Websites:
http://www.lifesciencesite.com
http://www.sciencepub.net

Emails:
editor@sciencepub.net
sciencepub@gmail.com

Life Science Journal

MARSLAND PRESS
Multidisciplinary Academic Journal Publisher

Analysis on Transcriptome of Genes Involved in Maize Seed Germination

Xueqin Hu, Juan Zou, Guoqing Dong, Wanyong Zeng, Jinhua Li and Dengxiang Du*

School of Biological and Pharmaceutical Engineering, Wuhan Polytechnic University, Wuhan 430023, China
* Correspondence: ddx@whpu.edu.cn

Abstract: Seed dormancy and germination are key to plant growth and development, and the transformation from dormancy to germination is closely related to agricultural production. Maize (*Zea mays* L.) is an important food crop, feed source, and industrial raw material. It is an important part of agricultural production, and the efficiency of seed germination after sowing is a key link. This study used non germinated maize as a control, and conducted transcriptome analysis of maize at 24, 48, 72, and 96 hours of germination. Through identification of differentially expressed genes (DEGs) and functional analysis of DEGs, it was found that a large number of differentially expressed genes were detected in the cell process, metabolic process, cell part, biological regulation, and response to stimuli during the germination stage of maize. These genes are significantly clustered in ribosomes, plant hormone signal transduction, carbon metabolism, amino acid metabolism, glutathione metabolism, amino and nucleotide sugar metabolism, fatty acid metabolism, and MAPK signaling pathways. In short, under the regulation of hormones, significant changes in the expression of corn seeds occur during the germination process, utilizing the metabolism of stored proteins, starch, and lipids, and utilizing the activation of metabolic pathways in the respiratory pathway to provide the necessary materials and energy for seed dormancy release and germination, as well as promoting the germination of corn seeds. Through the findings of this study, we have deepened our understanding of maize seed germination, and also contributed to the development of molecular based breeding, providing a theoretical basis for improving the production time of maize.
[Xueqin Hu, Juan Zou, Guoqing Dong, Wanyong Zeng, Jinhua Li and Dengxiang Du. Analysis on Transcri ptome of Genes Involved in Maize Seed Germination. Life Sci J 2023;20(5):18-31]. ISSN 1097-8135 (prin t); ISSN 2372-613X (online). http://www.lifesciencesite.com. 04. doi:10.7537/marslsj200523.04.

Keywords Seed Germination； Maize (*Zea mays* L.); Transcriptome; Plant hormone signal transduction

中图分类号：S5　　文献标识码 A

Seed is the main material and energy source for human survival. Cereal crops alone account for 90% of all plant seeds, providing more than half of the energy intake for human beings worldwide. Therefore, seed biology has become one of the key contents of plant physiology research (Yan & Chen., 2020). Seed dormancy not only improves the seed to survive the bad environment, but also effectively reduces the fierce competition for growth resources among individuals of the same species, and plays a positive role in the survival, continuation, evolution and germplasm protection of its species (Tuan et al., 2018). Because there are many reasons for seed dormancy, different plant seeds have different dormancy characteristics. On the basis of long-term research, different seed biologists have put forward several theories of seed dormancy regulation from different research perspectives, such as the theory of respiratory pathway regulation, hormone regulation, energy regulation, photosensitizer regulation, gene regulation, etc (Wu et al., 2020; Negi & Ayele., 2021).

Seed dormancy is a very complex phenomenon. So far, researchers all over the world have carried out systematic research on the dormancy characteristics of most plant seeds. There are many reasons for seed dormancy. Some seeds have special anatomical structures and characteristics, such as the high degree of lignification of the seed coat, which seriously affects the water absorption rate of the seeds, and the exchange of O_2 and CO_2 gas or mechanical resistance to the extension of the radicle (Graeber et al., 2012; Shu et al., 2015; Shu et al., 2016). Some seeds' dormancy is caused by some physiological reasons, such as immature embryo development, the existence of endogenous inhibitors in seeds, etc. Some are caused by the obstruction of some physiological metabolism in the seeds, such as the insufficient level of soluble metabolic substances, enzyme activities, plant hormones, etc. required for the growth and development of seed embryos (Cheng et al., 2022; Reed et al., 2022). But generally speaking, the dormancy of seeds is often the result of many factors.

Most plant seed dormancy in nature belongs to comprehensive dormancy (also known as combined dormancy or multiple dormancy), and there are many combinations of internal dormancy and external dormancy.

Transcriptomics has been applied in the research of seed dormancy and germination, and has achieved certain results. Li et al. (Li et al., 2015) used RNA-seq technology to sequence the transcriptome of different dormant peanut (*Arachis hypogaea* Linn.) seeds and screened 1026 differentially expressed unigenes, among which gibberellin 20 oxidase, abscisic acid 8 'hydroxylase, ACC oxidase, EREBP-like factor, heat shock protein, etc. comprehensively regulate the dormancy release and germination of peanut seeds. Yu et al. (Yu et al., 2016) analyzed the related mechanism of the wheat germination process by transcriptome analysis. The results showed that seed germination involves complex network regulations, such as substance transport, hormone metabolism, protein degradation, and signal transduction. Some differentially expressed genes and related metabolic pathways were also identified (Mangrauthia et al., 2016; Han et al., 2020).

Maize (*Zea mays* L.), belonging to the cereal corn plant, is one of the most important crops in the world, and plays an important role in agricultural production and economic life. Corn is the main food crop in many parts of the world and the main feed crop in most parts of the world. It is also an important industrial raw material. It is widely used to produce alcohol, ferment and provide plant fiber materials for bioenergy. With the development of corn production, China's corn planting area continues to move northward, and the risk of corn production also increases, and corn germination will also be affected by more complex environment. Strengthening the research of corn germination and improving the success rate of germination are important measures to improve the stable production of corn, which is conducive to promoting corn breeding.

1. Materials and Methods

1.1. Maize materials

The healthy and plump maize (*Zea mays* L.) seeds of the same size were selected, disinfected with 1% sodium hypochlorite, washed with distilled water for 3-4 times, and then placed in a petri dish (9 cm in diameter) covered with two layers of filter paper, with 30 seeds in each dish. The seeds treated above were placed in an incubator at 25℃ for cultivation. The untreated seeds were used as control, and the seeds germinated for 24 h, 48 h, 72 h and 96 h were selected for transcriptome analysis.

1.2. Experimental methods

1.2.1 RNA-seq library construction and sequencing

Total RNA was isolated from each sample from a pool of callus after morphological classify, extracted using the Invitrogen Trizol reagent according to the manufacturer's instructions and evaluated RNA integrity and concentration. RNA integrity was evaluated using the Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA), and RNA Integrity Number (RIN) ≥ 7 were subjected to the subsequent analysis. The libraries were constructed using TruSeq Stranded mRNA LTSample Prep Kit (Illumina, San Diego, CA, USA) according to the manufacturer's instructions. Briefly, the total mRNA was isolated with oligo (dT) beads. All of the mRNA was cut into short fragments (200 nt) by adding a fragmentation buffer. First-strand cDNA was generated using random hexamer-primed reverse transcription. Second-strand cDNA was then synthesized by DNA polymerase I and RNase H. Afterward, the synthesized cDNA fragments were purified and then subjected to end pairing, the addition of a single "A" base, and ligation with Illumina adapters. The ligation products were subsequently size-fractioned by agarose gel electrophoresis, after which the fragments were excised for PCR amplification and prepared using Illumina TrueSeq Stranded Total RNA HT Library Preparation Kit and as per the protocol described by the manufacturer (Illumina®, San Diego, California, USA).

