Comparative analysis of *MAPK* and *MKK* gene families reveals differential evolutionary patterns in *Brachypodium distachyon* inbred lines

Min Jiang[1,2,#,*], Peng Li[1#], and Wei Wang[1]

[1]Shanghai Key Laboratory of Plant Functional Genomics and Resources, Shanghai Chenshan Plant Science Research Center, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences (CAS), Shanghai Chenshan Botanical Garden, Shanghai, China
[2]Ministry of Education Key Laboratory for Biodiversity Science and Ecological Engineering, Insititute of Biodiversity Science, School of Life Sciences, Fudan University, Shanghai, China
# These authors contributed equally in this work.
* Correspondence: Min Jiang (yijinsha@126.com)

Authors' email addresses:

Min Jiang: yijinsha@126.com

Peng Li: 381778082@qq.com

Wei Wang: 1514034946@qq.com

Abstract

Background: Mitogen-activated protein kinase (MAPK) cascades are involved with signal transduction in almost every aspect of plant growth and development, as well as biotic and abiotic stress responses. The evolutionary analysis of MAPKs and MKKs in individual or entire plant species has been reported, but the evolutionary patterns in the diverse inbred lines of *Brachypodium distachyon* are still unclear.
Results: We conducted the systematical molecular evolutionary analysis of *B. distachyon*. A total of 799 MAPKs and 618 MKKs were identified from 53 *B. distachyon* inbred lines. Remarkably, only three inbred lines had 16 MPKs and most of those inbred lines lacked MPK7-2 members, whereas 12 MKKs existed in almost all *B. distachyon* inbred lines. Phylogenetic analysis indicated that MAPKs and MKKs were divided into four groups as previously reported, grouping them in the same branch as corresponding members. MPK21-2 was the exception and fell into two groups, which may be due to their exon-intron patterns, especially the untranslation regions (UTRs). We also found that differential evolution patterns of MKK10 paralogues from ancient tandem duplicates may have undergone functional divergence. Expression analyses suggested that *MAPKs* and *MKKs* likely played different roles in different genetic contexts within various tissues and with abiotic stresses.
Conclusion: Our study revealed that UTRs affected the structure and evolution of *MPK21-2* genes and the differential evolution of *MKK10* paralogues with ancient tandem duplication might

35 have functional divergences. Our findings provided new insights into the functional evolution of
36 genes in closely inbred lines.

Introduction

38 Mitogen-activated protein kinase (MAPK or MPK) signaling cascades play vital roles in the
39 stress response, cell division, and developmental regulation. They are divided into three highly-
40 conserved subfamilies that continuously act in a sequential manner in evolution and fundamental
41 signaling transduction pathways (Rodriguez et al., 2010; Xu and Zhang, 2015; Jagodzik et al.,
42 2018). The MAPK kinase kinases (MKKKs or MEKKs) are activated by extracellular cues and
43 subsequently phosphorylate and activate the S/T-X$_{3-5}$-S/T motif of downstream MAPK kinases
44 (MAPKKs or MKKs), which, in turn, phosphorylate and activate MAPKs at their TXY
45 activation loop (Rodriguez et al., 2010, Singh et al., 2012). Activated MPKs regulate
46 downstream cellular targets, including regulatory and metabolic enzymes and transcription
47 regulators (Joo et al., 2008, Guan et al., 2014).

48 *Brachypodium distachyon* (2n = 10) is an annual temperate grass with a close phylogenetic
49 relationship to other temperate cereals and an intermediate position within the Pooideae
50 subfamily (Soreng et al., 2015; Catalan et al., 2016). *B. distachyon* is desirable for its small
51 physical stature, rapid life cycle, ability to self-fertilize, and small diploid genome size (Draper et
52 al., 2001; Garvin et al., 2008). Highly efficient *Agrobacterium*-mediated transformation methods
53 in Brachypodium have also been established (Vain et al., 2008; Vogel and Hill, 2008). Therefore,
54 *B. distachyon* is widely used as a model plant for studying problems unique to cereals and
55 grasses (Vogel et al., 2010; Brkljacic et al., 2011; Mur et al., 2011; Catalan et al., 2014). The
56 morphological, molecular, and cytological analyses of diverse *B. distachyon* inbred lines have
57 been conducted (Filiz et al., 2009; Vogel et al., 2009) and their nuclear and plastid genomes have
58 been deep sequenced and annotated (Gordon et al., 2014; Gordon et al., 2017; Sancho et al.
59 2018). Further analysis showed that the inbred lines of *B. distachyon* are divided into three
60 different genomic groups, including a mostly Extremely Delayed Flowering (EDF+) clade, a
61 mostly Spanish (S+) clade, and a Turkish (T+) clade, based on their flowering phenotype and
62 geographical substructure (Sancho et al., 2018).

63 To date, the evolutionary mechanisms of MAPK cascades in plants have indicated a diverse
64 domain organization and novel activation loop variants (Mohanta et al., 2015) and/or distinct
65 expansion mechanism (Jiang and Chu, 2018). A variety of single-gene duplication types emerged
66 continuously and have involved in the plant's adaptation to dramatically changing environments
67 (Wang et al., 2009; Cuevas et al., 2016). However, whole-genome duplications (WGDs) are
68 considered to be a major force in the evolution of morphological and physiological diversity
69 (Soltis et al., 2009; Paterson et al.; 2010). The ancient tandem duplication event occurred at the
70 adjacent genes in the same chromosome, which are usually expanded or retained by an unequal
71 crossing (Freeling, 2009). Tandem duplication often displays less expression difference and
72 functional divergence than distant duplication (Makino and McLysaght, 2012; Ghanbarian and
73 Hurst, 2015). However, there is limited information on the gene expansion mechanism and

74   functional evolution of the MAPK cascades in diverse *B. distachyon* inbred lines, including for
75   Bd21 (Chen et al., 2012; Jiang et al., 2015). We studied the evolutionary patterns of MAPKs and
76   MKKs from different *B. distachyon* inbred lines. The phylogenetic relationships and the
77   identification of MAPKs and MKKs were determined for 53 *B. distachyon* inbred lines. We
78   investigated gene and domain construction patterns of the individual members with a certain
79   divergence and focused on the evolutionary history of *MKK10* paralogues in different *B.*
80   *distachyon* inbred lines. This revealed various conservative and divergent tandem gene clusters.
81   The expression patterns of these genes were analyzed in Bd21, BdTR8i, and Bd30-1 from three
82   genetic groups in various tissues and abiotic stresses, and their potential functions were also
83   investigated.

Methods and materials

Identification of *MAPK and MKK* gene family members

86   We downloaded gene information for *MAPK* and *MKK* from *B. distachyon* Bd21 from the
87   PLAZA platform (https://bioinformatics.psb.ugent.be/plaza/) (Van Bel et al., 2018). BLASTP
88   (Altschul et al. 1997) searches were conducted with a threshold of 90% identity; searches were
89   performed with orthologous protein sequences using BdMAPKs and BdMKKs as queries in
90   *BrachyPan* (https://brachypan.jgi.doe.gov/) (Gordon et al., 2017) to identify these genes in the
91   53 diverse *B. distachyon* inbred lines. Collected sequences were only accepted for scanning
92   using InterPro software (Mitchell et al. 2019) if they harbored MAPK or MKK consensus
93   sequences, including the activation loop TXY motif for MPKs, the active site motif D(L/I/V)K,
94   and the phosphorylation target site S/T-X$_5$-S/T within the activation loop for MKKs. The gene
95   identifier information of these sequences was collected and is listed in Table S1 and Table S2.

Gene structure and sequence alignments

97   The exon/intron structure of identified *MAPKs* and *MKKs* was performed using Gene Structure
98   Display Server (GSDS) software (http://gsds.cbi.pku.edu.cn/). All of the full-length amino acid
99   sequences were initially aligned using Clustal Omega with default parameters
100 (http://www.ebi.ac.uk/Tools/msa/clustalo/). The domains and motifs of MAPKs and MKKs were
101 scanned using InterProScan software (http://www.ebi.ac.uk/interpro/) (Jones et al., 2014). The
102 structural schematic of all members of MAPK and MKK were executed according to
103 InterProScan analysis results. The alignment logos of the protein conserved domain were
104 generated using the WebLogo3 application (http://weblogo.threeplusone.com/).

Synteny and phylogenetic analyses

106 The phylogenetic relationships of all 53 *B. distachyon* inbred lines were generated in the
107 *BrachyPan* project and visualized with the CorelDRAW X3 program. Phylogenetic trees were
108 created based on the alignment of all MAPKs or MKKs using the maximum likelihood (ML)
109 method with the Jones-Taylor-Thornton (JTT) model, 2,000 bootstrap values, and partial
110 deletion by the MEGA 6.0 software, respectively (Tamura et al., 2013). The Neighbor Joining
111 (NJ) Trees of the MAPKs or MKKs were also reconstructed with the same parameters using

112   MEGA 6.0. We obtained the synteny information of duplicate genes and the tandem (TD) data
113   from the PlantDGD database (http://pdgd.njau.edu.cn:8080/) (Qiao et al., 2019).

Plant sample preparation

115   We sowed BdTR8i, Bd21, and Bd30-1 seeds in 1/2 MS medium in the dark for 4 d at 25 °C and
116   then transferred them to a soil mix. Plants were grown in a greenhouse under 14 h light (21
117   °C)/10 h dark (18 °C) photoperiods. We harvested the root, stem, leaf blade, and leaf sheath at
118   the eight-to-nine leaf stage. Spikelet samples from *B. distachyon* were collected at the early
119   flowering stages according to their different flowering times (Figure S1). For the abiotic stress
120   treatment, 2-week-old *B. distachyon* seedlings were dipped in ½ MS liquid medium containing
121   20% PEG 6000 and 200 mM NaCl, and then plants were collected after treatment for 3 h and 6 h,
122   respectively. Moreover, seedlings were transferred to a growth chamber and heat-treated at 40 °C
123   for 3 h and 6 h. All samples were flash frozen in liquid nitrogen and stored at -80 °C for RNA
124   extraction.

