Genome-wide analysis, identification, evolution and genomic organization of Dehydration Responsive Element Binding (DREB) gene family in *Solanum tuberosum*

Nida Mushtaq [1] , **Faiza Munir** [Corresp., 1] , **Alvina Gul** [1] , **Rabia Amir** [1] , **Rehan Zafar Paracha** [2]

[1] Department of Plant Biotechnology, Atta-ur-Rahman School of Applied Biosciences, National University of Sciences and Technology, Islamabad, Pakistan

[2] Research Centre for Modelling & Simulation, National University of Sciences and Technology, Islamabad, Pakistan

Corresponding Author: Faiza Munir
Email address: faiza.munir@asab.nust.edu.pk

Background: Dehydration Responsive Element Binding (DREB) gene family plays a crucial role as transcription regulators and enhances plant tolerance to abiotic stresses. Although the DREB gene family has been identified and characterized in many plants, knowledge about it in *Solanum tuberosum* (Potato) is limited. **Results:** In the present study, StDREB gene family was comprehensively analyzed using bioinformatics approaches. We identified 66 StDREB genes through genome wide screening of Potato genome based on the AP2 domain architecture and amino acid conservation analysis (Valine at position 14th). Phylogenetic analysis divided them into six distinct subgroups (A1-A6). The categorization of StDREB genes into six subgroups was further supported by gene structure and conserved motif analysis. Potato DREB genes were found to be distributed unevenly across 12 chromosomes. Gene duplication proved that StDREB genes experienced tandem and segmental duplication events which led to the expansion of the gene family. The Ka/Ks ratios of the orthologous pairs also demonstrated the StDREB genes were under strong purification selection in the course of evolution. Interspecies synteny analysis revealed 45 and 36 StDREB genes were orthologous to *Arabidopsis* and *Solanum lycopersicum*, respectively. Moreover, subcellular localization indicated that StDREB genes were predominantly located within the nucleus and the StDREB family's major function was DNA binding according to gene ontology (GO) annotation. **Conclusions:** This study provides a comprehensive and systematic understanding of precise molecular mechanism and functional characterization of StDREB genes in abiotic stress responses and will lead to improvement in *Solanum tuberosum*.

1 **Genome-wide analysis, identification, evolution and genomic organization of Dehydration**
2 **Responsive Element Binding (DREB) gene family in *Solanum tuberosum***

4 **Nida Mushtaq[1], Faiza Munir[1*], Alvina Gul[1], Rabia Amir[1], Rehan Zafar Paracha[2]**

5 [1]Department of Plant Biotechnology, Atta-ur-Rahman School of Applied Biosciences,
6 National University of Sciences and Technology, Islamabad, Pakistan

7 [2]Research Centre for Modelling & Simulation, National University of Sciences and
8 Technology Islamabad, Pakistan.

9 ***Corresponding author's**
10 **Faiza Munir**
11 **E-mail:** faiza.munir@asab.nust.edu.pk
12 Telephone: +92-51-90856146

31 **Abstract**

32 **Background:** Dehydration Responsive Element Binding (DREB) gene family plays a crucial
33 role as transcription regulators and enhances plant tolerance to abiotic stresses. Although the
34 DREB gene family has been identified and characterized in many plants, knowledge about it in
35 *Solanum tuberosum* (Potato) is limited.

36 **Results:** In the present study, StDREB gene family was comprehensively analyzed using
37 bioinformatics approaches. We identified 66 StDREB genes through genome wide screening of
38 Potato genome based on the AP2 domain architecture and amino acid conservation analysis
39 (Valine at position 14th). Phylogenetic analysis divided them into six distinct subgroups (A1-A6).
40 The categorization of StDREB genes into six subgroups was further supported by gene structure
41 and conserved motif analysis. Potato DREB genes were found to be distributed unevenly across
42 12 chromosomes. Gene duplication proved that StDREB genes experienced tandem and
43 segmental duplication events which led to the expansion of the gene family. The Ka/Ks ratios of
44 the orthologous pairs also demonstrated the StDREB genes were under strong purification
45 selection in the course of evolution. Interspecies synteny analysis revealed 45 and 36 StDREB
46 genes were orthologous to *Arabidopsis* and *Solanum lycopersicum*, respectively. Moreover,
47 subcellular localization indicated that StDREB genes were predominantly located within the
48 nucleus and the StDREB family's major function was DNA binding according to gene ontology
49 (GO) annotation.

50 **Conclusions:** This study provides a comprehensive and systematic understanding of precise
51 molecular mechanism and functional characterization of StDREB genes in abiotic stress
52 responses and will lead to improvement in *Solanum tuberosum*.

53 **Keywords:** DREB, *Solanum tuberosum*, Bioinformatics, Phylogeny, Gene Duplication, GO
54 annotation

55 **Abbreviations:** DRE, dehydration responsive element; CRT, C- repeat; AP2, APETALA2

56

57

58

59

60

61

62

63

64

INTRODUCTION

Dehydration responsive element binding (DREB) is considered as one of the largest and best studied gene family involved in the abiotic stress mitigation by regulating the expression of genes involved in ABA-independent stress tolerance pathway (Lata & Prasad 2011). DREB belongs to the AP2 multigene family and possess a single conserved AP2 domain (Dietz et al. 2010). The DREB gene family is characterized by the presence of valine (V) at position 14th and glutamic acid (E) at position 19th respectively within the conserved AP2 domain (Zhou et al. 2010). The AP2 domain consists of 60-70 conserved amino acid residues, indispensable for plant's stress and defense response mechanisms (Wu et al. 2015). The AP2 domain comprises seven key amino acids that are necessary for DRE binding: one V residue, four R residues, and two W residues (Allen et al. 1998). In addition, a conserved Serine/Threonine rich region is located adjacent to the AP2 domain, which is responsible for phosphorylation of the DREB genes (Shen et al. 2003). Based on their structural characteristics, DREB genes can be further divided into six subgroups: A1-A6 (Zhou et al. 2010). DREB transcription factors are one of the most promising regulons for abiotic stress tolerance in plants that directly interact with DRE/CRT, which consist of core motifs ACCGAC/GCCGAC (Sharoni et al. 2011; Vazquez-Hernandez et al. 2017). These sequences are found in the promoter regions of drought and cold responsive genes for example KIN1 and KIN2 (cold responsive) (Kimotho et al. 2019), RD29A (drought responsive) (Zhao et al. 2013), LEA (Late Embryogenesis Abundant) (Liang et al. 2019), COR15A and COR15B (cold responsive) (Yu et al. 2018).

DREB gene family members are widely involved in abiotic stress responses such as drought, salt, cold, and heat (Du et al. 2018; Kudo et al. 2017; Wu et al. 2018). DREB genes were first identified in *Arabidopsis*. AtDREB1 and AtDREB2 can function as two independent proteins in two distinct signal transduction pathways under cold and dehydration stress, respectively (Liu et al. 1998). In *Arabidopsis*, studies have shown that ABA signaling does not normally mediate the expression of the A1 and A2 subgroups. On the contrary, ABI4, the only member of A3 subgroup is involved in sugar and ABA signaling (Shkolnik-Inbar & Bar-Zvi 2011). A4 subgroup includes the most studied members namely TINY which functions in drought tolerance and HARDY which participates in both salinity and drought tolerance (Xie et al. 2019). Drought and cold stress induce RAP2.1, member of A5 subgroup, in *Arabidopsis* (Dong & Liu 2010). RAP2.4 and RAP2.4B genes in the A6 subgroup are responsive to drought, heat and salinity stress respectively (Rae et al. 2011). DREB genes are reported to impart enhanced drought tolerance in many plant species. Overexpression of DREB1A led to heightened drought tolerance in *Arabidopsis* and tobacco. A similar phenotype is reported in wheat using the RD29A promoter. Likewise, DREB gene from cotton conferred enhanced tolerance to drought, heat and cold stress when expressed in wheat (Rabara et al. 2014). Constitutive overexpression of CsDREB gene led to enhanced drought and salt tolerance in transgenic *Arabidopsis* (Wang et al. 2017).

Subsequently, a number of DREB genes have been identified and characterized in a wide variety of plant species such as *Arabidopsis* (Hwang et al. 2012), rice (Dubouzet et al. 2003), bell pepper (Hong & Kim 2005), soybean (Chen et al. 2007), pearl millet (Agarwal et al. 2007), wheat (Egawa et al. 2006; Lucas et al. 2011), maize (Feng et al. 2003; Liu et al. 2013), chrysanthemum (Yang et al. 2009), tomato (Guo & Wang 2011; Hichri et al. 2016), lettuce (Park et al. 2020). This depicts their involvement in different abiotic stress responses. For instance, transgenic expression of DREB3a from a forage grass *Leymus chinensis* led to improved drought and salt

110   tolerance in *Arabidopsis* (Xianjun et al. 2011). Another study reported enhanced drought and
111   freezing stress tolerance by constitutive overexpression of *Arabidopsis* DREB1B gene in
112   transgenic potato (Movahedi et al. 2012). Functional analysis of *Broussonetia papyrifera*
113   DREB2 gene suggested its participation in drought and salinity stress responses. Furthermore,
114   transgenic expression of BpDREB2 gene in *Arabidopsis* demonstrated enhanced salt and cold
115   stress (Sun et al. 2014). Studies on *Medicago truncatula* revealed the role of DREB gene
116   members both in freezing and cold stress (Shu et al. 2016). A novel DREB gene has been
117   reported in maize namely ZmDBF3 which exhibits positive regulation relationship under
118   salinity, drought, heat and cold stress (Zhou et al. 2016).

119   The distribution of DREB genes varies greatly between plants. Polyploidization is thought to
120   have a major contribution in genome evolution and plant diversity (Wendel et al. 2016). Gene
121   duplication is considered as a crucial mechanism in the evolutionary history of plants. Plant
122   diversification is greatly aided by gene duplication events, which results in the generation of
123   novel genes necessary for plant evolution (McKain et al. 2016). According to a recent study,
124   polyploidization has played vital roles in the expansion of the DREB gene family specifically for
125   plants that have undergone recent whole genome duplication (WGD) events such as t for
126   Commelinid- specific and r for Poaceae (Wang et al. 2019). The DREB gene family has
127   expanded significantly, allowing for more extensive and complex functional distinction. For
128   instance, 30 DREB genes have been characterized in mung bean, five of which are highly
129   expressed under drought stress (Labbo et al. 2018). In barley, 41 DREB genes have been
130   discovered, many of which are expressed in drought and salt stress (Guo et al. 2016). Studies on
131   *Musa acuminata* and *Musa balbisiana* revealed 81 and 99 DREB genes, respectively (Lakhwani
132   et al. 2016). A large number of DREB genes (210) haven been identified in wheat genome.
133   Over-expression of TaDREB3-AI displayed enhanced tolerance to heat, dehydration, and salinity
134   stresses (Niu et al. 2020).