1.2.2 Sequencing analysis and differential expression analysis

We used the B73 reference genomic and annotated files as the database and identified the expression abundance of each protein-coding gene in each sample by the method of sequence similarity alignment. Raw data were processed using Trimmomatic, sequencing output raw data were first filtered to remove adaptor tags, low quality sequences (reads containing ploy-N) and tags with a copy number of 1 (probably sequencing error) were removed to obtain the clean reads. For annotation, all tags were mapped to reference genome using hisat2 and allowed no more than one nucleototide mismatch (Kim et al., 2019). All the tags mapped from multiple genes were filtered and the remaining tags were designed as unambiguous tags. The number of mapped clean reads for each gene was then counted and normalized into Fragments Per Million Reads (FPKM) for calculating gene expression using the HTSeq software to obtain the number of reads that were compared to the protein-coding genes in each sample, and CuffLinks software was used to calculate the expression amount of protein-coding genes FPKM (Zhao et al., 2021). Each gene was calculated using cuff links, and there adcounts of each gene were obtained by htseq-count. To obtain statistical confirmation of the differences in gene expression among the different state callus, we

then compared the FPKM read count using a threshold value of P≤0.001 and the absolute value of log2Ratio ≥1 based on the FDR < 0.05.The differentially expressed genes (DEGs) were screened with FPKM values using the DESeq package function sestimate Size Factors and nbinomTest. When comparing two groups, edgeR was used to analyze DEGs to correct for multiple testing, and the FDR was calculated to adjust the threshold of the p-value. Genes with a minimum 2-fold difference in expression, |log2FC| ≥ 1 and FDR ≤ 0.05 were were identified as differentially expressed genes (DEGs). The normalized sequencing data from two biological replicates were used for DEGs identification. Hierarchical cluster analysis of differentially expressed genes (DEGs) was performed to explore genes expression pattern. Gene Ontology (GO) enrichment analysis was performed using the WEGO 2.0, based on the Gene Ontology Database (http://www.geneontology.org/). KOG functional classification, and pathway annotation and enrichment analyses were based on the NCBI COG (http://www.ncbi.nlm.nih.gov/COG) (Ashburner et al., 2000). Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis was performed using the KOBAS 3.0 tools, based on the KEGG pathway (http://www.genome.jp/kegg/). For each KEGG pathway, the numbers of DEGs were compared to the entire reference gene set by hypergeometric tests to determine the pathways enriched for differentially regulated genes (Kanehisa and Goto., 2000). The p-values of the GO and KEGG enrichment analyses were adjusted using the Bonferroni correction, and a corrected p-value ≤ 0.05 was chosen as the threshold value for determining significantly enriched GO terms.

1.2.3 Real-time PCR validation

To validate the DEGs obtained from transcriptome sequencing, DNA-free RNA samples were reverse transcribed to cDNA using a cDNA synthesis kit (Thermo Fisher Science, USA). Gene-specific primers were designed using Primer Premier 5.0 (Table 1). Real-time fluorescence quantitative PCR (qRT-PCR) was performed using Maxima SYBR Green/ROX qPCR Master Mix (Chen et al., 2019), procedure: 95 °C for 5 min; 95 °C for 30 s, 58 °C for 30 s, 72 °C for 30 s (40 cycles), followed by 72 °C for 5 min. The experiments were set up with three biological replicates and three technical replicates. All reactions were performed on the CFX96 Real-time system (Bio-RAP, USA). The mean values of the three biological replicates were counted using the maize Actin gene (*GRMZM2G126010*) as the internal reference, and the data were analyzed by CFX manager software. Gene expression was quantified using the $2^{-\Delta\Delta CT}$ method, and the standard error among the three biological replicates was calculated.

Table 1. Primer sequence information for qRT-PCR.

Primer name	Primer sequence (5'→3')
Zea mays-Actin-F // *Zea mays*-Actin-R	TTACAGAGAGGCCCAACTGC // GCATCATCACCAGCGAAACC
Zm00001d001961 –F // *Zm00001d001961* –R	GGACAAAAGCATCTTGCCCG // AAATCACGAGGAACCGCCTT
Zm00001d002302 -F // *Zm00001d002302* -R	GGGACGCAAGGATTACCCAA // TCAACGGGGAGTAGGTTTGC
Zm00001d002374 -F // *Zm00001d002374* -R	AACGAGCTTGACCCGTGAAA // GCGGAATTGTTACACCAGGC
Zm00001d002004 -F // *Zm00001d002004* -R	GCCTGGTGTAACAATTCCGC // CTCGCGTACAACTTTGGCAG
Zm00001d002704 -F // *Zm00001d002704* -R	CTGCCAAAGTTGTACGCGAG // TTGTCGGGTTTGCTTCGGTA
Zm00001d005300 -F // *Zm00001d005300* -R	CGCCGAAGCTCCATGAAAAG // TACTGTTGGGGTCGTGCTTC
Zm00001d005547 -F // *Zm00001d005547* -R	GCCTGGTGTAACAATTCCGC // CTCGCGTACAACTTTGGCAG
Zm00001d010445 -F // *Zm00001d010445* -R	GATGATGAGTTGTGTGCGCC // CTGCTAGCTGCTCCGATACC
Zm00001d016105 -F // *Zm00001d016105* -R	AACTTCGATGTTTCGCGTGC // ACATCGACCACGAACGACTC

1.2.5 Statistical Analysis

Statistical analyses were performed using GraphPad Prism 9 (https://www.graphpad.com/). The experiments were performed with three biological replicates, and plant materials from ten seedlings were pooled for each biological replicate. Statistical signif-icance was determined using two-way ANOVA and Tukey's test. Differences were considered statistically significant at *** $P \leq 0.001$, ** $P \leq 0.01$, and * $P \leq 0.05$.

2. Results and Analysis

2.1. Transcriptome analysis during the seed germination

The RNA were isolated from the seeds after germinated and control, mixture 30 samples of each tissue to RNA-seq profiling. An overview of the transcriptome analysis reads was listed in **Table 2**. Raw reads obtained using the Illumina sequencing platform were processed to obtain high-quality clean reads by removing low-quality sequences and removing joint contamination. Five categories of 15 independent RNA libraries were prepared and sequenced. Overall, 41.05 Gb clean data was generated, and the volume of effective data for each sample ranged from 4.06 Gb to 4.12 Gb after removal of the adapter. The low-quality reads were filtered using the Trimmomatic software for obtaining the Q30 index, which ranged from 94.5% to 95.22%. In different tissue materials, total reads detected were between 39,411,524 to 47,792,140, with the average at 44,245,632. Correspondingly, the mapped reads obtained were between 35,418,029 and 45,926,926, respectively. Of these, only 4.84% (average) had multiple sites in the genome. More than 95% of the reads were mapped as unit points on the chromosomes.

Table 2 Statistical results of the RNA-Seq reads.

Sample	Total Reads	Mapped Reads	Mapping Rate	MultiMap Reads	MultiMap Rate
CK-1	46,499,240	44,867,694	0.9649	2,322,511	0.0499
CK-2	45,463,434	43,865,145	0.9648	2,948,311	0.0649
CK-3	44,134,864	42,584,859	0.9649	2,403,053	0.0544
G-24 h-1	47,792,140	45,926,926	0.961	2,462,499	0.0515
G-24 h-2	45,180,670	43,429,347	0.9612	2,234,932	0.0495
G-24 h-3	44,043,452	42,376,482	0.9622	2,208,442	0.0501
G-48 h-1	45,546,706	43,660,472	0.9586	2,188,648	0.0481
G-48 h-2	43,273,500	41,487,442	0.9587	2,015,319	0.0466
G-48 h-3	40,654,950	38,990,655	0.9591	2,182,609	0.0537
G-72 h-1	43,563,050	42,009,619	0.9643	2,003,777	0.046
G-72 h-2	42,838,314	41,196,565	0.9617	1,657,503	0.0387
G-72 h-3	44,970,248	43,348,876	0.9639	2,068,021	0.046
G-96 h-1	45,744,716	44,002,427	0.9619	2,278,373	0.0498
G-96 h-2	44,589,262	42,907,064	0.9623	1,825,904	0.0409
G-96 h-3	46,107,834	44,370,082	0.9623	1,912,018	0.0415

Sample, Sample name; **Total Reads**, Total number of read pairs obtained; **Mapped Reads**, Number of sequenced sequences that can be located on a genome; In general, if there is no contamination and the reference genome selection is appropriate, the percentage of this data is greater than 70%; **Mapping Rate**, Proportion of mapped reads in total reads; **MuliMap Reads**, Number of sequenced sequences with multiple alignment locations on the reference sequence; **MultiMap Rate**, Rate means mapped reads in total reads.

