Analysis of Soybean DREB Transcription Factors Family and DREB Time Expression Profile under Drought Stress

Yaxing Zhou
Inner Mongolia University for Nationalities
Wei Zhou
Inner Mongolia University for Nationalities
Hui Liu
Chifeng Lv Lu Seed Co.,Ltd
Peng Liu
Inner Mongolia University For Nationalities
Zhigang Li (breeding@126.com)
Inner Mongolia University for Nationalities

Research article

Keywords: Glycine max, DREB, Drought Stress, Transcriptome

Posted Date: October 8th, 2019

DOI: https://doi.org/10.21203/rs.2.15699/v1

Abstract

Background: DREB transcription factors regulate the expression of stress-responsive genes, and thus play an important role in plant stress response to abiotic stresses. To obtain a global expression profile of soybean DREB genes under drought stress, we first explored the soybean genome to identify DREB family genes, and then analyzed a set of transcriptome data of drought stress to verify their involvement in stress response.

Results: We identified 73 DREB family genes from the soybean genome. These DREB genes were further divided into six subgroups basing on the phylogenetic analysis. Gene structure analysis showed that most DREB genes have a single exon. Soybean DREB genes were unevenly distributed on 19 chromosomes. We further identified 186 putative target genes of soybean DREB proteins and found that these targeted genes were significantly enriched in metabolism pathways of fructose and mannose. The analysis of transcriptome data after 4 hours, 8 hours, and 12 hours of drought, the expression of DREB genes were constantly increased, indicating that the DREB family genes responded to drought stress.

Conclusion: We identified 73 DREB genes from soybean genome and analyzed their features including gene structures, protein motifs, and putative target genes.

Background

Transcription factors participate in many biological processes such as plant growth, development, metabolism, reproduction and differentiation [1, 2]. DREB (Dehydration responsive element binding protein) transcription factors are a subgroup of AP2/ERF (APETALA2/ethylene-Responsive factor) family, and widely involved in drought, high salt, cold, heat and other stress response [3]. The common feature of this group of proteins is that there is a 60–70 amino acid AP2 domain, without other obvious homology in the sequences of other regions [4]. DREB transcription factors enhance plant resistance to different abiotic stress by binding to a DRE/CRT (dehydration responsive element) *cis*-element in the promoter region of the stress-resistant gene, regulating a series of downstream gene (which contain DRE/CRT elements) expressions. DREB is one of the most promising transcription factor families for plant resistance to molecular breeding [5]. Many of plants have conducted resistance research of DREB. In Arabidopsis, DREB1A, DREB1B and DREB1C were induced by low temperature, and DREB2A and DREB2B were induced by drought, high salt and heat stress [6, 7]. But not all DREB2 genes respond to stress. *AtDREB2E* and *AtDREB2H* are reported to be unable to be induced by stress [8]. There are also induced expression of some DREB genes that can simultaneously respond to low temperature, drought and high salt stress. For example, SwDREB1 gene in the fibrous roots from sweet potato response to various stresses, such as drought, low temperature, high salt, and cadmium ion stress [9].

Soybean (*Glycine max)* is one of the most important commercial crop in the world [10], and a traditional grain oil crop in China. It not only has high nutritional value, but also has an important proportion in people's daily dietary structure, and has a high utilization value. The development of soybean production

is of great significance for improving the nutritional level of the people and promoting the development of the national economy. Recently, diseases, freezing and drought greatly affect the yield of soybeans. Although there are many studies on anti-physiological and biochemical reactions, there are few studies on the molecular basis of anti-resistance, the regulation of the expression of stress-related genes and the stress signals involved.

With the availability of soybean genome sequence and the advance of next-generation sequencing technologies, studies of *DREB* genes will greatly enhance our knowledge of how soybeans response to abiotic stress and will facilitate soybean breeding [11]. The genome-wide identification and mining of DREB transcription factor family genes in soybean under drought stress were systematically carried out. These results could serve as a foundation for further functional study of soybean *DREB* family genes.

Results

Identification and annotation of DREB gene family

The DREB transcription factor family is a subfamily of the ERF family and is mainly involved in plant abiotic stress. Through sequence alignment and cluster analysis, 73 *DREB* genes were identified from the soybean genome. These genes were unevenly distributed on 19 chromosomes (Supplementary Table 1). The largest number of *DREB* genes was fund on chromosome 2 (8 genes) while chromosome 20 had the smallest number (only 1 gene). All of DREB genes belong to ethylene-responsive transcription factor and dehydration-responsive element-binding protein based on swissProt database anotation. Specific information on these genes is shown in Supplementary Table 1.

Phylogenetic tree of *DREB* genes and collinear analysis

Phylogenetic analysis showed that 73 *DREB* genes were subdivided into six subgroups. To understand the mechanisms of expansion of the *DREB* gene family in soybean, tandem and segmental duplication events of DREB gene family were investigated through genome synteny analysis. Among the 73 DREB in soybean, we identified 8 genes within 4 pairs of synthetic blocks (Fig. 2).

Motif analysis of soybean DREB gene family

Structures analysis revealed (Fig. 3) that most soybean *DREB* genes had no introns. MEME [12] is a motif-based sequence analysis tools, and was used to perform motif sequence prediction. Motif analysis (Fig. 4) showed that soybean DREB proteins contained 10 conservative functional elements in total. The AP2 domain and motif 2 (PTPEMAARAYDVAALALKGPSARLNFPEL) were found in all proteins. Motif 8 (QVDDTPSSFIYQLQNPDAKLLGSLPHMEQTPSGFDYGLDFLKTVEPGDYN) only existed in four proteins: KRH14907.1, KRH14906.1, KRH73245.1, and KRH73244.1.

Targeted gene prediction of DREB and functional enrichment analysis

The DREB subfamily proteins recognize the DRE/CRT motif (core sequence A/GCCGAC). We extracted 2kb promoter regions of all genes and got 186 genes containing this motif (Supplementary Table 2). Enrichment analysis showed that these putative targeted genes were significantly enriched in pathways of fructose and mannose metabolism (Fig. 5).

