iTRAQ-Based Quantitative Glutelin Proteomic Analysis Reveals Differentially Expressed Proteins in Physiological Metabolism Process during Endosperm Development and their Impacts on Yield and Quality in Autotetraploid Rice

Lin Xian
South China Agricultural University
Yanxi Long
South China Agricultural University
Meng Yang
South China Agricultural University
Zhixiong Chen
South China Agricultural University
Jinwen Wu
South China Agricultural University
Xiangdong Liu
South China Agricultural University
Lan Wang (wanglan@scau.edu.cn)
South China Agriculture University https://orcid.org/0000-0003-2242-1028

Original article

Keywords: Rice, Endosperm, Seed storage protein, Glutelin, Proteomic

Posted Date: June 24th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-36944/v1

Version of Record: A version of this preprint was published at Plant Science on May 1st, 2021. See the published version at https://doi.org/10.1016/j.plantsci.2021.110859.

Abstract

Background

Autotetraploid rice, which is developed through chromosome set doubling using diploid rice, produces high-quality kernels that are rich in storage proteins. However, little information is available about the content of different proteins in autotetraploid rice and their proteomic analysis.

Results

The dynamic changes in four storage proteins, albumin, globulin, prolamin, and glutelin, were analyzed in the endosperm of autotetraploid rice (AJNT-4x) and in that of its diploid counterpart (AJNT-2x) for comparison. The contents of the four proteins were all higher during endosperm development in AJNT-4x than in AJNT-2x, but their change and composition were almost the same in the two materials. Then, iTRAQ was employed to analyze the glutelin profiles of AJNT-4x and AJNT-2x at 10 DAF, 15 DAF, and 20 DAF. A total of 1,326 proteins were identified in AJNT-4x and AJNT-2**x** using high-throughput LC-MS/MS. Among the 1,326 identified proteins, there were 362 DEPs in AJNT-4x compared with those in AJNT-2x and 372 DEPs between different development stages in AJNT-4x. Eight important upregulated proteins were identified by qRT-PCR, including B8AM24, B8ARJ0, B8AQM6, A2ZCE6, and Q40689. Among them, B8AM24 and B8ARJ0 were related to the lysine biosynthesis process. GO enrichment analysis revealed that the critical functions of DEPs exhibited little overlap between the 10, 15, and 20 DAF groups. Endosperm glutelin accumulation was regulated mainly by different DEPs during the late stage, and 15 DAF was a critical regulating point for glutelin accumulation. KEGG pathway analysis showed that ribosomal proteins were significantly higher in AJNT-4x than in AJNT-2x at 10 DAF, and the protein processing, biosynthesis, and metabolism of amino acids were higher and more active in AJNT-4x at 15 DAF, while the peroxisome was richer in AJNT-4x at 20 DAF. The PPI network showed that ribosomal proteins gradually decreased with increasing endosperm development.

Conclusions

These results provide new insights into dynamic glutelin expression differences during endosperm development in autotetraploid rice, which will help in developing rice cultivars with increased yield and improved grain nutritional quality.

Background

Rice (*Oryza sativa* L.) is one of the most widely cultivated food crops in the world, especially in China, where the total output in 2019 was nearly 210 million tons, making China the world's largest rice producer. As the critical location of nutrient storage, endosperm in rice grains contributes greatly to embryo development for plant growth and nutrition, as rice is a staple food for humankind. For example, biological processes and protein production, aggregation, and metabolism are crucial for embryo development and nutritional quality.

The main ingredients of endosperm are starch and storage proteins. Endosperm storage proteins constitute approximately 5.5%-12% of polished rice grains (Zhang et al. 2014). With the development of modern biotechnology, mass spectrometry (MS)-based proteomics analysis has become a powerful tool to investigate protein expression levels in a broad range of biological samples, and it has also been used to interpret the molecular details of polyploid plants, such as wheat, peanut, and black locust (Islam et al. 2003; Meng et al. 2016; Chen et al. 2018; Liu et al. 2018). In terms of rice, a number of proteomic studies have focused on different organs or tissues under multiple conditions, for example, pathogen infection (Wang et al. 2015), environmental challenges (Zhang et al. 2017), and physical stress (Chen et al. 2015), and proteomics and phosphoproteomics have also been applied to interpret the molecular details of early embryo development (Wiśniewski et al. 2009; Wang et al. 2016).

As a basic feature in nearly all plant species, polyploidization normally results in a higher yield than that of diploid plants and is regarded as a powerful tool in crop breeding (Ramanna and Jacobsen 2003). During the last decade, the application of polyploidization has been used in rice breeding, and several highly productive hybrid rice lines have been developed (Cai et al. 2007; He et al. 2011; Zhang et al. 2017; Ghaleb et al. 2020). Through comparative studies between diploid and autotetraploid rice cultivars, it was shown that there are higher glutelin contents in autotetraploid cultivars than in diploid cultivars (Cai 2011). Autotetraploid rice has a higher content of storage protein and better quality and nutritional value than those of its diploid counterpart.

Rice storage proteins, namely, albumin, globulin, prolamin, and glutelin, are classified into four fractions according to their differences in solubility. Glutelin and prolamin are the main storage proteins in endosperm. Glutelin constitutes approximately 60%-80% of total endosperm proteins and is the main focus in polished grains (Chen et al. 2016). Prolamin constitutes approximately 5%-10% of total endosperm proteins (Tanaka et al. 1980). Glutelin is easily digested (Zhang et al. 2008) and is rich in lysine (Ogawa et al. 1987). Therefore, glutelin has higher nutritional value than the other three endosperm proteins. Increasing the endosperm glutelin content will enhance the nutrient level of rice grains and increase the content of lysine, which is an essential amino acid for humans. This increase in glutelin content would be of profound significance for rice quality breeding in terms of cloning more novel glutelin genes and further studying their molecular mechanism.

Nevertheless, there is a lack of comparative investigations at the protein level between diploid and autotetraploid rice, which results in little understanding of the developmental process of endosperm after flowering. In this study, a diploid and an autotetraploid rice line were researched further to identify differentially expressed proteins (DEPs) in different developmental stages. First, the protein components during endoplasm development and maturation were monitored, and some critical developmental time points were found. Then, a high-throughput isobaric tag for relative and absolute quantitation (iTRAQ) label-based proteomics analysis was performed to identify and quantify glutelin profiles in both diploid and autotetraploid endosperms, and bioinformatics tools were used to interpret the biological meaning of the changes in protein expression. The integration of these phenotypic and molecular results furthered the understanding of biological processes, primarily the regulation of protein expression and metabolism, in the

development and maturation of rice endosperm. Furthermore, the differentially expressed genes will be further researched to elucidate the molecular mechanism of glutelin accumulation in autotetraploid rice.

Results

SDS-PAGE and phenotypic analysis of storage proteins in seed endosperm development

The seed setting rate of new double autotetraploid rice is usually very low because of the separation behavior of the homologous chromosome. To study further dynamic changes in protein content and DEPs during different developmental stages, we selected AJNT-4x and AJNT-2x as the research materials. In our research, the seed setting rate of AJNT-4x reached 49.94%. On the other hand, the agronomic traits, such as grain length, grain width, and grain thickness, of AJNT-4x were also better than those of its diploid (Table 1). The grain weight of AJNT-4x increased obviously, and there was a significant difference in grain weight between AJNT-4x and AJNT-2x. In addition, storage proteins in the seed endosperm were divided into albumin, globulin, prolamin, and glutelin fractions according to their solubility. SDS-PAGE analysis showed that our results were the same as those of a previous report. The molecular weights of the albumin and globulin fractions were 14-16 kDa and 26 kDa, respectively. The molecular weights of the prolamin fraction were 10 kDa, 13 kDa, and 16 kDa, but mainly 13 kDa. The molecular weights of the glutelin fraction were 57 kDa, 32 kDa, and 20 kDa; among these molecular weights, 57 kDa was a precursor protein, 32 kDa was an acidic subunit, and 20 kDa was a basic subunit. All the expression levels of the four seed proteins increased gradually with increasing flowering time. The extracted protein fractions were divided into two parts: one was used to measure their contents, and the other was subjected to MS analysis for glutelin protein evaluation.

Dynamic changes in storage protein contents during endosperm development

The contents and proportions of endosperm storage proteins affect the quality of rice (You et al. 2017). In this study, the contents of four storage proteins were measured by the BCA method (Noble and Bailey 2009). The four proteins all began to accumulate, and their contents were almost the same at 5 DAF in AJNT-2x and AJNT-4x.

The accumulation trends of albumin and globulin were almost the same in AJNT-4x and AJNT-2x. From 5 DAF to grain maturity, during the whole endosperm developmental period, the albumin and globulin contents increased only slightly (Figure 1B).

The prolamin content first decreased, followed by a slight increase, and then a rapid increase from 12 DAF. The content reached a high peak at 18 DAF in AJNT-4x, declined again, and finally stabilized. Eighteen days after flowering, slightly later than that in AJNT-2x, the prolamin content reached its maximum of 0.31 mg/mL in AJNT-4x. The prolamin content in AJNT-2x reached a peak of 0.18 mg/mL at 15 DAF. From 12 to 15 DAF, the prolamin content increased much more rapidly in AJNT-4x than in AJNT-2x (Figure 1C). The prolamin content was much higher in AJNT-4x than in AJNT-2x during all the developmental stages, and the prolamin content difference between AJNT-4x and AJNT-2x at maturity was highly significant.

Compared with those of albumin, globulin, and prolamin, the variation range of glutelin was the largest during endosperm development. Glutelin accumulation was almost the same in AJNT-4x and AJNT-2x. There were two rapid accumulation stages: one was from 9 to 11 DAF, and the other was from 18 to 21 DAF in AJNT-4x, presenting two peaks at 11 DAF and 21 DAF (Figure 1D). Then, the glutelin content declined gradually and stabilized. In general, the accumulation rate of glutelin in AJNT-4x was faster than that in AJNT-2x. The glutelin content was highly significantly different between AJNT-4x and AJNT-2x during the rapid accumulation period.

The change in the contents of the four proteins in AJNT-4x was almost the same as that in AJNT-2x, but the total content in AJNT-4x was much higher than that in AJNT-2x because the four proteins accumulated quickly in AJNT-4x (Figure 1E). Glutelin is rich in lysine, which is an essential amino acid for humans, so we further studied changes in the glutelin content during the developmental stages of both AJNT-4x and AJNT-2x. According to the change in glutelin content during endosperm development, samples from several critical developmental stages, namely, samples collected 10, 15, and 20 DAF, were subjected to proteomic analysis by using the iTRAQ technique.