2.2. Identified of DEGs

The correlation coefficient of different materials is 0.89–0.99 (Figure 1a). The correlation between repetitions of the same material is significantly higher than that between materials, indicating that each replicate could be used for subsequent analysis, which showed a high correlation, and the change in traits may be caused by the difference in the expression changes of certain genes. Expression genes were identified by transcriptome sequencing during the five stages, and the expression values of all genes were normalized as FPKM values to determine gene expression. The genes with FPKM ≥ 0.1 present in at least one sample were identified as expressed gene, a total of 67,257 genes were identified, five groups were divided based on the FPKM of the genes (**Figure 1b**). In all samples, low expression genes (about 42%) and medium expression genes (about 30%) occupied the largest proportion.

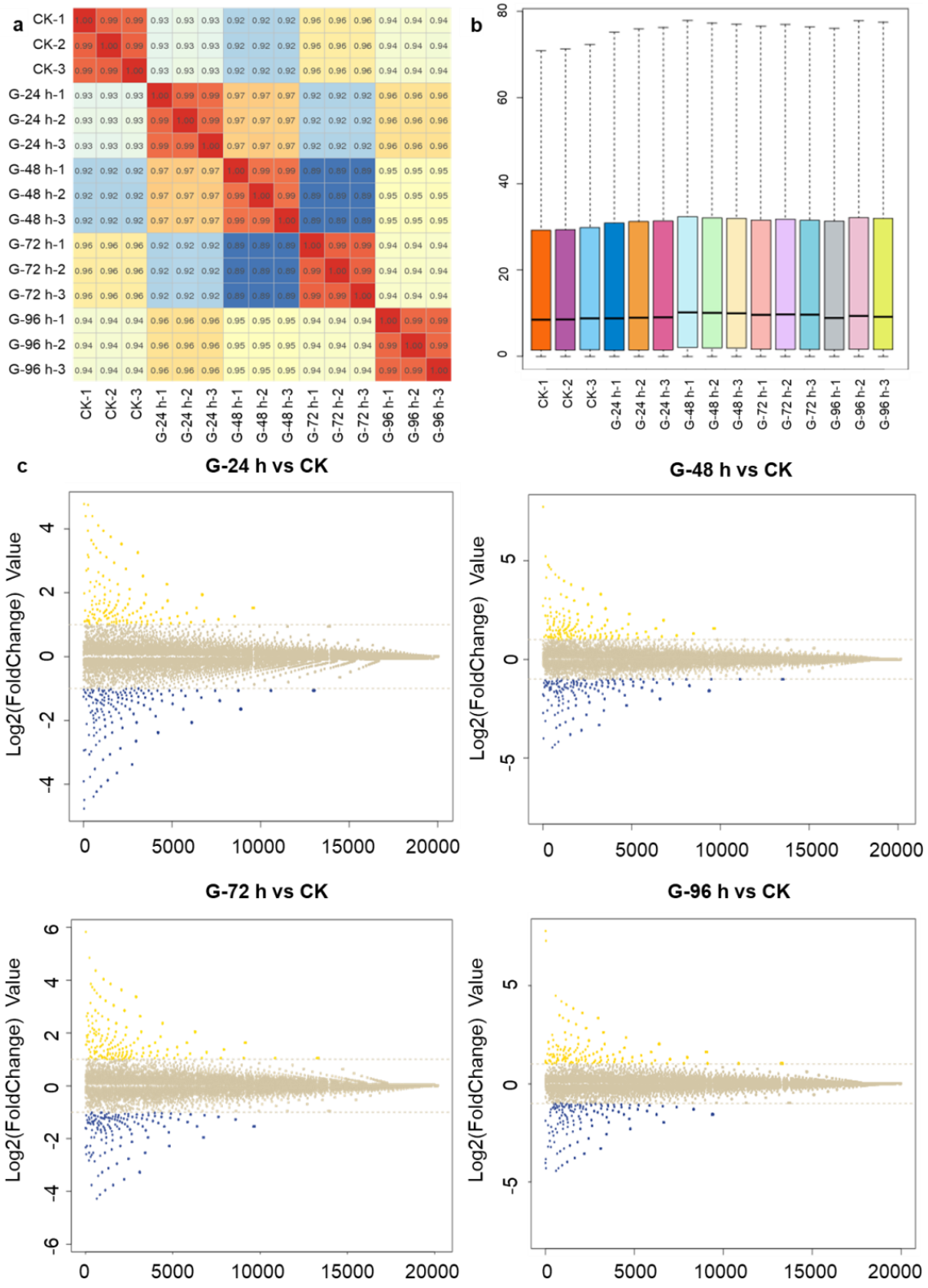

Figure 1. Gene expression during the seed germination, generated using high-throughput deep sequencing technology.

a, Relationship between samples in three biological replicates and summary of expressed genes during the seed germination. The abscissa represents the sample name; the ordinate, the corresponding sample name; and the color, the correlation coefficient. **b**, The boxplot of the expressions of 15 samples. Expression density of genes in twelve samples, represented by reads per kilobase of exon per million mapped reads (FPKM). The abscissa is the sample name, and the ordinate is log2 (TPM+1). The box chart for each region is divided into five statistics (top to bottom are the maximum, upper quartile, median, lower quartile, and minimum) **c**, Identification of differentially expressed genes (DEGs) between treatments, generated using high-throughput deep sequencing technology. The volcano plot presents the expression of the DEGs in different treatments, the red dots represent up-regulated genes, and the green dots represent down-regulated genes.

In this study, genes with a stable expression level more than twice between three replicates were selected as effectively expressed genes, and the average expression level between three replicates was calculated as the tissue expression level. The mean values in different tissue materials were referred to as CK, G-24 h, G-48 h, G-72 h and G-96 h. The DEGs were identified using Deseq2 v1.10.1. A p-value < 0.05, and fold change > 2 or fold change < 0.5 was set as the threshold for significantly differential expression (**Figure 1c**). A total of 6202 DEGs with significant variation were identified between G-24 h and CK, among which 3610 genes were up-regulated and 2592 genes were down-regulated. There were 1715 up-regulated genes and 2153 down-regulated genes between the G-48 h and CK groups and 3868 DEGs in total. There were 2574 DEGs between G-72 h and CK, including 1428 up-regulated genes and 1146 down-regulated genes. Between G-96 h and CK, 1402 genes with up-regulated expression and 1749 genes with depressed expression showed significant differential expression, and the total number of DEGs was 3150.

According to Z-score, the unigenes expression patterns were grouped and analyzed into 18 groups (Figure 2). The classification analysis reflected the trend of unigenes expression during seed dormancy release. Clusters 4, 7 and 10 showed a downward trend with induction. Clusters 11, 13, 16 and 17 showed an increasing trend with induction. For clusters 1, 2, 5, 6, 8, 9, 12, 14, 15 and 18, the genes were up-regulated and down-regulated, the significantly up-regulated in the early period of seed germination. The last group of cluster (Cluster 3) showed a down-regulated expression trend and then up-regulated expression trend. Further, we analyzed the number of unigenes and GO clusters contained in each cluster. In cluster 1, 695 genes had been identified and involved in the top 5 metabolic pathways as membrane, integral component of membrane, transferase activity, ATP binding and metal ion binding. In cluster 2, 526 genes had been identified and involved in the top 5 metabolic pathways as membrane, integral component of membrane, transferase activity, ATP binding and protein binding. In cluster 3, 163 genes had been identified and involved in the top 5 metabolic pathways as membrane, integral component of membrane, oxidation-reduction process, hydrolase activity and catalytic activity. In cluster 4, 263 genes had been identified and involved in the top 5 metabolic pathways as integral component of membrane, membrane, nucleus, protein binding and DNA binding. In cluster 5, 593 genes had been identified and involved in the top 5 metabolic pathways as membrane, integral component of membrane, protein binding, nucleus and nucleic acid binding. In cluster 6, 542 genes had been identified and involved in the top 5 metabolic pathways as membrane, integral component of membrane, protein binding, ATP binding and nucleic acid binding. In cluster 7, 637 genes had been identified and involved in the top 5 metabolic pathways as membrane, integral component of membrane, metal ion binding, oxidation-reduction process and oxidoreductase activity. In cluster 8, 610 genes had been identified and involved in the top 5 metabolic pathways as membrane, integral component of membrane, transferase activity, metal ion binding and protein binding. In cluster 9, 660 genes had been identified and involved in the top 5 metabolic pathways as membrane, integral component of membrane, transferase activity, ATP binding and hydrolase activity. In cluster 10, 523 genes had been identified and involved in the top 5 metabolic pathways as membrane, integral component of membrane, cytoplasm, golgi apparatus and vesicle−mediated transport. In cluster 11, 462 genes had been identified and involved in the top 5 metabolic pathways as membrane, integral component of membrane, protein binding, transferase activity and ATP binding. In cluster 12, 573 genes had been identified and involved in the top 5 metabolic pathways as oxidation-reduction process, metal ion binding, oxidoreductase activity, catalytic activity and metabolic process. In cluster 13, 435 genes had been identified and involved in the top 5 metabolic pathways as membrane, integral component of membrane, protein binding, regulation of transcription, DNA-templated and nucleus. In cluster 14, 557 genes had been identified and involved in the top 5 metabolic pathways as membrane, integral component of membrane, transferase activity, ATP binding and hydrolase activity. In cluster 15, 560 genes had been identified and involved in the top 5 metabolic pathways as membrane,

