

# SCIENTIFIC REPORTS

OPEN

## Expansion and stress responses of AP2/EREBP superfamily in *Brachypodium distachyon*

Lihong Chen<sup>1,\*</sup>, Jiapeng Han<sup>2,\*</sup>, Xiaomin Deng<sup>3,\*</sup>, Shenglong Tan<sup>4</sup>, Lili Li<sup>1</sup>, Lun Li<sup>1</sup>, Junfei Zhou<sup>1</sup>, Hai Peng<sup>1</sup>, Guangxiao Yang<sup>2</sup>, Guangyuan He<sup>2</sup> & Weixiong Zhang<sup>1,5</sup>

Received: 31 July 2015

Accepted: 26 January 2016

Published: 12 February 2016

APETALA2/ethylene-responsive element binding protein (AP2/EREBP) transcription factors constitute one of the largest and most conserved gene families in plant, and play essential roles in growth, development and stress response. Except a few members, the AP2/EREBP family has not been characterized in *Brachypodium distachyon*, a model plant of Poaceae. We performed a genome-wide study of this family in *B. distachyon* by phylogenetic analyses, transactivation assays and transcript profiling. A total of 149 AP2/EREBP genes were identified and divided into four subfamilies, i.e., ERF (ethylene responsive factor), DREB (dehydration responsive element binding gene), RAV (related to ABI3/VP) and AP2. Tandem duplication was a major force in expanding *B. distachyon* AP2/EREBP (BdAP2/EREBP) family. Despite a significant expansion, genomic organizations of BdAP2/EREBPs were monotonous as the majority of them, except those of AP2 subfamily, had no intron. An analysis of transcription activities of several closely related and duplicated BdDREB genes showed their functional divergence and redundancy in evolution. The expression of BdAP2/EREBPs in different tissues and the expression of DREB/ERF subfamilies in *B. distachyon*, wheat and rice under abiotic stresses were investigated by next-generation sequencing and microarray profiling. Our results are valuable for further function analysis of stress tolerant AP2/EREBP genes in *B. distachyon*.

The APETALA2/ethylene-responsive element binding protein (AP2/EREBP) superfamily is one of the largest groups of plant specific transcription factors. They play vital roles in plant growth, development and response to such diverse stresses as extreme temperature (freezing damage, and high temperature stress), drought, high salinity and pathogen infection<sup>1–7</sup>. They have also been implicated in various hormones-related signal transduction pathway including abscisic acid (ABA), ethylene, cytokinin and jasmonates (JAs)<sup>8–12</sup>.

Many AP2/EREBP genes have been identified in diverse plant species, such as *Arabidopsis* (145)<sup>13</sup>, poplar (200)<sup>14</sup>, soybean (98)<sup>15</sup>, grape (149)<sup>16</sup>, tomato (85)<sup>17</sup>, rice (163)<sup>7</sup>, potato (155)<sup>18</sup> and foxtail millet (171)<sup>19</sup>. These genes have been studied in domain conservation, features of amino acid sequence and biological functions. The AP2/EREBP genes share a highly conserved AP2 DNA binding domain with 50 to 70 amino acid residues, which consist of a three-stranded anti-parallel  $\beta$ -sheet and an  $\alpha$ -helix<sup>20</sup>. This AP2 domain was first discovered in the amino acid sequence of a homeotic gene *APETALA2* (*AP2*) when studying *Arabidopsis* flower and seed development<sup>21</sup>. The AP2/EREBP genes are divided into three families based on the number of AP2 domain and sequence similarity, i.e., the AP2 family, the RAV (related to ABI3/VP1) family and the ERF (ethylene response factor) family. The AP2 family, containing two AP2 domains and essential for plant development<sup>22,23</sup>, is further categorized into the AP2 and ANT (AINTEGUMENTA) groups in seed plants. The RAV family, consisting of a AP2 domain and a B3 DNA binding domain, functions as negative regulators in plant development and is important

<sup>1</sup>The Institute for Systems Biology, Jiangnan University, Wuhan 430056, China. <sup>2</sup>The Genetic Engineering International Cooperation Base of Chinese Ministry of Science and Technology, Chinese National Center of Plant Gene Research (Wuhan) HUST Part, College of Life Science and Technology, Huazhong University of Science & Technology (HUST), Wuhan 430074, China. <sup>3</sup>Ministry of Agriculture Key Laboratory of Biology and Genetic Resources of Rubber Tree/State Key Laboratory Breeding Base of Cultivation and Physiology for Tropical Crops, Rubber Research Institute, Chinese Academy of Tropical Agricultural Sciences, Danzhou 571737, China. <sup>4</sup>School of Information Engineering, Hubei University of Economics, Wuhan 430205, China. <sup>5</sup>Department of Computer Science and Engineering and Department of Genetics, Washington University, St. Louis, MO 36130, USA. \*These authors contributed equally to this work. Correspondence and requests for materials should be addressed to G.H. (email: hegy@mail.hust.edu.cn) or W.Z. (email: weixiong.zhang@wustl.edu)

for mediating abiotic and biotic stress responses<sup>24–27</sup>. The ERF family, hosting a single AP2 domain and being the largest in the AP2/EREBP superfamily, has been revealed to be critical regulators of diverse processes in plant. Based on the amino acid sequence of the AP2 domain, the ERF family can be further divided into the dehydration-responsive element-binding protein (DREB) and ERF subfamilies, or the groups A and B, each of which consists of six groups individually. Nevertheless, based on a phylogeny, the ERF family has been divided into 12 and 15 groups in *Arabidopsis* and rice, which is different from DREB and ERF, two major subfamilies used before<sup>28</sup>.

Much effort has been devoted to elucidation of the functions of the DREB and ERF subfamily. The DREB genes activate multiple stress-responsive genes by interacting with the C-repeat/dehydration responsive element (CRT/DRE), which has a core motif of A/GCCGAC<sup>29,30</sup>, in the promoters of target genes. They regulate stress-responsive genes under various abiotic stresses, including low temperature<sup>31–33</sup>, heat<sup>34,35</sup>, drought<sup>5,6,36,37</sup> and high salinity<sup>36</sup>. Thus they are excellent candidates for genetic breeding to enhance crop stress tolerance. The genes of the ERF subfamily, which play more diverse functions than the genes of the DREB subfamily, have been shown to induce or repress gene expression in response to external stimuli of ethylene, cytokinin and abiotic stresses, such as wounding, cold, high salinity, and drought<sup>9,38–40</sup> by recognizing the GCC-box<sup>20</sup>.

Coupled with a small genome, short life cycle, and easy cultivation and transformation, *B. distachyon*, becomes an excellent model plant for functional genomics of temperate cereals and forage grasses<sup>41</sup> and for developing sustainable energy and feed stocks<sup>42,43</sup>. With the genome of the diploid inbred line Bd21 being sequenced<sup>44–46</sup>, several gene families of *B. distachyon*, including *WRKY*<sup>47–49</sup>, *MAPK* and *MAPKK*<sup>50,51</sup>, *bZIP*<sup>52</sup>, and *NAC*<sup>53</sup>, have been studied. However, the AP2/EREBP superfamily remains largely unexplored in *B. distachyon*, although the C-repeat binding factors (CBFs or DREBs) have been identified and investigated<sup>54,55</sup> in *B. distachyon* and other monocotyledon plants<sup>56–61</sup>. Genomic comparison, an effective means to transferring knowledge of one taxon to another<sup>62</sup>, could be adopted to understand the expansion and functions of this superfamily of genes among *Brachypodium*, rice, wheat, and sorghum after divergence during evolution.

In this study, we systematically investigated the AP2/EREBP superfamily in *B. distachyon* by adopting and combining comparative and phylogenetic analyses, transactivation assay, and gene expression profiling. We report here our findings on expansion and conservation of this large and critical set of transcription factors in *B. distachyon* and their functions as gene regulators in response to abiotic stresses.

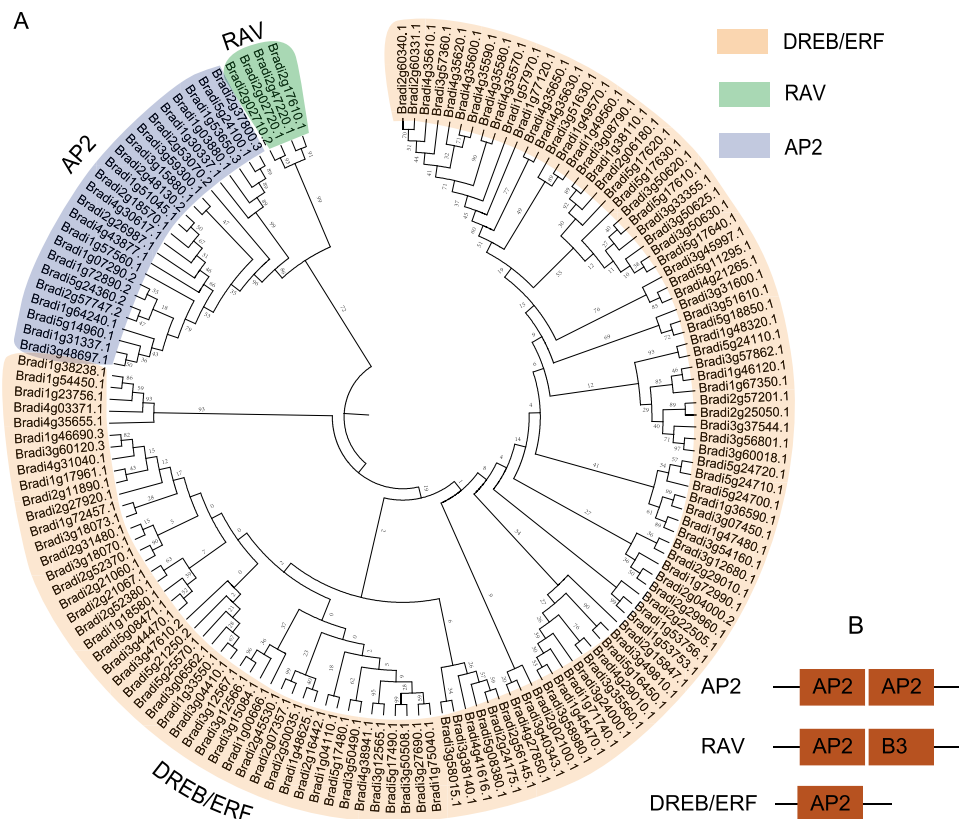
## Results and Discussion

**Identification of the AP2/EREBP superfamily in *B. distachyon*.** Based on a stringent HMM and using restrictive BLAST searches (see Methods), 189 putative AP2/EREBP proteins were identified in the *B. distachyon* reference genome (version 3.0) and represent 0.44% of the annotated proteins in *B. distachyon*<sup>19,44</sup>. This ratio is much closer to that of *Setaria italica* (0.44%)<sup>19</sup> and rice<sup>7</sup> (0.43%), but less than that of *Arabidopsis* (0.55%)<sup>19,28,63</sup>. Among the 189 putative AP2/EREBP proteins, 13 were splicing variants and thus discarded from further consideration, and 17 additional proteins (Table S1) were also excluded from further analysis because their AP2 domains were too short or divergent to fit into an acceptable phylogeny. The remaining 149 putative BdAP2/EREBP genes (Table S1) were used for further analysis.

