Transcriptome analysis and molecular mechanism of linseed (*Linum usitatissimum* L.) drought tolerance under repeated drought using single-molecule long-read sequencing

Wei Wang
Oil Crops Research Institute Chinese Academy of Agricultural Sciences https://orcid.org/0000-0003-1271-6674
Lei Wang
Oil Crops Research Institute Chinese Academy of Agricultural Sciences
Ling Wang
Oil Crops Research Institute Chinese Academy of Agricultural Sciences
Meilian Tan
Oil Crops Research Institute Chinese Academy of Agricultural Sciences
Collins O. Ogutu
Wuhan Botanical Garden
Ziyan Yin
Oil Crops Research Institute Chinese Academy of Agricultural Sciences
Jian Zhou
Wuhan Igenebook Biotechnology Co., Ltd
Jiaomei Wang
Oil Crops Research Institute Chinese Academy of Agricultural Sciences
Lijun Wang
Oil Crops Research Institute Chinese Academy of Agricultural Sciences
Xingchu Yan (yanxc@oilcrops.cn)

Research article

Keywords: transcriptome, Linseed, Repeated drought, SMRT, transcription factors

Posted Date: January 9th, 2020

DOI: https://doi.org/10.21203/rs.2.20411/v1

Version of Record: A version of this preprint was published at BMC Genomics on February 9th, 2021. See the published version at https://doi.org/10.1186/s12864-021-07416-5.

Abstract

Background

Oil flax (*Linum usitatissimum* L.) also as known as linseed is one of the most important oil crops in the world. Although linseed was reported to show better tolerance to abiotic stress conditions compared to other oil crops, the molecular mechanisms underlying linseed tolerance to drought stress are largely unknown. Moreover, as a result of climate change, drought dramatically reduces linseed yield and quality, but so far very little is known about how linseed coordinates the drought-resistant genes expression of response to different level of drought stress on the genome-wide level.

Results

To explore the transcriptional response of linseed to drought stress (DS) and repeated drought stress (RD), we first determined the drought tolerance of different linseed varieties. Then we performed full-length transcriptome sequencing of drought-resistant variety (Z141) and drought-sensitive variety (NY-17) using single-molecule real-time sequencing and RNA-sequencing under drought stress (DS) and repeated drought stress (RD) at the seedling stage. Gene Ontology (GO) enrichment analysis showed that compared with NY-17, the up-regulated genes of Z141 were enriched in more functional pathways related to plant drought tolerance under drought stress. In addition, the number of up-regulated genes in linseed under RD was more 30% than it under DS. In addition, a total of, 4,436 linseed transcription factors were identified, of these, 1,190 genes were responsive to stress treatments. Finally, the expression patterns of proline biosynthesis and DNA repair structural genes were verified by RT- PCR.

Conclusions

Drought tolerance of Z141 may be related to its specifically up-regulated drought tolerance genes under drought stress. Several variable physiological responses occurred in repeated than in sustained drought treatment. Sum up, this study provides a new perspective to understand the drought adaptability of linseed.

Background

Flax (Linum usitatissimum L.) is an important commercial crop in the family Linaceae with a history of cultivation dating back 7000 years. Large scale production is done in Russia, China, Canada, Germany, Argentina and Kenya (Database of FAO, http://www.fao.org/faostat). Recently, linseed has been associated with health-promoting and nutraceutical properties further highlighting its importance and increased demand. The Flax seed is mainly used for food, feed and industrial oil purposes [1, 2]. Based on morphology, its uses can be divided into fiber flax and oil flax (linseed) [3]. Linseed oil contains a variety of fatty acids, such as oleic acid, linoleic acid and α-linolenic acid (ALA). Linoleic acid concentration is nearly 75% of total fatty acids whereas ALA concentration is about 59% of total fatty acids and it can be metabolized into eicosapentaenoic acid (EPA) and docosahexaenoic acid (DHA) [4, 5].

The relatively high levels of α-linolenic acid (ALA), which is approximately 5 times higher than sources make linseed the only terrestrial food plant with the highest levels of omega-3 fatty acids [6].

Drought is an environmental stress that limit the growth and productivity of plants [7]. It can significantly influence plant survival and functioning by causing physiological and metabolic changes [8]. In addition, drought directly and indirectly affects photosynthesis and consequently dry matter production. In addition, In addition, it reduces leaf expansion, promotes leaf senescence and abscission. And leads to yield loss of more than 50% [9–12]. For example, drought has adversely affected more than 50% of global wheat cultivation area thus causing considerable yield loss by inhibiting photosynthesis [13, 14]. Similarly, linseed yields has seriously been by various abiotic stresses such as drought and global warming [15]. In China, linseed is widely grown in the western and northwestern provinces such as Gansu and Inner Mongolia which experience the highest drought frequency and the maximum duration in East Asia [16]. In addition, several linseed growing countries such as Indian, Kenya, and Chile experiences drought conditions [17–19].

To counter drought effects, the plants have evolved special mechanisms and undergone serious of physiological changes. For example genes involved in ABA, proline, glycine-betaine and sorbitol pathways were found up-regulated by drought stress in wheat [20]. Similarly, tolerant maize varieties exhibited more drastic changes in global gene expression which correlated with different physiological mechanisms of adaptation to drought [21]. In addition, transgenic maize with enhanced ZmVPP1 expression demonstrated improved drought tolerance which was attributed to enhanced photosynthetic efficiency and root development [22].

Unlike other oil crops, linseed has previously been associated with drought resistance. Moreover, different linseed varieties demonstrate varied degree of tolerance. Whole-genome sequence of linseed has been released [23], in addition, transcriptome analysis of the drought tolerant flax cultivar at different developmental stages has been reported [24]. However, a comprehensive relationship between linseed drought tolerance and gene expression or transcription factor (TF) regulation still remains unclear.

PacBio's SMRT (single molecule real time) sequencing (PacBio, http://www.pacificbiosciences.com/) provides a third-generation sequencing platform and its widely for long reads genome sequencing [25]. Due to its ability to obtain full-length transcripts without assembly, PacBio can provide direct comprehensive analysis for splice isoforms of each gene as well as improve annotation of existing gene models. SMRT technology is an ideal method for conducting plant genomes research, due to their highly repetitive nature as compared to vertebrate genomes [26–28]. Recently, Li et al., (2017) used Iso-Seq to analyze full-length (FL) splice isoforms in strawberry thus suggesting its suitability in uncovering mechanism of drought tolerance in linseed [29]. In this study, we first identified variation in drought tolerance of linseed varieties NY-17and Z141, then combined SMRT sequencing and short-read next generation sequencing technology to generate a more complete FL linseed transcriptome. In addition, comprehensive candidate gene identification was conducted for; drought (DS), re-watering (RW), and

repeated drought (RD) and analysis of partitioned expression patterns for homologous genes in linseed under different drought tolerance.

Results

Determination of drought tolerance in linseed varieties

In this study, we measured three drought-tolerant related phenotypic traits of Z141 and NY-17 (Supplementary Table S1). The result showed that Z141 consistently performed better than NY-17 under drought stress (Fig. 2a-d). In addition, we observed that under drought stress Z141 had lower plant height and biomass reduction rate compared with NY-17 under drought stress (Fig. 2e, f; Supplementary Table S2). The biomass reduction rate under drought stress was 30% and 46% in Z141 and NY-17 respectively. The LAWC and LRWC which are key components for drought-tolerance were significantly higher in Z141 than NY-17 under drought stress, thus suggesting Z141 can hold more water when it under drought stress. (Fig. 2g, h; Supplementary Table S3)

Table 1
Two-way ANOVAs to test the effects of different drought stress (Fixed effect), two linseed biotypes (random effects), and their interaction on plant height, biomass, leaf absolute water content (LAWC) and leaf relative water content (LRWC).

Trait	Drought			Linseed			D x L		
	df	F	p	df	F	p	df	F	p
Plant height	1	709.13	0.000	1	479.90	0.000	1	26.36	0.001
Biomass	1	108.03	0.000	1	302.29	0.000	1	41.29	0.000
LAWC	1	314.44	0.000	1	1.27	0.292	1	36.88	0.000
LRWC	1	949.63	0.000	1	5.22	0.05	1	6.48	0.03
The drought (D) had two levels (drought stress or non-drought stress) and linseed biotype (L) had two levels too.									

Two-way ANOVA result showed significant effects of the different varieties, different drought level treatments and their effects on plant height, biomass LAWC and LRWC (Table 1). NY-17 did not reveal any significant evidence of drought tolerance.

Analysis of linseed transcriptome by PacBio Iso-seq

The total RNA of Z141 and NY-17 was isolated from the control, DS, RW and RD treatments groups and quality checked. A total of 16 RNA samples were sent to Wuhan Frasergen Bioinformatics Co.,Ltd.

Genomic Service for sequencing using PacBio Sequel platform. This platform can generate sufficiently longer read lengths which covers most RNA transcripts in full length thus ensuring accurate reconstructed full-length (FL) splice variants are obtained. Over 2 million polymerase reads with mean length of ~30,000bp were generated after quality checking by Frasergen (Supplementary Table S5). After processing raw data we obtained more than 33 million filtered subreads with mean length ~2000bp (Supplementary Table S6). In addition we obtained 1,599,415 CCS reads which included 1,293,134 FL reads (Supplementary Table S7). De novo transcriptome reconstruction of the data was done using RNA-Seq reads and publically available flax sequences. To evaluate the density and length of isoforms, we compared the loci coverage of PacBio FLNC and swine SSC 10.2 annotation. In PacBio data set, a total of 1,093,282 high-quality FLNCs covered 108,579 isoforms and were allocated to 28,686 loci (Supplementary Table S9). Due to high base error of SMRT sequencing technology, a high-quality Illumina short reads was done using Provverad software to correct the error (Supplementary Table S10). In this study the before and after correction FLNC sequences were aligned to linseed genome sequence respectively through GMAP, finally we obtained 1,093,282 high-quality FLNC for further study (Supplementary Table S8).

Global comparisons of DS and RD related transcriptomes reveal their gene expression and functional groups difference

The mRNA populations were compared using principal component analysis (PCA) to provide a framework for understanding how linseed genes are regulated to respond to the drought stress. Transcriptomes of Z141 and NY-17 under DS, RW or RD were likely to share a great similarity in gene expression respectively, which variations forming three groups that were far deviated from the control (Figure 3a). The transcriptomes of DS exhibited distinct relationship from that of RD, suggesting the gene expression of transcriptome has a major shift between DS and RD.

