




Analyses of Sequence Features and Expression Patterns of *DOF* Transcription Factor Genes in Pearl Millet (*Pennisetum glaucum* (L.) R. Br.)

Yingwei Qu¹ · Shashi Kumar Gupta² · Ambika Dudhate³ · Harshraj Subhash Shinde⁴ · Shenkui Liu⁵ · Tetsuo Takano¹ · Daisuke Tsugama¹ 

Received: 30 November 2022 / Accepted: 8 June 2023 / Published online: 24 June 2023
© The Author(s) 2023

Abstract

Pearl millet (*Pennisetum glaucum* (L.) R. Br.) is a cereal crop that is tolerant to drought, high-temperature and poor-nutrient stresses throughout its life cycle. The *DOF* (*DNA binding with One Finger*) family genes encode plant-specific zinc finger transcription factors. These transcription factors have a variety of functions and can be involved in regulating the stress tolerance of pearl millet. In this study, we identified 12 *DOF* family genes (*PgDOFs*) in pearl millet. Ten of them were distributed on four chromosomes and the other two were on scaffolds (i.e., non-chromosomal sequences). Protein sequence analysis showed that *PgDOFs* have the *DOF* domain in their N-terminal regions and that other conserved motifs are also present in them. *PgDOFs* and *DOF* proteins from five other plant species were divided into seven groups, and the 12 *PgDOFs* could be classified into six of those seven groups. In gene expression analysis, most *PgDOFs* were upregulated by cold, heat, dehydration and salinity stress. These data can be useful for further characterization of the *PgDOFs*.

Keywords Pearl millet · DNA-binding one zinc finger transcription factor · Abiotic stress · Gene expression

Introduction

The *DOF* (DNA-binding one zinc finger) protein family is a plant-specific transcription factor family. A *DOF* protein was first identified in maize (*Zea mays*) (Yanagisawa

and Izui 1993), and later found in many other plant species. Unlike other zinc-finger domains, the *DOF* domain includes a C2C2-type zinc-finger-like motif (Yanagisawa 2002). Most *DOF* proteins consist of 200–400 amino acids and have the *DOF* domain with 50–52 amino acids at their N-terminal regions. The *DOF* domain consists of 50–52 amino acids and can bind a *cis*-regulatory elements with the core sequence 5'-T/AAAG-3' (Yanagisawa 2002). In contrast to the *DOF* domain, C-terminal regions of *DOF* proteins are variable.

DOF proteins regulate their target genes and thereby regulate various biological processes. There are 36 *DOF* genes in *Arabidopsis thaliana*, 30 in rice (*Oryza sativa*) (Lijavetzky et al. 2003), 28 in sorghum (Kushwaha et al. 2011), 34 in tomato (Cai et al. 2013), 37 in Chinese cabbage (Ma et al. 2015), 32 in potato (Venkatesh and Park 2015), 38 putative *DOF* genes in pigeon pea (Malviya et al. 2015), 36 in cucumber (Wen et al. 2016), 33 in pepper (Kang et al. 2016), 36 in common beans (Ito et al. 2017), 35 in foxtail millet (Zhang et al. 2017), 41 in poplar (Wang et al. 2017), 114 in cotton (Li et al. 2018), 29 in eggplant (Wei et al. 2018), 24 in durian (Khaksar et al. 2019), 40 in alfalfa (Cao et al. 2020),

Communicated by Ray Ming.

✉ Daisuke Tsugama
tsugama@g.ecc.u-tokyo.ac.jp

¹ Asian Research Center for Bioresource and Environmental Sciences (ARC-BRES), The University of Tokyo, Nishi-tokyo-shi, Tokyo, Japan

² International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Hyderabad, Telangana State, India

³ Sequencing and genome discovery center, Stowers Institute for Medical Research, Kansas City, MO, United States of America

⁴ Department of Animal and Food Sciences, University of Kentucky, Lexington, KY, United States of America

⁵ State Key Laboratory of Subtropical Silviculture, Zhejiang A and F University, Lin'an, Hangzhou, China

33 in tef (Mulat and Sinha 2020), 36 in watermelon (Zhou et al. 2020), 50 in *Cleistogenes songorica* (Wang et al. 2021), 39 in common walnut (Khan et al. 2021), 117 in rapeseed (Lohani et al. 2021), and 51 in olive (Mariyam et al. 2021). However, no *DOF* genes have been identified in pearl millet (*Pennisetum glaucum* (L.) R. Br., also known as *Cenchrus americanus* (L.) Morrone).

Pearl millet is a cross-pollinated diploid C4 crop with seven pairs of chromosomes ($2n=14$) and a ~1.79-Gb genome (Varshney et al. 2017). Pearl millet is one of the most cultivated millets in the world. Compared with other crops such as maize and wheat (*Triticum aestivum*), pearl millet exhibits higher tolerance to stressed conditions such as a drought, low soil fertility and a high temperature (Basavaraj et al. 2010). In this study, we performed a comprehensive analysis of pearl millet *DOF* genes (*PgDOFs*) as candidates for regulators of pearl millet stress tolerance.

Materials and Methods

Plant Materials

The pearl millet drought-tolerant genotype ICMB 843 (Dudhate et al. 2018) was provided by ICRISAT (International Crops Research Institute for the Semi-Arid Tropics, India) and used for this study. Approximately 60 seeds were sown on soil with fertilizers in pots. The resulting plants were grown under a long-day (16-hour light/8-hour darkness) condition at 28 °C in a growth chamber. For gene expression analysis using quantitative real time PCR (qRT-PCR), 28-day-old seedling were subjected to (a) dehydration stress induced by PEG-6000 15% (w/v), (b) salinity stress by 250mM NaCl, (c) high temperature stress of 42 °C, and (d) low temperature stress of 4 °C for 6, 12 and 24 h. Leaves and roots of these plants were sampled and stored at -80 °C for RNA isolation followed by cDNA synthesis.

Identification and Annotation of *PgDOFs* in Pearl Millet

To find *PgDOFs*, the *DOF* sequences of Arabidopsis, rice, and foxtail millet from the Uniprot (<https://www.uniprot.org>) database, using the keyword “DOF”. Alignment of these protein sequences were performed by Clustal W (Thompson et al. 1994). A profile of this multiple sequence alignment was generated by hmmbuild and used for hmmsearch on HMMER (v3.32) with the pearl millet protein sequences (Varshney et al. 2017) (<https://cegsb.icrisat.org/ipmgsc/genome.html>) as queries. This search returned 37 protein sequences with that profile. The presence of the *DOF* domains in these 37 protein sequences was confirmed

by CD-search (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) and Pfam (<https://pfam.xfam.org/family/PF02701>). These detected *DOF* domains in 12 of those 37 sequences, and these 12 proteins were regarded as *PgDOFs*. Logos for the *DOF* domain sequences were obtained by WebLogo 2.8.2 (<http://weblogo.berkeley.edu/logo.cgi>) (Crooks et al. 2004). The isoelectric point (pI) and molecular weight (MW) of *PgDOFs* were obtained by ProtParam (<https://web.expasy.org/protparam/>) (Gasteiger et al. 2005). Subcellular localization of *PgDOFs* was predicted by WoLF PSORT (<https://wolfpsort.hgc.jp>) (Horton et al. 2007). Annotations of *Arabidopsis thaliana* and rice homologs of *PgDOFs* were downloaded from TGIF-DB (<http://web-park2116.sakura.ne.jp/rlgpr/>) (Tsugama et al. 2021).

