryo developmental events [2, 3, 7-9, 32]			Proteomics events [28, 29]
			Shape and size	Cell division	Differentiation	
0	Double fertilization			Zygote formation		
1	Globular stage					
		Early	Spherical shape	Increase in cell number from 1 to 25 by rapid cell division	Determination of organ differentiation; establishment of organ position; control of embryo size	
2		Middle	Spherical shape; 50-55 μm in length	Increase in cell number from 25 to 150 by relatively slower cell division		
3		Late	Oblong shape; 100-110 μm in length	Increase in cell number from 150 to 800	Formation of dorsal-ventral axis; start of organ differentiation	
4	Organ Differentiation					
		Apical meristem differentiation	200-300 μm in length	Continue to increase cell number	Differentiation of coleoptile, shoot apical meristem, radical and scutellum	
5		Leaf primordium differentiation	Rapid enlargement of scutellum; 400-500 μm in length		First leaf primordium formation	Higher expression of proteins involved in cell growth and division, central carbon metabolism (glycolysis and TCA cycle) and protein turnover (protein synthesis, amino acid metabolism and proteolysis).
6			Rapid enlargement of scutellum; 0.8-1.1 mm in length			
7			Enlargement of epiblast; 1.3-1.5 mm in length		Second leaf primordium formation; protrusion of epiblast	
8					Third leaf primordium formation	
9		Organ enlargement	Column augmentation of differentiated organs; 1.5-1.6 mm in length			
10					Accomplishment of morphological differentiation in embryonic organs	
11-20	Maturation		Slight enlargement in embryo size		Rapid maturation	Higher expression of proteins related to protein folding/ modification, lipid/sterol metabolism and stress defense.
21	Dormancy and desiccation		Roughly 1.4 mm in length			

DAF, day after fertilization.

Table 3. Endosperm development

DAF	Endosperm developmental stages	Endosperm developmental events [6, 7, 11-16, 24, 32, 60, 61]				Proteomics events [16, 32]
		Endosperm Size	Mitosis	Endoreduplication	Differentiation	Accumulation of Storage Compounds	
0	DF		Formation of primary endosperm nucleus	Nuclear DNA content is 3C			
1	Syncytium formation						
2	Nuclear division		Multinucleate layer formation	Increase in nuclear DNA content to 6-12 C			Maximal accumulation of proteins mediating cell growth, division and morphogenesis and signal transduction. Higher expression of proteins involved in central carbon metabolism (glycolysis and TCA cycle) and protein turnover (protein synthesis, amino acid metabolism and proteolysis). Lower expression of proteins related to starch synthesis and alcoholic fermentation. Up-regulated expression of proteins involved in starch synthesis and alcoholic fermentation. Continued up-regulation of proteins associated with starch synthesis and alcoholic fermentation. Dynamic changes in the expression of proteins involved in ROS-mediated PCD.
3	Cellularization		Cell wall formation in the multinucleate layer; onset of cell division				
4	Cell division		Entire embryo sac cavity is filled with endosperm cells			Start rapid accumulation of starch; starch granules are visible	
5	Endoreduplication				Outmost endosperm cells differentiate into aleurone cells	Storage proteins begin to synthesize	
6	Cell enlargement			Increase in nuclear DNA content to 12-24 C		Beginning of proteoplast deposition	
7	Differentiation	Full length	Reach the maximum cell number		Formation of aleurone layer		
8-11	Storage phase (accumulation of storage compounds)						
12		Full width					
13-14							
15		Full thickness			Completion of morphological differentiation	Starch granules reach the maximum size	
16	Programmed cell death			Nuclear DNA content reach the highest value (30 C)			
17-20							
21-	Dormancy and desiccation						

DAF, day after fertilization; DF, double fertilization.

3 Proteomic techniques used in the study of seed development in rice

Over the past 20 years, proteomics has become a major tool for surveying protein profiles in rice seed development, providing useful information at the proteomic level. Proteomic researches into rice seed development employed the same experimental workflow and the same proteomic techniques as other proteomic studies in plants, especially rice, which have been well documented and discussed [17–20]. However, multiple technique improvements, such as extraction and purification methods for obtaining low-abundance seed proteins, have been made in the proteomic studies related to rice seeds, which have progressed seed proteomics and led to a deeper understanding of seed development in rice. Here, we describe the extraction, separation, and identification approaches used in the proteomic studies reviewed here and highlight their technical progresses and improvements.

3.1 Extraction of rice seed proteins with or without seed storage proteins

Mature rice seeds have a total protein content ranging from 5 to 12% [21, 22], most of which is contributed by seed storage proteins (SSPs), including glutelins, albumins, globulins, and prolamines. The predominant SSPs of rice seeds are glutelins, which constitute 80% or more of the total seed proteins. The remaining three types of SSPs, respectively, account for 1–5% (albumins), 4–15% (globulins), and 2–8% (prolamines) of the total seed proteins [23]. During seed development, SSPs are first synthesized as precursor proteins, then processed and finally deposited in protein bodies in rice seeds [24]. Therefore, SSPs are highly abundant in both developing and mature seeds.

Most studies of rice seeds used total proteins, including SSPs, as samples for proteomic analysis. Total proteins were first extracted from whole seeds [25, 26], endosperms [27], or embryos [27–30] by use of strong denaturing solutions containing 8 M urea and/or detergents and then subjected to 2DE directly or after further extraction and concentration by the trichloroacetic acid/acetone method. In addition, several other studies extracted only soluble proteins to eliminate the impacts of SSPs on further analysis. These studies used low-ionic-strength and neutral pH buffers for protein isolation, which minimized the extraction of SSPs because only albumins are water soluble, whereas glutelins, globulins, and prolamines are soluble in alkali or acid, salt, and aqueous alcohol solution, respectively. For example, through grinding materials in a simple extraction buffer with 20 mM Tris-HCl (pH 8.0), 20 mM NaCl, and 1 mM EDTA, followed by trichloroacetic acid/acetone precipitation, our laboratory has extracted low salt-soluble total proteins with a few SSPs from mature [31] and developing [32] rice seeds and developing rice endosperms [16, 33].