135   Potato, which originated from the Andean regions of Bolivia and Peru (Davies Jr et al. 2005), is
136   the third most significant agricultural crop worldwide after wheat and rice. Potato is highly
137   adaptable to a wide range of ecosystems. According to FAO, over 388 million tons of potatoes
138   were produced annually with consumption of over 239 million tons in 2017 (Handayani et al.
139   2019). Being a wholesome food, potato is rich in vitamins, minerals and complex carbohydrates
140   (Hussain 2016). However, potato is highly sensitive to various types of stress thus affecting its
141   sustainable production. The key factors that severely affect potato's yield are abiotic stresses,
142   including drought, low temperature, heat and salinity (Dahal et al. 2019). In previous studies,
143   characterization of two DREB genes (StDREB1 and StDREB2) showed a remarkable increase in
144   their expression by salinity stress. The ectopic expression of these genes in potato conferred
145   enhanced tolerance to salinity stress in transgenic potato lines (Bouaziz et al. 2015b; Bouaziz et
146   al. 2012). Furthermore, overexpression of StDREB1 and StDREB2 led to enhanced drought
147   tolerance in transgenic potato lines (Bouaziz et al. 2015a). In addition, overexpression of DREB
148   transcription factors imparted increased cadmium (Cd) stress tolerance in transgenic potato
149   cultivars (Charfeddine et al. 2017). Despite overexpression, little is known about characterization
150   of the DREB gene family in *S. tuberosum*. Given the vital role that DREB genes play in plant's
151   abiotic stress tolerance, it is essential to identify and study the DREB gene family in the potato
152   genome.

153   In the present study, we identified all potential DREB genes encoded in the *Solanum tuberosum*
154   genome. Further, we performed bioinformatics analysis for classification of DREB genes into
155   different subgroups, presence of characteristic motifs, exon/intron organization, chromosomal
156   distribution, and gene duplication events. Finally, we analyzed homology of DREB genes with
157   *A. thaliana* and *S. lycopersicum*, functional diversity, and subcellular localization of StDREB
158   genes. The results gained herein will provide useful insights for future studies on functional
159   characterization of DREBs in Potato.

MATERIALS AND METHODS

Retrieval and Identification of DREB Genes in *Solanum tuberosum* Genome

162   To perform a comprehensive identification of the DREB gene family members in *S. tuberosum*,
163   the amino acid sequences of *Arabidopsis* DREB proteins were retrieved from TAIR database
164   (https://www.arabidopsis.org/) (Huala et al. 2001) and used as queries for BLASTp homology
165   search against the Potato genome v4.03 with an e-value of 1×10^{-5} in Sol Genomics Network
166   (SGN) (https://solgenomics.net/) (Mueller et al. 2005) and Phytozome v12.1 database
167   (https://phytozome.jgi.doe.gov/pz/portal.html) (Goodstein et al. 2012), respectively. Since DREB
168   gene family consists of only one conserved AP2 domain, all retrieved amino acid sequences of
169   potato were scanned for the presence of AP2 domain using PFAM (http://pfam.xfam.org/)
170   (Mistry et al. 2021) and SMART (http://smart.embl-heidelberg.de/) (Letunic & Bork 2018)
171   programs. Furthermore, amino acid conservation analysis was performed by alignment of the
172   AP2 domain for each filtered DREB protein using MEGA X. The candidate protein sequences
173   were validated for the presence of valine (V) at position 14th and glutamic acid (E) at position
174   19th, especially valine, which has been a characteristic feature for DREB gene selection. Amino
175   acid positions were also assessed by ScanProsite server (https://prosite.expasy.org/scanprosite/)
176   (Hulo et al. 2006). Additionally, physiochemical properties of DREB proteins such as protein
177   length, molecular weight (kDa) and isoelectric point (pI) were computed by ExPASy ProtParam
178   tool (https://web.expasy.org/protparam/) (Gasteiger et al. 2003).

Phylogenetic Analysis StDREB proteins

180   Full length amino acid sequences of AtDREB and StDREB proteins were aligned in order to
181   construct the phylogenetic trees. Multiple sequence alignment (MSA) was executed by
182   MUSCLE algorithm with default parameters. Following alignment, phylogenetic trees were
183   constructed by neighbor-joining (NJ) method based on pairwise deletion and Poisson substitution
184   model with 1000 bootstrap replicates using MEGA X software (http://www.megasoftware.net)
185   (Kumar et al. 2018). Based on the classification schemes for AtDREB proteins, recently
186   identified StDREB proteins were characterized into distinct subgroups.

Gene Structure and Protein Motif Features

188   Genomic and complete coding sequences (CDS) of each StDREB gene were downloaded from
189   Phytozome v12.1 to analyze exon-intron structures. Gene Structure Display Server (GSDS 2.0)
190   (http://gsds.gao-lab.org/) was utilized for a detailed graphical illustration of the exon-intron
191   organization by comparing CDS sequences of the StDREB genes with their respective genomic
192   sequences (Guo et al. 2007). The MEME version 5.3.2
193   (https://memesuite.org/meme/tools/meme) was employed to detect conserved motifs in *S.*
194   *tuberosum* DREB proteins with the following parameters: (i) an optimal motif width of 6-50

195 amino acids, (ii) zero or one occurrence per sequence, and (iii) a maximum number of motifs set
196 to 15 (Bailey et al. 2015).

197 **In Silico Chromosomal Mapping, Gene Duplication Events and Synteny Analysis**
198 Chromosomal position information of each StDREB gene was determined from Phytozome
199 v12.1 database. Physical locations and relative distances of the StDREB genes were mapped on
200 to their respective potato chromosomes using MapChart 2.2 Software
201 (https://www.wur.nl/en/show/Mapchart.htm) (Voorrips 2002). To analyze gene duplication
202 events two mechanisms of gene expansion were considered: tandem duplications and segmental
203 duplications. Two gene pairs situated on same chromosomal fragment and segregated by five or
204 fewer gene loci were regarded as tandem duplications. To assess the effect of selective pressure
205 and divergence time of StDREB genes, the Ka (non-synonymous) and Ks (synonyms) values
206 were computed using ToolKit Biologists Tools (TB tools) software (https://github.com/CJ-
207 Chen/TBtools) (Chen et al. 2020). The approximate divergence time was estimated using the
208 formula $T=Ks/2x*MYA$, where $x=6.56 \times 10^{-9}$ and $MYA = 10^{-6}$ (He et al. 2016). For synteny
209 analysis, both the genomic and gff3 annotation files of *S. tuberosum*, *A. thaliana* and *S.
210 lycopersicum* were extracted. Circos Plot and synteny images were constructed and visualized
211 through advanced Circos and dual synteny plotter software (https://github.com/CJ-
212 Chen/TBtools) to examine segmental duplication gene pairs and orthologous gene conservation
213 of StDREBs with other plants species respectively.

214 **Gene Ontology (GO) Annotation and Subcellular Localization Prediction**
215 Gene ontology annotation analysis of StDREB genes was conducted through Blast2Go software
216 (https://www.blast2go.com/) (Conesa & Götz 2008). Transcript sequence of each StDREB gene
217 was uploaded to this software to determine biological processes (BP), cellular compartments
218 (CC) and, molecular functions (MF). To execute Blast2Go functional annotation BLASTx
219 search, InterPro Scan, mapping and annotation were carried out with default settings. The WoLF
220 PSORT tool (https://wolfpsort.hgc.jp/) (Horton et al. 2007) and CELLO Online server v.2.5
221 (http://cello.life.nctu.edu.tw/) were used to predict the subcellular localization of DREB genes.

222 ## RESULTS

223 **Identification and Characterization of Putative DREB Genes in *S. tuberosum***
224 **Genome**
225 To identify all potential DREB gene family members in Potato, DREB protein sequences from
226 the model plant *Arabidopsis thaliana* were used as queries in BLASTp homology search against
227 *S. tuberosum* genome v4.03. A total of 66 DREB genes designated as StDREB1 to StDREB66
228 were identified in Potato by removing other redundant hits with insignificant e-value and identity
229 percentage. Based on Phytozome v12.1 annotation, all of the StDREB genes belonged to
230 dehydration responsive element binding protein and ethylene-response factor **(Data S1)**.
231 Detailed information of all the StDREB genes aligned with their respective *Arabidopsis* ortholog
232 were predicted **(Data S2)**. Later, StDREB genes were verified by using SMART and Pfam
233 analysis in order to confirm the presence of one conserved AP2 domain as shown in **Table 1**
234 followed by amino acid conservation analysis. Valine has been indicated as the most significant
235 amino acid for binding affinity whereas glutamic acid might have some flexibility among
236 proteins. Among 66 StDREB proteins, 26 sequences presented only valine residue at the position

237 14th, while 40 sequences exhibited both V and E at position 14th and 19th within the conserved
238 AP2 domain, respectively. StDREB1 to StDREB66 genes were subjected to further analysis in
239 order to assess their physiochemical characteristics. Analysis depicted that the amino acid
240 lengths of 66 StDREB proteins varied greatly from 72 a.a residues (StDREB51) to 457 a.a
241 residues (StDREB31) with different range of molecular weights from 8.3 kDa to 52.29 kDa and
242 isoelectric points (pI) ranged from 4.17 (StDREB59) to 9.97 (StDREB51), respectively as
243 provided in **Data S2**.

Phylogeny and Group Classification of StDREB Genes

245 To investigate the evolutionary relationship and sequence homology between *Arabidopsis* and *S.*
246 *tuberosum* DREB protein sequences, a Neighbor- Joining (NJ) phylogenetic method with 1000
247 bootstrap replicates was used to generate a phylogenetic tree for DREB proteins among *S.*
248 *tuberosum* (66 DREB proteins) and *A. thaliana* (56 DREB proteins) **(Data S3)**. The phylogenetic
249 tree was constructed followed by multiple sequence alignment (MSA) of AP2 domain coding
250 protein sequences present in potato DREB sequences. All 66 genes of DREB family in potato
251 were categorized into six subgroups referred to as A1, A2, A3, A4, A5 and A6 with reference to
252 the classification of AtDREBs as shown in **Fig. 1**.

253 The largest subgroup A6 consisted of 19 members while the smallest subgroup A3 consisted of
254 only 2 StDREB members. A total of 11, 8, 18 and 8 protein members were assigned to A1, A2,
255 A4 and A5 subgroups, respectively.

256 The conserved amino acid sequence present in AP2 domain of StDREB members had valine at
257 position 14th and glutamic acid at position 19th but some members only had valine which plays
258 an important role in identification of the DREB gene family's various DNA binding sites. The
259 phylogenetic tree revealed that StDREB members (StDREB50, StDREB52, StDREB39,
260 StDREB40, StDREB56, StDREB65, StDREB61, StDREB66, and StDREB62) belonging to
261 group 6 exhibited greater homology with each other rather than with *Arabidopsis* DREB
262 sequences due to high structure similarity among them which was further investigated by gene
263 structural analysis and conserved motifs identification in potato.

Gene Structure and Motif Composition of StDREB Members

265 The organization of introns and exons is pivotal in the evolution of gene families. Gene structure
266 analysis was conducted through aligning the cDNA and gDNA sequences **(Data S4)** to gain
267 further intuition into the structural similarity and divergence of DREB genes in Potato. We also
268 constructed a phylogenetic tree as shown in **Fig. 2** but using only full-length protein sequences of
269 StDREB genes to relate it with exon/intron distribution and motif composition. Results showed
270 that eight StDREB genes (21, 34, 60, 20, 13, 39, 40, and 52) had one intron with the exception of
271 StDREB56 which contained 3 introns in its coding region while other fifty-seven StDREB genes
272 had only exons in their coding region as shown in **Fig. 2B**. MEME motif detection web software
273 was utilized to recognize the conserved motifs to further understand the diversification in
274 StDREB members of Potato **Fig. 2C**. We found fifteen conserved motifs with different amino
275 acid range from 9-50. Motif 12 was the shortest with 9 amino acid residues (DDDMSLWSY)
276 while motif 15 was the longest with 50 amino acid residues
277 (NNYJPYGFYPAVQYAEDISQNPQHSIQKQTFDDNYGFLDGETTKASGMIW). Motif
278 composition in different subgroups was different **(Figs. S1-S2)**. Motifs 1 and 2 were found
279 within the AP2 domain. Majority of the StDREB proteins observed motif 1 and motif 2 with

280 very few exceptions. Motif 13 was unique to three StDREB (5, 23 and, 57) members belonging
281 to A1 subgroup. Motif 15 was exclusively present in A3 subgroup (63, 64), whereas motif 8 was
282 found both in A3 subgroup and seven out of eight (7, 9, 10, 13, 15, 29, and 31) members of A2
283 subgroup proposing that proteins in these subgroups may possibly share a specific function.
284 Motif 9 was only encountered in five StDREB protein sequences (11, 22, 24, 28, and 46) of the
285 A4 subgroup. Motif 10, motif 11, and motif 14 were detected only within some members of A6
286 subgroup. Majority of StDREB proteins within the same subgroup exhibited similar intron-exon
287 structure and motif compositions which supported the phylogenetic analysis of DREB gene
288 family, while the difference among the distinct subgroups directed their diverse roles. The
289 similarities and structural variations among these motifs can be studied further to provide new
290 insights. Most StDREBs classified in the same subgroup generally had a similar motif
291 composition and might have similar functions. Details of 15 conserved motifs are summarized in
292 **Data S5**.