Chromosomal Localization and Gene Structure Analysis

Positions of *PgDOFs* on chromosomes were displayed by Map Gene 2 Chromosome (MG2C) v2 (http://mg2c.iask.in/mg2c_v2.0/) (Chao et al. 2021). The exon-intron structures of *PgDOFs* were obtained by comparing their cDNA sequences with their genomic sequences on Gene Structure Display Server 2.0 (<http://gsds.gao-lab.org>) (Hu et al. 2015).

Motifs Analysis and Phylogenetic Tree Construction

The MEME Suite 5.5.0 (<http://meme-suite.org/tools/meme>) (Bailey et al. 2015) was used to examine the conserved motifs in *PgDOFs*. The maximum number of motifs was set at 15 and the minimum and maximum width at 6 and 50 amino acids, respectively. Eukaryotic linear motifs (ELMs) (Kumar et al. 2022) for *PgDOF*1-12 were downloaded from TGIF-DB (Tsugama et al. 2021). A phylogenetic tree of *PgDOFs* and *DOF* proteins of *Setaria itatica* (foxtail millet), *Oryza sativa* (rice), sorghum, maize, and Arabidopsis was generated with the ClustalW multiple-alignment program and the neighbor-joining method, and evaluated by bootstrapping with 1000-time resampling. The resulting phylogenetic tree was displayed on iTOL v6 (Letunic and Bork 2019).

Synten Analysis

All amino acid sequences as well as general feature-formatted (GFF) genome annotation files of pearl millet, Arabidopsis, rice, maize and foxtail millet were downloaded from the International Pearl Millet Genome Sequencing Consortium website (<https://cegsb.icrisat.org/ipmgsc/>), The Arabidopsis Information Resource (TAIR) (<https://www.arabidopsis.org/index.jsp>), the Rice Genome Annotation Project website (<http://rice.uga.edu/>), the Ensembl Plants

Zea mays website (http://plants.ensembl.org/Zea_mays/Info/Index) and the Ensembl Plants *Setaria italica* website (https://plants.ensembl.org/Setaria_italica/Info/Index), respectively. BLASTP in the BLAST+ suite (Camacho et al. 2009) was run with all combinations of those sequences as a query and a database. For this analysis, the option was set as “-outfmt 6 -num_alignments 5” to use the tabular output format and five as the maximum alignment number per query. The resulting files and the above GFF files were concatenated and used as the input for MCScanX (Wang et al. 2012) to analyze colinear syntenic blocks of genes. Default options were used for MCScanX.

In silico Analysis of cis-acting Elements in Promoters of PgDOFs

Two thousand-bp upstream sequences from start codons were used as promoters as previously described (Zhang et al. 2021) for *PgDOFs*. Such sequences were extracted from the pearl millet genome sequence. Known *cis*-acting elements in these sequences were identified by PlantCare (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>) (Lescot et al. 2002).

RNA Extraction and Expression of PgDOFs Under Stressed Conditions

Total RNA was prepared from leaves and roots of the stressed 28-day-old pearl millet plants (see the above subsection “Plant materials”) with the NucleoSpin RNA Plant kit (MACHEREY-NAGEL, Germany). cDNA was prepared from 1 µg of the total RNA with PrimeScript Reverse Transcriptase (Takara Bio, Japan) and the oligo (dT) primer, and used as the template for qRT-PCR. StepOne Real-Time PCR System (Applied Biosystems, USA) and TB Green Premix Ex Taq (Takara Bio, Japan) were used to perform the qRT-PCR. The PCR cycle used was 95 °C for 30 s, 40 cycles of 95 °C for 5 s, and 60 °C for 30 s. Primers used for the

analysis are listed in Table S1. The pearl millet *Actin (ACT)* gene (Anup et al. 2017) was used as the internal control.

Results

Identification of PgDOFs

Twelve *PgDOFs* (*PgDOF1-12*) were identified by our analysis. Among *PgDOF1-12*, *PgDOF12* was the smallest with 262 amino acids, and *PgDOF4* was the largest with 720 amino acids. The pI of *PgDOF1-12* was between 4.62 (for *PgDOF7*) and 11.92 (for *PgDOF12*). *PgDOF1*, *PgDOF2*, *PgDOF4*, and *PgDOF12* were predicted to be localized to the plastids, whereas the other *PgDOFs* were predicted to be localized to the nucleus (Table 1).

PgDOF1-10 were located on four of the seven pearl millet chromosomes: Three on Chromosome 2, three on Chromosome 5, three on Chromosome 6, and one on Chromosome 7 (Fig. 1). *PgDOF11* and *PgDOF12* were located on scaffolds (sequences that are unassigned to chromosomes). These genes were physically distant from each other, suggesting that none of them is a tandemly duplicated gene. *PgDOF1-10* and genes in their proximity were present in colinear syntenic blocks with their homologs from rice, maize and foxtail millet, whereas *PgDOF11* and *PgDOF12* were not (Dataset S1). This result suggests that at least *PgDOF1-10* can be segmentally duplicated genes.

Phylogenetic Classification and Protein Sequence Alignment of PgDOFs

A phylogenetic tree was generated, including *PgDOFs* and *DOF* proteins from other five species: foxtail millet, rice, sorghum, maize, and Arabidopsis. The tree suggests that these *DOF* proteins could be divided into seven groups (Group I-VII) (Fig. 2). Group III did not contain any of *PgDOF1-12* but the other groups contained at least one of *PgDOF1-12*. The

Table 1 The 12 *PgDOFs* identified in pearl millet and their sequence features

Name	CDS length (bp)	Protein length (aa)	Gene ID	pI	Predicted subcellular localization	Molecular weight
<i>PgDOF1</i>	900	299	Pgl_GLEAN_10013736	9.43	Plastid	31109.13
<i>PgDOF2</i>	1398	465	Pgl_GLEAN_10022608	9.07	Plastid	49029.00
<i>PgDOF3</i>	813	270	Pgl_GLEAN_10032064	9.06	Nucleus	27174.23
<i>PgDOF4</i>	2163	720	Pgl_GLEAN_10032145	6.01	Plastid	76766.12
<i>PgDOF5</i>	867	288	Pgl_GLEAN_10016344	9.95	Nucleus	29557.79
<i>PgDOF6</i>	1302	433	Pgl_GLEAN_10002018	8.41	Nucleus	46091.65
<i>PgDOF7</i>	855	284	Pgl_GLEAN_10017694	4.62	Nucleus	30323.59
<i>PgDOF8</i>	1626	541	Pgl_GLEAN_10025744	8.88	Nucleus	58135.9
<i>PgDOF9</i>	1785	594	Pgl_GLEAN_10013279	4.98	Nucleus	64863.49
<i>PgDOF10</i>	1338	445	Pgl_GLEAN_10000709	8.15	Nucleus	47157.66
<i>PgDOF11</i>	948	315	Pgl_GLEAN_10010830	9.34	Nucleus	33229.01
<i>PgDOF12</i>	789	262	Pgl_GLEAN_10000298	11.92	Plastid	27954.51