Based on sequence similarities, the composition of domains, and the number of the AP2 domain, the 149 BdAP2/EREBP genes could be divided into three subfamilies. Specifically, 23 genes encoded two complete AP2 domains and were thus assigned to the AP2 subfamily; 4 genes had a single AP2 domain and a single B3 domain and were classified into the RAV subfamily; and the remaining 122 genes carried a single AP2/ERF domain and were assigned to the DREB/ERF subfamilies. The DREB/ERF subfamily members were further divided into DREB and ERF subfamilies based on the phylogeny in the following analysis. Since no standard nomenclature has been adopted for the AP2/EREBP family genes, each subfamily gene was named with a two-letter code corresponding to *B. distachyon* (Bd), followed by a subfamily designation (AP2, RAV, DREB or ERF), and a number based on the order of its chromosomal locations. This scheme followed the naming convention suggested for the AP2/EREBP family genes in *Setaria italica*<sup>19</sup>. Table S1 lists the detailed information of these *B. distachyon* AP2/EREBP genes, including accession numbers, number of amino acids, and expression sequence tag (ESTs).

**Phylogeny of the AP2/EREBP genes in *B. distachyon*.** In order to appreciate the phylogenetic relationship of the *BdAP2/EREBP* genes, a phylogenetic tree was constructed using the 149 AP2/EREBP protein sequences. The 149 genes clustered into distinct RAV, AP2 and DREB/ERF clades comprising of 4, 23 and 122 proteins, respectively (Fig. 1). The result was consistent with the above classification result based on domain compositions and the number of AP2 domains. The number of *B. distachyon* RAV subfamily genes was nearly the same as that in *Setaria italica*<sup>1</sup>, *rice*<sup>7</sup>, *Arabidopsis*<sup>13</sup>, *Populus trichocarpa*<sup>14</sup>, *Vitis vinifera*<sup>16</sup> and *Salix arbutifolia*<sup>64</sup>, showing a strong conservation of this subfamily and suggesting these RAV genes in all of these plants may share a common ancestor prior to the separation of *Brachypodium* from the other plants. However, the number of DREB/ERF subfamily members varied significantly in these plants, indicating that they might have independently undergone different expansions in evolution.

The AP2 domain sequences of the 122 ERF/DREBs and 25 published ERF/DREBs<sup>7,28</sup> were used to infer the phylogenetic history of the ERF/DREB genes. As shown in the phylogenetic tree in Fig. 2, the 122 BdERF/DREBs were further classified into two subfamilies based on the amino acid sequences of the AP2 domains: the DREB subfamily has 65 (53% of the 122) genes encoding DREB/CBF-like proteins and the ERF subfamily contains 57 (47%) genes encoding ERF-like proteins. Moreover, the DREB subfamily was further divided into six groups as in *Arabidopsis*<sup>28</sup>, among which the A1 group was the largest and the A2 group the smallest (Fig. 2). Likewise, the ERF subfamily was classified into seven groups as in rice<sup>7</sup>, with the B1 subgroup being the largest (Fig. 2). Interestingly, the ratio of the DREB subfamily members in the 122 BdERF/DREBs is higher than that of the ERF subfamily (Table S1). This case indicated each subfamily of AP2/EREBPs evolved with species-specific characteristics.

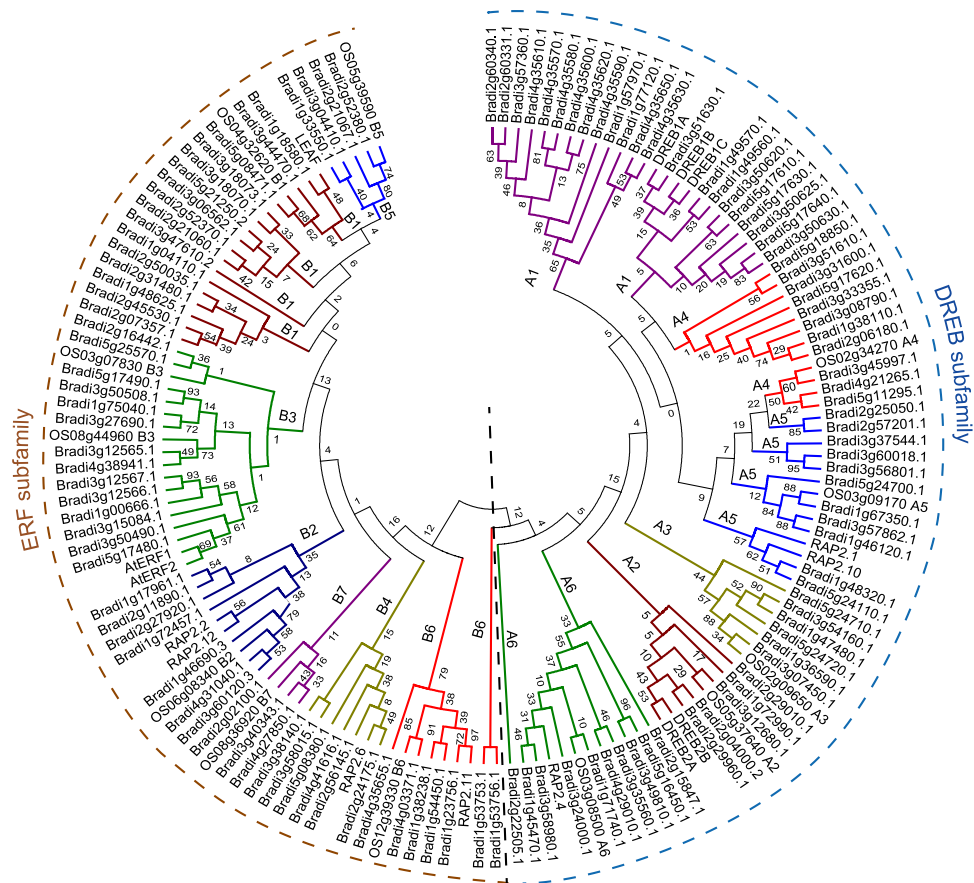


**Figure 1.** (A) Phylogeny of the *B. distachyon* AP2/ERE/EBP (APETALA2/ethylene-responsive element binding protein) family. (B) Domain structures of the AP2/ERE/EBP genes.

The DREB subfamily members (Fig. 2) included or overlapped with the BdCBF members reported in previous study<sup>54,55</sup>. In particular, all the 19 DREB/CBFs identified by Ryu *et al.*<sup>55</sup> fell into the A1 subgroup (Table S2), while six members (Bradi1g49560, Bradi1g49570, Bradi3g50630, Bradi5g17610, Bradi5g17630 and Bradi5g17640) were missed in the study by Li *et al.*<sup>54</sup>. Among the 18 DREB/CBFs reported in Li *et al.*<sup>54</sup>, 14 fell into A1 subgroup (Table S2); Bradi4g35640 which was not identified in our study nor in Ryu *et al.*<sup>55</sup>, had no obvious AP2 domain, two other CBFs, Bradi3g45997 and Bradi4g21265 in the A4 subgroup (Table S2), were missed by Ryu *et al.*<sup>55</sup>; and Bradi3g45997 was counted twice. These results suggested that our work was rigorous for identification of AP2/ERE/EBP genes in *B. distachyon*.

**Conservation and divergence of DREB/ERF subfamily genes during monocot evolution.** DREB/ERF subfamily genes are important regulators of stress response in plants. The size of this subfamily varies greatly across different species relative to the AP2 and RAV subfamilies (Table S3). To understand the expansion and variation of DREB/ERF subfamily in Poaceae, the phylogenetic tree of each group of the DREB/ERF subfamily in *B. distachyon* and several other plants of Poaceae was constructed based on their conserved AP2 domains. The Poaceae evolved into several subfamilies 55–70 million years ago (MYA)<sup>57</sup>. Compared to members of the Pooideae including the temperate cereals wheat, barley and oat, the subfamilies of Panicoideae (maize, sorghum) and Oryzaceae (rice) are preferentially distributed in tropical regions<sup>57</sup>. The AP2/ERE/EBP superfamily in rice<sup>7</sup>, sorghum<sup>59</sup> and wheat<sup>60</sup> has been identified recently. Therefore, the three species were used as representatives of the Oryzaceae, Panicoideae and Pooideae subfamilies, respectively, for phylogenetic analysis of the DREB/ERF subfamilies.

The genes with substantially divergent sequences were discarded in constructing the phylogenetic tree of each group of DREB/ERF subfamilies. As shown in Figure S1, the A1 group (CBF) of DREB subfamily could be classified into four subgroups (I, II, III and IV) as similarly done for wheat<sup>57</sup>. Subgroup I contains only one gene (*Ta.47061*, also named as *CBFlA-A11*) in wheat, two in rice, and four in *B. distachyon*, but none in sorghum, suggesting that CBFIA homologs in *B. distachyon* and rice may have expanded following divergence from wheat. The second subgroup (CBFII) contains three genes in wheat (Pooideae), one each in sorghum (Panicoideae) and *B. distachyon* (Pooideae), but none in rice (Oryzaceae). Such differences indicated that the subgroup II CBF genes of rice may have separated earlier from other subfamilies of the Poaceae. Within the four subgroups of the DREB A1 group, the third subgroup is the largest and can be further organized in four distinct branches of the phylogenetic tree, named as IIIa, IIIb, IIIc and IIId following an early report<sup>57</sup>. Group IIIa is the only one containing CBF homologs from the Oryzaceae, Panicoideae and Pooideae subfamilies. This result indicated that the CBFIIIa genes were already present before the separation of these subfamilies. Groups IIIb and IIIc were grouped together in the phylogenetic tree (Figure S1) and contained CBF homologs exclusively from the two subfamilies