Cluster analysis of differentially expressed genes further supported our observed results in PCA (Figure 3b). The ratio of up- or down-regulated genes overlap between Z141-RD and NY-17-RD was significantly higher than that of Z141-DS and Z141-RD, with 62.1% compared to 47.8 and 70.7% compared to 60.6% respectively (Figure 3c, d). In addition, in Z141 and NY-17 approximately 52.2% and 65.6% of up-regulated genes were uniquely responsive to RD respectively, 29.9% and 43.6% of up-regulated genes were uniquely responsive to DS. Specifically, in Z141 and NY-17 has 8005 (including 3245 for DS and 4760 for RD) and 6285 (including 2381 for DS and 3904 for RD) genes up-regulated under drought stress respectively. About 9104 (including 4041 for DS and 5063 for RD) and 7908 (3515 for DS and 4393 for RD) genes were down-regulated under drought stress in Z141 and NY-17. We also observed a higher proportion of stress responsive genes under RD compared to that under DS. In this study, 2275, 1343, 3067 and 2154 genes were up- and down-regulated when Z141 or NY-17 under drought stress respectively. Thereinto, 1007 and 1686 genes were significantly up- and down-regulated when Z141 and NY-17 under DS and RD

(Figure 3c, d). Taken together, this result suggest that the transcriptomes of DS and RD has fundamentally different.

Gene Ontology (GO) enrichment analysis was conducted to examine the functional distribution of the drought stress-related candidate genes identified in our study. A serial of GO categories exhibited significantly higher enrichments in the overlapped or unique up-regulated gene sets under DS, and RD treatments compared to the background. The GO terms of overlapped up-regulated genes between DS and RD in Z141 and NY-17 were mainly enriched in “proline biosynthetic process (GO: 0006561)” and “proline metabolic process (GO: 0006560)” (Figure 4a, b). Moreover, except for the amino acids biosynthesis and metabolism, more abiotic stress related GO terms e.g. “response to stress (GO: 0009650)” and “response to desiccation (GO: 0009269)” exhibited significant enrichment in Z141 up-regulated genes (Fig. 4A). Interestingly, GO terms related to flower development (GO: 0009908) were only significantly enriched in Z141 up-regulated genes which indicates that the drought avoidance mechanism of Z141 has been activated (Figure 4a). Drought stress inhibits plant photosynthesis, under DS and RD, in this study we observed the GO terms of photosynthesis (GO: 0015979) significantly enriched in down-regulated genes in Z141 and NY-17 (Supplementary Figure S1 a, b). Although the molecular mechanism between proline and plant drought tolerance is still unproven, our results suggest that proline may play an important role in drought tolerance of linseed.

The difference of linseed gene regulation pattern under DS and RD, suggest that under repeated drought stress linseed may have different molecular mechanisms for drought tolerance. Of the stress responsive GO terms, two distinct functional categories of DS specifically up-regulated genes inZ141 exhibited significantly higher enrichments, namely methylation and negative regulation. The first group included “histone H3-K36 demethylation (GO: 0070544)” and “macromolecule methylation (GO: 0043414)”, whereas the second group contained “negative regulation of biological process (GO: 0048519)” and “negative regulation of macromolecule metabolic process (GO: 0010605)” (Supplementary Figure S1c). The GO terms of up-regulated genes in Z141 under RD mainly enriched in “fatty acid oxidation (GO: 0019395)”, “fatty acid biosynthetic process (GO: 0006633)”, “fatty acid m metabolic process (GO: 0006631)” and “lipid metabolic process (GO: 0006629)” (Supplementary Figure S1d). The GO terms of unique down-regulated genes in Z141 under DS were mainly enriched in “carbohydrate metabolic process (GO: 0005975)”, “lignin biosynthetic process (GO:0009809)” and “lignin metabolic process (GO:0009808)”, whereas under RD the GO terms of unique down-regulated genes in Z141 were mainly enriched in “amide biosynthetic process (GO:0043604)” and “cellular amide metabolic process (GO:0043603)” (Supplementary Figure S1e, f). Overall, these functional categories indicated that epigenetic modifications might play a crucial role in the DS responsive process, although the exact functions of these genes remain to be elucidated. Meanwhile, drought stress may induced Z141 growth state from vegetative growth to reproductive growth.

The GO terms of unique up-regulated genes in NY-17 under DS were mainly enriched in RNA regulation including “RNA modification (GO: 0009451)”, “RNA processing (GO: 0006396)” and “ncRNA processing (GO: 0034470)”. Under RD, the GO terms of unique up-regulated genes were mainly enriched in

“transmembrane transport (GO: 0055085)” (Supplementary Figure S1g, h). The GO terms of down-regulated genes in NY-17 under DS were mainly enriched in flavonoid biosynthetic (GO: 0009813). Interestingly, The GO terms of down-regulated genes in NY-17 under RD were similarity with that in Z141 mainly enriched in “amide biosynthetic process (GO: 0043604)” and “cellular amide metabolic process (GO: 0043603)” (Supplementary Figure S1i, j).

Comparison of Z141 and NY-17 related transcriptomes reveals molecular mechanism of linseed drought tolerance

Although the transcriptomes of Z141 and NY-17 are very similar in overall gene expression, a set of stress responsive genes exhibited altered expression patterns specific to Z141 or NY-17 under drought stress, indicating distinguished functional categories could be impact on the drought tolerance of linseed. The GO terms of overlapped up-regulated genes between Z141 and NY-17 under DS were mainly enriched in two distinct functional categories, including proline biosynthesis and reproductive development. The proline biosynthesis category included “proline biosynthetic process (GO: 0006561)”, “proline metabolic process (GO: 0006560)”, “glutamine family amino acid biosynthetic process (GO: 0009084)” and “glutamine family amino acid metabolic process (GO:0009064)”, whereas the abiotic stress response category contained “reproductive system development (GO: 0061458) and “reproductive structure development (GO:0048608)” (Figure 5a). Under RD treatment, the GO terms of overlapped up-regulated genes between Z141 and NY-17 were also mainly enriched in proline biosynthesis category with “proline biosynthetic process (GO: 0006561)” and “proline metabolic process (GO: 0006560)” , and in abiotic stress response category with “response to abscisic acid (GO: 0009737)”, “response to desiccation (GO: 00009269)”, “response to acid chemical (GO: 0001101)” (Figure 5b). Without suspense, The GO terms of overlapped down-regulated genes between Z141 and NY-17 under DS and RD were mainly enriched in functional categories related to photosynthesis (Supplementary Figure S2a, b)

The GO terms for unique up-regulated genes in Z141 under DS were mainly enriched in “abscission (GO: 0009838)”, “defense response (GO: 0006952)” and “NADP biosynthetic process (GO: 0006741)” (Supplementary Figure S2c), whereas under RD, the GO terms were mainly enriched in “jasmonic acid biosynthetic process (GO: 0009695)” and “jasmonic acid metabolic process (GO: 0009694)” (Supplementary Figure S2d). The unique up-regulated genes showed more enrichment on pathways which closely related to plant drought resistance such as jamonic acid biosynthesis, abscission and NADP biosynthetic [30, 31]. In contrast, in NY-17 under DS the GO terms for unique up-regulated genes were mainly enriched in RNA regulation functional category with “ncRNA metabolic process (GO: 0034660)”, “ncRNA processing (GO: 0034470)”, “tRNA processing (GO: 0008033)” and so on (Supplementary Figure S2g). Under RD, the GO terms for unique up-regulated genes in NY-17 were mainly enriched in “phenylpropanoid biosynthetic process (GO: 0009699)” and “phenylpropanoid metabolic process (GO: 0009698)” (Supplementary Figure S2h).

Identification of temporarily up- and down-regulated transcription factors (TFs) in response to DS and RD

TFs have been proven to play irreplaceable roles in response to various abiotic stresses by modulating target gene expression [32]. To understand the essence of regulatory processes during DS and RD treatment, a domain searching method [33] was used to first predict transcription factors in Z141 and NY-17 on a whole-genome scale based on our identified non-redundant linseed unigenes. A total of 4,936 linseed TF genes distributed among 50 families were identified (Supplementary Table S11).

To profile stress-responsive TFome under DS and RD, we focused on TF genes exhibiting diverse expression patterns with stress changes, including continuous up-regulation, continuous down-regulation an early peak of expression and the late peak of expression patterns. AS a result, 1190 TFs distributed in 50 families were found to be differentially regulated in response to at least one stress. (Fold change ≥ 2 and FDR adjusted $p < 0.01$). Among which, eleven TF families accounted for approximately half of stress-responsive TF genes, including bHLH (9%), C2H2 (8%), NAC (8%), MYB (6%), ERF (6%), bZIP (5%), WRKY (5%) and MYB-related (4%) (Figure 6a)

Moreover, the 1190 TFs were further classified into 15 clusters according to their expression patterns by performing Mfuzz program analysis in R software [34]. Cluster5, 8,11 and 13 consists of 387 TFs mainly up-regulated by DS and RD, including *DREB*, *HSF* and *NF-YA10* which have been confirmed to be key regulator of plant abiotic resistance pathways [35-37] (Figure 6b and Supplementary Table S11).

Validation of isoforms by RT-PCR

Expression analysis of differentially expressed functional candidate genes, associated with DNA repair, MAPK signaling pathway, proline biosynthetic, and photosynthesis, selected from transcriptome data, and validated by RT-PCR. The results (Figure 7a-d) demonstrated that transcript abundances of selected genes were consistent with transcriptome analysis thereby validating the reliability of our annotated transcriptome data for future study.

Discussion

In this study, we regulated the absolute water content of soil by measuring the weight of soil. This method enabled us to investigate the phenotypic and gene expression change of plants under various treatment levels of drought stress, and these stresses are reproducible because the SWC is not dependent on the type of soil. Two linseed varieties of NY-17 and Z141 with varying drought tolerance phenotypes were selected in this study to detected differences in gene expression pattern under simulated severe drought (DS) and repeated severe drought (RD) stress. Our findings reveals the responses and acclimation of linseed plants with varying drought-tolerance to different levels of drought stress.