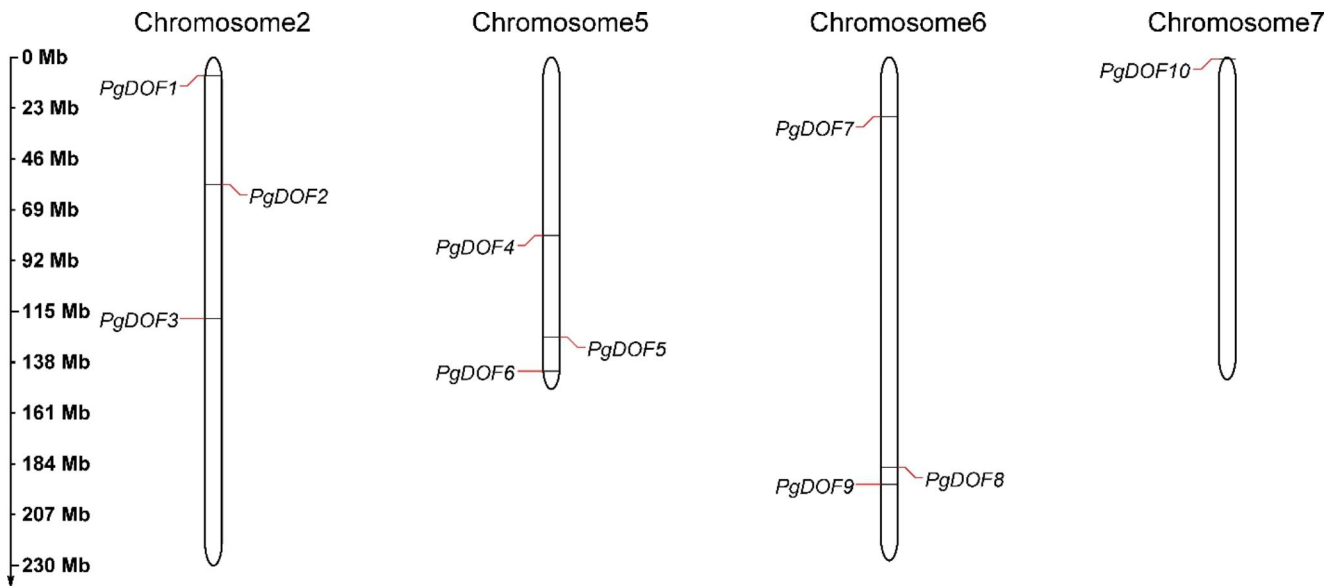
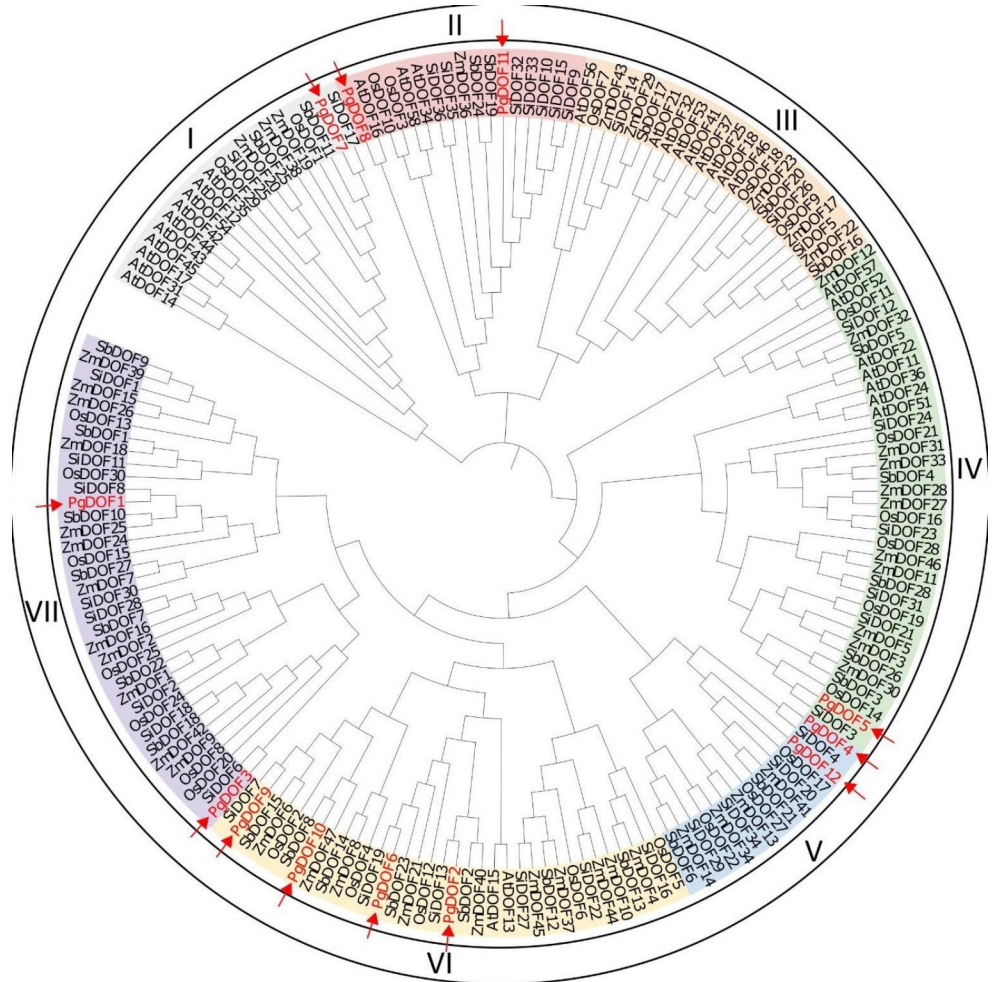


Fig. 1 Positions of *PgDOFs* in pearl millet chromosomes. The chromosome number is indicated at the top of each bar in red. The scale is represented in megabase (Mb)

Fig. 2 Neighbor-joining phylogenetic tree of DOF proteins from pearl millet (*PgDOF*), with foxtail millet (*SiDOF*), rice (*OsDOF*), sorghum (*SbDOF*), maize (*ZmDOF*), and Arabidopsis (*AtDOF*). *PgDOFs* are indicated by arrows. The details about the databases to download the sequences are listed in Table S4



alignment of protein sequences showed that PgDOFs have the conserved DOF domain at their N-terminal regions and their C-terminal regions are variable (Fig. S1).

Exon-intron Structure and Conserved Motif Analysis

PgDOF3 and *PgDOF10* contained only one exon, *PgDOF4* contained six exons, and the other *PgDOFs* contained two or three exons (Fig. 3a).