Proteome differential analysis at different endosperm developmental stages between AJNT-4x and AJNT-2x

To investigate differential alterations in glutelin expression during seed endosperm development, glutelin fractions from both AJNT-4x and AJNT-2x endosperm were extracted following a TCA-acetone method (Méchin et al. 2007) and subjected to iTRAQ-based global proteomics analysis by LC-MS/MS (Unwin et al. 2010). A total of 841,375 spectra were generated, of which 38,851 were matched to known peptides from the utilized database. Eventually, 4,880 peptides, 4,047 unique peptides and 1,326 protein groups were identified. Comparative analysis between AJNT-4x and AJNT-2x at 10, 15, and 20 DAF was conducted, and 125, 159 and 78 DEPs were found (fold change>1.2 or<0.83, p-value<0.05) (Table 2). There were many DEPs in the different developmental stages of rice endosperm analyzed. A total of 372 DEPs were measured at 10, 15, and 20 DAF AJNT-4x samples, which was much more than the 184 DEPs detected in the AJNT-2x samples.

Principal component analysis (PCA) showed that the autotetraploid rice endosperm had better separation than diploid endosperm, which indicated that more significant alterations in protein expression occurred in autotetraploid endosperm (Figure 2).

Differentially expressed protein GO analysis for AJNT-4x and AJNT-2x

All DEPs were annotated by GO. Comparing DEPs between AJNT-4x and AJNT-2x at 10, 15 and 20 DAF, it was surprising that no GO functions were mutual among the three comparison groups, but there were 15 common DEPs between 15 DAF and 20 DAF, and only one common DEP between 10 DAF and 15 DAF (Figure 3A), which indicated that endosperm glutelin accumulation was regulated mainly by different DEPs during the late stage, and 15 DAF was a critical regulating point for glutelin accumulation.

The GO enrichment results also showed that the critical functions of DEPs exhibited little overlap at 10, 15 and 20 DAF. For instance, at 10 DAF, all DEPs between AJNT-4x and AJNT-2x endosperm samples were

related to metabolic processes (Figure 3B), including those of RNA (GO: 0016072, 0034660), peptides (GO: 0006518), proteins (GO: 0019538) and phosphate (GO: 0019220). At 15 DAF, the DEPs AJNT-4x and AJNT-2x represented a much broader range of regulatory functions, such as growth regulation (GO: 0040007). Metabolic processes or functions were still annotated at 15 DAF, but few proteins for RNA, peptide and protein metabolism were enriched (Figure 3C). At 20 DAF, the number of DEPs between AJNT-4x and AJNT-2x endosperm was decreased compared to that at 10 or 15 DAF (Figure 3D). However, the annotated functions of these proteins were still abundant and mainly enriched in stress response (GO: 0006950), carbohydrate metabolism (GO: 1901135), nutrient reservoir (GO: 0045735) and so on. In summary, differences in protein synthesis and metabolism mainly occurred at 10 DAF, but decreased gradually during later developmental stages. DEPs involved in cell growth and development and amino acid biosynthesis were mainly observed at 15 DAF.

KEGG analysis of differentially expressed proteinsbetween AJNT-4x and AJNT-2x

To investigate metabolic processes and feasible signaling pathways, these DEPs were subjected to KEGG analysis. The KEGG pathway results (Figure 4) indicated that the biological pathways for AJNT-4x endosperm proteins were different from those for AJNT-2x endosperm proteins at 10, 15 and 20 DAF. Ribosomal proteins (Figure S1), which were significantly higher in AJNT-4x than in AJNT-2x endosperm, were the major difference at 10 DAF due to alterations in central carbon metabolism and cellular senescence. At 15 DAF, the most enriched pathway was protein processing in the endoplasmic reticulum (ER) (Figure S2), which contained 12 DEPs, and all of them were enriched in AJNT-4x endosperm compared to their expression levels in AJNT-2x. For example, protein disulfide-isomerase and heat shock protein 70 are both critical for protein folding (Gruber et al. 2006; Wisén and Gestwicki 2008). The biosynthesis and metabolism of amino acids was another important series of pathways enriched at 15 DAF, and the enriched proteins were involved in arginine biosynthesis; alanine, aspartate and glutamate metabolism; tyrosine metabolism, cysteine and methionine metabolism; arginine and proline metabolism; phenylalanine metabolism; and phenylalanine, tyrosine and tryptophan biosynthesis. These amino acid biosynthetic and metabolic pathways were much higher in AJNT-4x than in AJNT-2x (Table S1). This result suggested that not only were protein synthesis and processing more active in AJNT-4x endosperm at 15 DAF but there was also a higher level of amino acid accumulation. Interestingly, two upregulated differential lysine proteins (B8AM24 and B8ARJ0) were found in the lysine biosynthesis process (Figure S3), and four upregulated differential proteins (P37866, Q9S768, A0A0N7KTS9 and B8AEL7) were found to regulate alanine, aspartic acid, and glutamic acid metabolism processes (Figure S4).

In the late period (20 DAF) of endosperm development, the pathways of DEPs were primarily metabolic pathways, including fructose and mannose metabolism, methane metabolism, and starch and sucrose metabolism (Figure S5). However, the most enriched pathway at 20 DAF was the peroxisome pathway (Figure S6), in which 4 DEPs were downregulated in AJNT-4x when compared to their expression levels in AJNT-2x. As a crucial organelle that regulates large numbers of biological processes, for instance, metabolism and development (Hu et al., 2012), the lower level of the peroxisome pathway in AJNT-4x at this stage suggested reduced metabolic and developmental activity compared with that in AJNT-2x endosperm. Interestingly, KEGG pathway analysis also mapped DEPs to protein processing in the ER (Figure S7) at 20

DAF, but two proteins, namely, protein transport protein Sec61 and translocon-associated protein, had lower expression levels in AJNT-4x than in AJNT-2x.

Protein-protein interaction (PPI) network for DEP analysis

The protein-protein interaction (PPI) network is another important tool to interpret proteomics information. DEPs between AJNT-4x and AJNT-2x were annotated, and PPI networks were generated by the String database (Szklarczyk et al. 2015). As shown in the results, most connected and annotated proteins were ribosomal proteins and metabolic proteins (Figure 5). At 10 DAF, half of the DEPs in the PPI networks were ribosomal proteins, and the other half were metabolic proteins. With increasing endosperm development, ribosomal proteins gradually decreased, and metabolic proteins increased. Until 20 DAF, the proportion of ribosomal proteins was reduced to approximately 1/3, and no ribosomal DEPs were found in the PPI network.

WGCNA analysis of co-expression modules

Traditionally, the analysis of proteomics data has focused on DEPs, which were regarded as largely important participants in the comparison group. However, a large amount of “unchanged” proteins (fold change<1.2) and “nonsignificantly changed” proteins (*p*-value>0.05) might be ignored because of disadvantages in statistics, although they might result in an important impact on the phenotype. Therefore, a WGCNA approach was applied to construct a global co-expression network without any artificial cut-offs (Pei et al. 2017). First, the identified proteins were divided into 5 different modules (Figure 6A), and then GO and KEGG analysis of each module were performed. Briefly, the blue module contained ribosomal and oxidative phosphorylation-related proteins; the brown module contained proteins related to carbohydrate metabolism, including the metabolism of amino/nucleotide sugars, starch, sucrose, fructose and mannose, and glycolysis; the turquoise module contained proteins related to the proteasome and fatty acid biosynthesis; the yellow and gray modules were similar to the blue module but failed to enrich any KEGG pathways (Figure 6B). During the endosperm development process of AJNT-2x, it can be concluded that proteins in the blue and turquoise modules had high expression levels at 10 DAF that then decreased at 15 and 20 DAF. The proteins in the blue module were expressed much higher in AJNT-4x than in AJNT-2x at 10 DAF but had similar levels at 15 and 20 DAF (Figure 6C). These findings suggested that protein expression and oxidative phosphorylation, which provide energy to biological processes, were highly active in the early stage of development but quickly slowed. In contrast, carbohydrate metabolism processes were less active at 10 DAF than at 15 and 20 DAF but then increased in both AJNT-4x and AJNT-2x. Compared with the differences in expression levels between different developmental stages in AJNT-2x, those in AJNT-4x showed more significant differences.

Confirmation of proteomic data by qRT-PCR analysis

To validate the experimental proteomics data, thirteen randomly selected DEPs, including ten upregulated DEPs and three downregulated DEPs (fold change>1.2 or<0.83, *p*-value<0.05), were analyzed using real-time quantitative PCR, (Table 3). Among them, an ATP binding protein (A0A0N7KMN9) and a cell redox homeostasis protein (B8AJS5) were significantly downregulated, and a glutelin protein (Q40689) and cell

growth protein (A3AI97) were upregulated, contrasting with the results of iTRAQ; the results for the other nine proteins were consistent with the iTRAQ data (Figure 7). Thus, the proteomics data were reliable. Among the nine DEPs with qRT-PCR data correlating with those of iTRAQ, eight were upregulated proteins (Table 3). These proteins are very important to plant growth and development. For example, B8AM24 and B8ARJ0 are related to the lysine biosynthetic process, B8AQM6 regulates cell growth by extracellular stimulus, A2ZCE6 is related to endosperm development, and P55857 is a glutelin protein. These proteins will be subjected to functional experiments in subsequent work.

Discussion

The total storage protein content was higher in autotetraploid rice than in diploid

Rice contains four seed storage proteins primarily located in the endosperm and is rich in all the eight essential amino acids necessary for human health (Kannan et al. 2001). In addition, the rice endosperm plays crucial roles in nourishing the embryo during embryogenesis and seed germination. Previous studies have shown that protein constructs were almost the same between autotetraploid rice and diploid rice, and the protein contents were higher in autotetraploid rice than in the corresponding diploid rice (Xie et al. 2007; Wang et al. 2008; Gu et al. 2015). In our study, the protein construction of four seed storage protein fractions was almost the same between AJNT-4x and its AJNT-2x, but the protein content was higher in AJNT-4x, which was also reported in previous studies. Through self-crossing of fifty-two generations, the protein expression of the repetitive genome on homologous chromosomes tends to diploidize autotetraploid rice (Zhang et al. 2015), which presents similar protein construction. With increasing gene dosages, the protein expression content will be higher in autotetraploid rice than in the corresponding diploid rice. This is one of the causes of the increased seed setting rate of AJNT-4x.

The change in the protein contents was different between the four proteins in our study; the albumin and globulin content increased slowly, and the rate of increase was almost the same between AJNT-4x and AJNT-2x, differing only in the time at which the peak concentration was reached, i.e., their maximum concentrations were reached earlier in AJNT-4x than in AJNT-2x. The prolamin content increased more quickly in AJNT-4x than in AJNT-2x, but reached its maximum later in AJNT-4x than in AJNT-2x. The glutelin content began to increase from 6 DAF and reached its peak concentration at almost the same for AJNT-4x and AJNT-2x. These results differ from previous conclusions; for instance, prolamin was thought to be expressed beginning at 5 DAF, to be expressed abundantly at 10 DAF, and then to increase gradually until grains reached maturity (Juliano and Chemists 1985; Wang et al. 2009). This difference in results could be a consequence of different research materials. The beginning of endosperm cellularization was shown to be earlier in autotetraploids than in diploids (Wang et al. 2005), and protein accumulation may occur earlier in autotetraploids. The total storage protein content was higher in the autotetraploid than in the corresponding diploid rice in this study.