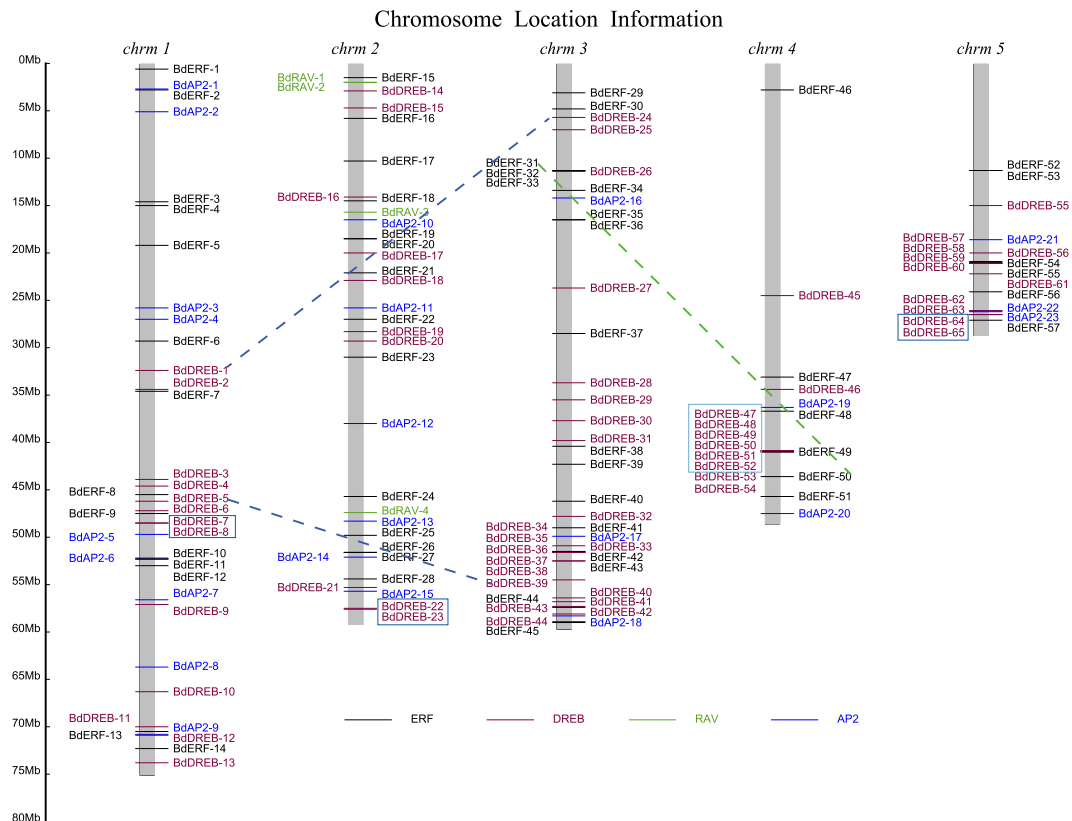


**Figure 2. Phylogeny of the *B. distachyon* DREB (dehydration responsive element binding gene) and ERF (ethylene responsive factor) subfamily proteins based on their conserved domain sequences. Different subgroups of DREB and ERF subfamily proteins are highlighted in different colors.**

of Oryzaceae and Pooideae. Interestingly, group IIIId contained CBF genes only from *B. distachyon* (11 BdCBFs, 7 were originated from gene duplication) and wheat (6 TaCBFs, 4 were originated from gene duplication) of Pooideae subfamily and the CBF genes from each of the two species clustered together, respectively. This result revealed an obvious monophyletic origin of the CBFIIIId genes and indicated that this group evolved following the appearance of the Pooideae. Compared with subgroup III, the CBF subgroup IV (CBFIV) had fewer members. It formed two clades; the big one included CBF genes exclusively from wheat of Pooideae subfamily, and the small clade contained CBF genes from rice of Oryzaceae and *B. distachyon* of Pooideae subfamilies. However, none of the CBFIV genes was found in the Panicoideae subfamily, as similarly described in wheat<sup>57</sup>. This result indicated that these CBFIV genes may evolve to serve some special function in the three species of wheat, rice and *B. distachyon* and evolve much faster in wheat than in other two species after their separation. For some other CBF genes, such as *Os09g35010*, that do not cluster together with any of the above described subgroups (Figure S1), we did not force to classify them, since they may have evolved with species specific characteristics or they may be more adequately grouped with some genes in other subfamilies of Poaceae to better reflect their evolutionary relationships. The other DREB/ERF subgroups were processed similarly. In general, the CBF group has more genes in wheat and *B. distachyon* than in the two other Poaceae species, and its subgroups IIIId and IV have expanded significantly (Figure S1), which might contributed to their specific function against stresses such as cold.

Beyond wheat and *B. distachyon*, the DREB family seemed to evolve differently in the two other Poaceae plants we examined. DREBs in sorghum expanded mainly in the subgroup II of the A2 group and the subgroup IV of the A5 group, which formed an independent branch in the phylogenetic tree (Figure S2). In comparison, DREBs in rice expanded primarily in the subgroup II of the A4 group and the subgroup I of the A5 group (Figure S2). The A3 and A4 group do not contain DREBs in wheat, while only two subgroups (subgroup III of A1 and subgroup I of A6) contain the DREBs in the four species we studied. These results showed the specificity and conservation of DREB subfamily genes after the separation of the four species.

Similar analysis was conducted on each subgroup of ERF subfamily. As shown in Figure S3, B1 group had the largest number, followed by B2 and B3 groups. Each group of ERF subfamily genes could form several hybrid or monophyletic clades. For example, some ERF genes from rice (subgroup II of B1) or sorghum (subgroup VI of B1 and subgroup IV of B4) developed into an independent branch as observed in the DREB genes. All the subgroups of B2 group contain the ERF homologs from the four species of wheat, *B. distachyon*, rice and sorghum, while the DREB subgroups have only two like this, which showed that these genes may have a common ancestor prior to



**Figure 3. Chromosomal distributions of the *B. distachyon* AP2/EREBP genes.** *chr*m is an abbreviation for chromosome. Tandem duplicated gene clusters were highlighted in boxes. The dotted lines, connecting the AP2/EREBP genes in different chromosomes, show the chromosomal segment duplication events. The genes in different subfamilies were color coded. AP2/EREBP is shorted for APÉTALA2/ethylene-responsive element binding protein.

the separation of the four species. The genes forming a monophyletic clade may evolve fast and specifically into their own subfamily.

In general, the DREB/ERF subfamily genes evolved into distinct lineages and expanded with their own specificity after the separation of the four monocot plants, showing their conservation and divergence during evolution and in the process of adaptation to the environment. These results would be helpful for understanding the expansion mechanism of this DREB/ERF subfamily in the monocot plants.

**Tandem and segmental duplications contributed to BdAP2/EREBP expansion.** The 149 BdAP2/EREBP genes scattered unevenly over the 5 chromosomes of *B. distachyon* (Fig. 3). A large portion (74.5%) of the 149 genes was located on chromosomes 1, 2 and 3. The majority of the BdAP2/EREBP genes on chromosomes 4 and 5 congregated closely to one end of the chromosomes. Interestingly, all of the BdRAV genes resided on chromosome 2. Such uneven distribution of these genes provided a clue to their evolution.

In plants, tandem and segmental duplications are known to be major forces for expansion of gene families<sup>65</sup>. Adopting the method and criteria we used previously<sup>50</sup>, we analyzed possible tandem and segmental duplication events in the *B. distachyon* AP2/EREBP family. We found that 12 BdDREB genes belonging to four gene clusters (genes in 4 boxes in Fig. 3, Table S4) were involved in tandem duplication. Two pairs of BdERF genes (*Bradi3g12565/Bradi4g38941* and *Bradi2g21060/Bradi2g52370*) and two pairs of BdDREBS genes (*Bradi1g47480/Bradi3g54160* and *Bradi1g36590/Bradi3g07450*) were implicated in segmental duplication. The largest gene cluster consisted of 6 tandem duplicated genes (the box on chromosome 4 in Fig. 3). Surprisingly, no duplication event was recognizable in the AP2 and RAV subfamilies under the relatively stringent criteria we used.

If the criteria for gene duplication were relaxed, i.e., with a lower similarity ( $\geq 40\%$ ) between amino acids sequences and a shorter sequence matching length ( $\geq 60\%$ ), more AP2/EREBP genes, including all members of the BdRAV subfamily and most genes clustered together in the phylogenetic tree of the BdAP2/EREBP proteins (Fig. 1), could be recognized as involved in segmental or tandem duplication. In other words, this result indicated that tandem and/or segmental duplications had great contributions to the expansion of the BdAP2/EREBP family. *BdDREB-53*, *BdDREB-54* and the 6 tandem duplicated genes could form a super gene cluster. Within this super cluster, *BdDREB-53* and *BdDREB-54* shared approximately 52% and 46% similarity of amino acids sequences with the other 6 tandem duplicated proteins, respectively, which were far below that within the 6 tandem duplicated genes. Interestingly, *BdDREB-51* of this super cluster together with the gene cluster consisting of *BdDREB-22* and *BdDREB-23* (the blue box on chromosome 2, Fig. 3) happened to be involved in segmental

gene duplication (Table S4), which may be the reason that these duplicated genes were originated from a common ancestral chromosomal segment but subsequently diverged under different selective environmental pressures.

**The exon/intron organizations of BdAP2/EREBP genes are relatively simple.** To obtain deep insight into the gene structures of the *B. distachyon* AP2/EREBP genes, their exon/intron organizations were analyzed and ranked based on the phylogenetic tree. All members of BdAP2 subfamily had four to nine introns, while the majority of members of BdRAV, BdDREB and BdERF subfamilies had no intron and the gene structures of them were relatively simple (Figure S4), which was corresponding with that in rice<sup>7</sup>. In particular, only 12 (18% of the total) DREBs and one (25%) RAV gene had introns. The number of introns of the two subfamilies members was small compared with that of BdAP2 subfamily and almost all of them with introns had only one intron except Bradi2g29960 and Bradi2g04000 which had 3 and 2 introns, respectively. For the BdERF subfamily, sixty-one percent of them had no intron and all the other members had only one intron. In addition, the exon/intron structures of 20 genes originated from tandem or segmental duplication and most of Ap2/EREBP members clustering together in the phylogenetic tree tended to share similar structure organizations (Figure S4).

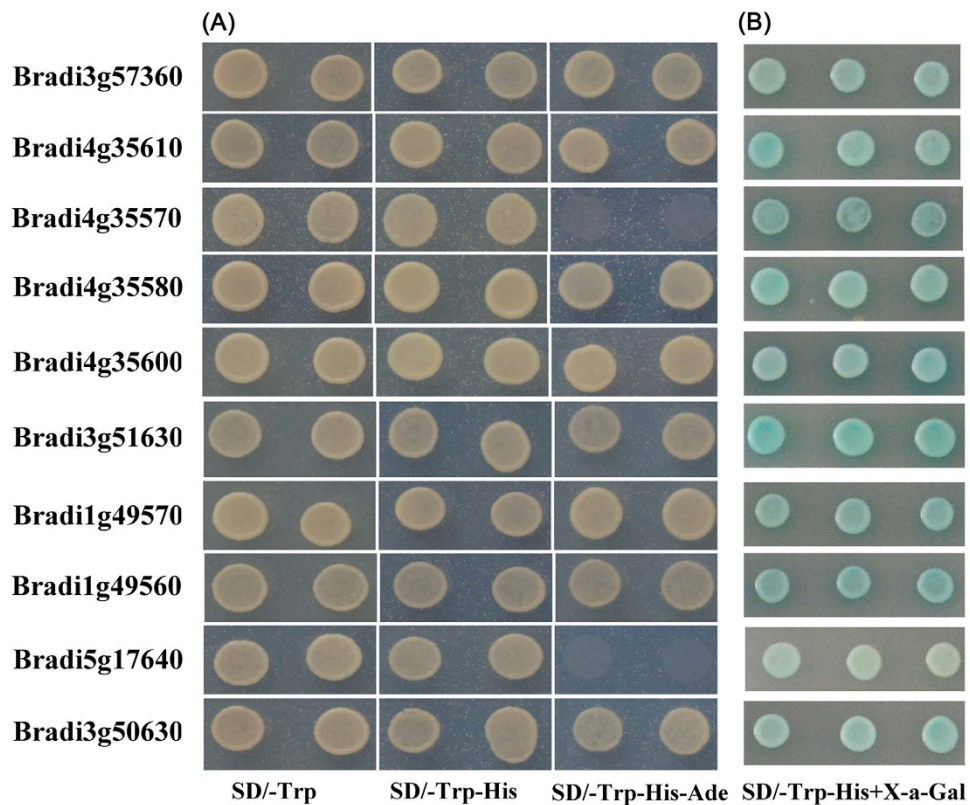
**Synteny of AP2/EREBP genes across *B. distachyon* and other grass species.** A genomic analysis of the synteny of this gene family across several grass plant species could provide insights to its evolution. We comparatively analyzed the synteny across *B. distachyon*, sorghum, maize and rice (Figure S5). The maximum orthology of the *B. distachyon* AP2/EREBP genes was found to be with sorghum (89, 60%), followed by maize (84, 56%) and rice (76, 51%); the result was consistent with the evolutionary relationships among these species. Several BdAP2/EREBP genes were syntenic with the same orthologs of the other grass plants. For example, both orthologs of a pair of duplicated genes (*Bradi4g38941/Bradi3g12565*) correspond to the same gene of sorghum, maize or rice, indicating the two genes may have evolved after *B. distachyon* was separated from the other three grass plants. Interestingly, many BdAP2/EREBP genes are biased towards particular chromosomes or chromosomal regions of sorghum, maize and rice, which may be caused by chromosomal rearrangement events, e.g., genomic duplication and inversion.