Analysis of transcriptome data of soybean drought stress

To further investigate the transcriptional changes of DREB genes in soybean under drought stress, we selected the previously published soybean drought stress transcriptome data (PRJNA285677, NCBI) for quantitative and differential expression analysis. The results showed that (Fig. 6) compared with corresponding control, 132, 1164 and 1477 genes were found differently expressed at ZT4, ZT8, and ZT12, respectively. In order to further explore the expression variation of DREB family genes in different samples, pheatmap was used to cluster expression of *DREB* genes (Fig. 7). Compared with control groups, *DREB* genes in treatment group had a big change. In the early stages of drought stress, gene21317 (protein ID KRH44298.1) and gene46233 (protein_ID KRH04169.1) were highly expressed in the group of ZT4. With the extension of drought time, gene13714 (protein ID *KRH52029.1, KRH52027.1, KRH52028.1),* gene36865(*KRH22495.1)* and gene53012(*KRG97200.1)* had high expression level in the group of ZT8. Further extension with the drought time, in the group of ZT12, seven different DREB genes were highly expressed (Supplementary Table3). The heatmap (Fig. 8) of corresponding targeted genes proved that targeted genes had significant changes.

Discussion

Transcription factors can specifically bind to the cis-acting elements of the promoter region, thereby regulating the expression of plant stress response genes, so more and more research focuses on the key role of transcription factors in plants under external environmental stress [13, 14]. Many studies on DREB family focused on gene types and expression characteristics, and transgenic studies. Limited studies on plant stress response and expression regulation and its complex regulatory pathways have constrained the application of DREB transcription factors in breeding practice [15–18]. Based on the whole genome information of soybean, combined with the transcriptome data of under drought stress, bioinformatics analysis was carried out in this study to analyze the soybean DREB transcription factor family. The obtained results may serve as a foundation to further understand the regulation network of soybean *DREB* genes.

We identified a total number of 73 DREB gene family members (Supplementary Table1). Based on phylogenetic tree analysis, these genes can be divided into six subgroups (Fig. 2). In soybean, the DREB family gene is more abundant than other species. A total of 60 *DREB* genes were identified in apples, 56

in Arabidopsis, 52 in rice, and 41 in sesame [19]. This may also be one of the causes of soybean drought resistance [20].

In many species, the *DREB* genes have a single exon [21]. This is consistent with our findings. 70 of the 73 genes are single exon structures (Fig. 3). We found 10 motif sequences from soybean DREB proteins. The AP2 domain was present in every protein as expected (Fig. 4). It is also because of the existence of this structure that it plays an important role in the process of plant resistance to abiotic stress.

Drought stress is what plants often cope with, and it has a great influence on the growth and development of plants. Plants regulate their physiological and biochemical indicators, and the changes in the expression of many genes in response to this stress [22–24]. DREB proteins participate in regulation and expression of stress response genes, and play an important role in coping with low temperature, drought, high salt and other stresses [25, 26]. It has been reported that under drought stress Arabidopsis plants with overexpressed *DREB1* gene grew well while the control group died [27].

In upland cotton, it was found that DREB genes *GhDBP3, GhDBP1,*and *GhDBP2* participated in the dehydration response pathway of ABA [28]. We found this was also true for soybean. Heatmap analysis (Fig. 7) showed that with the extension of drought time, the number and category of DREB genes in response to drought stress were changed (Fig. 7).

In the early stages of drought stress, gene21317 and gene46233, the soybean homologs of Arabidopsis *DREB3* and *ERF039,* respectively, were highly expressed in the group of leaves-ZT4. AtDREB3 and AtERF039 belong to Dehydration-responsive element-binding proteins and Ethylene-responsive transcription factors, respectively. AtDREB3 binds to the DRE element. It has been found that in tomato the expression of *LeDREB3* was significantly induced by NaCl, drought, low temperature and H_2O_2. Moreover, *LeDREB3* was slightly up regulated after treatment with ABA [29].

With the extension of drought time, three soybean genes including gene13714, the soybean homologs of Arabidopsis RAP2.1 had high expression level in the group of ZT8 (Supplementary Table 3). RAP2.1 protein sequence possessed an AP2 domain and an ERF-associated amphiphilic repression (EAR) motif. Arabidopsis plants overexpressing *RAP2.1* show enhanced resistance to cold and drought stresses [30]. This may explain why *RAP2.1* gene was highly expressed at ZT8. In the third stage of the experiment, in the group of leaves-ZT12h, sevenDREB genes highly express including RAP2–1(Supplementary Table3). However, the specific functions of some genes are still unclear and further research is needed.

Conclusion

In summary, we identified 73 DREB genes from soybean genome and analyzed their features including gene structures, protein motifs, and putative target genes. Transcriptome analysis showed that members of these genes did involved in drought stress, and expression of their putative target genes also changed. These results could serve as a foundation for further studies such as functional redundancy analysis, and finally facilitate breeders to improve soybean resistance to abiotic stress.

Methods

Identification of DREB gene in Soybean Genome

The Glycine max genome (Glycine_max v2.1, NCBI assemble GCA_000004515.4) was used to identify all members of *DREB* family genes. Since DREB subfamily member proteins contain only one AP2 domain, soybean proteins were searched against the pfam database (Pfam: the protein families database) with customized perl script to identify proteins containing only one AP2 conserved domain and not including the B3 domain.

Phylogenetic analysis and classification of DREB gene family

Multiple alignment of the protein sequences of identified DREB genes were conducted with ClustalW [31]. A neighbor-joining phylogenetic tree was then generated using MEGA7 with bootstrap repeated for 1000 times [32].

The localization of DREB genes in soybean genome was made using the software MapInspect (http://www.softsea.com/review/MapInspect.html citation). MEME was used to perform motif sequence prediction. The gene structures were analyzed with GSDS (citation; http://gsds.cbi.pku.edu.cn/).

Gene family collinear analysis

Proteins were queried using local BlastP of NCBI's blast package against all the genome-coded soybean proteins, with E value < 1e−5, identity >90%, and coverage > 75%. Matched protein pairs were considered as duplicators and further analyzed by MCScan (http://chibba.pgml.uga.edu/mcscan2/).