Different genes were differentially expressed at different times

The expression of storage proteins was considered to be regulated by multiple gene families. Albumin with a weight of 16 kDa contained a little gene family. Globulin contained 26 kDa and 16 kDa globulin (Pan and Reeck 1988). Prolamins genes were classified into three groups: those of 10 kDa prolamin (RP10), 13 kDa prolamin (RM1, RM2, RM4, and RM9), and 16 kDa prolamin (RP16) (Yamagata et al. 1982; Zhang et al. 2008). Glutelin is encoded by 15 genes classified into four subfamilies: GluA, GluB, GluC, and GluD (Kawakatsu et al. 2008). These storage protein genes were differentially expressed at different times during endosperm development. For example, GluA-3 expression was high, and GluA-1 and GluA-2 expression was low at 5 DAF. However, until 15 DAF, GluA-1 and GluA-2 expression increased approximately 3-fold, while that of GluA-3 remained fundamentally stable. GluA-3 expression peaked at 10 DAF and then began to decline (Cai et al. 2011). For prolamin, the first group encoding the Pro114 gene began to be expressed at 5 DAF, and the third group encoding the S23 gene began to be expressed at 8 DAF (Kawakatsu et al. 2008). The abundances of differentially expressed genes were different in different developmental stages, which resulted in fluctuations in the protein content.

The time of initial and maximum expression of these genes differed among the different materials, so the expression profiles were also different. Seed storage proteins are quantitatively inherited, and the gene expression of proteins in rice grains from 15 DAF to maturation is relatively easily influenced by environmental conditions (Shi et al. 2006; Shewry. 2007), which will result in more complex protein expression.

Proteins that control autotetraploid rice metabolism were identified

Endosperm development in rice was divided into four stages: the coenocyte stage (1-2 DAF), the cellularization stage (3-5 DAF), the storage product accumulation stage (6-20 DAF), and maturation (21-30 DAF) (Wu et al. 2016). The major phase of storage product accumulation occurred between 6 and 21 DAF (Wu et al. 2016). In this study, we systematically analyzed these differentially expressed genes in the glutelin accumulation process at 10, 15, and 20 DAF between AJNT-4x and AJNT-2x by GO and KEGG pathway annotation, PPI networks, and WGCNA co-expression modules. The DEPs regulating protein synthesis and metabolism mainly occurred at 10 DAF. In this stage, the levels of ribosomal proteins in autotetraploid rice were significantly higher than those in diploid rice. The expression gradually decreased with endosperm development, while the DEPs involved in cell growth and acid synthesis, such as B8AQM6, A0A0N7KMN9, B8AJS5, and P37833, rapidly increased at 15 DAF. Finally, the DEPs promptly decreased from 105 to 36, with endosperm development from 15 DAF to 20 DAF. This result also showed that 15 DAF was the critical period of endosperm protein accumulation and that the DEPs were higher in the autotetraploid than in the diploid. Therefore, glutelin accumulated earlier and more in autotetraploid rice than in its diploid counterpart. The mass spectrometric results corresponded with the measured protein contents.

The amylose content was negatively correlated with protein content

Although the proteomic profiles of endosperm development have rarely been reported, previous studies have been performed only at the transcriptome level (Zhou et al. 2013, Xu et al. 2016; Chen et al. 2018; Zhu et al. 2019). Proteases synthesized by some of the genes identified in transcriptome studies were also identified

in this study. For instance, granule-bound starch synthase I (waxy) (B1B5Z2), an important enzyme for amylose synthesis (Hanashiro et al. 2013), was identified and located in the gray WGCNA module, indicating that it was increasingly expressed in the late period of endosperm development. In the three different periods, the wax content of the autotetraploid was lower than that of the corresponding diploid, so the content of amylose synthase was also lower. The amylose content was negatively correlated with the protein content (Cheng et al. 2016), so the protein content of the autotetraploid was higher than that of the diploid.

Protein content was significantly negatively correlated with the number of grains per panicle and seed setting rate, and was positively correlated with grain length, grain width, and grain weight (Zou et al. 2001; Cheng et al. 2016). The number of grains per panicle and seed setting rate of the autotetraploid were all lower than those of the corresponding diploid, and the grain length, width and weight of the autotetraploid were also obviously higher than those of the corresponding diploid (Table 1).

However, some known genes involved in endosperm development, for example, KRP1 and some transcription factors, were unable to be identified in this dataset. This might be because of the acquisition of an incomplete proteomics dataset due to the data-dependent (DDA) acquisition method (Röst et al. 2014).

Changes in carbohydrate metabolism in the endosperm

Metabolism-related proteins were another main class of proteins that were identified in this study. As critical processes of energy metabolism, oxidative phosphorylation and glycolysis were highlighted in the proteomics results. Unlike oxidative phosphorylation, which was more active at 10 DAF, glycolysis-related proteins had relatively low expression levels at 10 DAF and showed different tendencies of increase in the diploid and autotetraploid endosperms (Figure 8). In the diploid endosperm, the expression of glycolysis-related proteins continued to increase from 15 to 20 DAF, and in the autotetraploid endosperm, the peak of glycolysis occurred at 15 DAF and decreased slightly during the last 5 days. The alteration in carbohydrate metabolism changed not only the biological process but also the storage of nutrients.

Conclusion

In conclusion, a total of 372 DEPs were observed in AJNT-4x, and glutelin accumulation was regulated mainly by different DEPs during the late stage. We found that 15 DAF was a critical regulatory period for glutelin accumulation, and DEPs related to protein synthesis and metabolism mainly occurred at 10 DAF, while the differential proteins involved in cell growth and development and amino acid biosynthesis mainly occurred at 15 DAF. On the other hand, some important DEPs were found in the autotetraploid, and eight important upregulated proteins were verified by qRT-PCR, including B8AM24, B8ARJ0, B8AQM6, and A2ZCE6. These proteins regulate lysine biosynthesis, cell growth, and endosperm development. Our results revealed alterations of protein expression in these important biological processes during the different developmental stages in rice. These results will be helpful for gaining a general understanding of metabolic changes and molecular mechanisms related to autotetraploid rice development. The future direction of this

study is to validate key proteins that are involved in developmental progress, and functional studies will also be performed on these key proteins, especially proteins that are still uncharacterized.

Materials And Methods

Grain weight and seed setting rate investigation and endosperm preparation

Rice lines (*O. sativa* L. *sp.* Aijiaonante (AJNT)-2x and AJNT-4x after self-crossing for 52 generations) were grown on farms at South China Agricultural University in the early and late seasons in 2019. Ten rows were planted in each of the test plots, and the row spacing was 18 cm×21 cm. During the maturation period, a total of 70 rows were measured for seed setting in the middle of the test plots (Li et al. 2011). Calculations and analyses were performed using SPASS software (Liu et al. 2016).

The grains were labeled at the time of flowering and were collected 5, 6, 7, 8, 9, 10, 11, 12, 15, 18, 20□21, and 24 days after flowering (DAF) and stored at -80 °C in a freezer.

Protein extraction and quantification

Rice embryos were removed from the stored grains, and then the endosperms were frozen in liquid nitrogen and ground with a mortar and pestle. Five volumes of TCA/acetone (1:9, v/v) were added to the powder and mixed by vortexing. The mixture was placed at -20 °C for 4 h and centrifuged at 6,000 g for 40 min at 4 °C. The supernatant was discarded, and precooled acetone was added to wash the residue three times. The precipitate was then air dried.

For extraction of globulin, 100 mg of dried precipitate was resuspended in 1 mL of globulin extraction buffer (60 mM Tris-HCl, pH 6.6, 0.7 M NaCl), mixed and placed in a 37 °C water bath for 20 min and then centrifuged at 11,000 rpm for 10 min. The supernatant was collected as the globulin fraction. The precipitate from the last step was washed with water, resuspended in 65% propanol and then incubated in a 37 °C water bath for 20 min. The supernatant was collected after centrifugation at 11,000 rpm for 10 min as the prolamin fraction. Finally, the glutelin fraction was extracted by a 1.2% lactic acid solution using the same method described above. To prepare samples for analysis via liquid chromatography-tandem mass spectrometry (LC-MS/MS), 30 volumes of SDT buffer were added to 20-30 mg of powder, mixed and boiled for 5 min. The lysate was sonicated and then boiled for 15 min. After centrifugation at 14,000 g for 40 min, the supernatant was filtered through a 0.22 μm filter. The filtrate was quantified with the BCA protein assay kit (Bio-Rad, USA). All samples were stored at -80 °C.

Protein digestion and iTRAQ labeling

Protein digestion was performed according to the FASP procedure described by Wiśniewski et al. (2009), and the resulting peptide mixture was labeled using the 8-plex iTRAQ reagent according to the manufacturer's instructions (AB Sciex). Briefly, 200 μg of protein for each sample was incorporated into 30 μL of STD buffer (4% SDS, 100 mM DTT, and 150 mM Tris-HCl; pH 8.0). The detergent, DTT, and other low-molecular-weight components were removed using SDT buffer (4% (w/v) SDS in 100 mM Tris-HCl; pH 7.6;

50 mM DTT) by repeated ultrafiltration (Microcon units, 30 kD). Then, 100 μL of 0.05 M iodoacetamide in UA buffer was added to block reduced cysteine residues, and the samples were incubated for 20 min in the dark. The filters were successively washed with 100 μL of SDT buffer three times and 100 μL of DS buffer (50 mM triethylammonium-bicarbonate at pH 8.5) twice. Finally, the protein suspensions were digested with 2 μg of trypsin (Promega) in 40 μL of DS buffer overnight at 37 °C, and the resulting peptides were collected in the filtrate. The peptide content was estimated by UV light spectral density at 280 nm using an extinction coefficient of 1.1 for a 0.1% (g/L) solution that was calculated on the basis of the frequency of tryptophan and tyrosine in vertebrate proteins.

For labeling, each iTRAQ reagent was dissolved in 70 μL of ethanol, added to the respective peptide mixture, and multiplexed and vacuum dried.

LC-MS/MS analysis

Experiments were performed on a Q Exactive mass spectrometer that was coupled to an Easy nLC system (Proxeon Biosystems, now Thermo Fisher Scientific). Ten microliters of each fraction was injected for nanoLC-MS/MS analysis. The peptide mixture (5 μg) was loaded onto a C18-reversed-phase column (Thermo Scientific Easy Column, 10 cm long, 75 μm inner diameter, 3 μm particle size) in buffer A (0.1% formic acid) and separated with a linear gradient of buffer B (80% acetonitrile and 0.1% formic acid) at a flow rate of 250 nL/min controlled by IntelliFlow technology over 140 min. MS data were acquired using a data-dependent top10 method that dynamically chose the most abundant precursor ions from the survey scan (300-1,800 m/z) for high-energy collisional dissociation (HCD) fragmentation. Determination of the target value was based on predictive automatic gain control (pAGC). The dynamic exclusion duration was 60 s. Survey scans were acquired at a resolution of 70,000 at m/z 200, and the resolution for HCD spectra was set to 17,500 at m/z 200. The normalized collision energy was 30 eV, and the underfill ratio, which specifies the minimum percentage of the target value likely to be reached at the maximum fill time, was defined as 0.1%. The instrument was run with peptide recognition mode enabled.