Unlike the animal, plants are sessile, therefore, plants have evolved to developed specific mechanisms in respond to drought stresses. In recent years, several studies have confirmed that plant drought-tolerance

in plants is related to their phenotype variety on which directly affect their drought tolerance [38, 39]. LWC and biomass are important phenotypic traits for identifying plant drought tolerance [10]. A large number of studies have shown that under drought stress high LWC values will reduce the yield loss [40, 41]. A previous study analyzed the patterns of plant biomass allocation in relation to drought stress and found that drought significantly increased the fraction of root mass but decreased that of stem, leaf, and reproductive mass [42]. Under drought stress, the plant biomass reduction rate is inversely proportional to its drought tolerance [43]. In addition, some studies reported that in the rice and maize, drought-related SNPs are usually associated with plant height [44, 45]. In the study of woody plants, the maximum plant size should be short under drought environment [46]. Thus, in this study, therefore, we measured the degrees of leaf wilting, leaf water content, plant height and biomass dry weight of Z141 and NY-17 under different drought levels. Compared with control, when SWC was lower than 10%, the biomass reduction rate in NY-17 was significantly more than that in Z141, but the relative leaf water content in Z141 was significantly higher than that in NY-17. Even in severe drought stress, the NY-17 variety had higher plant height than Z141, for example, the plant height of NY-17 was more than 32% higher than that of Z141 when SWC was 70% (Table 1 and Supplementary Table S2). Interestingly, almost all measured traits indicated that the level of drought tolerance in NY-17 was lower than that in Z141, despite previous wide cultivation of the former in Semi-arid areas of China such as Gansu, Inner Mongolia and other northwestern provinces where the average annual precipitation lower than 400 mm.

Under the drought stress the differences are between Z141 and NY-17 were not only in phenotype, but also in gene expression patterns. The number of up- or down-regulated genes in Z141 was significantly higher than in NY-17 under stress. For example, of the 3,245 identified up-regulated genes in Z141, 1693 genes accounting for 52.2% were specifically up-regulated under DS (Figure. 3a). In contrast, 2,381 genes were up-regulated in NY-17, of which only 829 genes, accounting for 34.8% were specifically up-regulated (Fig. 3c). In addition, these specifically regulated genes were functionally different. Under DS, the function of specifically up-regulated genes in Z141 were mainly associated with NADP biosynthetic, abscission, defense response and MAPK signal pathway (Supplementary Figure S2c), whist the function of specifically up-regulated genes in NY-17 were mainly enriched in RNA regulation (Supplementary Figure S2f). Previous studies have shown that expression of NADP biosynthesis genes are up-regulated when plants are under drought stress [34, 47], despite some studies suggesting that NADP genes could be related to ABA-mediated signaling [48], the mechanism still remains unclear and more studies are warranted Our study revealed an up-regulation of NADP biosynthesis genes in Z141 leaves under drought stress which suggests that NADP may compensate for the deficiency of CO_2 in the light-independent reactions caused by drought stress. Thus the specifically up- or down-regulated genes in Z141 maybe explain why this linseed variety has better drought tolerance than NY-17.

Most of the plant drought tolerance studies have been conducted by considering stress as a single event that happen once in the life of a plant, however little is known when recurrent drought episodes occur. Study in two shortgrass species found that drought timing and lack of previous drought exposure determines their sensitivity to water stress [49]. In contrast, some studies have found that plants exposed

to multiple drought cycles can develop a differential acclimation that potentiates their defense mechanisms, allowing them to be kept in an 'alert state' to successfully cope with further drought events [50, 51]. Compared with DS, under RD treatment there were more unique up- and down-regulated genes both in Z141 and NY-17 and those unique up- and down-regulated genes were enriched in more GO categories (Fig. 3c, d; Supplementary Figure S1d, f, h and j). This result suggest that several variable physiological responses occurred in repeated than in sustained drought treatment [52]. Moreover, in this study, we observed that the functional categories of unique down-regulated genes were significantly difference between Z141 and NY-17 under DS, but they were very similar under RD (Supplementary Figure S1e, f, i and j). The difference in linseed response to repeated drought stress suggests that linseed might develops drought stress "memory" thus changing its gene expression pattern to help adapt fast to future drought events [53].

Approximately 7% of the coding sequences were associated with TFs which play a central role in regulating gene responses to abiotic stresses in plant [54–56]. In this study, we predicted 4936 potential TFs in linseed genome, accounting for approximately 9% of total genes, and the number is nearly twice the 2481 TFs registered in plantTFDB (Supplementary Table S9). Numerous studies have shown the DREB have been demonstrated to be master regulators of gene networks in plant acclimation response to drought by regulating responsive gene expressions via binding to the cis-acting elements [57]. Similarly, one-third of DREB gene family members were significantly up-regulated under drought stress. Many other TFs, such as HSF and NF-YA10, were also specifically up-regulated under drought stress. Previous studies showed that HSF and NF-YA10 can increase plant high temperature and salt tolerance respectively [36, 37]. However, in this study we ensured that the temperature (~ 22 ℃) was suitable for linseed growth during drought treatment. This may suggests that the molecular mechanism of abiotic stress tolerance in plants such as drought, high temperature, saline-alkali is not independent, and there may be some interactions [58]. Moreover, some TFs were related to plant drought avoidance were also specifically up-regulated under drought stress in this study. NF-YC3 and WRKY75 has been proven to induce flowering or regulation root development when plants under abiotic stress [59–61]. Unexpectedly, some validated negative stress-regulators were also up-regulated under drought stress which make it more complicated to understand the molecular mechanisms underlying linseed tolerance to abiotic stress. For example, MYB102 has been proven delays leaf senescence and decreases abiotic stress tolerance in Arabidopsis thaliana [62]. This information indicates that even under abiotic stress the up-regulated TF may not necessarily help improve plant abiotic tolerance. In conclusion, our results indicate that TFs regulated linseed drought tolerance is considerably complex but it still helpful for us to understand the molecular mechanisms underlying linseed tolerance to drought stress.

Conclusions

Our results revealed that a group of genes involved in plant drought tolerance were only up-regulated in linseed variety with better drought tolerance under drought stress. In addition, compared with DS, more genes are involved in linseed response to drought stress under RD. Finally, some of the TFs involved in the response to high temperature stress were expressed in linseed under drought stress, indicating that

the linseed response to drought and high temperature stress was combinational rather than independent. Taken together, this study deepens our understanding of the molecular mechanism of linseed drought tolerance and orchestrated linseed responses to repeated drought stress, which frequently occurred under field condition and provides a new perspective to understand the drought adaptability of linseed. To our knowledge, this is the first study to compare and analyze gene expression patterns of linseed with different drought tolerance under different drought treatments on a genome-wide scale using single-molecule long-read sequencing. Therefore, our study will contribute to the current body of knowledge on drought tolerance gene identification and functional analysis in linseed.

Materials And Methods

Phenotyping for drought tolerance in linseed seedling stage

Linseed variety NY-17 was provided by Guyuan Branch of Ningxia Academy of Agriculture and Forestry Sciences, while Z-141 which introduced from Canada, Alberta was provided by Zhangjiakou Academy of Agricultural Sciences. Z141 and NY-17 seeds were surface-sterilized in 1% sodium hypochlorite for 5 min, rinsed in distilled water for six times, and then transferred into 7cm diameter plastic pots filled with mixture of cultivation soil and vermiculite in 1:1 ratio. After germination, the pots were transferred to culture room with a 16h photoperiod and 22℃ temperature.

Drought stress and re-watering began from 20d old after linseed germination. Seedlings were well-watered to keep the suitable soil moisture at (70% soil absolute water content [SAWC]) and to ensure normal plants growth followed by gradual water reduction until SAWC of ~10%. After two days when SAWC was 10%, the stressed plants were watered to reach 70% SAWC to help recovering. The drought stress and re-watering scheme treatments were repeated, and afterward, the irrigation was kept normally until the maturation stage. Leaf tissues were collected during each drought and re-watering treatment from six independent plants with three biological replicates (Figure 1).

Measurement of phenotypic traits

Three drought-tolerant related phenotypic traits such as plant height, biomass, leaf water content (including leaf absolute water content [LAWC] and leaf relative water content [LRWC]) (Supplementary Table S1). LAWC and LRWC were measured as described by Ghashghaie [63] and Yamasaki [64] respectively. Six plants per pot were put in three biological replicates, and the measurement averaged for each trait

RNA isolation, library preparation, and transcriptome sequencing

The total RNA from leaf tissues was extracted using TRIzol reagent (Invitrogen), according to the manufacturer's instructions. RNA concentration was measured using a NanoDrop 2000 spectrophotometer (ND-2000, Thermo Fisher Scientific, Inc., USA). Equimolar rations of 16 samples were pooled together. Approximately, 2μg of total RNA was used for cDNA synthesis using the optimized SMARTer PCR cDNA Synthesis Kit that has been optimized for preparing high-quality, full-length cDNAs (Takara Biotechnology, Dalian, China), followed by size fractionation (1-3 and >3 kb) using the BluePippin™ Size Selection System (Sage Science, Beverly, MA). Iso seq libraries were subsequently constructed using the protocol by (http:// www.pacb.com / wp-content/ uploads/ 2015 09/ Procedure Checklist Isoform Sequencing Iso Seq using the Clontech SMARTer PCR cDNA Synthesis Kit and the BluePippin Size Selection System. pdf). Each of single molecule real-time (SMRT) cells were sequenced using P6 C4 reagent on the PacBio RS II platform with 4 h sequencing movies.

Illumina RNA Seq library construction

mRNA was purified from the total RNA using poly T oligo-attached magnetic beads. Sequencing libraries were generated using the NEBNext® Ultra™ RNA Library Prep Kit for Illumina® (NEB, USA) following the manufacturer's recommendations. The library quality was assessed on the Agilent Bioanalyzer 2100 system.

Subreads processing and error correction

PacBio raw data was pre-processed using the SMRT Pipe analysis workflow of the PacBio SMRT Analysis software suite (http: //www. pacb. com/ products- andservices/ analytical- software/ smrt-analysis/). Circular consensus (CCS) read was obtained from the P_CCS model. Briefly, raw polymerase reads were filtered and trimmed to generate the subreads and read of inserts (ROIs) using the RS_Subreads protocol, requiring a minimum polymerase read length of 50 bp, a minimum polymerase read quality of 0.75, a minimum subread length of 50 bp and a minimum of one full pass. Full-length and non-chimeric (FLNC) reads were regarded as those containing a 5′ adapter, 3′ adapter and poly(A) tail in the expected arrangement with no additional copies of the adapter sequence within the ROI.

