Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

A Comparative Transcriptome Analysis Reveals New Insights into Pre-Harvest Sprouting (PHS) in Wheat

Xingyan Li
Inner Mongolia Agricultural University https://orcid.org/0000-0002-1268-5540
Hongyu Qiao
Inner Mongolia Agricultural University
Zeng Wang
Inner Mongolia Agricultural University
Bing Han
Inner Mongolia Agricultural University
Yanping Xing
Inner Mongolia Agricultural University
Yan Yang (yangyanchutao@126.com)
Inner Mongolia Agricultural University

Research

Keywords: pre-harvest sprouting, Triticum aestivum L., transcriptome sequencing, Chitinase.

Posted Date: September 27th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-910461/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Abstract

Background: pre-harvest sprouting (PHS) is a significant cause of yield loss in cereal crops, and is an important topic of study for the improvement of wheat quality. Many studies have focused on PHS in wheat during the last 10 years, especially on the involvement of abscisic acid (ABA) in PHS, however, a lot remained unknown about this topic.

Results: In this study, a PHS resistant line was isolated from an ethylmethane sulfonate (EMS) mutant population derived from the wheat cultivar 'Long 13-3778', namely 'LQ18'. The mutant line LQ18 showed highly significant resistance to PHS compared with the wild-type. Transcriptome sequencing was conducted to determine the differences between the LQ18 mutant and the wild-type at the level of gene expression. The results showed no conclusive evidence that the ABA biosynthesis and signaling pathways contribute to the differences in PHS between the mutant and the wild-type, and some genes and their alleles associated with PHS tolerance showed differential expression between the mutant and wild-type lines. The most interesting result of this study was that the expression levels of the chitinase family genes showed significant differences between the mutant and the wild-type as determined by GO enrichment analysis, and a subsequent analysis of differential expression profiling of the chitinase genes led to the same conclusion.

Conclusions: Transcriptomic analysis in this study have revealed the global transcriptome profiles of the PHS sensitive wheat cultivar 'Long 13-3778' and its PHS resistant mutants. Furthermore, this study has proposed a possible explanation of the connection between PHS and the chitinase family for the first time, which added to our understanding of PHS and seed dormancy in common wheat.

Background

Pre-harvest sprouting refers to the germination of grains on mature spikes before harvest, usually in wet weather, which generally causes yield losses and reduces the end use of cereals[1]. Accordingly, PHS tolerance has been an important research topic for quality improvement in wheat. It is thought that PHS occurs because of the lack of dormancy in cereal seeds, and the results of many studies have been used to explain the mechanisms of seed dormancy in an effort to relieve PHS.

Sensitivity to abscisic acid (ABA) has been frequently found to correlate with seed dormancy and to further affect PHS tolerance[2,3]. ABA content was originally thought to have little effect on seed dormancy, but the results of recent studies have shown that ABA metabolism may play a part in dormancy release in after-ripened grains[2,4,5,6]. Changes in ABA content have been found to correlate with the timing of germination of imbibed dormant seeds[2,4]. Gibberellic acid (GA) was found to have an effect on promoting germination, and there is evidence that a high concentration of GA can release seeds from dormancy. Studies have shown that the GA concentration is negatively correlated with ABA levels, and the balance between the two hormones determines the timing of dormancy release[1]. Considering the complex crosstalk among phytohormones, this could be another complicated research topic.

Recent studies of the effect of ABA on dormancy further confirmed the earlier findings. Studies involving PHS in rice found multiple genes that regulate seed dormancy through ABA pathways. These genes include *PHS9*, which regulates PHS by adjusting the ABA sensitivity[5]; *OsMFT2*, an ABA signaling gene that regulates the sensitivity to ABA by interacting with three OsbZIP transcription factors[6]; and *OsDOG1L-3*, a gene that is induced by ABA and positively regulates ABA pathways. Over-expression of *OsDOG1L-3* results in enhanced seed dormancy[4]. Recent studies have provided evidence that GA participates in seed dormancy through the action of the microRNA *miR156*, which is an important regulator of GA pathways[1].

Other than the regulation of seed dormancy that is controlled by the antagonistic action of the phytohormones ABA and GA, quantitative trait locus (QTL) analysis has identified a correlation between coat color and seed dormancy in wheat[7–11]. In *Sorghum bicolor*, although the correlation between coat color and germination could not be rigorously confirmed, studies still found relationships between seed coat color and germination in tannin-less sorghums[3].

QTL localization has identified several major QTLs that control PHS in wheat. A wheat *MOTHER OF FT AND TFL* (*MFT*) homolog was identified in a wheat microarray assay of mature seeds grown at different temperatures as a candidate gene that may contribute to the high dormancy level of seeds grown at low temperature. Transient expression studies confirmed the important role of *MFT* in the control of seed dormancy[12]. Later studies in rice further defined the role of *MFT* and its effect on seed dormancy. Studies have shown that *OsMFT2* expression enhances rice seed dormancy by interacting with the transcription factors OsbZIP23/66/72 and altering the sensitivity to ABA[6]. *Phs1*, another major locus responsible for seed dormancy in wheat, was identified by QTL analysis, and a candidate gene that may regulate the dormancy functions of the QTL *phs1* is a *MITOGEN-ACTIVATED PROTEIN KINASE KINASE 3* (*MKK3*) gene, designated *TaMKK3-A*. Transformation of the *TaMKK3-A* allele from a non-dormant cultivar to a dormant cultivar significantly relieved its seed dormancy[13]. Studies in wheat also showed that the expression of an *ABSCISIC ACID-INSENSITIVE 5* (*ABI5*) gene, *TaABI5*, was much higher in the PHS resistant cultivar SHW-L1 than in PHS susceptible cultivar 'Chuanmai 32'. Four expression quantitative trait loci (eQTL) of *TaABI5* are located on chromosomes 2DS, 4DS, 6DS, and 7DL, and explained 13%-46% of the total phenotypic variation[14]. The *Delay of Germination 1* (*DOG1*) gene is responsible for seed dormancy in Arabidopsis, and ectopic overexpression of its homologs from wheat and barley, *TaDOG1L1* and *HvDOG1L1*, respectively, significantly enhanced seed dormancy in Arabidopsis[15]. In addition, the rice homolog of Arabidopsis *DOG1*, *OsDOG1L-3*, is responsible for the function of the QTL *qsd-1-1*, which has a large effect on seed dormancy in rice. Expression of *OsDOG1L-3* enhanced seed dormancy and was enhanced by ABA[4].

A number of previous studies have reported research results on PHS and seed dormancy in wheat; however much remains to be discovered about the regulation of PHS. In our study, we obtained two individual M_2 plants carrying PHS resistant mutations, namely LQ1813 and LQ1815, that were isolated from the cultivar 'Long 13-3778' by EMS treatment. The mutant plants showed significantly lower levels of PHS compared to the wild-type. A comprehensive transcriptomic analysis was performed to discover the molecular basis of PHS-resistance in the mutants. Our objective is to provide directions for additional focused research. This study aims to broaden our overall understanding of seed dormancy and PHS, and to provide resources for future studies.

Materials And Methods

Plant materials and sampling

The wheat cultivar 'Long 13-3778' was obtained from the Heilongjiang Academy of Agricultural Sciences. The mutants were generated by EMS treatment and the M_2 mutant pool was constructed with 3,000 mutant individuals. The wild-type and the M_2 mutants were planted in an experimental plot at Inner Mongolia Agricultural University in April 2018. Row spacing was 20 cm, and the seeds were sown 15 cm apart. Fully mature seeds were collected at 100–105 d after sowing. The seeds used for determining germination index (GI were collected and treated according a published method (Yang et al. 2007)[34]. The seeds used for transcriptome sequencing were frozen in liquid nitrogen immediately after sampling, and stored at -80°C prior to RNA extraction.

Agronomic trait statistics

The seeds of 23 selected M_3 plants (LQ01-LQ23) and the wild-type 'Long13-3778' were planted in an experimental plot at Inner Mongolia Agricultural University in April 2019 using the same methods described above. The entire plants were removed from the plot when fully mature, and the values for the following agronomic traits were recorded; tiller number, plant height, GI, first stem length and width, second stem length and width, and third stem length and width.