Data processing

MS/MS spectra were searched using the MASCOT engine (Matrix Science, London, UK; version 2.2) embedded into Proteome Discoverer 1.3 (Thermo Electron, San Jose, CA.) against the UniProt *O. sativa* database (215,215 sequences downloaded on March 15, 2018) and the decoy database. For protein identification, the following parameters were used: peptide mass tolerance, 20 ppm; MS/MS tolerance, 0.1 Da; enzyme, trypsin; missed cleavage, 2; fixed modification, carbamidomethyl (C), iTRAQ4/8plex (K), and iTRAQ4/8plex (N-term); variable modification, oxidation (M); and false discovery rate (FDR), ≤0.01.

Bioinformatics

Sequence data of the selected DEPs were retrieved in batches from the UniProtKB *O. sativa* database in FASTA format. The retrieved sequences were locally searched against the SwissProt database (*O. sativa*) using NCBI BLAST plus client software to find homologous sequences from which the functional annotation could be transferred to the studied sequences. In this work, the top 10 BLAST hits with E-values

less than 1e-3 for each query sequence were retrieved and loaded into Blast2GO2 (Version 3.0.1) (Conesa et al. 2005) for Gene Ontology (GO) mapping and annotation. Following annotation and annotation augmentation steps, the studied proteins were also searched against the Kyoto Encyclopedia of Genes and Genomes (KEGG) GENES database to retrieve their KEGG orthologies (Kos) and were subsequently mapped to pathways in KEGG.

Protein interaction data of the studied proteins were retrieved from the IntAct molecular interaction database6 by their gene symbols. The results were downloaded in XGMML format and imported into Cytoscape 3.7.0 for further analysis. A weighted gene co-expression network analysis (WGCNA) co-expression network was generated by the R package (Langfelder and Horvath, 2008).

Quantitative PCR

Quantitative PCR was conducted using the manufacturer instructions of the 2×miRNA qPCR masterMix kit of Sangon Biotech (Shanghai) Co., Ltd. A 0.2 mL thin-walled 96-well PCR plate was acquired and numbered. The 2×miRNA qPCR masterMix dye (10 μL) was added to a tube; then, 1 μL of each of the forward and reverse primers (primer concentration: 10 μmol/L) was added, and 1 μL of the mixed cDNA was added to the tube. Each tube was supplemented with ddH_2O to 20 μL, centrifuged at 800 rpm for 30 s, and placed in a LightCycler 480 Real-time PCR System. A two-step cycle was carried out: 95 °C for 30 s, 95 °C for 5 s, 60 °C for 1 min, and the last two processes for 40 cycles and 40 °C for 1 min. The 2×miRNA qPCR masterMix kit was purchased from Sangon Biotech (Shanghai) Co., Ltd.

Abbreviations

iTRAQ: Isobaric tag for relative and absolute quantitation; LC-MS/MS: liquid chromatography-tandem mass spectrometry; DEPs: differentially expressed proteins; HCD: High-energy collisional dissociation; GO: Gene Ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes; WGCNA: weighted gene co-expression network analysis; PPI: The protein-protein interaction.

Declarations

Acknowledgments

Not applicable.

Authors' Contributions

LX and YXL performed the experiments. XDL and LW conceived and designed the experiments. LX and LW analyzed the data and wrote the paper. MY, ZXC, JWW, XDL supervised and complemented the manuscript. All authors read and approved the final manuscript.

Funding

The study was supported by Guangdong Provincial Natural Science Foundation Program (2019A1515011826), and Guangdong provincial Science and Technology Innovation Program (2013KJCX0035).

Availability of Data and Materials

The data sets supporting the conclusions of this article are included within the article and its additional files.

Ethics Approval and Consent to Participate

Not applicable.

Consent for Publication

Not applicable.

Competing Interests

The authors declare that they have no competing interests

Author details

[1]State Key Laboratory for Conservation and Utilization of Subtropical Agro-Bioresources, College of Agriculture, South China Agricultural University, Guangzhou 510642, China; [2]Guangdong Provincial Key Laboratory of Plant Molecular Breeding, College of Agriculture, South China Agricultural University, Guangzhou 510642, China

References

1. Zhang X, Shi L L, Ding D L, Wang S W, Cui J (2014) The relationship between rice protein related character and the RVA characteristic profile and palatability character. Food Science and Technology 39(10): 189-191
2. Islam N, Tsujimoto H, Hirano H (2003) Proteome analysis of diploid, tetraploid and hexaploid wheat: towards understanding genome interaction in protein expression. Proteomics 3(4), 549-557
3. Meng F J, Luo Q X, Wang Q Y, Zhang X L, Qi Z H, Xu F L, Lei X, Cao Y, Chow W S, Sun G Y (2016) Physiological and proteomic responses to salt stress in chloroplasts of diploid and tetraploid black locust (*Robinia pseudoacacia* L.). Sci Rep 6: 23098
4. Chen S L, Chen J, Hou F, Feng Y G, Zhang R Q (2018) iTRAQ-based quantitative proteomic analysis reveals the lateral meristem developmental mechanism for branched spike development in autotetraploid wheat (*Triticum turgidum* L.). BMC genomics 19(1): 228
5. Liu H, Li H F, Gu J Z, Deng L, Ren L, Hong Y B, Lu Q, Chen X P, Liang X Q (2018) Identification of the Candidate Proteins Related to Oleic Acid Accumulation during Peanut (*Arachis hypogaea* L.) Seed Development through Comparative Proteome Analysis. Int J Mol Sci 19(4): 1235

6. Wang B, Hajano J U D, Ren Y D, Lu C T, Wang X F (2015) iTRAQ-based quantitative proteomics analysis of rice leaves infected by *Rice stripevirus* reveals several proteins involved in symptom formation. Virol J 12(1): 99
7. Zhang H Y, Lei G, Zhou H W, He C, Liao J L, Huang Y J (2017) Quantitative iTRAQ-based proteomic analysis of rice grains to assess high night temperature stress. Proteomics 17(5): 1600365
8. Chen C, Song Y F, Zhuang K, Li L, Xia Y, Shen Z G (2015) Proteomic analysis of copper-binding proteins in excess copper-stressed roots of two rice (*Oryza sativa* L.) varieties with different Cu tolerances. PLoS One 10(4), e0125367
9. Wiśniewski J R, Zougman A, Nagaraj N, Mann M (2009) Universal sample preparation method for proteome analysis. Nat methods 6: 359-362
10. Wang S Z, Chen W Y, Yang C D, Yao J, Xiao W F, Xin Y, Qiu J R, Hu W M, Yao H G, Ying W, Fu Y P, Tong J X, Chen Z Z, Ruan S L, Ma H S (2016) Comparative proteomic analysis reveals alterations in development and photosynthesis-related proteins in diploid and triploid rice. BMC Plant Biol 16(1), 199
11. Ramanna M S, Jacobsen E (2003) Relevance of sexual polyploidization for crop improvement–A review. Euphytica 133(1): 3-8
12. Cai D T, Chen J G, Chen D L, Dai B C, Zhang W, Song Z J, Yang Z F, Du C Q, Tang Z Q, He Y C, Zhang D S,He G C, Zhu Y G (2007) The breeding of two polyploid rice lines with the characteristic of polyploid meiosis stability. Science in China Series C: Life Sciences 50(3): 356-366
13. He Y C, Ge J, Wei Q, Jiang A M, Gan L, Song Z J, Cai D T (2011) Using a polyploid meiosis stability (PMeS) line as a parent improves embryo development and the seed set rate of tetraploid rice hybrid. Can J Plant Sci 91(2): 325-335
14. Zhang X H, Zuo B, Song Z J, Wang W, He Y C, Liu Y H, Cai D T (2017) Breeding and study of two new photoperiod-and thermo-sensitive genic male sterile lines of polyploid rice (*Oryza sativa* L.). Sci Rep 7(1): 14744
15. Ghaleb M A A, Li C, Shahid M Q, Yu H, Liang J H,Chen R X, Wu,J W, Liu X D (2020) Heterosis analysis and underlying molecular regulatory mechanism in a wide-compatible neo-tetraploid rice line with long panicles. BMC Plant Biol 20:83
16. Cai Z Z (2011) Differential expression of the genes related storage proteins in endosperms of rice with different ploidy. The master's thesis of Hubei University, Wuhan in China p35
17. Chen D G, Zhou X Q, Liu C G, Li L J, Li J C, Chen Y D (2016) Breeding of indica rice lines with low glutelin content by molecular marker-assisted selection. Molecular Plant Breeding 14(7): 1753-1758
18. Tanaka K, Sugimoto T, Ogawa M, Kasai Z (1980) Isolation and characterization of two types of protein bodies in the rice endosperm. Agric Biol Chem 44(7):1633-1639
19. Zhang W W, Bi J C, Chen L M, Zheng L N, Ji S L, Xia Y M, Xie K, Zhao Z G, Wang Y H, Liu L L, Jiang L, Wan J M (2008) QTL mapping for crude protein and protein fraction contents in rice (*Oryza sativa* L.). J Cereal Sci 48(2): 539-547
20. Ogawa M, Kumamaru T, Satoh H, Iwata N, Omura T, Kasai Z, Tanaka K (1987) Purification of protein body-I of rice seed and its polypeptide composition. *Plant Cell Physiol* 28(8):1517–1527

