Jestr

Journal of Engineering Science and Technology Review 17 (5) (2024) 135-144

Research Article

JOURNAL OF
Engineering Science and Technology Review

www.jestr.org

Differential Proteomics of Endosperm During Dormancy Release of *Tilia miqueliana* Seeds

Jieqing Wang[1,2] and Yongbao Shen[1,*]

[1]*Department of Landscape Plants and Ornamental Horticulture, Nanjing Forestry University,Nanjing, 210037, China*
[2]*Gold Mantis School of Architecture, Soochow University, Suzhou 215123, China*

Received 12 March 2024; Accepted 22 October 2024

Abstract

Seed dormancy affords plants long-term effective resistance to severe environments. Various regulatory factors and protection mechanisms that are involved in seed dormancy and dormancy release and valuable to the use and release of seed dormancy were analyzed from the molecular perspective. *Tilia miqueliana* has a good ability to improve the environment and has a broad application prospect, but its seeds undergo dormancy. The differential proteomics of endosperm during dormancy release of *T. miqueliana* seeds were investigated in this study by combining two-dimensional electrophoresis and mass spectrometry to explore the protein changes and protein–protein interaction during dormancy release of *T. miqueliana* seeds and thereby disclose the molecular mechanism of dormancy release. A total of 16 differentially expressed proteins were identified, most of which were related to the activity of the nutrient bank, programmed plant cell death, and redox reactions. Subsequently, bioinformatics analysis of protein-related data was carried out. The results demonstrate that in the key stage of dormancy release of *T. miqueliana* seeds, the number of proteins, which have oxidoreductase activity, antioxidant activity, and combined enrichment with transition metal ions, is the highest in molecular functions.They mainly participate in metabolism processes in organelles, the mitochondria, peroxisome, and cell catabolism. According to the results of Kyoto Encyclopedia of Genesand Genomes(KEGG) pathways, some pathways, including glycolysis/gluconeogenesis, biosynthesis of secondary metabolites, biosynthesis of amino acids, and glyoxylate and dicarboxylate metabolism, play a dominant role during dormancy release of *T. miqueliana* seeds. Finally, important proteins (genes) during the dormancy release of *T. miqueliana* seeds were screened through weighed gene co-expression network analysis (WGCNA), including SPS1, SDRA, ENO2, At1g64255, At4g17670, and At2g28490. This study provides a technical method and data support for the dynamic analysis of proteomics during the process of seed dormancy release, and the research results further elucidated the molecular mechanism of dormancy release of *T. miqueliana* seeds.

Keywords: *Tilia miqueliana*, Dormancy release of seeds, Proteomics, Bioinformatics

1. Introduction

Seed dormancy is a biological characteristic of adapting to environmental changes that plants acquired for survival and evolution of species [1,2]. However, such seeds could only be germinated after dormancy release, which bring many barriers against agricultural and forestry production and forest breeding. Understanding the dormancy release mechanism of seeds is conducive to the dormancy release of seeds by using appropriate methods.

Seeds with deep physiological embryo dormancy could be germinated through various processing stimuli including, laminating cold storage [3,4]. The hormone imbalance between germination inhibitor (e.g., abscisic acid) and growth stimulant (e.g. gibberellin) is an important cause of dormancy release of seeds [5–7]. Dormancy release of seeds includes a series of physiological processes, including breaking the dormancy mechanism of seeds, triggering metabolism in seeds, and regulating germination time[8,9]. Protein is the executor of physiological functions. Studying protein changes during the dormancy release of seeds under certain conditions reflects not only the gene expression levels when genomes are activated but also the changes in intra-seed physiological metabolism from relative static state to active state. The results are of important significance to release the dormancy of seeds and discuss the relevant molecular mechanism.

Proteome refers to the set of proteins contained in a specific research object, such as the set of proteins encoded by genomes of specific species or the set of proteins expressed by specific tissues, organs, and cells. The method that investigates proteomes is called proteomics [10,11]. This technology provides a good means to study dynamic changes in proteins during the dormancy release of seeds, and the core technology includes two-dimensional gel electrophoresis (2-DE) and mass spectrometry (MS) [12,13]. Research on the dormancy release mechanism of seeds based on proteomics technology has been applied successfully to various plants. For example, Rashid analyzed the protein composition of the aleurone layer of Arabidopsis thaliana through 2-DE electrophoresis and liquid chromatography-tandem mass spectrometry (LC-MS) [14]. Su He identified 550 nonredundant proteins through the 2-DE method by studying differential proteomics during the dormancy release of *Epimedium wushanense* seeds, which were mainly related to metabolism [15]. Lan Hai et al. analyzed the proteomic differential expression of seeds of strong-dormancy maize inbred line 08-641 in dormancy and after dormancy release. The detected differentially expressed proteins are mainly divided into three classes: storage

*E-mail address: ybshen@njfu.com.cn
ISSN: 1791-2377
doi:10.25103/jestr.175.18

proteins, proteins participating in substance metabolism, and proteins participating in protein structure and cell functional regulation. Excessive expression or deficiency of these proteins is speculated to inhibit normal germination of seeds [16]. Hybrid-2D-linear quadrupole ITFT-ICR (LTQ-FTICR) MS [17] was applied during protein identification of *Acer platanoides*, *Firmiana platanifolia*, and *Fagus sylvatica* seeds [17]. MALDI-TOF MS was applied to analyze the dormancy of *A. thaliana* seeds [18]. The dormancy of *Prunus campanulatal* seeds was also investigated by MALDI-TOF MS/MS and ESI-MS/MS [19]. In addition, many scholars have investigated proteomes in seeds of soybean, rice, *Buchloe dactyloides*, wheat, and *Fagopyrum esculentum* [20-24]. However, most of these studies identified differentially expressed proteins during dormancy release of seeds and analyzed the functions of specific proteins. The results demonstrated that energy metabolism, proteasome, transcription, protein synthesis, signal transduction, and methionine metabolism protein have unique importance. However, none of these studies have carried out bioinformatics analysis in depth by combining relevant data.

Many studies on the dormancy characteristics and germination mechanism of seeds have been recently reported and achieved some outcomes. The dormancy and germination mechanisms of seeds of some plants have been disclosed clearly. Existing results have provided references for scientific dormancy release of seeds. However, the dormancy and germination mechanisms of seeds are conservative, to some extent, in different plants. Due to some specificity between monocotyledon and dicotyledon plants and among different plants, the dormancy mechanism of seeds of each plant species must be studied specifically.

Tilia miqueliana Maxim. (Tiliaceae) belongs to the deciduous tree, which has beautiful structure, good appreciation, and strong stress tolerance to harmful gases, and can improve the environment very well. It is not only a good honey crop; it could also be used as a medicinal material and material. Moreover, due to the deep origin with Buddhist culture, it is called as the “Bodhi tree” [25,26]. *T. miqueliana* seeds have deep dormancy characteristics that restrict the industrialization process of seedling breeding significantly.

Existing studies on *T. miqueliana* seeds mainly focused on the causes of dormancy and dormancy release methods. Shi and Wu et al. believed that the inhibitors and mechanical barriers of shells and testa of *T. miqueliana* seeds, endosperm inhibitors, hormone regulation, and material metabolism are all causes of seed dormancy, and it is comprehensive dormancy type [27–29]. Wang et al. concluded from a research that the dormancy of *T. miqueliana* seeds belongs to physical and physiological dormancy, and the mechanical barrier of endosperm cap and the permeability of the outer cells of testa and endosperm are major causes of dormancy of *T. miqueliana* seeds[30]. Yang investigated various physiological and biochemical changes in endosperm during fast dormancy release of *T. miqueliana* seeds and recognized the optimal dormancy release method[31]. To sum up, although studies on the dormancy release of *T. miqueliana* seeds have achieved breakthrough processes, the molecular mechanism remains unknown.