RNA-sequencing and library construction

Fully mature seeds of the mutants and the wild-type 'Long13-3778' were collected and stored as described above. Seeds of the M2 mutants LQ1813 and LQ1815 were chosen for transcriptome sequencing, and three biological replicates were used. Transcriptome sequencing and analysis were performed by OE Biotech Co., Ltd. (Shanghai, China). First, total RNA was

extracted using the mirVana™ miRNA ISOlation Kit (Ambion-1561, also capable of mRNA extraction) following the manufacturer's instructions. The total RNA was treated with DNase to remove contaminating DNA, and mRNA was enriched by using Oligo DT magnetic beads. The mRNA was broken into short fragments and then used as a template for the synthesis of first-strand cDNA using six-base random primers. Second-strand cDNA was synthesized, and the double-stranded cDNA was purified, end-repaired, and ligated with a poly-A tail and the adapter to produce the sequencing library. An Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA) was used to determine the quality of the libraries. RNA-sequencing was conducted on an Illumina HiSeqTM platform (Illumina, San Diego, CA). Trimmomatic[35] software was used for quality control, including removing the adaptor sequences and low-quality reads. Statistics for the raw and clean reads, percent quality scores Q30, and the percent GC were analyzed.

Mapping and gene expression analysis

The clean reads were mapped to the wheat genome reference from IWGSC (Triticum_aestivum. IWGSC.dna.toplevel) using hisat2[36]. Htseq-count[37] and cufflinks[38] were used for gene expression analysis and to identify differentially expressed genes, respectively. DESeq[39] was used for the normalization of read counts, fold-change calculation, and negative binomial distribution tests. Gene expression comparisons with adjusted p value < 0.05 and foldchange > 2 were determined to be differentially expressed.

Members of the *TaDOG*, *TaSdr*, *PM19*, and *TaQsd1* gene families were identified in IWGSC based on the existing references. Certain members of these gene families have been reported to affect seed dormancy and PHS. Expression profiles of the genes are given in Table 1.

Members of the *NCED* family, the abscisic acid receptor *PYL* family, and the *ABI 5-like* family were identified. The numbers of DEGs in the gene families in each comparison were counted. The DEG ratios were calculated and compared with the DEG ratios for all the genes.

GO and KEGG analyses

GO and KEGG enrichment analyses were conducted. Annotation files from IWGSC were used as annotation references (Triticum_aestivum.IWGSC.43.gff3). The genes that showed differential expression between the wild-type and both mutants were designated as the intersection genes. We performed a GO classification analysis of the intersection genes and the total genes. The ratios of each GO term at level 2 were calculated and compared between the two clusters. DAGs were generated for the GO enrichment in the three major GO categories “biological process”, “cellular component”, and “molecular function” of the intersection genes. A KEGG pathway classification was also performed, and the KO (KEGG Ontology) ratios were compared between the intersection genes and the total genes at level 2, similar to the GO classification.

Results

Agronomic traits of the LQ18 mutants and the wild-type

The average length of the first, second ,and third stems on M_2 plants of the mutant LQ18 were 4.81 cm, 10.74 cm, and 15.8 cm, respectively, which were significantly higher than in the wild-type (3.19 cm, 7.57 cm, and 12.19 cm). The average germination index (GI) value of the mutants was 0.04, which was significantly lower than that of the wild-type (0.43). There were no significant differences in agronomic traits between the mutant strain and the wild-type other than stem length and GI (Fig. 1, additional file 1).

RNA-sequencing of the wild-type ‘Long 13-3778’ and its mutants

An RNA-sequencing was performed to identify the global differences between the transcriptomes of the mutants and the wild-type. Two individual M_3 selfings of LQ18 (namely LQ1815 and LQ1813, respectively) and the wild-type ‘Long 13-3778’ were chosen for sampling. Three replicates were used in the experiment. An average of 11.30 GB of clean bases was created for each sample, with the Q30 scores > = 93.92%. 64.96% of the total reads were mapped uniquely in average.

A principal component analysis was conducted based on the gene expression levels of each sample (Fig. 2A). Legends that represented the replicates of the same sample were closely plotted, and according to the results of the analysis, the global expressional differences between the mutants LQ1815 and LQ1813 was smaller than the differences between the wild-type and each mutant. The results indicated a good repeatability for the sequencing, and the global transcription profile of the two mutants were comparatively similar. Additionally, a sample-to sample cluster analysis has come to a similar result (Fig. 2B).

ABA biosynthesis and signaling

A comparative transcriptome analysis was conducted to identify the deferentially expressed genes (DEGs) in the ABA synthesis and signaling pathways between the mutants and the wild-type. The results came out that the DEG ratio for the *9-cis-EPOXYCAROTENOID DIOXYGENASE* (*NCED*) gene family, which encodes the key enzyme in ABA biosynthesis, was arguably higher than in the total genes (Fig. 3A), which suggests that ABA biosynthesis may play a more important role in the mutant LQ18 than in the wild-type. However the chi-square test results, due to the insufficient sample number (P = 0.05), showed that the expression profiles of the *9-NCED* genes were not significantly different between the two mutants and the wild-type. In addition, no strong evidence of differences in ABA signaling between the wild-type and the mutants could be found based on the expression profiles of genes for abscisic acid receptor PYL families and abscisic acid-insensitive 5-like (ABI5-like) protein families, which are key components of ABA signaling(Fig. 3A). However, the expression of one *ABI5*-like gene (ID = TraesCS2A02G099400) was significantly higher in both mutant plants compared with the wild-type (additional file 2). Additionally, the ratio of DEGs of the genes that contained the gene ontology (GO) term "abscisic acid" was calculated, which showed no significant differences compared with the ratio of DEGs of the total genes (Fig. 3B). In summary, no conclusive evidence could be found that the ABA biosynthesis and signaling pathways contribute to the differences in PHS between the mutant and the wild-type in this study.

PHS-related gene expression profiling

The expression profiles of the genes in wheat that were previously reported to have obvious effects on PHS regulation[12–15, 16–21] are shown in Table 1. Nine genes and their alleles located on the wheat A, B, and D genomes were chosen based on the results of previous studies. Most genes did not show a particular pattern of expression, and no firm conclusions could be made; however, a few genes and their alleles did show differential levels of transcript abundance between the mutants and the wild-type.

• TaDOG

The expression of the alleles of the *TaDOG* gene located on chromosomes 3A (TraesCS3A02G103500) and 3D (TraesCS3D02G105800) were higher in the wild-type than in both of the mutants. Transcription of all *TaDOG* alleles remained at a comparatively low level.

• TaSdr

The expression of *TaSdr* alleles located on chromosomes 2A (TraesCS2A02G191400), 2B (TraesCS2B02G215300), and 2D (TraesCS2D02G196200) showed consistently lower expression in mutant plants compared to the wild-type. Transcription of all alleles of *TaSdr* remained at a comparatively low level.

• PM19

Expression of the *PM19* alleles located on chromosome 4A (TraesCS4A02G313300 and TraesCS4A02G313400) was much higher in the wild-type than in the mutants. Of these, TraesCS4A02G313300 showed a significant difference in transcript levels between the mutants and the wild-type ($p < 0.05$, fold change > 2).

• TaQsd1

The alleles of TaQsd1,TraesCS5B02G214700 and TraesCS5D02G224200, located on chromosomes 5B and 5D, respectively, showed significantly higher expression levels in the mutant LQ1815 ($p < 0.05$, fold change > 2) than in wild-type. Although no

scientific conclusions could be reached from this, the expression of the two alleles in LQ1813 was also higher than in wild-type.