For subproteomic analysis of rice seeds, the highly abundant SSPs substantially mask low-abundance seed proteins, thus affecting the effectiveness of proteomic studies with shotgun or 2DE gel-based approaches [34]. Therefore, selective removal of SSPs from seed protein extracts is a key step in sample preparation. However, SSPs cannot be completely removed by differential fractionation because protein bodies (in which SSPs are stored) and various SSP precursor-containing organelles (e.g. vesicles, Golgi apparatus, and ER) have different sedimentation coefficients [35]. SSPs may be isolated or removed in light of their physical and chemical properties. For example, acid or alkali treatment can eliminate most glutelins from seed proteins. However, these methods have not been commonly used in rice seed proteomics. In a rice nuclear subproteome study, Peng et al. first eliminated starch granules and parts of SSPs sticking on starch granules by cotton filtration after isolating nuclei from developing rice seeds at 9 DAF, then removed the major remaining SSPs (i.e. 38 and 55 kD glutelin) by directly cutting out the two protein bands after 1DE separation of the extracted proteins from the isolated nuclei [35]. The resulting 1DE gels had a relatively large proportion of nuclear proteins and were ground into powder and extracted with phenol to restore nuclear proteins [35]. Using this two-step removal method, the authors obtained about tenfold more proteins than in total protein extracts of rice seeds. More importantly, the obtained proteins included many low-abundance transcriptional factors [35]. Affinity chromatography is another powerful method for removing SSPs and enriching low-abundance seed proteins. Motoki and associates fractionated RNA-binding proteins (RBPs) from mature dry rice seeds by single-stranded DNA (ssDNA) affinity column chromatography [36]. This one-step affinity purification removed all SSPs detected in the total protein extracts [36]. Okita's group applied poly(U)-sepharose affinity columns and streptavidin magnetic beads bound with biotinylated prolamine mRNAs to isolate general and prolamine RBPs, respectively, from cytoskeleton-enriched proteins of developing rice seeds [37, 38]. To identify general cytosolic RBPs, the authors used differential centrifugation and sucrose gradients to remove chromatin and nuclei and enrich cytoskeleton-associated fractions, then poly(U)-sepharose affinity purification to obtain the cytoskeleton and poly(U)-binding fractions [37]. The resulting proteins were separated by 2DE with both pH 6–11 and pH 3–10 nonlinear immobilized pH gradient strips [37]. After MS and peptide analysis, 124 proteins with basic net charges and 38 proteins with acidic isoelectric points were identified from the pH 6–11 and pH 3–10 2DE gels, respectively [37]. Besides 21 SSPs among the 162 identified proteins, 148 cytoplasmic localized, cytoskeleton, and nucleic acid-binding proteins were obtained ultimately [37]. To capture prolamine RBPs, the enriched cytoskeleton-associated fractions were incubated with streptavidin magnetic beads bound with the biotinylated RNA oligos containing the prolamine 5'CDS zipcode sequence [38]. This capture method led to the isolation of 138

cytoplasmic localized, cytoskeleton, and prolamine mRNA-binding proteins [39].

3.2 Separation and identification of rice seed proteins

The frequently used separation and identification technologies are widely applied in rice seed proteomic studies. Almost all of the studies we review involved gel-based approaches, most combining 2DE separation and MS identification. Until now, gel-free-based quantitative approaches have been rarely used to analyze rice seed protein extracts.

In an early study comparing seed proteomes between wild-type (WT) rice and the *dwarf1* mutant, total embryo, and endosperm proteins extracted from developing embryos and endosperms, respectively, were separated on 2DE gels and electroblotted onto PVDF membranes, then identified by Edman sequencing on a gas-phase protein sequencer [27]. Because Edman sequencing is slow and requires large amounts of samples, the identification technique for protein spots excised from 2DE gels has currently switched to MS. Developing rice seed proteins [32], endosperm proteins [40], embryo proteins [29,40,41], and mature seed RBPs [36] have been successfully identified by MALDI-TOF-MS. Quadrupole-TOF-MS (Q-TOF-MS) is also used in 2DE-based approaches; e.g. developing rice seed RBPs from silver-stained gels were identified by Q-TOF-MS [37].

In addition to the improvement in identification, a more sensitive separation method with high reproducibility and low demand for sample amounts, namely 2D-DIGE, has been recently used to investigate endosperm proteomes during starch accumulation [16] and its day-night cycle [33]. Although 1DE has a lower resolution than 2DE, it is quick, simple, and easy to perform, providing a useful complement in some proteomic studies. For RBP isolation, cytoskeleton-associated prolamine mRNA-binding proteins were separated from developing rice seeds by 1DE before LC-MS/MS identification [38]. In analyzing rice nuclear proteomes, 1DE was used to separate storage proteins from seed nuclear proteins. Storage protein bands were then excised from the 1DE gel and the remaining proteins were recovered for gel-free LC/LC-MS/MS by the shotgun method [35].

4 Advances and challenges in proteomic analysis of rice seed development

4.1 Proteomic studies of different seed developmental stages

The development of rice seeds is a highly coordinated process that relies on different sets of proteins expressed at the right time. Proteomics has offered powerful approaches for establishing a global view of protein expression at different seed developmental stages, helping to elucidate the control and coordination mechanism underlying seed development. Through 2DE or 2D-DIGE combined with MS, several studies have compared the proteomes of rice embryos [28, 29], endosperms [16, 33], or whole seeds [32] at different developmental stages, especially during the accumulation of reserves, which have largely broadened our understanding of the molecular regulation and metabolic networks during seed development.

Owning to the importance of reserve accumulation for both seed development and yield improvement in rice, our laboratory has performed several proteomic investigations to elucidate the molecular and metabolic mechanisms related to this process. First, we utilized 2DE and MS to analyze the differences in whole seed proteomes among eight developmental stages covering the full period of reserve accumulation, which led to the identification of 345 differentially expressed proteins (DEPs) [32]. The 345 DEPs were classified into nine functional categories, including metabolism (45%), protein synthesis/destination (20%), defense response, and cell growth/division. The metabolism category was further divided into 11 subcategories containing glycolysis, alcoholic fermentation, tricarboxylic acid (TCA) cycle, starch synthesis, amino acid metabolism, etc. Using our newly established method, called the digital expression profile analysis, we revealed that the accumulation of proteins in different categories or subcategories is associated with the cellular, molecular, and metabolic events that occur in a timely and orderly manner during seed development. From 2 to 8 DAF, proteins mediating cell growth, division, and morphogenesis, such as cytoskeleton proteins (tubulin, actin) and their associated proteins (actin-binding protein profilin), showed maximal accumulation, which supports the rapid increase in seed size at the early developmental phase. Meanwhile, at this phase, proteins related to central carbon metabolism (glycolysis and TCA cycle) showed high expression and those related to alcoholic fermentation showed low expression. In addition, the high expression of proteins involved in protein synthesis, amino acid metabolism, and proteolysis, such as ubiquitin/26S proteasome components, indicates an active turnover of proteins during the early developmental stages. At 8–12 DAF, when the growth rate of developing seeds slowed, the expression of proteins related to growth was decreased, with an increase in expression of starch synthesis-related proteins, which demonstrates a transition from cell growth to reserve accumulation from this stage. Proteins implicated in central carbon metabolism were downregulated, whereas those involved in alcoholic fermentation became upregulated from this stage and remained upregulated until seeds matured. At 12–20 DAF, accompanied by the fast deposition of storage compounds, a large number of starch synthesis-related proteins were upregulated; protein folding and modification-related proteins showed a similar tendency, which indicates the importance of protein folding and modification in the accumulation of starch and storage proteins. This proteomic study indicates that the switch from central carbon metabolism to alcoholic fermentation is

essential for starch synthesis and accumulation during seed development.