Chromosomal Localization and Gene Duplication Analysis of StDREBs

294 Chromosomal positions of all StDREB genes were obtained from Phytozome database v.12.1.
295 Physical location of all the identified StDREB genes were mapped by using MapChart 2.2
296 software on to their corresponding chromosomes as depicted in **Fig. 3**. Chromosomal
297 localization indicated that the 66 StDREB genes were heterogeneously distributed onto 12
298 chromosomes across the Potato genome which indicated that genetic variation occurred during
299 the evolutionary process. Largest numbers of StDREB genes i.e. 13 were positioned on
300 chromosome 3 (chr03). Conversely, chromosome 2 (chr02) and chromosome 9 (chr09) had the
301 least number of StDREB genes i.e. 1. Two chromosomes (chr04 and chr11) hosted 5 StDREB
302 genes. In addition, chromosome 8 (chr08) harbored 11 StDREB genes, chromosome 1 had 7
303 StDREB genes, chromosome 5 (chr05) and chromosome 7 (chr07) had 3 and 2 StDREB genes,
304 respectively.

305 Tandem and segmental duplications contribute to the expansion of new gene family members
306 and novel functions in the evolution of plant genome. To investigate gene duplication events
307 within the potato genome, we investigated tandem and segmental duplications during evolution
308 of StDREB gene family. A total of 10 StDREB gene pairs were confirmed to be tandem
309 duplications. According to gene duplication analysis, it was observed that chr01, chr10, chr11
310 and chr12 each experienced one tandem duplicated gene pair while chr3 and chr8 harbored three
311 tandem duplicated gene pairs illustrated as red lines in **Fig. 3**. Besides tandem duplication, 8
312 segmental duplication events were identified by constructing Circos Plot illustrated as red color
313 in **Fig. 4**. To determine the selective evolutionary pressure on StDREB gene divergence after
314 duplication, Ka and Ks values were computed for the duplicated StDREB gene pairs using KaKs
315 calculator. Generally, KaKs = 1 implies neutral selection, Ka/Ks >1 implies positive selection
316 and, Ka/Ks < 1 implies purification selection. Each duplicated StDREB gene pair had Ka/Ks < 1,
317 which was indicative of purification selection during evolution. Furthermore, duplication events
318 of 18 gene pairs were estimated to have occurred between 6.15 and 223.0 million years ago
319 **(Data S6)**.

Synteny Relationship of StDREB Genes

321 To further infer the evolutionary relationship, DREB genes were compared to identify
322 orthologous StDREB gene pairs between *S. tuberosum*, *A. thaliana* and *S. lycopersicum*.
323 According to the synteny analysis, 36 of 66 StDREB genes had collinear gene pairs in A.

324 thaliana and 45 of 66 StDREB genes had corresponding orthologs in *S. lycopersicum*,
325 respectively as provided in **Data S7**. The syntenic maps showed high evolutionary homology
326 relationship of StDREB genes with *A. thaliana* and *S. lycopersicum* DREB genes implying that
327 they might have related functions as depicted in **Fig. 5**.

328 **Gene Ontology (GO) Annotation and Subcellular Localization Prediction of**
329 **StDREB Genes**
330 To explore the functions of StDREB genes in different biological processes, molecular functions
331 and cellular compartment building, GO functional annotation was performed as depicted in **Fig.**
332 **6** and gene ontology number was also identified during the analysis as provided in **Data S8**.
333 Biological processes showed that the StDREB genes were involved in cell metabolism, defense
334 response and ethylene activated signaling. Majority of the StDREB genes were involved in
335 sequence specific DNA-binding functions and transcriptional regulation, indicating their
336 significance as a transcription factor for potato growth and development, as well as regulating
337 the expression of abiotic stress responsive genes. The cellular compartment study revealed that
338 most of the StDREB genes were concentrated in cell nucleus with the exception of StDREB47
339 and StDREB53 which were localized in cytoplasm. In addition, some StDREB genes were found
340 in cell membrane to control the signal transduction inside the cell.

341 To predict the subcellular localization of StDREB proteins, two different tools were employed.
342 Results confirmed that StDREB proteins were predominantly localized in the nucleus followed
343 by mitochondria and chloroplast, respectively as indicated in the heat map (**Fig. 7; Data S9**.

DISCUSSION

345 Potato is one of the world's most important crops, providing food to over 100 countries. Potato is
346 vulnerable to several abiotic stresses, and its steady development is jeopardized by recurrent
347 stress outbreaks (Mirzaei et al. 2020). The DREB gene family plays an important role in plant's
348 abiotic stress signaling. DREB gene family is well known for its highly conserved AP2 domain
349 that can specifically bind to the DRE/CRT cis-acting elements to activate expression of several
350 stress tolerance genes, thereby enhancing plant tolerance (Chen et al. 2016b). With the
351 availability of whole-genome sequence, members of the DREB gene family have been identified
352 in many plant species. Until now, bioinformatic analysis of DREB gene family has not been
353 reported in potato. Therefore, our current study furnished a detailed data on genome-wide
354 characterization of DREB genes in *Solanum tuberosum*. By using *Arabidopsis thaliana* DREB
355 genes as query sequences, we identified 66 putative DREB genes in *S. tuberosum* **(Data S2)** and
356 categorized them into six subgroups corresponding to *A. thaliana* DREB orthologs. The amino
357 acid conservation (valine at 14th and glutamic acid at 19th) confirmed the presence of DREB
358 genes in *S. tuberosum* genome.

359 The genome-wide analysis revealed the difference in number of genes and genome size of *S.*
360 *tuberosum* and *A. thaliana*. The identified number of DREB genes in *S. tuberosum* (66) is greater
361 than that in *Arabidopsis* (56) which may be due to the difference in their genome sizes. When
362 compared to previous studies, greater numbers of DREB genes were reported in *S. spontaneum*
363 (110) (Huang et al. 2020), *P. trichocarpa* (75) (Chen et al. 2013), soybean (73) (Zhou et al.
364 2020), and malus (68) (Zhao et al. 2012), but less DREB genes were found in sesame (41)
365 (Dossa et al. 2016), maize (51) (Du et al. 2014), mulberry (30) (Liu et al. 2015), grape (38)

366 (Zhao et al. 2014), common bean (54) (Konzen et al. 2019), *Phyllostachys edulis* (27), and
367 pineapple (20) (Chai et al. 2020). In order to comprehend the potential functions of proteins, it is
368 inevitable to compute the physiochemical characteristics of plant protein families (Salih et al.
369 2019). The DREB proteins in potato demonstrated diversification in terms of physiochemical
370 properties, which indicated that these genes might play various roles in plant development and
371 defense response. The lengths of all DREB proteins had a range of 72-457 amino acid residues.
372 The molecular mass largely varied between 8.3 and 52.29 kDa, and the isoelectric point ranged
373 from 4.17 to 9.97 **(Data S2)**. Moreover, our phylogenetic analysis of *S. tuberosum* and *A.
374 thaliana* DREB proteins demonstrated that the six subgroups (A1-A6) identified previously in
375 *Arabidopsis* (Akhtar et al. 2012) were also present in *S. tuberosum*. The total number of genes
376 for each group was 11, 8, 2, 18, 8, and 19, corresponding to the A1 to A6 subgroups. This
377 distribution was quite different from *A. thaliana* due to different genetic patterns in both plants
378 **(Fig. 1)**. Gene number similar to StDREB A3 subgroup had been reported in *Salix arbutifolia* (2)
379 (Rao et al. 2015) and *P. trichocarpa* (2). However, 18 DREB genes were categorized in A3
380 subgroup of foxtail millet (Shi et al. 2018), five in *S. spontaneum*, one in mulberry, one in
381 *Phaseolus vulgaris*, and zero in pineapple.

382 To gain insights into the structural diversity of the StDREB genes, their gene structures were
383 analyzed. Formerly, it was thought that DREB genes only contain coding region exons, without
384 any intron as in *Arabidopsis*. But, later studies revealed that DREB genes contain both the exon
385 and intron in their gene structure. This finding was further confirmed by gene structure analysis
386 in wheat (Sazegari & Niazi 2012), rice (Matsukura et al. 2010) and maize. Our study
387 demonstrated nine StDREB genes consist of intronic regions along with exons. Out of nine
388 StDREB genes, StDREB56 contain three introns whereas remaining 8 genes contain one intron
389 only. While 76% of the StDREB genes showed a single exon **(Fig. 2B)**. A previous study
390 conducted on pineapple had also demonstrated highest number of introns as three in one of its 20
391 DREB genes. There was a strong connection of exon/intron structure between *S. tuberosum* and
392 other species (*S. spontaneum*, soybean, pineapple, and populas) due to the presence of both exon
393 and intron. Also, some studies showed that a compact gene structure with few or no introns
394 enhanced timely response to various abiotic stresses in plants (Jeffares et al. 2008). The variation
395 found in gene structure of StDREB genes elucidates the functional diversification which might
396 be due to climatic or evolutionary changes in the plant genome profile (Li et al. 2020). Recent
397 studies revealed the presence of some conserved motif sequences in transcription factors of *F.
398 tataricum* located on the same chromosome (Liu et al. 2019), similar profile was observed in *A.
399 thaliana* (Sakuma et al. 2002) which may be due to polyploid changes that occurred in the
400 genome. Same results were obtained in *S. tuberosum* DREB genes through identification of
401 conserved motif sequences. We identified 15 conserved motifs in StDREB genes along with the
402 conserved AP2 domain **(Fig. 2C)**. The AP2 domain has one alpha helix and three beta sheets at
403 the N-terminus (Wang et al. 2011). The identified motifs in *S. tuberosum* DREB genes indicated
404 high similarity with AP2 structure pattern. All conserved motifs had varied composition **(Figs.
405 S1-S2)**. Transcription factor domains and motifs are frequently linked to DNA binding,
406 transcriptional activity, and protein interaction (Liu et al. 1999). The gene structure and motif
407 analysis of the same subgroup were alike, thus validating the reliability of the phylogenetic tree
408 classification. This finding was in line with the previous studies on DREB, which found that
409 broad similarities existed in motifs and exon/intron structure between members of the same
410 subgroup.