Table 1
Expression profile of PHS associated genes and their alleles.

gene_ID	gene_family	FPKM_wildtype	FPKM_LQ1815	FPKM_LQ1813	Sig_CvT15	Sig_CvT13
TraesCS3D02G004100	*MFT*	14.49 ± 4.74	9.97 ± 1.59	9.26 ± 1.85	–	–
TraesCS3B02G010100	*MFT*	13.28 ± 3.69	6.76 ± 1.26	7.67 ± 2.11	Down	–
TraesCS3A02G006600	*MFT*	57.56 ± 8.20	41.82 ± 6.59	49.20 ± 7.38	–	–
TraesCS5B02G565100	*Mkk3*	2.06 ± 0.16	2.87 ± 0.31	2.84 ± 0.46	–	–
TraesCS5D02G549600	*Mkk3*	1.45 ± 0.14	1.76 ± 0.20	1.79 ± 0.29	–	–
TraesCS3A02G063600	*Mkk3*	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	–	–
TraesCS3A02G417300	*TaVp*	7.12 ± 0.47	10.45 ± 0.46	9.07 ± 1.02	–	–
TraesCS3B02G452200	*TaVp*	20.71 ± 2.24	23.99 ± 2.10	24.85 ± 1.17	–	–
TraesCS3D02G412800	*TaVp*	15.67 ± 0.75	20.55 ± 1.44	17.58 ± 1.03	–	–
TraesCS3B02G120900	*TaDOG*	0.09 ± 0.16	0.00 ± 0.00	0.06 ± 0.07	–	–
TraesCS3D02G105800	*TaDOG*	1.11 ± 0.61	0.07 ± 0.05	0.10 ± 0.02	Down	Down
TraesCS3A02G103500	*TaDOG*	1.70 ± 0.32	0.06 ± 0.05	0.20 ± 0.02	Down	Down
TraesCS2A02G191400	*TaSdr*	5.83 ± 0.65	1.92 ± 0.38	2.34 ± 0.37	Down	Down
TraesCS2B02G215300	*TaSdr*	3.31 ± 0.61	1.67 ± 0.25	1.34 ± 0.12	Down	Down
TraesCS2D02G196200	*TaSdr*	1.66 ± 0.27	0.72 ± 0.14	0.78 ± 0.41	Down	Down
TraesCS3D02G468400	*Tamyb10*	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	–	–
TraesCS3B02G515900	*Tamyb10*	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	–	–
TraesCS5D02G558800	*PM19*	535.99 ± 23.38	654.22 ± 35.25	518.10 ± 16.75	–	–
TraesCS4A02G313300	*PM19*	61.28 ± 5.75	8.82 ± 1.35	6.67 ± 0.47	Down	Down
TraesCS4A02G313400	*PM19*	81.80 ± 7.26	167.90 ± 14.24	119.06 ± 10.12	–	–
TraesCS3B02G404300	*ABI5*	35.87 ± 4.44	48.80 ± 4.90	45.03 ± 0.91	–	–
TraesCS3B02G404400	*ABI5*	5.46 ± 0.80	3.20 ± 0.21	5.46 ± 0.08	Down	–
TraesCS3B02G404500	*ABI5*	6.49 ± 2.82	11.03 ± 0.55	6.69 ± 0.47	–	–
TraesCS3B02G404200	*ABI5*	5.18 ± 0.65	3.59 ± 0.20	4.34 ± 0.12	–	–
TraesCS3A02G371800	*ABI5*	31.47 ± 4.64	62.03 ± 5.90	48.40 ± 2.86	–	–
TraesCS5A02G216200	*TaQsd1*	25.90 ± 0.64	58.63 ± 2.91	48.84 ± 1.40	–	–
TraesCS5D02G224200	*TaQsd1*	18.13 ± 1.09	53.40 ± 4.92	31.43 ± 1.92	Up	–
TraesCS5B02G214700	*TaQsd1*	15.25 ± 0.99	39.77 ± 2.19	28.71 ± 1.58	Up	–

"Sig_CvT15" and "Sig_CvT13" indicate the significance of the expression of LQ1815 and LQ1813 compared to the wild-type, respectively. "UP" and "DOWN" indicate up-regulation and down regulation (P = 0.05).

Results of the GO and KEGG analysis

The genes that showed differential expression between the wild-type and both mutants (LQ1815 and LQ1813) were designated as the intersection genes, which reduced the impact of deviations and better represented the differences between the wild-type and the mutants. A GO classification was conducted to compare the GO enrichment at level 2 (The second layer of GO terms) between the intersection genes and the total genes. A total of 2,967 and 67,855 GO terms were enriched in the target gene group and the total gene group, respectively. The percentage of each GO enrichment to the total GO enrichment was calculated for the intersection genes and the total genes and compared to the other to determine the overall differences in GO enrichment between the two groups. As shown in Fig. 4A, the results indicated a matched enrichment level for most of the GO terms in “biological process” and “cellular component”, while there was an obvious difference in the “molecular function” category. The GO terms that were highly different between the two groups were “cell killing”, “localization”, “nucleoid”, “electron carrier activity”, “molecular transducer activity”, “nutrient reservoir activity”, “protein binding transcription factor activity”, “protein tag”, “receptor activity”, and “structural molecule activity”.

Directed acyclic graphs (DAGs) of the GO enrichment analysis of the intersection genes were generated (additional file 3) to illustrate the differences between the GO enrichment in the mutants and the wild-type in the three major GO categories “biological process”, “cellular component:, and “molecular function”. Based on the DAGs for “biological process” and “molecular function”, we concluded that the pathways involved in chitin catabolism were significantly enriched, which indicates a major difference in chitin catabolism activity between the two PHS mutants and the wild-type.

A Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway classification analysis was performed similar to the GO classification. A total of 27,732 and 732 KEGG ontology (KO) terms were enriched in the intersection genes group and the total genes group at level 2, respectively. The numbers and percentages of genes enriched for 23 KOs for the two groups are shown in Fig. 4B. Some KOs showed noticeable differences between the two groups. These were: translation, transcription, xenobiotics biodegradation and metabolism, metabolism of other amino acids, lipid metabolism, carbohydrate metabolism, and amino acid metabolism.

Expression profiling of the chitinase family genes and the WRKY transcription factor genes in the target gene group

Because a significant GO enrichment was found for genes involved in chitin catabolism, the expression profiles of the chitinase family genes in the intersection genes will be studied in greater depth. Additionally, the expression profiles of genes encoding WRKY transcription factors in the intersection genes will also be studied, as WRKYs have been frequently reported to be associated with plant defense and the stress response, and function as regulators of chitinase gene expression. In this study, we identified 26 chitinase genes and 18 WRKY transcription factor genes in the target gene group, all of which consistently showed significantly higher expression levels in the wild-type than in the mutants (Fig. 5).

Discussion

Chitinase like proteins (CLPs) are members of the glycosyl hydrolase 18 and 19 families and are grouped into six classes. These proteins are usually known to function in plant defense against pathogens and biotic stresses. CLPs from several plant species, such as chickpea and field bean, have been shown to have anti-fungal activities[22–24]. There is also evidence that CLPs may play a role in plant tolerance to abiotic stresses such as heat, salt, and drought. In *Arabidopsis*, the *hot2* mutant, which has a mutation in a CPL gene, showed tolerance to salt and drought conditions, possibly by preventing over-accumulating Na^+ ions[25].

CLPs also take part in plant organ development, particularly in cellulose metabolism and lignin deposition, thus, they have a significant effect on primary and secondary cell wall development[22]. This effect may also influence the architecture of seed coats, as both lignin and cellulose are involved in seed coat composition. There is evidence to show that lignin deficiency may correlate with seed coat fragility. Another study identified an EMS-derived peanut mutant with a significantly shorter germination time and showed seed-coat cracking, which represented weaker dormancy capacity. The mutant showed lower levels of lignin, anthocyanins, and proanthocyanidins compared to the wild-type[26]. Lignin was also found to affect the permeability of seed coat, hence affecting the longevity of seeds[27]. The transcription profile of the CLP family in our study

showed a significant difference between the wild-type and the mutants. The two wheat mutants in our study showed deficiencies in the expression of multiple CLP genes, which may lead to the abnormal deposition of lignin and cellulose (Fig. 6), and eventually affect seed dormancy by changing the permeability of the seed coat (Fig. 7).