Target gene prediction of DREB proteins

Members of the DREB subfamily proteins recognize drought and cold-inducible response elements (DRE/CRT, core sequence A/GCCGAC). A customized perl script was used to search for DRE elements in the promoter regions (2kb upstream of start codens) of all genes in the soybean genome.

Transcriptome analysis of soybean drought stress

A set of RNA-seq data for soybean leaf samples deposited in NCBI's Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc = GSE69469) was retrieved to analysis expression pattern of *DREB* subfamily genes. This data set is the leaf transcriptome results of soybean leaf samples collected at six time-points in a 4-hour interval during a 24-hour course under water deficit stress. Reads

were aligned to soybean genome with Hisat2 [33]. Gene quantification and subsequent differential expression analysis were carried out using feature Counts [34] The FPKM (Fragments per Kilobase Million) values were calculated with Cufflinks [35].

Abbreviations

DREB: Dehydration responsive element binding protein; AP2/ERF: APETALA2/ethylene-Responsive factor; DRE/CRT: dehydration responsive element; EAR: ERF-associated amphiphilic repression

Declarations

Ethics approval and consent to participate

This study did not involve any endangered or protected species and followed all relevant ethical guideline. The samples examined in this study were used as agricultural plant in China.

Consent for publication

Not applicable.

Availability of data and material

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Competing interests

All authors declare that they have no financial or other conflicts of interest in relation to this research and its publication.

Funding

This research was supported by Inner Mongolia Autonomous Region Science and Technology Plan (KJJH1702); Tongliao City and Inner Mongolia University for Nationalities Cooperation Project (SXYB2012050), and Inner Mongolia Autonomous Region Science and Technology Reserve Project (2018MDCB02).

Authors' contributions

LZ as the corresponding author contributed to the conception of the study; YZ as the first author contributed significantly to analysis and manuscript preparation; ZW as the second author performed the data analyses and helped to write the manuscript; LH as the third author wrote the manuscript and helped to prepare the material; LP as the fourth author helped perform the analysis with constructive discussions. All authors read and approved the final manuscript.

Acknowledgements

This study is grateful to Zhigang Li from Agricultural College of Inner Mongolia University for Nationalities.

References

1.Yanagisawa S. A novel DNA-binding domain that may form a single zinc finger motif. *Nucleic acids research* 1995; 23:3403–3410.

2.Kim JH, Tsukaya H. Regulation of plant growth and development by the GROWTH-REGULATING FACTOR and GRF-INTERACTING FACTOR duo. *Journal of experimental botany* 2015; 66:6093–6107.

3.Xu ZS, Chen M, Li LC, Ma YZ. Functions and application of the AP2/ERF transcription factor family in crop improvement. *Journal of integrative plant biology* 2011; 53:570–585.

4.Kizis D, Lumbreras V, Pages M. Role of AP2/EREBP transcription factors in gene regulation during abiotic stress. *FEBS letters* 2001; 498:187–189.

5.Lata C, Prasad M. Role of DREBs in regulation of abiotic stress responses in plants. *Journal of experimental botany* 2011; 62:4731–4748.

6.Ederli L, Bianchet C, Paolocci F, Alqurashi M, Gehring C, Pasqualini S. Drought stress induces a biphasic NO accumulation in Arabidopsis thaliana. *Plant signaling & behavior* 2019; 14:e1573098.

7.Li X, Liang Y, Gao B, Mijiti M, Bozorov TA, Yang H, Zhang D, Wood AJ. ScDREB10, an A–5c type of DREB Gene of the Desert Moss Syntrichia caninervis, Confers Osmotic and Salt Tolerances to Arabidopsis. *Genes* 2019; 10.

8.Sakuma Y, Liu Q, Dubouzet JG, Abe H, Shinozaki K, Yamaguchi-Shinozaki K. DNA-binding specificity of the ERF/AP2 domain of Arabidopsis DREBs, transcription factors involved in dehydration- and cold-inducible gene expression. *Biochemical and biophysical research communications* 2002; 290:998–1009.

9.Kim YH, Yang KS, Ryu SH, Kim KY, Song WK, Kwon SY, Lee HS, Bang JW, Kwak SS. Molecular characterization of a cDNA encoding DRE-binding transcription factor from dehydration-treated fibrous roots of sweetpotato. *Plant physiology and biochemistry: PPB* 2008; 46:196–204.

10.Yu X, Yuan F, Fu X, Zhu D. Profiling and relationship of water-soluble sugar and protein compositions in soybean seeds. *Food chemistry* 2016; 196:776–782.

11.Velasco R, Zharkikh A, Affourtit J, Dhingra A, Cestaro A, Kalyanaraman A, Fontana P, Bhatnagar SK, Troggio M, Pruss D *et al.* The genome of the domesticated apple (Malus x domestica Borkh.). *Nature genetics* 2010; 42:833–839.

12.Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS. MEME SUITE: tools for motif discovery and searching. *Nucleic acids research* 2009; 37:W202–208.

13.RK G, D T, N C, D G-PF, C W, BE E, M F, A L, NA F. Analysis of MAPK and MAPKK gene families in wheat and related Triticeae species. *BMC genomics* 2018; 19:178.

14.Y W, M M, D W, Q Y, F Y, Mandlaa, G L, R W. Identification of the WRKY gene family and functional analysis of two genes in Caragana intermedia. *BMC plant biology* 2018; 18:31.

15.M K, S K, T Y, J M, D T, AR F, K S, K Y-S. Double overexpression of DREB and PIF transcription factors improves drought stress tolerance and cell elongation in transgenic plants. *Plant biotechnology journal* 2017; 15:458–471.

16.H L, Y Z, Q G, W Y. Molecular characterisation of a DREB gene from Sophora moorcroftiana, an endemic species of plateau. *Protoplasma* 2017; 254:1735–1741.

17.H L, D Z, X L, K G, H Y. Novel DREB A–5 subgroup transcription factors from desert moss (Syntrichia caninervis) confers multiple abiotic stress tolerance to yeast. *Journal of plant physiology* 2016; 194:45–53.