Expression analysis

126   Total RNA was extracted from samples using Trizol reagent and 1-2 µg was reverse-transcribed
127   into cDNA using PrimeScript RT Master Mix Perfect Real Time (TaKaRa) according to the
128   manufacturer′s instructions. The quality of total RNA was detected using Nanodrop1000 and its
129   integrity was estimated by electrophoresis in 1.5 % (w/v) agarose gel. The real-time quantitative
130   polymerase chain reaction (RT-qPCR) was carried out in 10 µl reactions with 5-50 ng of first-
131   stand cDNA products (4 µl), 5 pmol of each primer (0.4 µl), 5 µl SYBR green master mix (2X),
132   0.2 µl ROX as a passive reference standard to normalize the SYBR fluorescent signal. The
133   conditions for RT-qPCR were: initial activation at 95 °C for 5 min followed by 45 cycles of 95
134   °C for 30 s, and 60 °C for 30 s. Subsequently, the specificity of PCR products was monitored
135   using a melting curve analysis (61-95 °C with fluorescence read every 0.5 °C). The *B.*
136   *distachyon actin* (gene locus: *Bradi2g24070*) gene was used as an internal control for all RT-
137   qPCR analyses; specific primers for *MAPK* and *MKK* were listed in Table S3. Three independent
138   biological replicates were conducted for each experiment. The relative expression of *MAPK* and
139   *MKK* genes was calculated using the $2^{-\Delta\Delta Ct}$ method.

Results

Identification and annotation of MPKs and MKKs in 53 diverse *B. distachyon* inbred lines

142   We identified the two gene families by searching homologous Bd21 sequences in the public
143   *BrachyPan* database (Chen et al., 2012) to determine the conservation and divergence of MPKs
144   and MKKs in 53 diverse *B. distachyon* inbred lines. All predicted MPKs and MKKs were named
145   based on the similarity of their orthologous protein to that of *A. thaliana* and *B. distachyon*
146   (Ichimura et al., 2002; Chen et al., 2012). Ultimately, a total of 799 MPKs and 618 MKKs were
147   retrieved (Table 1, Table S4, and Table S5). We found that most *B. distachyon* inbred lines had
148   14 or 15 MPKs apart from Bd21, BdTR3c, and Bd18-1, which had 16 members (Table 1).
149   Further analysis showed that only seven inbred lines had the *MPK7-2* gene, including Bd21,

150   Bd2-3, BdTR3c, Bd18-1, S8iiC, Mur1, and Foz1 (Table S1). This may be the result of an
151   incomplete annotation of the genome sequence or the long sequence of the MPK7-2 protein,
152   which usually consists of 1,708 amino acid (aa) residues. Most *B. distachyon* inbred lines
153   harbored 12 MKK members except Tek-4 (11), Bd3-1 (7), Adi-10 (10), Gaz-8 (5), ABR5 (11),
154   Foz1 (11), and Jer1 (11) (Table 1 and Table S5). The incomplete assembly of Tek-4 (77.82%),
155   Bd3-1 (89.52%), Adi-10 (89.52%), and Gaz-8 (88.31%) may be the reason that relatively few
156   MKK members have been identified (Gordon et al., 2017). Further analysis showed that the *B.*
157   *distachyon* inbred lines lacked a particular MKK member; for example, MKK10-5 of Tek-4,
158   MKK10-4 of ABR5, MKK4 of Foz1, MKK5 of Jer1, MKK10-3 and -4 of Adi-10 (Table S2).
159   We also incorporated the available genomic detailed information from MPKs and MKKs (Table
160   S1 and Table S2).

Phylogenetic classification of *B. distachyon* inbred lines MAPKs and MKKs

162   To investigate the phylogenetic relationship of MPK proteins in diverse *B. distachyon* inbred
163   lines, the phylogeny of all identified 799 MPK protein sequences were performed using ML and
164   NJ methods, respectively. As expected, all homologues for each of the 16 Bd21 MPKs (BdMPKs)
165   were divided into four groups (A, B, C, and D) and clustered on the corresponding branch except
166   Tek-4MPK16 (Figures 1, S2 and S3). Tek-4MPK16 consisted of only 183 aa, while the other
167   MPK16 members had 544 aa (Table S1). MPK21-2 had two branches, designated as type I and II
168   (Table S6), indicating that it may have a certain functional divergence. In addition, MPK7-1 and
169   MPK7-2 were located on same branch with a large discrepancy in their lengths (Figure 1),
170   suggesting functional divergence, which is supported by previous functional studies (Jiang et al.,
171   2015).

172   We analyzed a total of 618 MKKs for their phylogenetic relationship with corresponding protein
173   sequences using ML and NJ methods, respectively. Almost all of the orthologous genes for each
174   of the 12 Bd21 MKKs (BdMKKs) had similar clustering patterns with corresponding branches.
175   The exceptions fell into four groups: A, B, C, and D (Figures 2, S4 and S5). However,
176   BdTR10cMKK10-3, Jer1MKK10-4, Mur1MKK10-5, and BdTR13cMKK10-5 were branched
177   out from the other members, suggesting that these gene members may have diverged (Figure 2).
178   It is noteworthy that Mur1MKK10-5 (188 aa) and BdTR13cMKK10-5 (208 aa) had shorter
179   amino acids than other members, which usually contained 332 aa (Table S2). BdTR10cMKK10-
180   3 (163 aa) was also shorter relative to the normal MKK10-3 (344 aa). In contrast, Jer1MKK10-4
181   (424 aa) was longer when compared with the typical MKK10-4 (341 aa) (Table S2). These
182   situations may affect their evolutionary relationship with MKKs from other *B. distachyon* inbred
183   lines.

Exon-intron compositions and length variations of MPKs and MKKs in *B. distachyon*
inbred lines

186   The abundance of non-protein-coding DNA within a genome, such as an intron, increased
187   consistently with the genome complexity (Taft et al., 2007). Intron pattern analyses can enhance
188   our understanding of the structure and evolution of genes (Zhang et al., 2014). We also surveyed

the exon-intron architecture of different MPKs and MKKs using GSDS software to elucidate the relationship or divergence among paralogues and orthologues. Most members showed similar exon-intron patterns with the intron number, exon length, and intron phase. The intron number was found to be relatively constant in three genetic groups, with the exceptions of MPK20-3, MPK21-2, and MKK3-3 (Figure 3). Remarkably, the number of introns from MPK11, MPK21-1, and MKK1 in all *B. distachyon* inbred lines consistently contained 5, 10, and 8 introns, respectively (Figure 3). Almost all MPK20-3 had three introns in group T+ apart from BdMPK20-3, Bd21-3MPK20-3, BdTR13aMPK20-3, BdTR3cMPK20-3, and Bd18-1MPK20-3 which had eight introns, while other two genetic groups of MPK20-3 displayed three or eight introns (Figures 3A and S6). Moreover, the number of introns of MKK3-3 was highly consistent in group S+, and T+ usually contained nine introns. In contrast, the intron numbers in group EDF+ were highly variable; for example, ABR9MKK3-3, Bd29-1MKK3-3, Tek-4MKK3-3 had eight, six, and four introns, respectively (Figures 3B and S7).

The exon-intron patterns of MPK21-2 fell into two groups, which coincided with their phylogenetic relationship mentioned above (Figures 1 and 4). However, the phylogenetic relationship was not completely consistent between the reconstructed full length coding sequence (CDS) and their exon-intron patterns (Figure 4). Further analysis showed that all type II MPK21-2s harbored UTRs, while type I MPK21-2s had no UTR except for RON2MPK21-2 and Tek-4MPK21-2 (Figure 4). Most MPK21-2s had seven introns in group EDF+ and S+ or nine introns in group T+ except for BdMPK21-2, Bd3-1MPK21-2, and Adi-10MPK21-2 (Figure 3A). These results suggested that the structure and evolution of these genes were influenced by intron patterns and may be affected by UTRs.

Common conserved domain analysis of *B. distachyon* MPKs and MKKs

Previous research has reported that MAPKs had several conserved domains or signature sequences with vital structural or functional roles, including the GxGxxG motif in the nucleotide binding (NB) domain (Mohanta et al., 2015), the TXY motif in the activation loop (Xu and Zhang, 2015), D(I/L/V)K motif in the active site (Goyal et al., 2018), and the common docking (CD) domain in the C-terminal extension region outside the catalytic domain (Tanoue et al., 2000). Specifically, the threonine or tyrosine in the TXY motif as the activation loop plays pivotal roles in the signal transduction pathway. Remarkably, the average abundance of threonine or tyrosine (TXY), the most important enzymes in the *B. distachyon* inbred lines, were 4.67 and 3.91, respectively, and 4.62 and 3.81 in Bd21, respectively, (Table S7) suggesting that these amino acids were relatively constant. We also found that the conserved domains of eleven MPKs reported in comprehensive plant species were highly conserved in individual MPKs of *B. distachyon* inbred lines (Figures S8 and S9). These analyses revealed that the activation loop TEY in groups A, B, and C, and TDY in group D, were consistent with results from previous studies. Most notably, the activation loop TEY motif in all *B. distachyon* MPK11s was replaced by the MEY motif (Figure 1). MAPKs harbored a CD domain featuring a cluster of negatively-charged amino acids with consensus sequences M/L-L-A/V-F-D-P-X2-R-P/I-T/S-A/V-X-E-A-L (Figure 1) that bind the basic residues at the N-terminus of the docking site in MAPK-interaction

229 proteins (Jiang et al., 2018). MPK7-2s belonged to leucine-rich repeat receptor kinases (LRR-
230 RKs) that had LRR domains and an NB-ARC domain, only appearing in seven kinds of *B.*
231 *distachyon* inbred lines (Figure 1). Moreover, the DLK active site within the MPK7-1s signature
232 was conserved; however, Luc1MPK7-1 and ABR6MPK7-1 were replaced with a DLN motif.
233 Foz1MPK7-2 and Mur1MPK7-2 with DLN motifs were treated similarly (Figure S8).
234 Specifically, all MPK11s had a DLR motif instead of a DLK active site (Figure S8). Furthermore,
235 an elongation factor hand (EF-hand) calcium binding protein (CBP) with the consensus
236 sequences "NDNNEQHASDQER" was observed in all *B. distachyon* MPK17s at their C-
237 terminal end (Figures 1, 5A, and 5B). Further analysis indicated that 13 MPK17 members (group
238 T+: BdTR13cMPK17, Bis-1MPK17, Bd21-3MPK17, BdMPK17, Adi-2MPK17 and
239 BdTR13aMPK17; group EDF+: Tek-2MPK17, Tek-4MPK17, Bd29-1MPK17, ABR9MPK17,
240 BdTR8iMPK17, BdTR7aMPK17, and Bd1-1MPK17) had a mutation in which E changed to D,
241 which only resulted from a single nucleotide substitution of A changed to C (Figure 5C). This
242 mutation may be have no effect on the function due to the canonical EF-hand (for example the
243 calcineurin B-like (CBL) protein), that are characterized by a conserved Asp (D) and Glu (E)
244 residue with completely constant spacing (Kolukisaoglu et al., 2004; Jiang et al., 2020).
245 Remarkably, the MPK17s exhibited the conserved domain specific to *B. distachyon* members
246 compared to other plant species, especially eudicots (Figure 5D).