Based on the observations that starch accumulation mainly occurs in the endosperm between 12 and 18 DAF and most starch synthesis-related proteins are upregulated also in this period [32], we further examined the cytological, physiological, and proteomic alterations of developing endosperm at 12, 15, and 18 DAF by combining biochemical and cytological detection with 2D-DIGE analysis, the improved 2DE technology allowing for better quantitative comparison between samples [16]. The biochemical and cytological detection showed that rice endosperm completed the deposition of starch molecules and the formation of organized starch granules in the central part at 15 DAF and in nearly the entire portion except the outmost part at 18 DAF. Accompanied by starch accumulation, reactive oxygen species (ROS), represented by hydrogen peroxide, started to appear in the innermost part of the endosperm at 12 DAF, disappeared from the central part but burst into the remaining parts at 15 DAF, and were present only in the outermost part at 18 DAF. Lifeless endosperm cells were present in the part where starch granule packaging was accomplished and hydrogen peroxide had disappeared. These observations demonstrate the emergence of hydrogen peroxide, the completion of starch accumulation and PCD occur successively in the same endosperm region; whereas all of the three developmental events initiate from the inner endosperm cells, then expand toward the periphery endosperm cells. ROS including hydrogen peroxide are mainly formed as byproducts of the activity of oxidases such as respiratory burst oxidase homolog (Rboh). After ROS are produced, their levels are adjusted by ROS-scavenging enzymes and the remaining ROS act as key signaling molecules inducing redox-dependent regulation of diverse molecular or metabolic processes during plant development [42]. Real-time RT-PCR analysis revealed two rice *Rboh* genes, *OsRbohB* and *OsRbohH*, showed the highest transcriptional level at 15 DAF, and another *Rboh* gene, *OsRbohD*, was greatly upregulated at 15 DAF, with the peak level at 18 DAF. Thus, upregulation of ROS-producing enzymes is chronologically consistent with the burst of hydrogen peroxide. Proteomic analysis detected 317 DEPs among the three developmental stages, which included 22 ROS-scavenging proteins showing diverse expression regulation. Among them, thioredoxin (one of the most important oxidoreductases in redox regulation), GSH peroxidase (a key ROS scavenger), phospholipid hydroperoxide glutathione peroxidases, and trypanothione-dependent peroxidases showed the highest expression levels at 15 DAF. Mondehydroascorbate reductase showed the lowest level at 15 DAF. Four isoforms of superoxide dismutase and peroxiredoxin showed peak expression at 18 DAF. Four isoforms of ascorbate peroxidase were downregulated from 12 to 18 DAF, and four of the five identified glyoxalase I isoforms showed opposite expression patterns. In addition, two-thirds of the 317 DEPs are potential targets of thioredoxin, 70% of which function in metabolism (glycolysis, TCA cycle, alcoholic fermentation, starch metabolism, and amino acid metabolism), protein synthesis/destination (protein synthesis and protein folding), cell structure, and cell growth/division. Taken together, this research reveals that the ROS-induced redox-mediated pathway is activated in rice endosperms, playing essential roles in the regulation and coordination of starch accumulation and PCD during seed development.

The previous two studies have shown the proteome changes during different developmental days; however, whether and how the developing rice endosperm regulates its proteome to response to the light-dark rhythm in the same day is still unclear. Therefore, we further compared the proteomes of 10-day-old endosperms at the light phase, the light-dark transition phase, and the dark phase by use of 2D-DIGE and MS, leading to identification of 91 proteins with significant changes in expression [33]. The 91 DEPs showed different light-dark regulation patterns; however, proteins belonging to a certain functional category or subcategory displayed similar diurnally changed expression patterns. Enzyme proteins participating in the conversion of leaf-importing sucroses to endosperm starches, such as UDP-Glc pyrophosphorylase, pullulanase, and ADP-glucose pyrophosphorylase large chain 2 (a subunit of the first key regulatory enzyme in the starch biosynthetic pathway), were downregulated in the light cycle but upregulated in the dark cycle. This diurnal-regulated expression of proteins related to starch synthesis was consistent with the following two cytological observations. First, in developing endosperms, the levels of sucrose, as opposed to glucose or fructose, were highest at the beginning of the light cycle, lowered to a plateau in the middle of the light cycle, stable at the plateau the light-dark transition, and reduced rapidly in the dark cycle. Second, growth rings, representing alternating amorphous and semicrystalline regions of starch granules, were visible in endosperm grown under alternating light-dark cycles but invisible in endosperm grown under constant light conditions. These results strongly suggest that endosperm starches are mainly synthesized and accumulated in the dark phase during rice seed development. Furthermore, proteins involved in cell division, such as cell division control protein 48 homolog E and two isoforms of DNA repair protein RAD23, and those implicated in protein synthesis, such as elongation factor 1-beta and 40S ribosomal protein S12, showed increased accumulation in the light phase, with a peak at the light-dark switch and a decrease in the dark phase. ROS-scavenging proteins, including glyoxalase I, superoxide dismutase (SOD) 1 [Cu-Zn], chloroplastic SOD [Cu-Zn] isoform, peroxiredoxin-2E-1, and glutaredoxin-C6, had one expression peak in the light phase and the other peak in the dark phase; mitochondrial SOD [Mn] showed one expression peak in the dark phase. Similarly, most proteins related to protein folding and proteolysis, amino acid metabolism, and TCA cycle showed two accumulation peaks, similar to ROS-scavenging proteins. Therefore, the circadian clock control of starch accumulation requires the coordination of diverse cellular and molecular processes that include cell division and enlargement, redox regulation, protein

synthesis, folding, and proteolysis, TCA cycle, and amino acid metabolism.

In 2012, two groups published their proteomic studies of embryo development. Wang's group analyzed the difference in embryo proteomes among five developmental stages, namely first leaf primordium formation (5 DAF), second and third leaf primordium formation (7 DAF), morphological completion (14 DAF), maturation (21 DAF), and desiccation (30 DAF), resulting in identification of 275 DEPs [28]. Liu's group compared the proteomic profiles of embryos at three developmental phases that approximately correspond to the three middle stages detected by Wang's group, namely the morphologically changing stage (6 DAF) and two morphologically quiescent stages (12 and 18 DAF); and finally they identified 53 protein spots showing difference in expression levels [29]. Analogous to the proteome of endosperms [16], the largest functional categories of the DEPs in the embryo proteome of different developmental stages were metabolism (30.5%) and protein synthesis/destination (17.8%); however, proteins involved in starch synthesis were much less abundant in the embryo than in the endosperm [28], which suggests that the regulation of metabolism and protein synthesis is necessary and essential not only for endosperm growth and starch accumulation but also for embryo development. The changes in expression of different proteins were closely associated with the cellular, molecular, and metabolic events occurring during embryo development [28, 29], similar to the situation in the dynamic proteomes of developing endosperms [16] and whole seeds [32]. The two studies of rice embryos identified most enzymes involved in the glycolytic pathway, which displayed high-level expression at early developmental stages, with a peak at 7 DAF, and decreased levels afterwards [28, 29]. Proteins related to the TCA cycle, such as aconitase, malate dehydrogenase, and succinate dehydrogenase were upregulated until the peak level was attained at 14 DAF and were downregulated at later developmental stages [28]. Moreover, proteins of the above two metabolic pathways were expressed at the lowest level at 30 DAF [28]. This expression pattern of proteins involved in central carbon metabolism (glycolysis and TCA cycle) has also been observed in the proteomic study of rice whole seeds [32], which suggests that high metabolic activity maintained by glycolysis and TCA cycle is important for early embryo and endosperm developmental events such as cell division and differentiation, but reduction in metabolic activity at late developmental stages might be required for embryo maturation and dormancy as well as starch accumulation in the endosperm. Late-embryogenesis-abundant proteins (LEAs), including LEA D-34, LEA group 1, LEA group 3, and LEA domain-containing proteins, showed peak expression at 14 or 21 DAF, which is consistent with their roles in protecting embryo cells against water-deficit stress during seed desiccation [28]. In addition to LEAs, globulin might offer a similar protection function. Liu et al. noted that more than half of the identified differential spots (38 of 53) corresponded to nine rice embryo globulins (REG) or their isoforms with different molecular weight and isoelectric point. Among them, 13 spots were identified as globulin Os03g46100 (REG 100), 11 spots were Os03g57960 (REG 960), and ten spots were Os08g03410 (REG 410) [29]. During rice embryo development from 6 to 18 DAF, the three major globulins, namely REG 100, REG 960, and REG 410, showed an increase in the number of isoforms and most of these isoforms showed upregulated expression [29]. Further investigation demonstrated that globulin isoforms with different molecular weight were deliberately and orderly produced by embryogenesis-controlled proteolysis, and the degraded isoforms were associated with protein complexes during embryo development, although the three major globulins contributed differently to the formation of protein complexes [29]. Thus, these results suggest that the cleaved REG isoforms might protect embryo proteins by formatting protein complexes with protected proteins during embryo maturation and desiccation.