integral component of membrane, protein binding, ATP binding and hydrolase activity. In cluster 16, 891 genes had been identified and involved in the top 5 metabolic pathways as membrane, integral component of membrane, oxidation-reduction process, transferase activity and hydrolase activity. In cluster 17, 577 genes had been identified and involved in the top 5 metabolic pathways as membrane, integral component of membrane, protein binding, regulation of transcription, DNA-templated and nucleus. In cluster 18, 619 genes had been identified and involved in the top 5 metabolic pathways as membrane, integral component of membrane, transferase activity, metal ion binding and protein binding.

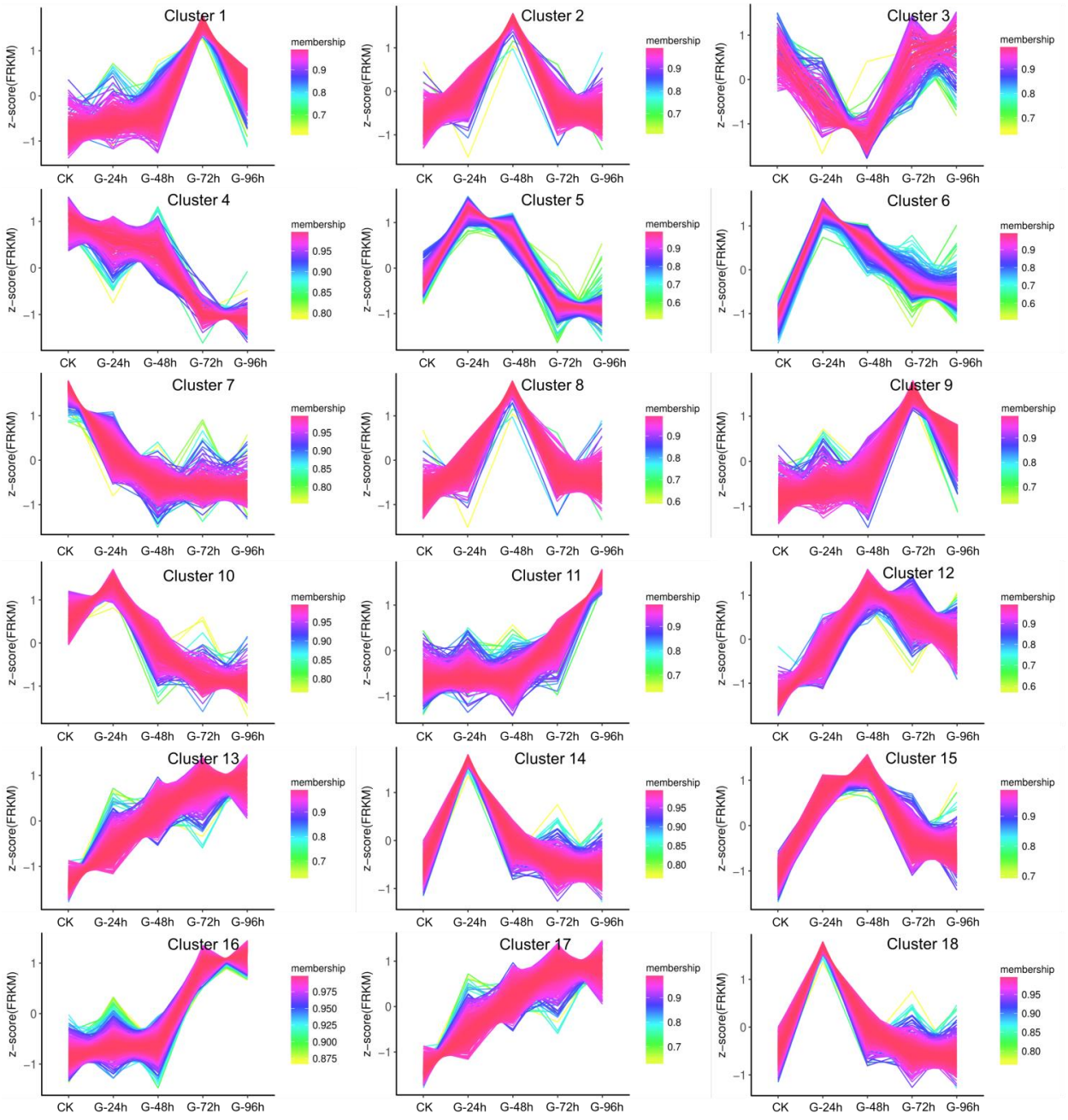

Figure 2. Unigenes expression pattern of different clusters in five samples.
The x-axis represents germination time. The y-axis represents normalized z-value RNA-seq expression data.

2.3. Functional analysis of DEGs

The differentially expressed genes identified at two stages of seed germination were compared with the Gene Ontology (GO) database to obtain detailed annotations of differentially expressed genes. Results showed that 15,795 differentially expressed genes were annotated in the GO database. A total of 54 GOs were detected into three categories: biological process (24), cellular component (17) and molecular function (13). Figure 3 showed the top 30 GOs with the most aggregated genes, including cellular component contains 9 GO, biological process contains 12 GO, mobile function contains 9 GOs. In the four stages after treatment, the GO with the highest concentration of differentially expressed genes were cell part, cellular process, binding, metabolic process, catalytic activity, organelle, membrane, biological regulation, membrane part and response to stimulation.