21. You C C, Chen L, He H B, Wu L Q, Wang S H, Ding Y F, Ma C X (2017) iTRAQ-based proteome profile analysis of superior and inferior spikelets at early grain filling stage in japonica rice. BMC Plant Biol 17:100
22. Noble J E and Bailey M J A (2009) Methods in Enzymology, Chapter 8 Quantitation of Protein 463: 73-95
23. Méchin V, Damerval C, Zivy M (2007) Total protein extraction with TCA-acetone. In Plant Proteomics, Humana Press: p1-8
24. Unwin R D, Griffiths J R, Whetton A D (2010) Simultaneous analysis of relative protein expression levels across multiple samples using iTRAQ isobaric tags with 2D nano LC–MS/MS. Nat Protoc 5(9):1574-1582
25. Gruber C W, Čemažar M, Heras B, Martin J L, Craik D J (2006) Protein disulfide isomerase: the structure of oxidative folding. Trends Biochem Sci 31(8), 455-464
26. Wisén S and Gestwicki J E (2008) Identification of small molecules that modify the protein folding activity of heat shock protein 70. Anal Biochem 374(2), 371-377
27. Hu J P, Baker A, Bartel B, Linka N, Mullen R T, Reumann S, Zolman B K (2012) Plant peroxisomes: biogenesis and function. The Plant Cell 24(6), 2279-2303
28. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, Simonovic M, Roth A, Santos A, Tsafou K P, Kuhn M, Bork P, Jensen L J, von Mering C (2015) STRING v10: protein–protein interaction networks, integrated over the tree of life. Nucleic Acids Res 43: 447-452
29. Pei G, Chen L, Zhang W (2017) Chapter nine, WGCNA application to proteomic and metabolomic data analysis. Methods in Enzymology 585: 135-158
30. Kannan S, Nielsen S S, Mason A C (2001) Protein digestibility-corrected amino acid scores for bean and bean-rice infant weaning food products. J Agric Food Chem 49(10):5070-5074
31. Xie H B, Huang Q C, Li G P, Wang X L, Ye C Y, Qin G Y (2007) Differential expression of the proteins in endosperms of rice with different chromosome sets. Hereditas 29(3):360-364
32. Wang X L, Xie H B, Huang Q C, Dai X M, Hu X M, Qin G Y (2008) Effects of Chromosome Doubling on Morphology and endosperm protein content of rice. J Henan Agricultural Sci 02:21-24
33. Gu Y L, Dai X M, Li J X (2015) Rice quality analysis of different ploidy rice. J Zhengzhou Univ (Nat Sci Ed) 47(4):81-85
34. Zhang J, Liu Y, Xia E H, Yao Q Y, Liu X D, Gao L Z (2015) Autotetraploid rice methylome analysis reveals methylation variation of transposable elements and their effects on gene expression. Proc Natl Acad Sci U S A 112(50):7022-7029
35. Juliano B O, Chemists A A C (1985) Rice: chemistry and technology. 2nd Ed, St Paul, American Association of Cereal Chemists, St Paul MN:324-326
36. Wang Y H, Zhu S S, Liu S J, Jiang L, Chen L M, Ren Y L, Han X H, Liu F, Ji S L, Liu X, Wan J M (2009) The vacuolar processing enzyme OsVPE1 is required for efficient glutelin processing in rice. The Plant Journal 58(4):606-617

37. Wang L, Liu X D, Lu Y G, Feng J H, Xu X B (2005) Endosperm development in autotetraploid rice. Rice Sci 12(2):83-91
38. Pan S J, Reeck G R (1988) Isolation and characterization of riceα-globulin. Cereal Chem 65:316 -319
39. Yamagata H, Sugimoto T, Tanaka K, Kasai Z (1982) Biosynthesis of storage proteins in developing rice seeds. Plant Physiol 70:1094-1100
40. Zhang W W, Bi J C, Chen L M, Zheng L N, Ji S L, Xia Y M, Xie K, Zhao Z G, Wang Y H, Liu L L (2008) QTL mapping for crude protein and protein fraction contents in rice (*Oryza sativa* L.). J Cereal Sci 48(2):539-547
41. Kawakatsu T, Yamamoto M P, Hirose S, Yano M, Takaiwa F (2008) Characterization of a new rice glutelin gene GluD-1 expressed in the starchy endosperm. J Exp Bot 59 (15): 4233-4245
42. Shi C H, Ge G K, Wu J G, Ye J, Wu P (2006) The dynamic gene expression from different genetic systems for protein and lysine contents of *indica* rice. Genetica 128:297-306
43. Shewry P R (2007) Improving the protein content and composition of cereal grain. J Cereal Sci 46(3):239-250
44. Wu X B, Liu J X, Li D Q, Liu C M (2016) Rice caryopsis development Ⅱ: Dynamic changes in the endosperm. J Integr Plant Biol 58(9):786-798
45. Zhou S R, Yin L L, Xue H W (2013) Functional genomics based understanding of rice endosperm development. Curr Opin Plant Biol 16(2), 236-246
46. Xu H H, Liu S J, Song S H, Wang R X, Wang W Q, Song S Q (2016) Proteomics analysis reveals distinct involvement of embryo and endosperm proteins during seed germination in dormant and non-dormant rice seeds. Plant Physio Biochem 103:219-242
47. Chen P L, Shen Z K, Ming L C, Li Y B, Dan W H, Lou G M, Peng B, Wu B, Li Y H, Zhao D, Gao G J, Zhang Q L, Xiao J H, Li X H, Wang G W, He Y Q (2018) Genetic basis of variation in rice seed storage protein (albumin, globulin, prolamin, and glutelin) content revealed by genome-wide association analysis. Front Plant Sci 9:612
48. Zhu J P, Ren Y L, Wang Y L, Liu F, Teng X, Zhang Y Y, Duan E C, Wu M M, Zhong M S, Hao Y Y, Zhu X P, Lei J, Wang Y F, Yu Y F, Pan T, Bao Y Q, Wang Y H, Wan J M (2019) OsNHX5-mediated Ph homeostasis is required for post-Golgi trafficking of seed storage proteins in rice endosperm cells. BMC Plant Biol 19:295
49. Hanashiro I, Itoh K, Kuratomi Y, Yamazaki M, Igarashi T, Matsugasako J I, Takeda Y (2008) Granule-bound starch synthase I is responsible for biosynthesis of extra-long unit chains of amylopectin in rice. Plant Cell Physiol 49(6): 925-933
50. Cheng W M, Liu B M, Ye Y F, Tao L Z, Wu Y J (2016) Screening of Amylose and Protein Mutants and Correlation Analysis of Agronomic Character in Rice[J]. Acta Laser Biology Sinica 25(04):356-361
51. Zou Q F, Liu Y B, Pan X Y, Zhu J, Xie P (2001) Genetic correlation analysis of yield and quality characters of rice (*Oryza sativa* L.) in Multiple Environments[J]. Acta Agriculturae Universitis Jiangxiensis 23(01):16-23

52. Röst H L, Rosenberger G, Navarro P, Gillet L, Miladinović S M, Schubert O T, Wolski W, Collins B C, Malmström J, Malmström L, Aebersold R (2014) OpenSWATH enables automated, targeted analysis of data-independent acquisition MS data. Nat Biotechnol 32(3), 219-223
53. Li Y J, Fang S H, Zhang H S, Shen J C, Liu X D (2011) Genetic Analysis on Agronomic Traits in Autotetraploid Rice. Chinese Agricultural Science Bulletin 27(03):56-59
54. Liu X, Zhang C C, Wang X R, Liu Q Q, Yuan D Q, Pan G, Sun S S M, Tu J M (2016) Development of high-lysine rice via endosperm-specific expression of a foreign gene[J]. BMC Plant Biol 16(1):147
55. Conesa A, Götz S, García-Gómez J M, Terol J, Talón M, Robles M (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21(18), 3674-3676
56. Langfelder P, Horvath S (2008) WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics 9(1): 559

Tables

Table 1 Table 1 The traits difference for AJNT-4x and AJNT-2x

	Seed setting rate (%)	Grain length (mm)	Grain width (mm)	Grain thickness (mm)	Grains weight (g/1000 grains)
AJNT-2x	71.18±0.04	8.73±0.30	3.19±0.13	2.06±0.09	28.95±0.25
AJNT-4x	49.94±0.04	10.48±0.35	3.61±0.14	2.44±0.10	42.05±0.65**

"**" indicates high significant difference at the P=0.01 level.

Table 2 Differentially expressed proteins (DEPs) between AJNT-4x and AJNT-2x

Comparisons	No. of upregulated proteins	No. of downregulated proteins	Total DEPs
2n-10/2n-15/2n-20	__	__	185
4n-10/4n-15/4n-20	__	__	372
4n-10/2n-10	86	39	125
4n-15/2n-15	85	74	159
4n-20/2n-20	35	43	78

Table 3 Real-time quantitative PCR analysis for thirteen DEPs.

Protein	Gene	Up-/downregulated	DAF	iTRAQ (A/D)	qRT-PCR (A/D)	Go Annotation
B8AM24	*OsI_11183*	Up	15	1.64	3.39	Copper ion binding; L,L-diaminopimelate aminotransferase activity; pyridoxal phosphate binding
B8ARJ0	*OsI_15102*	Up	15	1.81	1.17	Lysine biosynthetic process via diaminopimelate
P55857	*SUMO1*	Up	15	1.45	2.32	Ubiquitin-like protein ligase binding
Q40689	*Gt2*	Down	15	0.73	0.10	Glutelin
A1YQG5	N/A	Down	15	0.04	-0.49	Glutelin
B8AQM6	*OsI_11879*	Up	15	1.45	1.26	Regulation of cell growth by extracellular stimulus; brassinosteroid mediated signaling pathway
A0A0N7KMN9	*Os06g0701100*	Up	15	1.54	-0.47	ATP binding; nucleic acid binding
B8AJS5	*OsI_11975*	Up	15	1.53	-0.32	Cell redox homeostasis
P37833	*Os01g0760600*	Up	15	1.85	0.51	Aspartate aminotransferase; cytoplasmic; involved in nitrogen metabolism and in aspects of carbon and energy metabolism
A2ZCE6	*OsI_35452*	Up	15	1.95	4.55	Disulfide-isomerase; related with endosperm development
A2XNE7	*OsI_14085*	Up	15	2.07	1.81	Metallopeptidase activity; zinc ion

						binding
A3AI97	*OsJ_10966*	Down	10	0.68	0.74	Xylan biosynthetic process
Q42971	*ENO1*	Up	10	2.10	11.01	Catalytic activity

Notes: ubiquitin was the control protein and was analyzed in triplicate.