CLP genes in plants have been shown to be regulated by transcription factors from families such as WRKY, NAC, and MYB. These transcription factors have similar patterns of regulation according to related studies; they bind directly to the W-box *cis*-regulatory elements and positively regulate CLP gene activity in the plant defense response against pathogens. Gao et. al isolated a MYB TF, BjMYB1, from Brassica juncea that specifically binds to the W-box-like-4 element in the promoter of the CLP gene *BjCHI1*, which is responsive to fungal infection[28]. In another study, an NAC transcription factor in rice, *OsNAC111*, was found to positively regulate several defense-related genes, including two CLP genes. Over-expression of *OsNAC111* resulted in enhanced resistance to the rice blast fungus *Magnaporthe oryzae*[29]. WRKYs were also found to regulate CLP gene expression. Three WRKY TFs were found to bind specifically to the two W-boxes present in the CLP gene *CHN48* in tobacco. Subsequent studies showed that *CHN48* is positively regulated by these WRKYs in response to elicitor treatment, suggesting a possible involvement of *NtWRKY*s in the defense response via the regulation of CLP gene transcription[30]. Recent findings have provided more evidence for the relationship between WRKY TFs and CLP gene regulation than for other TFs. Further studies on the regulatory roles of WRKY TFs in wheat may provide an explanation for the significant differences in expression profiles of the CLP family genes between the mutants and the wild-type. In addition, over-expression of a *WRKY* gene in rice, *OsWRKY71*, resulted in a significant up-regulation of the expression of multiple CLP genes and the defense response. In this study, the over-expression of *OsWRKY71* in rice cells resulted in the up-regulation of 200 genes, 146 of which could also be up-regulated by chitin oligosaccharide elicitor treatment. Sixty genes were down-regulated, 21 of which were also down-regulated by chitin oligosaccharide elicitor treatment. Seven rice CLP genes showed the most significant up-regulation of all the up-regulated genes. Together with the other findings mentioned above, these results indicate that there is a strong relationship between WRKY TFs and CLP genes, and that certain WRKYs may function in the generic regulation of multiple CLP genes[31]. In our study, all the *WRKY* genes in the target gene group showed reduced transcription in the mutants compared to the wild-type. These results suggest a possible explanation for the transcriptional profile of CLP genes; the deficiency of WRKYs and/or expression of a certain *WRKY* gene in the mutants has led to the down-regulation of the CLP genes.

In our study, GO analysis also identified a significant enrichment of genes involved in nodulation. This result is consistent with the discovery that CLPs participate in the development of nodules. A related study demonstrated that a CLP gene, *Srchi24*, is involved in the onset of nodulation in *Sesbania rostrata*[22].

ABA and the corresponding synthesis/signaling pathways have been shown to be key regulators of seed dormancy and PHS. However, in our study, we found no significant evidence to show that ABA contributes to the enhanced seed dormancy in the mutants. From the transcriptome data, however, we identified one *ABI5*-like gene that showed differential expression between the mutants and the wild-type. Further research will be needed to elucidate the possible role of this gene in seed dormancy.

Other transcriptomic studies of seed dormancy in cereals have identified a diversity of dormancy control mechanisms, with variable contributions from the activities of ABA. Consistent with our findings, a previous transcriptomic study of two rice cultivars with different levels of seed dormancy, found no obvious differences in the ABA synthesis and signaling pathways. The seed maturation pathways are significantly correlated with seed dormancy[32]. In another study, however, the strongly dormant rice cultivar N22 and two mutants derived from it, Q4359 and Q4646, which showed weaker dormancy, were used in a microarray study. Genes involved in ABA signaling and GA biosynthesis/signaling showed differential expression between Q4359 and the wild-type N22, indicating that ABA and GA are involved in the seed dormancy process[33]. These findings suggest that the mechanisms of seed dormancy control could be complicated, with ABA signals only partially controlling the regulation.

It should also be noted that although a possible pathway of dormancy regulation involving CLPs has been proposed, we cannot determine whether the overall differential expression of the CLP gene family causes the variation in seed dormancy or merely correlates with it based on our current findings. The group changes in the CLP family and *WRKY* genes indicate that

there could be a mutation in a key upstream regulatory factor in the two wheat mutants, which might be the initiator of the changes in seed dormancy. Follow-up studies will focus on finding the causes of the differential expression of the CLP genes, and possibly the key upstream regulator.

Conclusions

A PHS-resistant wheat strain, LQ18, was isolated from an EMS population. The mutant showed significantly stronger PHS tolerance and longer stem length compared to the wild-type. A transcriptome sequencing project was performed based on the wild-type and two independent M_2 lines derived from LQ18, LQ1815 and LQ1813. RNA-seq results showed no significant contributions from ABA biosynthesis and signaling pathways to the transcriptional differences between the mutants and the wild-type. The expression profile of nine genes and their alleles, which have been confirmed as functioning in dormancy release and PHS resistance, were studied, and the alleles of *TaDOG*, *TaSdr*, *PM19*, and *TaQsd1* showed differential expression between the mutants and the wild-type. A GO classification analysis of the intersection genes and the total genes at level 2 was performed to determine the GO enrichment in the two groups. The overall results were similar between the two groups, while an obvious difference was found in GO terms in the “molecular function” category. A KEGG classification analysis of the intersection genes and the wild-type at level 2 was conducted similar to the GO classification analysis. Some KOs showed differential enrichment between the two groups. We generated DAGs of the GO enrichment analyses of the intersection genes in the “biological process”, “cellular component”, and “molecular function” GO categories, and found a significant enrichment in chitin catabolism genes. Subsequent expression profiling showed consistently higher expression levels of chitinase family genes in the wild-type in the target gene group. These results add to our overall understanding of seed dormancy and PHS tolerance in cereal crops. Future studies will be conducted to further confirm the role of the chitinase genes in PHS, and to explore the reasons for the changes in expression found in the chitinase gene family members in the mutants.

Abbreviations

PHS: pre-harvest sprouting

ABA: abscisic acid

EMS: ethylmethane sulfonate

GA: gibberellic acid

QTL: quantitative trait locus

GI: germination index

DEG: deferentially expressed gene

GO: Gene Ontology

DAG: directed acyclic graph

KEGG: Kyoto Encyclopedia of Genes and Genomes

KO: KEGG ontology

CLP: chitinase like protein

Declarations

Ethics approval and consent to participate: not applicable.

Consent for publication: not applicable.

Availability of data and materials: the transcriptome datasets generated and analysed during the current study are available in the GSA repository, accession number: PRJCA006574. Other data generated or analysed during this study are included in this published article and its supplementary information files.

Competing interests: the authors declare that they have no competing interests.

Funding: this research was funded by the National Natural Science Foundation of China, grant number 31760382; the Inner Mongolia Natural Science Foundation, grant number 2019MS03016 and 2019MS03067.

Authors' contributions: XY performed the experiments and wrote the manuscript. HY and Z assisted in performing experiments. B , YP and Y designed the experiments and assisted in writing the manuscript. All of the authors have read and approved the final manuscript.

Acknowledgements: not applicable.