Error correction of FLNC reads with the high-quality Illumina short reads was performed using Proovread version 2.12 with the default parameters. The quality of Illumina short reads was examined using FastQC (v0.11.5; http://www. bioinformatics. babraham. ac.uk/projects/fastqc). Sequencing adaptors and low-quality bases in short reads were trimmed before the error correction of FLNC reads. FLNC reads before and after error correction were respectively mapped to the IWGSC RefSeq v1.0 using GMAP (version 2016-09-14; https: //github. com/ juliangehring/ GMAP-GSNAP).

Identification of gene loci and isoforms

Based on read-genome alignments, FLNC reads with the same splicing junctions were collapsed into one isoform. The isoforms with shorter 5' terminal region but shared the introns and splicing sites in the remaining region compared to other isoforms, were regarded as transcripts degraded at the 5' terminal region and were filtered out. For the remaining isoforms, the supporting evidence was examined. We retained isoforms supported with at least two FLNC reads, or one FLNC read with PID higher than 99%, or all junction sites that were fully supported by Illumina reads or annotations of the IWGSC RefSeq v1.0. Isoforms that overlapped by at least 20% of their length on the same strand were considered to be from the same gene locus. Newly discovered loci and isoforms were compared the reference genome annotation using the same criterion as for loci and isoform identification. AS events were classified and characterized by comparing different isoforms of the same gene loci using As profile.

Expression abundance of genes and isoforms

For each samples, the trimmed short reads were mapped to the linseed reference genome (https://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org_Lusitatissimum) using Tophat (v2.1.1; https://ccb.jhu.edu/software/tophat). RSEM (v1.3.0; https:// deweylab. github. io/ RSEM) was used to calculate the isoform-level expression in terms of FPKM and TPM (transcripts per million) (Supplementary Table S13).

Identification of differentially expressed genes and differentially spliced genes

To carry out differential expression analysis, transcript quantification results generated by RESM were processed and refined in successive steps. Firstly, transcript and gene read counts were generated from TPM data correcting for possible gene length variations across samples that mainly derived from differential transcript usage, using the tximport 1.10.0 R package with the option “lengthScaledTPM”. Secondly, the corrected read count data of genes were used to estimate their expression in terms of FPKM. Thirdly, the corrected read count data of genes were imported into the R package EdgeR to identify differentially expressed genes (DEGs) with the criteria of a fold change ≥ 2.0, a false discovery rate [FDR]-adjusted p-value < 0.05 and expression (FPKM≥1) in at least one sample for each comparison (Supplementary Table S14 and S15).

Gene set enrichment and Transcription factors (TFs) analysis

The GO descriptions were obtained by BLAST and BLAST2GO searches and GO enrichment analysis using R package clusterprofiler. Transcription factors prediction was based on Zheng's method using software iTAK to predict the TF by the protein sequence [65].

PCA and heatmap analysis

PCA analysis was based on the principal component analysis of all samples' FPKM values. The first principal component and second principal component values of each sample were calculated and plotted using the R package ggplot2.

We selected the differentially expressed genes from all comparison groups, and then use the expression levels of these differentially expressed genes in all samples to perform hierarchical clustering. Finally, plot the heatmap was plotted using the R package pheatmap.

Statistical analysis

For the phenotypic traits measurement experiment, data from the different drought stresses treatment were analyzed separately. The significant effects of different varieties as fixed effects and different drought stress treatments as random effects were tested in ANOVA. For all comparisons involving pairs of means (Z141 *versus* NY-17), we used independent *t*-test. Statistical analyses were performed using the software package SPSS ver. 21.0 for Windows (IBM Inc., New York, USA).

Validation by RT-PCR

To further evaluate the reliability of our transcriptome data, total RNAs of all the treated samples was extracted using TRIeasy™ Total RNA Extraction Reagent (YEASEN, Shanghai, China) and the first-strand cDNA synthesis was performed by Hifair® 1st Strand cDNA Synthesis Kit (gDNA digester plus) (YEASEN, Shanghai, China) according to manufacturer's protocol. Subsequently, expressions of Glyceraldehyde-3-phosphate dehydrogenase (GAPDH, Lus10014603) and seven candidate genes, including one DNA repair related gene (Lus10021585), two MAPK signaling pathway associated genes (Lus10012962 and Lus10001832), two proline biosynthetic dependent genes (Lus10004697 and Lus10001016), and two photosynthesis related genes (Lus10038490 and Lus10027966) were detected by RT-PCR using the first-strand cDNA of eight treatment materials. The coding sequences of all selected genes were used to design specific amplification primers (Supplementary Table S2) in Primer Premier 6.0 software. All primers were synthesized by Sangon (Shanghai, China). The 20 μL RT-PCR verification system contained 1.0 μL cDNA template, 1.0 μL each of the forward and reverse primers (10 μM), 10 μL 2×Hifair Canace® Gold PCR Master Mix (containing 1.0 U/50 μL polymerase, 1.5 mM Mg2+, and 200 μM dNTP) (YEASEN, Shanghai, China), and 7 μL ddH2O. Double distilled water was used as blank control template. The amplification conditions were as follows: an initial denaturing for 98 °C for 5 min; followed by 34 cycles of 98 °C for 10 s, 60 °C for 20 s, and 72 °C for 30 s; and a final extension at 72 °C for 5 min. Finally, the PCR products were checked by 2.0% agarose gel electrophoresis.

Abbreviations

DS: Drought stress; RW: Re-watering; RD: Repeated drought; DEGs: differentially expressed genes; GO: Gene ontology; RT-PCR: Reverse transcriptional RCR; ALA: α-linolenic acid; EPA: eicosapentaenoic acid; DHA: docosahexaenoic acid; SMRT: Single-molecule real-time; SAWC: soil absolute water content; FL: full-length; LAWC: leaf absolute water content; LRWC: leaf relative water content; CCS: Circular consensus; ROIs: read of inserts; FLNC: Full-length and non-chimeric; DA: Drought avoidance; DT: drought tolerance

Declarations

Acknowledgements

We would like to thanks Guyuan Branch of Ningxia Academy of Agriculture and Forestry Sciences and Zhangjiakou Academy of Agricultural Sciences for kindly supplying the experimental material.

Funding

This research was supported by China Agriculture Research System (CARS-14-1-17), National Infrastructure for Crop Germplasm Resources (NICGR2019-014), National Program for Crop Germplasm Protection of China (2019NWB033), the Fundamental Research Funds for Central Non-profit Scientific Institution (2019CGZH10, 1610172019010)

Availability of data and materials

The raw data has been uploaded to the National Center for Biotechnology Information (NCBI), but we have not received the "review link" yet.

Authors' contributions

WW and XY designed the experiment. WW, LW, LW MT, ZY, JZ, JW and LW performed the experiments and analyzed the data. WW and XY wrote the manuscript. CO checked the manuscript

Competing interests

The authors declare that they have no competing interests

References

1. Zohary D: **Monophyletic and polyphyletic origin of the crops on which agriculture was formed in the Near East**. *Genetic Resources and Crop Evolution* 1999, **46**:133–142.
2. Zuk M, Richter D, Matuła J, Szopa J: **Linseed, the multipurpose plant**. *Industrial Crops and Products* 2015, **75**:165-177.
3. Xie D, Dai Z, Yang Z, Tang Q, Sun J, Yang X, Song X, Lu Y, Zhao D, Zhang L *et al*: **Genomic variations and association study of agronomic traits in flax**. *BMC Genomics* 2018, **19**(1):512.

4. Thambugala D, Duguid S, Loewen E, Rowland G, Booker H, You FM, Cloutier S: **Genetic variation of six desaturase genes in flax and their impact on fatty acid composition**. *Theor Appl Genet* 2013, **126**(10):2627-2641.
5. Soto-Cerda BJ, Duguid S, Booker H, Rowland G, Diederichsen A, Cloutier S: **Association mapping of seed quality traits using the Canadian flax (Linum usitatissimum L.) core collection**. *Theor Appl Genet* 2014, **127**(4):881-896.
6. Kariuki LW, Masinde P, Githiri S, Onyango AN: **Effect of water stress on growth of three linseed (Linum usitatissimum L.) varieties**. *Springerplus* 2016, **5**(1):759.
7. Stallmann J, Schweiger R, Muller C: **Effects of continuous versus pulsed drought stress on physiology and growth of wheat**. *Plant Biol (Stuttg)* 2018, **20**(6):1005-1013.
8. Jin R, Wang Y, Liu R, Gou J, Chan Z: **Physiological and Metabolic Changes of Purslane (Portulaca oleracea L.) in Response to Drought, Heat, and Combined Stresses**. *Front Plant Sci* 2015, **6**:1123.
9. Pinheiro C, Chaves MM: **Photosynthesis and drought: can we make metabolic connections from available data?** *J Exp Bot* 2011, **62**(3):869-882.
10. Hu H, Xiong L: **Genetic Engineering and Breeding of Drought-Resistant Crops**. *Annual Review of Plant Biology* 2014, **65**(1):715-741.
11. Iseki K, Homma K, Shiraiwa T, Jongdee B, Mekwatanakarn P: **The effects of cross-tolerance to oxidative stress and drought stress on rice dry matter production under aerobic conditions**. *Field Crops Research* 2014, **163**:18-23.
12. Kim HJ, Nam HG, Lim PO: **Regulatory network of NAC transcription factors in leaf senescence**. *Curr Opin Plant Biol* 2016, **33**:48-56.
13. Venuprasad R, Lafitte HR, Atlin GN: **Response to Direct Selection for Grain Yield under Drought Stress in Rice**. *Crop Science* 2007, **47**(1):285-293.
14. Ashraf M, Harris PJC: **Photosynthesis under stressful environments: An overview**. *Photosynthetica* 2013, **51**(2):163-190.
15. Soto-Cerda BJ, Cloutier S, Gajardo HA, Aravena G, Quian R: **Identifying drought-resilient flax genotypes and related-candidate genes based on stress indices, root traits and selective sweep**. *Euphytica* 2019, **215**(3):41.
16. Zhang L, Zhou T: **Drought over East Asia: A Review**. *Journal of Climate* 2015, **28**(8):3375-3399.
17. Minetti JL, Vargas WM, Poblete AG, de la Zerda LR, Acuña LR: **Regional droughts in southern South America**. *Theoretical and Applied Climatology* 2010, **102**(3-4):403-415.
18. Rojas O, Vrieling A, Rembold F: **Assessing drought probability for agricultural areas in Africa with coarse resolution remote sensing imagery**. *Remote Sensing of Environment* 2011, **115**(2):343-352.
19. Das PK, Dutta D, Sharma JR, Dadhwal VK: **Trends and behaviour of meteorological drought (1901-2008) over Indian region using standardized precipitation-evapotranspiration index**. *International Journal of Climatology* 2016, **36**(2):909-916.