In this study, changes in the differential proteomics of endosperm during dormancy release of *T. miqueliana* seeds were investigated using proteomics technology and bioinformatics analysis on the basis of previous research. The key proteins during dormancy release of *T. miqueliana* seeds were screened by weighed gene co-expression network analysis (WGCNA) method to discuss the relevant molecular mechanism.

2. Materials and Methods

2.1 Research materials

The seeds of *T. miqueliana* were collected from fruits that fell naturally onto the ground in Anhui Huangzangyu National Forest Park in China. In the experiment, the shells of seeds were eliminated, and full seeds were chosen as experimental materials.

2.2 Experimental methods

2.2.1 Seed treatment

The dormancy release of *T. miqueliana* seeds was realized with concentrated sulfuric acid. The first sampling was performed on the day of seed treatment, and 3 × 150 seeds were collected randomly and the endospermwas stripped for protein extraction, the extracted protein was frozen at - 80℃. The endosperm was collected to extract proteins, which were then frozen at －80 ℃. Subsequently, the samples were collected every 15 d. Sampling was performed at about 30 and 60 d of cold stratification to avoid missing the key stages during the dormancy release of *T. miqueliana* seeds. The sampling times were set at 0, 15, 30, 35, 40, 45, 60, and 70 d. Random sampling was performed, and the endosperm proteins of *T. miqueliana* seeds were extracted for 2-DE.

2.2.2 2-DE of endosperm proteins

The endosperm proteins of *T. miqueliana* seeds were extracted by phenol extraction method, which was appropriately improved with references to the method of Isaacson et al. [32]. The protein concentration was tested by referring to the method of Bradford [33], and 2-DE was performed using PH3-10, 24 cm-nonlinear adhesive tapes (Fig.1). The isoelectric focusing parameters included temperature (20℃) and maximum current (50 μA/adhesive tapes). The temperature and electrophoresis of SDS-PAGE electrophoresis water bath circulator were set 15 ℃ and 100 V×45 min, respectively. The SDS-PAGE electrophoresis was continued under 200 V until bromophenolan’s leading edge ran out of gel. Gel dying was performed using Coomassie brilliant blue staining method, and samples were scanned after de-coloration by using the GS-800 calibrated densimeter (Bio Rad Company, USA). A matched analysis of the gained images was carried out using the PDQuest software (version 8.0, Bio Rad Company, USA).

2.2.3 Mass spectral identification and bioinformatics analysis

The endosperm proteins of *T. miqueliana* seeds at 0 d were used as the control group, and the spectra at different stages were compared. Protein spots with differences over two times and $P < 0.05$ were viewed differentially and chosen for MS. The Q Exactive Plus-combined quadrupole Orbitrap mass spectrometer (Thermo Scientific, Fig. 2) was used for analysis. The acquired level-1 and level-2 mass data were integrated. Proteome Discoverer 2.3 (Thermo Fisher) was used for the analysis of MS data and protein identification. Given that *T. miqueliana* is a non-modal species, homogenous mapping was conducted onto the closest model

species, namely, A. thaliana (L.), followed by bioinformatics analysis.

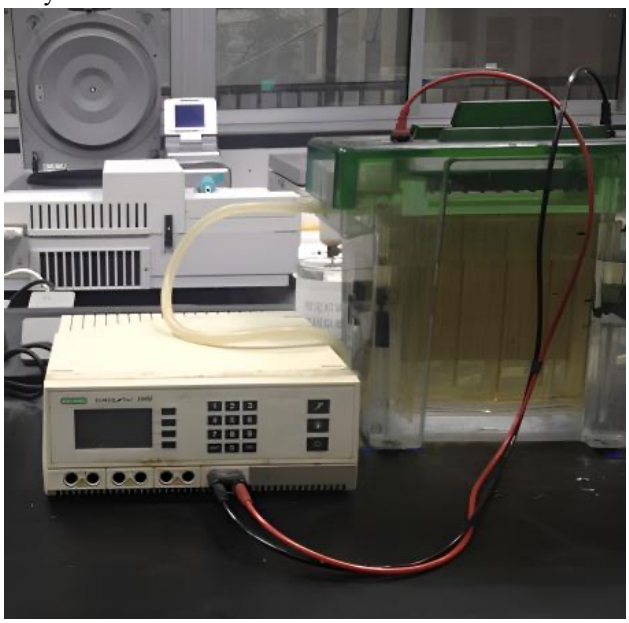


Fig. 1. SDS-PAGE electrophoresis system

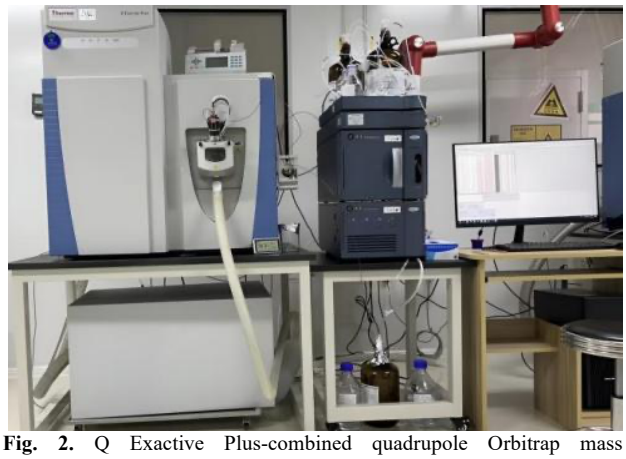


Fig. 2. Q Exactive Plus-combined quadrupole Orbitrap mass spectrometer

2.2.4 Screening of important proteins based on WGCNA

WGCNA is a typical systematic biological algorithm that establishes the gene co-expression network. Based on the high-throughput expression data, the WGCNA algorithm first hypothesizes that the gene network obeys to scale-free network, defines the gene co-expression correlation matrix and adjacent function formed by the gene network, and then calculates the coefficient of difference of different nodes and recognizes the gene set module related to seed dormancy[34,35]. The protein expression spectra at each timepoint were analyzed using the WGCNA package[36]version 1.61 in R R3.4.1, thus obtaining the correlations of each protein and module with dormancy release of seeds. The correlation formula is as follows:

The Pearson correlation coefficient of each protein pair (m and n) could be calculated as $S_{mn} = |cor_{(m,n)}|$. On this basis, the protein co-expression correlation matrix was formed.

The adjacent function of power function is as follows: $a_{mn} = power_{(S_{mn},\beta)}$

The correlation matrix S_{mn} was transformed into the adjacent matrix a_{mn}. The relations of each gene and all other genes in the analysis were included into consideration, and the adjacent matrix was transformed into the topology matrix $\Omega = w_{mn}$ [13]. The elements in the matrix were expressed as follows:

$$w_{mn} = \frac{l_{mn} + a_{mn}}{min\{k_m, k_n\} + 1 - a_{mn}}, \quad (1)$$

where l_{mn} represents the sum of area of adjacency coefficient of node, which is connected with genes m and n;km is the sum of adjacencies of the node, which is connected independently with the gene m.