247 As same as the MAPKs, MKKs also contained some important domains or motifs including the
248 activation or T-loop $S/T-X_5-S/T$ motif (Asai et al., 2002), the docking site $(K/R_{2-3}X_{1-5}L/IXL/I)$ in
249 the N-terminal domain (Bardwell and Shah, 2006; Jiang and Chu, 2018), the GxGxxGxV motif
250 in the NB domain and $HK-X_6-ALK$ motif in the ATP binding site (Hadiarto et al., 2006), and the
251 active site D(I/L/V)K motif (Goyal et al., 2018). A detailed analysis of the conserved sequences
252 of MKKs was displayed in alignment of individual MKKs (Figures S10 and S11). Groups A, B,
253 and C MKKs had the typical T-loop $S/T-X_5-S/T$ motif, while group D MKKs (MKK10s) had a
254 part mutation in the phosphorylation site which coincided with a wide range of plant species
255 (Figures 2 and S10) (Jiang and Chu, 2018). Interestingly, the average abundance of serine or
256 threonine $(S/T-X_5-S/T)$, which were the most crucial amino acids in *B. distachyon* inbred lines,
257 were 7.2 and 3.32, respectively, while the same results were 7.2 and 3.32 in Bd21 (Table S7),
258 respectively. This suggests that these amino acids remained constant. We speculated that the
259 MKKs may have experienced fundamental functional conservation. Further analysis showed that
260 variations of some the conserved MKKs were also present. For instance, in addition to
261 BdTR13CMKK4 being replaced by a DIL motif, the D(I/L/V)K active site within the signature
262 of MKKs was conserved despite occasional variations (Figure S10). We found the
263 HRPTGRCYALK motif in the ATP binding site of MKK5 members, however, BdTR3cMKK5
264 was replaced by the HRPPGRCYALK motif (Figure S10). Furthermore, our data showed that
265 the nuclear transport factor 2 (NTF2) domains existed in all MKK3s from *B. distachyon* inbred
266 lines (Figure 2).

267 **The differential evolution of *MKK10* paralogs with tandem duplications**

268   Gene duplication was the necessary material source for evolutionary novelty, leading to the gene
269   responsible for the gene families (Lynch and Conery, 2000). In addition, some tandem
270   duplication was observed in monocot *MKK10* paralogues, such as in *B. distachyon* (Jiang and
271   Chu, 2018). A 6,7-dimethyl-8-ribityllumazine (*DMRL*) synthase gene was observed between two
272   *MKK* members (Jiang and Chu, 2018). We surveyed the duplicated genes in different *B.*
273   *distachyon* inbred lines genome to further comprehend the duplication and evolutionary events of
274   the *B. distachyon MKK10* paralogues. As expected, most of *MKK10* paralogues with the
275   exception of five kinds of *B. distachyon* inbred lines (Mon3, Bd3-1, Adi-10, BdTR10c, and Gaz-
276   8) presented tandem duplication in the canonical form of the *MKK-DMRL-MKK* model with
277   occasional variations (Figure 6). In addition, Tek-4 had the *PNN* (pinin) gene instead of the
278   *DMRL* gene between two *MKK* gene members (Figure 6). Twenty *B. distachyon* inbred lines had
279   the same canonical model with Bd21 in the form of the *MKK-DMRL-MKK-MKK* model. Eight *B.*
280   *distachyon* inbred lines possessed the tandem duplication in the form of the *MKK-DMRL-MKK*
281   model (Figure 6). Moreover, some small variations were also found in other *B. distachyon* inbred
282   lines. For example, we also found the tandem gene clusters with the *MKK-DMRL-UDPGT-*
283   *UDPGT-MKK* (*UDPGT*: UDP-glucoronosyl and UDP-glucosyl transferase) model in Mur1,
284   *MKK-DMRL-MKK-ChaC-MKK* (*ChaC*: ChaC-like protein) model in BdTR12c, and *MKK-*
285   *DMRL-PK-PK-MKK* (*PK*: protein tyrosine kinase) model in Sig2 (Figure 6). These results
286   indicated that the tandem *MKK10* gene clusters in *B. distachyon* inbred lines originated in the
287   common ancestral genomic contexts and a certain variation developed in order to adapt to the
288   environment differences including light, temperature, or elevation.

289   **Expression variation in the *MPK* and *MKK* gene family in three selected genomes**

290   The expression profiles for five different tissues (root, stem, leaf blade, leaf sheath, and spikelet)
291   and abiotic stresses (salt, drought, and heat) of *MPK* and *MKK* genes were performed (Table S8)
292   to explore the expression variation of the *MPK* and *MKK* gene family in Bd21, BdTR8i, and
293   Bd30-1. These families belong to three different genetic groups. We observed the different
294   expression levels of 15 *MPK* and 11 *MKK* genes apart from the *MKK10-4* gene (Figures 7 and 8),
295   which may be involved with the barely detectable low expression level of the *MKK10-4* gene
296   observed in previous research (Jiang and Chu, 2018). Among these *MPK* and *MKK* genes, most
297   genes showed distinct quantitative expression patterns in their different genetic backgrounds. For
298   instance, the expression level of *MPK4s* at root was higher in BdTR8i and moderate in Bd21
299   compared with in Bd30-1 (Figure 7). *MPK3s* had the same reaction after 6 h of drought
300   treatment (Figure 8). *MKK10-3s* had higher expression in spikelet's in BdTR8i compared with in
301   Bd21 and Bd30-1 (Figure 7). Moreover, these genes also had similar expression patterns in
302   different genetic backgrounds, such as *MPK16s*, *MPK20-4s*, *MKK3-2s,* and *MKK10-1s* in
303   spikelet, *MKK5s* in leaf sheath, *MPK6s* and *MKK10-2s* in leaf blade, *MPK16s* and *MKK10-5s* in
304   stem, *MPK16s,* and *MPK20-5s* and *MKK10-1s* in root (Figure 7). *MKK4s* had a higher
305   expression under drought conditions in BdTR8i compared with in Bd21 and Bd30-1 (Figure 8).
306   *MPK3s* had a relatively low expression under heat treatment in three backgrounds, while the
307   expressions of *MPK14s* and *MPK20-1s* were opposite (Figure 8). Expression variations were

308 also observed in the different tissues and/or abiotic stresses. For example, *MKK10-1s* were more
309 highly expressed in three genetic backgrounds in roots than in other tissues (Figure 7). Almost all
310 genes had higher expression levels after 6 h of exposure to three abiotic stresses in BdTR8i
311 (Figure 8). We further found that certain genes had unique expression profiles in specific tissues
312 of a particular genetic background. For example, *BdTR8iMPK4* was highly expressed in only the
313 root (Figure 7) and *BdTR8iMKK1* was highly expressed in heat, salt, and drought stresses (Figure
314 8). The striking variations of *MPK* and *MKK* gene family members expressed in different genetic
315 contexts increases the diversity of the potential biological functions of these genes.

316 **Discussion**

317 **Exon-intron compositions with conservative and divergent patterns**

318 The conservation of exon length was associated with constraints of the gene function of
319 organisms (Davila-Velderrain et al., 2014). The non-coding regions, such as the intron, may
320 affect gene functions by a gradual deletion, which may be the result of recombination throughout
321 the evolution of the intron (Hu, 2006). Therefore, we investigated the exon-intron composition of
322 the corresponding *MPKs* and *MKKs*. Our results showed that the exon-intron architecture,
323 including lengths and numbers of intron, intron phase, and lengths of UTR, was generally
324 conserved in corresponding orthologs (Figures S6 and S7). However, some variability was also
325 found. For example, type II *MPK21-2s* harbored 5'-UTR and 3'-UTR which were absent in type
326 I *MPK21-2s* (Figure 4), although they were in agreement with their phylogenetic relationships
327 (Figure 1), indicating that they may have a difference in expression and functional divergence.
328 The UTR length-dependent functional specificity significantly increases the coding capacity of
329 the genome that regulates multiple plant process, including nutrient homeostasis, stress responses,
330 and plant growth and development (Srivastava et al., 2018). In addition, there was a large
331 difference in the intron lengths and numbers among *B. distachyon MPK7-1s* and *MPK20-3s*
332 (Figure S6). A detailed analysis showed that the fourth intron of *BdTR8iMPK20-3* was shorter
333 than *BdMPK20-3* and *Bd30-1MPK20-3* (Figure S6). Moreover, *BdTR8iMPK20-3* had a lower
334 expression level in roots and spikelets compared to the other corresponding members as
335 described in a previous study that intron lengths were correlated with gene expression (Rose et
336 al., 2016). Our findings indicated that the exon-intron composition affect the evolutionary
337 patterns and expression efficiency of *MPK* and *MKK* orthologs.

338 **Tandem duplications contributed to *MKK10s* gene expansion**

339 Our analysis suggested that *MKK10* paralogs undergo an ancient tandem duplication event with
340 differential evolution. Further examination of tandem *MKK10* gene clusters revealed that a
341 *DMRL* gene often occurred (Figure 6). These results are supported by previous studies (Jiang and
342 Chu, 2018), indicating that they were derived from common ancestral genomic contexts.
343 However, some variations have also been found among *B. distachyon* inbred lines such as an
344 insertion of *RRM* (RNA recognition motif protein) instead of the *DMRL* gene (Figure 6). This
345 may result in a difference of gene expression. Indeed, tandem duplicates generally show more
346 similar expression patterns than remote duplicates (Dai et al., 2014, Lan and Pritchard, 2016) and

347 preferentially retain the cis-PPIs (protein–protein interactions) after WGD (Makino and
348 McLysaght, 2008, 2012). Therefore, ancient tandem duplications of *MKK10s* may have
349 contributed to gene expansion and function conservation and/or divergence during the evolution
350 process of monocots.