20. Aprile A, Mastrangelo AM, De Leonardis AM, Galiba G, Roncaglia E, Ferrari F, De Bellis L, Turchi L, Giuliano G, Cattivelli L: **Transcriptional profiling in response to terminal drought stress reveals differential responses along the wheat genome**. *BMC Genomics* 2009, **10**:279.
21. Hayano-Kanashiro C, Calderon-Vazquez C, Ibarra-Laclette E, Herrera-Estrella L, Simpson J: **Analysis of gene expression and physiological responses in three Mexican maize landraces under drought stress and recovery irrigation**. *PLoS One* 2009, **4**(10):e7531.
22. Wang X, Wang H, Liu S, Ferjani A, Li J, Yan J, Yang X, Qin F: **Genetic variation in ZmVPP1 contributes to drought tolerance in maize seedlings**. *Nat Genet* 2016, **48**(10):1233-1241.
23. Kariuki LW, Masinde P, Githiri S, Onyango AN: **Effect of water stress on growth of three linseed (Linum usitatissimum L.) varieties**. *Springerplus* 2016, **5**(1):759.
24. Dash PK, Rai R, Mahato AK, Gaikwad K, Singh NK: **Transcriptome Landscape at Different Developmental Stages of a Drought Tolerant Cultivar of Flax (Linum usitatissimum)**. *Front Chem* 2017, **5**:82.
25. Vembar SS, Seetin M, Lambert C, Nattestad M, Schatz MC, Baybayan P, Scherf A, Smith ML: **Complete telomere-to-telomere de novo assembly of the Plasmodium falciparum genome through long-read (>11 kb), single molecule, real-time sequencing**. *DNA Res* 2016, **23**(4):339-351.
26. Deschamps S, Campbell MA: **Utilization of next-generation sequencing platforms in plant genomics and genetic variant discovery**. *Molecular Breeding* 2009, **25**(4):553-570.
27. Zhang T, Hu Y, Jiang W, Fang L, Guan X, Chen J, Zhang J, Saski CA, Scheffler BE, Stelly DM *et al*: **Sequencing of allotetraploid cotton (Gossypium hirsutum L. acc. TM-1) provides a resource for fiber improvement**. *Nat Biotechnol* 2015, **33**(5):531-537.
28. Li F, Fan G, Lu C, Xiao G, Zou C, Kohel RJ, Ma Z, Shang H, Ma X, Wu J *et al*: **Genome sequence of cultivated Upland cotton (Gossypium hirsutum TM-1) provides insights into genome evolution**. *Nat Biotechnol* 2015, **33**(5):524-530.
29. Li Y, Dai C, Hu C, Liu Z, Kang C: **Global identification of alternative splicing via comparative analysis of SMRT- and Illumina-based RNA-seq in strawberry**. *Plant J* 2017, **90**(1):164-176.
30. Fang L, Su L, Sun X, Li X, Sun M, Karungo SK, Fang S, Chu J, Li S, Xin H: **Expression of Vitis amurensis NAC26 in Arabidopsis enhances drought tolerance by modulating jasmonic acid synthesis**. *J Exp Bot* 2016, **67**(9):2829-2845.
31. Hýskováet DV, Miedzińska L, Dobrá J, Vankova R, Ryšlavá H: **Phosphoenolpyruvate carboxylase, NADP-malic enzyme, and pyruvate, phosphate dikinase are involved in the acclimation of Nicotiana tabacum L. to drought stress**. *J Plant Physiol* 2014, **171**(5):19-25.
32. Zhang X, Liu X, Zhang D, Tang H, Sun B, Li C, Hao L, Liu C, Li Y, Shi Y *et al*: **Genome-wide identification of gene expression in contrasting maize inbred lines under field drought conditions reveals the significance of transcription factors in drought tolerance**. *PLoS One* 2017, **12**(7):e0179477.
33. Jin J, Zhang H, Kong L, Gao G, Luo J: **PlantTFDB 3.0: a portal for the functional and evolutionary study of plant transcription factors**. *Nucleic Acids Res* 2014, **42**(Database issue):D1182-1187.

34. Kumar L, Futschik M: **Mfuzz: A software package for soft clustering of microarray data**. *Bioinformation* 2007, **2**:5-7.
35. Akhtar M, Jaiswal A, G.Taj, Jaiswal JP, Qureshi MI, Singh NK: **DREB1/CBF transcription factors: their structure, function and role in abiotic stress tolerance in plants**. *Indian Academy of Sciences* 2012, **91**:385–395.
36. Scharf KD, Berberich T, Ebersberger I, Nover L: **The plant heat stress transcription factor (Hsf) family: structure, function and evolution**. *Biochim Biophys Acta* 2012, **1819**(2):104-119.
37. Ma X, Zhu X, Li C, Song Y, Zhang W, Xia G, Wang M: **Overexpression of wheat NF-YA10 gene regulates the salinity stress response in Arabidopsis thaliana**. *Plant Physiol Biochem* 2015, **86**:34-43.
38. Witt S, Galicia L, Lisec J, Cairns J, Tiessen A, Araus JL, Palacios-Rojas N, Fernie AR: **Metabolic and phenotypic responses of greenhouse-grown maize hybrids to experimentally controlled drought stress**. *Mol Plant* 2012, **5**(2):401-417.
39. Ghanem ME, Marrou H, Sinclair TR: **Physiological phenotyping of plants for crop improvement**. *Trends Plant Sci* 2015, **20**(3):139-144.
40. Soltys-Kalina D, Plich J, Strzelczyk-Zyta D, Sliwka J, Marczewski W: **The effect of drought stress on the leaf relative water content and tuber yield of a half-sib family of 'Katahdin'-derived potato cultivars**. *Breed Sci* 2016, **66**(2):328-331.
41. Larkunthod P, Nounjan N, Siangliw JL, Toojinda T, Sanitchon J, Jongdee B, Theerakulpisut P: **Physiological Responses under Drought Stress of Improved Drought-Tolerant Rice Lines and their Parents**. *Notulae Botanicae Horti Agrobotanici Cluj-Napoca* 2018, **46**(2):679-687.
42. Eziz A, Yan Z, Tian D, Han W, Tang Z, Fang J: **Drought effect on plant biomass allocation: A meta-analysis**. *Ecol Evol* 2017, **7**(24):11002-11010.
43. Nguyen KH, Mostofa MG, Li W, Van Ha C, Watanabe Y, Le DT, Thao NP, Tran L-SP: **The soybean transcription factor GmNAC085 enhances drought tolerance in Arabidopsis**. *Environmental and Experimental Botany* 2018, **151**:12-20.
44. Lu Y, Xu J, Yuan Z, Hao Z, Xie C, Li X, Shah T, Lan H, Zhang S, Rong T *et al*: **Comparative LD mapping using single SNPs and haplotypes identifies QTL for plant height and biomass as secondary traits of drought tolerance in maize**. *Molecular Breeding* 2011, **30**(1):407-418.
45. Wallace JG, Zhang X, Beyene Y, Semagn K, Olsen M, Prasanna BM, Buckler ES: **Genome-wide Association for Plant Height and Flowering Time across 15 Tropical Maize Populations under Managed Drought Stress and Well-Watered Conditions in Sub-Saharan Africa**. *Crop Science* 2016, **56**(5):2365-2378.
46. Olson ME, Soriano D, Rosell JA, Anfodillo T, Donoghue MJ, Edwards EJ, Leon-Gomez C, Dawson T, Camarero Martinez JJ, Castorena M *et al*: **Plant height and hydraulic vulnerability to drought and cold**. *Proc Natl Acad Sci U S A* 2018, **115**(29):7551-7556.
47. Ghannoum O, Caemmerer Sv, Conroy JP: **The effect of drought on plant water use efficiency of nine NAD - ME and nine NADP - ME Australian C4 grasses**. *Functional Plant Biology* 2002, **29**(11):1337-1348.

48. Schroeder JI, Kwak JM, Allen GJ: **Guard cell abscisic acid signalling and engineering drought hardiness in plants.** *Nature* 2001, **410**:327-330.
49. Lemoine NP, Griffin-Nolan RJ, Lock AD, Knapp AK: **Drought timing, not previous drought exposure, determines sensitivity of two shortgrass species to water stress**. *Oecologia* 2018, **188**(4):965-975.
50. Menezes-Silva PE, Sanglard L, Avila RT, Morais LE, Martins SCV, Nobres P, Patreze CM, Ferreira MA, Araujo WL, Fernie AR *et al*: **Photosynthetic and metabolic acclimation to repeated drought events play key roles in drought tolerance in coffee**. *J Exp Bot* 2017, **68**(15):4309-4322.
51. Tombesi S, Frioni T, Poni S, Palliotti A: **Effect of water stress "memory" on plant behavior during subsequent drought stress**. *Environmental and Experimental Botany* 2018, **150**:106-114.
52. Vandegeer RK, Tissue DT, Hartley SE, Glauser G, Johnson SN: **Physiological acclimation of a grass species occurs during sustained but not repeated drought events**. *Environmental and Experimental Botany* 2020, **171**.
53. Fleta-Soriano E, Munne-Bosch S: **Stress Memory and the Inevitable Effects of Drought: A Physiological Perspective**. *Front Plant Sci* 2016, **7**:143.
54. Udvardi MK, Kakar K, Wandrey M, Montanari O, Murray J, Andriankaja A, Zhang JY, Benedito V, Hofer JM, Chueng F *et al*: **Legume transcription factors: global regulators of plant development and response to the environment**. *Plant Physiol* 2007, **144**(2):538-549.
55. Nuruzzaman M, Sharoni AM, Kikuchi S: **Roles of NAC transcription factors in the regulation of biotic and abiotic stress responses in plants**. *Front Microbiol* 2013, **4**:248.
56. Golldack D, Li C, Mohan H, Probst N: **Tolerance to drought and salt stress in plants: Unraveling the signaling networks**. *Front Plant Sci* 2014, **5**:151.
57. Lata C, Prasad M: **Role of DREBs in regulation of abiotic stress responses in plants**. *J Exp Bot* 2011, **62**(14):4731-4748.
58. Liu Z, Qin J, Tian X, Xu S, Wang Y, Li H, Wang X, Peng H, Yao Y, Hu Z *et al*: **Global profiling of alternative splicing landscape responsive to drought, heat and their combination in wheat (Triticum aestivum L.)**. *Plant Biotechnol J* 2018, **16**(3):714-726.
59. Devaiah BN, Karthikeyan AS, Raghothama KG: **WRKY75 transcription factor is a modulator of phosphate acquisition and root development in Arabidopsis**. *Plant Physiol* 2007, **143**(4):1789-1801.
60. Kumimoto RW, Zhang Y, Siefers N, Holt BF, 3rd: **NF-YC3, NF-YC4 and NF-YC9 are required for CONSTANS-mediated, photoperiod-dependent flowering in Arabidopsis thaliana**. *Plant J* 2010, **63**(3):379-391.
61. Ranjan A, Sawant S: **Genome-wide transcriptomic comparison of cotton (Gossypium herbaceum) leaf and root under drought stress**. *3 Biotech* 2015, **5**(4):585-596.
62. Piao W, Sakuraba Y, Paek N-C: **Transgenic expression of rice MYB102 (OsMYB102) delays leaf senescence and decreases abiotic stress tolerance in Arabidopsis thaliana**. *BMB Reports* 2019, **52**(11):653-658.