References

1. Miao, C.; Wang, Z.; Zhang, L.; Yao, J.; Hua, K.; Liu, X.; Shi, H.; Zhu, J.K. The grain yield modulator miR156 regulates seed dormancy through the gibberellin pathway in rice. *Nat Commun.* 2019;10:3822, doi:10.1038/s41467-019-11830-5.
2. Gubler, F.; Millar, A.A.; Jacobsen, J.V. Dormancy release, ABA and pre-harvest sprouting. *Curr Opin Plant Biol.* 2005;8:183-187, doi:10.1016/j.pbi.2005.01.011.
3. Benech-Arnold, R.L.; Rodríguez, M.V. Pre-harvest Sprouting and Grain Dormancy in Sorghum bicolor: What Have We Learned? *Front Plant Sci.* 2018;9:811, doi:10.3389/fpls.2018.00811.
4. Wang, Q.; Lin, Q.; Wu, T.; Duan, E.; Huang, Y.; Yang, C.; Mou, C.; Lan, J.; Zhou, C.; Xie, K., et al. *OsDOG1L-3* regulates seed dormancy through the abscisic acid pathway in rice. *Plant Sci.* 2020;298:110570, doi:10.1016/j.plantsci.2020.110570.
5. Xu, F.; Tang, J.; Gao, S.; Cheng, X.; Du, L.; Chu, C. Control of rice pre-harvest sprouting by glutaredoxin-mediated abscisic acid signaling. *Plant J.* 2019;100:1036-1051, doi:10.1111/tpj.14501.
6. Song, S.; Wang, G.; Wu, H.; Fan, X.; Liang, L.; Zhao, H.; Li, S.; Hu, Y.; Liu, H.; Ayaad, M., et al. OsMFT2 is involved in the regulation of ABA signaling-mediated seed germination through interacting with OsbZIP23/66/72 in rice. *Plant J.* 2020;103:532-546, doi:10.1111/tpj.14748.
7. Gupta, P.K.; Balyan, H.S.; Sharma, S.; Kumar, R. Genetics of yield, abiotic stress tolerance and biofortification in wheat (*Triticum aestivum* L.). *Theor Appl Genet.* 2020;133:1569-1602, doi:10.1007/s00122-020-03583-3.
8. Kulwal, P.L.; Kumar, N.; Gaur, A.; Khurana, P.; Khurana, J.P.; Tyagi, A.K.; Balyan, H.S.; Gupta, P.K. Mapping of a major QTL for pre-harvest sprouting tolerance on chromosome 3A in bread wheat. *Theor Appl Genet.* 2005;111:1052-1059, doi:10.1007/s00122-005-0021-4.
9. Mohan, A.; Kulwal, P.; Singh, R.; Kumar, V.; Mir, R.R.; Kumar, J.; Prasad, M.; Balyan, H.S.; Gupta, P.K. Genome-wide QTL analysis for pre-harvest sprouting tolerance in bread wheat. *Euphytica.* 2009;168:319-329, doi:10.1007/s10681-009-9935-2.
10. Kumar, J.; Mir, R.R.; Kumar, N.; Kumar, A.; Mohan, A.; Prabhu, K.V.; Balyan, H.S.; Gupta, P.K. Marker-assisted selection for pre-harvest sprouting tolerance and leaf rust resistance in bread wheat. *Plant Breeding.* 2010;129:617-621, doi:https://doi.org/10.1111/j.1439-0523.2009.01758.x.
11. AMBALAMAATIL, S.; LUKOW, O.M.; MALCOLMSON, L.J. QUALITY ATTRIBUTES OF CANADIAN HARD WHITE SPRING WHEAT. *Journal of Food Quality.* 2006;29:151-170, doi:https://doi.org/10.1111/j.1745-4557.2006.00064.x.
12. Nakamura, S.; Abe, F.; Kawahigashi, H.; Nakazono, K.; Tagiri, A.; Matsumoto, T.; Utsugi, S.; Ogawa, T.; Handa, H.; Ishida, H., et al. A wheat homolog of MOTHER OF FT AND TFL1 acts in the regulation of germination. *Plant Cell.* 2011;23:3215-3229, doi:10.1105/tpc.111.088492.

13. Torada, A.; Koike, M.; Ogawa, T.; Takenouchi, Y.; Tadamura, K.; Wu, J.; Matsumoto, T.; Kawaura, K.; Ogihara, Y. A Causal Gene for Seed Dormancy on Wheat Chromosome 4A Encodes a MAP Kinase Kinase. *Curr Biol.* 2016;26:782-787, doi:10.1016/j.cub.2016.01.063.
14. Zhou, K.; Yang, J.; Wang, Z.-X.; Wang, J.-R. Sequence analysis and expression profiles of *TaABI5*, a pre-harvest sprouting resistance gene in wheat. *Genes & Genomics.* 2017;39:161-171, doi:10.1007/s13258-016-0483-6.
15. Ashikawa, I.; Abe, F.; Nakamura, S. Ectopic expression of wheat and barley *DOG1*-like genes promotes seed dormancy in Arabidopsis. *Plant Sci.* 2010;179:536-542, doi:10.1016/j.plantsci.2010.08.002.
16. Abe, F.; Haque, E.; Hisano, H.; Tanaka, T.; Kamiya, Y.; Mikami, M.; Kawaura, K.; Endo, M.; Onishi, K.; Hayashi, T., et al. Genome-Edited Triple-Recessive Mutation Alters Seed Dormancy in Wheat. *Cell Rep.* 2019;28:1362-1369.e1364, doi:10.1016/j.celrep.2019.06.090.
17. Himi, E.; Noda, K. Red grain colour gene (*R*) of wheat is a Myb-type transcription factor. *Euphytica.* 2005; 143:239-242, doi:10.1007/s10681-005-7854-4.
18. Luo, W.; Liu, J.; Ding, P.; Li, C.; Liu, H.; Mu, Y.; Tang, H.; Jiang, Q.; Liu, Y.; Chen, G., et al. Transcriptome analysis of near-isogenic lines for glume hairiness of wheat. *Gene.* 2020;739:144517, doi:10.1016/j.gene.2020.144517.
19. Onishi, K.; Yamane, M.; Yamaji, N.; Tokui, M.; Kanamori, H.; Wu, J.; Komatsuda, T.; Sato, K. Sequence differences in the seed dormancy gene *Qsd1* among various wheat genomes. *BMC Genomics.* 2017;18:497, doi:10.1186/s12864-017-3880-6.
20. Utsugi, S.; Nakamura, S.; Noda, K.; Maekawa, M. Structural and functional properties of Viviparous1 genes in dormant wheat. *Genes Genet Syst.* 2008;83:153-166, doi:10.1266/ggs.83.153.
21. Zhang, Y.; Miao, X.; Xia, X.; He, Z. Cloning of seed dormancy genes (*TaSdr*) associated with tolerance to pre-harvest sprouting in common wheat and development of a functional marker. *Theor Appl Genet.* 2014;127:855-866, doi:10.1007/s00122-014-2262-6.
22. Kesari, P.; Patil, D.N.; Kumar, P.; Tomar, S.; Sharma, A.K.; Kumar, P. Structural and functional evolution of chitinase-like proteins from plants. *Proteomics.* 2015;15:1693-1705, doi:10.1002/pmic.201400421.
23. Vogelsang, R.; Barz, W. Purification, characterization and differential hormonal regulation of a beta-1,3-glucanase and two chitinases from chickpea (*Cicer arietinum* L.). *Planta.* 1993;189:60-69, doi:10.1007/bf00201344.
24. Ye, X.Y.; Wang, H.X.; Ng, T.B. Dolichin, a new chitinase-like antifungal protein isolated from field beans (*Dolichos lablab*). *Biochem Biophys Res Commun.* 2000;269:155-159, doi:10.1006/bbrc.2000.2115.
25. Kwon, Y.; Kim, S.-H.; Jung, M.-S.; Kim, M.-S.; Oh, J.-E.; Ju, H.-W.; Kim, K.-i.; Vierling, E.; Lee, H.; Hong, S.-W. Arabidopsis *hot2* encodes an endochitinase-like protein that is essential for tolerance to heat, salt and drought stresses. *The Plant Journal.* 2007;49:184-193, doi:10.1111/j.1365-313X.2006.02950.x.
26. Wan, L.; Li, B.; Pandey, M.K.; Wu, Y.; Lei, Y.; Yan, L.; Dai, X.; Jiang, H.; Zhang, J.; Wei, G., et al. Transcriptome Analysis of a New Peanut Seed Coat Mutant for the Physiological Regulatory Mechanism Involved in Seed Coat Cracking and Pigmentation. *Front Plant Sci.* 2016;7:1491, doi:10.3389/fpls.2016.01491.
27. Renard, J.; Martínez-Almonacid, I.; Sonntag, A.; Molina, I.; Moya-Cuevas, J.; Bissoli, G.; Muñoz-Bertomeu, J.; Faus, I.; Niñoles, R.; Shigeto, J., et al. *PRX2* and *PRX25*, peroxidases regulated by COG1, are involved in seed longevity in Arabidopsis. *Plant Cell Environ.* 2020;43:315-326, doi:10.1111/pce.13656.
28. Gao, Y.; Jia, S.; Wang, C.; Wang, F.; Wang, F.; Zhao, K. BjMYB1, a transcription factor implicated in plant defence through activating *BjCHI1* chitinase expression by binding to a W-box-like element. *J Exp Bot.* 2016;67:4647-4658, doi:10.1093/jxb/erw240.
29. Yokotani, N.; Tsuchida-Mayama, T.; Ichikawa, H.; Mitsuda, N.; Ohme-Takagi, M.; Kaku, H.; Minami, E.; Nishizawa, Y. OsNAC111, a blast disease-responsive transcription factor in rice, positively regulates the expression of defense-related genes. *Mol Plant Microbe Interact.* 2014;27:1027-1034, doi:10.1094/mpmi-03-14-0065-r.
30. Yamamoto, S.; Nakano, T.; Suzuki, K.; Shinshi, H. Elicitor-induced activation of transcription via W box-related cis-acting elements from a basic chitinase gene by WRKY transcription factors in tobacco. *Biochim Biophys Acta.* 2004;1679:279-

287, doi:10.1016/j.bbaexp.2004.07.005.