3. Results analysis

3.1 Mass spectral identification of the differentially expressed proteins of the endosperm during dormancy release of *T. miqueliana* seeds

2-DE was performed three times on the endosperm proteins of the *T. miqueliana* seeds. The repeatability of the test was ensured by extracting these proteins at different stages, and 24 images of the 2-DE atlas were obtained. Different groups of protein expression atlas were similar, and the matching rate of a repeated adhesive was higher than 90%. The 2-DE atlas of the samples collected on July 21 was used as the control group, and protein spots with a mean differential expression of over two times in three repeated atlases were marked onto the overlapping atlas. The spots were numbered by sampling time as follows: 0d, 15d, 30d, 35d, 40d, 45d, 60d, and 70d. The 2D electrophoretograms of proteins in eight periods were analyzed using PD Quest version 8.0.1 software, and a total of 101 differentially expressed protein spots with abundance variation of over two times, P of <0.05, and good repeatability were detected. Through MS detection, 49 protein spots were matched successfully, and 16 differentially expressed proteins were acquired, which mainly included storage, programmed cell death (PCD)-related, and defense-related proteins.

Protein expression profiles with significantly similar expression patterns over time were clustered at eight different time points with STEM version 1.3.11 software (significant similarity threshold: P < 0.05). Three clusters with significantly similar expression patterns were obtained, including 18, 8, and 5 proteins. The protein expression trend in set 1 continuously decreased with time (Fig. 3A). In the cluster heat map, the color gradually changed from red to light and then to dark green, mainly including MuDR family transposase (At1g64255), vicilin-like seed storage protein At2g28490 (At2g28490), P-loop-containing nucleoside triphosphate hydrolases superfamily protein (At3g12020), senescence-associated family protein (At4g17670), cation/calcium exchanger 1 (AtCCX1; CCX1), probable mediator of RNA polymerase II transcription subunit 37c (heat shock 70 kDa protein 4; MED37C), short-chain dehydrogenase/reductase (SDRA), and sucrose phosphate synthase 1 (SPS1).

The expression trend of proteins in set 2 (Fig. 3B) continues to increase over time. As shown in the clustering heat map, the color of these proteins gradually changes from green to light, and then to gradually dark red. These proteins mainly include vicilin-like seed storage protein At3g22640 (PAP85), 12S seed storage protein CRD (CRD), catalase (CAT2), KDEL-tailed cysteine endopeptidase CEP1(CEP1), glyceraldehyde-3-phosphate dehydrogenase GAPC2 (GAPC2), and At2g28490.

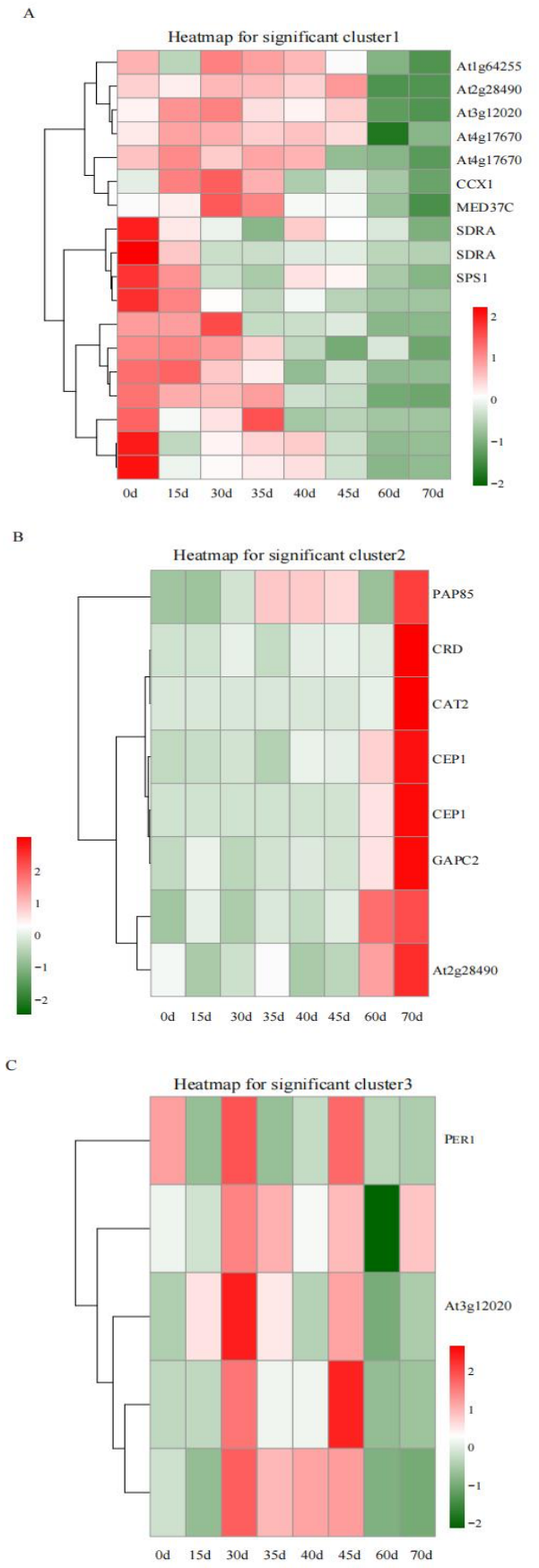


Fig. 3. Clustering maps of protein expression levels in three clusters with significant similarity in expression obtained by STEM
Note: A–C represents three clustering maps of proteins in different clusters

3.2 Gene Ontology (GO) functional annotation of differentially expressed proteins

The significantly correlated GO annotation classification of proteins with significant variation in expression levels was carried out using DAVID 6.8 [37], including biological process (BP), cell composition (CC), and molecular function (MF, $P < 0.05$ was used as the significance threshold). A total of 253 annotations were acquired. Specifically, BP accounted for 32.02% (81); CC, for 34.78% (88); and MF, 33.20% (84). The results(Fig. 4) were obtained by sequencing the enrichment analysis results of BP, CC, and MF from left to right in accordance with the significance of differences and the summary of information of the top 10 classes. Fig. 4 shows that organelles and intracellular organelles in CC had the highest number of enriched differentially expressed proteins . Significant enrichment in BP was the response to inorganic matter. In CC, the driving protein compound had the most significant enrichment. In MF, nutrient bank activity had the most significant enrichment of differences.

As can be seen from Fig. 5A, there were 14 BPs significantly enriched for differential proteins during the first 15 days of seed treatment, including developmental programmed cell death, developmental maturation, and cell development. Developmental maturation was the most significant, involving proteins, such as the defense-related protein PER1, SDRA, and seed storage protein CRD.

The MFs in which proteins are annotated and significantly enriched at this stage are: nutrient reservoir activity, cysteine-type endopeptidase activity, thioredoxin thioredoxin peroxidase activity (Fig.5B). The nutrient reservoir activity was the most significant, and the involved proteins were seed storage proteins, such as At2g18540, CRD, At2g28490, and PAP85. The CCs in which differential proteins were annotated and significantly enriched were plant-type vacuoles, lysosomes, lytic vacuoles, and plant-type cell wall (Fig.5C). The most significant difference was observed in CEP1 and CRD in plant-type vacuoles.