Fifteen conserved motifs (Motif 1–15) were identified *de novo* in PgDOF1–12 by the MEME program (Table S2). Motif 1 (Fig. S2a) and Motif 3 (Fig. S2b) were found in N-terminal regions of most of the PgDOFs. PgDOF12 contained only one (Motif 1) of those motifs, whereas PgDOF9 contained 10 of those motifs (Fig. 3b). The differences in motif distribution between PgDOF1–12 may be relevant to their functional divergence. ELMs, which are motifs known to be involved in regulating protein functions such as protein modification, protein-protein interactions and subcellular localization (Kumar et al. 2022), were also identified in PgDOF1–12 (Dataset S2).

Analysis of *cis*-acting Elements in Promoters of *PgDOFs*

In the 2000-bp promoters of *PgDOFs*, *cis*-acting elements were identified such as light-responsive elements (TCT-motif and G-box), plant hormone-responsive elements (including abscisic acid-responsive elements, methyl jasmonate-responsive elements, gibberellin-responsive elements, and salicylic acid-responsive element), stress-responsive elements (including anaerobic-induction elements, low temperature-responsive elements, MYB-binding elements, and W-Box (WRKY-binding elements), plant growth associated elements (including the CAT-box for meristem-specific gene expression and the RY-element for seed-specific expression and the circadian rhythm) (Fig. 4 and Table S3).

Expression Pattern of *PgDOFs* Under Various Abiotic Stresses

The expression patterns of *PgDOF1–12* were analyzed by qRT-PCR. All *PgDOFs* were upregulated by PEG (polyethylene glycol)-induced dehydration stress in leaves. *PgDOF4* expression was ~100 times stronger 6 and 12 h after the stress treatment was initiated than 0 h. *PgDOF8* expression was ~400 times stronger 12 and 24 h after the treatment was initiated. In contrast, *PgDOF4*, *PgDOF5* and *PgDOF8* were suppressed by the same treatment in roots (Fig. 5a). Under a salinity (NaCl)-stressed condition, all the *PgDOFs* except *PgDOF8* were upregulated in leaves. Among them, *PgDOF5* was most greatly upregulated (with 400-fold increase 12 h after the stress

was imposed). However, only seven out of 12 *PgDOFs* were upregulated by the salinity stress treatment in roots. *PgDOF5* was even suppressed in the roots (Fig. 5b). A high temperature (42°C) upregulated all the *PgDOFs* in both leaves and roots (Fig. 5c). A low temperature (4°C) upregulated all the *PgDOFs* except *PgDOF2* and *PgDOF12* in leaves, and upregulated all of them in roots (Fig. 5d). Thus, most of the *PgDOFs* were upregulated by the stress treatments in leaves and/or roots. This result is consistent with the finding that *PgDOFs* have stress-responsive *cis*-elements in their promoters (Fig. 4 and Table S3).

Discussion

The DOF gene family contains approximately 30 members in *Arabidopsis*, rice, maize and other species. However, this study identified only 12 putative *DOF* genes, which could be classified into six of those seven groups, in the current version of the pearl millet genome sequence. Some *DOF* genes in pearl millet may have been lost in the course of evolution, or the current pearl millet genome sequence may have insufficient or incorrect parts.

Most *PgDOFs* contain zero-two introns (Fig. 3a). This number is consistent with the number of introns of *DOF* genes of other plant species. *PgDOF1–10* were present in a synteny block with their homologs in rice, maize and foxtail millet (Dataset S1). These findings support the idea that these *DOF* genes are evolutionarily conserved, and their functions can also be conserved. For example, *ZmDOF1*, a maize *DOF* gene, regulates carbohydrate metabolism (Yanagisawa 2004). *TaDOF1*, a wheat *DOF* gene, also regulates the carbon metabolism (Chen et al. 2005). *AtDOF4.7*, an *Arabidopsis DOF* gene, regulates the floral organ abscission (Wei et al. 2010). *PpDOF1*, a *Physcomitrella patens DOF* gene, regulates the growth of nutrient-dependent filament (Sugiyama et al. 2012). *BnCDF1*, a *Brassica napus DOF* gene, regulates flowering time and freezing tolerance (Xu and Dai 2016). *OsDOF3*, a rice *DOF* gene, regulates the biosynthesis of gibberellins (Li et al. 2009). *OsDOF12*, *OsDOF23*, *OsDOF24* and *OsDOF25*, four rice *DOF* genes, regulate flowering time, gene expression in seeds, and carbon and nitrogen metabolism (Yamamoto et al. 2006). Sorghum *DOF* (*SbDOF*) genes are responsive to light and hormones, and regulate endosperm-specific gene expression (Kushwaha et al. 2011). *PgDOFs* may have those functions. Among the 15 conserved motifs identified *de novo*, Motif 10 was present in only PgDOF4 and PgDOF5 (Fig. 3b). It can be interesting to determine whether Motif 10 has a specific function.

PgDOF2, *PgDOF6*, *PgDOF8*, *PgDOF9*, and *PgDOF10* were upregulated by the stress treatments (Fig. 5). These are all homologs of cycling DOF factor (CDF) genes of *Arabidopsis*