411   Chromosomal positions demonstrated uneven scattering of 66 DREB genes over 12 potato
412   chromosomes. The asymmetrical arrangement of genes has been suggested to reveal information
413   about their evolution (Chen et al. 2016a). Whole genome duplications (WGD) or
414   polyploidization have been considered as the major causes of evolution that gives rise to novel
415   traits and new transcriptional regulatory sites that can alter expression patterns (Panchy et al.
416   2016). In total, 10 homologous pairs were confirmed to be produced by tandem duplication
417   events **(Fig. 3)**. In contrast, eight paralogous gene pairs were produced through segmental
418   duplication events **(Fig. 4)**. Moreover, gene pairs with tandem duplications were found on
419   chromosomes of the same origin. The most recent tandem duplication events were estimated to
420   be six million years ago for two pairs of genes belonging to A3 and A6 subgroup, respectively
421   **(Data S6)**. *S. spontaneum*, *P. trichocarpa,* and soybean genome had undergone both tandem and
422   segmental duplications whereas only tandem duplications were found in *P. vulgaris*. Tandem
423   duplications are also known to be adaptively important in the development and function of
424   abiotic stress responsive genes. Previous studies reported that tandem repeats often share
425   common cis-acting elements, and may perform similar functions (Flagel & Wendel 2009).
426   Hence, our study also emphasizes that StDREB tandem gene duplication pairs may share similar
427   functions and regulatory elements in their promoter region. To further investigate the potential
428   evolutionary mechanisms of the *S. tuberosum* DREB gene family, interspecies synteny was
429   inferred. The number of orthologous events of StDREB genes was greater with SlDREB genes
430   as compared to AtDREB genes, which may be due to close evolutionary link between *S.*
431   *tuberosum* and *S. lycopersicum* **(Fig. 5)**. GO enrichment has been considered as a powerful tool
432   for enhancing the understanding of functional genomics and underlying molecular mechanism.
433   GO annotation verified the DNA-binding ability of StDREB genes which was consistent with the
434   findings in *P. vulgaris* **(Fig. 6)**. All the StDREB genes were concentrated inside the nucleus.
435   Furthermore, our results from the subcellular localization also indicated that StDREB proteins
436   were primarily present inside the nucleus **(Fig. 7)**.

CONCLUSIONS

437
438   In conclusion, we performed genome-wide identification and characterization of the DREB gene
439   family in *S. tuberosum* and conducted a detailed investigation of their evolutionary relationship,
440   genome organization, duplication events, and functional annotation using bioinformatics tools. In
441   total, 66 DREB genes having an AP2 domain and conserved amino acid residues (Valine at
442   position 14th) were identified in the Potato genome and unevenly mapped across 12
443   chromosomes. Based on the sequence alignment and phylogenetic analysis, StDREB genes were
444   classified into six subgroups (A1-A6) corresponding to previous report of *Arabidopsis*. The
445   results of gene structure and conserved motif analysis were found consistent with the
446   phylogenetic classification. Gene structures and motif patterns showed that the StDREB
447   members in the same subgroup displayed broad similarities. The expansion of the DREB gene
448   family in Potato was aided greatly by tandem and segmental duplications. Evolutionary
449   divergence analysis (Ka/Ks) suggested that the StDREBs were under strong purification
450   selection during plant evolution. Synteny relationship analysis indicated that 35 and 46 StDREB
451   genes were orthologous to *Arabidopsis* and *S. lycopersicum*, respectively. Furthermore,
452   subcellular localization revealed that the StDREB genes were primarily located inside the
453   nucleus. Through gene ontology (GO) annotation, we found most of the StDREB genes had
454   DNA binding function, which suggested their role as important transcriptional activators.

455 Functional enrichment indicated pivotal roles of StDREB genes in cell development, defense
456 responses, and hormone signaling. Overall, our data delineated the evolutionary characteristics
457 and genome duplication events along with biological and molecular functions of DREB genes in
458 *S. tuberosum*. Taken together, our results will provide a foundation for unraveling the molecular
459 mechanisms and further functional characterization of the StDREB gene family, thus providing
460 sources for plant breeding and genetic engineering.

461 ## REFERENCES

462 Agarwal P, Agarwal PK, Nair S, Sopory S, and Reddy M. 2007. Stress-inducible DREB2A transcription
463 factor from Pennisetum glaucum is a phosphoprotein and its phosphorylation negatively
464 regulates its DNA-binding activity. *Molecular Genetics and Genomics* 277:189-198. DOI
465 10.1007/s00438-006-0183-z

466 Akhtar M, Jaiswal A, Taj G, Jaiswal J, Qureshi M, and Singh N. 2012. DREB1/CBF transcription factors:
467 their structure, function and role in abiotic stress tolerance in plants. *Journal of Genetics* 91:385-
468 395. DOI 10.1007/s12041-012-0201-3

469 Allen MD, Yamasaki K, Ohme-Takagi M, Tateno M, and Suzuki MJTEj. 1998. A novel mode of DNA
470 recognition by a β-sheet revealed by the solution structure of the GCC-box binding domain in
471 complex with DNA. *The EMBO Journal* 17:5484-5496. DOI 10.1093/emboj/17.18.5484

472 Bailey TL, Johnson J, Grant CE, and Noble WS. 2015. The MEME suite. *Nucleic acids research* 43:W39-
473 W49. DOI 10.1093/nar/gkv416

474 Bouaziz D, Charfeddine M, Jbir R, Saidi MN, Pirrello J, Charfeddine S, Bouzayen M, and Gargouri-Bouzid
475 R. 2015a. Identification and functional characterization of ten AP2/ERF genes in potato. *Plant
476 Cell, Tissue and Organ Culture (PCTOC)* 123:155-172. DOI 10.1007/s11240-015-0823-2

477 Bouaziz D, Jbir R, Charfeddine S, Saidi MN, and Gargouri-Bouzid R. 2015b. The StDREB1 transcription
478 factor is involved in oxidative stress response and enhances tolerance to salt stress. *Plant Cell,
479 Tissue and Organ Culture (PCTOC)* 121:237-248. DOI 10.1007/s11240-014-0698-7

480 Bouaziz D, Pirrello J, Amor HB, Hammami A, Charfeddine M, Dhieb A, Bouzayen M, and Gargouri-Bouzid
481 R. 2012. Ectopic expression of dehydration responsive element binding proteins (StDREB2)
482 confers higher tolerance to salt stress in potato. *Plant Physiology and Biochemistry* 60:98-108.
483 DOI 10.1016/j.plaphy.2012.07.029

484 Chai M, Cheng H, Yan M, Priyadarshani S, Zhang M, He Q, Huang Y, Chen F, Liu L, and Huang X. 2020.
485 Identification and expression analysis of the DREB transcription factor family in pineapple
486 (Ananas comosus (L.) Merr.). *Peer J* 8:e9006. DOI 10.7717%2Fpeerj.9006

487 Charfeddine M, Charfeddine S, Bouaziz D, Messaoud RB, and Bouzid RG. 2017. The effect of cadmium on
488 transgenic potato (Solanum tuberosum) plants overexpressing the StDREB transcription factors.
489 *Plant Cell, Tissue and Organ Culture (PCTOC)* 128:521-541. DOI 10.1007/s11240-016-1130-2

490 Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, and Xia R. 2020. TBtools: an integrative toolkit
491 developed for interactive analyses of big biological data. *Molecular plant* 13:1194-1202. DOI
492 10.1016/j.molp.2020.06.009

493 Chen L, Han J, Deng X, Tan S, Li L, Li L, Zhou J, Peng H, Yang G, and He G. 2016a. Expansion and stress
494 responses of AP2/EREBP superfamily in Brachypodium distachyon. *Scientific Reports* 6:1-14.
495 https://doi.org/10.1038/srep21623

496 Chen M, Wang Q-Y, Cheng X-G, Xu Z-S, Li L-C, Ye X-G, Xia L-Q, and Ma Y-Z. 2007. GmDREB2, a soybean
497 DRE-binding transcription factor, conferred drought and high-salt tolerance in transgenic plants.
498 *Biochemical and Biophysical Research Communications* 353:299-305. DOI
499 10.1016/j.bbrc.2006.12.027

500 Chen Y, Huang L, Yan H, Zhang X, Xu B, and Ma X. 2016b. Cloning and characterization of an ABA-
501   independent DREB transcription factor gene, HcDREB2, in Hemarthria compressa. *Hereditas*
502   153:1-7. DOI 10.1186/s41065-016-0008-y
503 Chen Y, Yang J, Wang Z, Zhang H, Mao X, and Li C. 2013. Gene structures, classification, and expression
504   models of the DREB transcription factor subfamily in Populus trichocarpa. *The Scientific World*
505   *Journal* 2013. DOI 10.1155/2013/954640
506 Conesa A, and Götz S. 2008. Blast2GO: a comprehensive suite for functional analysis in plant genomics.
507   *International journal of plant genomics* 2008. DOI 10.1155/2008/619832
508 Dahal K, Li X-Q, Tai H, Creelman A, and Bizimungu B. 2019 Improving potato stress tolerance and tuber
509   yield under a climate change scenario–a current overview. *Frontiers in Plant Science* 10:563. DOI
510   10.3389/fpls.2019.00563
511 Davies Jr FT, Calderón CM, Huaman Z, and Gómez R. 2005. Influence of a flavonoid (formononetin) on
512   mycorrhizal activity and potato crop productivity in the highlands of Peru. *Scientia Horticulturae*
513   106:318-329. DOI 10.1016/j.scienta.2005.04.013
514 Dietz K-J, Vogel MO, and Viehhauser A. 2010. AP2/EREBP transcription factors are part of gene
515   regulatory networks and integrate metabolic, hormonal and environmental signals in stress
516   acclimation and retrograde signalling. *Protoplasma* 245:3-14.
517 Dong C-J, and Liu J-Y. 2010. The Arabidopsis EAR-motif-containing protein RAP2. 1 functions as an active
518   transcriptional repressor to keep stress responses under tight control. *BMC Plant Biology* 10:1-
519   15. DOI 10.1186/1471-2229-10-47
520 Dossa K, Wei X, Li D, Fonceka D, Zhang Y, Wang L, Yu J, Boshou L, Diouf D, and Cissé N. 2016. Insight into
521   the AP2/ERF transcription factor superfamily in sesame and expression profiling of DREB
522   subfamily under drought stress. *BMC Plant Biology* 16:1-16.
523 Du H, Huang M, Zhang Z, and Cheng S. 2014. Genome-wide analysis of the AP2/ERF gene family in maize
524   waterlogging stress response. *Euphytica* 198:115-126. DOI 10.1007/s10681-014-1088-2
525 Du X, Li W, Sheng L, Deng Y, Wang Y, Zhang W, Yu K, Jiang J, Fang W, and Guan Z. 2018. Over-expression
526   of chrysanthemum CmDREB6 enhanced tolerance of chrysanthemum to heat stress. *BMC Plant*
527   *Biology* 18:1-10. DOI 10.1186/s12870-018-1400-8
528 Dubouzet JG, Sakuma Y, Ito Y, Kasuga M, Dubouzet EG, Miura S, Seki M, Shinozaki K, and
529   Yamaguchi-Shinozaki K. 2003. OsDREB genes in rice, Oryza sativa L., encode transcription
530   activators that function in drought-, high-salt-and cold-responsive gene expression. *The Plant*
531   *Journal* 33:751-763. DOI 10.1046/j.1365-313x.2003.01661.x
532 Egawa C, Kobayashi F, Ishibashi M, Nakamura T, Nakamura C, and Takumi S. 2006. Differential regulation
533   of transcript accumulation and alternative splicing of a DREB2 homolog under abiotic stress
534   conditions in common wheat. *Genes & Genetic Systems* 81:77-91. DOI 10.1266/ggs.81.77
535 Feng Q, Jie L, Gui-You Z, Jun Z, Shou-Yi C, and Qiang L. 2003. Isolation and structural analysis of DRE-
536   binding transcription factor from maize (Zea mays L.). *Journal of Integrative Plant Biology*