A total of 2 – 19 BPS were annotated and significantly enriched with differential proteins at each stage within 35 days after seed treatment. Developmental programmed cell death showed the most significant difference, and CEP1 was the only protein involved. Its protein expression was always up-regulated in contrast to that in the initial stage of seed treatment and reached the highest value after 35 – 40 days. During the same period, proteins were significantly enriched in lysosomes, plant-type vacuoles, plant-type cell walls, cell walls and extracellular encapsulation structures, and the most significant difference was observed in lysosomes. The related protein was CEP1.The protein in the MF was significantly enriched in nutrient reservoir activity and cysteine-type endopeptidase activity. The proteins involved were seed storage proteins PAP85, CRD, and programmed cell death protein CEP1.

Overall, the differential proteins obtained at each time point were involved in 2 BPs, both of which were related to anther development and anther wall tapetum development. Death, cell death, and programmed cell death were also involved throughout the whole stage of dormancy release.

Within 35 days after seed treatment, the BP with significant differences and the largest amount of enriched proteins were mainly developmental processes, such as unicellular and multicellular organism process, multicellular organism development, single-organism developmental process and developmental processes. The cell components with significant differences and the largest amounts of enriched proteins were: cell periphery, extracellular region

part, vacuole, and cell wall. It mainly performs MFs, such as cysteine-type peptidase activity.

At 35 – 60 days, BP showed significant differences, and a large number of enriched proteins concentrated on BPs related to reproductive structure, such as reproductive system development, reproductive structure development, and single organism reproductive process. They mainly occurred in organelles, organelle lumina, nonmembrane-bounded organelles, mitochondrion, and peroxisomes. MFs, such as antioxidant activity and transition metal ion binding had the highest concentrations of proteins.

Approximately 70 days after seed treatment, the difference was significant. The BP with high protein enrichment was mainly related to developmental process and stress response, such as the development of multicellular organisms and response to abiotic stimulus. In CC, the cell wall, external encapsulating structures, vacuoles, cell junctions, and cell – cell junctions, symplasts, mitochondria, and plasmodesmas had the largest number of enriched proteins, and the significantly different MF was nutrient reservoir activity.

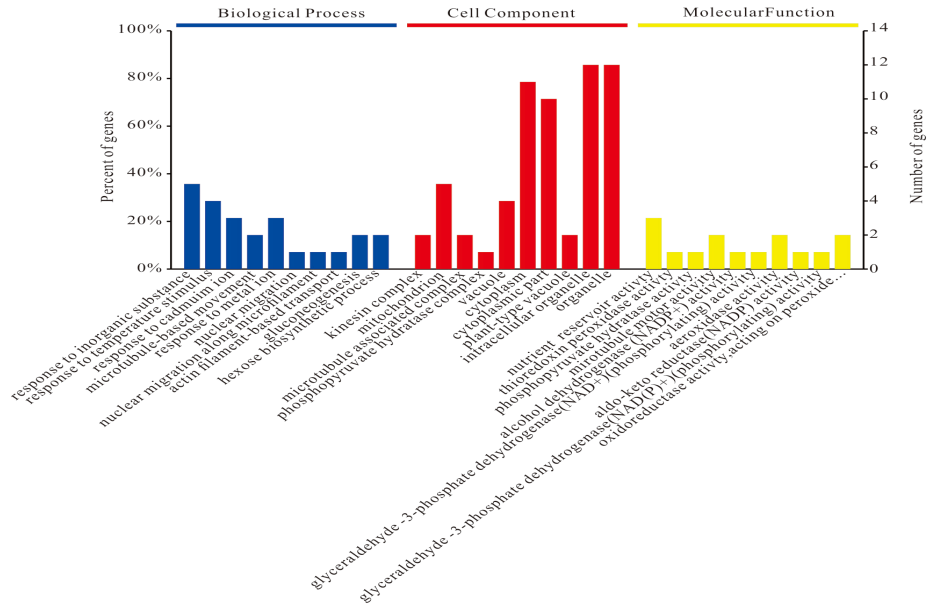
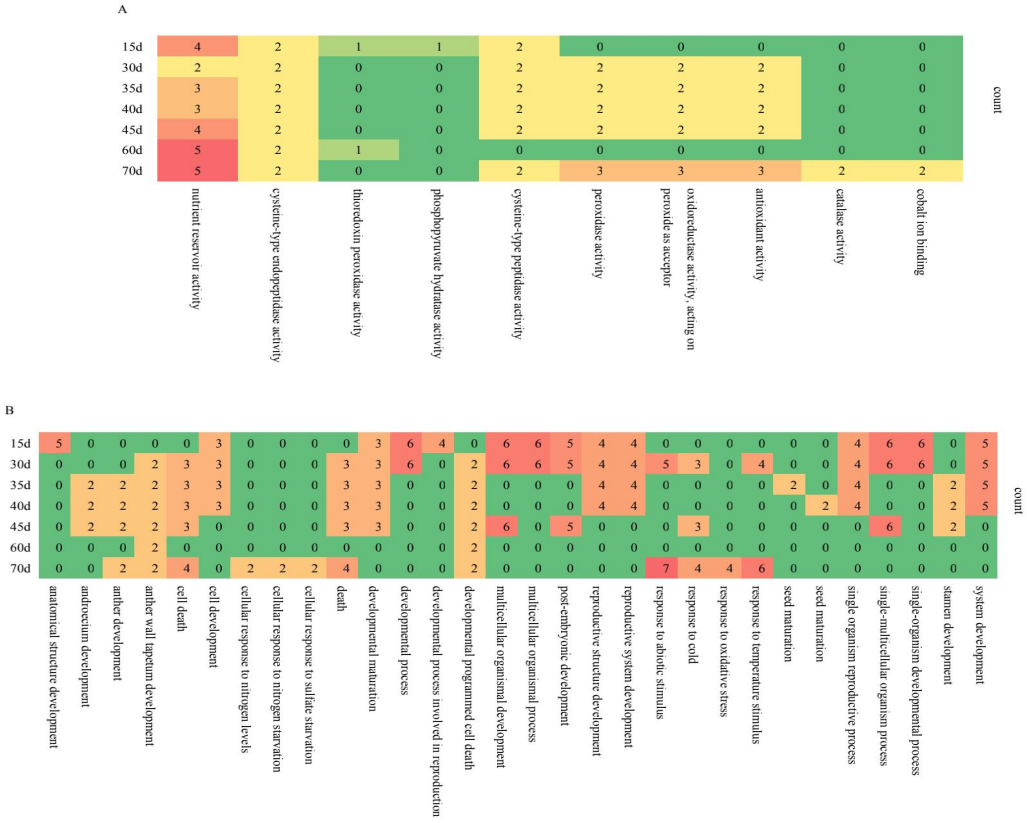


Fig. 4. Analysis of differentially expressed proteins identified by GO functional classification



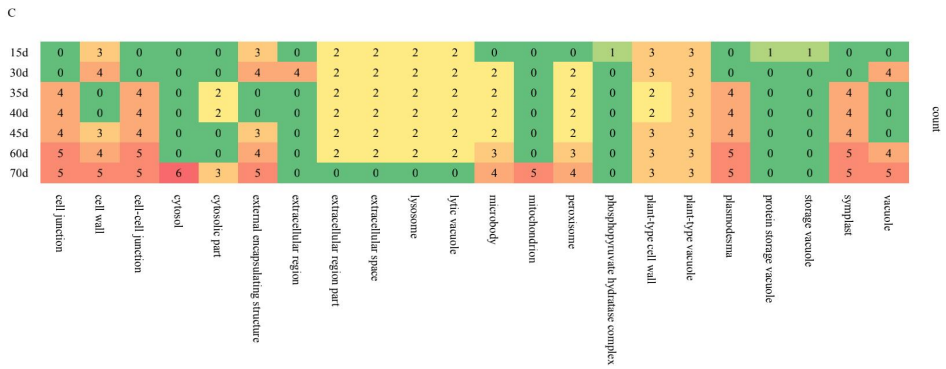


Fig. 5. GO analysis of differentially expressed proteins during the dormancy release of *T. miqueliana* seeds
Note: A:BP, B:MF, C:CC

3.3 Kyoto Encyclopedia of Genesand Genomes (KEGG) pathway annotations of differentially expressed proteins

By screening different sampling points, we acquired the endosperm differentially expressed protein information and mapped it onto the KEGG database to obtain enriched KEGG pathways. Comprehensive statistics on KEGG pathway annotations, which were significantly correlated with differentially expressed proteins at different time points and the number of differentially expressed proteins included in each pathway, were carried out (Fig. 6). The differentially expressed proteins at all time points participated in carbon metabolism and peroxisome pathways, indicating that the two pathways were closely related to the dormancy release of *T. miqueliana* seeds. The major proteins involved included ENO2, CAT2, GAPC2, and SDRA.