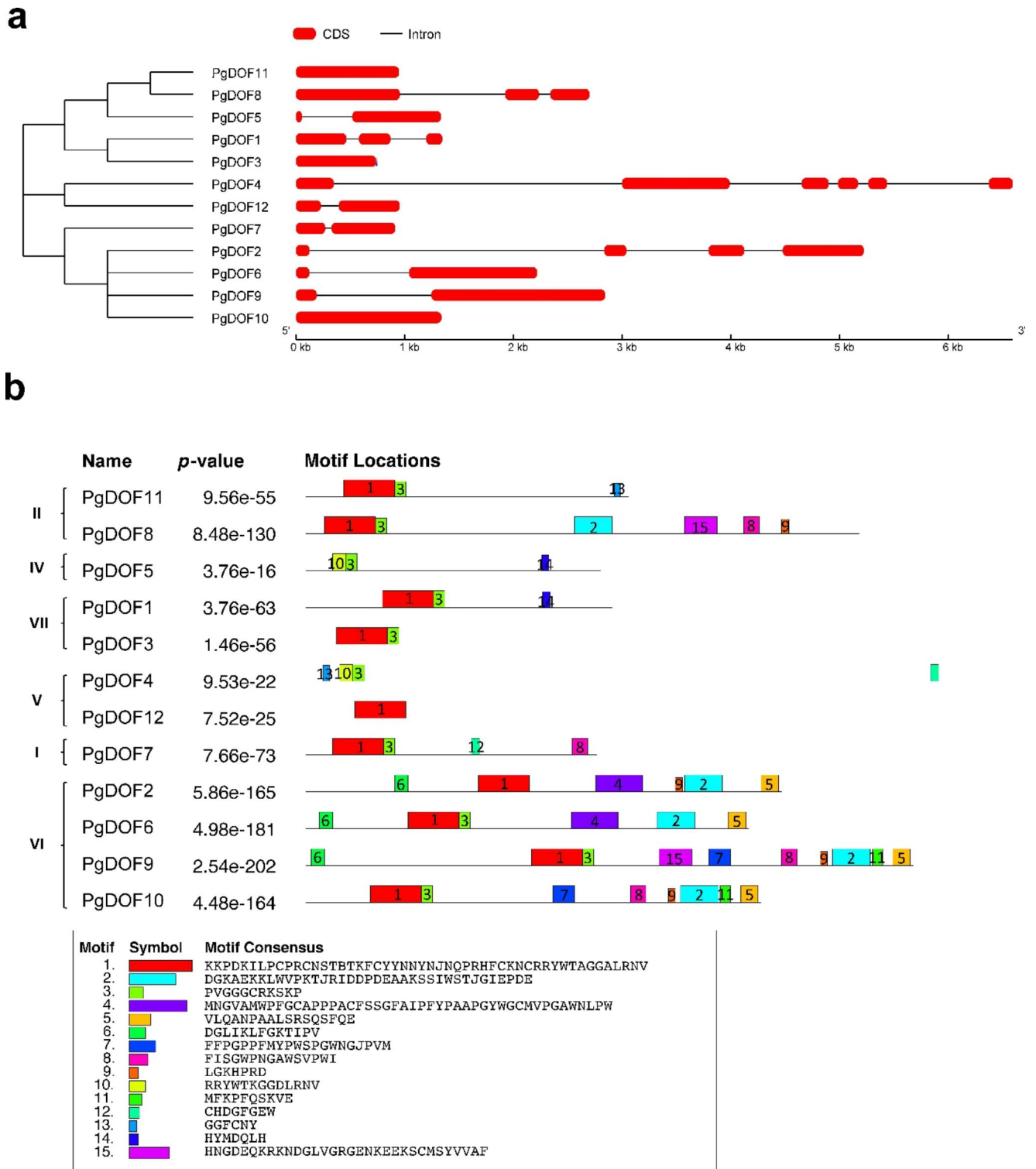


Fig. 3 (a) Exon-intron structures of *PgDOFs*. Exons and introns are indicated by red rectangles and black horizontal lines, respectively. (b) Motifs identified in *PgDOFs*. Motifs indicated by boxes were detected by MEME and the number in boxes (1 to 15) represents Motif 1 to 15, respectively. Box sizes indicate the length of the motifs, and the consensus sequences of these motifs are presented in Table S2

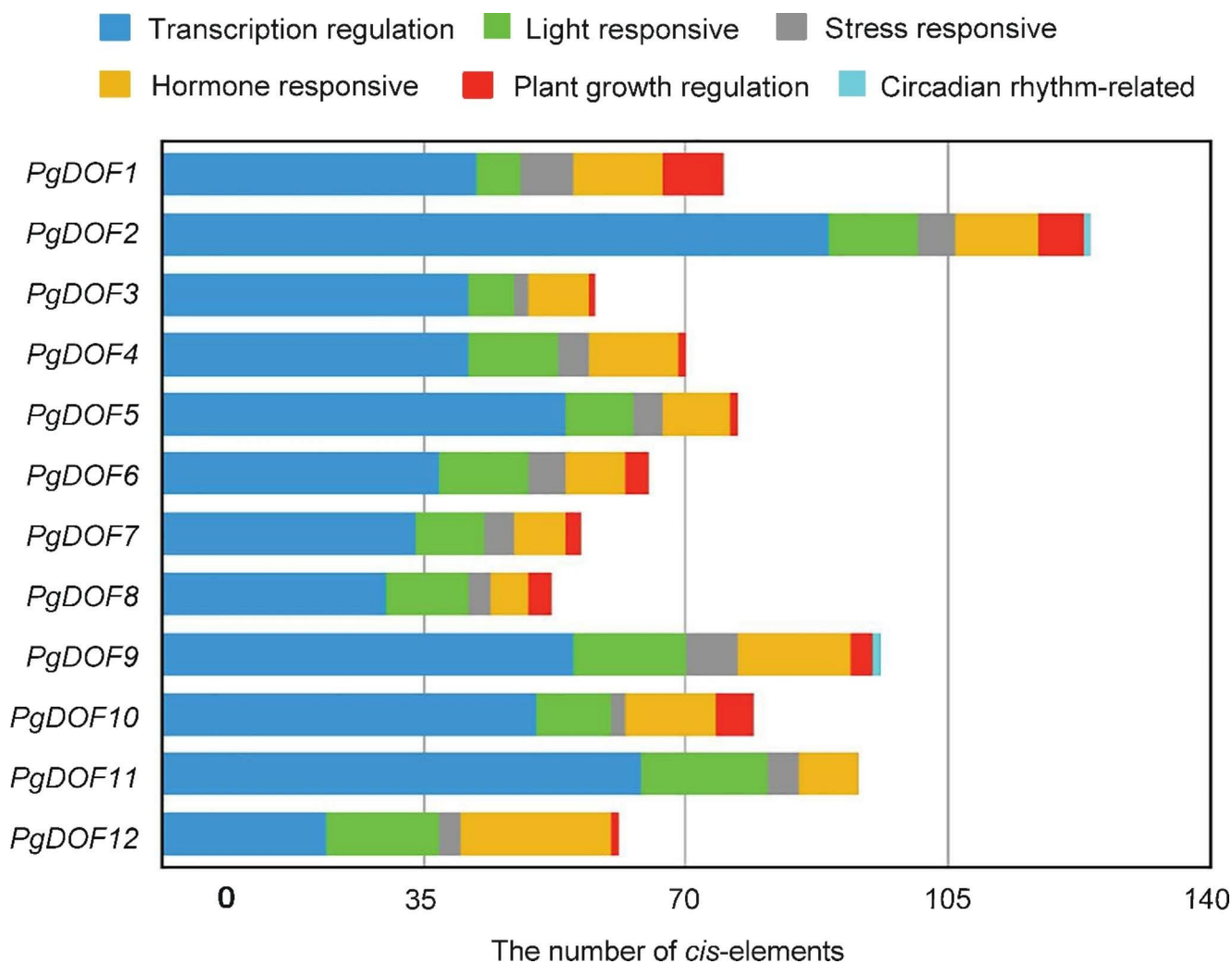
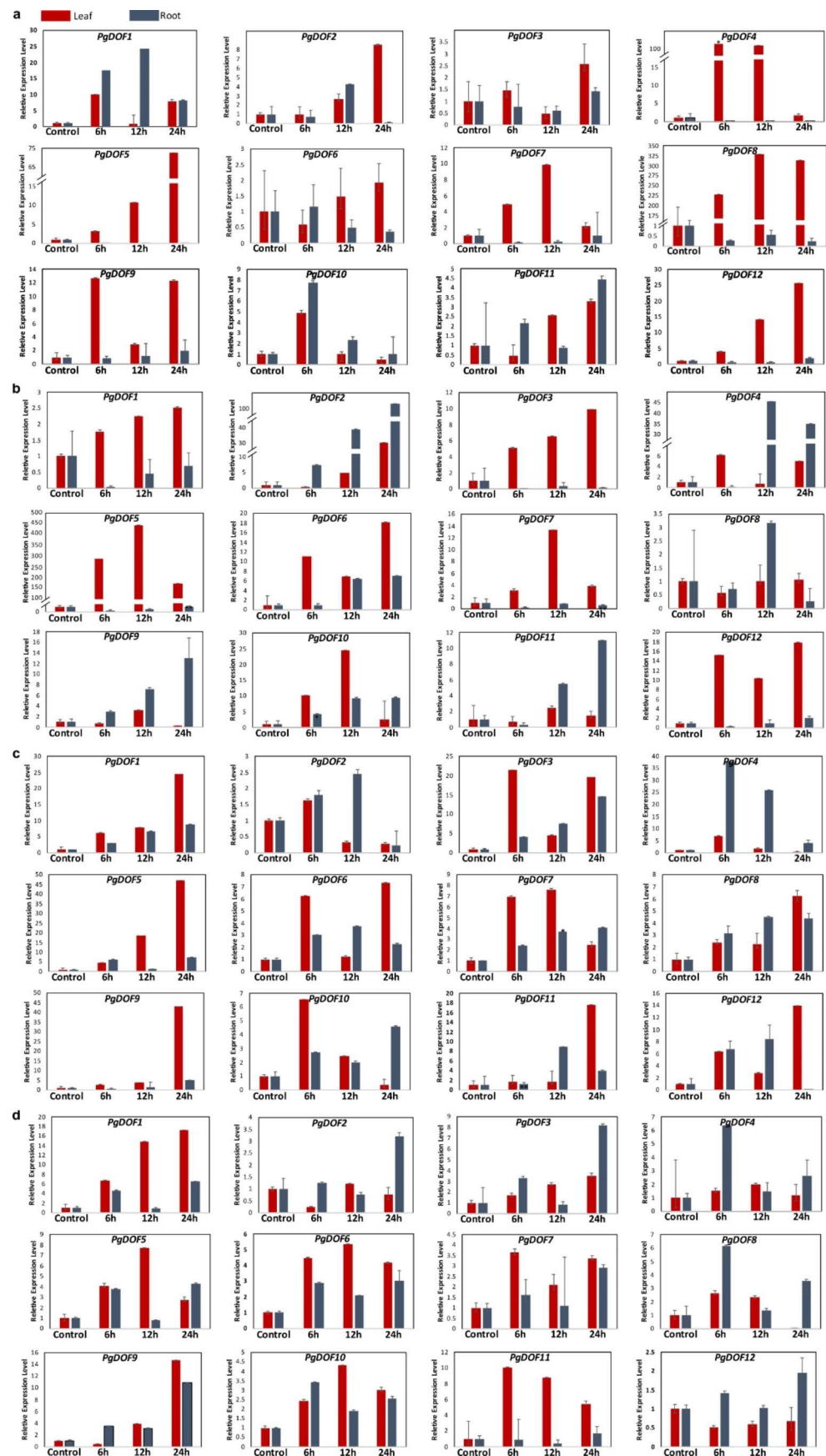