537   45:331.
538 Flagel LE, and Wendel JF. 2009 Gene duplication and evolutionary novelty in plants. *New Phytologist*
539   183:557-564. DOI 10.1111/j.1469-8137.2009.02923.x
540 Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, and Bairoch A. 2003. ExPASy: the proteomics
541   server for in-depth protein knowledge and analysis. *Nucleic acids research* 31:3784-3788. DOI
542   10.1093/nar/gkg563
543 Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, Mitros T, Dirks W, Hellsten U, and
544   Putnam N. 2012. Phytozome: a comparative platform for green plant genomics. *Nucleic acids*
545   *research* 40:D1178-D1186. DOI 10.1093/nar/gkr944

546 Guo A-Y, Zhu Q-H, Chen X, and Luo J-C. 2007. GSDS: a gene structure display server. *Yi chuan= Hereditas*
547   29:1023-1026. DOI 10.1360/yc-007-1023
548 Guo B, Wei Y, Xu R, Lin S, Luan H, Lv C, Zhang X, Song X, and Xu R. 2016. Genome-wide analysis of
549   APETALA2/ethylene-responsive factor (AP2/ERF) gene family in barley (Hordeum vulgare L.).
550   *PLoS One* 11:e0161322. DOI 10.1371/journal.pone.0161322
551 Guo J, and Wang M-H. 2011. Expression profiling of the DREB2 type gene from tomato (Solanum
552   lycopersicum L.) under various abiotic stresses. *Horticulture Environment and Biotechnology*
553   52:105. DOI 10.1007/s13580-011-0125-5
554 Handayani T, Gilani SA, and Watanabe KN. 2019. Climatic changes and potatoes: How can we cope with
555   the abiotic stresses? *Breeding Science* 69:545-563. DOI 10.1270/jsbbs.19070
556 He Y, Liu X, Ye L, Pan C, Chen L, Zou T, and Lu G. 2016. Genome-wide identification and expression
557   analysis of two-component system genes in tomato. *International Journal of Molecular Sciences*
558   17:1204. DOI 10.3390/ijms17081204
559 Hichri I, Muhovski Y, Clippe A, Žižková E, Dobrev PI, Motyka V, and Lutts S. 2016. SlDREB2, a tomato
560   dehydration-responsive element-binding 2 transcription factor, mediates salt stress tolerance in
561   tomato and A rabidopsis. *Plant, Cell & Environment* 39:62-79. DOI 10.1111/pce.12591
562 Hong J-P, and Kim WT. 2005. Isolation and functional characterization of the Ca-DREBLP1 gene encoding
563   a dehydration-responsive element binding-factor-like protein 1 in hot pepper (Capsicum
564   annuum L. cv. Pukang). *Planta* 220:875-888. DOI 10.1007/s00425-004-1412-5
565 Horton P, Park K-J, Obayashi T, Fujita N, Harada H, Adams-Collier C, and Nakai K. 2007. WoLF PSORT:
566   protein localization predictor. *Nucleic acids research* 35:W585-W587. DOI 10.1093/nar/gkm259
567 Huala E, Dickerman AW, Garcia-Hernandez M, Weems D, Reiser L, LaFond F, Hanley D, Kiphart D, Zhuang
568   M, and Huang W. 2001. The Arabidopsis Information Resource (TAIR): a comprehensive
569   database and web-based information retrieval, analysis, and visualization system for a model
570   plant. *Nucleic acids research* 29:102-105. DOI 10.1093/nar/29.1.102
571 Huang X, Song X, Chen R, Zhang B, Li C, Liang Y, Qiu L, Fan Y, Zhou Z, and Zhou H. 2020. Genome-wide
572   analysis of the DREB subfamily in Saccharum spontaneum reveals their functional divergence
573   during cold and drought stresses. *Frontiers in Genetics* 10:1326. DOI 10.3389/fgene.2019.01326
574 Hulo N, Bairoch A, Bulliard V, Cerutti L, De Castro E, Langendijk-Genevaux PS, Pagni M, and Sigrist CJ.
575   2006. The PROSITE database. *Nucleic acids research* 34:D227-D230. DOI 10.1093/nar/gkj063
576 Hussain T. 2016 Potatoes: ensuring food for the future. *Advances in Plants & Agriculture Research* 3:178-
577   182. DOI 10.15406/apar.2016.03.00117
578 Hwang JE, Lim CJ, Chen H, Je J, Song C, and Lim CO. 2012. Overexpression of Arabidopsis dehydration-
579   responsive element-binding protein 2C confers tolerance to oxidative stress. *Molecules and Cells*
580   33:135-140. DOI 10.1007/s10059-012-2188-2
581 Jeffares DC, Penkett CJ, and Bähler J. 2008. Rapidly regulated genes are intron poor. *Trends in Genetics*
582   24:375-378. DOI 10.1016/j.tig.2008.05.006
583 Kimotho RN, Baillo EH, and Zhang Z. 2019. Transcription factors involved in abiotic stress responses in
584   Maize (Zea mays L.) and their roles in enhanced productivity in the post genomics era. *Peer J*
585   7:e7211. DOI 10.7717/peerj.7211
586 Konzen ER, Recchia GH, Cassieri F, Caldas DGG, Berny Mier y Teran JC, Gepts P, and Tsai SM. 2019. DREB
587   genes from common bean (Phaseolus vulgaris L.) show broad to specific abiotic stress responses
588   and distinct levels of nucleotide diversity. *International Journal of Genomics* 2019. DOI
589   10.1155/2019/9520642
590 Kudo M, Kidokoro S, Yoshida T, Mizoi J, Todaka D, Fernie AR, Shinozaki K, and Yamaguchi-Shinozaki K.
591   2017. Double overexpression of DREB and PIF transcription factors improves drought stress
592   tolerance and cell elongation in transgenic plants. *Plant Biotechnology Journal* 15:458-471. DOI
593   10.1111%2Fpbi.12644

594   Kumar S, Stecher G, Li M, Knyaz C, and Tamura K. 2018. MEGA X: molecular evolutionary genetics
595   analysis across computing platforms. *Molecular biology and evolution* 35:1547-1549. [DOI](#)
596   [10.1093/molbev/msy096](#)
597   Labbo AM, Mehmood M, Akhtar MN, Khan MJ, Tariq A, and Sadiq I. 2018. Genome-wide identification of
598   AP2/ERF transcription factors in mungbean (Vigna radiata) and expression profiling of the
599   VrDREB subfamily under drought stress. *Crop and Pasture Science* 69:1009-1019.
600   Lakhwani D, Pandey A, Dhar YV, Bag SK, Trivedi PK, and Asif MH. 2016. Genome-wide analysis of the
601   AP2/ERF family in Musa species reveals divergence and neofunctionalisation during evolution.
602   *Scientific Reports* 6:1-17. [DOI 10.1038/srep18878](#)
603   Lata C, and Prasad M. 2011. Role of DREBs in regulation of abiotic stress responses in plants. *Journal of*
604   *Experimental Botany* 62:4731-4748. [DOI 10.1093/jxb/err210](#)
605   Letunic I, and Bork P. 2018. 20 years of the SMART protein domain annotation resource. *Nucleic acids*
606   *research* 46:D493-D496. [DOI 10.1093/nar/gkx922](#)
607   Li P, Chai Z, Lin P, Huang C, Huang G, Xu L, Deng Z, Zhang M, Zhang Y, and Zhao X. 2020. Genome-wide
608   identification and expression analysis of AP2/ERF transcription factors in sugarcane (Saccharum
609   spontaneum L.). *BMC Genomics* 21:1-17. [DOI 10.1186/s12864-020-07076-x](#)
610   Liang Y, Kang K, Gan L, Ning S, Xiong J, Song S, Xi L, Lai S, Yin Y, and Gu J. 2019. Drought-responsive
611   genes, late embryogenesis abundant group3 (LEA 3) and vicinal oxygen chelate, function in lipid
612   accumulation in Brassica napus and Arabidopsis mainly via enhancing photosynthetic efficiency
613   and reducing ROS. *Plant Biotechnology Journal* 17:2123-2142. [DOI 10.1111/pbi.13127](#)
614   Liu L, White MJ, and MacRae TH. 1999. Transcription factors and their genes in higher plants: functional
615   domains, evolution and regulation. *European Journal of Biochemistry* 262:247-257. [DOI](#)
616   [10.1046/j.1432-1327.1999.00349.x](#)
617   Liu M, Sun W, Ma Z, Zheng T, Huang L, Wu Q, Zhao G, Tang Z, Bu T, and Li C. 2019. Genome-wide
618   investigation of the AP2/ERF gene family in tartary buckwheat (Fagopyum Tataricum). *BMC*
619   *Plant Biology* 19:1-19. [DOI 10.1186/s12870-019-1681-6](#)
620   Liu Q, Kasuga M, Sakuma Y, Abe H, Miura S, Yamaguchi-Shinozaki K, and Shinozaki K. 1998. Two
621   transcription factors, DREB1 and DREB2, with an EREBP/AP2 DNA binding domain separate two
622   cellular signal transduction pathways in drought-and low-temperature-responsive gene
623   expression, respectively, in Arabidopsis. *The Plant Cell* 10:1391-1406. [DOI](#)
624   [10.1105/tpc.10.8.1391](#)
625   Liu S, Wang X, Wang H, Xin H, Yang X, Yan J, Li J, Tran L-SP, Shinozaki K, and Yamaguchi-Shinozaki K.
626   2013. Genome-wide analysis of ZmDREB genes and their association with natural variation in
627   drought tolerance at seedling stage of Zea mays L. *PLoS Genetics* 9:e1003790. [DOI](#)
628   [10.1371/journal.pgen.1003790](#)
629   Liu X, Zhu J, Wei C, Guo Q, Bian C, Xiang Z, and Zhao A. 2015. Genome-wide identification and
630   characterization of the DREB transcription factor gene family in mulberry. *Biologia Plantarum*
631   59:253-265. [DOI 10.1007/s10535-015-0498-x](#)
632   Lucas S, Durmaz E, Akpınar BA, and Budak H. 2011 The drought response displayed by a DRE-binding
633   protein from Triticum dicoccoides. *Plant Physiology and Biochemistry* 49:346-351. [DOI](#)
634   [10.1016/j.plaphy.2011.01.016](#)
635   Matsukura S, Mizoi J, Yoshida T, Todaka D, Ito Y, Maruyama K, Shinozaki K, and Yamaguchi-Shinozaki K.
636   2010. Comprehensive analysis of rice DREB2-type genes that encode transcription factors
637   involved in the expression of abiotic stress-responsive genes. *Molecular Genetics and Genomics*
638   283:185-196. [DOI 10.1007/s00438-009-0506-y](#)
639   McKain MR, Tang H, McNeal JR, Ayyampalayam S, Davis JI, Depamphilis CW, Givnish TJ, Pires JC,
640   Stevenson DW, and Leebens-Mack JH. 2016. A phylogenomic assessment of ancient polyploidy