4.2 Proteomic studies of stress response and tolerance in rice seeds

The development of rice seeds is not only controlled by hereditary but also shaped by environment. Environmental factors such as temperature and water influence the expression of specific proteins, thus affecting the metabolism of carbohydrates and proteins and thereby the yield and quality of rice seeds. The discovery and identification of proteins responding to environmental stresses are the first step toward understanding the relationship between growth environment and seed development, which is notably valuable for yield and quality improvement of rice by biological engineering [43]. Recently, several comparative proteomic investigations have used 2DE combined with MS and bioinformatics to systematically characterize rice seed proteins responsive to abiotic environmental stresses such as high temperature, drought, cold, and high salinity [25, 26, 30].

Through comparing the proteome of developing rice seeds grown under normal environmental conditions (control) and under high temperature conditions, two research groups analyzed the effect of high temperature on seed proteins related to rice grain quality and on those involved in stress response [25, 26]. Lur's group identified 54 DEPs responding to continuous high temperature during seed development [26]. Among them, 21 and 14 proteins were associated with carbohydrate and protein metabolism, respectively, nine proteins were involved in stress response, nine proteins played other diverse roles, and one protein had unknown function (NCBI accession number: AU082974). Zeng's group was more concerned with the different effects resulting from day and night high temperature (DHT and NHT) treatments. They identified 61 proteins differentially expressed in treated seeds as compared with controls, including 23 implicated in carbohydrate metabolism, seven in protein metabolism, ten in stress tolerance, four in signal

transduction, 12 in other cellular processes, and five with unknown function [25]. When the 61 identified proteins were compared between DHT and NHT treatments, nine proteins were less abundant under NHT than DHT at the early stage of seed development (5 DAF) but more abundant at the middle (10 and 15 DAF) or late (20 DAF) stages; seven proteins showed opposite accumulation pattern to these nine proteins; 15 proteins were upregulated under NHT only at 1–3 stages; 18 proteins were downregulated under NHT only at 1 or 2 stages; 12 proteins showed similar expression levels at every stage [25].

Both groups found rice quality-related proteins from the identified DEPs involved in carbohydrate or protein metabolism, providing possible molecular mechanisms for the negative effects of high temperature stress on rice quality. The cooking and taste quality of rice is largely influenced by the ratio of the two types of starch, amylose, and amylopectin. Granule-bound starch synthase (GBSS) and pullulanase (a type of starch-debranching enzyme) are the key enzymes for the biosynthesis of amylose [44] and amylopectin [45], respectively. Six isoforms of GBSS were detected in developing rice seeds, but only the high-molecular-weight isoforms showed an accumulation pattern positively correlated with amylose content, which suggests that these large isoforms are the major enzymes involved in amylose biosynthesis [26]. Under high temperature, GBSS expression was decreased at the translational level instead of the transcriptional level [26]. In contrast, the five pullulanase isoforms in developing rice seeds showed upregulated expression after high temperature treatment for 5, 10, and 15 days, although they differed in their responses to DHT and NHT [25]. The downregulation of GBSS and the upregulation of pullulanase caused by high temperature treatment may lead to reduced amylose content and increased amylopectin content, respectively, in seed starches, thus resulting in poor taste of edible rice endosperms. Furthermore, high temperature suppressed the accumulation of four cytosolic isoforms of pyruvate orthophosphate dikinase B (PPDKB), and three isoforms showed higher repression under DHT than NHT at 5 and 10 DAF, which is consistent with the rice grown under DHT, NHT, and normal conditions displaying orderly decreasing chalkiness in mature seeds [25]. These results, combined with the observation that mutation in one of the PPDKB isoforms caused the formation of endosperms with a white core [46], indicate that high temperature downregulates PPDKB, which in turn leads to seed chalkiness, thus ultimately affecting the appearance quality of rice grain. Besides PPDKB, the functionally unknown protein (AU082974) with 74% GC content in DNA sequence, might also be related to chalkiness induced by high temperature, because this protein was downregulated by high temperature and showed lower abundance in chalky parts than in translucent parts of rice seeds [26]. Moreover, Lur's group revealed that one type of storage protein, glutelin, was phosphorylated and glycosylated during seed development and another type of storage protein, prolamine, showed increased accumulation under high temperature [26]. These observations provide new insight into factors influencing the nutritional quality of rice grains.

The two studies also identified a number of differentially expressed stress-responsive proteins, most of which were upregulated by high temperature treatment. Examples included HSP 70, 3 low-molecular-weight HSPs, 2-Cys peroxiredoxin, two isoforms of 1-Cys peroxiredoxin, and Cu/Zn superoxide dismutase [25, 26]. Upregulation of these stress-responsive proteins is consistent with their critical roles in protecting seed proteins, nucleic acids, or RNAs against thermal damage. However, several stress-responsive proteins were downregulated or showed complex regulation under high temperature. For instance, DHT and NHT treatments decreased the expression of two glyoxalases I isoforms at 5 and 10 DAF but increased their expression at 15 and 20 DAF [25]. Ascorbate peroxidases displayed lowered accumulation under both NHT and DHT. The expression of thioredoxin H-type was suppressed by DHT but enhanced by NHT. Thus, stress tolerance in rice seeds may be achieved by more complex mechanisms besides upregulation of a few stress-responsive proteins.

In addition to the above two researches, Pagès et al. designed a distinct strategy to study stress tolerance in rice seeds. They compared mature embryo proteomes among three varieties sensitive to stress (drought, salt, and cold) and three tolerant to stress and identified 28 DEPs: 16 implicated in the metabolism of reserve proteins or starches, six in stress defense or tolerance, and six in other cellular processes [30]. Among the 16 proteins involved in reserve metabolism, the glutelin type-B 2 precursor was much more abundant in the embryo proteome of the drought-tolerant variety than in that of the drought-sensitive variety, suggesting this protein might possibly act as a protector against drought stress besides being a type of storage protein. The six stress-defense proteins represented four LEA groups, namely, group 1 (embryonic abundant protein 1), dehydrins (three isoforms of the water-stress inducible protein Rab21, with different isoelectric point), group 3 (a putative embryonic abundant protein similar to AtLEA3–1), and group 7 (a putative LEA similar to AtLEA7–1). The accumulation of the six LEA proteins was highly variable among the six rice varieties, not matching the expected results that stress-defense proteins should show consistent upregulation in the stress-tolerant varieties as compared with the stress-sensitive ones. However, the three identified Rab21 isoforms as well as the other four Rab21 isoforms visible on the 2DE gels were less phosphorylated in the stress-tolerant varieties than in the sensitive ones. In addition, the embryo proteomes of the three stress-tolerant varieties contained an unphosphorylated Rad21 isoform with an approximate isoelectric point (p*I*) of 6.8, which was invisible on the 2DE gels but detected by Western blot analysis. These results clearly indicate that posttranslational modification such as phosphorylation, rather than regulation of gene expression at the translational level, may contribute more to stress tolerance in rice seeds.