Fig. 4 *Cis*-acting elements in 2000-bp promoters of *PgDOFs*. The information on these *cis*-elements is given in Table S3

and rice (Dataset S3 and S4), although *PgDOF8* is distant from *PgDOF2*, *PgDOF6*, *PgDOF9* and *PgDOF10* in the phylogenetic tree (Fig. 2). In a previous study, overexpression of tomato (*Solanum lycopersicum*) *CDF* genes, *SICDF1* and *SICDF3*, in *Arabidopsis* activated stress-responsive genes, *COR15*, *RD29A* and *ERD10*, and enhanced its tolerance to drought and salinity stresses (Corrales et al. 2014). In another study, a transfer DNA (T-DNA) insertion mutation in an *Arabidopsis* *CDF* gene, *CDF3*, decreased plant tolerance to drought and cold stresses, and *CDF3* overexpression activated *COR15*, *RD29A* and *ERD10* and their upstream transcription factor genes, *CBF1-3*, *DREB2A*, *ZAT10*, and *ZAT12* to enhance the tolerance to drought and cold stresses, suggesting that *CDF3* functions as a positive regulator of drought and cold stress tolerance upstream of those transcription factors (Corrales et al. 2017). In a previous transcriptome analysis, among *PgDOF1-12*, *PgDOF6* as well as *PgDOF12*, which has no close homologs in either

Arabidopsis or rice, was found to be expressed more strongly in the drought-tolerant line ICMB 843 than in the drought-sensitive line ICMB 863 in roots under a drought-stressed condition (Dudhate et al. 2018; Fig. S3). Thus, *PgDOF6* and the other putative *CDF* genes of pearl millet (i.e., *PgDOF2*, *PgDOF8*, *PgDOF9*, and *PgDOF10*) are candidates for *DOF* genes that regulate the pearl millet stress tolerance. *Arabidopsis* *CDFs* repress the expression of a circadian clock-related gene, *CONSTANS*, to inhibit photoperiod-dependent flowering (Imaizumi et al. 2005; Fornara et al. 2009), and both the transcript and protein levels of the *Arabidopsis* *CDFs* are regulated by components of the circadian clock (Imaizumi et al. 2005; Sawa et al. 2007; Fornara et al. 2009). It will also be interesting to determine whether *PgDOF2*, *PgDOF6*, *PgDOF8*, *PgDOF9*, and *PgDOF10* are involved in the circadian rhythm and flowering. The circadian rhythm-related *cis*-elements in the promoters of *PgDOF2* and *PgDOF9* (Fig. 4 and Table S3) may be relevant to their expression patterns during a day.

Fig. 5 (a) The expression levels of *PgDOFs* in leaves and roots in the presence of dehydration (15% (w/v) PEG) stress. (b) The expression levels of *PgDOFs* in leaves and roots in the presence of salinity (250 mM NaCl) stress. (c) The expression levels of *PgDOFs* in leaves and roots in the presence of heat (42 °C) stress. (d) The expression levels of *PgDOFs* in leaves and roots in the presence of cold (4 °C) stress. Four-week-old plants were treated with PEG, NaCl, 42 °C or 4 °C for 6 h, 12 h, and 24 h. The expression levels were calculated by the comparative cycle threshold method. The *Actin* (*ACT*) gene was used as an internal control. Data are means \pm SD.



Abbreviations

CDS	Coding sequence
DOF	DNA-binding one zinc finger
PEG	Polyethylene glycol
qRT-PCR	Quantitative reverse transcription-PCR

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s12042-023-09337-x>.

Acknowledgements We are grateful to ICRISAT, India for providing the pearl millet seeds for this study.

Author Contributions All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Yingwei Qu, Shashi Kumar Gupta, Ambika Dudhate and Harshraj Subhash Shinde. The first draft of the manuscript was written by Yingwei Qu and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding Open access funding provided by The University of Tokyo. This work was supported by JSPS (Japan Society for the Promotion of Science) KAKENHI Grant [grant number JP 19H02928].