63. Ghashghaie J, Brenckmann F, Saugier B: **Water relations and growth of rose plants cultured in vitro under various relative humidities**. *Plant Cell, Tissue and Organ Culture (PCTOC)* 1992, **30**:51-57.
64. Yamasaki S, Dillenburg L: **Measurements of leaf relative water content in araucaria angustifolia**. *Revista Brasileira de Fisiologia Vegetal* 1999, **11** 69-75.
65. Zheng Y, Jiao C, Sun H, Rosli HG, Pombo MA, Zhang P, Banf M, Dai X, Martin GB, Giovannoni JJ *et al*: **iTAK: A Program for Genome-wide Prediction and Classification of Plant Transcription Factors, Transcriptional Regulators, and Protein Kinases**. *Mol Plant* 2016, **9**(12):1667-1670.

Additional Files

Additional file 1: Table S1. Methodology for measuring the drought-tolerant related traits.

Additional file 2: Table S2. Effect of drought stress (SAWS=10%) on drought tolerant related traits in Z141 and NY-17

Additional file 3: Table S3. Effect of drought stress on LAWC and LRWC in Z141 and NY-17.

Additional file 4: Table S4. Primers designed for RT-PCR validation.

Additional file 5: Table S5. Sequence summary of PacBio SMRT Cells.

Additional file 6: Table S6. Sequence summary of PacBio subreads.

Additional file 7: Table S7. Full length evaluation

Additional file 8: Table S8. Classification of FLNC sequences with genome alignment.

Additional file 9: Table S9. Gene structure annotation

Additional file 10: Table S10. Illumina RNA-seq data of each stress.

Additional file 11: Table S11. Comparison of our predicated TFs with that released by PlantTFDB

Additional file 12: Table S12. Detail lists of 15 clusters of differentially expressed transcription factors.

Additional file 13: Table S13. All sample FPKM sheet

Additional file 14: Table S14. All sample reads count sheet

Additional file 15: Table S15. The list of Z141 and NY-17 DEGs under DS or RD

Additional file 16: Figure S1. Bubble diagram showing the GO classification of differentially expressed transcripts between DS and RD in Z141 or NY-17. (a, b)The GO terms of overlapped down-regulated genes between DS and RD in Z141 (a) or NY-17 (b) treatment. (c-f)The GO terms of unique up- (c, d) or

down-regulated (e, f) genes in Z141 under DS or RD respectively. (g-j) The GO terms of unique up- (g, h) or down-regulated (i, j) genes in NY-17 under DS or RD respectively.

Additional file 17: Figure S2. Bubble diagram showing the GO classification of differentially expressed transcripts between Z141 and NY-17 under DS or RD treatment. (a, b)The GO terms of overlapped down-regulated genes between Z141 and NY-17 under DS (a) or RD (b) treatment. (c-f)The GO terms of unique up- (c, d) or down-regulated (e, f) genes in Z141 or NY-17 under DS. (g-j) The GO terms of unique up- (g, h) or down-regulated (i, j) genes in Z141 or NY-17 under RD.

Figures

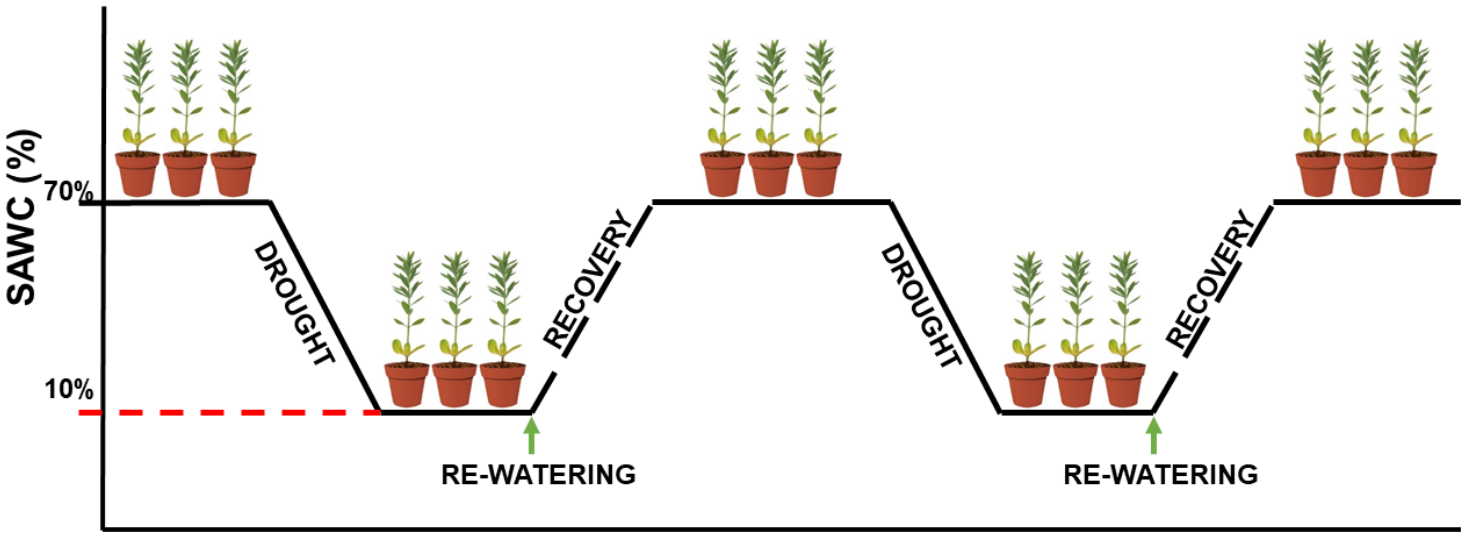

Figure 1

Schematic representation of the cycles of dehydration and re-watering. Each drought cycle consisted of a dehydration phase followed by a recovery period. Dehydration was imposed by suspending irrigation until the soil water content reached approximately SAWC was 10% (red dashed line) and the plants were kept at this soil water status for 2 days, after which the pots were re-watered until the SAWC was 70%. The plants were then maintained during also for 2 days.

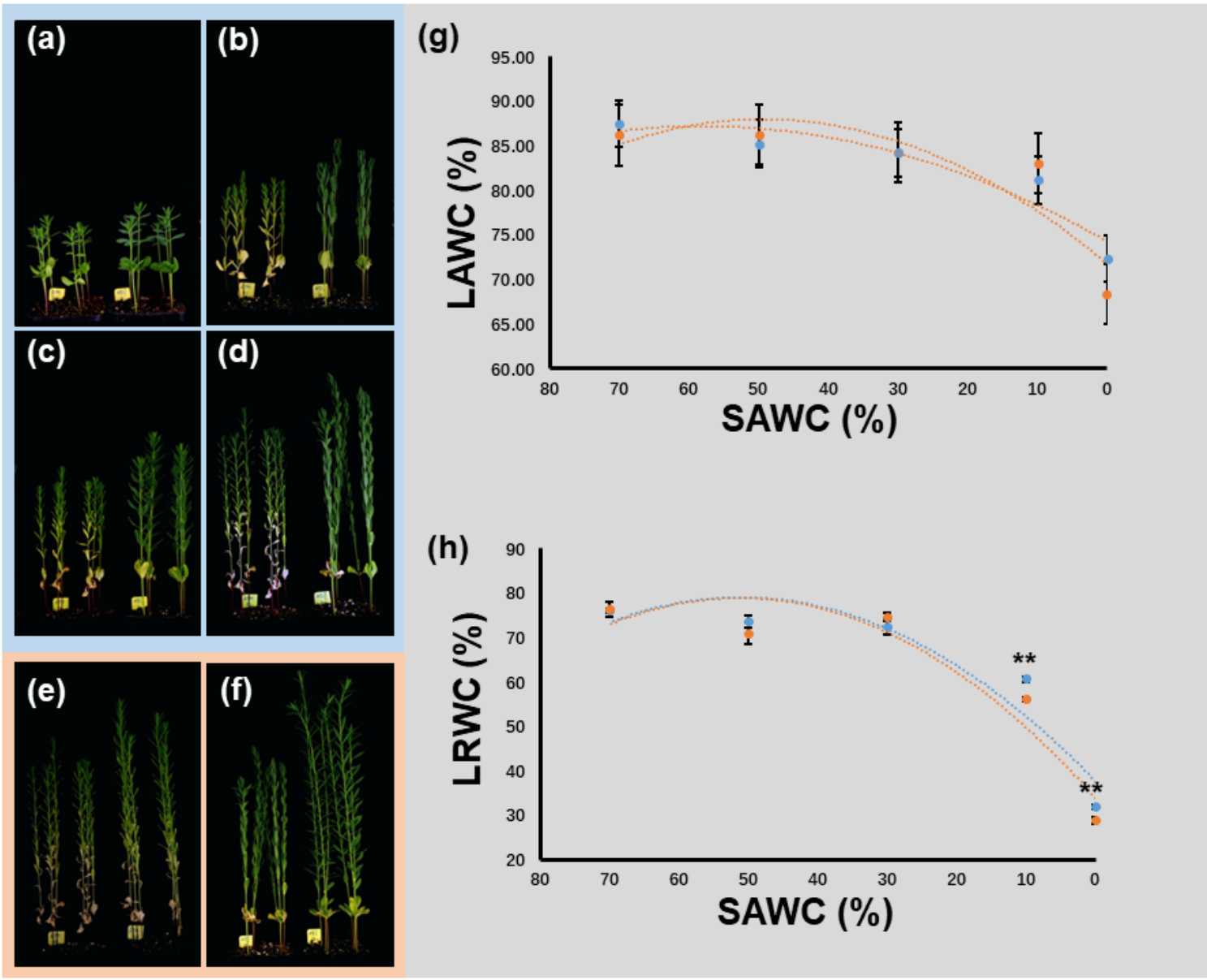

Figure 2

Identification of linseed drought tolerance. (a-d) Z141 (left) and NY-17 (right) phenotype difference under normal water content (CK), drought stress (DS), re-watering (RW), and repeated drought (RD) respectively. (e, f) Z141 and NY-17 phenotypic differences between drought stress (left) and controls (right) respectively. (g, h) Z141 and NY-17 LAWC and LRWC with means and SE (n= 3) respectively. The abscissa indicates SAWC and the ordinates indicates LAWC (g) and LRWC (h) respectively. Blue dots indicates Z141 and orange dots indicates NY-17. **, $p<0.01$, see table 1 for ANOVA and Supporting information table S2 and table S3 for a summary of these drought tolerant related traits.