Expression divergence of *MAPKs* and *MKKs* within three *B. distachyon* genetic groups

352 Tissue-specific expression patterns of *MAPK* and *MKK* genes have been characterized with
353 corresponding functions in plant growth and development. For instance, the expression levels of
354 *AtMKK10* are high in pollen but do not appear in shoot apices, mesophyll cells, or mature leaves
355 (Yoo et al., 2008), indicating a potential role in flower tissues (Jiang and Chu, 2018). *CaMPK19-*
356 *2* genes are highly expressed in roots and stems in pepper, while *CaMPK1* is highly expressed in
357 in leaves (Liu et al., 2015), which indicates that these genes are expressed preferentially in
358 different tissues and developmental stages (Wei et al., 2014). We investigated the tissue-specific
359 expression profiles in different *B. distachyon* inbred lines. The result indicated that most *MPK*
360 and *MKK* genes had quantitative distinct expression patterns among the three different genetic
361 contexts in different tissues and various abiotic stresses. For example, *MPK17* had higher
362 expression levels in the root, stem, leaf blade, and salt treatment in Bd30-1 compared with Bd21
363 and BdTR8i (Figures 7 and 8). These results are indicative of the distinct function of *MPK17s*,
364 which may result from the nonsynonymous substitutions at some pivotal amino acid sites in EF-
365 hand CBP motif in their C-terminal extensions (Figure 5) as described previously (Yang et al.,
366 2019). Moreover, *MKK10-3* and *MKK10-5* had similar expression patterns in the leaf blade in
367 Bd30-1 and BdTR8i and distinct profiles in Bd21 (Figure 7). These results coincide with the
368 tandem gene cluster model (Figure 6) and are supported by previous reports that physically
369 linked genes (tandem duplicates) usually had less expression differences than distant genes
370 (Ghanbarian and Hurst, 2015; Lan and Pritchard, 2016). Furthermore, the *MKK3-2* gene had
371 similar patterns under heat and salt condition (Figure 8). Taken together, these results suggest
372 that *MAPKs* and *MKKs* had an expression divergence which was correlated with the differential
373 evolution in *B. distachyon* inbred lines.

Conclusion

375 A total of 799 *MPK* and 618 *MKK* genes were retrieved from 53 kinds of *B. distachyon* inbred
376 lines based on their conserved TXY or S/T-X$_5$-S/T domain, respectively, using bioinformatics
377 approaches. Phylogenetic analyses showed that most MAPKs and MKKs clustered into same
378 branch, with the exception of MPK21-2s, which was divided into two groups, designated as type
379 I and II. Further analysis found that the divergence of MPK21-2 may be involved with the
380 presence of UTRs. *MKK10s* expanded during the evolutionary process by ancient tandem
381 duplications with a differential model. This may have resulted in expression differences and
382 functional divergence. We discovered that the expression of the *MPK* and *MKK* gene members
383 varied in different tissues and across abiotic stresses in three different genetic contexts,
384 suggesting that these genes may have diverse biological functions. Taken together, our results

385 revealed a more comprehensive understanding of the function and evolutionary patterns of
386 MAPKs and MKKs in diverse *B. distachyon* inbred lines.

387 **References**

388 Altschul, SF, Madden, TL, Schaffer, AA, Zhang, JH, Zhang, Z, Miller, W, and Lipman, DJ. 1997. Gapped BLAST and
389 PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Research **25**:3389-3402.
390 Asai, T, Tena, G, Plotnikova, J, Willmann, MR, Chiu, WL, Gomez-Gomez, L, Boller, T, Ausubel, FM, and Sheen, J.
391 2002. MAP kinase signalling cascade in Arabidopsis innate immunity. Nature **415**:977-983.
392 Bardwell, L, and Shah, K. 2006. Analysis of mitogen-activated protein kinase activation and interactions with
393 regulators and substrates. Methods **40**:213-223.
394 Brkljacic, J, Grotewold, E, Scholl, R, Mockler, T, Garvin, DF, Vain, P, Brutnell, T, Sibout, R, Bevan, M, Budak, H,
395 Caicedo, AL, Gao, CX, Gu, Y, Hazen, SP, Holt, BF, Hong, SY, Jordan, M, Manzaneda, AJ, Mitchell-Olds, T,
396 Mochida, K, Mur, LAJ, Park, CM, Sedbrook, J, Watt, M, Zheng, SJ, and Vogel, JP. 2011. Brachypodium as a
397 Model for the Grasses: Today and the Future. Plant Physiology **157**:3-13.
398 Catalan, P, Chalhoub, B, Chochois, V, Garvin, DF, Hasterok, R, Manzaneda, AJ, Mur, LAJ, Pecchioni, N, Rasmussen,
399 SK, Vogel, JP, and Voxeur, A. 2014. Update on the genomics and basic biology of Brachypodium
400 International Brachypodium Initiative (IBI). Trends in Plant Science **19**:414-418.
401 Catalan, P, Lopez-Alvarez, D, Bellosta, C, and Villar, L. 2016. Updated taxonomic descriptions, iconography, and
402 habitat preferences of *Brachypodium distachyon*, B-stacei, and B-hybridum (Poaceae). Anales Del Jardin
403 Botanico De Madrid **73**:e028.
404 Chen, LH, Hu, W, Tan, SL, Wang, M, Ma, ZB, Zhou, SY, Deng, XM, Zhang, Y, Huang, C, Yang, GX, and He, GY. 2012.
405 Genome-Wide Identification and Analysis of *MAPK* and *MAPKK* Gene Families in *Brachypodium distachyon*.
406 Plos One **7**:e46744.
407 Cuevas, HE, Zhou, CB, Tang, HB, Khadke, PP, Das, S, Lin, YR, Ge, ZX, Clemente, T, Upadhyaya, HD, Hash, CT, and
408 Paterson, AH. 2016. The Evolution of Photoperiod-Insensitive Flowering in Sorghum, A Genomic Model for
409 Panicoid Grasses. Molecular Biology and Evolution **33**:2417-2428.
410 Dai, ZM, Xiong, YY, and Dai, XH. 2014. Neighboring Genes Show Interchromosomal Colocalization after Their
411 Separation. Molecular Biology and Evolution **31**:1166-1172.
412 Davila-Velderrain, J, Servin-Marquez, A, and Alvarez-Buylla, ER. 2014. Molecular Evolution Constraints in the Floral
413 Organ Specification Gene Regulatory Network Module across 18 Angiosperm Genomes. Molecular Biology
414 and Evolution **31**:560-573.
415 Draper, J, Mur, LAJ, Jenkins, G, Ghosh-Biswas, GC, Bablak, P, Hasterok, R, and Routledge, APM. 2001.
416 *Brachypodium distachyon*. A new model system for functional genomics in grasses. Plant Physiology
417 **127**:1539-1555.
418 Filiz, E, Ozdemir, BS, Budak, F, Vogel, JP, Tuna, M, and Budak, H. 2009. Molecular, morphological, and cytological
419 analysis of diverse *Brachypodium distachyon* inbred lines. Genome **52**:876-890.
420 Freeling, M. 2009. Bias in Plant Gene Content Following Different Sorts of Duplication: Tandem, Whole-Genome,
421 Segmental, or by Transposition. Annual Review of Plant Biology **60**:433-453.
422 Garvin, DF, Gu, YQ, Hasterok, R, Hazen, SP, Jenkins, G, Mockler, TC, Mur, LAJ, and Vogel, JP. 2008. Development of
423 genetic and genomic research resources for *Brachypodium distachyon*, a new model system for grass crop
424 research. Crop Science **48**:S69-S84.