Declarations

Competing Interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Anup PC, Melvin P, Kini KR (2017) Reference gene selection and validation for gene expression studies in downy mildew infected pearl millet by quantitative real-time PCR. *Australasian Plant Pathol* 46:441–452. <https://doi.org/10.1007/s13313-017-0506-z>
- Bailey TL, Johnson J, Grant CE, Noble WS (2015) The MEME suite. *Nucleic Acids Res* 43:W39–W49. <https://doi.org/10.1093/nar/gkv416>
- Basavaraj G, Rao PP, Bhagavatula S, Ahmed W (2010) Availability and utilization of pearl millet in India. *J SAT Agrirc Res* 8:1–6
- Cai X, Zhang Y, Zhang C, Zhang T, Hu T, Ye J, Zhang J, Wang T, Li H, Ye Z (2013) Genome-wide analysis of plant-specific Dof transcription factor family in tomato. *J Integr Plant Biol* 55:552–566. <https://doi.org/10.1111/jipb.12043>
- Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL (2009) BLAST+: architecture and applications. *BMC Bioinformatics* 10:421. <https://doi.org/10.1186/1471-2105-10-421>
- Cao B, Cui Y, Lou K, Luo D, Liu Z, Zhou Q (2020) Genome-wide identification and expression analysis of the *Dof* gene family in *Medicago sativa* L. under various abiotic stresses. *DNA Cell Biol* 39:1976–1989. <https://doi.org/10.1089/dna.2020.5652>
- Chao J, Li Z, Sun Y, Aluko OO, Wu X, Wang Q, Liu G (2021) MG2C: a user-friendly online tool for drawing genetic maps. *Mol Hortic* 1:16. <https://doi.org/10.1186/s43897-021-00020-x>
- Chen R, Ni Z, Qin Y, Nie X, Lin Z, Dong G, Sun Q (2005) Isolation and characterization of TaDof1 transcription factor in wheat (*Triticum aestivum* L). *DNA Seq* 16:358–363. <https://doi.org/10.1080/10425170500272940>
- Corrales AR, Nebauer SG, Carrillo L, Fernández-Nohales P, Marqués J, Renau-Morata B, Granell A, Pollmann S, Vicente-Carbajosa J, Molina RV, Medina J (2014) Characterization of tomato Cycling Dof factors reveals conserved and new functions in the control of flowering time and abiotic stress responses. *J Exp Bot* 65:995–1012. <https://doi.org/10.1093/jxb/ert451>
- Corrales AR, Carrillo L, Lasierra P, Nebauer SG, Dominguez-Figueroa J, Renau-Morata B, Pollmann S, Granell A, Molina RV, Vicente-Carbajosa J, Medina J (2017) Multifaceted role of cycling DOF factor 3 (CDF3) in the regulation of flowering time and abiotic stress responses in Arabidopsis. *Plant Cell Environ* 40:748–764. <https://doi.org/10.1111/pce.12894>
- Crooks GE, Hon G, Chandonia JM, Brenner SE (2004) WebLogo: a sequence logo generator. *Genome Res* 14:1188–1190. <https://doi.org/10.1101/gr.849004>
- Dudhate A, Shinde H, Tsugama D, Liu S, Takano T (2018) Transcriptomic analysis reveals the differentially expressed genes and pathways involved in drought tolerance in pearl millet [*Pennisetum glaucum* (L.) R. Br]. *PLoS ONE* 13:e0195908. <https://doi.org/10.1371/journal.pone.0195908>
- Fornara F, Panigrahi KC, Gissot L, Sauerbrunn N, Rühl M, Jarillo JA, Coupland G (2009) Arabidopsis DOF transcription factors act redundantly to reduce *CONSTANS* expression and are essential for a photoperiodic flowering response. *Dev Cell* 17:75–86. <https://doi.org/10.1016/j.devcel.2009.06.015>
- Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A (2005) Protein identification and analysis tools on the ExPASy server. In: Walker JM (ed) *The Proteomics Protocols Handbook*. Humana Press, Totowa, pp 571–607
- Horton P, Park KJ, Obayashi T, Fujita N, Harada H, Adams-Collier CJ, Nakai K (2007) WoLF PSORT: protein localization predictor. *Nucleic Acids Res* 35:W585–W587. <https://doi.org/10.1093/nar/gkm259>
- Hu B, Jin J, Guo AY, Zhang H, Luo J, Gao G (2015) GSDS 2.0: an upgraded gene feature visualization server. *Bioinformatics* 31:1296–1297. <https://doi.org/10.1093/bioinformatics/btu817>
- Imaizumi T, Schultz TF, Harmon FG, Ho L, Kay S (2005) FKF1 F-box protein mediates cyclic degradation of a repressor of *CONSTANS* in Arabidopsis. *Science* 309:293–297
- Ito TM, Trevizan CB, dos Santos TB, de Souza SGH (2017) Genome-wide identification and characterization of the *Dof* transcription factor gene family in *Phaseolus vulgaris* L. *Am J Plant Sci* 8:3233–3257. <https://doi.org/10.4236/ajps.2017.812218>
- Kang WH, Kim S, Lee HA, Choi D, Yeom SI (2016) Genome-wide analysis of Dof transcription factors reveals functional characteristics during development and response to biotic stresses in pepper. *Sci Rep* 6:33332. <https://doi.org/10.1038/srep33332>
- Khaksar G, Sangchay W, Pinsorn P, Sangpong L, Sirikantaramas S (2019) Genome-wide analysis of the Dof gene family in durian reveals fruit ripening-associated and cultivar-dependent Dof transcription factors. *Sci Rep* 9:12109. <https://doi.org/10.1038/s41598-019-48601-7>
- Khan H, Ullah I, Zeb U, Ullah S, Zhao P (2021) Identification, characterization and expression profiles of Dof transcription factors in common walnut (*Juglans regia* L). *Pol J Environ Stud* 30:2149–2160. <https://doi.org/10.15244/pjoes/125518>
- Kumar M, Michael S, Alvarado-Valverde J, Mészáros B, Sámáno-Sánchez H, Zeke A, Dobson L, Lazar T, Örd M, Nagpal A, Farahi

