

```

Ha  MALGFSDRNTE MSPYSQOSNTSAVPMDS SKKRKTRSRKEAPEKVAATLAKWIEYNKANE - 59
Dv  -----MTNSDQPSSTPSSSMDS PRKRKRSRSDGPKGVAETLAKWKEYNNKIDS 49
Le  -----MAIMDEAANMVCVPLDY SRKRKRSRSDRTKNVEETLAKWKEYNEKLDN 49
Cr  -----MALLDQASNLSMPDFDTRKRKSR-RRDGTKNVAETLAKWKEYNEKLEDA 48
At  -----MAVYDQSGDRNRTQIDT SRKRKRSRSGDGT-VAERLKRWKEYNETVEE 48

                                     • •
Ha  -DGKAK-TRKAPAKGSKKGC MKGKGGPENSRTNFRGVQRQTWGWVVAEIREPNRGRRLWL 117
Dv  LDEKAKPARKVPAKGSKKGCMKGKGGPENSRCFRGVQRQTWGWVVAEIREPNRGSRLWL 109
Le  -EGKGPVVRKVPAGSKKGC MKGKGGPENWRCKYRGVQRQTWGWVVAEIREPKRGSRLWL 108
Cr  -LDGGKPAKVPAGSKKGC MKGKGGPENSHCKYRGVQRQTWGWVVAEIREPNRGSRLWL 107
At  ---VSTKKRKVPAGSKKGC MKGKGGPENSRC SFRGVQRQTWGWVVAEIREPNRGSRLWL 105

                                     ▼ ▼
Ha  GTFGSAVEAALAYDEAARAMYGTSARLNLPNC RSKNDILPLLVTN-PASSCDSTTTCSYS 176
Dv  GTFGSAVEAALAYDEAARVMYGS CARLNLPNC RSMNEYSQMVPVNGSASSCDSTTTCSHS 169
Le  GTFGTAIEAALAYDDAARAMYGPCARLNLPNYACDSVSWATTSSAS--ASASDCTVASGFG 166
Cr  GTFRNAIEAALAYDEAARAMYGPCARLNLPNYRASEE--SSSLPT--TSGSDTTTASGIS 163
At  GTFPTAQEAASAYDEAAKAMYGPLARLNFRSDASEVTSTSSQSE--VCTVETPGCVHVK 163

Ha  EVDATHDS-----KPGPAVFPSVKHEESVQVKHEP--- 206
Dv  LEDSKASSS-----RQMIKQDEGDSKREDTIVKHEVKEE 203
Le  EVCPVDG-----ALHEADTPLSSVKDEGTAM 192
Cr  EVSVYEDKKFTP-----VVSGLKQDDKGESLESADSKPQLLDAGTPMSAVKEEPKEY 216
At  TEDPDCESKPFSGGVEPMYCLENGAEEMKRGVKADKHWLSEFEHNYWSDILKEKEKQKEQ 223

Ha  -----EIVAKEEHCVDN-----NDLGFDIGD-- 227
Dv  PVEEFKSIRIKDEPVEITN-----EDISLPHVN-- 230
Le  DIVEPTS---IDEDTLKS-----GWDCLDKLNMD-- 218
Cr  QVMDSQSEGGFGDEEPPSKLVCKEVD FGDQAVVPAVKNAEEMGGEMGGDILKGCSSL-- 274
At  GIVETCQQQQQDSL SVADYG-----WPNDVDQSHLDSS 256

Ha  EMFDLEELLGAVEDSNPE-----AGSGDG 251
Dv  EMFDMEELLEMMGQRKSQNTGNRMGQNGFDEPSGVSQPWIKIEPQSEYQQDTYGTGGGFG 290
Le  EMFDVDELLAMLDSTPVFTK-----DYNSDG 244
Cr  EMFDVDELLSVLDSTPLHAS-----DFQHGM 300
At  DMFDVDELLRDLNGDDVFAG-----LNQDR 281

Ha  YDGRFVN-----CADANENMQMYMQP--DPCVSG--QDYSEDFLQPGRPEDCS 295
Dv  FDFSMPKQPAMDQKLGDMGWFDQEPMPNYTRPKWEDKCDVGAVADYGFDFLMPGRPEDEGN 350
Le  KHNNMVSDS-----QCQEP-----NAVVDPMTVDYGFDFLQPGRQEDLN 283
Cr  GNGNVKAEAAANYAPSWDSAFQLQNDKLGSGQ-QHMAQTPPEINSGTDFLQPGRQEDSY 359
At  YPGNSVAN-----CSYRPE SQQS-----GFDPLQSLNYGTPPFQLEGKDCNG 323

Ha  FTLEELG-LGLDAELDLYSN 314
Dv  FTLQELG-FDLGADFGF--- 366
Le  FSSDDLAFIDLDSELVV--- 300
Cr  FTLGDLDFLDLGAELGL--- 376
At  F-FDDL SYLDLEN----- 335

```

Additional File 6 The predicted amino acid sequence of HaDREB2 shows unique features that are conserved in some DREB2 factors, but not in AtDREB2A.

Alignment, using the Clustal X program [Larkin et al., Bioinformatics 23(21): 2947-2948 (2007)], of the predicted HaDREB2 protein (Ha) with other subgroup A-2 DREB proteins [9]: DvDREB2A (ABR23508, Dv), LeDREB1 (AAN77051, Le), CrORCA1 (CAB93939, Cr), and AtDREB2A (At). The highly conserved AP2 domain is underlined. Within the AP2 domain, the Valine-14 and Aspartic acid-19 residues characteristic of DREB proteins [9] are indicated by dots. Some AP2 domain residues conserved in the HaDREB2, DvDREB2A, LeDREB1, and CrORCA1 proteins, but not in AtDREB2A, are indicated by filled triangles. On the AtDREB2A protein, and next to AP2 domain, the sequences (amino acids 136 to 165 [18]) involved in its instability are boxed. These sequences are not present in HaDREB2, DvDREB2A, LeDREB1, or CrORCA1 proteins. The carboxyl-

terminal region is less conserved and enriched in acidic residues that could participate in transcriptional activation [10, 18]. On the alignment we indicate a carboxyl-terminal motif that is conserved in all five proteins (box with thin line). This motif is included in the sequences required for transcriptional activation by AtDREB2A in protoplast assays; the 81-amino acid, terminal, fragment between positions 254 and 335 in AtDREB2A [18]. Similarly, transcription activation by ZmDREB2A required the carboxyl-terminal 83 amino acids. In this case the necessary fragment included amino acid motifs specific for monocot DREB2 proteins [14]. We also indicate a second motif conserved in the HaDREB2, DvDREB2A, LeDREB1, and CrORCA1 proteins, but not in AtDREB2A (box with thick line). Therefore, DREB factors from different plant groups could use dissimilar motifs for transcriptional activation; these motifs would be located in the carboxyl-terminal 80-90 amino acids. The conserved amino acid residues are highlighted with black (when identical) or gray (when similar) background. Dashes show gaps in the amino acid sequences introduced to optimize alignment.