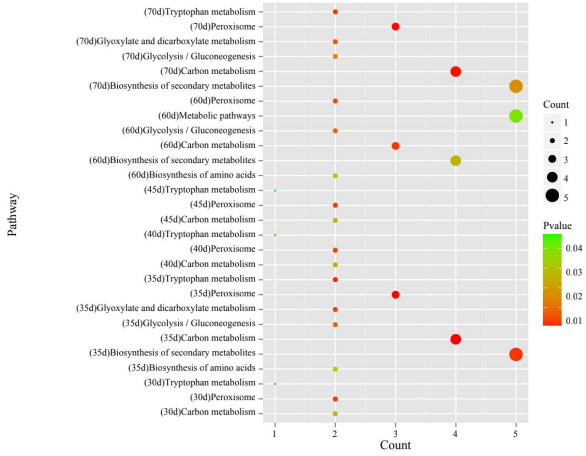


Fig.6. KEGG annotation diagram of the comparative protein significantly associated at each timepoint
Note: The horizontal axis represents the number of differentially expressed proteins in each KEGG pathway, and the longitudinal axis represents a significantly correlated KEGG node. The color of point from green to red indicates that the significant p value changes from small to large, and the size of point represents the number of genes.

3.4 Construction of differentially expressed protein interaction network

The interactions of differentially expressed proteins acquired at each time point were searched using the STRING version 10.5 database [38]. The interaction network of differentially expressed proteins at each time point was visualized through Cytoscape3.3[39-40]. STRING analysis of the differentially expressed proteins interaction network showed that the number of interaction proteins' increased and the complexity of the network increased with the extension of cold stratification during the dormancy release of *T. miqueliana* seeds (Fig. 7).

3.5 Screening of important proteins through WGCNA

The power when the squares of correlation coefficient reached 0.9 for the first time was selected. That is, power = 9. Then, we calculated the dissimilarity coefficient of the proteins and drew the system's clustering tree based on it. Subsequently, cutHeight was set at 0.95 in accordance with the standards of a hybrid dynamic shear tree. Five related modules were screened.

The correlations of the expression of each protein and each module with samples at different time points were calculated. The correlation coefficients of each module and dormancy release of seeds ranged between 0.5 and 0.9. The significant P value of the general state and dormancy release state of the seeds was 1e−05. The proteins of the cyan module had the strongest correlation with seed dormancy release (Fig. 8). The proteins involved in the cyan module included SPS1, SDRA, ENO2, At1g64255, CEP1, At4g17670, and At2g28490.

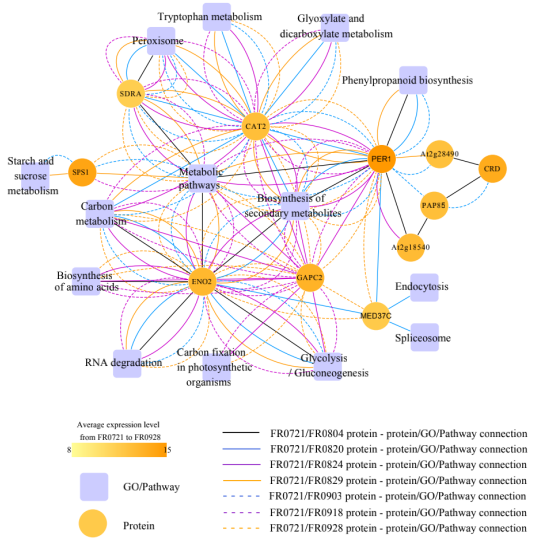


Fig.7. Integrated network diagram of differentially expressed protein interactions at each timepoint
Note: Orange circles and purple squares indicate differentially expressed protein nodes and significantly related GO or pathways, respectively. The color of the protein node is expressed from shallow to deep, which shows the expression changes in the process from the starting point 0721 to 0928(15d-70d after seed treatment). Different

colors and linear connections indicate the connection between protein and protein/GO/pathway at different times.

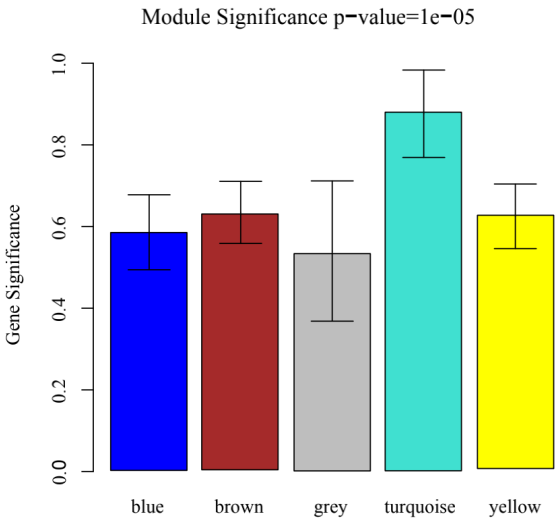


Fig. 8. Module bar chart of the suspended dormancy state of *T. miqueliana* seeds

4. Discussion

4.1Functional analysis of differential proteins during the dormancy release of *T. miqueliana* seed

4.1.1 Storage proteins

Storage proteins showed close relationship with seed germination. The activation and membrane repair of various enzyme systems require storage proteins or their degradation products. Degraded amino acid products can be used for seed germination through the gluconeogenic pathway [41]. GA can weaken the obstacle tissues of seeds by inducing hydrolase and facilitate the germination of seeds by inducing storage proteins in seeds and stimulating the extension of the embryo[42]. In this experiment, the seeds were stratified after treatment with GA3. Thus, the storage protein began to degrade in the early stage of stratification. The expression of some storage proteins was complex and fluctuated[42-43]. Further studies on these proteins would provide valuable information on physiological and molecular mechanisms involved in the germination of *T. miqueliana* seeds.

4.1.2 Programmed cell death (PCD)-related proteins

PCD means that cells commit “suicide” positively in accordance with a preset program to maintain environmental stability in cells to adapt to the environment[44]. During seed germination, aleurone layer cells secrete glucanase and xylanase to the external environments of cells to decompose endosperm cell walls and secrete various hydrolases to decompose endosperm cellular contents, thus providing nutrients to the growth and development of the embryo. In this process, endosperm cells degrade and die, followed by the successive deaths of aleurone layer cells[45]. PCD was related to seed dormancy release during *Oyama sieboldii* seed germination. During the dormancy release of *T. miqueliana* seeds, three protein spots were identified as KDEL-tailed cysteine endopeptidase, and they were related to the PCD of plants. These proteins were always up-regulated during the seed dormancy release, which may play an important role for the dormancy release of *T. miqueliana* seeds.