31. Chujo, T.; Kato, T.; Yamada, K.; Takai, R.; Akimoto-Tomiyama, C.; Minami, E.; Nagamura, Y.; Shibuya, N.; Yasuda, M.; Nakashita, H., et al. Characterization of an elicitor-induced rice WRKY gene, *OsWRKY71*. *Biosci Biotechnol Biochem.* 2008;72:240-245, doi:10.1271/bbb.70553.
32. Huh, S.M.; Hwang, Y.S.; Shin, Y.S.; Nam, M.H.; Kim, D.Y.; Yoon, I.S. Comparative transcriptome profiling of developing caryopses from two rice cultivars with differential dormancy. *J Plant Physiol.* 2013;170:1090-1100, doi:10.1016/j.jplph.2013.03.003.
33. Wu, T.; Yang, C.; Ding, B.; Feng, Z.; Wang, Q.; He, J.; Tong, J.; Xiao, L.; Jiang, L.; Wan, J. Microarray-based gene expression analysis of strong seed dormancy in rice cv. N22 and less dormant mutant derivatives. *Plant Physiol Biochem.* 2016;99.27-38, doi:10.1016/j.plaphy.2015.12.001.
34. Yang, Y., Zhang, C. L., He, Z. H., Xia, L. Q. Advances on Resistance to Pre -harvest Sprouting in Wheat. *Journal of Plant Genetic Resources.* 2007;8:503-509, doi:10.3969/j.issn.1672-1810.2007.04.025.
35. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 2014;30:2114-2120, doi:10.1093/bioinformatics/btu170.
36. Kim, D.; Langmead, B.; Salzberg, S.L. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods.* 2015;12:357-360, doi:10.1038/nmeth.3317.
37. Anders, S.; Pyl, P.T.; Huber, W. HTSeq–a Python framework to work with high-throughput sequencing data. *Bioinformatics.* 2015;31:166-169, doi:10.1093/bioinformatics/btu638.
38. Roberts, A.; Pimentel, H.; Trapnell, C.; Pachter, L. Identification of novel transcripts in annotated genomes using RNA-Seq. *Bioinformatics.* 2011;27:2325-2329, doi:10.1093/bioinformatics/btr355.
39. Anders, S.; Huber, W. Differential expression analysis for sequence count data. *Genome Biol* 2010, *11*, R106, doi:10.1186/gb-2010-11-10-r106.

Figures

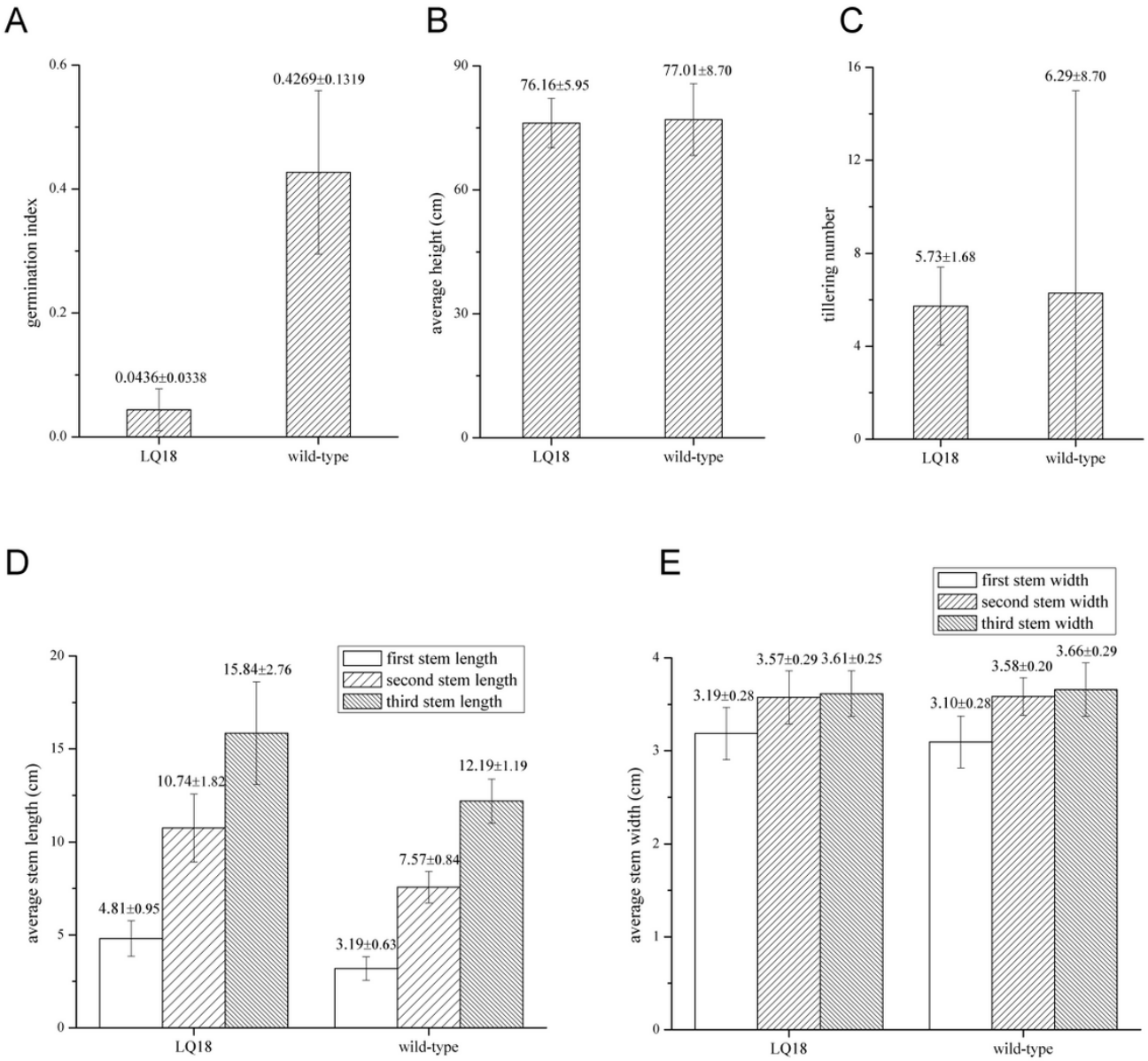

Figure 1

Agronomic traits of the LQ18 mutants and the wild-type 'Long 13-3778'. (a) germination index. (b) average height. (c) tillering number. (d) the first, second and third stem length. (e) average width of the first, second and third stem width.