**Transactivation activity analyses of BdDREBs.** The DREB proteins are important transcription factors for gene regulation in response to various environmental stimuli and the A1 group members of the DREB subfamily expand greatly in *B. distachyon*. To investigate the functional divergence of these genes, 10 closely related and representative members from the A1 subgroup were chosen for transactivation activity analysis (see Methods). As shown in Fig. 4, eight of the ten members tested grew well on the SD/-Trp-His-Ade medium, while Bradi5g17640 and Bradi4g35570 failed to show sustainable growth, indicating that they might have weak transactivation activities. Furthermore, the transactivation activities of the 10 BdDREBs were further assessed according to the growth status and color changes of yeast cells on the SD-Trp-His-X-gal medium for 3 d. All these proteins on the selective medium containing X-gal appeared blue (Fig. 4B), showing that they have obvious transactivation activities, but the single colony color of Bradi5g17640 was much weaker than the other nine, which was consistent with the result from the experiment of yeast survivability discussed above. Nevertheless, these results should not be interpreted as the members detected having definitive transactivation activities in plant. To address this issue, their transactivation activities should be further confirmed by other biological technologies such as ChIP (Chromatin Immunoprecipitation) or transient expression assays in *Nicotiana benthamiana*<sup>66</sup>.

**Involvement of BdAP2/EREBP genes in *B. distachyon* development.** Profiling gene expression can gain insights into gene functions. We investigated the tissue-specific expressions of the 149 BdAP2/EREBP genes in three tissues, namely roots, stems and leaves, and under four abiotic stress conditions. The expressions of some of these genes varied significantly across the three tissues (Fig. 5, Table S5). For example, nine BdDREB genes such as *Bradi1g49560*, *Bradi1g49570* and *Bradi3g37544* had higher expression in root than in the other two tissues, suggesting they may play vital roles in root development. In addition, the genes in the same subfamily had quite different expression profiles. For example, *Bradi3g04410*, *Bradi2g21060*, *Bradi2g52370*, *Bradi3g06562*, *Bradi5g21250* and *Bradi5g25570* in the ERF subfamily were expressed abundantly, whereas the expressions of several other BdERF genes such as *Bradi3g47610*, *Bradi3g40343* and *Bradi2g16442* were barely detectable in the 3 tissues (Fig. 5, Table S5). Similar observations could be made on the other BdAP2/EREBP subfamily genes. These results indicated that the dynamic expression change of each BdAP2/EREBP gene was correlated with its specific function. In addition, closely related genes generally showed highly similar expression profiles (Table S5), further supporting that they may have similar or overlapping functions.

Moreover, to investigate the reliability of deep sequencing based expression profiling, the expressions of 32 representative members of the BdDREB subfamily, which play import roles in development and in response to external stimuli and expand greatly in *B. distachyon*, were analyzed by RT-PCR. Among the 32 BdDREB genes, twelve were selected from the biggest subgroup A1 and four were chosen from each of the other five subgroups of the DREB subfamily. However, all the members of the A3 subgroup could not be detected by RT-PCR, in agreement with their low abundance detected by RNA-seq profiling. The remaining 28 representatives were profiled by RT-PCR. The result showed that the majority of the 28 BdDREB genes had similar expression profiles in all three tissues using the two profiling methods (Figure S6). However, several genes exhibited different expression patterns between the two methods. For example, *Bradi4g35580*, *Bradi4g35600*, *Bradi1g57970*, *Bradi3g45997* and *Bradi2g25050* were expressed at a low level in some tissues using RT-PCR, but were barely detected by deep sequencing. This discrepancy reflected the difference and biases of the two profiling platforms. Nevertheless, despite such difference in profiling, the results were overall consistent.

**Many BdAP2/EREBP genes were functional under abiotic stresses.** To appreciate the potential functions of the BdAP2/EREBP genes under environmental stresses, we investigated their expressions under cold,



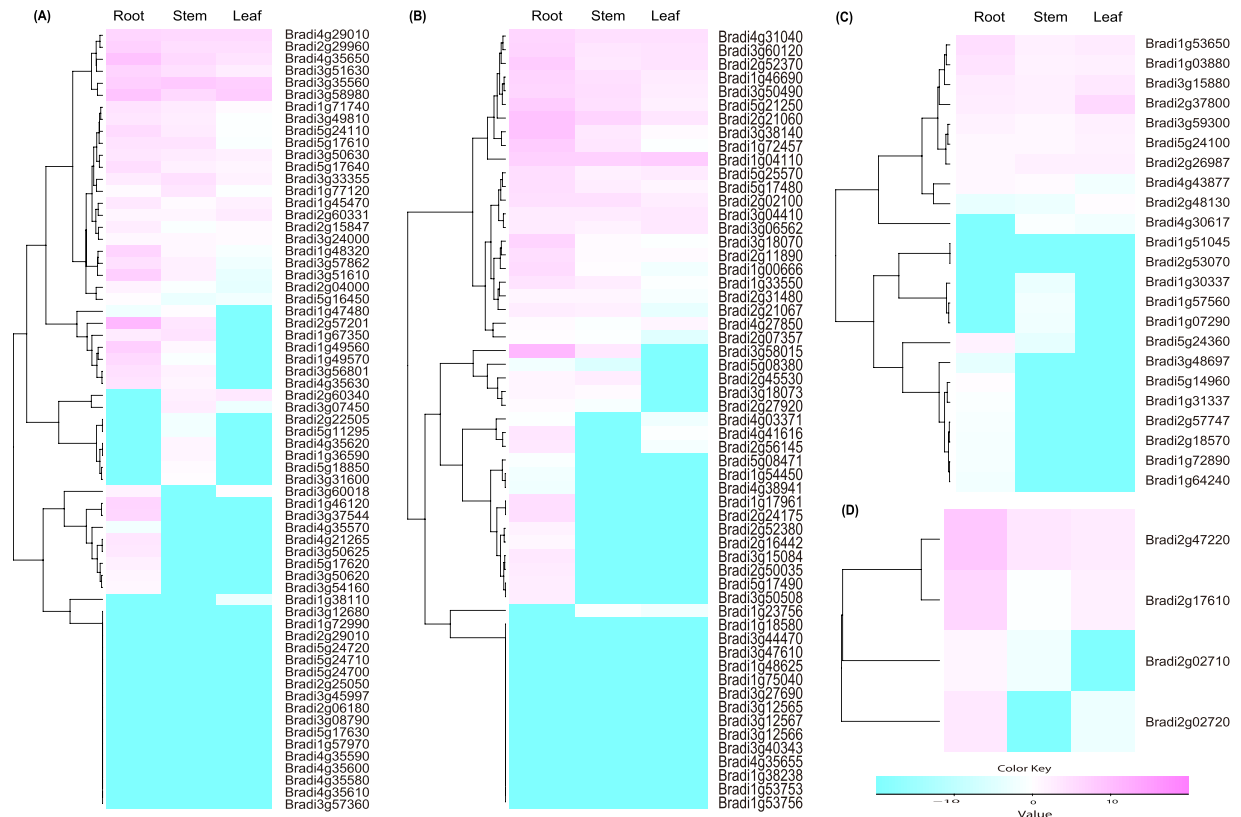
**Figure 4. Transactivation activity assays of 10 *B. distachyon* DREB proteins.** AH109 yeast cells, transformed with plasmid containing different BdDREB genes individually, were cultivated in nutritional selective medium (A) SD/-Trp, SD/-Trp-His, SD/-Trp-His-Ade or (B) SD/-Trp-His-X- $\alpha$ -gal. The empty vector pGBKT7 was negative control. Photos were taken after incubating at 28 °C for three days. DREB is shorted for dehydration responsive element binding gene.

heat, drought and high-salinity conditions. A total of 147 (98.7% of the 149) BdAP2/EREBP genes (56 BdERFs, 64 BdDREBs, 4 BdRAVs and 23 BdAP2s) were subjected to expression analyses under these stresses, as they were probed on the Affymetrix *Brachypodium* Genome Array (BradiAR1b520742 chip, see Methods).

The result revealed that the BdDREB subfamily genes could be classified into 3 subgroups based on their expression profiles (Fig. 6A, Table S6). Except the ones in the middle subgroup, the genes in the other two subgroups were all differentially expressed under at least one of the stresses considered. Among the 64 BdDREBs, 12 might be involved in cold signaling pathways (Fig. 6B), consistent with a previous report<sup>1</sup>. For example, the tandem duplicated genes of *Bradi2g60340*, *Bradi2g60331*, *Bradi4g35570*, *Bradi4g35590*, *Bradi4g35600*, *Bradi4g35610*, *Bradi4g35620* and *Bradi4g35630* in the first subgroup were significantly up-regulated at all the first three time points after cold treatment. *Bradi4g35600*, particularly, exhibited quick cold responses with the highest expression abundance at 1 h point after cold stress, indicating that it may be an early response factor of cold. *Bradi4g35630*, also known as Cold-Binding Factor 1 (BdCBF1), has been shown to be significantly up-regulated under cold stress<sup>55</sup>. Transgenic *Brachypodium* plants overexpressing *Bradi4g35630* also exhibited enhanced resistance to drought and salt stresses as well as low temperatures<sup>55</sup>. These results, plus *Bradi4g35630* being a transcription factor, evidently showed *Bradi4g35630* as a regulator of stress response, particularly of cold response. Moreover, not only a single BdCBF1, i.e., *Bradi4g35630*, more than half of BdCBF genes in the A1 subgroups (Fig. 2) respond to cold stress at different developmental stages of *B. distachyon*<sup>54,67</sup>, showing their critical roles in cold response.