4.1.3 Defense-related proteins

PER1 is a dormancy regulator and antioxidant and can sense a harsh external environment. It keeps seeds dormant, inhibits seed germination under unfavorable conditions [46], protects the embryo, and alters cells from oxidative damage during imbibition and water absorption in the late and early stages of seed development [47]. The expression of PER1 in *T. miqueliana* seeds fluctuates during dormancy release, early down-regulation of degradation is conducive to dormancy release, and late up-regulation is conducive to seed germination.

SDRA has redox activity and participates in many biological pathways, such as metabolism of fatty and indolic acids and root hair elongation [41]. It is the precursor that inhibits active growth hormones for root growth [48]. SDRA can transform the ABA precursor lutein into ABA aldehyde and thereby change it into ABA to inhibit seed germination [49]. The expression of SDRA showed a decreasing trend during the dormancy release of *T. miqueliana* seeds, indicating that the degradation of short-chain dehydrogenase is conducive to the release of seed dormancy.

Heat shock protein (HSP) plays an important role in seed growth, storage, and germination. A previous study has pointed out that the HSP content in dry and mature seeds was quite high, but the HSP content decreased gradually with the germination of seeds until it disappeared. This finding proved that HSP is a protein that is essential for seeds to maintain dormancy [50]. And the mass synthesis of HSP could increase the vitality of seeds, which is conducive to the maturity of protein complexes and the degradation of damaged or misfolding polypeptides [51]. A probable mediator of RNA polymerase II transcription subunit 37c (MED37C) was identified in this study, it is a heat-shock 70 kDa protein 4, belonging to HSP70. This protein showed abundant expression levels after the *T. miqueliana* seeds were processed, reaching the peak at about 30 d and then decreasing continuously. It is indicated that the dormancy is gradually released at this time.

4.1.4 Energy metabolism‐related proteins

Seed germination requires a substantial amount of material and energy. During germination, stored nutrients, such as starch, fat, and protein, are mobilized to provide energy for the initial seed germination. One protein spot was identified as GAPC2. It is a key enzyme of EMP, and it is the first step to transform d-glyceraldehyde 3-phosphoric acid into the 3-phosphoric acid-d-glycerophosphate catalysis pathway [41]. This protein is closely related to the dormancy release of Arabidopsis seeds [52]. The expression abundance of GAPC2 was up-regulated after seed treatment, which contributed to seed germination (Fig. 3B).

The activity of sucrose phosphate synthase (SPS) is positively related to the accumulation of saccharose, which is a main product of photosynthesis of higher plants and the main form of carbon transportation. Saccharose provides carbon and energy source for the growth and development of cells. Under adverse conditions, such as low temperature, drought, or high salt, SPS activity in plants may often be strengthened, changing the distributions of carbon photosynthate in plants, increasing the content of soluble sugar (e.g., saccharose) in plants, and changing the osmotic pressure of cells to resist the influences of stress environments [53]. *Magnolia sieboldii* seeds stored abundant saccharose in the early stage of cold stratification. The activity of invertase increased quickly, and the accumulated

saccharose in the seeds was hydrolyzed, thus resulting in the sharp reduction in the content and increase in glucose and fructose content [54]. Similar phenomenon has been discovered after seed germination was promoted by increase in GA3 content and anabolism-related protein expression levels in *Phellodendron chinense* through the exogenous GA3 [55]. In this research, one protein spot was identified as SPS1, and its expression basically decreased. This finding was speculated to be related to the dormancy release of *T. miqueliana* seeds.

4.2 Proteomic bioinformatics analysis of *T. miqueliana* seeds during dormancy release

By analyzing proteomic data through bioinformatics, we can predic the structures and functions of proteins and study the interaction of proteins and their mechanisms of action. At the same time, we can discover proteins that may have biological significance for further study. [55,56].

In the co-expression network, proteins are mainly related with biological processes, such as photosynthesis (GO:0015979), glucose metabolic process (GO:0006006), response to metal ion (GO:0010038), and so on. Moreover, they participate in biosynthesis of plant hormones (ath01070), Glycolysis / Gluconeogenesis (ath00010), etc. This proves that these biological processes and pathways are closely related with dormancy release of Tilia miqueliana seeds. Deep studies on these biological processes and pathways can further disclose the molecular mechanism of dormancy release of Tilia miqueliana seeds.

5. Conclusions

In this study, endosperm proteins of *T. miqueliana* seeds were separated and differentially expressed proteins were identified by using the two-dimensional electrophoresis combined with mass spectrometry (2DE-MALDI-TOF/TOF). Moreover, bioinformatics analysis was carried out. Key proteins during dormancy release of *T. miqueliana* seeds were screened and some major conclusions could be drawn:

(1) A total of 101 differentially expressed protein spots are identified by mass spectroscopy. Finally, 49 protein spots are identified successfully and most of identified proteins are related with energy metabolism as well as PCD and stress of PLANTS.

(2) GO annotation and classification are carried out to differentially expressed proteins in endosperm which are acquired during dormancy of *T. miqueliana* seeds. Results found that it is the key stage during dormancy release of *T. miqueliana* seeds at about 30d-40d. At this moment, the number of proteins which have oxidoreductase activity, antioxidant, and combined enrichment with transition metal ions is the highest in molecular functions. These proteins mainly participate in metabolism processes in organelles, cell cycles, non-membrane bounded organelles, mitochondria and peroxisome as well as catabolism processes of cells. It mainly involves proteins, including CEP1, PER1, SDRA, CAT2, ENO2, GAC2 and At3g12020.

(3) According to KEGG pathway enrichment analysis results, there are 4 more differentially expressed protein clustering pathways during dormancy release of *T. miqueliana* seeds compared to those at dormancy state. These are glycolysis/gluconeogenesis, secondary metabolic biosynthesis, amino acid biosynthesis as well as glyoxalic acid and diuronic acid metabolism. After dormancy release, Tilia miqueliana seeds begin to germinate in a large scale (after 60d of lamination) and two pathways are added, including metabolism pathway and amino acid biosynthesis. It is speculated that these pathways play the key role in dormancy release and germination of *T. miqueliana* seeds.

(4) In this experiment, proteins which are screened by WGCNA and have significant correlations with dormancy release of seeds include SPS1, SDRA, ENO2, At1g64255, At4g17670 and At2g28490.

Dormancy and germination of seeds are keys in breeding of plants. They involve a series of complicated physiological and biochemical reactions and conduction of many signal molecules. With the development and progresses in proteomics and bioinformatics, they provide good means to study dormancy and germination mechanism of seeds. In this study, proteomics and bioinformatics were applied to analyze differentially expressed proteins of endosperm which are gained during dormancy release of *T. miqueliana* seeds. Specific proteins which may be able to control dormancy of *T. miqueliana* seeds are identified. They provide new clues for further studying the dormancy release of *T. miqueliana* seeds.

In future, functions of the screened differentially expressed proteins will be further verified to determine their specific roles during dormancy release of *T. miqueliana* seeds.

Acknowledgements

The authors are grateful for the support provided by the Science Project of Jiangsu Forestry Science and Technology Innovation and Promotion Project, China(LYKJ[2021]03,LYKJ[2022]17).