641   and genome evolution across the Poales. *Genome Biology and Evolution* 8:1150-1164. DOI
642   10.1093%2Fgbe%2Fevw060
643 Mirzaei K, Bahramnejad B, and Fatemi S. 2020. Genome-wide identification and characterization of the
644   bZIP gene family in potato (Solanum tuberosum). *Plant Gene* 24:100257. DOI
645   10.1016/j.plgene.2020.100257
646 Mistry J, Chuguransky S, Williams L, Qureshi M, Salazar GA, Sonnhammer EL, Tosatto SC, Paladin L, Raj S,
647   and Richardson LJ. 2021. Pfam: The protein families database in 2021. *Nucleic acids research*
648   49:D412-D419. DOI 10.1093/nar/gkaa913
649 Movahedi S, Tabatabaei BS, Alizade H, Ghobadi C, Yamchi A, and Khaksar G. 2012. Constitutive
650   expression of Arabidopsis DREB1B in transgenic potato enhances drought and freezing
651   tolerance. *Biologia Plantarum* 56:37-42. DOI 10.1007/s10535-012-0013-6
652 Mueller LA, Solow TH, Taylor N, Skwarecki B, Buels R, Binns J, Lin C, Wright MH, Ahrens R, and Wang Y.
653   2005. The SOL Genomics Network. A comparative resource for Solanaceae biology and beyond.
654   *Plant physiology* 138:1310-1317. DOI 10.1104/pp.105.060707
655 Niu X, Luo T, Zhao H, Su Y, Ji W, and Li H. 2020. Identification of wheat DREB genes and functional
656   characterization of TaDREB3 in response to abiotic stresses. *Gene* 740:144514. DOI
657   10.1016/j.gene.2020.144514
658 Panchy N, Lehti-Shiu M, and Shiu S-H. 2016. Evolution of gene duplication in plants. *Plant physiology*
659   171:2294-2316. DOI 10.1104%2Fpp.16.00523
660 Park S, Shi A, and Mou B. 2020. Genome-wide identification and expression analysis of the CBF/DREB1
661   gene family in lettuce. *Scientific Reports* 10:1-14. DOI 10.1038/s41598-020-62458-1
662 Rabara RC, Tripathi P, and Rushton PJ. 2014. The potential of transcription factor-based genetic
663   engineering in improving crop tolerance to drought. *OMICS: A Journal of Integrative Biology*
664   18:601-614. DOI 10.1089%2Fomi.2013.0177
665 Rae L, Lao NT, and Kavanagh TA. 2011. Regulation of multiple aquaporin genes in Arabidopsis by a pair
666   of recently duplicated DREB transcription factors. *Planta* 234:429-444. DOI 10.1007/s00425-011-
667   1414-z
668 Rao G, Sui J, Zeng Y, He C, and Zhang J. 2015. Genome-wide analysis of the AP2/ERF gene family in Salix
669   arbutifolia. *FEBS Open Bio* 5:132-137. DOI 10.1016/j.fob.2015.02.002
670 Sakuma Y, Liu Q, Dubouzet JG, Abe H, Shinozaki K, and Yamaguchi-Shinozaki K. 2002. DNA-binding
671   specificity of the ERF/AP2 domain of Arabidopsis DREBs, transcription factors involved in
672   dehydration-and cold-inducible gene expression. *Biochemical and Biophysical Research*
673   *Communications* 290:998-1009. DOI 10.1006/bbrc.2001.6299
674 Salih H, Odongo MR, Gong W, He S, and Du X. 2019. Genome-wide analysis of cotton C2H2-zinc finger
675   transcription factor family and their expression analysis during fiber development. *BMC Plant*
676   *Biology* 19:1-17. DOI 10.1186/s12870-019-2003-8
677 Sazegari S, and Niazi A. 2012. Isolation and molecular characterization of wheat ('Triticum aestivum')
678   Dehydration Responsive Element Binding Factor (DREB) isoforms. *Australian Journal of Crop*
679   *Science* 6:1037-1044.
680 Sharoni AM, Nuruzzaman M, Satoh K, Shimizu T, Kondoh H, Sasaya T, Choi I-R, Omura T, and Kikuchi S.
681   2011. Gene structures, classification and expression models of the AP2/EREBP transcription
682   factor family in rice. *Plant and Cell Physiology* 52:344-360. DOI 10.1093/pcp/pcq196
683 Shen Y-G, Zhang W-K, He S-J, Zhang J-S, Liu Q, and Chen S-Y. 2003. An EREBP/AP2-type protein in
684   Triticum aestivum was a DRE-binding transcription factor induced by cold, dehydration and ABA
685   stress. *Theoretical and Applied Genetics* 106:923-930. DOI 10.1007/s00122-002-1131-x
686 Shi S, Zhang R, Zhao Z, Yang L, and Ge W. 2018. Genome-wide analysis of DREBs subfamily in foxtail
687   millet. *Genomics and Applied Biology* 37:827-835.

688 Shkolnik-Inbar D, and Bar-Zvi D. 2011. Expression of ABSCISIC ACID INSENSITIVE 4 (ABI4) in developing
689   Arabidopsis seedlings. *Plant Signaling & Behavior* 6:694-696. DOI 10.4161%2Fpsb.6.5.14978
690 Shu Y, Liu Y, Zhang J, Song L, and Guo C. 2016. Genome-wide analysis of the AP2/ERF superfamily genes
691   and their responses to abiotic stress in Medicago truncatula. *Frontiers in Plant Science* 6:1247.
692   DOI 10.3389/fpls.2015.01247
693 Sun J, Peng X, Fan W, Tang M, Liu J, and Shen S. 2014. Functional analysis of BpDREB2 gene involved in
694   salt and drought response from a woody plant Broussonetia papyrifera. *Gene* 535:140-149. DOI
695   10.1016/j.gene.2013.11.047
696 Vazquez-Hernandez M, Romero I, Escribano MI, Merodio C, and Sanchez-Ballesta MT. 2017. Deciphering
697   the role of CBF/DREB transcription factors and dehydrins in maintaining the quality of table
698   grapes cv. autumn royal treated with high CO2 levels and stored at 0 C. *Frontiers in Plant Science*
699   8:1591. DOI 10.3389%2Ffpls.2017.01591
700 Voorrips R. 2002. MapChart: software for the graphical presentation of linkage maps and QTLs. *Journal
701   of heredity* 93:77-78. DOI 10.1093/jhered/93.1.77
702 Wang L, Ma H, and Lin J. 2019. Angiosperm-wide and family-level analyses of AP2/ERF genes reveal
703   differential retention and sequence divergence after whole-genome duplication. *Frontiers in
704   Plant Science* 10:196. DOI 10.3389/fpls.2019.00196
705 Wang M, Zhuang J, Zou Z, Li Q, Xin H, and Li X. 2017. Overexpression of a Camellia sinensis DREB
706   transcription factor gene (CsDREB) increases salt and drought tolerance in transgenic
707   Arabidopsis thaliana. *Journal of Plant Biology* 60:452-461. DOI 10.1007/s12374-016-0547-9
708 Wang X, Chen X, Liu Y, Gao H, Wang Z, and Sun G. 2011. CkDREB gene in Caragana korshinskii is involved
709   in the regulation of stress response to multiple abiotic stresses as an AP2/EREBP transcription
710   factor. *Molecular Biology Reports* 38:2801-2811. DOI 10.1007/s11033-010-0425-3
711 Wendel JF, Jackson SA, Meyers BC, and Wing RA. 2016. Evolution of plant genome architecture. *Genome
712   Biology* 17:1-14. DOI 10.1186/s13059-016-0908-1
713 Wu H, Lv H, Li L, Liu J, Mu S, Li X, and Gao J. 2015. Genome-wide analysis of the AP2/ERF transcription
714   factors family and the expression patterns of DREB genes in Moso Bamboo (Phyllostachys
715   edulis). *PLoS One* 10:e0126657. DOI 10.1371/journal.pone.0126657
716 Wu Z, Liang J, Zhang S, Zhang B, Zhao Q, Li G, Yang X, Wang C, He J, and Yi M. 2018. A canonical DREB2-
717   type transcription factor in lily is post-translationally regulated and mediates heat stress
718   response. *Frontiers in Plant Science* 9:243. DOI 10.3389/fpls.2018.00243
719 Xianjun P, Xingyong M, Weihong F, Man S, Liqin C, Alam I, Lee B-H, Dongmei Q, Shihua S, and Gongshe L.
720   2011. Improved drought and salt tolerance of Arabidopsis thaliana by transgenic expression of a
721   novel DREB gene from Leymus chinensis. *Plant Cell Reports* 30:1493-1502. DOI 10.1007/s00299-
722   011-1058-2
723 Xie Z, Nolan T, Jiang H, Tang B, Zhang M, Li Z, and Yin Y. 2019. The AP2/ERF transcription factor TINY
724   modulates brassinosteroid-regulated plant growth and drought responses in Arabidopsis. *The
725   Plant Cell* 31:1788-1806. DOI 10.1105%2Ftpc.18.00918
726 Yang Y, Wu J, Zhu K, Liu L, Chen F, and Yu D. 2009. Identification and characterization of two
727   chrysanthemum (Dendronthema× moriforlium) DREB genes, belonging to the AP2/EREBP family.
728   *Molecular Biology Reports* 36:71-81. DOI 10.1007/s11033-007-9153-8
729 Yu Z, Wang X, and Zhang L. 2018. Structural and functional dynamics of dehydrins: a plant protector
730   protein under abiotic stress. *International Journal of Molecular Sciences* 19:3420. DOI
731   10.3390%2Fijms19113420
732 Zhao K, Shen X, Yuan H, Liu Y, Liao X, Wang Q, Liu L, Li F, and Li T. 2013. Isolation and characterization of
733   dehydration-responsive element-binding factor 2C (MsDREB2C) from Malus sieversii Roem.
734   *Plant and Cell Physiology* 54:1415-1430. DOI 10.1093/pcp/pct087

735   Zhao T, Liang D, Wang P, Liu J, and Ma F. 2012. Genome-wide analysis and expression profiling of the
736   DREB transcription factor gene family in Malus under abiotic stress. *Molecular Genetics and*
737   *Genomics* 287:423-436. DOI 10.1007/s00438-012-0687-7
738   Zhao T, Xia H, Liu J, and Ma F. 2014. The gene family of dehydration responsive element-binding
739   transcription factors in grape (Vitis vinifera): genome-wide identification and analysis,
740   expression profiles, and involvement in abiotic stress resistance. *Molecular Biology Reports*
741   41:1577-1590. DOI 10.1007/s11033-013-3004-6
742   Zhou M-L, Ma J-T, Pang J-F, Zhang Z-L, Tang Y-X, and Wu Y-M. 2010. Regulation of plant stress response
743   by dehydration responsive element binding (DREB) transcription factors. *African Journal of*
744   *Biotechnology* 9:9255-9269. DOI 10.1093/jxb/erx118
745   Zhou W, Jia C-G, Wu X, Hu R-X, Yu G, Zhang X-H, Liu J-L, and Pan H-Y. 2016. ZmDBF3, a novel
746   transcription factor from maize (Zea mays L.), is involved in multiple abiotic stress tolerance.
747   *Plant Molecular Biology Reporter* 34:353-364. DOI 10.1007/s11105-015-0926-2
748   Zhou Y, Zhou W, Liu H, Liu P, and Li Z. 2020. Genome-wide analysis of the soybean DREB gene family:
749   Identification, genomic organization and expression profiles in response to drought stress. *Plant*
750   *Breeding* 139:1158-1167. DOI 10.1111/pbr.12867

751

Figure 1

Phylogenetic tree of DREB proteins among *Arabidopsis* and *S. tuberosum*

StDREB proteins were assigned to six distinct subgroups (A1-A6) based on the classification of *Arabidopsis*. The six subgroups were highlighted with different colors.

Figure 2

Analysis of phylogenetic evolutionary relationship, intron-exon organization and protein motif patterns of StDREBs

(A) Phylogenetic tree of 66 StDREB members was constructed with neighbor-joining method using Mega X. The 66 StDREB proteins were then classified into six subgroups (A1-A6). **(B)** Gene structure was visualized using an online tool Gene Structure Display Server 2.0. The maroon boxes represented exons and the black line represented intron. The scale at the bottom showed the exon sizes. The numbers 0, 1, and 2 depicted the intron splicing phase. **(C)** Conserved motifs StDREB proteins were identified using MEME software. Fifteen predicted motifs were represented by distinct colored boxes and the grey lines indicated non-conserved regions.