425 Ghanbarian, AT, and Hurst, LD. 2015. Neighboring Genes Show Correlated Evolution in Gene Expression. Molecular
426   Biology and Evolution **32**:1748-1766.
427 Gordon, SP, Contreras-Moreira, B, Woods, DP, Marais, DLD, Burgess, D, Shu, SQ, Stritt, C, Roulin, AC, Schackwitz, W,
428   Tyler, L, Martin, J, Lipzen, A, Dochy, N, Phillips, J, Barry, K, Geuten, K, Budak, H, Juenger, TE, Amasino, R,
429   Caicedo, AL, Goodstein, D, Davidson, P, Mur, LAJ, Figueroa, M, Freeling, M, Catalan, P, and Vogel, JP. 2017.
430   Extensive gene content variation in the *Brachypodium distachyon* pan-genome correlates with population
431   structure. Nature Communications **8**:2184.
432 Gordon, SP, Priest, H, Marais, DLD, Schackwitz, W, Figueroa, M, Martin, J, Bragg, JN, Tyler, L, Lee, CR, Bryant, D,
433   Wang, WQ, Messing, J, Manzaneda, AJ, Barry, K, Garvin, DF, Budak, H, Tuna, M, Mitchell-Olds, T, Pfender,
434   WF, Juenger, TE, Mockler, TC, and Vogel, JP. 2014. Genome diversity in *Brachypodium distachyon*: deep
435   sequencing of highly diverse inbred lines. Plant Journal **79**:361-374.
436 Goyal, RK, Tulpan, D, Chomistek, N, Fundora, DGP, West, C, Ellis, BE, Frick, M, Laroche, A, and Foroud, NA. 2018.
437   Analysis of *MAPK* and *MAPKK* gene families in wheat and related Triticeae species. BMC Genomics **19**:178.
438 Guan, YF, Meng, XZ, Khanna, R, LaMontagne, E, Liu, YD, and Zhang, SQ. 2014. Phosphorylation of a WRKY
439   Transcription Factor by MAPKs Is Required for Pollen Development and Function in Arabidopsis. Plos
440   Genetics **10**:e1004384.
441 Hadiarto, T, Nanmori, T, Matsuoka, D, Iwasaki, T, Sato, K, Fukami, Y, Azuma, T, and Yasuda, T. 2006. Activation of
442   Arabidopsis MAPK kinase kinase (AtMEKK1) and induction of AtMEKK1-AtMEK1 pathway by wounding.
443   Planta **223**:708-713.
444 Hu, KJ. 2006. Intron exclusion and the mystery of intron loss. Febs Letters **580**:6361-6365.
445 Ichimura, K, Shinozaki, K, Tena, G, Sheen, J, Henry, Y, Champion, A, Kreis, M, Zhang, SQ, Hirt, H, Wilson, C, Heberle-
446   Bors, E, Ellis, BE, Morris, PC, Innes, RW, Ecker, JR, Scheel, D, Klessig, DF, Machida, Y, Mundy, J, Ohashi, Y,
447   Walker, JC, and Grp, M. 2002. Mitogen-activated protein kinase cascades in plants: a new nomenclature.
448   Trends in Plant Science **7**:301-308.
449 Jagodzik, P, Tajdel-Zielinska, M, Ciesla, A, Marczak, M, and Ludwikow, A. 2018. Mitogen-Activated Protein Kinase
450   Cascades in Plant Hormone Signaling. Frontiers in Plant Science **9**:1387.
451 Jiang, M, and Chu, ZQ. 2018. Comparative analysis of plant *MKK* gene family reveals novel expansion mechanism of
452   the members and sheds new light on functional conservation. BMC Genomics **19**:407.
453 Jiang, M, Niu, C, Cao, JM, Ni, DA, and Chu, ZQ. 2018. In silico-prediction of protein-protein interactions network
454   about MAPKs and PP2Cs reveals a novel docking site variants in *Brachypodium distachyon*. Scientific
455   Reports **8**:15083.
456 Jiang, M, Wen, F, Cao, JM, Li, P, She, J, and Chu, ZQ. 2015. Genome-wide exploration of the molecular evolution
457   and regulatory network of mitogen-activated protein kinase cascades upon multiple stresses in
458   *Brachypodium distachyon*. BMC Genomics **16**:228.
459 Jiang, M, Zhao, CL, Zhao, MF, Li, YZ, and Wen, GS. 2020. Phylogeny and Evolution of Calcineurin B-Like (CBL) Gene
460   Family in Grass and Functional Analyses of Rice CBLs. Journal of Plant Biology **63**:117-130.
461 Jones, P, Binns, D, Chang, HY, Fraser, M, Li, WZ, McAnulla, C, McWilliam, H, Maslen, J, Mitchell, A, Nuka, G, Pesseat,
462   S, Quinn, AF, Sangrador-Vegas, A, Scheremetjew, M, Yong, SY, Lopez, R, and Hunter, S. 2014. InterProScan
463   5: genome-scale protein function classification. Bioinformatics **30**:1236-1240.
464 Joo, S, Liu, Y, Lueth, A, and Zhang, SQ. 2008. MAPK phosphorylation-induced stabilization of ACS6 protein is
465   mediated by the non-catalytic C-terminal domain, which also contains the cis-determinant for rapid

466   degradation by the 26S proteasome pathway. Plant Journal **54**:129-140.

467   Kolukisaoglu, U, Weinl, S, Blazevic, D, Batistic, O, and Kudla, J. 2004. Calcium sensors and their interacting protein
468   kinases: Genomics of the Arabidopsis and rice CBL-CIPK signaling networks. Plant Physiology **134**:43-58.

469   Lan, X, and Pritchard, JK. 2016. Coregulation of tandem duplicate genes slows evolution of subfunctionalization in
470   mammals. Science **352**:1009-1013.

471   Liu, ZQ, Shi, LP, Liu, YY, Tang, Q, Shen, L, Yang, S, Cai, JS, Yu, HX, Wang, RZ, Wen, JY, Lin, YQ, Hu, J, Liu, CL, Zhang,
472   YW, Mou, SL, and He, SL. 2015. Genome-wide identification and transcriptional expression analysis of
473   mitogen-activated protein kinase and mitogen-activated protein kinase kinase genes in *Capsicum annuum*.
474   Frontiers in Plant Science **6**:780.

475   Lynch, M, and Conery, JS. 2000. The evolutionary fate and consequences of duplicate genes. Science **290**:1151-
476   1155.

477   Makino, T, and McLysaght, A. 2008. Interacting gene clusters and the evolution of the vertebrate immune system.
478   Molecular Biology and Evolution **25**:1855-1862.

479   Makino, T, and McLysaght, A. 2012. Positionally biased gene loss after whole genome duplication: Evidence from
480   human, yeast, and plant. Genome Research **22**:2427-2435.

481   Mitchell, AL, Attwood, TK, Babbitt, PC, Blum, M, Bork, P, Bridge, A, Brown, SD, Chang, HY, El-Gebali, S, Fraser, MI,
482   Gough, J, Haft, DR, Huang, HZ, Letunic, I, Lopez, R, Luciani, A, Madeira, F, Marchler-Bauer, A, Mi, HY,
483   Natale, DA, Necci, M, Nuka, G, Orengo, C, Pandurangan, AP, Paysan-Lafosse, T, Pesseat, S, Potter, SC,
484   Qureshi, MA, Rawlings, ND, Redaschi, N, Richardson, LJ, Rivoire, C, Salazar, GA, Sangrador-Vegas, A, Sigrist,
485   CJA, Sillitoe, I, Sutton, GG, Thanki, N, Thomas, PD, Tosatto, SCE, Yong, SY, and Finn, RD. 2019. InterPro in
486   2019: improving coverage, classification and access to protein sequence annotations. Nucleic Acids
487   Research **47**:D351-D360.

488   Mohanta, TK, Arora, PK, Mohanta, N, Parida, P, and Bae, H. 2015. Identification of new members of the *MAPK* gene
489   family in plants shows diverse conserved domains and novel activation loop variants. BMC Genomics
490   **16**:58.

491   Mur, LAJ, Allainguillaume, J, Catalan, P, Hasterok, R, Jenkins, G, Lesniewska, K, Thomas, I, and Vogel, J. 2011.
492   Exploiting the Brachypodium Tool Box in cereal and grass research. New Phytologist **191**:334-347.

493   Paterson, AH, Freeling, M, Tang, HB, and Wang, XY. 2010. Insights from the Comparison of Plant Genome
494   Sequences. Annual Review of Plant Biology **61**:349-372.

495   Qiao, X, Li, QH, Yin, H, Qi, KJ, Li, LT, Wang, RZ, Zhang, SL, and Paterson, AH. 2019. Gene duplication and evolution in
496   recurring polyploidization-diploidization cycles in plants. Genome Biology **20**:38.

497   Rodriguez, MCS, Petersen, M, and Mundy, J. 2010. Mitogen-Activated Protein Kinase Signaling in Plants. Annual
498   Review of Plant Biology **61**:621-649.

499   Rose, AB, Carter, A, Korf, I, and Kojima, N. 2016. Intron sequences that stimulate gene expression in Arabidopsis.
500   Plant Molecular Biology **92**:337-346.

501   Sancho, R, Cantalapiedra, CP, Lopez-Alvarez, D, Gordon, SP, Vogel, JP, Catalan, P, and Contreras-Moreira, B. 2018.
502   Comparative plastome genomics and phylogenomics of Brachypodium: flowering time signatures,
503   introgression and recombination in recently diverged ecotypes. New Phytologist **218**:1631-1644.

504   Singh, R, Lee, MO, Lee, JE, Choi, J, Park, JH, Kim, EH, Yoo, RH, Cho, JI, Jeon, JS, Rakwal, R, Agrawal, GK, Moon, JS,
505   and Jwa, NS. 2012. Rice Mitogen-Activated Protein Kinase Interactome Analysis Using the Yeast Two-
506   Hybrid System. Plant Physiology **160**:477-487.

507 Soltis, DE, Albert, VA, Leebens-Mack, J, Bell, CD, Paterson, AH, Zheng, CF, Sankoff, D, dePamphilis, CW, Wall, PK,
508   and Soltis, PS. 2009. Polyploidy and Angiosperm Diversification. American Journal of Botany **96**:336-348.

509 Soreng, RJ, Peterson, PM, Romaschenko, K, Davidse, G, Zuloaga, FO, Judziewicz, EJ, Filgueiras, TS, Davis, JI, and
510   Morrone, O. 2015. A worldwide phylogenetic classification of the Poaceae (Gramineae). Journal of
511   Systematics and Evolution **53**:117-137.

512 Srivastava, AK, Lu, YM, Zinta, G, Lang, ZB, and Zhu, JK. 2018. UTR-Dependent Control of Gene Expression in Plants.
513   Trends in Plant Science **23**:248-259.

514 Taft, RJ, Pheasant, M, and Mattick, JS. 2007. The relationship between non-protein-coding DNA and eukaryotic
515   complexity. Bioessays **29**:288-299.

516 Tamura, K, Stecher, G, Peterson, D, Filipski, A, and Kumar, S. 2013. MEGA6: Molecular Evolutionary Genetics
517   Analysis Version 6.0. Molecular Biology and Evolution **30**:2725-2729.

518 Tanoue, T, Adachi, M, Moriguchi, T, and Nishida, E. 2000. A conserved docking motif in MAP kinases common to
519   substrates, activators and regulators. Nature Cell Biology **2**:110-116.

520 Vain, P, Worland, B, Thole, V, McKenzie, N, Opanowicz, M, Fish, LJ, Bevan, MW, and Snape, JW. 2008.
521   Agrobacterium-mediated transformation of the temperate grass *Brachypodium distachyon* (genotype
522   Bd21) for T-DNA insertional mutagenesis. Plant Biotechnology Journal **6**:236-245.

523 Van Bel, M, Diels, T, Vancaester, E, Kreft, L, Botzki, A, Van de Peer, Y, Coppens, F, and Vandepoele, K. 2018. PLAZA
524   4.0: an integrative resource for functional, evolutionary and comparative plant genomics. Nucleic Acids
525   Research **46**:D1190-D1196.

526 Vogel, J, and Hill, T. 2008. High-efficiency Agrobacterium-mediated transformation of *Brachypodium distachyon*
527   inbred line Bd21-3. Plant Cell Reports **27**:471-478.