Figures

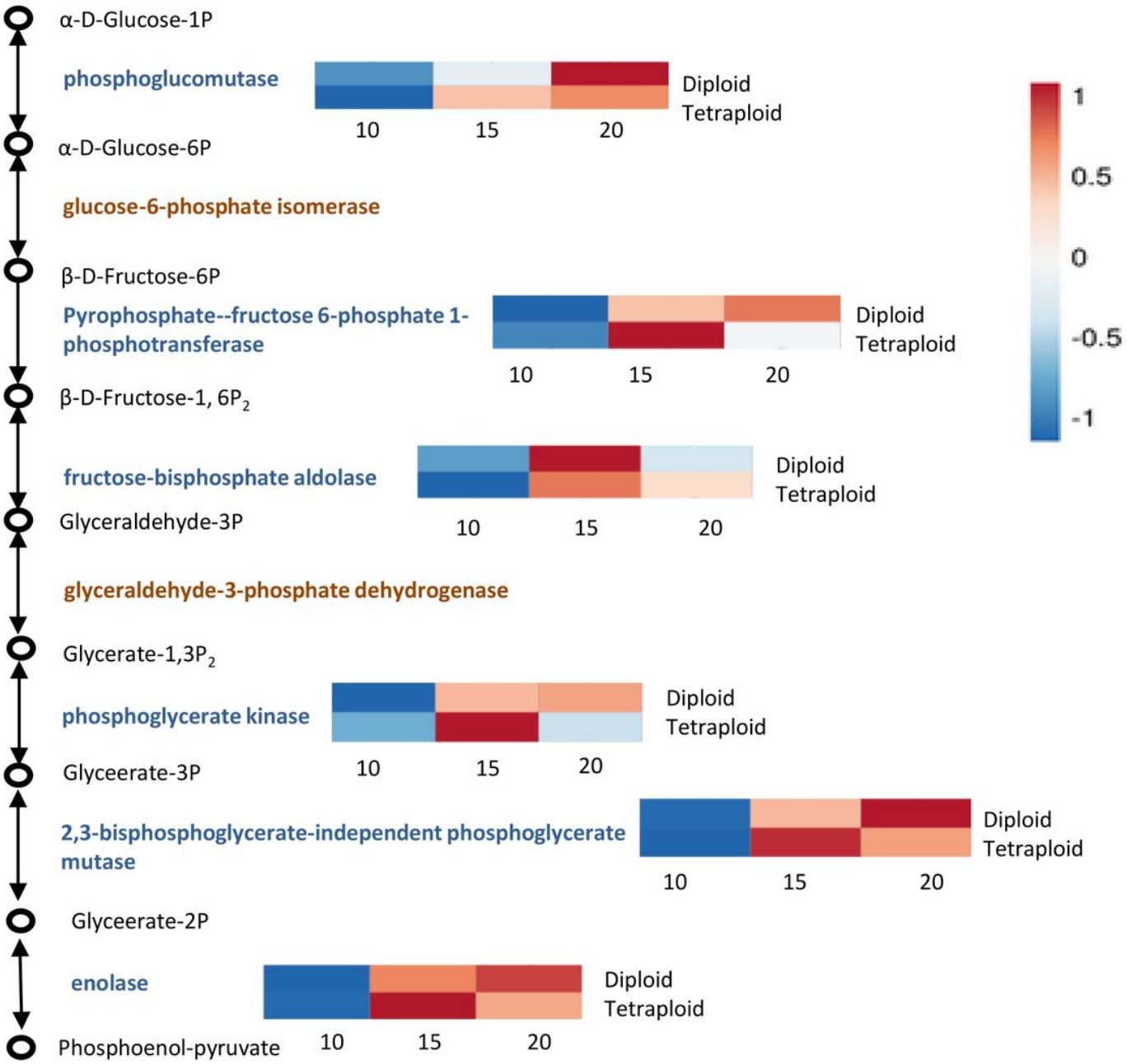


Figure 1

Glycolysis pathway. Each circle represents a metabolite, and proteins have been identified in blue according to the associated heatmap located on the right. In the heatmap, the top row, labeled Diploid, represents AJNT-2x, and the bottom row, labeled Tetraploid, represents AJNT-4x. Each column represents a trait. The heatmap is color-coded by correlation according to the color legend.

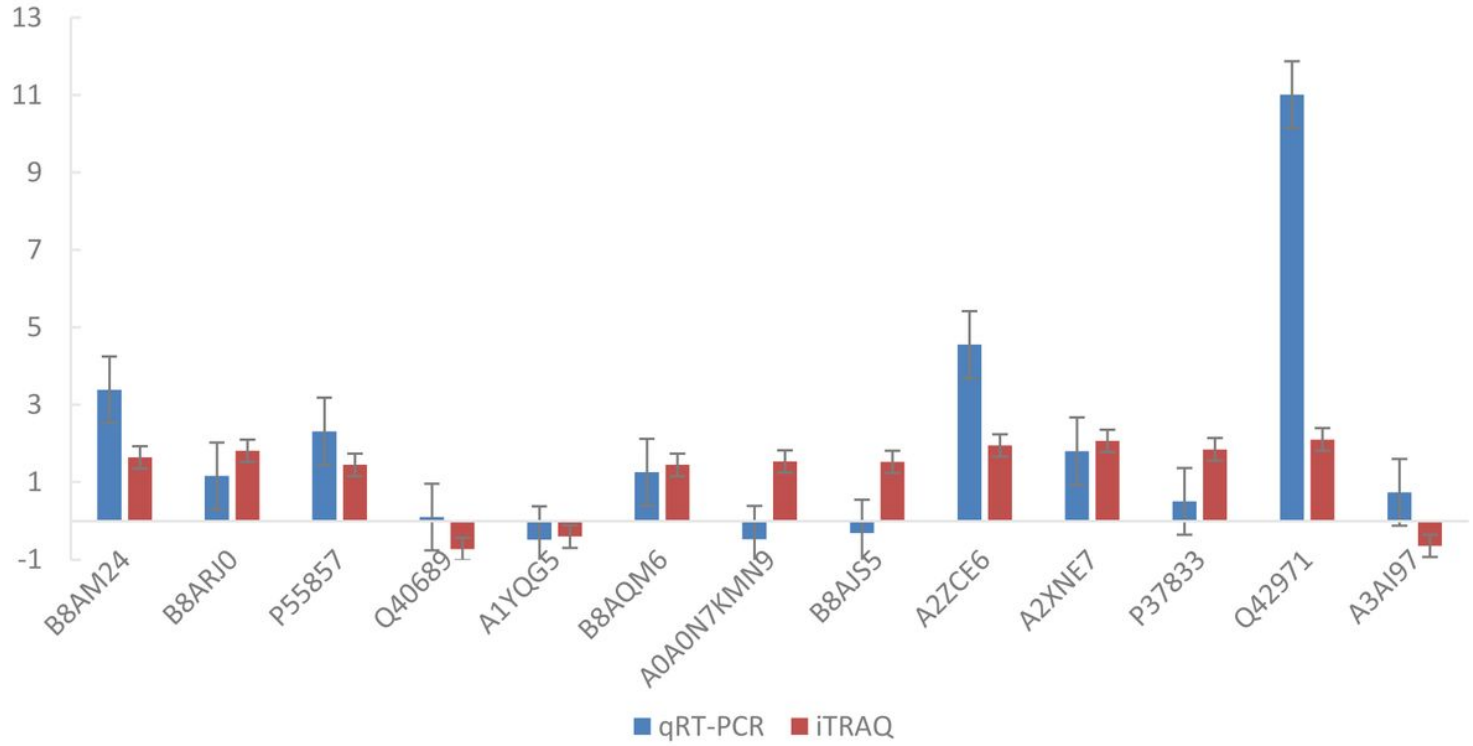


Figure 2

Expression difference of DEPs between AJNT-4x vs AJNT-2x Blue indicates qRT-PCR analysis of DEPs; red indicates iTRAQ analysis of DEPs. Ubiquitin acted as the control protein.

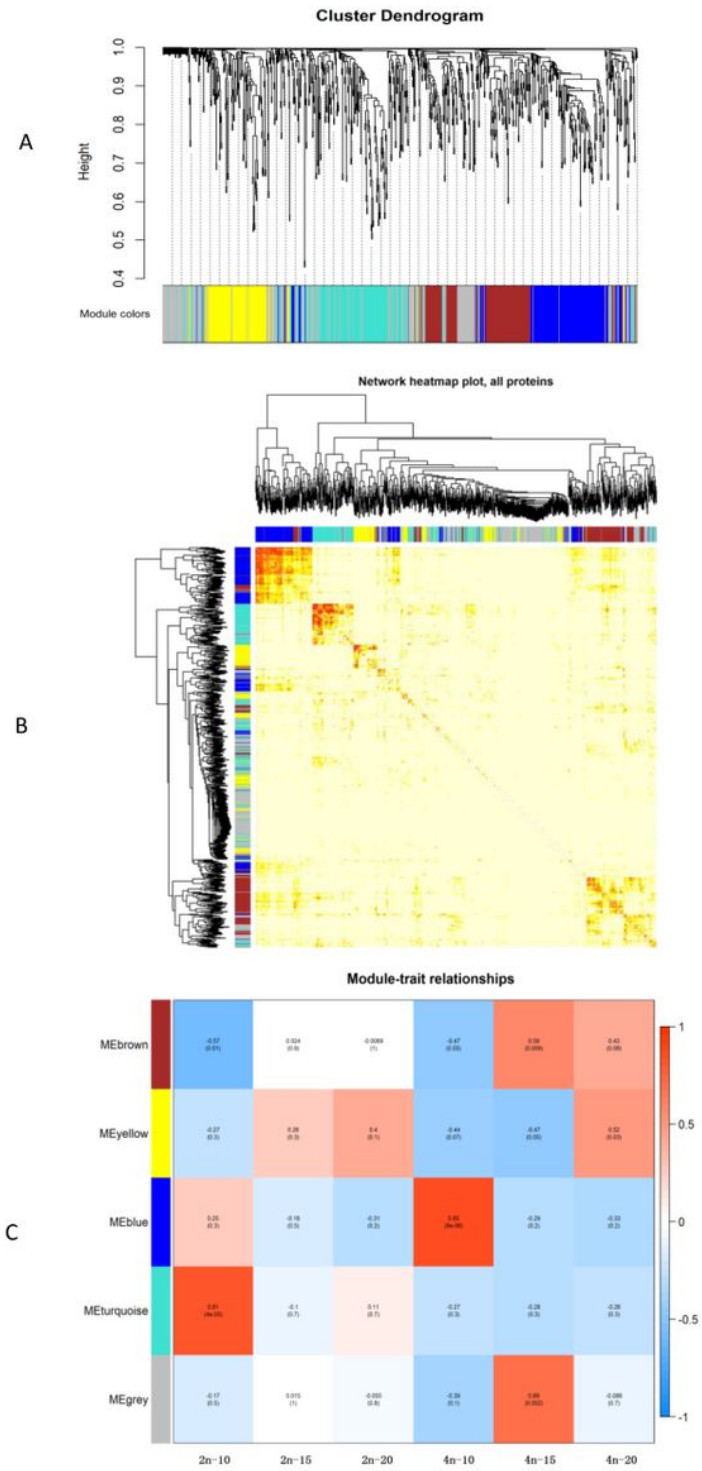


Figure 3

Weighted gene co-expression network analysis (WGCNA) of DEPs. A, cluster dendrogram; B, network heatmap plot of all proteins; C, module-trait relationships. The brown module constitutes proteins relation to carbohydrate metabolism; the yellow, blue and gray modules constitute ribosomal and oxidative phosphorylation-related proteins; the turquoise module constitutes proteins related to the proteasome and fatty acid biosynthesis.