A

B

C

Legend:
Exon   Intron   0 1 2: intron phase

1. 2. 3. 4. 5. 6. 7. 8.
9. 10. 11. 12. 13. 14. 15.

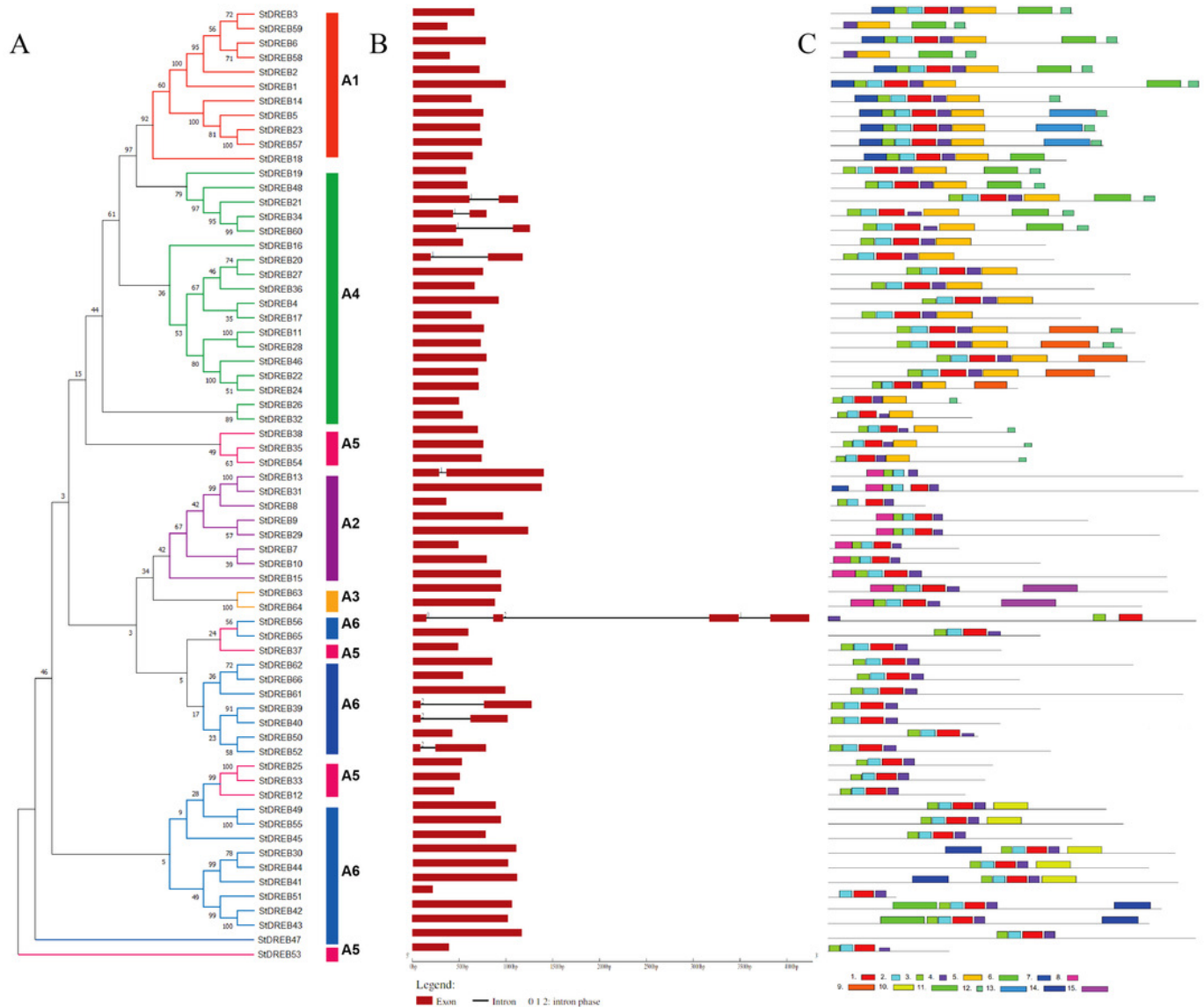



Figure 3

Genomic distribution of 66 StDREB genes across 12 potato chromosomes

Vertical bars indicated locus of StDREB genes on their respective chromosomes. The x-axis scale represented the chromosome length while tandem duplicated gene pairs harbored by various chromosomes were indicated in red.

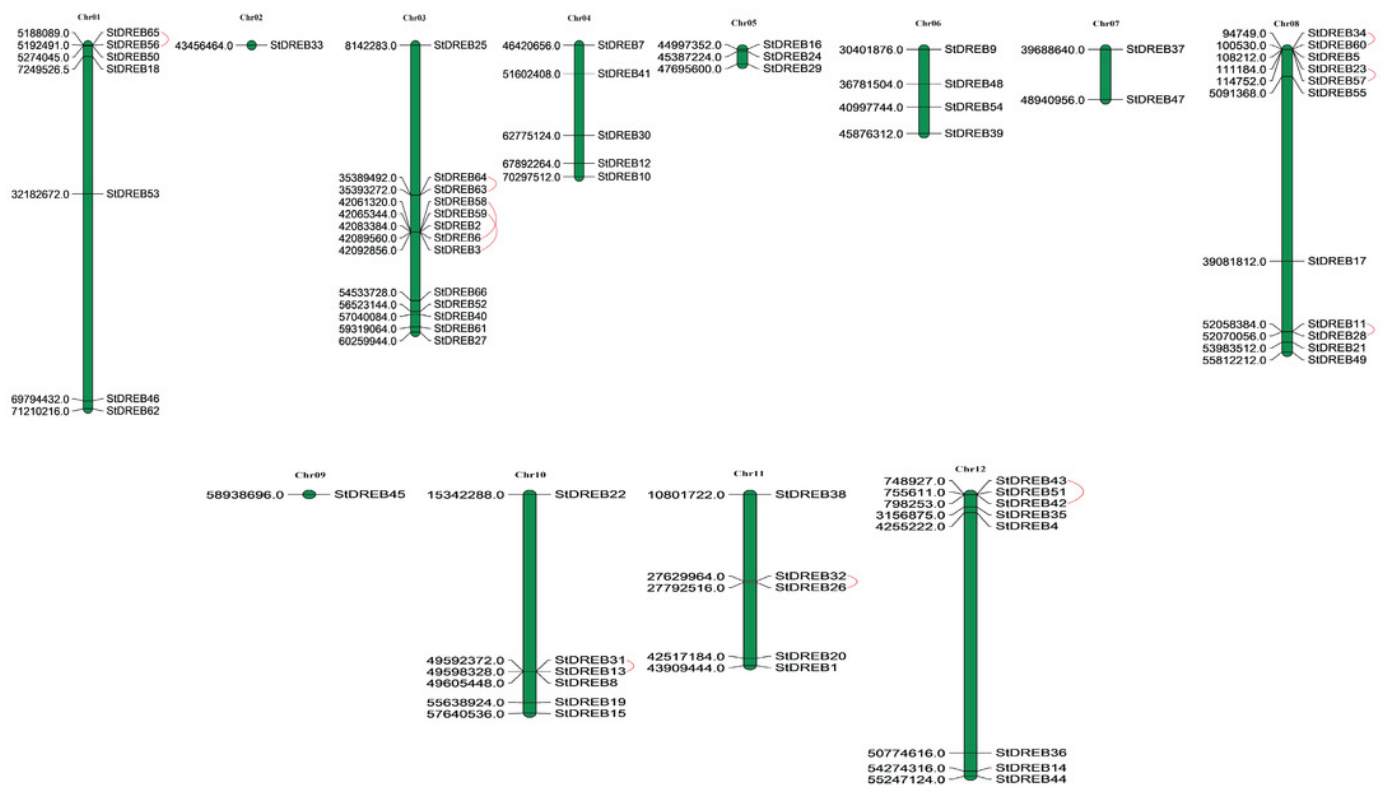



Figure 4

Segmental gene duplication events exhibited by StDREB genes across 12 potato chromosomes

Red lines represented segmental duplication of StDREB gene pairs. Chromosome number is located at top of each green block

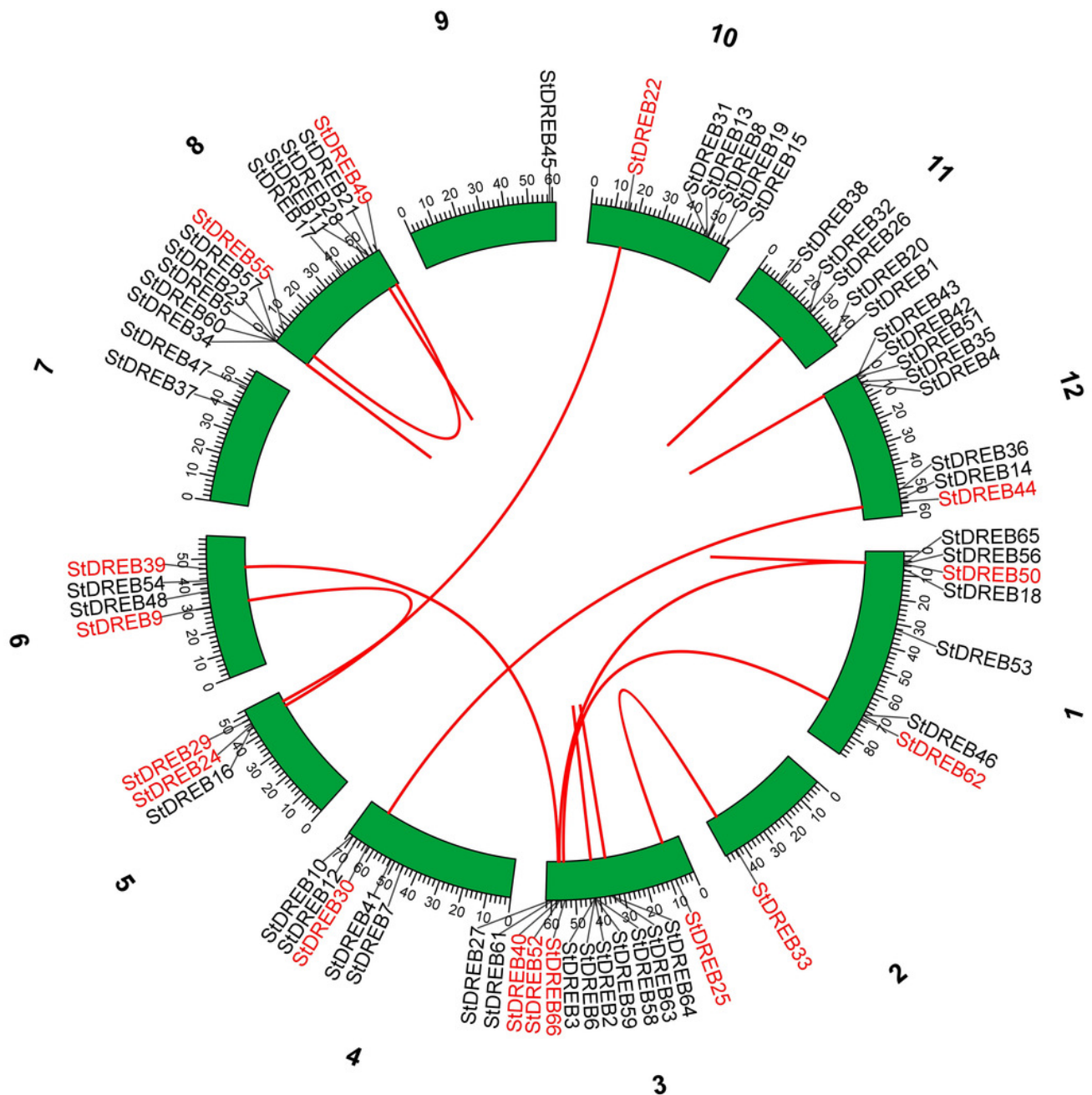



Figure 5

Synteny relationship analysis of DREB genes between *S. tuberosum*, *A. thaliana* and *S. lycopersicum*

Red and green bars represented the chromosomes while red lines indicated the homology and evolutionary link of StDREB genes with *A. thaliana* and *S. lycopersicum*, respectively. In addition, grey lines indicated all the collinear blocks present in their respective genomes.