528 Vogel, JP, Garvin, DF, Mockler, TC, Schmutz, J, Rokhsar, D, Bevan, MW, Barry, K, Lucas, S, Harmon-Smith, M, Lail, K,
529   Tice, H, Grimwood, J, McKenzie, N, Huo, NX, Gu, YQ, Lazo, GR, Anderson, OD, You, FM, Luo, MC, Dvorak, J,
530   Wright, J, Febrer, M, Idziak, D, Hasterok, R, Lindquist, E, Wang, M, Fox, SE, Priest, HD, Filichkin, SA, Givan,
531   SA, Bryant, DW, Chang, JH, Wu, HY, Wu, W, Hsia, AP, Schnable, PS, Kalyanaraman, A, Barbazuk, B, Michael,
532   TP, Hazen, SP, Bragg, JN, Laudencia-Chingcuanco, D, Weng, YQ, Haberer, G, Spannagl, M, Mayer, K, Rattei,
533   T, Mitros, T, Lee, SJ, Rose, JKC, Mueller, LA, York, TL, Wicker, T, Buchmann, JP, Tanskanen, J, Schulman, AH,
534   Gundlach, H, de Oliveira, AC, Maia, LD, Belknap, W, Jiang, N, Lai, JS, Zhu, LC, Ma, JX, Sun, C, Pritham, E,
535   Salse, J, Murat, F, Abrouk, M, Bruggmann, R, Messing, J, Fahlgren, N, Sullivan, CM, Carrington, JC,
536   Chapman, EJ, May, GD, Zhai, JX, Ganssmann, M, Gurazada, SGR, German, M, Meyers, BC, Green, PJ, Tyler,
537   L, Wu, JJ, Thomson, J, Chen, S, Scheller, HV, Harholt, J, Ulvskov, P, Kimbrel, JA, Bartley, LE, Cao, PJ, Jung,
538   KH, Sharma, MK, Vega-Sanchez, M, Ronald, P, Dardick, CD, De Bodt, S, Verelst, W, Inze, D, Heese, M,
539   Schnittger, A, Yang, XH, Kalluri, UC, Tuskan, GA, Hua, ZH, Vierstra, RD, Cui, Y, Ouyang, SH, Sun, QX, Liu, ZY,
540   Yilmaz, A, Grotewold, E, Sibout, R, Hematy, K, Mouille, G, Hofte, H, Pelloux, J, O'Connor, D, Schnable, J,
541   Rowe, S, Harmon, F, Cass, CL, Sedbrook, JC, Byrne, ME, Walsh, S, Higgins, J, Li, PH, Brutnell, T, Unver, T,
542   Budak, H, Belcram, H, Charles, M, Chalhoub, B, Baxter, I, and Initiative, IB. 2010. Genome sequencing and
543   analysis of the model grass *Brachypodium distachyon*. Nature **463**:763-768.

544 Vogel, JP, Tuna, M, Budak, H, Huo, NX, Gu, YQ, and Steinwand, MA. 2009. Development of SSR markers and
545   analysis of diversity in Turkish populations of *Brachypodium distachyon*. BMC Plant Biology **9**:88.

546 Wang, XY, Gowik, U, Tang, HB, Bowers, JE, Westhoff, P, and Paterson, AH. 2009. Comparative genomic analysis of
547   C4 photosynthetic pathway evolution in grasses. Genome Biology **10**:R68.

548   Wei, CJ, Liu, XQ, Long, DP, Guo, Q, Fang, Y, Bian, CK, Zhang, DY, Zeng, QW, Xiang, ZH, and Zhao, AC. 2014.
549   Molecular cloning and expression analysis of mulberry *MAPK* gene family. Plant Physiology and
550   Biochemistry **77**:108-116.
551   Xu, J, and Zhang, SQ. 2015. Mitogen-activated protein kinase cascades in signaling plant growth and development.
552   Trends in Plant Science **20**:56-64.
553   Yang, Q, Han, XM, Gu, JK, Liu, YJ, Yang, MJ, and Zeng, QY. 2019. Functional and structural profiles of *GST* gene
554   family from three Populus species reveal the sequence-function decoupling of orthologous genes. New
555   Phytologist **221**:1060-1073.
556   Yoo, SD, Cho, YH, Tena, G, Xiong, Y, and Sheen, J. 2008. Dual control of nuclear EIN3 by bifurcate MAPK cascades in
557   C2H4 signalling. Nature **451**:789-U781.
558   Zhang, GJ, Li, C, Li, QY, Li, B, Larkin, DM, Lee, C, Storz, JF, Antunes, A, Greenwold, MJ, Meredith, RW, Odeen, A, Cui,
559   J, Zhou, Q, Xu, LH, Pan, HL, Wang, ZJ, Jin, LJ, Zhang, P, Hu, HF, Yang, W, Hu, J, Xiao, J, Yang, ZK, Liu, Y, Xie,
560   QL, Yu, H, Lian, JM, Wen, P, Zhang, F, Li, H, Zeng, YL, Xiong, ZJ, Liu, SP, Zhou, L, Huang, ZY, An, N, Wang, J,
561   Zheng, QM, Xiong, YQ, Wang, GB, Wang, B, Wang, JJ, Fan, Y, da Fonseca, RR, Alfaro-Nunez, A, Schubert, M,
562   Orlando, L, Mourier, T, Howard, JT, Ganapathy, G, Pfenning, A, Whitney, O, Rivas, MV, Hara, E, Smith, J,
563   Farre, M, Narayan, J, Slavov, G, Romanov, MN, Borges, R, Machado, JP, Khan, I, Springer, MS, Gatesy, J,
564   Hoffmann, FG, Opazo, JC, Hastad, O, Sawyer, RH, Kim, H, Kim, KW, Kim, HJ, Cho, S, Li, N, Huang, YH,
565   Bruford, MW, Zhan, XJ, Dixon, A, Bertelsen, MF, Derryberry, E, Warren, W, Wilson, RK, Li, SB, Ray, DA,
566   Green, RE, O'Brien, SJ, Griffin, D, Johnson, WE, Haussler, D, Ryder, OA, Willerslev, E, Graves, GR, Alstrom,
567   P, Fjeldsa, J, Mindell, DP, Edwards, SV, Braun, EL, Rahbek, C, Burt, DW, Houde, P, Zhang, Y, Yang, HM,
568   Wang, J, Jarvis, ED, Gilbert, MTP, Wang, J, and Consortium, AG. 2014. Comparative genomics reveals
569   insights into avian genome evolution and adaptation. Science **346**:1311-1320.
570

Figure 1

Phylogenetic distribution of 799 MPKs among diverse *B. distachyon* inbredlines.

MPKs are divided into four clades (A, B, C and D). Unique to *B. distachyon* MPK17s, an EF-hand CBP domain is found in their C-terminal extensions. The group A, B and C carry TEY T-loops and group D MPKs carry TDY motifs, with the exception of MPK11s which carry an MEY motif. The CD domain is conserved in the MAPK family. The negatively charged amino acids (D and E) are expected to be exposed to the surface of the molecules. MPK7-2 also contained NB-ARC domain and LRR domain. NB: nucleotide-binding; LRR: leucine-rich repeat. ARC: APAF-1, R gene and CED-4.

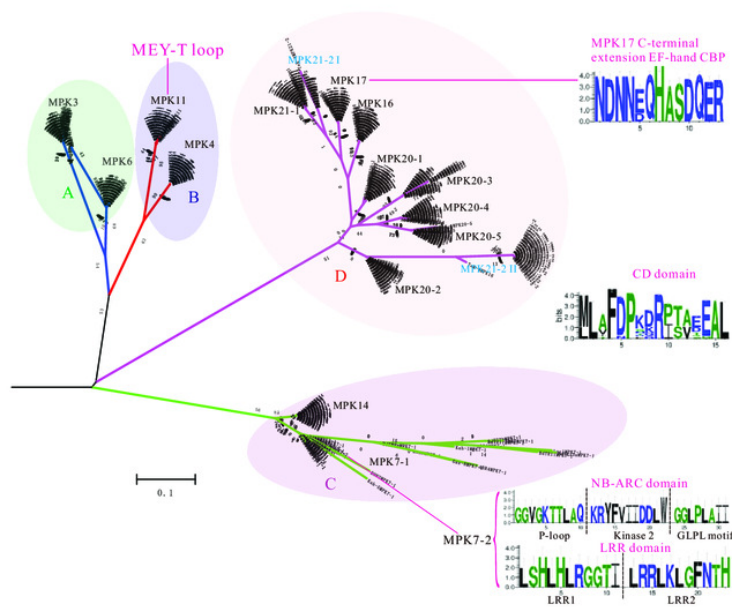



Figure 2

Phylogenetic distribution of 618 MKKs among diverse *B. distachyon* inbred lines

MKKs are divided into four clades (A, B, C and D). Sequence features shown in the form of web logos representing the NTF-2 domain conserved in the C-terminal extensions of MKK3s. Web logos analysis shows amino acid distribution of conserved S/T-X$_5$-S/T motif in MKKs (groups A, B and C) and the docking site conserved in N-terminal extension of MKKs.

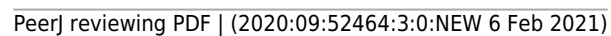



Figure 3

Introns number polymorphisms of MPKs from 53 diverse *B. distachyon* inbred lines

Different colors represent different genomic groups.

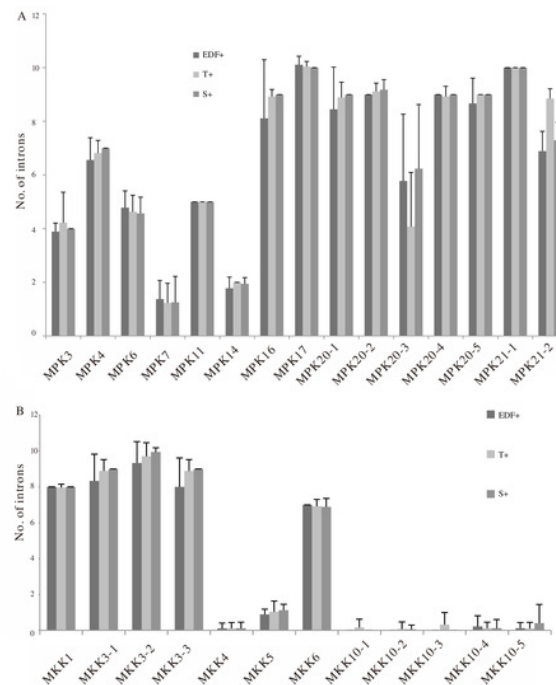



Figure 4

Gene structure of *MPK21-2* genes.