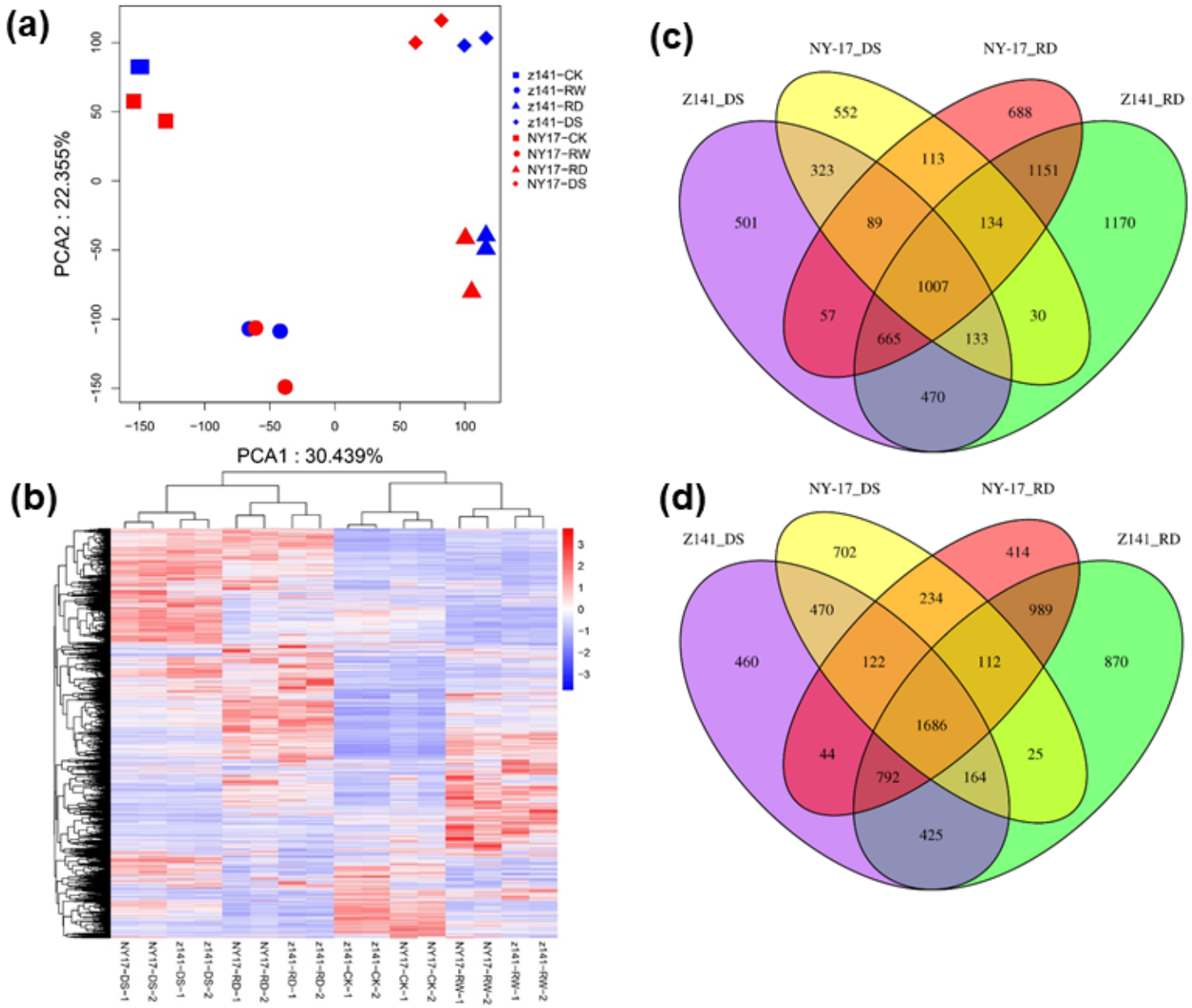

Figure 3

Comparative analysis of transcriptome profiles of linseed seedling leaves under DS and RD. (a) Principal component analysis (PCA) of mRNA populations from control, DS, RW and RD, each sample contained two replicates. Principal components (PCs) 1 and 2 account for 30 % and 22 % of the variance, respectively. PCA plot shows four distinct groups of mRNA populations. Group 1: Z141 CK (blue square) and NY-17 CK (red square); Group II: Z141 DS (blue diamond) and NY-17 DS (red diamond); Group III: Z141 RW (blue circle) and NY-17 RW (red circle) and Group IV: Z141 RD (blue triangle) and NY-17 RD (red triangle). (b) Hierarchical clustering of DEGs with altered expression levels in response to CK, DS, RW and RD The color scale of blue (low), white (medium) and red (high) represents the normalized expression levels of differentially expressed DEGs. (c, d) Venn diagrams showing overlap of up- (c) or down-regulated (d) genes in response to the four assayed abiotic stresses at Z141-DS (purple), NY-17-DS (yellow), Z141-RD (green) and NY-17-RD (red).

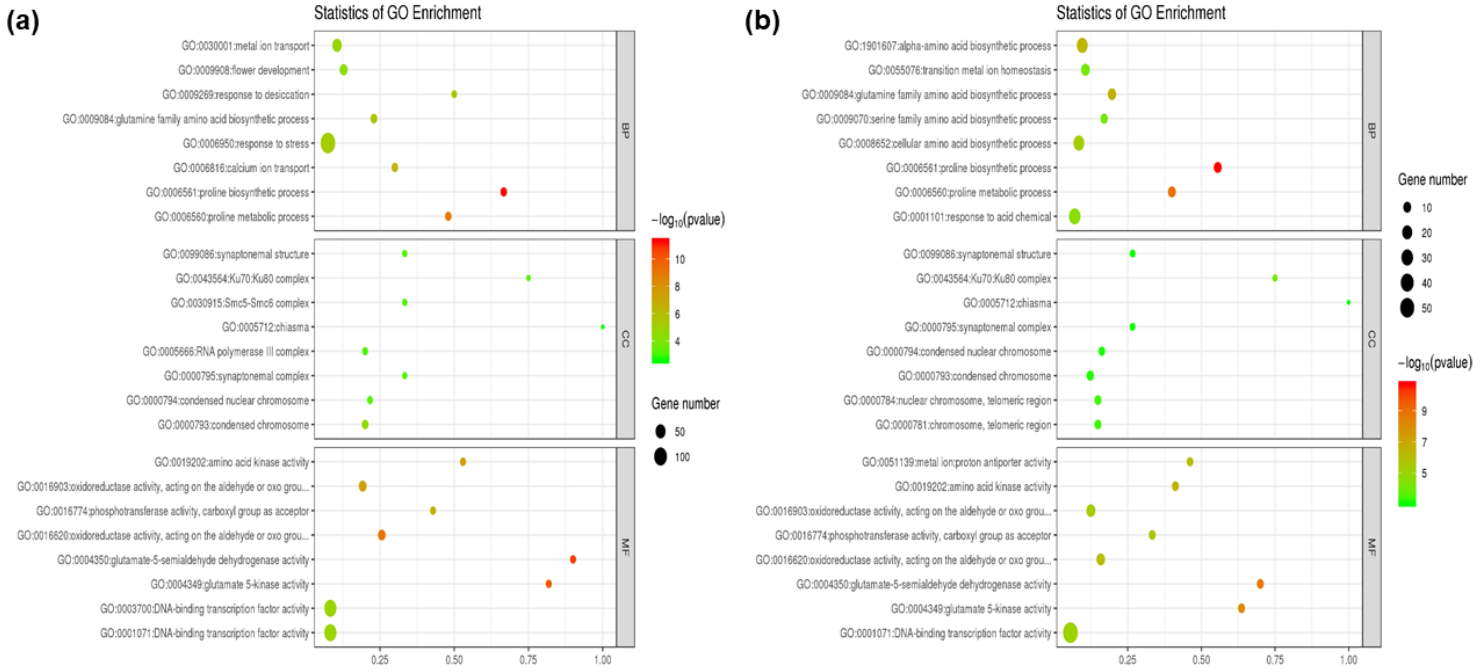

Figure 4

Bubble diagram showing the Gene Ontology (GO) classification of overlapped up-regulated genes between DS and RD in Z141 or NY-17. The GO terms of overlapped up-regulated genes between DS and RD in Z141 (a) or in NY-17 (b) respectively. The three main GO categories are (from top to bottom): biological process, cellular component and molecular function.