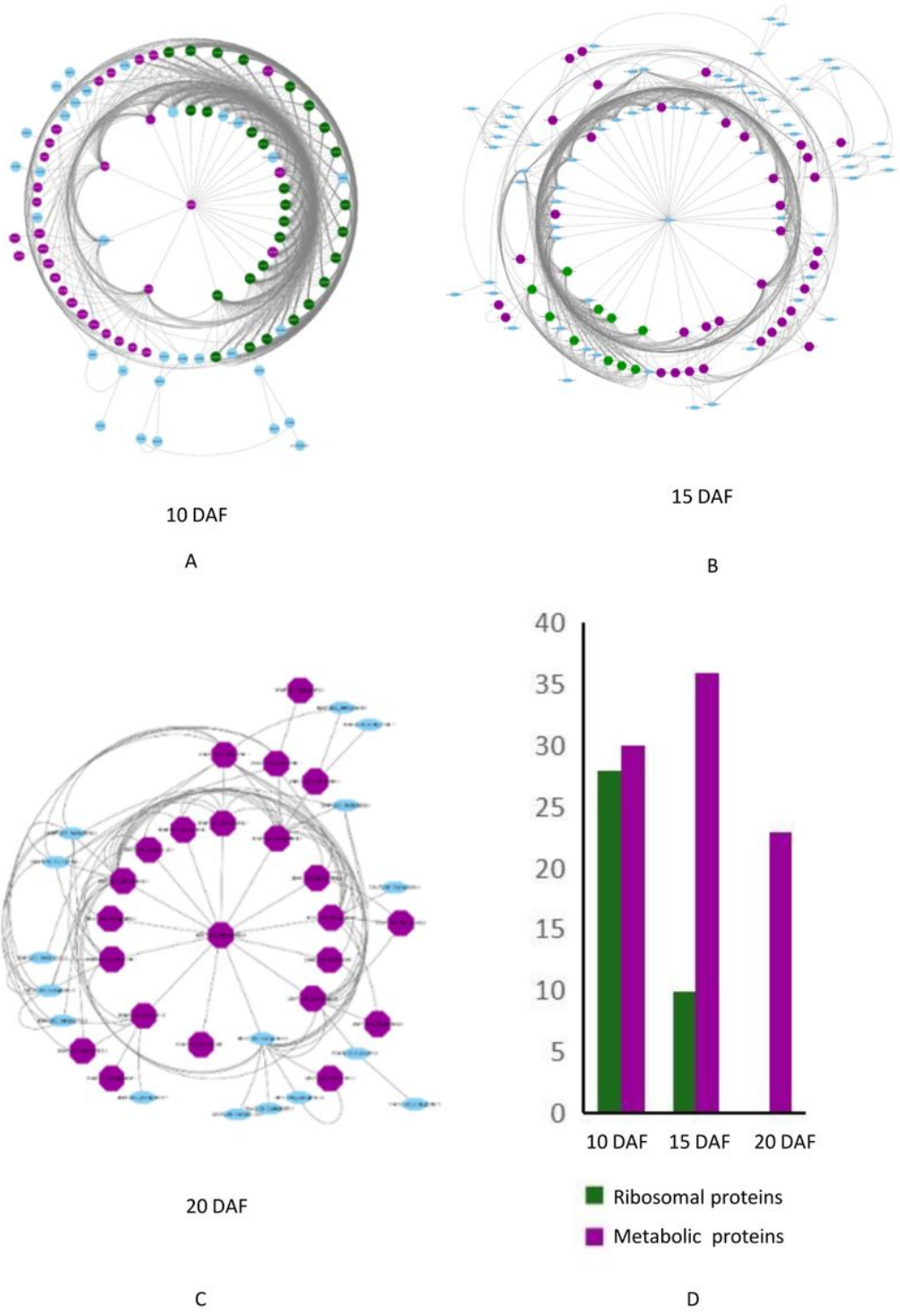


Figure 4

Protein-protein interaction (PPI) networks of DEPs at 10, 15, and 20 DAF between AJNT-4x and AJNT-2x. A, PPI network of DEPs at 10 DAF between AJNT-4x and AJNT-2x; B, PPI network of DEPs at 15 DAF between AJNT-4x and AJNT-2x; C, PPI network of DEPs at 20 DAF between AJNT-4x and AJNT-2x; D, bar chart of protein types in the PPIs of AJNT-4x vs AJNT-2x at 10, 15, and 20 DAF. Green circles represent ribosomal proteins, and purple circles represent metabolic proteins.

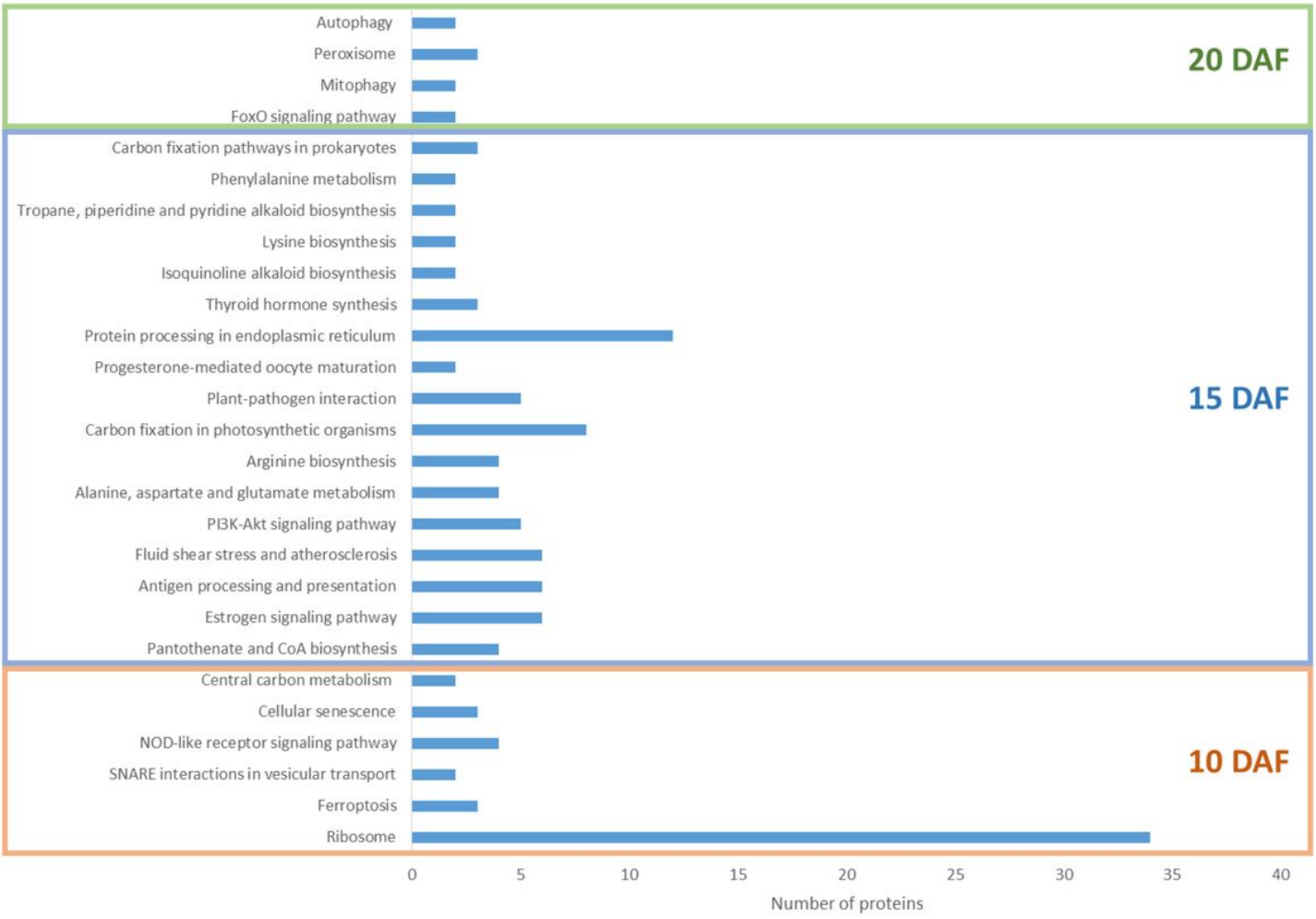
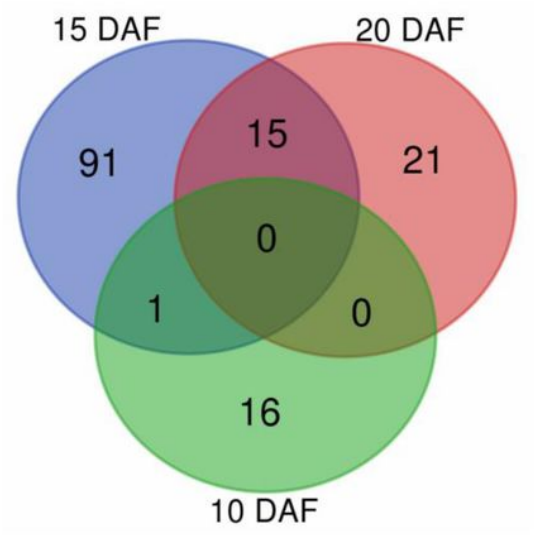
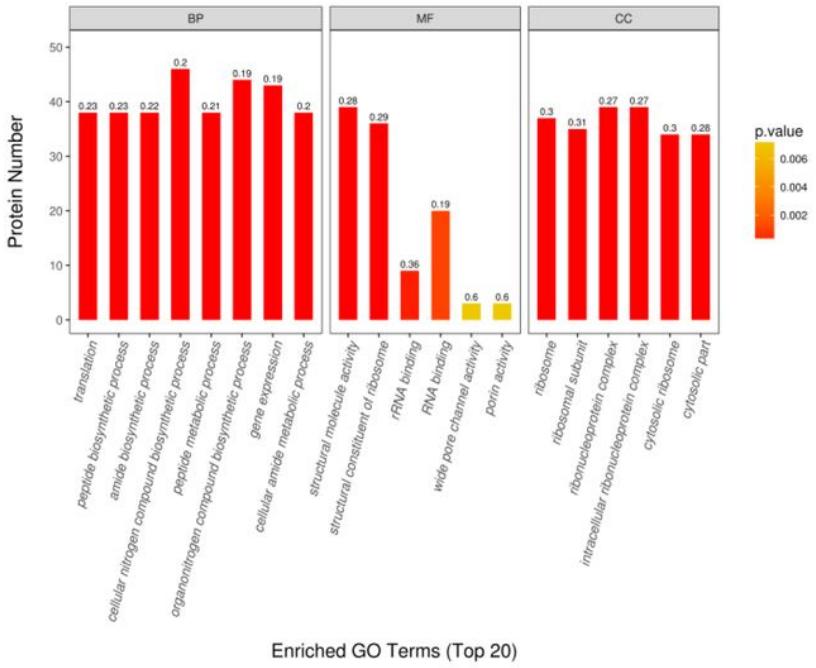


Figure 5

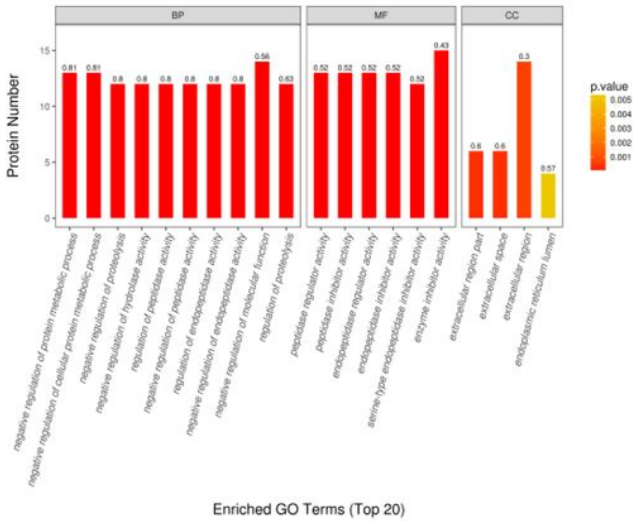
Enriched KEGG pathways for AJNT-4x vs AJNT-2x at 10, 15, and 20 DAF; the blue bar length represents the number of proteins. The orange box indicates comparisons at 10 DAF between AJNT-4x and AJNT-2x; The blue box indicates comparisons at 15 DAF between AJNT-4x and AJNT-2x; The green box indicates comparisons at 20 DAF between AJNT-4x and AJNT-2x.