Compared with the differentially expressed genes (DEG) of the BdDREB subfamily induced mainly by cold, the DEGs of the BdERF subfamily were induced mainly by drought and salt stresses (Fig. 6B), suggesting their potential roles in resistance to drought and salt. In particular, the genes (*Bradi2g56145*, *Bradi2g24175*, *Bradi5g21250*, *Bradi2g21060*, *Bradi5g25570*, *Bradi3g18070*, *Bradi4g41616* and *Bradi3g58015*) in the last subgroup of the heatmap were significantly up-regulated during the whole period of drought stress profiled except *Bradi2g56145* at the 1 h point. Interestingly, *Bradi4g41616* and *Bradi3g58015* were also significantly up-regulated at all the time during the salt stress. The two genes may be involved in the same pathway or have redundant or complementary functions in response to the two stresses. For these BdERF genes with differential expressions, they might be selected as potential candidate genes for further functional validation and for utilization in crop improvement plants.

RAV genes have been shown to be participated in plant growth and development<sup>68</sup>, leaf maturation, and senescence<sup>69</sup>, and to control the flowering time under long-day growth conditions<sup>70</sup>. They have also been shown to be responsive to biotic and abiotic stresses<sup>25,27,71–73</sup>, which were supported by our results as well. For example,



**Figure 5.** Expression profiles of BdAP2/EREBP (APETALA2/ethylene-responsive element binding protein) genes in three tissues of roots, stems and leaves. Expression profiles (in log<sub>2</sub> based values) of the BdDREB (A), BdERF (B), BdAP2 (C) and BdRAV (D) genes in three tissues. Bd is shorted for *B. distachyon*. DREB, ERF, AP2 and RAV were shorted for dehydration responsive element binding gene, ethylene responsive factor, APETALA2 and related to ABI3/VP, respectively.

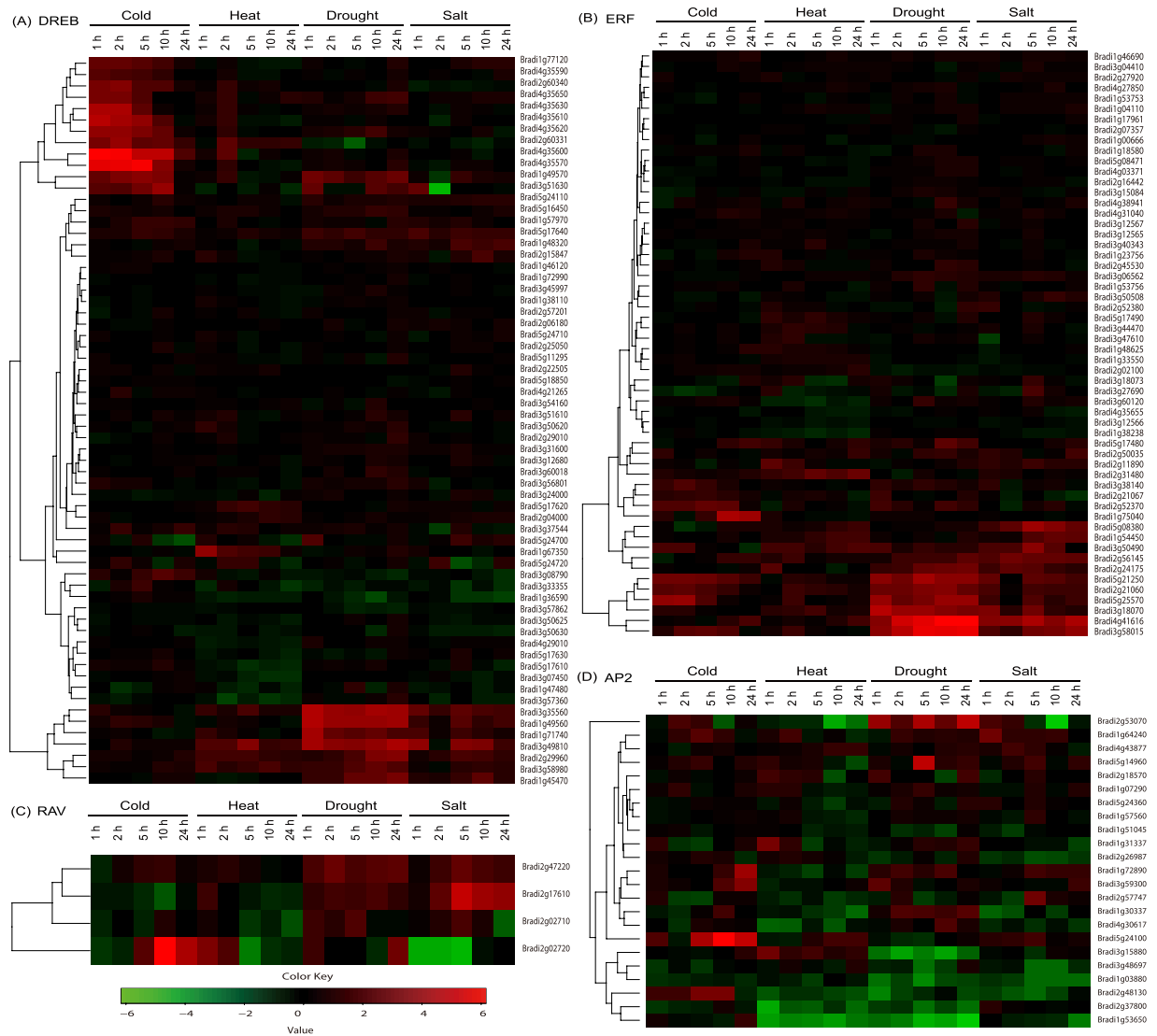
*Bradi2g17610* was significantly induced at 5 h point after salt treatment (Fig. 6C); similar observation can be made on its paralogs *Bradi2g47220* under salt stress condition, indicating that the two genes might be positive regulators of salt stress in *B. distachyon* as their orthologs in *Capsicum annum*<sup>71</sup> and *Arabidopsis*<sup>72</sup>. For the two other BdRAV genes, the expression of *Bradi2g02720* gradually increased over time and reached a peak at 10 h point after cold treatment, and was significantly down-regulated by the first three points of salt treatment. However, the expression of *Bradi2g02710* had little change across the four stress conditions except that at 24 h point after salt treatment. Such diverse responsive behavior of the four BdRAVs suggested their distinct functions and their functions in different signaling pathways.

In contrast to the effort into the three subfamilies discussed above, the work on BdAP2 genes in response to abiotic stress was limited. As was shown in the heatmap (Fig. 6D), the DEGs of the BdAP2 subfamily were also few under the four stress conditions. Several genes were significantly down-regulated by at least one stress. For example, the expression of *Bradi1g53650* decreased at all the time points after heat and drought stresses. However, there was no single gene significantly induced at all the time points after the corresponding stress treatment.

**Comparison of DREB/ERF subfamily gene expressions in three monocot plants.** DREB/ERF subfamilies play important roles in response to abiotic stresses such as cold, drought, salt and heat. The expression profiling of the genes in the two subfamilies in the four representatives (rice, sorghum, wheat (Winter Manitou) and *B. distachyon*) of the Poaceae would help understand the role of these transcription factors in response to abiotic stresses. However, the transcriptome data of sorghum under abiotic stresses were not publicly available. At last three species of wheat, rice and *B. distachyon* were selected for the expression analysis of DREB/ERF subfamily genes.

Expression profiling of wheat DREB/ERF subfamily genes was examined under cold, drought and heat stresses. A total of 20 TaDREBs and 22 TaERFs corresponding to 42 probe sets were identified in the three experiments. Details of the analytic result were provided in Table S7. As shown in Figure S7, the heatmaps of the expressions of both TaDREBs and TaERFs were grouped into three groups. All TaERF genes of group I (Figure S7B) were significantly up-regulated at 42 d point after cold treatment. Three TaDREBs (*Ta.41731*, *Ta.41203* and *Ta.34828*) of DREB group I were significantly up-regulated at least at one time point of cold treatment (Figure S1). The paralogs of the three up-regulated genes, *TaCBF14* and *TaCBF15*, happened to be involved in the cold acclimation process and in frost tolerance<sup>74</sup>. The expression of *Ta.45422* in the group IIIId of A1 group also increased, although less than twofold compared with the control. Interestingly, its orthologs in *B. distachyon* grouped





**Figure 6.** Expression profiles (in log<sub>2</sub> based fold change) of BdDREB (A), BdERF (B), BdRAV (C) and BdAP2 (D) genes under four abiotic stress conditions. Bd is shorted for *B. distachyon*. DREB, ERF, AP2 and RAV were shorted for dehydration responsive element binding gene, ethylene responsive factor, APETALA2 and related to ABI3/VP, respectively.

together in the subgroup IIIId (Figure S1) were all up-regulated significantly at least at three points after cold stress; while the expression of its paralogs in wheat were not profiled since the probe sets of these genes were not on the microarray used. These results suggested functional conservation and divergence of DREBs during the Pooideae evolution. In addition, the transcripts of three DREBs (*Ta.13471*, *Ta.920* and *Ta.57800*) and five TaERFs of group III (*Ta.17378*, *Ta.33090*, *Ta.41713*, *Ta.14000* and *Ta.41713*) (Figure S7) significantly declined at least at one time point of cold treatment, especially *Ta.13471*. However, none of TaDREB/ERF genes have significant differential expression under drought stress. There are only two TaDREBs (*Ta.49353* and *Ta.5364*) and two TaERFs (*Ta.33090* and *Ta.35532*) with obvious change in the expression under the heat stress. Generally, wheat had more differentially expressed genes under cold stress than under heat and drought stresses, which may be correlated with that winter wheat need to develop more DREBs to survive the cold winter.

Similarly, the two subfamily genes were also investigated in rice using the available microarray gene expression data under cold, drought and salt stress conditions. Thirty OsDREB/ERF genes (22.4% of 134 OsEREB/ERF members) including 12 OsDREBs and 18 OsERFs were significantly induced under drought stress (Figure S8 and Table S8), while only 3 OsDREB and one OsERF genes were repressed in this stress condition. Of the 12 induced OsDREB genes, five (*Os06g07030*, *Os04g34970*, *Os02g52670*, *Os08g31580* and *Os03g09170*) happened to be significantly induced by salt. A similar observation was made on OsERF subfamily. Particularly, five OsERF members (*Os01g54890*, *Os02g43790*, *Os05g34730*, *Os11g06770* and *Os05g29810*) were significantly induced under drought and salt stresses, suggesting their vital roles in regulating stress response. In addition, of the 18 OsERF genes induced by drought, 6 were evenly distributed across all subgroups of B1 group (Figure S3) except the subgroup VI, and 7 genes were in the B4 group. Interestingly, four of the 7 induced genes clustered together

with their orthologs in *B. distachyon* in the subgroup I of B4 group (Figure S3) and these orthologs were also significantly induced by drought in *B. distachyon*, while their only one orthologs in wheat did not show similar expression change, indicating the functional conservation and divergence of these orthologs in evolution. Compared with the differentially expressed genes under drought stress, the number (12) of OsDREB/ERF genes with differential expression was relatively small under salt stress and no single OsDREB/ERF gene was significantly down-regulated.