(A) Maximum Likelihood phylogenetic trees of the full CDS sequences of genes encoding
MPK21-2 from diverse *B. distachyon* inbred lines. (B) The exon/intron structure of each
MPK21-2 gene was displayed. Yellow boxes represent exons, gray lines represent introns and
blue boxes represent UTRs. The exons are drawn to scale.

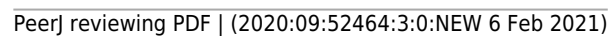



Figure 5

Conservation and consensus pattern of the EF-hand motif of *B. distachyon* inbred lines MPK17s.

(A) Maximum Likelihood phylogenetic trees of the full sequences of genes encoding MPK17s from *B. distachyon* inbred lines. (B) ClustalW multiple-sequence alignment of the region containing the EF-hand motif within MPK17s. Conserved residues are shown in dark colors. (C) Consensus pattern and sequence logo of the EF-hand motif generated using the Weblogo3 application *(* http://weblogo.threeplusone.com/ *)*. The overall height of each stack indicates the sequence conservation at that position (measured in bits), whereas the height of symbols within the stack reflects the relative frequency of the corresponding base at that position. (D) ClustalW multiple-sequence alignment of the region containing the EF-hand motif within MPK17 in different plant species. The specific domain was marked by thick line. At: *Arabidopsis thaliana*; Bd: *Brachypodium distachyon*; Cru: *Capsella rubella*; Cc: *Citrus clememtina*; Cs: *Cucumis sativus*; Eg: *Erythranthe guttata*; Egr: *Eucalyptus grandis*; Gr: *Gossypium raimondii*; Me: *Manihot esculenta*; Os: *Oryza sativa*; Pt: *Populus trichocarpa*; Sb: *Sorghum bicolor*; Si: *Setaria italica*; Tc: *Theobroma cacao*; Zm: *Zea mays*.

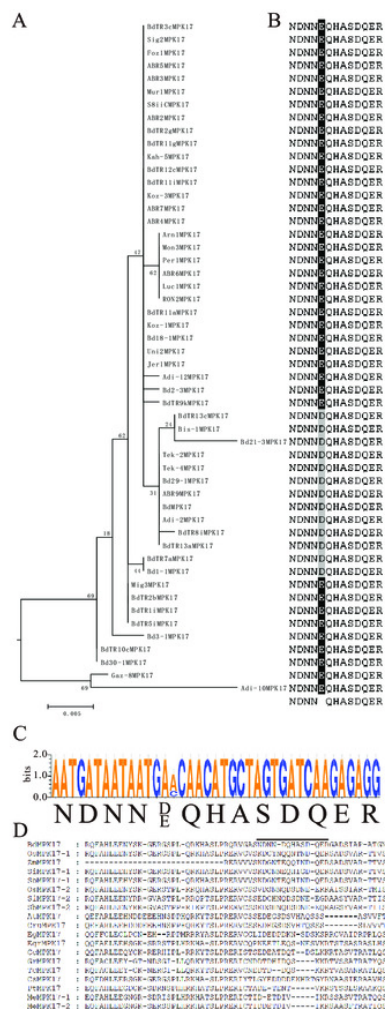



Figure 6

The diversity and evolution of the fate of an ancestral locus having *B. distachyon* inbred lines *MKK10* genes in tandem position.

A. Phylogenetic relationships among 53 diverse *B. distachyon* inbred lines were investigated. The phylogenetic tree is modified from BrachyPan (https://brachypan.jgi.doe.gov/). The variants of ancestral tandem *MKK10* gene clusters in *B. distachyon* inbred lines are shown on the right. The red asterisk indicates the gene cluster models of these inbred lines are same as Bd21, while blue pound represents no tandem duplication events. Gene or protein names: MKK (MAPK kinase 10); DMRL (DMRL synthase); Peptidase (Peptidase_C48); PNN (pinin); ChaC (ChaC-like protein); G-patch (glycine rich nucleic binding domain); ATP-synt (ATP synthase subunit C); RRM (RNA recognition motif protein); UDPGT (UDP-glucoronosyl and UDP-glucosyl transferase); ABC (ABC transporter); PK (Protein tyrosine kinase); AP2 (AP2/EREBP transcription factor); DNA_pol_B (DNA polymerase family B); GCH (Predicted glycine cleavage system H protein); MTL23 (Methyltransferase-like protein 23); Ulp1 (Ulp1 protease); SWIM (SWIM zinc finger); Str (Strictosidine synthase).

Summary of Brachypodium lines phylogeny   tandem MKK10 gene cluster

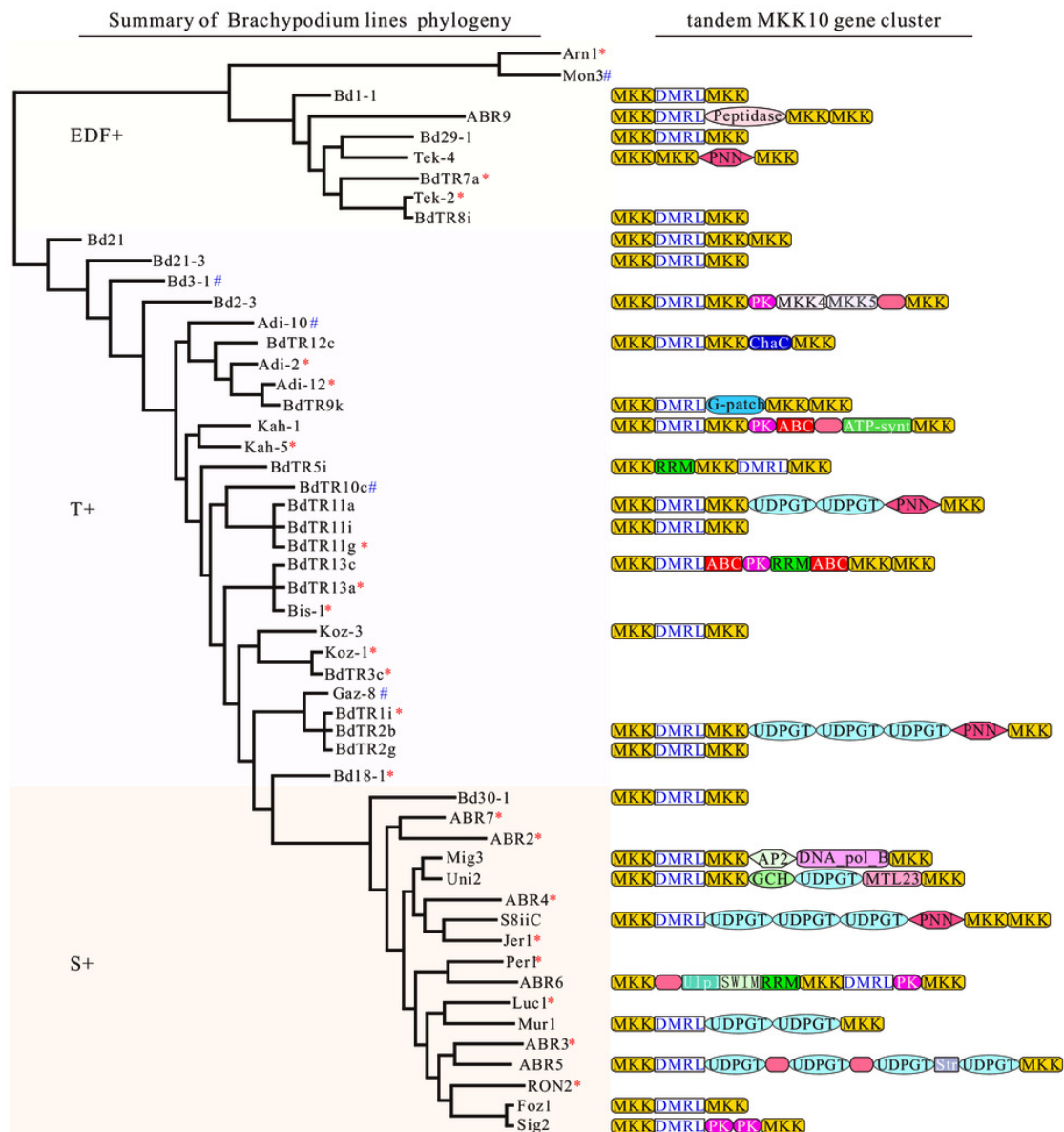



Figure 7

Quantitative expression variation of three diverse *B. distachyon* inbred lines *MPK* and *MKK* genes in five different tissues.

Root, stem, leaf blade and leaf sheath samples were collected at the 8-9 leaf stage of

BdTR8i, Bd21 and Bd30-1 seedling, respectively. While spikelet samples were harvested at *B.*

distachyon flowering one to two weeks according to their different flower time.