18.Y C, L H, H Y, X Z, B X, X M. Cloning and characterization of an ABA-independent DREB transcription factor gene„ in. *Hereditas* 2016; 153:3.

19.K D, X W, D L, D F, Y Z, L W, J Y, L B, D D, N C *et al.* Insight into the AP2/ERF transcription factor superfamily in sesame and expression profiling of DREB subfamily under drought stress. *BMC plant biology* 2016; 16:171.

20.C C, Y W, T J, B V, S P, L M, D X, HT N. Soybean transcription factor ORFeome associated with drought resistance: a valuable resource to accelerate research on abiotic stress resistance. *BMC genomics* 2015; 16:596.

21.T Z, D L, P W, J L, F M. Genome-wide analysis and expression profiling of the DREB transcription factor gene family in Malus under abiotic stress. *Molecular genetics and genomics: MGG* 2012; 287:423–436.

22.H K, MR K. Abiotic stress signalling pathways: specificity and cross-talk. *Trends in plant science* 2001; 6:262–267.

23.W C, NJ P, J G, F K, HS C, T E, F M, S L, G Z, SA W *et al.* Expression profile matrix of Arabidopsis transcription factor genes suggests their putative functions in response to environmental stresses. *The Plant cell* 2002; 14:559–574.

24.M S, M N, J I, T N, M F, Y O, A K, M N, A E, T S *et al.* Monitoring the expression profiles of 7000 Arabidopsis genes under drought, cold and high-salinity stresses using a full-length cDNA microarray. *The Plant journal: for cell and molecular biology* 2002; 31:279–292.

25.N A, F H. CBF/DREB transcription factor genes play role in cadmium tolerance and phytoaccumulation in Ricinus communis under molybdenum treatments. *Chemosphere* 2018; 208:425–432.

26.Eckardt NA. DREB Duo Defines Distinct Drought and Cold Response Pathways. *The Plant cell* 2019; 31:1196–1197.

27.Q L, M K, Y S, H A, S M, K Y-S, K S. Two transcription factors, DREB1 and DREB2, with an EREBP/AP2 DNA binding domain separate two cellular signal transduction pathways in drought- and low-temperature-responsive gene expression, respectively, in Arabidopsis. *The Plant cell* 1998; 10:1391–1406.

28.M C, QY W, XG C, ZS X, LC L, XG Y, LQ X, YZ M. GmDREB2, a soybean DRE-binding transcription factor, conferred drought and high-salt tolerance in transgenic plants. *Biochemical and biophysical research communications* 2007; 353:299–305.

29.MS I, MH W. Expression of dehydration responsive element-binding protein–3 (DREB3) under different abiotic stresses in tomato. *BMB reports* 2009; 42:611–616.

30.CJ D, JY L. The Arabidopsis EAR-motif-containing protein RAP2.1 functions as an active transcriptional repressor to keep stress responses under tight control. *BMC plant biology* 2010; 10:47.

31.Hung JH, Weng Z. Sequence Alignment and Homology Search with BLAST and ClustalW. *Cold Spring Harbor protocols* 2016; 2016.

32.Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Molecular biology and evolution* 2016; 33:1870–1874.

33.Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nature biotechnology* 2019; 37:907–915.

34.Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics (Oxford, England)* 2014; 30:923–930.

35.Ghosh S, Chan CK. Analysis of RNA-Seq Data Using TopHat and Cufflinks. *Methods in molecular biology (Clifton, NJ)* 2016; 1374:339–361.