Expression analysis of the DREB/ERF subfamily genes in wheat, rice and *B. distachyon* under cold, high-salinity, drought or heat stresses revealed diverse differential expression of these genes and their possible roles in response to these stresses. Especially in *B. distachyon*, 11 BdDREB genes that clustered together in the subgroup III of A1 group (Figure S1) and were all significantly induced by cold could be further studied and be exploited to enhance stress tolerance. In summary, these results would help expand our knowledge of this important class of transcription factors in plant stress resistance. Some of the results might be utilized to explore the mechanisms of gene regulation under abiotic stresses.

**Functional divergence of duplicated BdAP2/EREBP genes in evolution.** Gene duplication is an essential source of genetic novelty that can lead to evolutionary innovation. Three possible outcomes have been suggested for the evolution of gene duplication: subfunctionalization, neofunctionalization and non-functionalization<sup>75</sup>. Profiling gene expression may provide an ample amount of information of the mode and tempo of duplicated genes. We thus compared the expression patterns of duplicated genes under cold, heat, drought and salt stresses.

Our analysis focused on 19 (95%) of the 20 duplicated genes that were profiled by the microarray gene expression data. Among the 19 genes, a gene cluster consisting of five tandem duplicated genes (the box on chromosome 4, Fig. 3) had highly similar expression patterns (Table S6), indicating subfunctionalization during evolution. Specifically, all the pairs of the five duplicated genes had positive expression correlations, measured by the Pearson correlation coefficients (PCCs), being greater than 0.5, and there was one extreme case of the *Bradi4g35570/Bradi4g35610* pair with the PCC value greater than 0.96 (Fig. 7A). Such high correlations among the five duplicated genes suggested subfunctionalization of the paralogs through division of labor by retaining different parts (subfunctions) of their original ancestral functions through gene duplication.

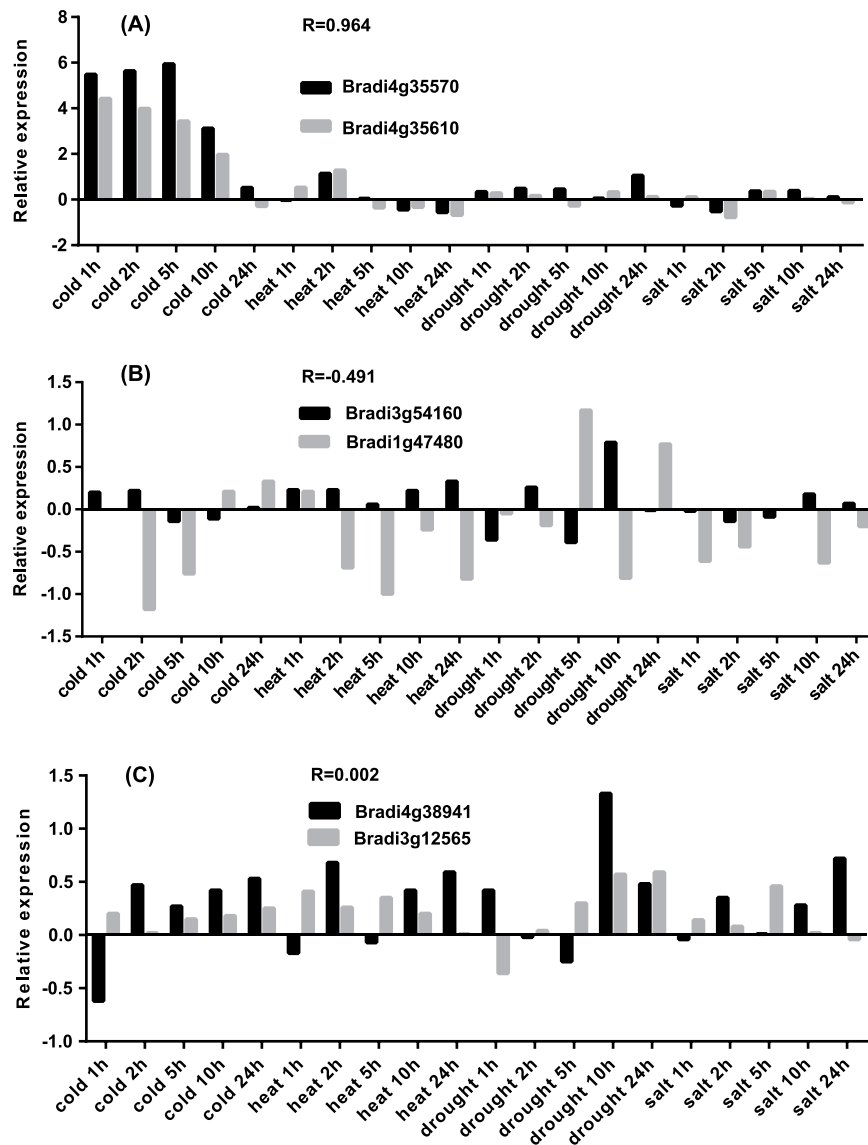
Note that out of the 17 pairs of the duplicated genes, 13 were positively correlated with positive PCC values. However, the most noticeable were the two pairs, *Bradi3g54160/Bradi1g47480* (PCC, -0.491) and *Bradi5g24720/Bradi5g24710* (PCC, -0.156), that had negative PCC values (Fig. 7B). In addition, two pairs of duplicated genes, *Bradi1g36590/Bradi3g07450* (PCC, 0.361) and *Bradi4g38941/Bradi3g12565* (PCC, 0.002), had small PCC values (Fig. 7C), exhibiting little correlation in gene expression across the four abiotic stresses. This implied a neofunctionalization as one copy of the pair acquired, after duplication, a novel and beneficial function that was preserved by natural selection, with the other copy retaining its ancestral function. Moreover, under loose criteria for gene duplication, more BdAP2/EREBP genes would be identified as being produced via gene duplication. For example, under the relaxed criteria of a lower similarity ( $\geq 40\%$ ) between amino acid sequences and a shorter sequence matching length ( $\geq 60\%$ ) as mentioned earlier, *Bradi4g35630* and *Bradi4g35650* would be paralogs. Some of these duplicated genes still had obviously positive correlation in gene expression, although their amino acid sequences had a lower similarity. For example, the expressions of *Bradi4g35630* and *Bradi4g35650* with 47% similarity in amino acid sequences had a PCC of 0.89, indicating the functional conservation of the two genes.

In short, the expression correlation analysis on duplicated genes revealed their functional roles for subfunctionalization and neofunctionalization in the expansion of BdAP2/EREBP superfamily. In other words, the functions of the superfamily genes were expanded and enhanced through gene duplication.

## Materials and Methods

**Database search and sequence retrieval.** The *B. distachyon* genome assembly and annotation version 3.0 (<http://genome.jgi.doe.gov/pages/dynamicOrganismDownload.jsf?organism=Bdistachyon>) were used, and the putative genes of the AP2/EREBP family in *B. distachyon* were extracted following the strategy for *Setaria italica*<sup>19</sup> and *Vitis vinifera*<sup>16</sup>. The AP2/EREBP family gene sequences in *Arabidopsis* and rice were downloaded from the *Arabidopsis* Information Resource (TAIR, <http://www.arabidopsis.org/>, release 10.0) and the rice genome annotation database (<http://rice.plantbiology.msu.edu/>, release 7.0), respectively. Blast search was then performed using sequences of *Arabidopsis* and rice AP2/EREBP genes as queries to search for putative AP2/EREBP genes in *B. distachyon* as described in *Vitis vinifera*<sup>16</sup>. At the same time, we also used the Hidden Markov Model (HMM) profile of the AP2/EREBP domain (PF00847, from the Pfam database <http://pfam.xfam.org/>) to search the *B. distachyon* protein sequences. All hits with expected (E) values less than 1.0<sup>19</sup> were retained and further manually analyzed using InterProScan program (<http://www.ebi.ac.uk/Tools/InterProScan/>). Finally, the resulting putative AP2/EREBP genes identified by the above two strategies were further verified with available ESTs ([http://www.plantgdb.org/download/download.php?dir=/PublicPlantSeq/Dump/B/Brachypodium\\_distachyon/FASTA](http://www.plantgdb.org/download/download.php?dir=/PublicPlantSeq/Dump/B/Brachypodium_distachyon/FASTA)) (Table S1). For phylogenetic analysis of this superfamily in Poaceae, the AP2/EREBP family gene sequences in sorghum were downloaded from Phytozome ([http://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org\\_Sbicolor\\_er](http://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org_Sbicolor_er)). The *T. aestivum* EST sequences were retrieved from the website of the National Center for Biotechnology Information (Ta.seq.all.gz), and then assembled and translated into amino acids as described earlier as for wheat<sup>60</sup>.

**Phylogenetic analysis and sequence alignment.** The identified AP2/EREBP sequences or domain sequences of *B. distachyon* and other plants were aligned against that of *Arabidopsis* and rice using Clustal 2.0<sup>76</sup> with default parameters. Phylogenetic trees were constructed based on the bootstrap neighbor-joining (NJ) method and bootstrap analysis (1,000 replicates) by MEGA<sup>77</sup>.



**Figure 7.** Comparisons of expression profiles of three representative pairs of duplicated AP2/EREBP genes in *B. distachyon* across 20 time points (in hours) after cold, heat, drought and high-salinity stress treatments. Represented in y-axes are the expressions, in log<sub>2</sub> based fold changes of the corresponding treatments. The correlation (R) between two duplicated genes is measured by the Pearson correlation coefficient of the expressions of the pair of genes. AP2/EREBP is shorted for APETALA2/ethylene-responsive element binding protein.

**Chromosomal distribution, gene structure, gene duplication and synteny analysis.** The exon/intron structures and chromosomal distribution of *B. distachyon* AP2/EREBP genes were determined using the Blat search tool<sup>78</sup> and CD-HIT<sup>79</sup>. Gene duplication events were analyzed according to our previous description<sup>50</sup>. Tandem duplicated genes were defined as neighboring homologous AP2/EREBP genes on the same chromosome of *B. distachyon*, with no more than one intervening gene<sup>80</sup>. For synteny analysis of this gene family across sorghum, maize, rice and *Brachypodium*, synteny blocks among them were downloaded from the Plant Genome Duplication Database<sup>81</sup> and those containing BdAP2/EREBP genes were identified. The orthologs relationships among these four plants were displayed using Circos<sup>82</sup>.