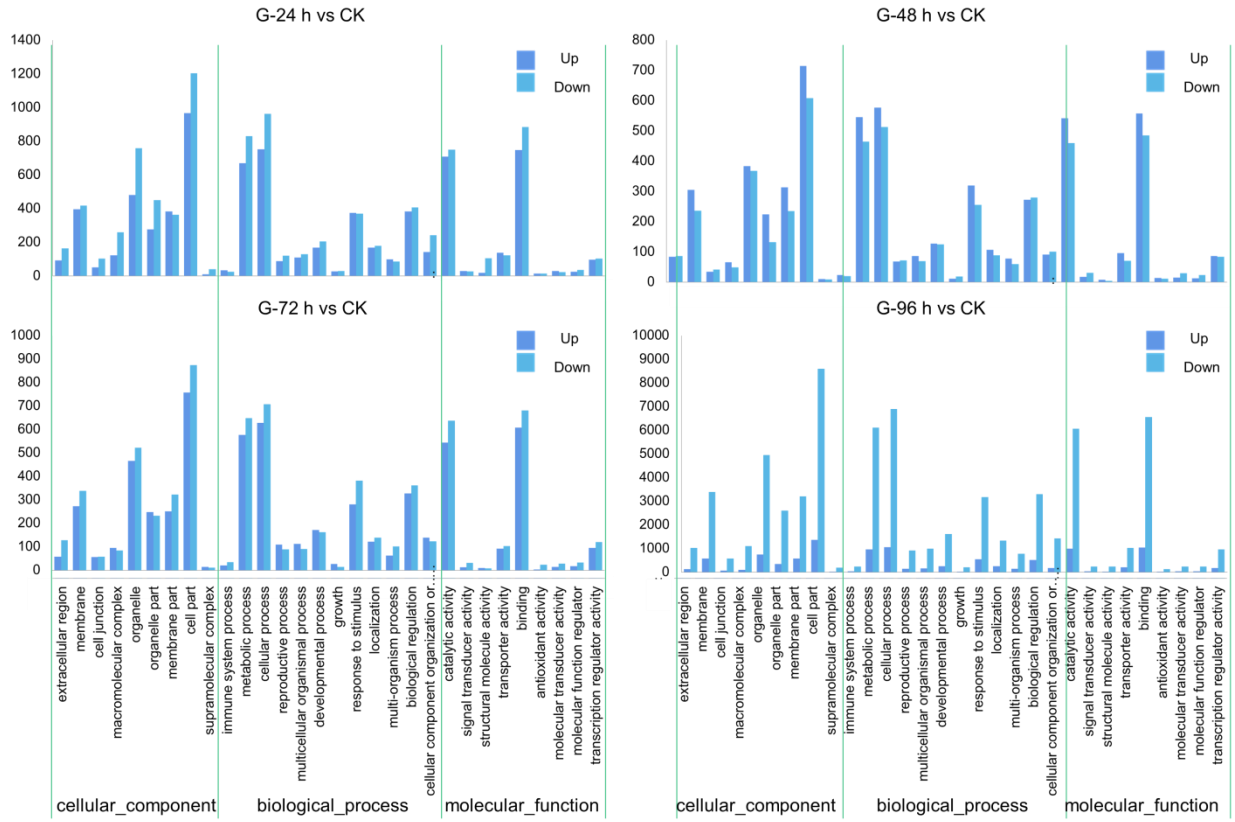

Figure 3. The GOs classification of all DEGs of germination seeds and controls.
The abscissa is the type of GO enrichment, and the ordinate is the number of differentially expressed genes in different GOs.

KEGG (Kyoto Encyclopedia of Genes and Genomes) pathways were divided into cellular processes, environmental information processing and genetic information processing, metabolism and organismal systems. A Q-value less than 0.05 was used as the significant enrichment threshold for all DEGs in the KEGG pathway. The DEGs of seed germination were enriched in 113 pathways. The 20 pathways with the largest number of genes enriched to differentially expressed genes were ribosome (75), plant hormone signal transduction (41), phenylpropanoid biosynthesis (34), carbon metabolism (30), plant-pathogen interaction (29), starch and sucrose metabolism (26), glutathione metabolism (25), amino sugar and nucleotide sugar metabolism (24), biosynthesis of amino acids (24), protein processing in endoplasmic reticulum (24), ubiquitin mediated proteolysis (21), pyrimidine metabolism (20), pentose and glucuronate interconversions (19), fatty acid metabolism (19), endocytosis (19), MAPK signaling pathway-plant (19), ribosome biogenesis in eukaryotes (17), cysteine and methionine metabolism (16), glycerophospholipid metabolism (15), glycolysis / gluconeogenesis (15).

2.4. Comparison of DEG Tag Data with qRT-PCR

To verify the expression profile obtained by transcriptome analysis, real-time PCR was performed. Nine DEGs were identified that expressed significantly differently between samples, they belonged to auxin-responsive protein, peroxidase, myb proto-oncogene protein and abscisic acid receptor

family, respectively. Overall, eight DEGs showed the same trend in transcriptomes and qRT-PCR (Figure 4). The coincidence rate of RNA-seq and qRT-PCR was 88.89%, indicating that RNA-seq had high accuracy and the identified pathways and candidate genes were reliable. No gene expression level was detected in some samples, indicating that DEG was not expressed at this stage.

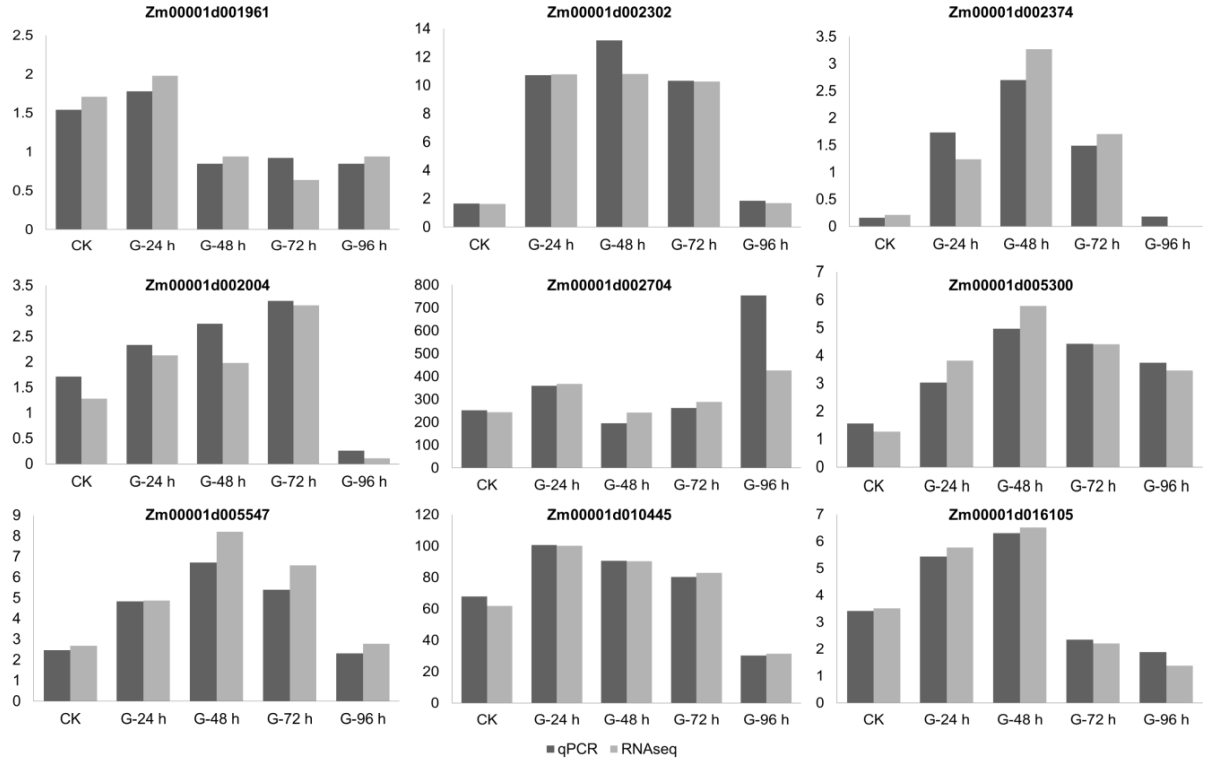

Figure 4. Expression profiles of DEGs which were selected based on the transcriptome analysis and qRT-PCR. The abscissa is the type of GO enrichment, and the ordinate is the number of differentially expressed genes in different GOs.

2.5. Analysis of hormone related genes expression during dormancy release in maize seeds

Hormones play a very important role in dormancy and germination. ABA has a positive regulatory effect on inducing and maintaining seed dormancy, while GA plays an important role in terminating seed dormancy and promoting germination. During seed imbibition and germination, the increase in ABA synthesis and the decrease in GA content maintain seed dormancy, thereby promoting seed germination. Based on transcriptome analysis results, compared with imbibited dormant seeds (CK) at 24h, 48h, 72h, 96h after germination, at least one set of comparative data from the three groups met the requirements of RFPM >1 and $P \leq 0.05$. Significant differences in the expression of 40 genes related to ABA, 15 genes related to GA, 60 genes related to ETH, and% genes related to auxin were obtained through comparison. The gene expression variations of key enzymes in some hormone signaling pathways are shown in Figure 5 and Figure 6.