Figures

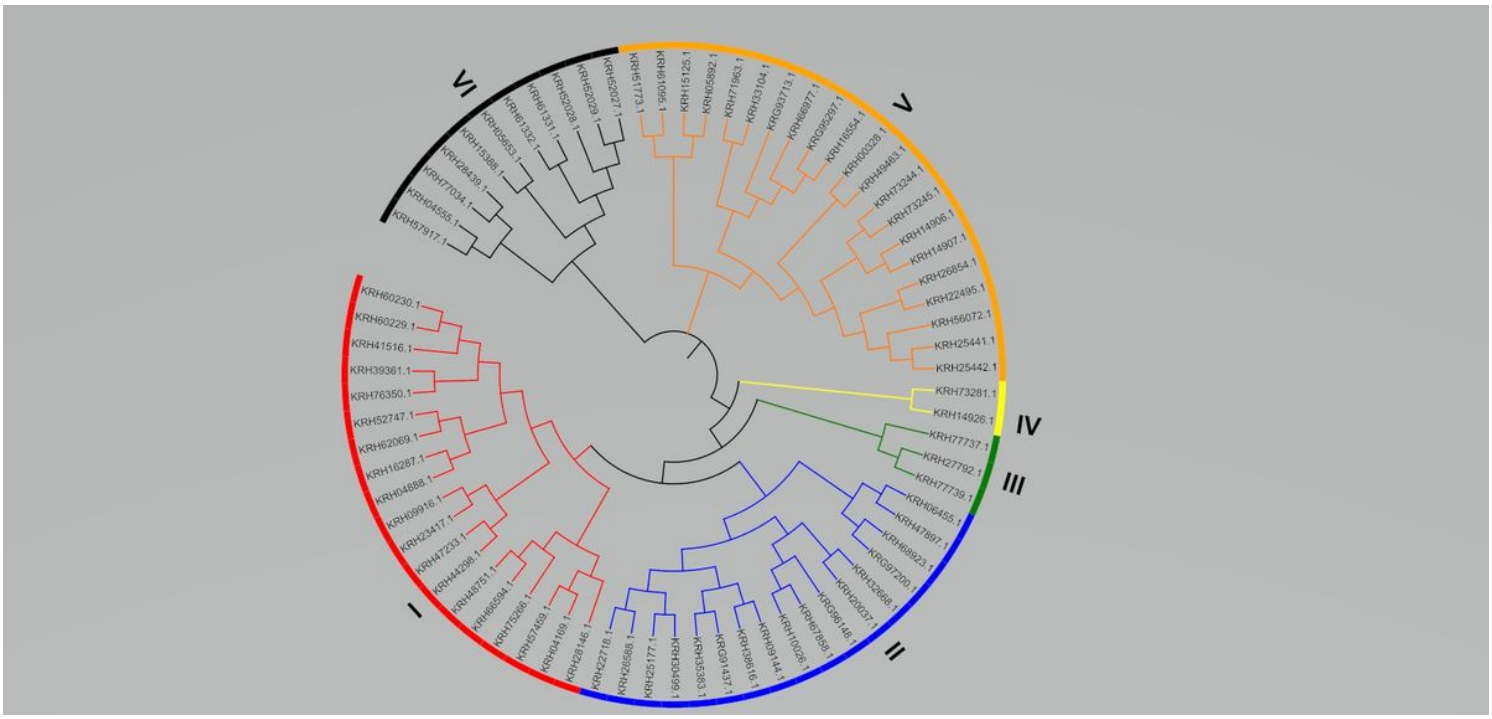

Figure 1

Phylogenetic analysis of DREB family members in soybean. The phylogenetic tree depicts 73 predicted DREB in soybean. All DREBs were classified into one of six groups (groups I-Ⅵ); the distinct groups are shown by differently colored branches.

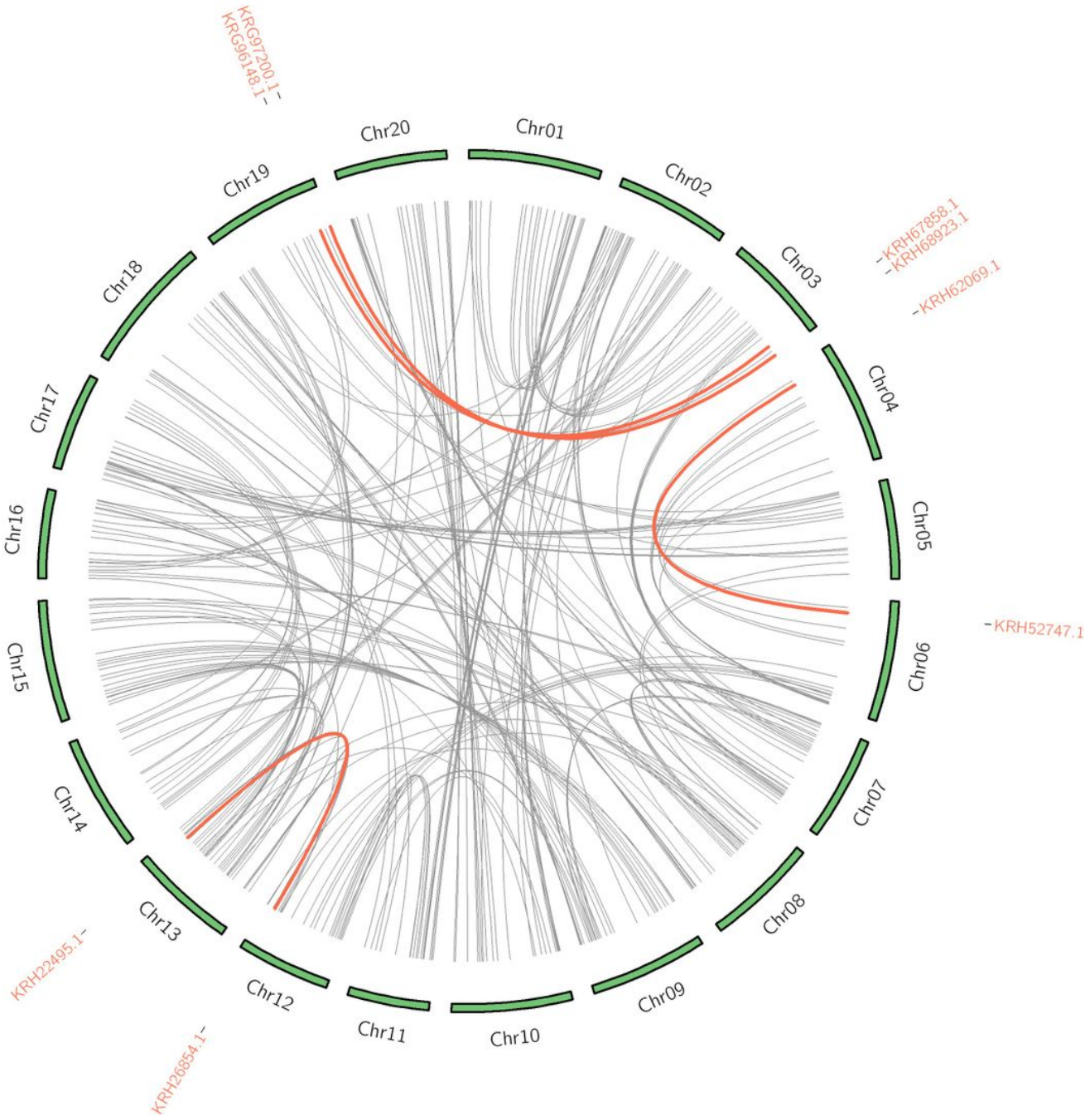

Figure 2

Synteny analysis of DREB. Chromosomes of soybean are shown in a circular form. Approximate positions of ordinary genes are labeled with a short gray line on the circle. Collinear DREB genes(listed by protein ID) is shown with red curves.