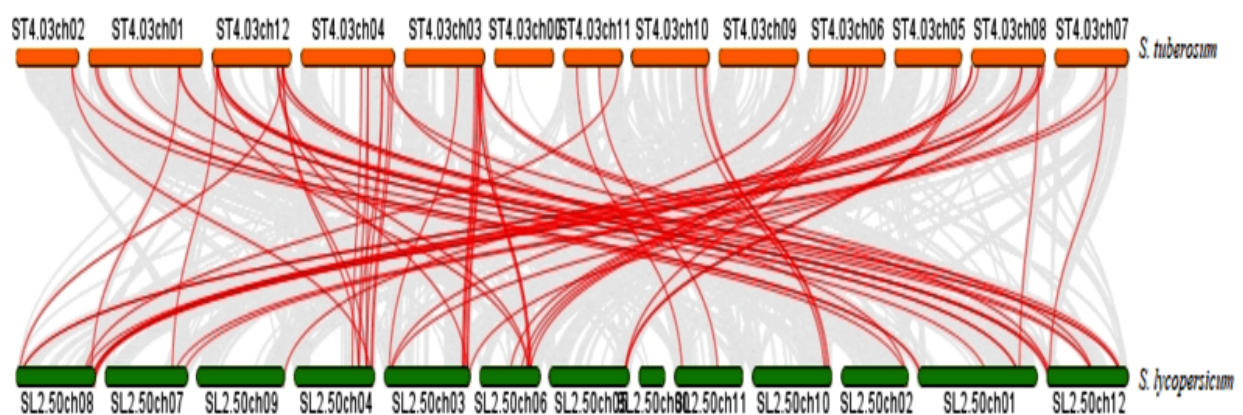
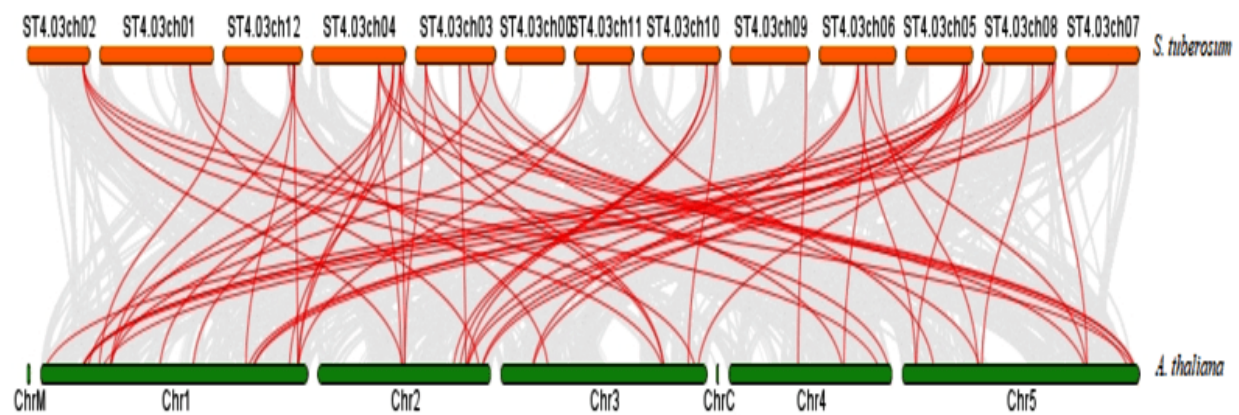



Figure 6

Gene Ontology (GO) and functional enrichment analysis of StDREB genes by using Blast2GO software

(A) GO distribution by level processes. **(B)** Direct count GO. Green bars represented biological processes (BP), blue bars represented molecular functions (MF), and yellow bars represented cellular compartments (CC)

(a)

Biological processes

Molecular Functions

Cellular Compartments

(b)

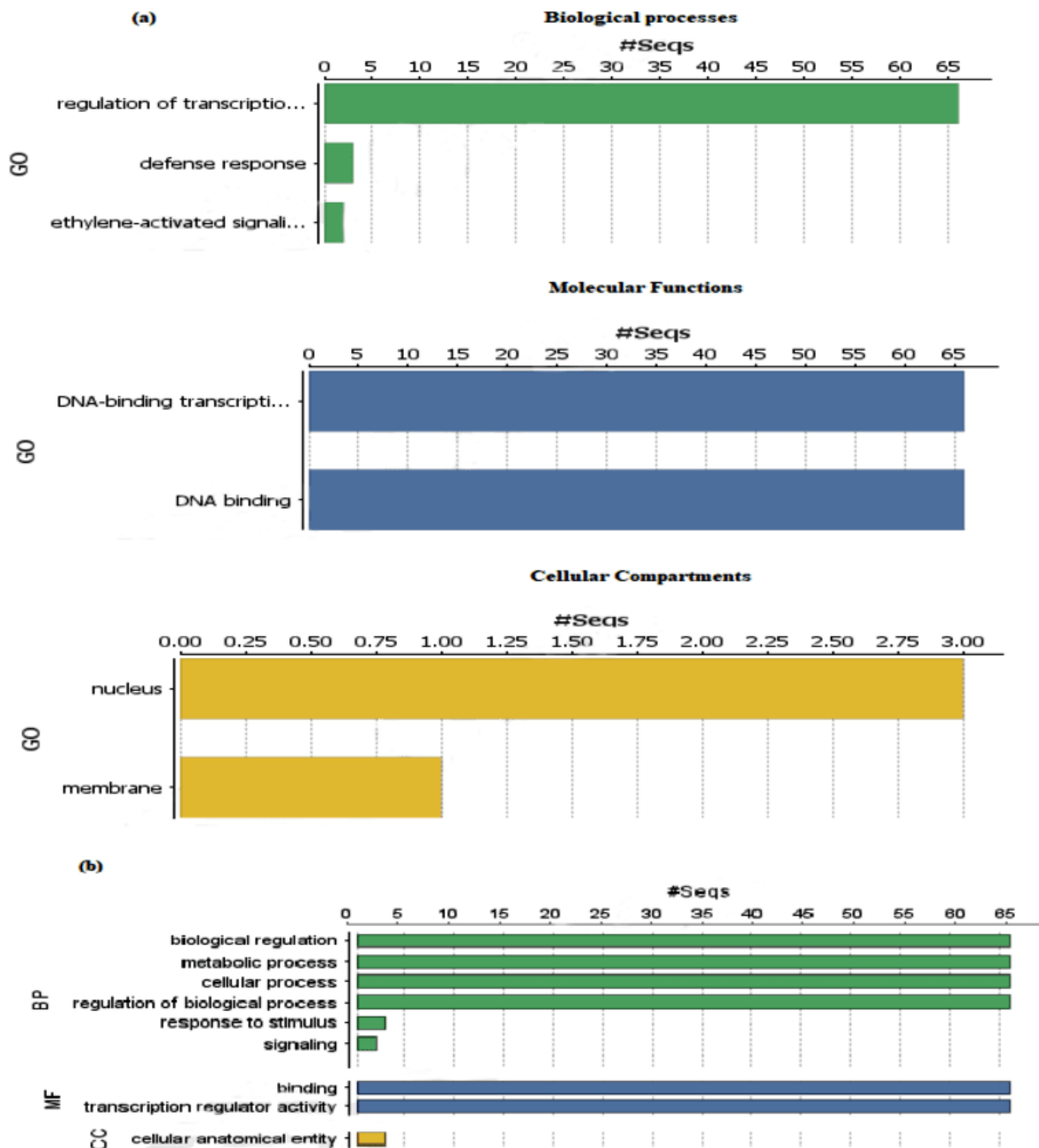
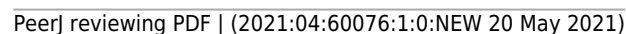



Figure 7

Heat map of subcellular localization predictions for 66 StDREB proteins

The color scale represents abundance of StDREBs in various cellular compartments.

Conserved AP2 domain sequences from six representative StDREB genes

1 **Table 1:**

2 **AP2 domain sequences identified in StDREB genes by Pfam.** Different colors indicated alignment of
3 the AP2 domain. First row depicted consensus region of the Hidden Markov Model (HMM); second row
4 depicted the match between query protein sequences and HMM; third row demonstrated the degree of
5 confidence in each individually aligned residue; whereas fourth row demonstrated the query sequence.

Gene ID	PF00847 AP2 Domain	Sequence Alignment	HMM Length
StDREB3	#HMM pkikGVrydkkrgrWvAewsk.ngkrkkkrfsvgkyGf.eeAkraAiaarkkleg #MATCH p+++GVr+ +g+Wv+e ++ n+ k+r+++g++ + e+A+ra++ a+ +l+g #PP 78********.5******999775...9**********************998 #SEQ PVYRGVRMRN-SGKWVCEVREpNK--- KTRIWLGTFPTaEMAARAHDVAAIALRG	60-110	54
StDREB12	#HMM ikGVrydkkrgrWvAewsk.ngkrkkkrfsvgkyGf.eeAkraAiaarkkleg #MATCH ++G+r +k +g+WvAe ++ n+ r+++g y + A+ra++ a ++l+g #PP 9*******.9******999775...9*********************9998 #SEQ YRGIRMRK--WGKWVAEIREpNK---RSRIWLGSYSSpVAAARAYDTAVFYLRG	18-66	54
StDREB13	#HMM pkikGVrydkkrgrWvAews #MATCH +k++GVr++ +g+WvAe + #PP 79********.9******65 #SEQ CKYRGVRQRT-WGKWVAEIR	69-87	54
StDREB28	#HMM pkikGVrydkkrgrWvAewsk.ngkrkkkrfsvgkyGf.eeAkraAiaarkklege #MATCH p+ GVr++ +g+Wv e ++ + kk r+++g++ + e+A+ra++ a+ ++ g+ #PP 679*******.9******97744...69********************99986 #SEQ PIYHGVRKRS-WGKWVSEIREpR--- KKSRIWLGTFSTaEMAARAHDVAAIAIKGH	58-109	54
StDREB47	#HMM ikGVrydkkrgrWvAe...wskngkrkkkrfsvgkyGf.eeAkraAiaarkkleg #MATCH ++GVr+++ +g+WvAe + ++ ++r ++g++++ e A++a+++ ++kl+g #PP 9*******.9******5554..55...8*******************9999987 #SEQ YRGVRQRH-WGKWVAEirlP--RN--- RTRLWLGTFDTaEDAAMAYDREAYKLRG	184-232	54
StDREB63	#HMM kikGVrydkkrgrWvAewsk.ngkrkkkrfsvgkyGf.eeAkraAiaark #MATCH +++GVr++ +g+WvAe ++ + k++r ++g++ + e A+ra+++a+ #PP 69*******.9******87744...69*******************996 #SEQ RYRGVRQRS-WGKWVAEIREpR---KRTRRWLGTFATaEDAARAYDRAAI	64-109	54