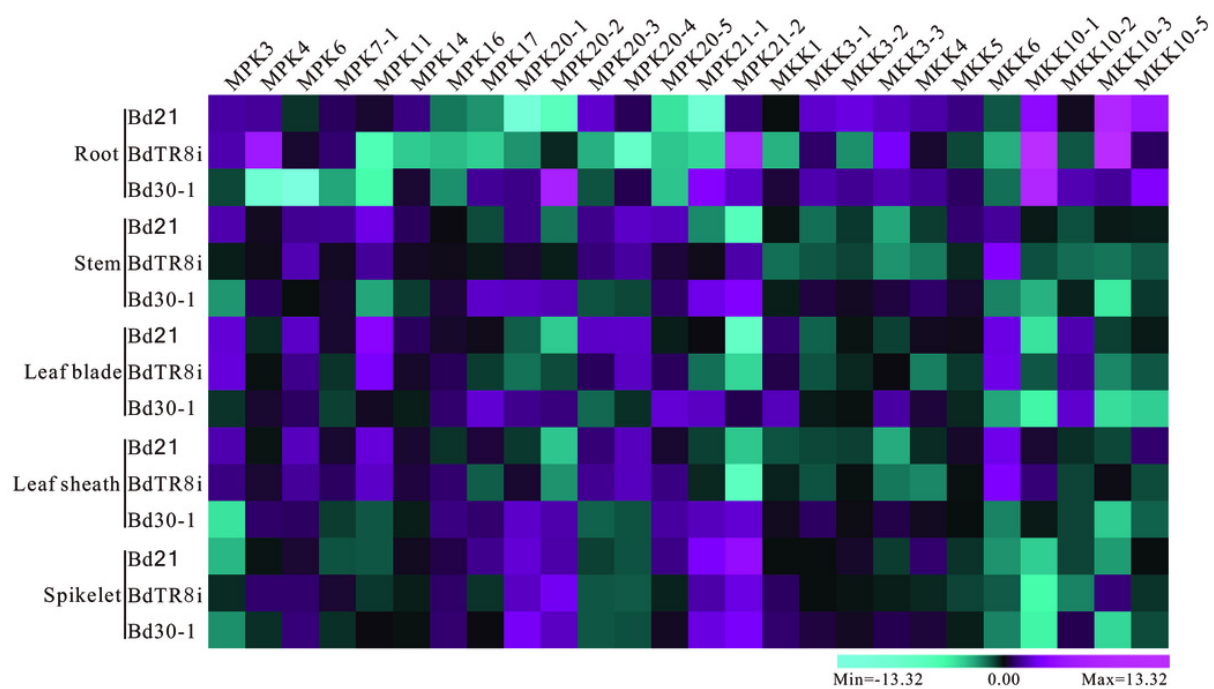
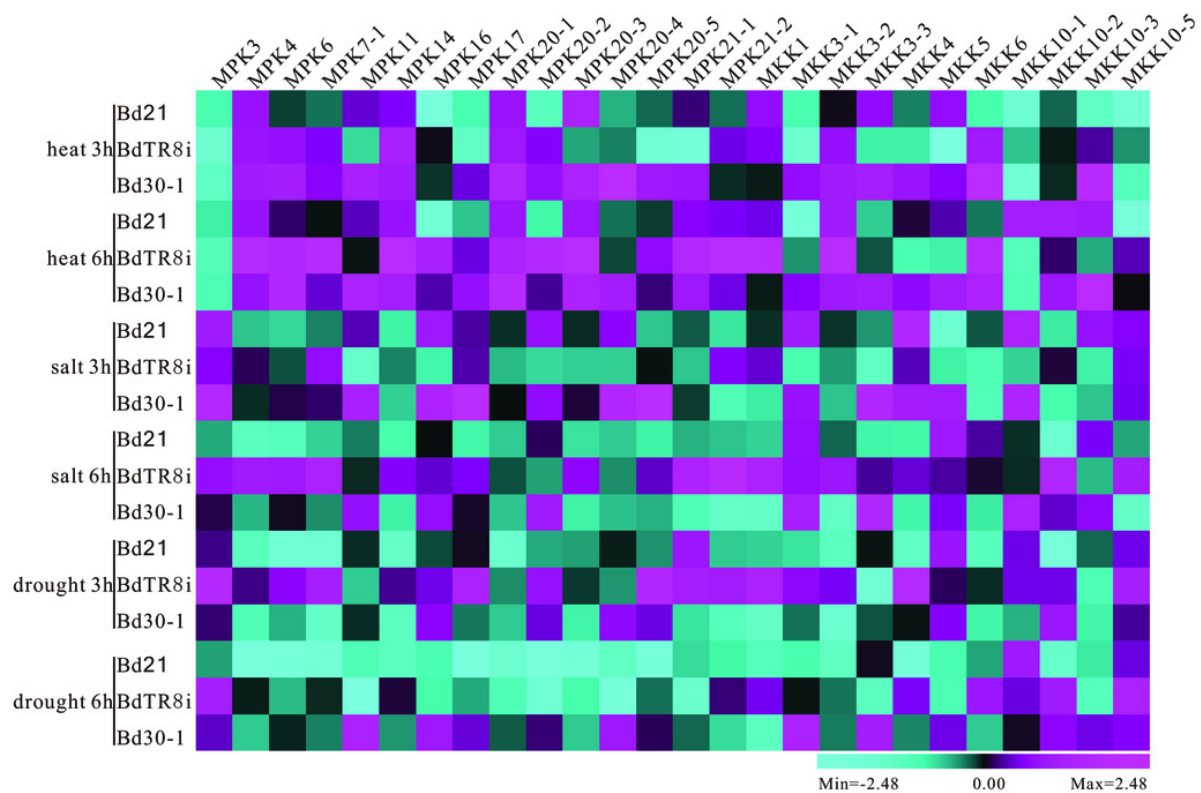



Figure 8

Expression patterns *MPK* and *MKK* genes in Bd21, BdTR8i and Bd30-1 seedlings under different abiotic stresses.

Number of *B. distachyon* inbred lines MPKs and MKKs identified genes from the *BrachyPan* database and their associated information.

1 **Table 1** Number of Brachypodium *MPK* and *MKK* genes identified from the *BrachyPan*

2 database and their associated information.

Genetic groups	Inbred line	Latitude	Longitude	Elevation (m)	Ploidy	MPKs	MKKs
	Arn1	42° 15' 23.44" N	0° 43' 47.46" E	681	-	15	12
	Mon3	41° 39' 4.75" N	0° 12' 37.51" W	515	diploid	15	12
	Bd1-1	39° 11' 27.44" N	27° 36' 28.59" E	141	diploid	14	12
	ABR9	-	-	-	-	15	12
EDF+	Bd29-1	44° 30' 55" N	33° 33' 23" E	260	diploid	15	12
	Tek-4	41° 0' 40.1"N	27° 31' 8.8"E	20	diploid	14	11
	BdTR7a	39° 44'53.45"N	34° 39'1.15"E	1035	diploid	15	12
	Tek-2	41° 0' 40.1"N	27° 31' 8.8"E	20	diploid	15	12
	BdTR8i	37° 6'31.87"N	34° 4'17.06"E	2385	diploid	15	12
	Bd21	33° 45' 39.18" N	44° 24' 11.07" E	42	diploid	16	12
	Bd21-3	33° 45' 39.19" N	44° 24' 11.08" E	43	diploid	14	12
	Bd3-1	33° 45' 39.19" N	44° 24' 11.08" E	43	diploid	14	7
	Bd2-3	33° 45' 39.18" N	44° 24' 11.07" E	42	diploid	14	12
T+	Adi-10	37° 46' 14.5"N	38° 21' 8.2"E	510	diploid	15	10
	BdTR12c	39° 44'53.45"N	34° 39'1.15"E	1035	diploid	15	12
	Adi-2	37° 46' 14.5"N	38° 21' 8.2"E	510	diploid	15	12
	Adi-12	37° 46' 14.5"N	38° 21' 8.2"E	510	diploid	15	12
	BdTR9k	39° 45'10.62"N	30° 47'19.07"E	932	diploid	15	12

	Kah-1	37° 44' 2.3"N	38° 32' 0.2"E	665	diploid	14	12
	Kah-5	37° 44' 2.3"N	38° 32' 0.2"E	665	diploid	15	12
	BdTR5i	40° 23'37.13"N	32° 59'7.32"E	1596	diploid	15	12
	BdTR10c	37° 46'41.64"N	31° 53'5.68"E	1288	diploid	15	12
	BdTR11a	38° 25'0.42"N	28° 1'52.75"E	986	diploid	14	12
	BdTR11i	39° 44'17.39"N	28° 2'24.71"E	363	diploid	15	12
	BdTR11g	41° 25'17.86"N	27° 28'36.81"E	124	diploid	15	12
	BdTR13c	39° 24'46.28"N	32° 59'17.24"E	1192	diploid	15	12
	BdTR13a	39° 45'23.35"N	32° 25'56.46"E	787	diploid	15	12
	Bis-1	37° 52' 35.6"N	41° 0' 54.3"E	529	diploid	15	12
	Koz-3	38° 9' 8.2.6"N	41° 36' 34.8"E	853	diploid	14	12
	Koz-1	38° 9' 8.2.6"N	41° 36' 34.8"E	853	diploid	15	12
	BdTR3c	36° 46'58.92"N	32° 57'46.71"E	1957	diploid	16	12
	Gaz-8	37° 7' 39.8"N	37° 23' 26.9"E	891	diploid	15	5
	BdTR1i	38° 5'35.03"N	28° 34'59.02"E	841	diploid	15	12
	BdTR2b	40° 4'55.55"N	31° 19'52.01"E	667	diploid	15	12
	BdTR2g	40° 23'37.13"N	32° 59'7.32"E	1596	diploid	15	12
	Bd18-1	39° 22' 4.25" N	33° 43' 48.91" E	1101	diploid	16	12
	Bd30-1	36° 59' 25.76" N	3° 33' 31.44" W	1220	diploid	15	12
	ABR5	42° 34' 23.45" N	0° 33' 49.39" W	828	diploid	15	11
S+	Mig3	42° 8' 52.76" N	0° 11' 41.89" W	572	diploid	15	12
	Uni2	42° 7' 3.98" N	0° 26' 42.81" W	480	diploid	15	12
	Mur1	42° 06' 18" N	0° 51' 23" E	487	diploid	14	12

Foz1	42º 38' 11.44" N	1º 18' 17.42" W	434	diploid	15	11
ABR2	43° 36' 15.343" N	3° 15' 46.580" E	371	diploid	14	12
ABR3	42° 10' 49.8" N	0° 4' 23.2" W	1928	diploid	14	12
ABR4	42° 15' 45.54" N	0° 43' 0.48" E	480	-	15	12
ABR6	42° 34' 27.48" N	2° 11' 5.39" W	484	-	15	12
ABR7	41° 35' 23.86" N	4° 45' 24.26" W	725	-	14	12
S8iiC	41° 36' 19.3" N	0° 08' 38.4" E	144	-	15	12
Jer1	42° 3' 16.56" N	0° 0' 44.57" W	418	-	14	11
Per1	42° 44' 13.34" N	1° 44' 58.6" W	742	-	14	12
Luc1	42° 36' 36.18" N	0° 53' 35.48" W	597	-	15	12
RON2	42° 46' 50" N	0° 57' 48" W	594	-	15	12
Sig2	42° 36' 46.55" N	1° 0' 52.38" W	524	-	14	12