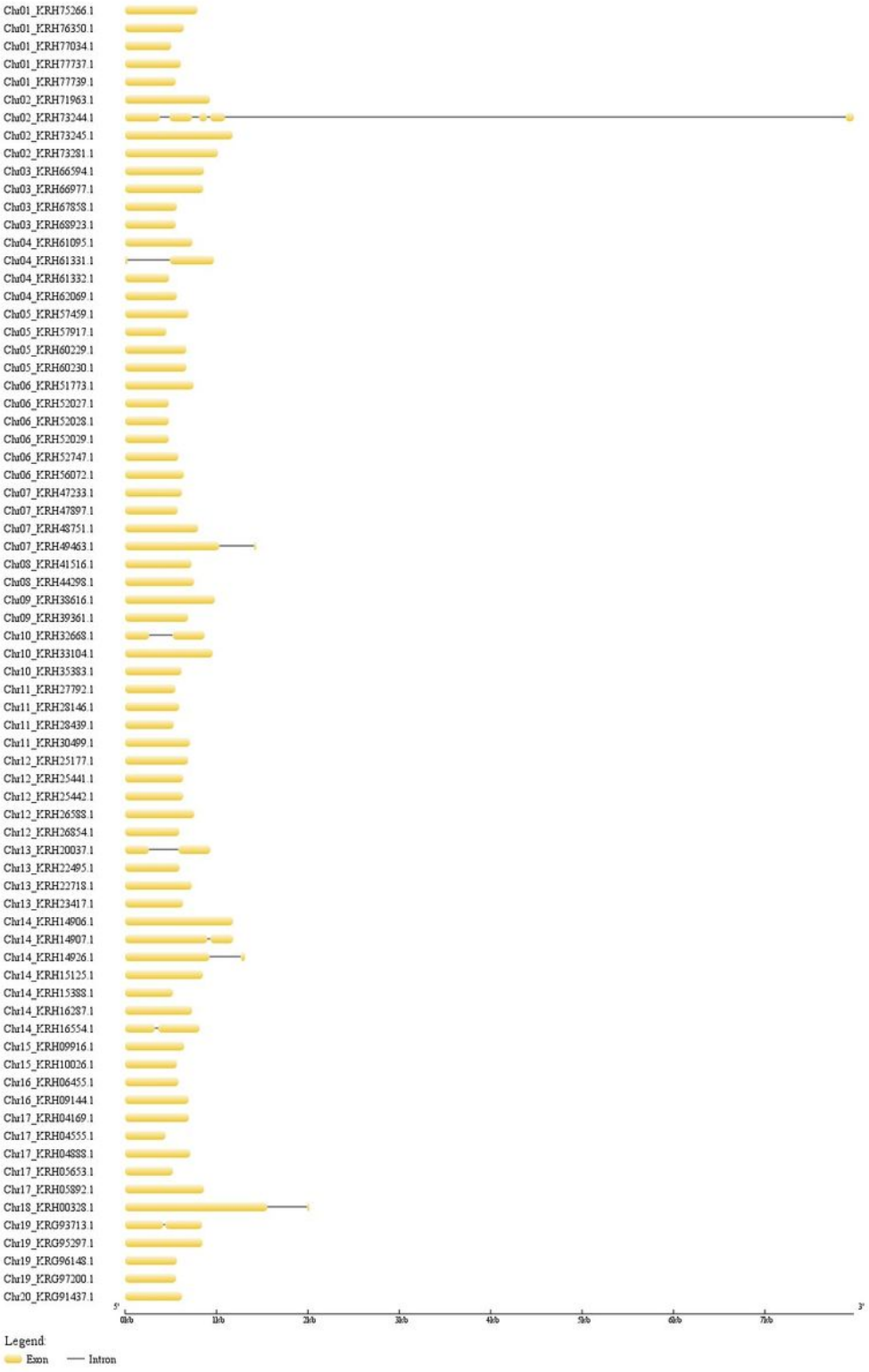

Figure 3

DREB gene structure and location on the chromosome

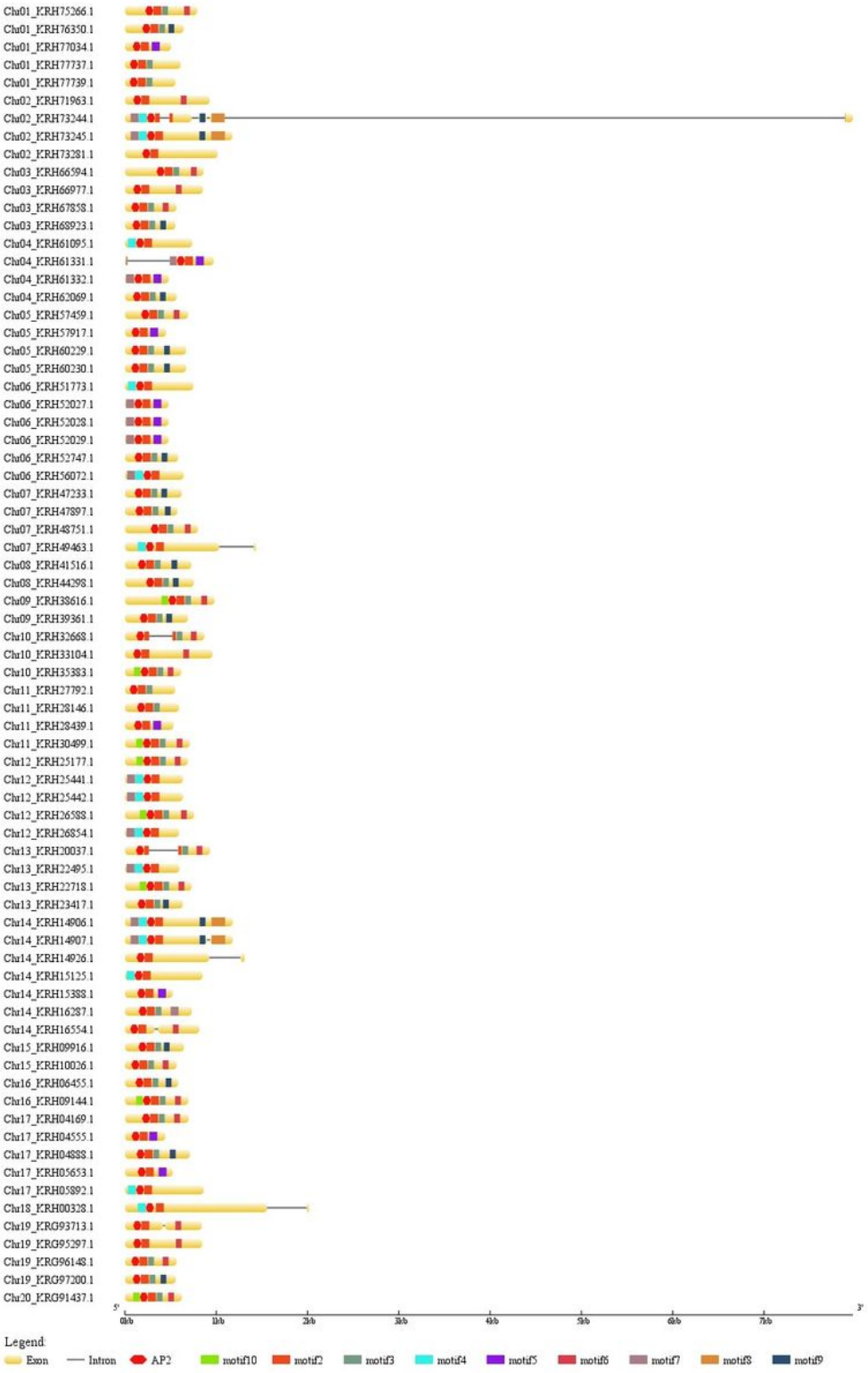

Figure 4

DREB motif distribution

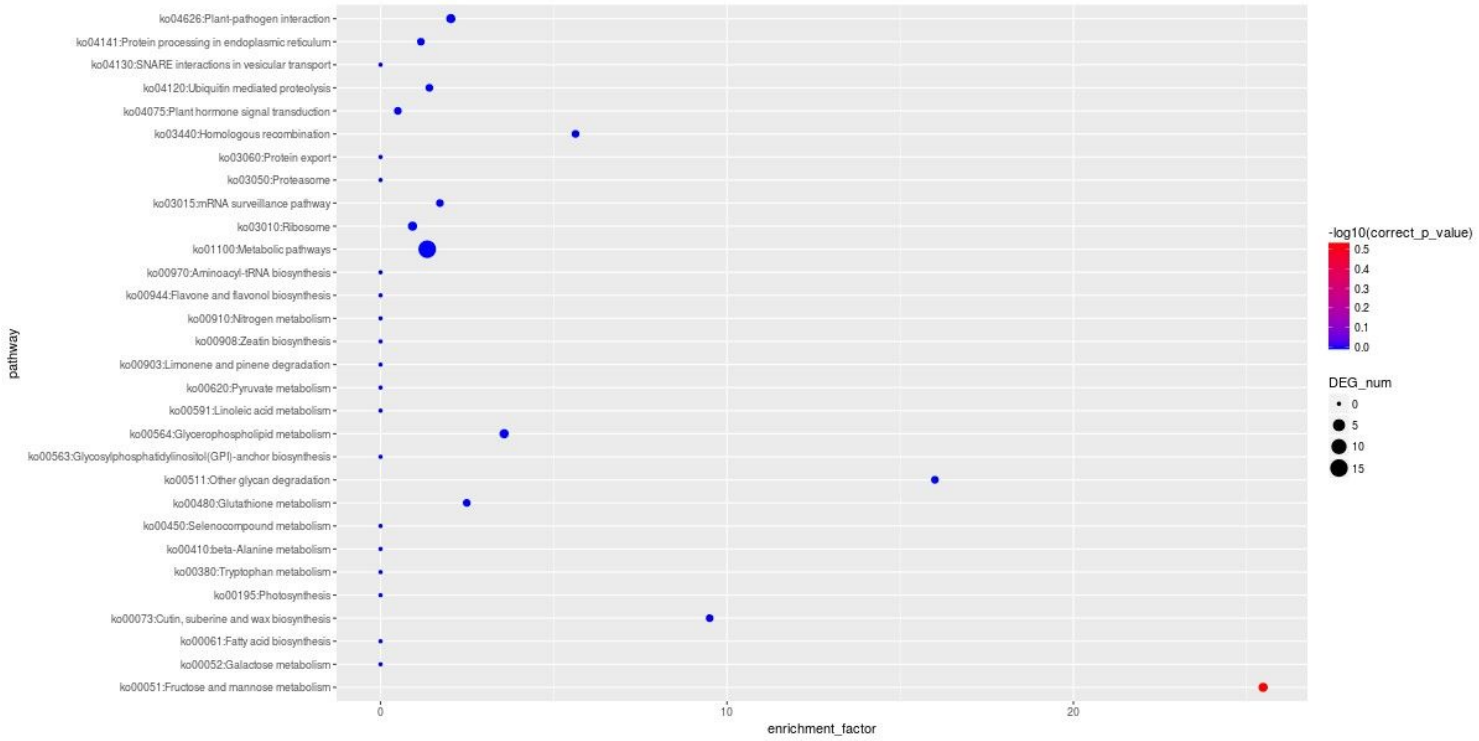

Figure 5

The picture shows the bubble map of the KEGG enrichment result. The ordinate is the name of the keg metabolic pathway, the abscissa is the enrichment factor, and the size of the point represents the number of target genes distributed in the pathway. The color is from red to blue, representing the FDR value from 0 to 1.

Figure 6

Histogram of differentially expressed genes. ZT4, ZT8, and ZT12 represent soybean leaves after 4 hours, 8 hours, and 12 hours of drought. ZT4-control, ZT8-control, and ZT12-control are corresponding control groups, respectively. Red represent up -regulated genes. Green represent down-regulated genes.