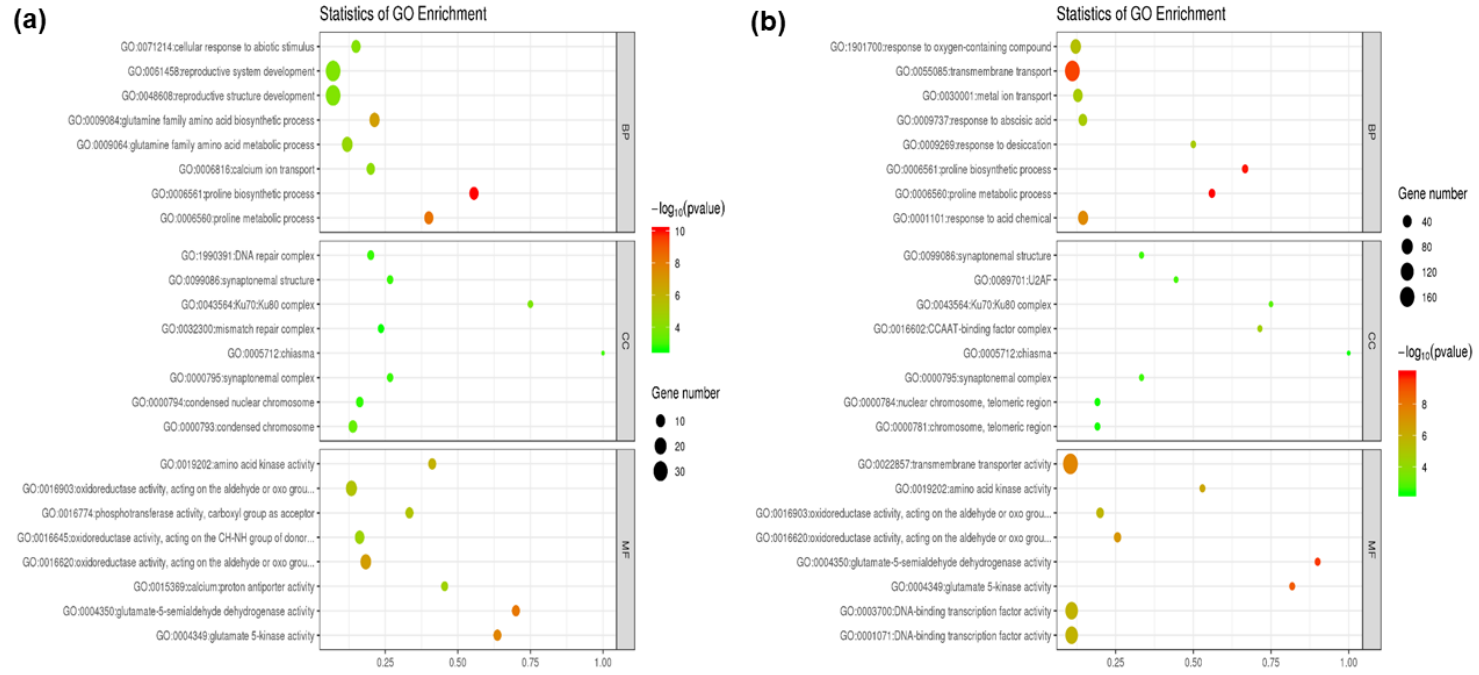

Figure 5

Bubble diagram showing the P value significance of enriched GO categories for Z141 and NY-17 overlapped up-regulated responsive genes under DS or RD. The GO terms of overlapped up-regulated

genes between Z141 and NY-17 under DS (a) or RD (b) treatment respectively.

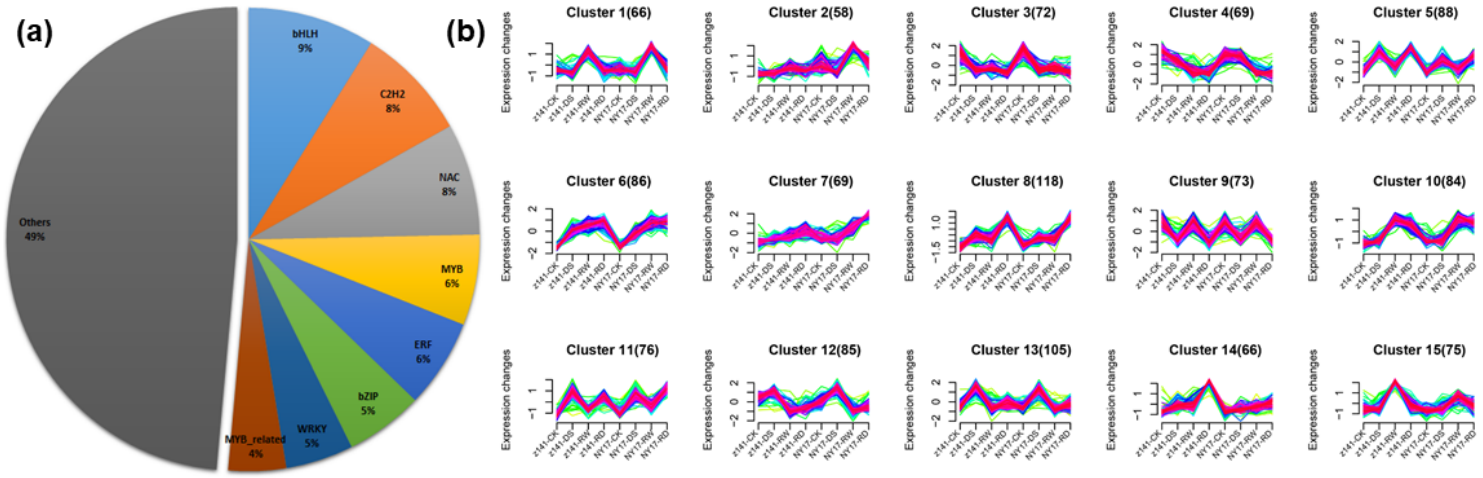

Figure 6

Clustering analysis of DS and HD responsive TFs. (A) Pie chart showing top 8 TF families which contain approximately 50 % of differentially expressed TF genes. (B) Clustering of the differentially expressed TFs based on their expression patterns in response to Z141 and NY-17 under CK, DS, RW and RD. 15 clusters comprising of 1190 TFs are exhibited here, the numbers in parentheses indicate TF amount in corresponding clusters. X axis represents treatment conditions and y axis represents centralized and normalized expression value. The red lines represent the mean expression trend of TFs (green lines) belonging to each cluster.

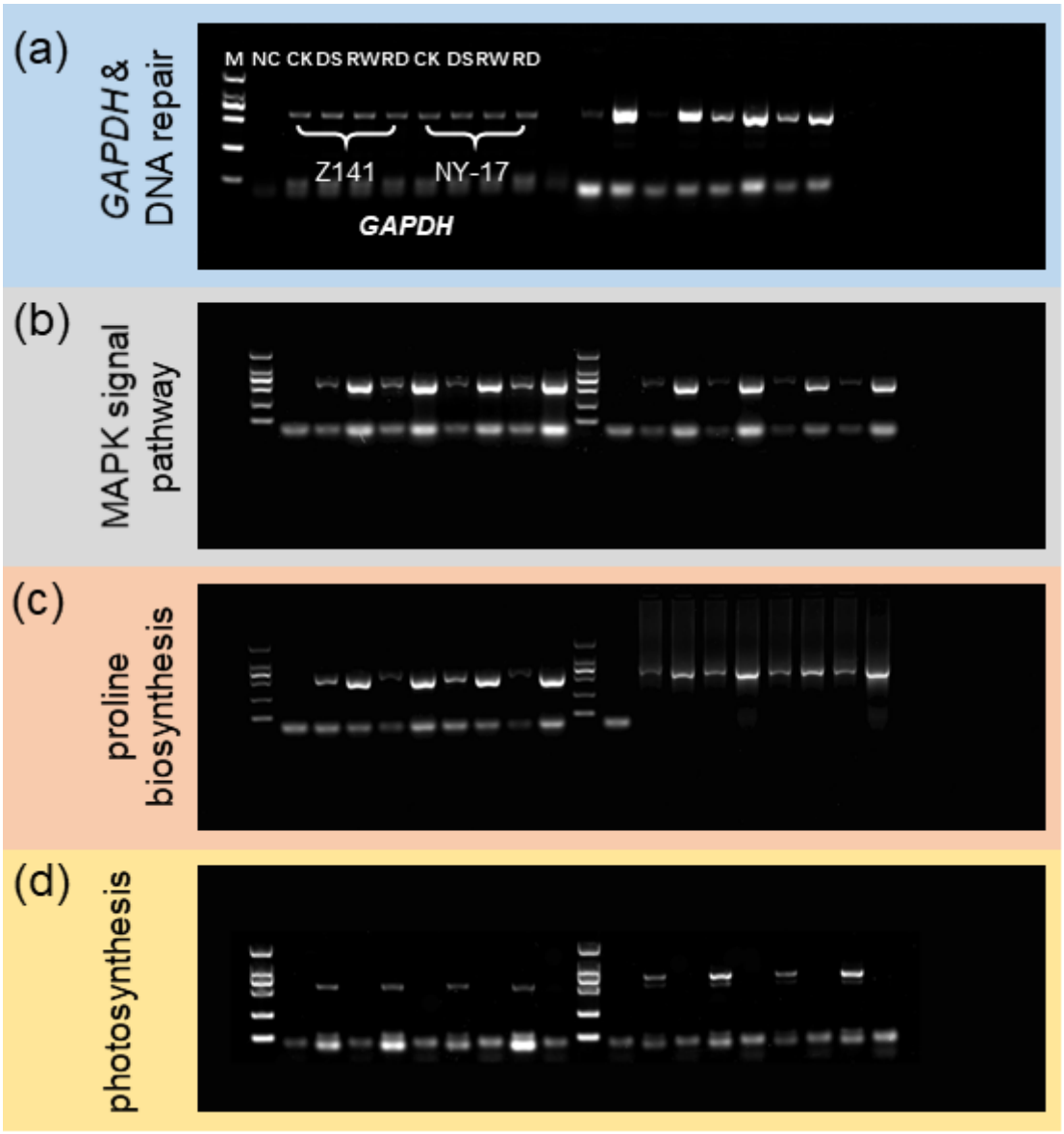

Figure 7

Validation of differential expressed genes by RT-PCR. (a-d) The gene expression changes of GAPDH, DNA repair, MAPK signaling pathway, proline biosynthesis and photosynthesis genes under CK, DS, RW and RD in Z141 and NY-17. M indicates the DL2000 DNA marker; NC indicates negative control and CK, DS, RW, RD indicates control check, drought stress, re-watering and repeat drought respectively.

Supplementary Files

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

- TableS13.xls
- TableS1S11.docx
- TableS14.xls

- TableS12.xlsx
- TableS15.xlsx
- SupplementaryFigureS1.rar
- SupplementaryFigureS2.rar