4.3 Proteomic studies of rice seed mutants

Comparison of proteome between WT and genetic mutants is an alternative and powerful strategy for screening proteins affected by gene mutations. Komatsu et al. applied 2DE to analyze differences in seed proteomes between the WT and rice *dwarf1* (*d1*) mutants, which showed multifaceted phenotypes (including dwarfism, small and round seeds, and dark green leaves) caused by mutations in the α subunit of a heterotrimeric G protein [27]. Compared to WT, the *d1* mutants exhibited no significant differences in the endosperm proteome. However, seven embryo proteins, namely 1 rice receptor for activated C-kinase (RACK) and 6 rice embryo globulin-2s (REG2s), showed reduced expression. In the WT, RACK and the 6 REG2 proteins were accumulated during embryo development and maturation. RACK initiated its degradation from the early stage of seed maturation and REG2s began to degrade until seed germination. Reduction in level of RACK and REG2s in the *d1* mutant occurred throughout seed maturation and even during seed germination. Treatment with abscisic acid during seed germination increased the protein level of RACK in WT rather than the *d1* mutant. These results show the linkage between the G-protein subunit and its mutation-affected proteins (RACK and REG2s) and suggest that these proteins are involved in embryo development. This study demonstrates that through comparing seed proteomes between WT and mutants defective in seed phenotypes, comparative proteomics can become an effective approach for identification of new proteins associated with certain molecular pathways in seed development.

4.4 Subproteomic studies of mature and developing rice seeds

Subproteomics focuses on investigating proteomes from a certain tissue or organelle (spatial subproteome) or a subset of proteins with specific structural or functional properties (functional subproteome) [47]. Subproteomic analysis involves specific fractionation and enrichment procedures before or during protein extraction [47], thus enabling proteomic identification of low abundance but often biologically important proteins. Over the past decade, subproteomics has been increasingly used to address biological questions in seed development; however, related studies of rice seeds are limited because of the challenges in establishing respective and effective purification and enrichment methods for different samples and applications, as well as the difficulties in removing the interference of highly abundant storage starches and proteins. In this section, we review recent progress in rice seed subproteomics, including one study of nuclear proteins in developing seeds [35] and three studies of RBPs, one in mature dry seeds [36] and the other two in developing seeds [37, 38].

Genetic and molecular studies have shown that nuclear proteins such as OsCCS52A [48], Orysa; KRP3 [49], gw5 [50], MEA/FIS1, FIS2, and FIE/FIS3 [51] play important roles in endoreduplication, syncytium, cell division, and epigenetic regulation during rice embryo and endosperm development. Therefore, cataloguing nuclear proteins and establishing their posttranslational modification map in developing seeds are highly useful for elucidating mechanisms underlying the molecular regulation of seed development. A major challenge for nuclear subproteome study in rice seeds is to deplete highly abundant starches and storage proteins. Peng et al. removed most starch granules and SSPs from nuclear protein extracts of developing rice seeds simply by cotton filtration and gel excision and recovered the remaining proteins by phenol extraction (detailed in Section 3.1) [35]. Through this nuclei purification, and nuclear protein enrichment and recovery procedure, they identified 468 proteins, including 220 nuclear-localized proteins, from the recovered nuclear fractions. In contrast, only 28 proteins, including 24 storage proteins, were identified from the total protein extracts of rice seeds at the same developmental stage. The 220 nuclear-localized proteins contained not only chromosome structural proteins (e.g. all the 24 histones and their variants) but also low-abundance nuclear proteins such as X1, X2, bZIP, and bHLH transcription factors. These nuclear proteins might be involved in various biological processes important for seed development such as DNA metabolism (29.4%), protein metabolism (19.3%), transcription (19%), cell organization, and biogenesis (15.2%), cell death (0.5%), and cell cycle (0.5%). This nuclear proteome of developing rice seeds also contained 208 hypothetical proteins, among which 86 had nuclear localization signals, which indicates a large number of unknown knowledge about nuclear proteins involved in seed development. In addition, Peng et al. analyzed the MS/MS mass spectral data with Bioworks software to characterize arginine and lysine methylation and acetylation of these seed nuclear proteins and found higher than expected frequency of modification. In all, 59 proteins contained acetylated amino acids and 40 had methylated residues. Among them, two proteins (Brittle-1 protein and Glutelin type-A 3 precursor) were both acetylated and methylated. In this subproteomic study, Peng's group provided the global characterization of nuclear proteins implicated in rice seed development and gave new information about the posttranslational methylation and acetylation modification of these seed nuclear proteins, thus offering many new targets for genetic and molecular studies in seed development and its regulation.

RBPs are a large class of nucleic acid-binding proteins that have one or more RNA-binding domains such as the widely spreading RNA recognition motif (RRM) and the K-homology (KH) domain [52]. Extensive studies in yeast and metazoans have demonstrated that RBPs perform diverse molecular functions in all aspects of RNA metabolism, including RNA exportation from the nucleus to the cytoplasm, mRNA expression and regulation events that occur in the nucleus (e.g. transcription, splicing, and polyadenylation) or in the cytoplasm (e.g. localization, storage, translation, and degradation) [52]. Due to the absence of plant derived in vitro

experimental systems for studying RNA metabolism [52] as well as the presence of numerous plant-specific RBPs lacking yeast and animal homologues, considerably less information has been accumulated on plant RBPs, especially those involved in seed development and maturation. However, the nucleic acid-binding activity of RBPs facilitates the purification and enrichment of RBPs from crude protein extracts, which led to the identification of 11 RBPs in mature rice seeds [36] and 257 RBPs in developing rice seeds [37–39] through proteomic approaches.

Using one-step affinity chromatography with denatured ssDNA cellulose column after extraction of total proteins from mature and dry rice seeds, Kanekatsu et al. effectively removed high-abundance storage proteins and obtained an ssDNA-binding fraction [36]. Although the protein recovery yield was low by use of this affinity purification method and only 18 proteins were identified from the ssDNA-binding fraction after 2DE separation and MS analysis, 11 of the identified proteins were RBPs, including one glycine-rich protein, three ribosomal proteins, three proteins with RRM, and four proteins with KH domains. By comparing the quantity of the 18 ssDNA-binding proteins in dry seeds and geminated seeds, the authors found that during seed germination, the three RRM-containing RBPs, the glycine-rich RBP and the two non-RBPs with ssDNA-binding motifs showed greatly reduced expression, whereas the three KH domain-containing RBPs showed almost no change in expression. Inhibition of seed germination by abscisic acid or cold treatment suppressed the downregulated expression of the three RRM RBPs and the glycine-rich RBP rather than the two non-RBPs. Therefore, seed germination was accompanied by the degradation of the RRM-containing RBPs and the glycine-rich RBP, which suggests that these RBPs may stabilize long-lived mRNAs stored in mature dry rice seeds for rapid protein synthesis at the beginning of seed germination.