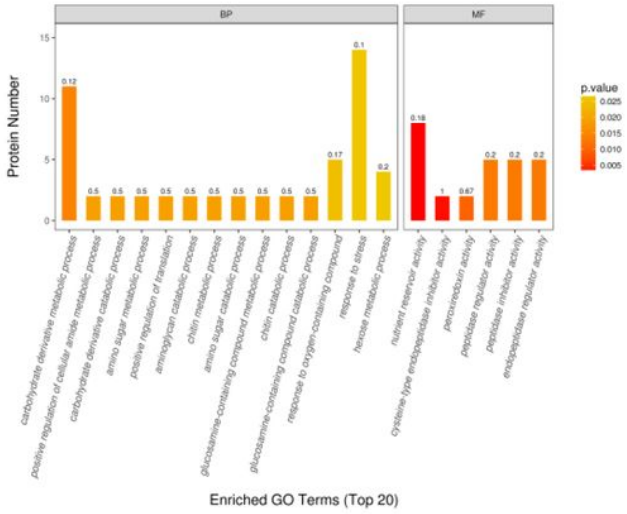
A



B



C



D

Figure 6

DEP analysis between AJNT-4x and AJNT-2x. A, Venn diagram of enriched Gene Ontology (GO) functions of AJNT-4x vs AJNT-2x at 10, 15, and 20 DAF; B, enriched GO functions at 10 DAF between AJNT-4x and AJNT-2x; C, enriched GO functions at 15 DAF between AJNT-4x and AJNT-2x; D, enriched GO functions at 20 DAF between AJNT-4x and AJNT-2x. BP: biological process; MF: molecular function; CC: cell component.

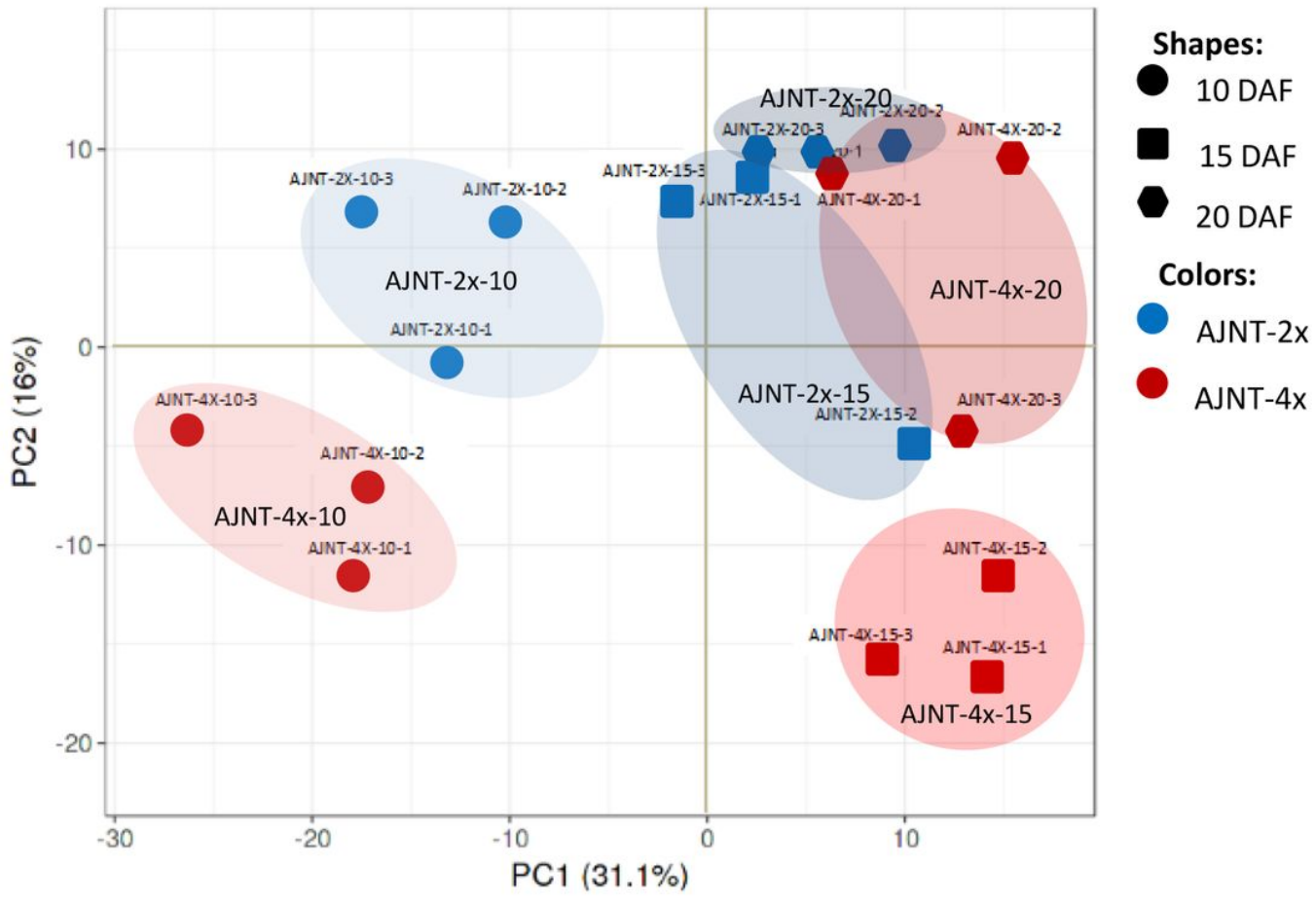


Figure 7

PCA score plot of AJNT-4x (red) and AJNT-2x (blue) endosperm proteins at 10, 15 and 20 DAF. Blue shapes indicate AJNT-2x; red shapes indicate AJNT-4x; circles indicate differentially expressed proteins (DEPs) at 10 DAF between AJNT-4x and AJNT-2x; squares indicate DEPs at 15 DAF between AJNT-4x and AJNT-2x; hexagons indicate DEPs at 20 DAF between AJNT-4x and AJNT-2x.

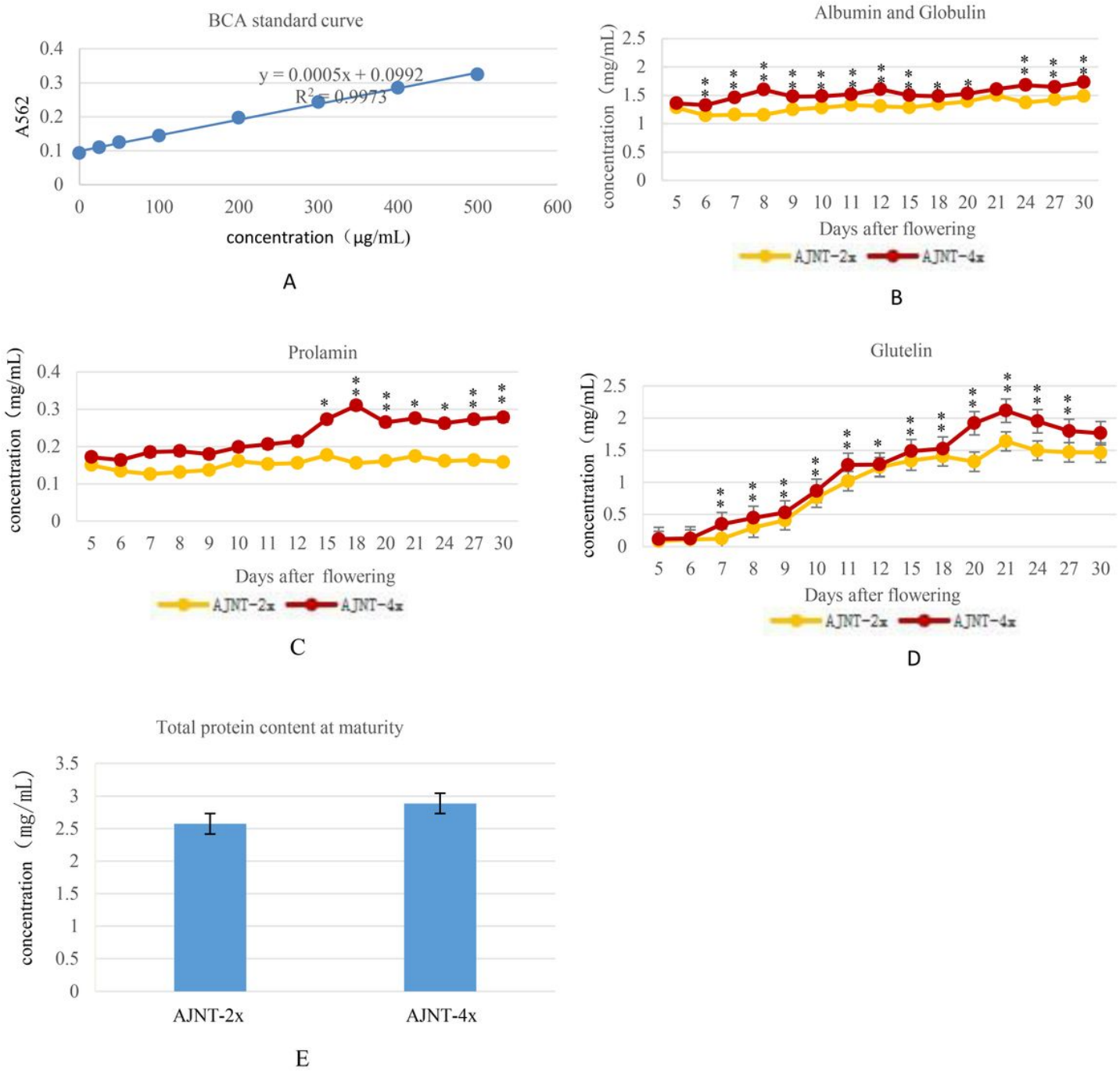


Figure 8

Endosperm protein content analysis of different developmental stages between Aijiaonante (AJNT)-4x and AJNT-2x A, standard curve; B, albumin and globulin content variation; C, prolamin content variation; D, glutelin content variation; E, total storage protein content at maturity. “*” indicates a significant difference at the P=0.05 level; “**” indicates a highly significant difference at the P=0.01 level.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementTableandFigures.rar