NCED is a key enzyme in the ABA synthesis pathway, during the process of dormancy release in maize seeds, two genes (*Zm00001d007876*, *Zm00001d009286*) of NCED were detected. The expression level of *Zm00001d007876* showed a rapid downward trend. During the process of breaking dormancy, at 96 hours, its expression level rapidly increased, reaching 2.5 folds that of the non-germinating stage, and 7-9.8 folds that of the initial stage of germination, indicating that the gene was significantly suppressed during the process of breaking dormancy in seeds. Although the expression level of *Zm00001d009286* gene is at a relatively low level, it also shows a significant upward expression trend at the germination stage, indicating that the enhancement of this gene has a significant promoting effect on maize seed germination. CYP707A is a key enzyme in the ABA metabolism pathway, the expression of the two CYP707A genes has changed, but the change in their expression trend also shows significant differences. The expression of *Zm00001d005889* slowly decreases as the seeds break dormancy, and is significantly up-regulated at 96 h, reaching 7 folds as high as in the seeds, and even 30 folds as high as in the germination stage (48 h relative

to germination). The expression of *Zm00001d017762* gene was significantly up-regulated (6.1-8.3 fold) during the germination stage (24-72 h), and returned to the expression level in the seeds after germination, indicating that the gene may be specifically induced to express during the germination stage. Eight ABA receptor genes, including four PYL genes and five IDH3 genes, were detected and their expression levels were significantly up-regulated during the germination stage, regardless of the initial expression level. It is also related to ABA promoting seed dormancy breaking. Seventeen ABF genes were detected and showed rich variations, indicating that complex regulation exists during seed germination. It was also observed that most of the gene expression levels were significantly upregulated, consistent with existing research bases at the germination stage.

GA20x (*Zm00001d034898*) is a key gene in the GA synthesis pathway. During the dormancy release process of maize seeds, significant up-regulation of genes expression was detected. In the GA signal transduction pathway, GID1 is a GA receptor, and the expression of GID1 genes is induced by exogenous hexafluorouracil. Two GID1 genes (*Zm00001d010308*, *Zm00001d038165*) were detected during maize seed germination, and their expression was significantly up-regulated during seed germination. At 96h, their expression levels returned to normal levels. This indicates that GID1 gene has a significant regulatory role in inhibiting GA3 and promoting germination.

During the process of dormancy release of maize seeds under the action of exogenous hormones, DELLA protein is a class of nuclear proteins belonging to the GRAS family of transcription regulators, and is a very important inhibitor of GA signal transduction. In this study, seven DELLA protein genes were detected, two of which had high expression levels, and two of them were at low expression levels. During seed germination, four genes were significantly up-regulated and one gene was slightly down-regulated.

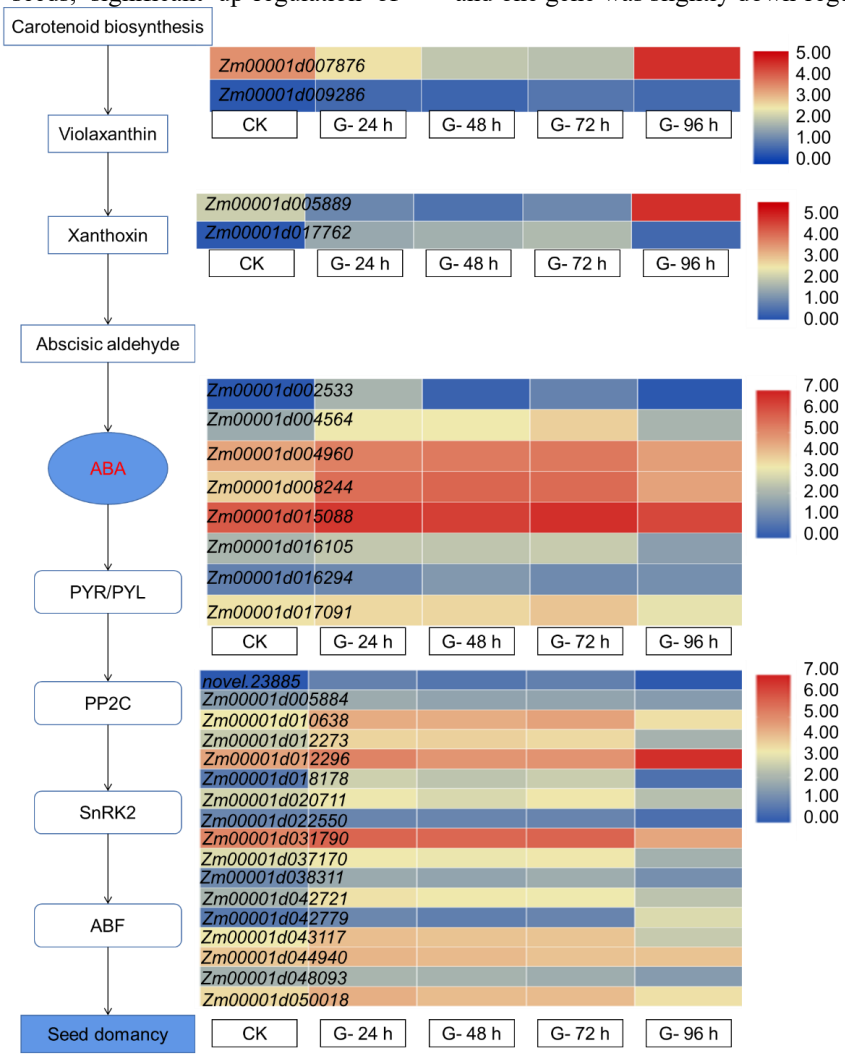

Figure 5. Expression trend of genes related to ABA during the process of seed dormancy release.

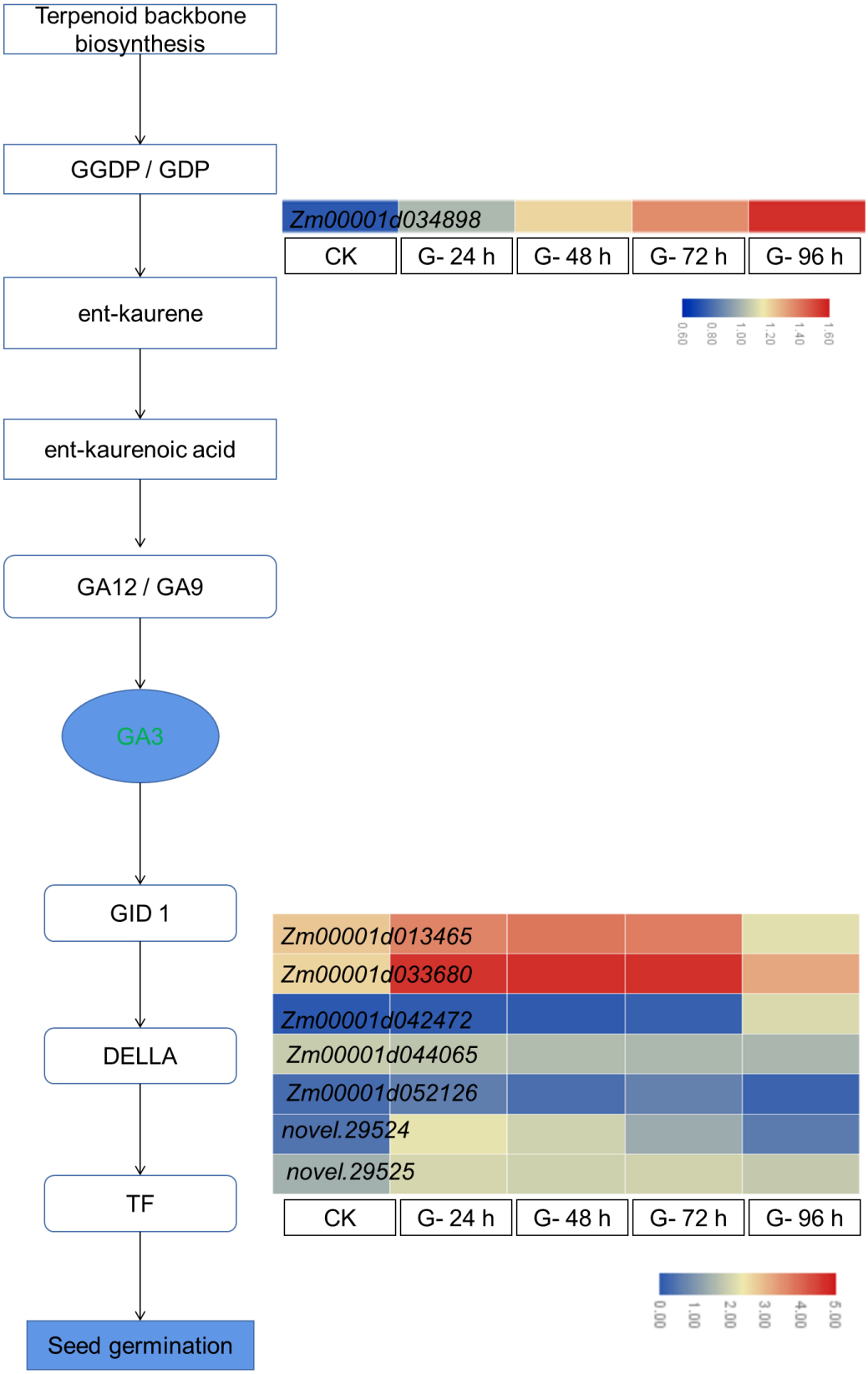

Figure 6. Expression trend of genes related to GA during the process of seed dormancy release.