In addition to the above investigation on mature seed RBPs, Okita's group has conducted a large amount of research on RBPs implicated in RNA sorting within the cytoplasm during rice seed development. Through genetic, cytological, and molecular approaches, the group found that RBPs in developing rice seeds, such as OsTudor-SN [53, 54], can recognize and bind the cis-acting localization elements (or zipcodes) within the mRNA sequence of storage proteins, forming large macromolecular complexes. These protein-RNA complexes are transported from the nucleus to the cytoplasm along the cell cytoskeleton and then are targeted and anchored properly to the cortical ER where storage proteins are synthesized. These findings demonstrate RBPs participate in regulating the biosynthesis of storage proteins during seed development, by controlling multiple molecular events occurring during mRNA transportation and localization to the ER. However, considering that mRNA sorting is a multistep process and the two major types of rice SSPs, namely, glutelins and prolamines, are localized to distinct ER subdomains, a large number of RBPs should be involved besides the few identified ones. To identify RBPs involved in mRNA transportation and localization during seed development, Okita et al. first searched for proteins able to bind cytoskeleton as well as nucleic acids in developing rice seeds by combining poly(U)-sepharose affinity purification with 2DE and MS, resulting in the identification of 148 putative cytosolic RBPs (see details in Section 3.1) [37]. The 148 proteins included the known RBPs involved in storage protein RNA sorting, such as OsTudor-SN, and 20 RNA-binding domain-containing proteins (i.e. nucleolar protein NOP5 and DEAD box RNA helicase [also identified in the nuclear subproteomic study], KH domain-containing protein LOC_Os02g57640 [also identified in the rice dry seed RBP study] and 17 newly identified proteins) that are the most probable RBP candidates for involvement in RNA metabolism. Other proteins were speculated to play roles mainly in protein translation (38 proteins), protein processing, turnover, and transport (23 proteins), cell structure and energy production (15 proteins), signaling and stress response (16 proteins), carbohydrate or lipid metabolism (22 proteins) and other processes.

To further identify RBPs involved in transportation and localization of prolamine mRNAs, Okita et al utilized the zipcode sequence in the prolamine 5'CDS region to capture proteins bound with prolamine mRNAs from the cytoskeleton-enriched fractions of developing rice seeds [38]. A highly charged polyanionic molecule, heparin, was added in all capture experiments at a high concentration to prevent nonspecific binding. Via this purification procedure, 18 distinct proteins were identified after 1DE separation and LC–MS/MS analysis. Among them, 15 proteins, which included three members of the heterogeneous nuclear ribonucleoprotein family (hnRNPs, NCBI accession number: AK105751, AK059225, AK100904), bound specifically to the prolamine zipcode sequence. The remaining three proteins had general RNA-binding activities, as shown by their association with both the 5'CDS zipcode sequence and the control RNA sequence present in the nonzipcode region of the prolamine open reading frame. Further molecular analysis provided strong evidence that one of the three hnRNPs, namely AK105751, may be involved in prolamine and glutelin mRNA exportation from the nucleus and subsequent localization to the ER, because this hnRNP (i) can bind with prolamine and glutelin mRNA in vivo, (ii) had a protein expression pattern similar to that of prolamine and glutelin, and (iii) localized not only in the nucleus but also on the cisternal ER membranes and cell cytoskeleton. Furthermore, because of the observation that heparin inhibited the binding of OsTudor-SN (the characterized prolamine and glutelin RBP) to its target RNAs, the prolamine zipcode capture assays were also performed without heparin, which revealed 132 putative prolamine RBPs [38, 39]. Among them, 77 proteins were newly identified RBP candidates, except for the 12 proteins found in the stringent prolamine zipcode capture experiment and the 43 proteins obtained by poly(U)-sepharose affinity purification [39]. In conclusion, Okita's group obtained 148 cytoskeleton-associated RBPs and 138 cytoskeleton-associated prolamine mRNA-binding proteins

from developing rice seeds through isolation with affinity chromatography followed by identification with proteomic approaches. They have documented these RBPs and a variety of related information in the RiceRBP website (http://www.bioinformatics2. wsu. edu/RiceRBP), which provides an elaborate and comprehensive database of RBPs accessible to the public [55].

5 Concluding remarks

In the past decade, proteomic studies have provided many new and detailed insights into the mechanisms of metabolic and molecular control during embryo and endosperm development in rice, mostly via combination of protein fractionation, 2DE separation and MS-based identification. Global comparison of proteome changes among different seed developmental stages results in revealing new aspects of the regulation mechanisms underlying seed development. Although genome-wide transcriptional analysis of developing rice seeds have demonstrated the coordinate regulation of expression of genes associated with seed growth and reserve accumulation during seed development [56], there remain significant gaps in our understanding of this coordination. At present, large-scale analysis of proteome changes associated with rice seed development show that: (i) endosperm growth, embryogenesis, and starch accumulation involve the coordination of various metabolic, molecular, and cellular processes, including the switch from central carbon metabolism to alcoholic fermentation, ROS-induced redox regulation of PCD, diurnal light-dark control of starch accumulation and cell growth; (ii) the expression pattern of proteins involved in specific metabolic, molecular, or cellular processes is relatively consistent during embryo and endosperm development, which is highly coordinated with the timing and sequence of seed developmental events (Tables 2 and 3). Comparing seed proteomes between stress-treated and -untreated rice or between stress-tolerant and -sensitive rice varieties provides new information about the proteins and mechanisms related to the effects of environmental stimulus on rice quality and those related to stress response and tolerance. Molecular studies have shown that the reduction of rice taste quality by high temperature might result from modification of the starch structure by the expressional regulation of enzymes related to starch synthesis, such as GBSS [57]. Proteomic studies further reveal that although these starch synthesis-related enzymes may have several isoforms, some but not all are involved in high temperature-induced reduction of rice grain quality. For example, six isoforms of GBSS are present in the proteome of developing rice seeds, but only the high-molecular-weight isoforms of GBSS are downregulated by high temperature. Stress-responsive proteins such as HSPs and LEAs are known to play protective roles under environmental stresses, and upregulation of their expression is an important mechanism for stress tolerance in plants [43]. Comparison of embryo proteomes in stress-tolerant and -sensitive rice varieties offers new and direct evidence that stress tolerance in rice seeds can result from not only upregulation but also posttranslational modification (e.g. phosphorylation) of stress-responsive proteins. When 2DE is used as a tool to compare WT and mutant proteomes or to analyze spatial and functional subproteomes, proteomics can reveal novel proteins involved in certain molecular and/or cellular processes, which helps resolve complex biological questions related to rice seed development. The typical example of proteomic comparison between WT and mutants is the analysis of differences between WT and *d1* seed proteomes, which led to the identification of seven new proteins (i.e. RACK and six REG2s) involved in G-protein-mediated pathways during embryo development. The localization and functional domain of a protein are closely related to its function. Identification and functional analysis of proteins with specific localization or domains can provide much information for understanding the molecular mechanisms involved in seed development. Compared with the low-efficient single-protein study by genetic and molecular approaches, proteomics provides a possible strategy for large-scale identification and analysis of proteins containing specific functional domains or proteins located in specific tissues, organelles, or regions of rice seeds. However, these proteins are often accumulated at lower levels in rice seeds, causing them invisible in 2DE gels if crude protein extracts are used. Innovative improvement in protein fractionation and sample preparation has led to the isolation and identification of low-abundance seed proteins, including 220 nuclear-localized proteins and 268 proteins containing RNA-binding domains, offering a large number of candidate proteins that may play specific and important roles in rice seed development.