- N, Käser M, Kraleti R, Davey NE, Pancsa R, Chemes LB, Gibson TJ (2022) The eukaryotic Linear Motif resource: 2022 release. *Nucleic Acids Res* 50:D497–D508. <https://doi.org/10.1093/nar/gkab975>
- Kushwaha H, Gupta S, Singh VK, Rastogi S, Yadav D (2011) Genome wide identification of *Dof* transcription factor gene family in sorghum and its comparative phylogenetic analysis with rice and Arabidopsis. *Mol Biol Rep* 38:5037–5053. <https://doi.org/10.1007/s11033-010-0650-9>
- Lescot M, Déhais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, Rouzé P, Rombauts S (2002) PlantCARE, a database of plant *cis*-acting regulatory elements and a portal to tools for *in silico* analysis of promoter sequences. *Nucleic Acids Res* 30:325–327. <https://doi.org/10.1093/nar/30.1.325>
- Letunic I, Bork P (2019) Interactive tree of life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res* 47:W256–W259. <https://doi.org/10.1093/nar/gkz239>
- Li D, Yang C, Li X, Gan Q, Zhao X, Zhu L (2009) Functional characterization of rice *OsDof12*. *Planta* 229:1159–1169. <https://doi.org/10.1007/s00425-009-0893-7>
- Li H, Dou L, Li W, Wang P, Zhao Q, Xi R, Pei X, Liu Y, Ren Z (2018) Genome-wide identification and expression analysis of the *Dof* transcription factor gene family in *Gossypium hirsutum* L. *Agronomy* 8:186. <https://doi.org/10.3390/agronomy8090186>
- Lijavetzky D, Carbonero P, Vicente-Carbajosa J (2003) Genome-wide comparative phylogenetic analysis of the rice and Arabidopsis *Dof* gene families. *BMC Evol Biol* 3:17. <https://doi.org/10.1186/1471-2148-3-17>
- Lohani N, Babaei S, Singh MB, Bhalla PL (2021) Genome-wide *in silico* identification and comparative analysis of *Dof* gene family in *Brassica napus*. *Plants* 10:709. <https://doi.org/10.3390/plants10040709>
- Ma J, Li MY, Wang F, Tang J, Xiong AS (2015) Genome-wide analysis of *Dof* family transcription factors and their responses to abiotic stresses in chinese cabbage. *BMC Genomics* 16:33. <https://doi.org/10.1186/s12864-015-1242-9>
- Malviya N, Gupta S, Singh VK, Yadav MK, Bisht NC, Sarangi BK, Yadav D (2015) Genome wide *in silico* characterization of *Dof* gene families of pigeonpea (*Cajanus cajan* (L) Millsp). *Mol Biol Rep* 42:535–552. <https://doi.org/10.1007/s11033-014-3797-y>
- Mulat MW, Sinha VB (2020) Identification and characterization of *Dof* in *tef* [*Eragrostis tef* (Zucc.) Trotter] using *in silico* approaches. *Gene Rep* 19:100590. <https://doi.org/10.1016/j.genrep.2020.100590>
- Sawa M, Nusinow DA, Kay SA, Imaizumi T (2007) FKF1 and GIGANTEA complex formation is required for day-length measurement in Arabidopsis. *Science* 318:261–265. <https://doi.org/10.1126/science.1146994>
- Sugiyama T, Ishida T, Tabei N, Shigyo M, Konishi M, Yoneyama T, Yanagisawa S (2012) Involvement of PpDof1 transcriptional repressor in the nutrient condition-dependent growth control of protonemal filaments in *Physcomitrella patens*. *J Exp Bot* 63:3185–3197. <https://doi.org/10.1093/jxb/ers042>
- Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22:4673–4680. <https://doi.org/10.1093/nar/22.22.4673>
- Tsugama D, Takano T (2021) TGIF-DB: terse genomics interface for developing botany. *BMC Res Notes* 14:181
- Varshney RK, Shi C, Thudi M, Mariac C, Wallace J, Qi P, Zhang H, Zhao Y, Wang X, Rathore A, Srivastava RK, Chitkineni A, Fan G, Bajaj P, Punhuri S, Gupta SK, Wang H, Jiang Y, Couderc M, Katta M et al (2017) Pearl millet genome sequence provides a resource to improve agronomic traits in arid environments. *Nat Biotechnol* 35:969–976. <https://doi.org/10.1038/nbt.3943>
- Venkatesh J, Park SW (2015) Genome-wide analysis and expression profiling of DNA-binding with one zinc finger (*dof*) transcription factor family in potato. *Plant Physiol Biochem* 94:73–85. <https://doi.org/10.1016/j.plaphy.2015.05.010>
- Wang Y, Tang H, Debarry JD, Tan X, Li J, Wang X, Lee TH, Jin H, Marler B, Guo H, Kissinger JC, Paterson AH (2012) MCS-canX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res* 40:e49. <https://doi.org/10.1093/nar/gkr1293>
- Wang H, Zhao S, Gao Y, Yang J (2017) Characterization of *Dof* transcription factors and their responses to osmotic stress in poplar (*Populus trichocarpa*). *PLoS ONE* 12:e0170210. <https://doi.org/10.1371/journal.pone.0170210>
- Wang P, Yan Z, Zong X, Yan Q, Zhang J (2021) Genome-wide analysis and expression profiles of the *Dof* family in *Cleistogenes songorica* under temperature, salt and ABA treatment. *Plants* 10:850. <https://doi.org/10.3390/plants10050850>
- Wei PC, Tan F, Gao XQ, Zhang XQ, Wang GQ, Xu H, Li LJ, Chen J, Wang XC (2010) Overexpression of AtDOF4.7, an Arabidopsis DOF family transcription factor, induces floral organ abscission deficiency in Arabidopsis. *Plant Physiol* 153:1031–1045. <https://doi.org/10.1104/pp.110.153247>
- Wei Q, Wang W, Hu T, Hu H, Mao W, Zhu Q, Bao C (2018) Genome-wide identification and characterization of *Dof* transcription factors in eggplant (*Solanum melongena* L). *PeerJ* 6:e4481. <https://doi.org/10.7717/peerj.4481>
- Wen CL, Cheng Q, Zhao L, Mao A, Yang J, Yu S, Weng Y, Xu Y (2016) Identification and characterisation of *Dof* transcription factors in the cucumber genome. *Sci Rep* 6:23072. <https://doi.org/10.1038/srep23072>
- Xu J, Dai H (2016) *Brassica napus* *Cycling Dof Factor1* (*BnCDF1*) is involved in flowering time and freezing tolerance. *Plant Growth Regul* 80:315–322. <https://doi.org/10.1007/s10725-016-0168-9>
- Yamamoto MP, Onodera Y, Touno SM, Takaiwa F (2006) Synergism between RPBFDof and RISBZ1 bZIP activators in the regulation of rice seed expression genes. *Plant Physiol* 141:1694–1707. <https://doi.org/10.1104/pp.106.082826>
- Yanagisawa S (2002) The *Dof* family of plant transcription factors. *Trends Plant Sci* 7:555–560. [https://doi.org/10.1016/s1360-1385\(02\)02362-2](https://doi.org/10.1016/s1360-1385(02)02362-2)
- Yanagisawa S (2004) *Dof* domain proteins: plant-specific transcription factors associated with diverse phenomena unique to plants. *Plant Cell Physiol* 45:386–391. <https://doi.org/10.1093/pcp/pch055>
- Yanagisawa S, Izui K (1993) Molecular cloning of two DNA-binding proteins of maize that are structurally different but interact with the same sequence motif. *J Biol Chem* 268:16028–16036
- Zhang L, Liu B, Zheng G, Zhang A, Li R (2017) Genome-wide characterization of the *SiDof* gene family in foxtail millet (*Setaria italica*). *Bio Syst* 151:27–33. <https://doi.org/10.1016/j.biosystems.2016.11.007>
- Zhang J, Lin LM, Cheng WW, Song X, Long YH, Xing ZB (2021) Genome-wide identification and expression analysis of glycosyltransferase gene family 1 in *Quercus robur* L. *J Appl Genet* 62:559–570. <https://doi.org/10.1007/s13353-021-00650-3>
- Zhou Y, Cheng Y, Wan C, Li J, Yang Y, Chen J (2020) Genome-wide characterization and expression analysis of the *Dof* gene family related to abiotic stress in watermelon. *PeerJ* 8:e8358. <https://doi.org/10.7717/peerj.8358>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.