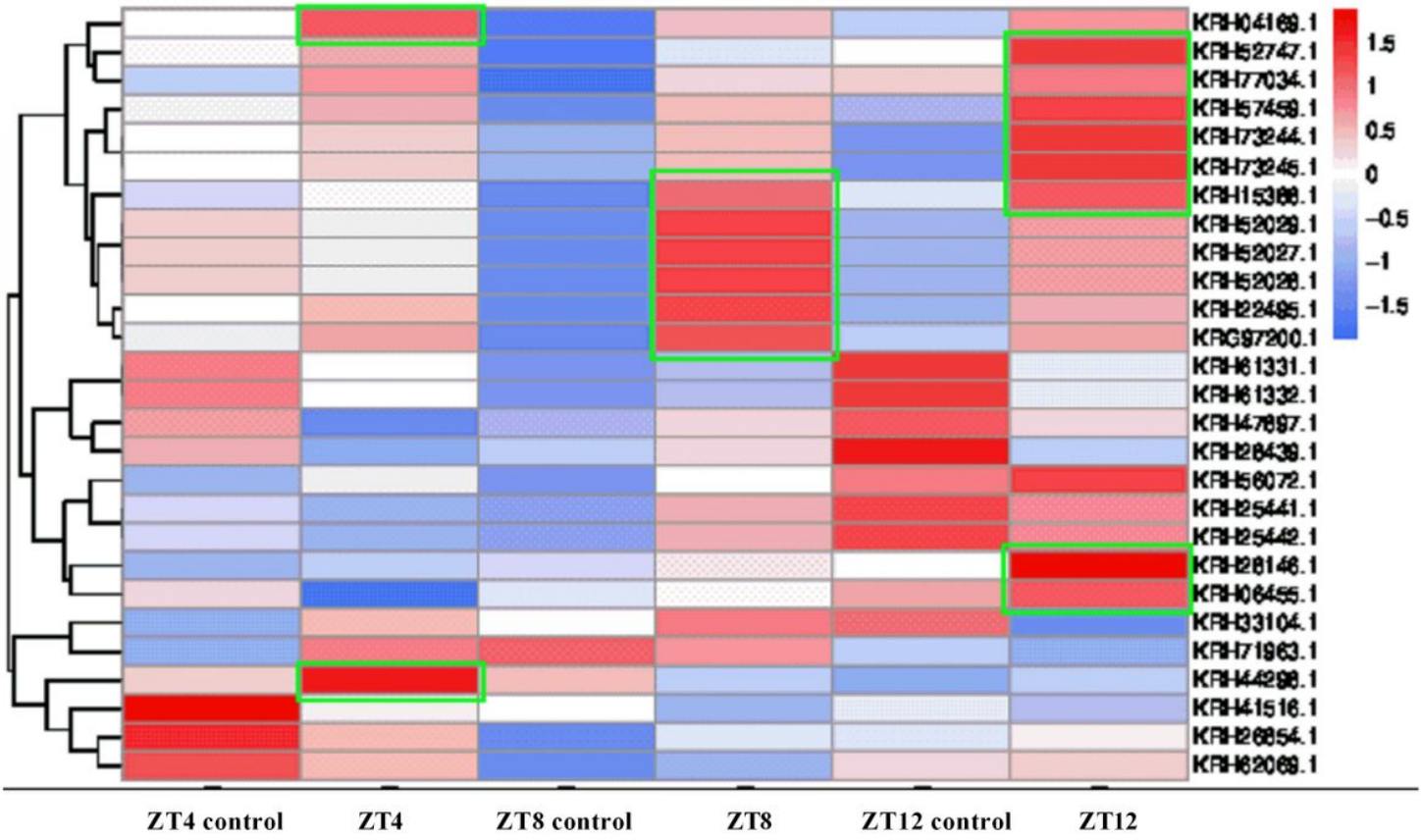

Figure 7

Expression profiles of DREBs in response to drought stress in leaves of ZT4, ZT8, ZT12 and corresponding control. Log2 based fold change was used to create the heat map. Fold changes in gene expression are shown in color as the scale.

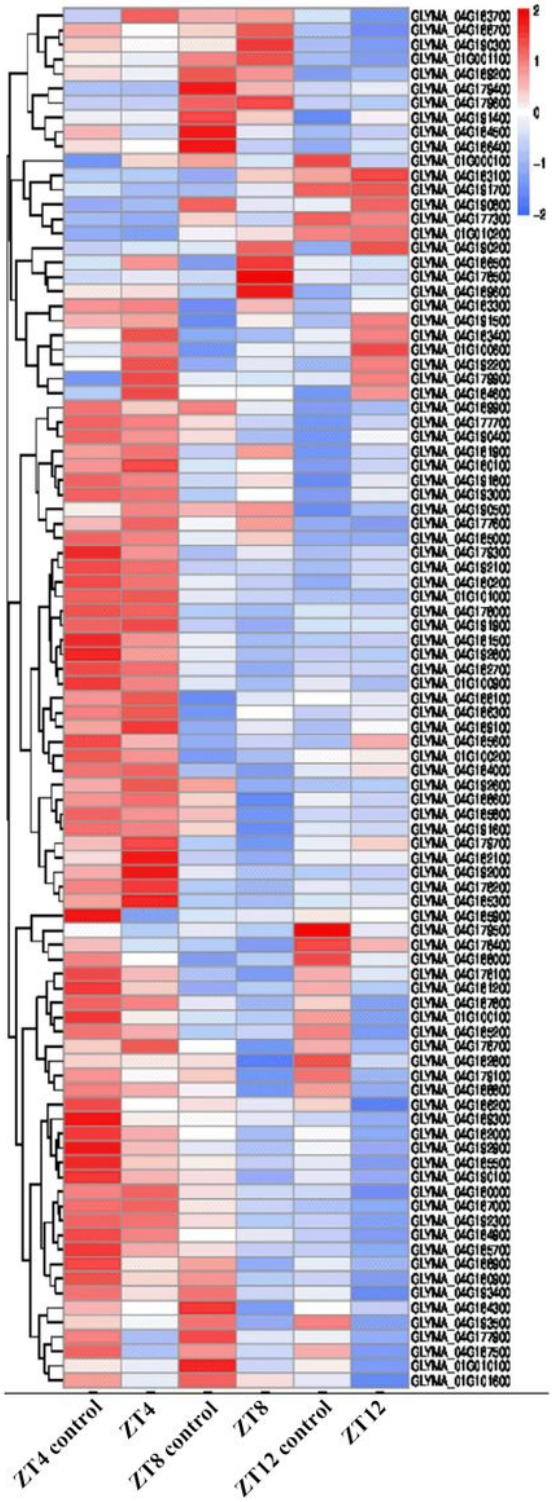

Figure 8

Expression profiles of DREB target gene in response to drought stress in leaves of ZT4, ZT8, ZT12 and corresponding control. Log2 based fold change was used to create the heat map. Fold changes in gene expression are shown in color as the scale.

Supplementary Files

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

- SupplementaryTable2.xls
- SupplementaryTable1.xlsx
- SupplementaryTable3.xlsx