In the near future, the proteomic study of rice seed development will become more exciting with the technological and knowledge advances. The extensive use of the advanced proteomic technologies such as DIGE and iTRAQ 2D-LC-MS/MS and the improvement in technologies themselves will offer increased accuracy and sensitivity in protein identification and quantification. Further improvement in protein fractionation and purification techniques for subproteomic studies will result in the large-scale identification of more low abundance but biologically important proteins. MALDI-imaging MS, an emerging technique enabling MS imaging of biological molecules in plant tissue sections [58, 59], can be used to analyze the spatial distribution patterns of metabolites and proteins in developing rice seeds. The application of MALDI-imaging MS will lead to the identification of protein markers expressed specifically in certain organs, cell types, or regions of rice seeds at certain developmental stages, as well as provide new insights into the metabolic and regulatory networks in seed development. Combining proteomic tools with genetic, biochemical, cytological, and molecular approaches will allow us to obtain biologically relevant answers regarding the control mechanisms of embryo and endosperm development in rice. Benefiting from these advances, proteomics will predictably improve and update our knowledge of

mechanisms underlying rice seed development and its regulation, providing helpful information for molecular breeding and improvement of quality, yield, and stress tolerance in rice and other cereals.

This work was supported by grants from the National Science Foundation of China (grant no. 31000694) and the Chinese Ministry of Science and Technology (grant no. 2012CB910504 and 2011CB100202).

The authors have declared no conflict of interest.

6 References

[1] Khush, G. S., What it will take to feed 5.0 billion rice consumers in 2030. *Plant Mol. Biol.* 2005, *59*, 1–6.

[2] Itoh, J.-I., Nonomura, K.-I., Ikeda, K., Yamaki, S. et al., Rice plant development: from zygote to spikelet. *Plant Cell Physiol.* 2005, *46*, 23–47.

[3] Sato, Y., Hong, S. K., Tagiri, A., Kitano, H. et al., A rice homeobox gene, OSH1, is expressed before organ differentiation in a specific region during early embryogenesis. *Proc. Natl. Acad. Sci. USA* 1996, *93*, 8117–8122.

[4] Cooper, B., Clarke, J. D., Budworth, P., Kreps, J. et al., A network of rice genes associated with stress response and seed development. *Proc. Natl. Acad. Sci. USA* 2003, *100*, 4945–4950.

[5] Peck, S. C., Update on proteomics in Arabidopsis. Where do we go from here? *Plant Physiol.* 2005, *138*, 591–599.

[6] Sabelli, P. A., Larkins, B. A., The development of endosperm in grasses. *Plant Physiol.* 2009, *149*, 14–26.

[7] Hoshikawa, K., Studies on the development of endosperm in rice: 1. process of endosperm tissue formation. *Jpn. J. Crop Sci.* 1967, *36*, 151–161.

[8] Kitano, H., Tamura, Y., Satoh, H., Nagato, Y., Hierarchical regulation of organ differentiation during embryogenesis in rice. *Plant J.* 1993, *3*, 607–610.

[9] Yoshida, K. T., Wada, T., Koyama, H., Mizobuchi-Fukuoka, R., Naito, S., Temporal and spatial patterns of accumulation of the transcript of myo-inositol-1-phosphate synthase and phytin-containing particles during seed development in rice. *Plant Physiol.* 1999, *119*, 65–72.

[10] Ramachandran, C., Raghavan, V., Changes in nuclear DNA content of endosperm cells during grain development in rice (*Oryza sativa*). *Ann. Bot.* 1989, *64*, 459–468.

[11] Ishimaru, T., Matsuda, T., Ohsugi, R., Yamagishi, T., Morphological development of rice caryopses located at the different positions in a panicle from early to middle stage of grain filling. *Funct. Plant Biol.* 2003, *30*, 1139–1149.

[12] Hoshikawa, K., Studies on the development of endosperm in rice: 4. differentiation and development of the aleuron layer. *Jpn. J. Crop Sci.* 1967, *36*, 216–220.

[13] Hoshikawa, K., Studies on the development of endosperm in rice: 2. process of endosperm tissue formation with special reference to the enlargement of cells. *Jpn. J. Crop Sci.* 1967, *36*, 203–209.

[14] Hoshikawa, K., Studies on the development of endosperm in rice: 10. electron microscopic studies on the development of starch granules in the endosperm cells. *Jpn. J. Crop Sci.* 1968, *37*, 97–106.

[15] Hoshikawa, K., Studies on the development of endosperm in rice: 11. development of starch granules in endosperm tissue. *Jpn. J. Crop Sci.* 1968, *37*, 207–216.

[16] Xu, S. B., Yu, H. T., Yan, L. F., Wang, T., Integrated proteomic and cytological study of rice endosperms at the storage phase. *J. Proteome Res.* 2010, *9*, 4906–4918.

[17] Rakwal, R., Agrawal, G. K., Rice proteomics: current status and future perspectives. *Electrophoresis* 2003, *24*, 3378–3389.

[18] Agrawal, G. K., Jwa, N. S., Iwahashi, Y., Yonekura, M. et al., Rejuvenating rice proteomics: facts, challenges, and visions. *Proteomics* 2006, *6*, 5549–5576.

[19] Agrawal, G. K., Rakwal, R., Rice proteomics: a cornerstone for cereal food crop proteomes. *Mass Spectrom. Rev.* 2006, *25*, 1–53.

[20] Agrawal, G. K., Jwa, N. S., Rakwal, R., Rice proteomics: ending phase I and the beginning of phase II. *Proteomics* 2009, *9*, 935–963.

[21] Krishnan, H. B., White, J. A., Morphometric analysis of rice seed protein bodies: implication for a significant contribution of prolamine to the total protein content of rice endosperm. *Plant Physiol.* 1995, *109*, 1491–1495.

[22] Villareal, R. M., Juliano, B. O., Properties of glutelin from mature and developing rice grain. *Phytochemistry* 1978, *17*, 177–182.

[23] Houston, D. F., Iwasaki, T., Mohammad, A., Chen, L., Radial distribution of protein by solubility classes in the milled rice kernel. *J. Agric. Food Chem.* 1968, *16*, 720–724.

[24] Yamagata, H., Sugimoto, T., Tanaka, K., Kasai, Z., Biosynthesis of storage proteins in developing rice seeds. *Plant Physiol.* 1982, *70*, 1094–1100.

[25] Li, H., Chen, Z., Hu, M., Wang, Z. et al., Different effects of night versus day high temperature on rice quality and accumulation profiling of rice grain proteins during grain filling. *Plant Cell Rep.* 2011, *30*, 1641–1659.

[26] Lin, S.-K., Chang, M.-C., Tsai, Y.-G., Lur, H.-S., Proteomic analysis of the expression of proteins related to rice quality during caryopsis development and the effect of high temperature on expression. *Proteomics* 2005, *5*, 2140–2156.

[27] Komatsu, S., Abbasi, F., Kobori, E., Fujisawa, Y. et al., Proteomic analysis of rice embryo: an approach for investigating Gα protein-regulated proteins. *Proteomics* 2005, *5*, 3932–3941.

[28] Xu, H., Zhang, W., Gao, Y., Zhao, Y. et al., Proteomic analysis of embryo development in rice (*Oryza sativa*). *Planta* 2012, *235*, 687–701.

[29] Zi, J., Zhang, J., Wang, Q., Lin, L. et al., Proteomics study of rice embryogenesis: discovery of the embryogenesis-dependent globulins. *Electrophoresis* 2012, *33*, 1129–1138.