References

[1] A. M. Staszak, M. Guzicka, and T. A. Pawłowski, "Signalling regulators of abscisic and gibberellic acid pathways are involved in dormancy breaking of Norway maple(*Acer platanoides* L.)seeds," *Acta Physiol. Plant.*, vol. 39, Oct. 2017, Art. no. 251, doi: 10.1007/s11738-017-2544-0

[2] A. Soltani, *et al*, "The genetics and physiology of seed dormancy, a crucial trait in common bean domestication," *BMC Plant Biol*., vol. 21, Jan. 2021, Art. no. 58, doi:10.1186/s12870-021-02837-6.

[3] F. Chen, *et al.*, "New insights into the molecular mechanism of low-temperature stratification on dormancy release and germination of *Saposhnikovia divaricata* seeds,"*Braz. J. Bot*, vol.45, no.4, pp.1183-1198, Nov. 2022, doi: 10.1007/s40415-022-00849-6

[4] M.V. Arana, R.S. Tognacca, M. Estravis-Barcalá, R. A. Sánchez, and J. F. Botto, "Physiological and molecular mechanisms underlying the integration of light and temperature cues in *Arabidopsis thaliana* seeds," *Plant Cell Environ*., vol.40, no.12, pp. 3113-3121, Sep. 2017, doi:10.1111/pce.13076

[5] Z. Vondráková, *et al*., "Dormancy breaking in *Fagus sylvatica* seeds is linked to formation of abscisic acid-glucosyl ester," *New For.*, vol.51, no.4, pp.671–688, Sep. 2019, doi: 10.1007/s11056-019-09751-8

[6] A. H. Li, *et al*., "Molecular mechanism of seed dormancy release induced by fluridone compared with cod stratification in *Notopterygium incisum*," *BMC Plant Biol.*, vol.18, Art. no.116, Jun. 2018, doi: 10.1186/s12870-018-1333-2

[7] P. Tomasz, "Proteome analysis of Norway maple (*Acer platanoides* L.) seeds dormancy breaking and germination: influence of abscisic and gibberellic acids,"*BMC Plant Biol*., vol. 9, May 2009, Art. no. 48, doi: 10.1186/1471-2229-9-48

[8] Y. Q. Niu, *et al.*, "Studies on molecular mechanism of plant seed dormancy," *Fujian. J. Agri. Sci*., vol.32, no.10, pp.1156-1164, Dec. 2017.

[9] H. Nonogaki, "Seed germination and dormancy: The classic story, new puzzles, and evolution,"*J. Integr. Plant Biol*.," vol. 61, no. 5, pp. 541-563, Dec. 2018, doi:10.1111/jipb.12762

[10] M. R. Wilkins, *et al.,* "From proteins to proteome: large scale protein identification by two dimensional electrophoresis and amino acid analysis," *Nat. Biotechnol*., vol.14, pp.61-65, Jan. 1996, doi: 10.1038/nbt0196-61

[11] M. R. Wikins, *et al*., "Progress with proteome projects:why all proteins expressed by a genome should be identified and how to do it," *Biotechnol. Genet. Eng. Rev*., vol.13, no.1, pp.19-50, Jan.1996, doi: 10.1080/02648725.1996.10647923

[12] T. A. Pawłowski, "Proteomic approach to analyze dormancy breaking of tree seeds," *Plant Mol. Biol*., vol.73, pp.15-25, Mar. 2010, doi: 10.1007/s11103-010-9623-6

[13] K. Gallardo, *et al*., "Proteomics of Arabidopsis seed germination.A comparative study of wild-type and gibberellin-deficient seeds," *Plant Physiol*., vol.129, no. 2., pp. 823-837, Jun. 2002, doi: 10.1104/pp.002816

[14] A. Rashid, A. Badhan, M. Deyholos, and N. Kav, "Proteomic profiling of the aleurone layer of mature *Arabidopsis thaliana* Seed,"*Plant Mol. Biol. Rep*., vol. 31, pp. 464-469, Aug. 2012, doi: 10.1007/s11105-012-0498-3

[15] H. Su, "Study on the dormancy mechanisms and dormancy releasing methods in *Epimediumn wushanense* (Berberidaceae) seed,"M.S. thesis, Dept. For. China Agri. Univ., Beijing, China, 2016.

[16] H. Lan, Y.F. Leng, Sh. F. Zhou, J. Liu, and T. Z.Rong, "Proteomic analysis of storage substances during after-ripening of dormant seeds with dry ripening process in Maize Inbred Line 08-641," *J. Plant Genet. Resour.*, vol.16, no.1, pp.23-28, Jan. 2015, doi: 10.13430/j.cnki.jpgr.2015.01.004

[17] T. A. Pawłowski, "Proteomics of European beech (*Fagus sylvatica* L.) seed dormancy breaking: Influence of abscisic and gibberellic acids," *Proteomics*, vol. 7, no. 13, pp. 2246-2257, Jun. 2007, doi: 10.1002/pmic.200600912

[18] K. Chibani, S. Ali-Rachedi, C. Job, D. Job, M Jullien and P. Grappin, "Proteomic analysis of seed dormancy in Arabidopsis," *Plant Physiol*., Vol. 142, no. 4, pp.1 493–1510, Oct. 2006, doi: 10.1104/pp.106.087452.

[19] C. Lee, C. Chien, C. Lin, Y. Y. Chiu, and Y. S. Yang, "Protein changes between dormant and dormancy-broken seeds of *Prunus cam-panulata* Maxim," *Proteomics*, vol.6, no.14, pp.4147-4154, Jul. 2006, doi: 10.1002/ pmic.200500118

[20] R. Zheng, X. Y. Xu, C. M. Li, J. Y. Gai, and D. Y. Yu, "Proteomic Analysis of Differentially Expressed Proteins in Developmental Soybean (*Glycine max*(L.)Merr.) Seed," *Soybean Sci*., no.4, pp.556-563, Aug. 2008, doi: 10.1007/s10499-007-9164-4.

[21] H. Xu, *et al*., "Proteomic analysis reveals different involvement of proteins during the maintenance and release of rice seed dormancy," *Mol. Breed*., vol. 39, Mar. 2019, Art no. 60, doi:10.1007/s11032-019-0963-x.

[22] P. Li, J. Sun, and Y. Sun, "Study on the proteomics of dormancy-breaking mechanismin buffalograss seeds," *Chin. J. Grassl*., vol.35, no.3, pp.45-50, May, 2013, doi:10.3969/j.issn.1673-5021.2013.03.008

[23] G. Z. Sun., X. Y. Zhang, and C. S. Yan, "Proteomic analysis of embryonic dormancy in wheat," *Acta Agri. Boreali-Sin*., vol.24, no.2, pp.1-7, Apr. 2009, doi:10.7668/hbnxb.2009.02.001.

[24] A. Hashiguchi, H. Yoshioka, and S. Komatsu, "Proteomic analysis of temperature dependency of buckwheat seed dormancy and quality degradation," *Theor. Exp. Plant Physiol*., vol. 30, pp.77-88, May 2018, doi: 10.1007/s40626-018-0104-7

[25] F. H. Shi, "Molecular systematics of Tilia and conservation genetics of *Tilia miqueliana*," Ph. D. dissertation, Dept. For., Nanjing For. Univ., Nanjing, China, 2009.