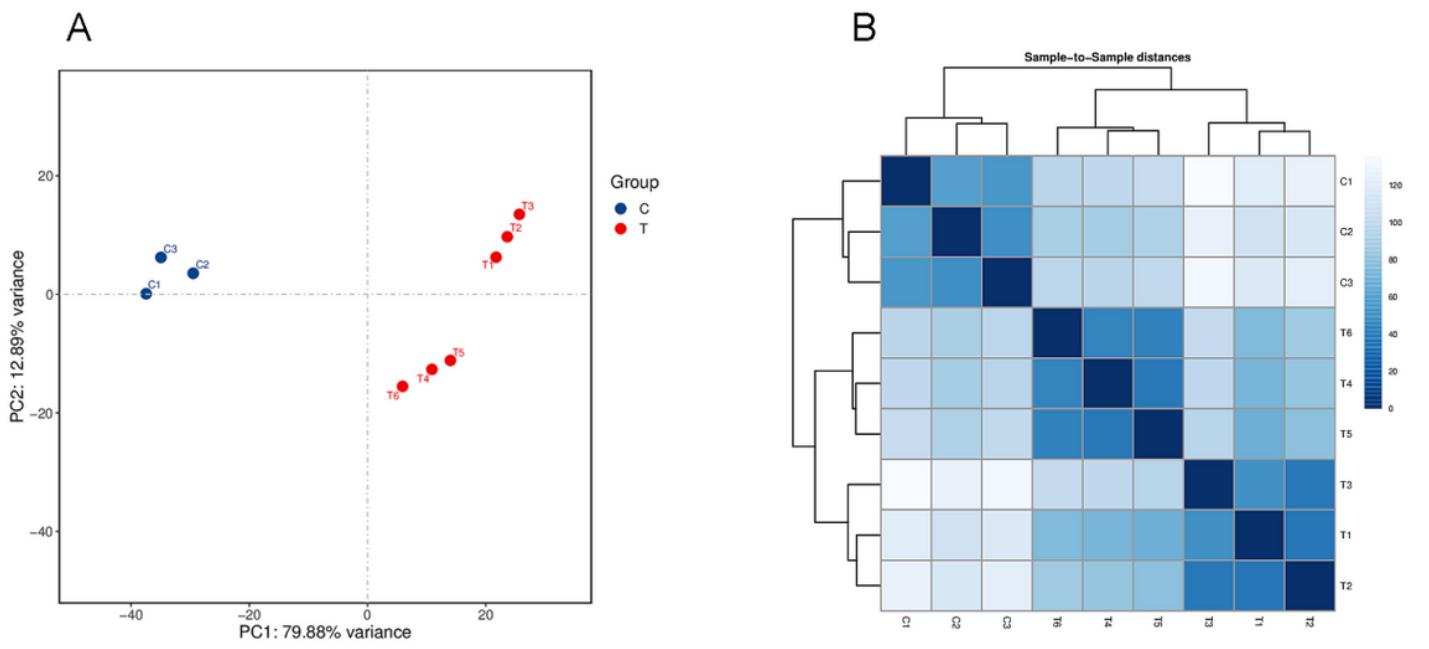

Figure 2

validity analysis of the transcriptome data. (a) PCA analysis. (b) sample-to-sample cluster. Samples C1-C3: wild-type samples. Samples T1-T3: LQ1815 samples. Samples T4-T6: LQ1813 samples.

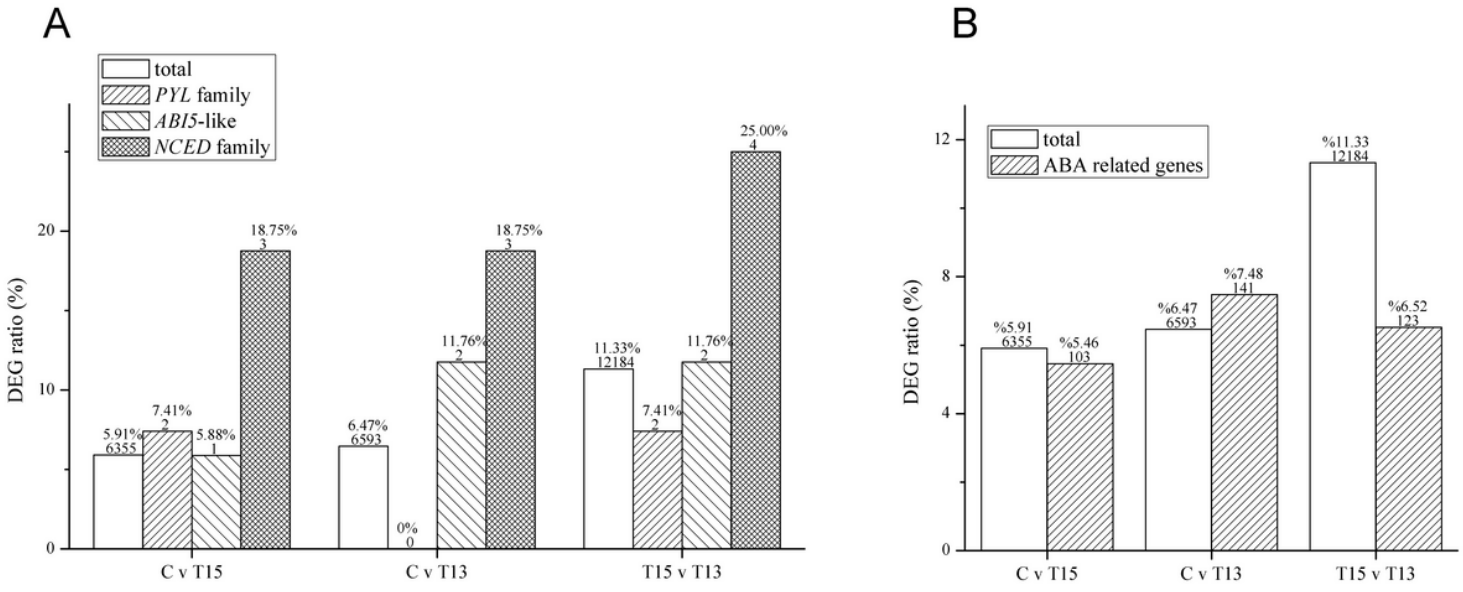

Figure 3

Comparison of DEG ratio in ABA pathway genes and the total genes. (a) Comparison between the ABA synthesis and signaling genes and the total genes. (b) Comparison between genes with the GO term “abscisic acid” and the total genes.

Figure 4

GO and KEGG classification analysis. (a) GO classification. (b) KEGG classification.