[30] Farinha, A. P., Irar, S., de Oliveira, E., Oliveira, M. M., Pagès, M., Novel clues on abiotic stress tolerance emerge from embryo proteome analyses of rice varieties with contrasting stress adaptation. *Proteomics* 2011, *11*, 2389–2405.

[31] Gong, C. Y., Li, Q., Yu, H. T., Wang, Z., Wang, T., Proteomics insight into the biological safety of transgenic modification of rice as compared with conventional genetic breeding and spontaneous genotypic variation. *J. Proteome Res.* 2012, *11*, 3019–3029.

[32] Xu, S. B., Li, T., Deng, Z. Y., Chong, K. et al., Dynamic proteomic analysis reveals a switch between central carbon metabolism and alcoholic fermentation in rice filling grains. *Plant Physiol.* 2008, *148*, 908–925.

[33] Yu, H. T., Xu, S. B., Zheng, C. H., Wang, T., Comparative proteomic study reveals the involvement of diurnal cycle in cell division, enlargement, and starch accumulation in developing endosperm of *Oryza sativa*. *J. Proteome Res.* 2012, *11*, 359–371.

[34] Miernyk, J. A., Hajduch, M., Seed proteomics. *J. Proteomics* 2011, *74*, 389–400.

[35] Li, G., Nallamilli, B. R. R., Tan, F., Peng, Z., Removal of high-abundance proteins for nuclear subproteome studies in rice (*Oryza sativa*) endosperm. *Electrophoresis* 2008, *29*, 604–617.

[36] Masaki, S., Yamada, T., Hirasawa, T., Todaka, D., Kanekatsu, M., Proteomic analysis of RNA-binding proteins in dry seeds of rice after fractionation by ssDNA affinity column chromatography. *Biotechnol. Lett.* 2008, *30*, 955–960.

[37] Doroshenk, K. A., Crofts, A. J., Morris, R. T., Wyrick, J. J., Okita, T. W., Proteomic analysis of cytoskeleton-associated RNA binding proteins in developing rice seed. *J. Proteome Res.* 2009, *8*, 4641–4653.

[38] Crofts, A., Crofts, N., Whitelegge, J., Okita, T., Isolation and identification of cytoskeleton-associated prolamine mRNA binding proteins from developing rice seeds. *Planta* 2010, *231*, 1261–1276.

[39] Doroshenk, K. A., Crofts, A. J., Morris, R. T., Wyrick, J. J., Okita, T. W., RiceRBP: a resource for experimentally identified RNA binding proteins in *Oryza sativa*. *Front. Plant Sci.* 2012, *3*, 90.

[40] Xie, Z., Wang, J., Cao, M., Zhao, C. et al., Pedigree analysis of an elite rice hybrid using proteomic approach. *Proteomics* 2006, *6*, 474–486.

[41] Wang, W., Meng, B., Ge, X., Song, S. et al., Proteomic profiling of rice embryos from a hybrid rice cultivar and its parental lines. *Proteomics* 2008, *8*, 4808–4821.

[42] Buchanan, B. B., Balmer, Y., Redox regulation: a broadening horizon. *Annu. Rev. Plant Biol.* 2005, *56*, 187–220.

[43] Zou, J., Liu, C., Chen, X., Proteomics of rice in response to heat stress and advances in genetic engineering for heat tolerance in rice. *Plant Cell Rep.* 2011, *30*, 2155–2165.

[44] Sano, Y., Differential regulation of waxy gene expression in rice endosperm. *Theor. Appl. Genet.* 1984, *68*, 467–473.

[45] Kubo, A., Fujita, N., Harada, K., Matsuda, T. et al., The starch-debranching enzymes isoamylase and pullulanase are both involved in amylopectin biosynthesis in rice endosperm. *Plant Physiol.* 1999, *121*, 399–410.

[46] Kang, H.-G., Park, S., Matsuoka, M., An, G., White-core endosperm floury endosperm-4 in rice is generated by knockout mutations in the C4-type pyruvate orthophosphate dikinase gene (OsPPDKB). *Plant J.* 2005, *42*, 901–911.

[47] Finnie, C., Sultan, A., Grasser, K. D., From protein catalogues towards targeted proteomics approaches in cereal grains. *Phytochemistry* 2011, *72*, 1145–1153.

[48] Su'udi, M., Cha, J.-Y., Jung, M., Ermawati, N. et al., Potential role of the rice OsCCS52A gene in endoreduplication. *Planta* 2012, *235*, 387–397.

[49] Mizutani, M., Naganuma, T., Tsutsumi, K., Saitoh, Y., The syncytium-specific expression of the Orysa;KRP3 CDK inhibitor: implication of its involvement in the cell cycle control in the rice (*Oryza sativa* L.) syncytial endosperm. *J. Exp. Bot.* 2010, *61*, 791–798.

[50] Weng, J., Gu, S., Wan, X., Gao, H. et al., Isolation and initial characterization of GW5, a major QTL associated with rice grain width and weight. *Cell Res.* 2008, *18*, 1199–1209.

[51] Luo, M., Bilodeau, P., Dennis, E. S., Peacock, W. J., Chaudhury, A., Expression and parent-of-origin effects for FIS2, MEA, and FIE in the endosperm and embryo of developing Arabidopsis seeds. *Proc. Natl. Acad. Sci. USA* 2000, *97*, 10637–10642.

[52] Lorković, Z. J., Role of plant RNA-binding proteins in development, stress response and genome organization. *Trends Plant Sci.* 2009, *14*, 229–236.

[53] Sami-Subbu, R., Choi, S.-B., Wu, Y., Wang, C., Okita, T. W., Identification of a cytoskeleton-associated 120 kDa RNA-binding protein in developing rice seeds. *Plant Mol. Biol.* 2001, *46*, 79–88.

[54] Wang, C., Washida, H., Crofts, A. J., Hamada, S. et al., The cytoplasmic-localized, cytoskeletal-associated RNA binding protein OsTudor-SN: evidence for an essential role in storage protein RNA transport and localization. *Plant J.* 2008, *55*, 443–454.

[55] Morris, R. T., Doroshenk, K. A., Crofts, A. J., Lewis, N. et al., RiceRBP: a database of experimentally identified RNA-binding proteins in *Oryza sativa* L. *Plant Sci.* 2011, *180*, 204–211.

[56] Xue, L.-J., Zhang, J.-J., Xue, H.-W., Genome-wide analysis of the complex transcriptional networks of rice developing seeds. *PloS One* 2012, *7*, e31081.

[57] Jiang, H., Dian, W., Wu, P., Effect of high temperature on fine structure of amylopectin in rice endosperm by reducing the activity of the starch branching enzyme. *Phytochemistry* 2003, *63*, 53–59.

[58] Zaima, N., Goto-Inoue, N., Hayasaka, T., Setou, M., Application of imaging mass spectrometry for the analysis of *Oryza sativa* rice. *Rapid Commun. Mass Spectrom.* 2010, *24*, 2723–2729.

[59] Kaspar, S., Peukert, M., Svatos, A., Matros, A., Mock, H.-P., MALDI-imaging mass spectrometry—an emerging technique in plant biology. *Proteomics* 2011, *11*, 1840–1850.

[60] Hoshikawa, K., Studies on the development of endosperm in rice: 3. observations on the cell division. *Jpn. J. Crop Sci.* 1967, *36*, 210–215.

[61] Hoshikawa, K., Studies on the development of endosperm in rice: XII. development of the protein—forming plastids. *Jpn. J. Crop Sci.* 1970, *39*, 295–300.