[26] H. L. Wang, L. J. Yan, and X. Huang, "Genetic diversity and genetic structure of *Tilia Miqueliana* population,"*J. Nanjing. For. Univ*. (Nat. Sci. Ed.), vol. 47, no. 1, pp. 145-153, Feb. 2023, doi: 10.12302/j.issn.1000-2006.202110049

[27] F. H. Shi, "Study on biological characteristics and dormancy mechanism of *Tilia Miqueliana* seeds," M.S. thesis, Dept. For., Nanjing For. Univ., Nanjing, China, 2006

[28] F. H. Shi, Y. B. Shen, and J. B. Shi, "Study on germination in hibitor of Nanjing Linden seeds," *J. Fujian Coll. For.* vol.27, no.3, pp.222-225, Jul. 2007, doi: 10.3969/j.issn.1001-389X.2007.03.007

[29] Y. Wu and Y. B. Shen, "The structural and chemical characteristics of the pericarp are important in *Tilia miqueliana* seed dormancy," *New For.*, vol.52, no.5, pp.875-888, Jan. 2021, doi: 10.1007/s11056-020-09820-3.

[30] Y. H. Wang, "Study on seed dormancy mechanism of *Tilia Miqueliana*," M.S. thesis, Dept. For., Nanjing For. Univ., Nanjing, China, 2007.

[31] Y. H. Yang, "Study on endosperm metabolism of *Tilia Miqueliana* seed during dormancy release," M.S. thesis, Dept. For., Nanjing For. Univ., Nanjing, China, 2009.

[32] T. Isaacson, *et al*., "Sample extraction techniques for enhanced proteomic analysis of plant tissues," *Nat. Protoc*., vol. 1, no. 2, pp.769-774, Jul. 2006, doi: 10.1038/nprot.2006.102

[33] M. Bradford, "A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding," *Anal. Biochem.*, vol.72, no.1-2, pp.248-254, Feb. 1976, doi: 10.1016/0003-2697(76)90527-3.

[34] M. C. Oldham, *et al*., "Functional organization of the transcriptome in human brain," *Nat. Neurosci*., vol. 11, pp. 1271-1282, Oct. 2008, doi: 10.1038/nn.2207.

[35] Q. Liao, *et al*., "Large-scale prediction of long non-coding RNA functions in a coding-non-coding gene co-expression network," *Nucleic. Acids Res*., vol.39, no.9, pp.3864-3878, Jan. 2011, doi: 10.1093/nar/gkq1348.

[36] P. Langfelder and S. Horvath, "WGCNA: an R package for weighted correlation network analysis," *BMC Bioinformatics*, vol. 9, Dec. 2008, Art. no. 559, doi: 10.1186/1471-2105-9-559

[37] D.W. Huang, B.T. Sherman, and R.A. Lempicki, "Systematic and integrative analysis of large gene lists using DAVID Bioinformatics Resources," *Nat. Protoc.*, vol. 4, pp. 44-57, Dec. 2008, doi: 10.1038/nprot.2008.211

[38] D. Szklarczyk, *et al*., "The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible," *Nucleic. Acids Res.*, vol. 45, no. D1, pp. D362-D368, Oct. 2016, doi: 10.1093/nar/gkw937

[39] D. Szklarczyk, *et al.* "STRING v10: protein-protein interaction networks, integrated over the tree of life," *Nucleic Acids Res.*, vol.43, no. 1, pp. 447-452, Oct. 2014, doi: 10.1093/nar/gku1003

[40] P. Shannon, "Cytoscape: a software environment for integrated models of biomolecular interaction networks," *Genome Res*., vol.13, no.11, pp.2498-2504, Nov. 2003, doi: 10.1101/gr.1239303

[41] L.Taiz and E.Zeiger, (2010) *Plant Physiology*. Sinauer Associates Inc., Sunderland, MA., U.S.A.

[42] M. Mei, "Carbohydrate changes and screening of related differential proteins during germination of *Magnolia sieboldii*," M.S. thesis, Dept. For., Shenyang Agri. Univ., Shenyang, China, 2015.

[43] Y. P. Gao, "Study on Physiological and biochemical changes and molecular mechanism of Seed dormancy release of *Cercis chinensis*," Ph.D. dissertation, Dept. For., Nanjing For. Univ., Nanjing, China, 2020.

[44] J. F. Keer, A. H. Wyllie, and R. Carriea, "Apoptosis: A basic biolclgical phenomenon with wide-ranging implication in tissue kinetics," *Br. J. Cancer*, vol.26, no.4, pp. 239–257, Aug. 1972.

[45] H. M. Wu, J. Li, and G. H. Wang, "Advances in programmed cell death in cereal aleurone," *Pratacult. Sci*., vol.30, no.5, pp.777-782, May, 2013.

[46] C. Haslekås, *et al.*, "Seed 1-cysteine peroxiredoxin antioxidants are not involved in dormancy, but contribute to inhibition of germination during stress," *Plant Physiol*., vol. 133, no.3, pp.1148-1157, Nov. 2 003, doi: 10.1104/pp.103.025916.

[47] R. B. Aalen, H, G. Opsahl-Ferstad, C. Linnestad, and Odd ǀǀ rne Olsen, "Transcripts encoding an oleosin and a dormancy-related protein are present in both the aleurone layer and the embryo of developing barley *(Hordeum vulgare* L.) seeds," *Plant J.*, vol.5, no.3, pp.385-396, Mar.1994, doi: 10.1111/j.1365-313X.1994.00385.x

[48] Z. C. Yin, "Optimization of seed germination conditions and proteomic analysis of *Cupressus gigantea*," M.S. thesis, Dept. For. ,Northeast For. Univ., Ha erbin,China, 2020.

[49] W. H. Cheng, *et al.,* "A unique short-chain dehydrogenase/reductase in Arabidopsis glucose signaling and abscisic acid biosynthesis and functions," *Plant Cell*, vol.14, no.11, pp.2723-2743, Nov. 2002, doi: 10.1105/tpc.006494.

[50] J. Li, X. J. Lu, M. Mei, X. L. Zhang, and Y.Ma. "Proteomics of Seed Development and Germination Process," *Seed*, vol.35, no.5, pp. 59-64, May, 2016, doi:10.16590/j.cnki.1001-4705.2016.05.059

[51] C. A. Downs, J. S. Coleman, and S. A. Heckathorn, "The chloroplast 22-Ku heat-shock protein: A luminal protein that associates with the oxygen evolving complex and protects photosystem Ⅱ during heat stress," *Plant Physiol*., vol.155, no.5, pp.477-487, Oct. 1999.

[52] L. Guo, *et al*., "Cytosolic glyceraldehyde-3-phosphate dehydrogenases interact with phospholipase Ddelta to transduce hydrogen peroxide signals in the Arabidopsis response to stress," *Plant Cell*, vol.24, no.5, pp.2200-2212, May, 2012, doi:10.1105/tpc.111.094946

[53] J. Y. Wu, M. Chen, and Y. L. Yao, "Research Progress of Sucrose Phosphate Synthetase (SPS) Gene in Plants,"*Mol. Plant Breed*., Vol.20, No.8, pp.2597– 2606, Feb. 2022.

[54] Y. Y. Liu, "The relationship between material changes and seed germination during seed stratification in *Magnolia sieboldii*," M.S. thesis, Dept. For., Shenyang Agri. Univ., Shenyang, China, 2010.

[55] L. Yang, *et al*. "Study on the mechanism of exogenous gibberellin promoting seed germination of Phellodendron chinense Schneid," *Plant Physiol. J*., vol.58, no.11, pp.2163-2172, Sep. 2022.

[56] P. S. Gromov and J. E. Celis, "From genomics to proteomics," *Mol. Biol.*, vol.34, no.4, pp.508-520, Jul. 2000.