3. Discussion

3.1. Construction of dormancy release transcriptome library of maize seeds

In recent years, with the continuous development of biotechnology, "omics" research has become a hot field of molecular biology, and its application in plants has become more and more extensive. Many scholars have conducted transcriptomics research on dormancy and germination of various plant seeds (Duan et al., 2023; Liao et al., 2022). The differential genes are mainly concentrated in plant hormone signal transduction, carbohydrate metabolism, starch and sucrose metabolism, ABC transporter, glycolysis, post-translational modification, ribosomal structure and biosynthesis and other biological processes, indicating that the germinated seeds can't provide enough energy for seed germination due to the obvious decline in respiratory efficiency, resulting in germination failure (Zaynab et al., 2021).

In this study, a new generation of high-throughput sequencing technology was used to sequence the transcriptome of the dormant and dormant samples of maize seeds. 41.05 Gb clean data was generated, and the volume of effective data for each sample ranged from 4.06 Gb to 4.12 Gb after removal of the adapter. A total of 15,795 DEGs with significant variation were identified among different stages of seeds, and 54 GOs were detected. In the four stages after treatment, the GO with the highest concentration of differentially expressed genes were cell part, cellular process, binding, metabolic process, catalytic activity, organelle, membrane, biological regulation, membrane part and response to stimulation. Further, there were ribosome, plant hormone signal transduction, carbon metabolism, starch and sucrose metabolism, glutathione metabolism, amino sugar and nucleo-tide sugar metabolism, fatty acid metabo-lism and MAPK signaling pathway-plant pathways that are more abundant in DEGs, and these pathways exhibit a high degree of similarity to research findings in other species.

3.2. Plant hormone signal transduction pathway participates in dormancy release of maize seeds

A large number of studies have found that plant hormones play an important regulatory role in the process of seed dormancy and germination, of which GA and ABA are the most important regulatory hormones (Shu et al., 2016). GAs promote seed germination, and ABA inhibit seed germination, the two have antagonistic effects, and affect each other's biological synthesis and signal transduction pathway (Zuo & Xu., 2020; Sano & Marion-Poll., 2021; Ali et al., 2022). In addition, other endogenous hormones, such as auxin (IAA), cytokinin (CTK), ethylene (ETH) and jasmonic acid (JA), regulate seed dormancy and germination by directly or indirectly participating in ABA or GA signal pathway (Sano & Marion-Poll., 2021).

In this study, 453 genes related to plant hormone metabolism (ABA, GA, IAA, etc.) were screened by comparing the transcriptome data of dormant and dormant maize seed samples, among which 157 genes were significantly differentially expressed. According to the significantly up-regulated expression of gibberellin biosynthesis gene (GA20ox3) and gibberellin receptor gene (GID, EDLLA) in GA metabolic signal transduction pathway, the significantly down-regulated expression of ABA biosynthesis (NCED6) and the significantly up-regulated expression of ABA catabolism (CYP701) gene in ABA metabolic signal transduction pathway. It is consistent with the continuous decrease of ABA content and the continuous increase of GA content during the dormancy of maize seeds. It shows that GA can promote the dormancy release of maize seeds, and ABA can antagonize GA and inhibit the germination of seeds. In the IAA metabolic signal transduction pathway, the expression of IAA biosynthesis gene (AUX1) and IAA receptor gene (TIR1) were significantly up-regulated, and the IAA content in its seeds also showed a gradual upward trend. Therefore, IAA may antagonize ABA, eliminate its inhibitory effect on germination, and promote seed germination. Genes in JA metabolic signal transduction pathway in maize seeds were significantly up-regulated, and their content gradually increased during dormancy release. Therefore, the dormancy release of maize seeds may be related to the change of JA content. In the maize seed transcriptome, the gene expression in the metabolic signal transduction pathway of CTK, ETH and BR was significantly up-regulated. We speculated that the content of related hormones increased and promoted the release of seed dormancy. In conclusion, plant hormone signal transduction pathway participates in the dormancy release of maize seeds and plays an important regulatory role.

3.3. Respiratory pathway metabolic pathway participates in dormancy release of maize seeds

Seeds undergo a complex process from maturity to germination and subsequent seedling establishment. Many factors affect seed germination, including the regulation of physiological and biochemical indicators, genes, and proteins (Chin et al., 2021; Yu et al., 2021). Protein metabolism, sugar metabolism, and lipid metabolism are all important metabolic processes during seed germination, providing an important material basis for seed germination (Li et al., 2022). Respiratory metabolism is the center of life activities, and its intensity and pathway changes affect and regulate the growth and development of plant organs. The three major respiratory pathways in plants, the tricarboxylic acid cycle (TCA) (Che-Othman et al.,

2020), the glycolytic pathway (EMP) (He et al., 2019), and the pentose phosphate pathway (PPP) (Lin et al., 2020), provide different energy and reducing power for their life activities. Therefore, seed dormancy is closely related to changes in germination and respiratory metabolic pathways.

In this study, by comparing transcriptome data from maize seed dormancy and dormancy release samples, a total of 764 genes related to the metabolism of the three major respiratory pathways were selected, including a total of 24 PGI genes, which are key enzymes in the glycolysis pathway, with significant upregulation of expression; There are 18 MDH genes that are key enzymes in the tricarboxylic acid cycle pathway, of which 12 are significantly up-regulated and 6 are significantly down-regulated; A total of 24 genes encoding the key enzyme G6PDH in the pentose phosphate pathway were significantly upregulated. In summary, the activation of metabolic pathways in the respiratory pathway provides the necessary materials and energy for dormancy release and germination of seeds, and promotes the germination of corn seeds.

4. Conclusion

Seeds serve a crucial function as the reproductive system and play a vital role in the life cycle of higher plants. Seed dormancy and germination are key to plant growth and development, and the transformation from dormancy to germination is closely related to agricultural production. In this study, by comparing transcriptome data from maize seed germination and non-germination samples, a total of 41.05 Gb clean data was generated. A total of 15795 DEGs were identified and compared with non-germinated maize seeds during the four germination stages, indicating significant expression and regulation changes in maize seeds during the germination stage. 54 GOs were detected, among which a large number of gene clusters were detected in cell part, cellular process, binding, metabolic process, catalytical activity, organelle, memory, biological regulation, memory part, and response to stimulation. These genes are significantly clustered in pathways such as ribosomes, plant hormone signal transduction, carbon metabolism, start and cross metabolism, glutathione metabolism, amino sugar and nucleo-tide sugar metabolism, fat acid metabolism, and MAPK signaling path path path paths. In short, under the regulation of hormones, maize seeds have undergone significant expression changes during the germination process, using the metabolism of stored proteins, starch, and lipids, and using the activation of metabolic pathways in the respiratory pathway to provide the necessary material and energy for seed dormancy release and germination, and promoting the germination of maize seeds.

Author Contributions: Conceived and designed the experiments: Dengxiang Du and Guoqing Dong. Performed the experiments: Xueqin Hu, Juan Zou and Jinhua Li. Analyzed the data: Wanyong Zeng and Dengxiang Du. Wrote the paper: Dengxiang Du and Guoqing Dong.

Conflicts of Interest: The authors declare no conflict of interest.