**cDNA cloning, plasmid construction and transcriptional activation analysis in yeast cells.** Ten closely related and representative BdDREBs were selected to be amplified by RT-PCR, with primers designed based on the predicted sequences and cloned into pMD18-T vectors for sequencing. They were then fused to the GAL4 DNA binding-domain vector pGBKT7 to construct plasmid pGBKT7-BdDREBs for transactivation assays. All of the primers used for RT-PCR, cDNA cloning and expression studies were listed in Table S9. Transcriptional activation assays were performed in the yeast strain AH109 with *Ade2* and *His* reporter genes. The PCR products of these genes were cloned separately into the pGBKT7 vector. The constructed BdDREB vectors along with the

pGBKT7 plasmid as a negative control were transformed into yeast strain AH109 following the manufacturer's protocol (Clontech, USA). Transformed strains were confirmed by PCR and then plated on selective synthetic dropout (SD) media without tryptophan (SD/-Trp), without tryptophan and histidine (SD/-Trp-His), or without tryptophan, histidine, and adenine (SD/-Trp-His-Ade), to determine their survivability. In addition, the transcription activation of these genes was also evaluated based on growth status and color variation of yeast cells on incubating plates with SD/-Trp-His-X-gal for 3 d. All transcriptional activation assays were performed in triplets.

**Tissue-specific expression profiling using RNA-seq and RT-PCR.** Roots, stems and leaves of over three-week old Bd21 were collected and stored separately for RNA isolation for tissue-specific expression analysis. Ten samples for each tissue randomly selected from ten individual plants were pooled together for RNA extraction. Sequencing using SOLiD™ 5500 next-generation sequencing platform, with main reagents and supplies in SOLiD™ Total RNA-Kit (Life Technologies), and preliminary data acquisition were performed in house. The alignment of RNA-seq reads to *B. distachyon* genome (version 3.0) was performed by TopHat (v2.0.13)<sup>83</sup>. The cuffquant and cuffnorm functions of Cufflinks (v2.2.1)<sup>84</sup> were applied to compute fragments per kilobase of exon per million fragments mapped (RPKM) of each transcript and the cuffdiff function was used to determine the significantly differential expressed genes (q-value  $\leq 0.05$ ). Tissue-specific expression profiles were visualized in heatmaps using R packages (version 2.15.1), based on the log<sub>2</sub>-based RPKM values of genes in different tissues.

For expression analysis of some BdAP2/EREBPs using RT-PCR, total RNA was isolated with Trizol reagent (Invitrogen) following the manufacturer's protocols and purified using an RNeasy Plant Kit (Qiagen). The integrity and purity of RNA samples were checked by gel electrophoresis and OD 260/280 nm absorption ratios and the RNA concentration was quantified using a micro-volume spectrophotometer (Quawell Technology). RT-PCR analysis was performed as described previously<sup>50</sup>.

**Expression analysis of AP2/EREBP genes in *B. distachyon*, rice and wheat under abiotic stresses.** Gene expression profiling of *B. distachyon*, wheat, and rice under drought, high-salinity, cold, or heat stress conditions<sup>67</sup> were downloaded from <ftp://brachypodium.org/brachypodium.org/Stress/> (dataset ID BradiAR1b520742) and Gene Expression Omnibus (dataset ID: GSE23889, GSE70443, GSE60351 and GSE6901). These data from Affymetrix microarray experiments were analyzed following the method previously described<sup>67</sup>. Briefly, the probes were aligned to the *B. distachyon* reference genome (version 3.0), cDNA sequences of rice AP2/EREBP genes or unigenes of wheat by Burrows-Wheeler Aligner (BWA)<sup>85</sup>, and the probes that had no mismatches in the alignment and were uniquely mapped to the exonic region of a unique gene were retained for further probe-set analysis. Probe set summarization and expression estimates for each gene were conducted using the apt-probeset-summarize tool (version 1.16.0) from Affymetrix<sup>67</sup>. Expression intensities under the normal, drought, high-salinity, cold, and heat conditions were used for expression analysis of AP2/EREBP genes under abiotic stresses. A gene was called up- or down-regulated across two conditions (e.g., cold vs normal condition) if the log<sub>2</sub>-based fold change of its expression abundance across the two conditions was greater than 1 or less than -1 as done in our previous report<sup>50</sup>. The significance levels of the differential expression genes were estimated by an R package 'samr' and genes with q-values  $\leq 0.05$  were deemed as significant. Hierarchical clustering of expression profiles of *B. distachyon* AP2/EREBP genes was performed using R packages (version 2.15.1).

## References

- Dubouzet, J. G. *et al.* OsDREB genes in rice, *Oryza sativa* L., encode transcription activators that function in drought-, high-salt- and cold-responsive gene expression. *Plant J* **33**, 751–763 (2003).
- Gutterson, N. & Reuber, T. L. Regulation of disease resistance pathways by AP2/ERF transcription factors. *Curr Opin Plant Biol* **7**, 465–471 (2004).
- Kume, S. *et al.* Differential and coordinated expression of Cbf and Cor/Lea genes during long-term cold acclimation in two wheat cultivars showing distinct levels of freezing tolerance. *Genes Genet Syst* **80**, 185–197 (2005).
- Yamaguchi-Shinozaki, K. & Shinozaki, K. Transcriptional regulatory networks in cellular responses and tolerance to dehydration and cold stresses. *Annu Rev Plant Biol* **57**, 781–803 (2006).
- Sakuma, Y. *et al.* Dual function of an *Arabidopsis* transcription factor DREB2A in water-stress-responsive and heat-stress-responsive gene expression. *Proc Natl Acad Sci USA* **103**, 18822–18827 (2006).
- Lim, C. J. *et al.* Over-expression of the *Arabidopsis* DRE/CRT-binding transcription factor DREB2C enhances thermotolerance. *Biochem Biophys Res Commun* **362**, 431–436 (2007).
- Sharoni, A. M. *et al.* Gene structures, classification and expression models of the AP2/EREBP transcription factor family in rice. *Plant Cell Physiol* **52**, 344–360 (2011).
- Ohme-Takagi, M. & Shinshi, H. Ethylene-inducible DNA binding proteins that interact with an ethylene-responsive element. *Plant Cell* **7**, 173–182 (1995).
- Rashotte, A. M. & Goertzen, L. R. The CRF domain defines cytokinin response factor proteins in plants. *BMC Plant Biol* **10**, 74 (2010).
- Shen, Y. G. *et al.* An EREBP/AP2-type protein in *Triticum aestivum* was a DRE-binding transcription factor induced by cold, dehydration and ABA stress. *Theor Appl Genet* **106**, 923–930 (2003).
- Pauwels, L. & Goossens, A. Fine-tuning of early events in the jasmonate response. *Plant Signal Behav* **3**, 846–847 (2008).
- Hu, Y., Jiang, L., Wang, F. & Yu, D. Jasmonate regulates the inducer of cbf expression-C-repeat binding factor/DRE binding factor1 cascade and freezing tolerance in *Arabidopsis*. *Plant Cell* **25**, 2907–2924 (2013).
- Feng, J. X. *et al.* An annotation update via cDNA sequence analysis and comprehensive profiling of developmental, hormonal or environmental responsiveness of the *Arabidopsis* AP2/EREBP transcription factor gene family. *Plant Molecular Biology* **59**, 853–868 (2005).
- Zhuang, J. *et al.* Genome-wide analysis of the AP2/ERF gene family in *Populus trichocarpa*. *Biochem Biophys Res Commun* **371**, 468–474 (2008).
- Zhang, G. *et al.* Phylogeny, gene structures, and expression patterns of the ERF gene family in soybean (*Glycine max* L.). *J Exp Bot* **59**, 4095–4107 (2008).
- Licausi, F. *et al.* Genomic and transcriptomic analysis of the AP2/ERF superfamily in *Vitis vinifera*. *BMC Genomics* **11**, 719 (2010).