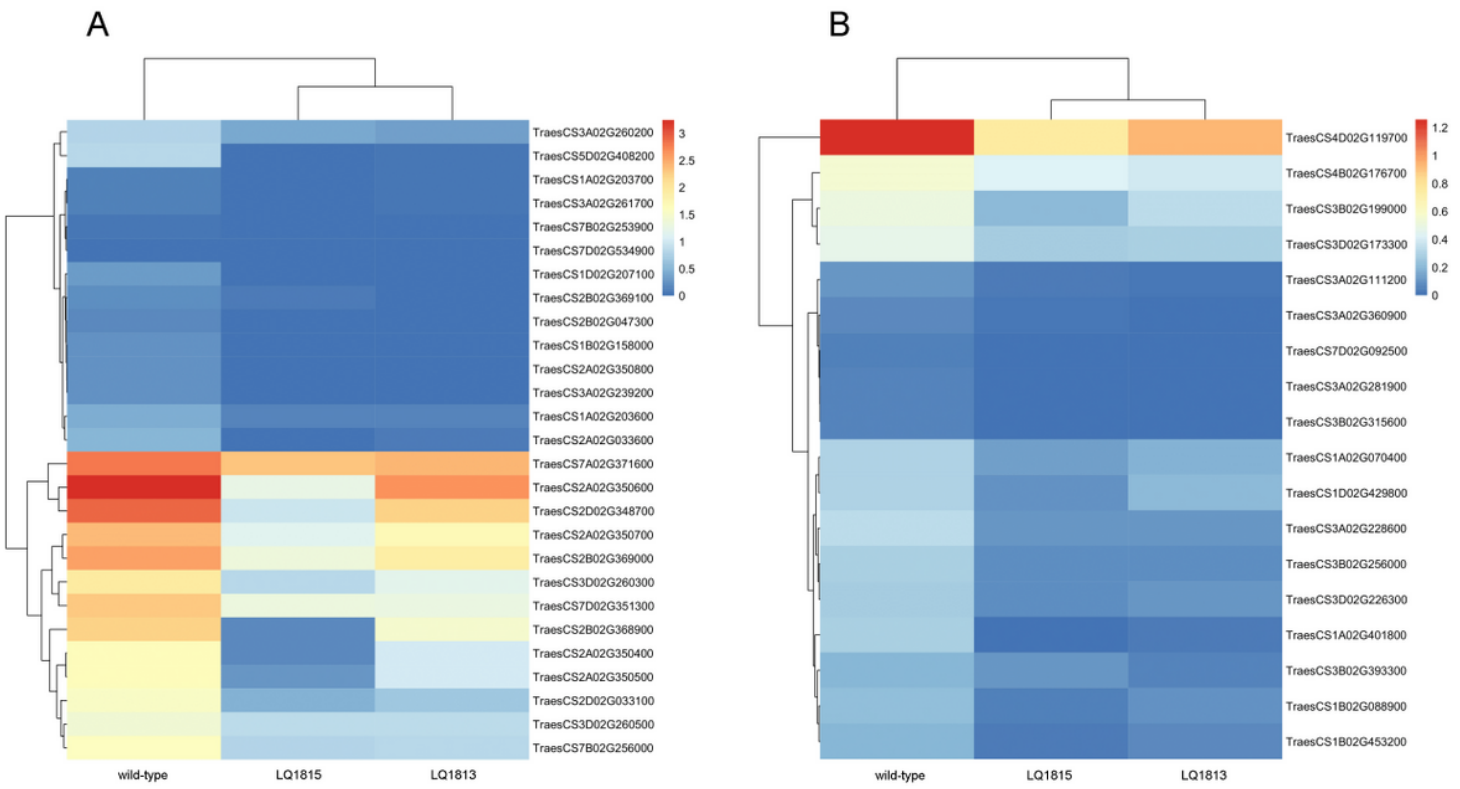

Figure 5

Expression profile of the chitinase and WRKY gene family of the RNA-sequencing. (a) Expression profile of the chitinase gene family. (b) Expression profile of the WRKY gene family. The values of log10 (FPKM+1) were used as the indicator of expression level.

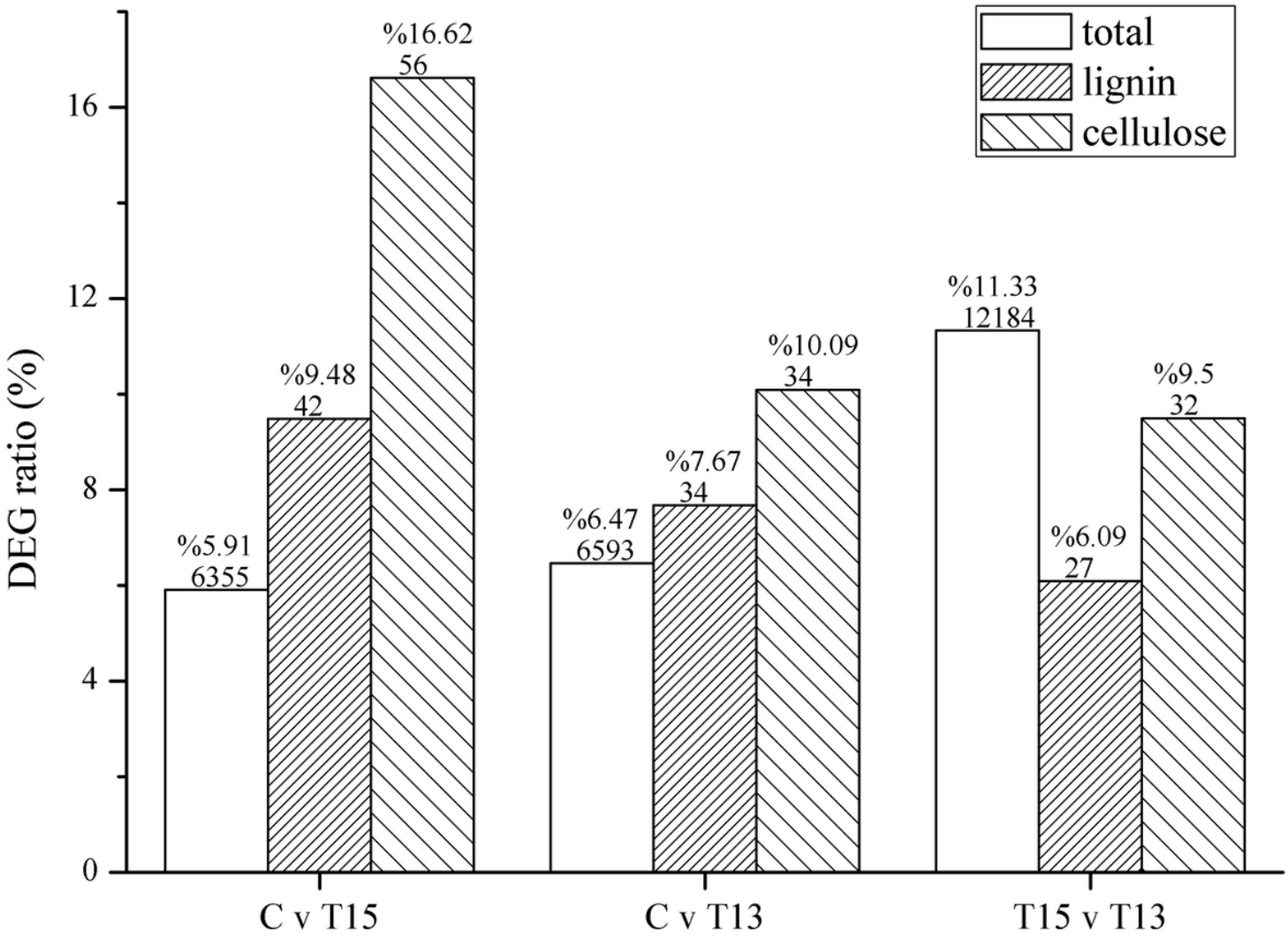

Figure 6

Comparison of DEG ratio in lignin associated genes, cellulose associated genes and the total genes. Lignin and cellulose associated genes were designated as genes with the GO term “lignin” and “cellulose”, respectively.

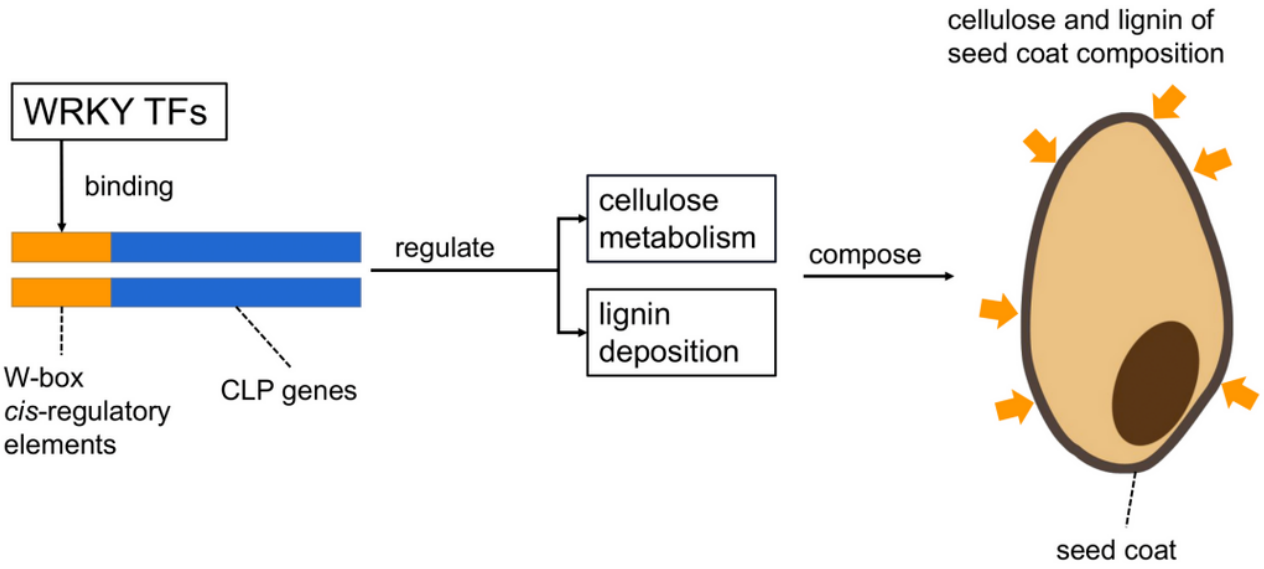

Figure 7

Possible regulation pathway of lignin and cellulose in seed coat composition involving WRKY transcription factors and CLP genes.

Supplementary Files

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

- Additionalfiles.docx