References

[1]. Ali F, Qanmber G, Li F, Wang Z. 2022. Updated role of ABA in seed maturation, dormancy, and germination [J]. Journal of Advanced Research, 35: 199-214. doi.org/10.1016/j.jare.2021.03.011

[2]. Ashburner M, Ball C A, Blake J A, Botstein D, Butler H, Cherry J M, ..., Sherlock G. 2000. Gene ontology: tool for the unification of biology [J]. Nature genetics, 25(1): 25-29. doi.org/10.1038/75556

[3]. Che-Othman M H, Jacoby R P, Millar A H, Taylor N L. 2020. Wheat mitochondrial respiration shifts from the tricarboxylic acid cycle to the GABA shunt under salt stress [J]. New Phytologist, 225(3): 1166-1180. doi.org/10.1111/nph.15713

[4]. Chen C, He B S, Liu X X, et al. 2019. Pyrophosphate-fructose 6-phosphate 1-phosphotransferase (PFP1) regulates starch biosynthesis and seed development via heterotet-ramer formation in rice (*Oryza sativa* L.) [J]. Plant Biotechnology Journal, 18(1): 83-95. doi.org/10.1111/pbi.13173

[5]. Cheng X, Gao C, Liu X, Xu D, Pan X, Gao W, ..., Ma C. 2022. Characterization of the wheat VQ protein family and expression of candidate genes associated with seed dormancy and germination [J]. BMC Plant Biology, 22(1):1-18. doi.org/10.1186/s12870-022-03430-1

[6]. Chin J M, Lim Y Y, Ting A S Y. 2021. Biopolymers for biopriming of *Brassica rapa* seeds: A study on coating efficacy, bioagent viability and seed germination [J]. Journal of the Saudi Society of Agricultural Sciences, 20(3): 198-207. doi.org/10.1016/j.jssas.2021.01.006

[7]. Duan X, Jiang W, Wu K, Chen J, Li Y, Tao Z. 2023. Integrating Transcriptomics and Hormones Dynamics Reveal Seed Germination and Emergence Process in Polygonatum cyrtonema Hua [J]. International Journal of Molecular Sciences, 24(4): 3792.doi.org/10.3390/ijms24043792

[8]. Graeber K A I, Nakabayashi K, Miatton E, Leubner-Metzger G E R H A R D, Soppe W J. 2012. Molecular mechanisms of seed dormancy [J]. Plant, Cell & Environment, 35(10): 1769-1786. doi.org/10.1111/j.1365-3040.2012.02542.x

[9]. Han Z, Wang B, Tian L, Wang S, Zhang J, Guo S, ..., Chen Y. 2020. Comprehensive dynamic transcriptome analysis at two seed germination stages in maize (*Zea mays* L.) [J]. Physiologia Plantarum, 168(1): 205-217. doi.org/10.1111/ppl.12944

[10]. He Y, Cheng J, He Y, Yang B, Cheng Y, Yang C, ..., Wang Z. 2019. Influence of isopropylmalate synthase *OsIPMS 1* on seed vigour associated with amino acid and energy metabolism in rice [J]. Plant Biotechnology Journal, 17(2): 322-337. doi.org/10.1111/pbi.12979

[11]. Kanehisa M, Goto S. 2000. KEGG: kyoto encyclopedia of genes and genomes [J]. Nucleic acids research, 28(1): 27-30. doi.org/10.1093/nar/28.1.27

[12]. Kim D, Paggi J M, Park C, Bennett C, Salzberg S L. 2019. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype [J]. Nature Biotechnology, 37(8): 907-915. doi.org/10.1038/s41587-019-0201-4

[13]. Koramutla M K, Negi M, Ayele B T. 2021. Roles of glutathione in mediating abscisic acid signaling and its regulation of seed dormancy and drought tolerance [J]. Genes, 12(10): 1620. doi.org/10.3390/genes12101620

[14]. Li H, Chen J, He L, Zhu H, Huang Z, Zhu M, ..., Yan J. 2022. Transcriptome Analyses Reveal the Role of Light in Releasing the Morphological Dormancy of Celery Seed by Integrating Plant Hormones, Sugar Metabolism and Endosperm Weakening [J]. International Journal of Molecular Sciences, 23(17): 10140. doi.org/10.3390/ijms231710140

[15]. Li S S, Wang L S, Shu Q Y, Wu J, Chen L G, Shao S, Yin D D. 2015. Fatty acid composition of developing tree peony (*Paeonia section Moutan* DC.) seeds and transcriptome analysis during seed development [J]. BMC Genomics, 16: 1-14. doi.org/10.1186/s12864-015-1429-0

[16]. Liao D, Chen Y, Qi J, Zhang H, Sun P, Chen C. 2022. Temporal transcriptomics reveal the molecular mechanism of dormancy and germination regulated by temperature in Paris polyphylla seed [J]. Horticultural Plant Journal. doi.org/10.1016/j.hpj.2022.06.003

[17]. Lin D, Huang Y, Zhao J, Wu Z, Liu S, Qin W, ..., Zhang Q. 2020. Evaluation of seed nitrate assimilation and stimulation of phenolic-linked antioxidant on pentose phosphate pathway and nitrate reduction in three feed-plant species [J]. BMC Plant Biology, 20: 1-12. doi.org/10.1186/s12870-020-02453-w

[18]. Mangrauthia S K, Agarwal S, Sailaja B, Sarla N, Voleti S R. 2016. Transcriptome analysis of *Oryza sativa* (rice) seed germination at high temperature shows dynamics of genome expression associated with hormones signalling and abiotic stress pathways [J]. Tropical Plant Biology, 9(4): 215-228. doi.org/10.1007/s12042-016-9170-7

[19]. Reed R C, Bradford K J, Khanday I. 2022. Seed germination and vigor: ensuring crop sustainability in a changing climate [J]. Heredity, 1-10. doi.org/10.1038/s41437-022-00497-2

[20]. Sano N, Marion-Poll A. 2021. ABA metabolism and homeostasis in seed dormancy and germination [J]. International Journal of Molecular Sciences, 22(10): 5069. doi.org/10.3390/ijms22105069

[21]. Shu K, Liu X D, Xie Q, He Z H. 2016. Two faces of one seed: hormonal regulation of dormancy and germination [J]. Molecular Plant, 9(1), 34-45. doi.org/10.1016/j.molp.2015.08.010

[22]. Shu K, Meng Y J, Shuai H W, Liu W G, Du J B, Liu J, Yang W Y. 2015. Dormancy and germination: How does the crop seed decide? [J]. Plant Biology, 17(6): 1104-1112. doi.org/10.1111/plb.12356

[23]. Tuan P A, Kumar R, Rehal P K, Toora P K, Ayele B T. 2018. Molecular mechanisms underlying abscisic acid/gibberellin balance in the control of seed dormancy and germination in cereals [J]. Frontiers in Plant Science, 9: 668. doi.org/10.3389/fpls.2018.00668

[24]. Wu M, Wu J, Gan Y. 2020.The new insight of auxin functions: transition from seed dormancy to germination and floral opening in plants [J]. Plant Growth Regulation, 91: 169-174. doi.org/10.1007/s10725-020-00608-1

[25]. Yan A, Chen, Z. 2020. The control of seed dormancy and germination by temperature, light and nitrate [J]. The Botanical Review, 6(1): 39-75. doi.org/10.1007/s12229-020-09220-4

[26]. Yu F, Li M, He D, Yang P. 2021. Advances on post-translational modifications involved in seed germination [J]. Frontiers in Plant Science, 12: 362. doi.org/10.3389/fpls.2021.642979

[27]. Yu Y, Zhen S, Wang S, Wang Y, Cao H, Zhang Y, ..., Yan Y. 2016. Comparative transcriptome analysis of wheat embryo and endosperm responses to ABA and H_2O_2 stresses during seed germination [J]. BMC Genomics, 17(1): 1-18. doi.org/10.1186/s12864-016-2416-9

[28]. Zaynab M, Pan D, Fatima M, Sharif Y, Chen S, Chen W. 2021. Proteomics analysis of Cyclobalanopsis gilva provides new insights of low seed germination [J]. Biochimie, 180: 68-78. doi.org/10.1016/j.biochi.2020.10.008

[29]. Zhao Y, Li M C, Konaté M M, Chen L, Das B, Karlovich C, ..., McShane L M. 2021. TPM, FPKM, or normalized counts? A comparative study of quantification measures for the analysis of RNA-seq data from the NCI patient-derived models repository [J]. Journal of Translational Medicine, 19(1): 1-15. doi.org/10.1186/s12967-021-02936-w

[30]. Zuo Y, Xu Y. 2020. Research Progress on the Mechanism of GA and ABA during Seed Germination [J]. Molecular Plant Breeding, 11. doi.org/10.5376/mpb.2020.11.0020.

5/22/2023