17. Sharma, M. K. *et al.* Identification, phylogeny, and transcript profiling of ERF family genes during development and abiotic stress treatments in tomato. *Mol Genet Genomics* **284**, 455–475 (2010).
18. Charfeddine, M. *et al.* Genome-wide analysis and expression profiling of the ERF transcription factor family in potato (*Solanum tuberosum* L.). *Mol Biotechnol* **57**, 348–358 (2014).
19. Lata, C. *et al.* Genome-wide investigation and expression profiling of AP2/ERF transcription factor superfamily in foxtail millet (*Setaria italica* L.). *PLoS One* **9**, e113092 (2014).
20. Allen, M. D. *et al.* A novel mode of DNA recognition by a beta-sheet revealed by the solution structure of the GCC-box binding domain in complex with DNA. *EMBO J* **17**, 5484–5496 (1998).
21. Jofuku, K. D., den Boer, B. G., Van Montagu, M. & Okamura, J. K. Control of *Arabidopsis* flower and seed development by the homeotic gene APETALA2. *Plant Cell* **6**, 1211–1225 (1994).
22. Okamura, J. K. *et al.* The AP2 domain of APETALA2 defines a large new family of DNA binding proteins in *Arabidopsis*. *Proc Natl Acad Sci USA* **94**, 7076–7081 (1997).
23. Shigyo, M. & Ito, M. Analysis of gymnosperm two-AP2-domain-containing genes. *Dev Genes Evol* **214**, 105–114 (2004).
24. Feng, C. Z. *et al.* *Arabidopsis* RAV1 transcription factor, phosphorylated by SnRK2 kinases, regulates the expression of ABI3, ABI4, and ABI5 during seed germination and early seedling development. *Plant J* **80**, 654–668 (2014).
25. Fu, M. *et al.* A subset of *Arabidopsis* RAV transcription factors modulates drought and salt stress responses independent of ABA. *Plant Cell Physiol* **55**, 1892–1904 (2014).
26. Matias-Hernandez, L. *et al.* RAV genes: regulation of floral induction and beyond. *Ann Bot* **114**, 1459–1470 (2014).
27. Li, X. J. *et al.* Overexpression of cotton RAV1 gene in *Arabidopsis* confers transgenic plants high salinity and drought sensitivity. *PLoS One* **10**, e0118056 (2015).
28. Nakano, T., Suzuki, K., Fujimura, T. & Shinshi, H. Genome-wide analysis of the ERF gene family in *Arabidopsis* and rice. *Plant Physiol* **140**, 411–432 (2006).
29. Huang, G. T. *et al.* Signal transduction during cold, salt, and drought stresses in plants. *Mol Biol Rep* **39**, 969–987 (2012).
30. Yamaguchi-Shinozaki, K. & Shinozaki, K. A novel cis-acting element in an *Arabidopsis* gene is involved in responsiveness to drought, low-temperature, or high-salt stress. *Plant Cell* **6**, 251–264 (1994).
31. Jaglo-Ottosen, K. R. *et al.* *Arabidopsis* CBF1 overexpression induces COR genes and enhances freezing tolerance. *Science* **280**, 104–106 (1998).
32. Qin, F. *et al.* Cloning and functional analysis of a novel DREB1/CBF transcription factor involved in cold-responsive gene expression in *Zea mays* L. *Plant Cell Physiol* **45**, 1042–1052 (2004).
33. Kobayashi, F. *et al.* Regulation by Vrn-1/Fr-1 chromosomal intervals of CBF-mediated Cor/Lea gene expression and freezing tolerance in common wheat. *J Exp Bot* **56**, 887–895 (2005).
34. Novillo, F., Alonso, J. M., Ecker, J. R. & Salinas, J. CBF2/DREB1C is a negative regulator of CBF1/DREB1B and CBF3/DREB1A expression and plays a central role in stress tolerance in *Arabidopsis*. *Proc Natl Acad Sci USA* **101**, 3985–3990 (2004).
35. Mizoi, J., Shinozaki, K. & Yamaguchi-Shinozaki, K. AP2/ERF family transcription factors in plant abiotic stress responses. *Biochim Biophys Acta* **1819**, 86–96 (2012).
36. Qin, F. *et al.* Regulation and functional analysis of ZmDREB2A in response to drought and heat stresses in *Zea mays* L. *Plant J* **50**, 54–69 (2007).
37. Chen, H. *et al.* *Arabidopsis* DREB2C functions as a transcriptional activator of HsfA3 during the heat stress response. *Biochem Biophys Res Commun* **401**, 238–244 (2010).
38. Fujimoto, S. Y. *et al.* *Arabidopsis* ethylene-responsive element binding factors act as transcriptional activators or repressors of GCC box-mediated gene expression. *Plant Cell* **12**, 393–404 (2000).
39. Cutcliffe, J. W., Hellmann, E., Heyl, A. & Rashotte, A. M. CRFs form protein-protein interactions with each other and with members of the cytokinin signalling pathway in *Arabidopsis* via the CRF domain. *J Exp Bot* **62**, 4995–5002 (2011).
40. Zwack, P. J. *et al.* Vascular expression and C-terminal sequence divergence of cytokinin response factors in flowering plants. *Plant Cell Physiol* **53**, 1683–1695 (2012).
41. Draper, J. *et al.* *Brachypodium distachyon*, a new model system for functional genomics in grasses. *Plant Physiology* **127**, 1539–1555 (2001).
42. Vogel, J. & Hill, T. High-efficiency *Agrobacterium*-mediated transformation of *Brachypodium distachyon* inbred line Bd21-3. *Plant Cell Rep* **27**, 471–478 (2008).
43. Thole, V. & Vain, P. *Agrobacterium*-mediated transformation of *Brachypodium distachyon*. *Methods Mol Biol* **847**, 137–149 (2012).
44. IBI. Genome sequencing and analysis of the model grass *Brachypodium distachyon*. *Nature* **463**, 763–768 (2010).
45. Catalan, P. *et al.* Update on the genomics and basic biology of *Brachypodium*: International *Brachypodium* Initiative (IBI). *Trends Plant Sci* **19**, 414–418 (2014).
46. Gordon, S. P. *et al.* Genome diversity in *Brachypodium distachyon*: deep sequencing of highly diverse inbred lines. *Plant J* **79**, 361–374 (2014).
47. Tripathi, P. *et al.* The WRKY transcription factor family in *Brachypodium distachyon*. *BMC Genomics* **13**, 270 (2012).
48. Wen, F. *et al.* Genome-wide evolutionary characterization and expression analyses of WRKY family genes in *Brachypodium distachyon*. *DNA Res* **21**, 327–339 (2014).
49. Sun, J. *et al.* The *Brachypodium distachyon* BdWRKY36 gene confers tolerance to drought stress in transgenic tobacco plants. *Plant Cell Rep* **34**, 23–35 (2015).
50. Chen, L. *et al.* Genome-wide identification and analysis of MAPK and MAPKK gene families in *Brachypodium distachyon*. *PLoS One* **7**, e46744 (2012).
51. Jiang, M. *et al.* Genome-wide exploration of the molecular evolution and regulatory network of mitogen-activated protein kinase cascades upon multiple stresses in *Brachypodium distachyon*. *BMC Genomics* **16**, 228 (2015).
52. Liu, X. & Chu, Z. Genome-wide evolutionary characterization and analysis of bZIP transcription factors and their expression profiles in response to multiple abiotic stresses in *Brachypodium distachyon*. *BMC Genomics* **16**, 227 (2015).
53. You, J. *et al.* Systematic analysis and identification of stress-responsive genes of the NAC gene family in *Brachypodium distachyon*. *PLoS One* **10**, e0122027 (2015).
54. Li, C. *et al.* Comparative analyses reveal potential uses of *Brachypodium distachyon* as a model for cold stress responses in temperate grasses. *BMC Plant Biol* **12**, 65 (2012).
55. Ryu, J. Y. *et al.* Molecular and functional characterization of cold-responsive C-repeat binding factors from *Brachypodium distachyon*. *BMC Plant Biol* **14**, 15 (2014).
56. Xiong, Y. & Fei, S. Z. Functional and phylogenetic analysis of a DREB/CBF-like gene in perennial ryegrass (*Lolium perenne* L.). *Planta* **224**, 878–888 (2006).
57. Badawi, M. *et al.* The CBF gene family in hexaploid wheat and its relationship to the phylogenetic complexity of cereal CBFs. *Mol Genet Genomics* **277**, 533–554 (2007).
58. Skinner, J. S. *et al.* Structural, functional, and phylogenetic characterization of a large CBF gene family in barley. *Plant Mol Biol* **59**, 533–551 (2005).
59. Yan, H. W. *et al.* A genome-wide analysis of the ERF gene family in sorghum. *Genet Mol Res* **12**, 2038–2055 (2013).
60. Zhuang, J. *et al.* Discovery and expression profile analysis of AP2/ERF family genes from *Triticum aestivum*. *Mol Biol Rep* **38**, 745–753 (2011).

61. Campoli, C. *et al.* Comparative expression of Cbf genes in the *Triticeae* under different acclimation induction temperatures. *Mol Genet Genomics* **282**, 141–152 (2009).
62. Zhang, Y. *et al.* Genome-wide identification and analysis of grape aldehyde dehydrogenase (ALDH) gene superfamily. *PLoS One* **7**, e32153 (2012).
63. Sakuma, Y. *et al.* DNA-binding specificity of the ERF/AP2 domain of *Arabidopsis* DREBs, transcription factors involved in dehydration- and cold-inducible gene expression. *Biochem Biophys Res Commun* **290**, 998–1009 (2002).
64. Rao, G. *et al.* Genome-wide analysis of the AP2/ERF gene family in *Salix arbutifolia*. *FEBS Open Bio* **5**, 132–137 (2015).
65. Cannon, S. B. *et al.* The roles of segmental and tandem gene duplication in the evolution of large gene families in *Arabidopsis thaliana*. *BMC Plant Biol* **4**, 10 (2004).
66. Badawi, M. *et al.* Structure and functional analysis of wheat ICE (inducer of CBF expression) genes. *Plant Cell Physiol* **49**, 1237–1249 (2008).
67. Priest, H. D. *et al.* Analysis of global gene expression in *Brachypodium distachyon* reveals extensive network plasticity in response to abiotic stress. *PLoS One* **9**, e87499 (2014).
68. Zhao, L. *et al.* Roles for a soybean RAV-like orthologue in shoot regeneration and photoperiodicity inferred from transgenic plants. *J Exp Bot* **63**, 3257–3270 (2012).
69. Woo, H. R. *et al.* The RAV1 transcription factor positively regulates leaf senescence in *Arabidopsis*. *J Exp Bot* **61**, 3947–3957 (2010).
70. Osnato, M., Castillejo, C., Matias-Hernandez, L. & Pelaz, S. TEMPRANILLO genes link photoperiod and gibberellin pathways to control flowering in *Arabidopsis*. *Nat Commun* **3**, 808 (2012).
71. Sohn, K. H. *et al.* Expression and functional roles of the pepper pathogen-induced transcription factor RAV1 in bacterial disease resistance, and drought and salt stress tolerance. *Plant Mol Biol* **61**, 897–915 (2006).
72. Mittal, A. *et al.* Related to ABA-Insensitive3 (ABI3)/Viviparous1 and AtABI5 transcription factor coexpression in cotton enhances drought stress adaptation. *Plant Biotechnol J* **12**, 578–589 (2014).
73. Li, C. W. *et al.* Tomato RAV transcription factor is a pivotal modulator involved in the AP2/EREBP-mediated defense pathway. *Plant Physiol* **156**, 213–227 (2011).
74. Soltesz, A. *et al.* Transgenic barley lines prove the involvement of TaCBF14 and TaCBF15 in the cold acclimation process and in frost tolerance. *J Exp Bot* **64**, 1849–1862 (2013).
75. Lynch, M. & Conery, J. S. The evolutionary fate and consequences of duplicate genes. *Science* **290**, 1151–1155 (2000).
76. Larkin, M. A. *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948 (2007).
77. Tamura, K. *et al.* MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol* **30**, 2725–2729 (2013).
78. Kent, W. J. BLAT—the BLAST-like alignment tool. *Genome Res* **12**, 656–664 (2002).
79. Fu, L. *et al.* CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* **28**, 3150–3152 (2012).
80. Zhu, Y. *et al.* Soybean (Glycine max) expansin gene superfamily origins: segmental and tandem duplication events followed by divergent selection among subfamilies. *BMC Plant Biol* **14**, 93 (2014).
81. Lee, T. H., Tang, H., Wang, X. & Paterson, A. H. PGDD: a database of gene and genome duplication in plants. *Nucleic Acids Res* **41**, D1152–1158 (2013).
82. Krzywinski, M. *et al.* Circos: an information aesthetic for comparative genomics. *Genome Res* **19**, 1639–1645 (2009).
83. Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**, 1105–1111 (2009).
84. Trapnell, C. *et al.* Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol* **28**, 511–515 (2010).
85. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).

## Acknowledgements

The work was supported in part by Natural Science Foundation of China (31300999), Youth and Technology Morning Program in Wuhan (2014072704011257), the Talent Development Program of Wuhan, the municipal government of Wuhan, Hubei, China (2014070504020241), an internal research grant of Jiangnan University (14042), Wuhan, China, and by United States National Institutes of Health (R01GM100364) and United States National Science Foundation (DBI-0743797). We thank the contributors of the *B. distachyon* genome database for convenient search of BdAP2/EREBP genes, and Priest *et al.* for guidance on processing the microarray data of the BdAP2/EREBP genes under abiotic stresses.

## Author Contributions

L.C. perceived the research and designed the experiments; J.H. and X.D. grew *B. distachyon*, and performed the RT-PCR validation and transactivation assay. S.T. and L.C. analyzed the gene distribution, chromosome location, gene structure and orthologs relationships among four grass species. J.Z., H.P. and L.L. constructed sequencing libraries and performed sequencing profiling; L.L. and S.T. carried out computational analysis; L.C., S.T. and J.H. created all the figures and tables; G.Y. provided the *B. distachyon* seeds used in this study; L.C., J.H. and W.Z. analyzed the data and results and wrote the paper; L.C., G.H. and W.Z. revised the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Chen, L. *et al.* Expansion and stress responses of AP2/EREBP superfamily in *Brachypodium distachyon*. *Sci. Rep.* **6**, 21623; doi: 10.1038/srep21623 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>