INNSPUB

International Journal of Biosciences | IJB |
ISSN: 2220-6655 (Print), 2222-5234 (Online)
http://www.innspub.net
Vol. 17, No. 6, p. 103-123, 2020

RESEARCH PAPER OPEN ACCESS

Genome-Wide Bioinformatics Analysis of *DOF* Transcription Factor Gene Family of Asparagus and Its Comparative Phylogenetic Assessment with Arabidopsis

Humera Amin, Shumaila Dastgir, Muhammad Shafiq*, Muhammad Arshad Javed, Muhammad Saleem Haider

Institute of Agriculture Sciences, University of the Punjab, Lahore, Pakistan

Key words: DNA binding with One Finger, Asparagus, Plant specific transcription factor, Genome-wide.

http://dx.doi.org/10.12692/ijb/17.6.103-123 Article published on December 12, 2020

Abstract

Asparagus officinalis L has been cultivated and harvested from the wild for thousands of years for medicinal purposes. The family of *DOF gene* (DNA binding with One Finger) is highly conserved and member of this gene family plays an important role in many regulatory mechanisms in plants including plant growth and development. We identified 7 putative *DOF* genes through the genome mining of asparagus (*A. officinalis* L.) and distributed unevenly among 5 chromosomes. The *DOF* gene family in asparagus was phylogenetically analyzed with *Arabidopsis thaliana DOF* genes and classified into 5 subfamilies. The exon-intron organization of *DOF* gene in asparagus showed the presence of intron and *AoDOF5* (2intron) contain a maximum no of introns. The *DOF* gene on chromosome in *A.officinalis is observed in 4 chromosomes*. Maximum 3 *DOF* genes were observed on chromosome 8 and a single *DOF* gene was found on other chromosomes. The segmental gene duplication was predominant over tandem duplication which might be the cause of *DOF* gene family expansion in *A.officinalis*. The *cis*-regulatory element analysis revealed the presence of light-responsive, abscisic acid responsiveness, endosperm-specific, anaerobic induction, root-specific expression, gibberellin-responsive, meristem-specific and stress-responsive elements. Comprehensive phylogenetic analysis of *DOF* genes of *A.officinalis* with Arabidopsis revealed several orthologs and paralogs assisting in understanding the putative functions of Ao*DOF* genes. The bioinformatics-based genome-wide assessment of *DOF* gene family of *A.officinalis* attempted in the present study could be a significant step for deciphering novel *DOF* genes based on genome-wide expression profiling.

* **Corresponding Author:** Muhammad Shafiq ✉ shafiq.iags@pu.edu.pk

Introduction

The *DOF* (DNA binding with One Finger) is a plant-specific transcription factor (PSTFs) family and plays a vital role in the regulation of plant growth and development (Liu *et al*., 2017). PSTFs *DOF* has been recently reviewed in eggplant, (Wei *et al*., 2018) chickpea, (Nasim *et al*., 2016) tomato, (Cai *et al*., 2013; Corrales *et al*., 2014) and wheat (Chen *et al*., 2005; Dong *et al*., 2007) using structural, functional and bioinformatics tool (Gupta *et al*., 2015). Proteins in the *DOF* PSTFs gene family possess a highly conserved 50–52 amino acids (AA) domain referred to as *DOF* domain. At the amino-terminal of the *DOF* domain, four cysteine residue (C2C2-type) with zinc finger (ZF) motif is present for sequence-specific DNA binding (DB) (Yanagisawa, 1995;Yanagisawa and Schmidt, 1999;Negi *et al*., 2013). Conserved DB domain at N-terminal and a transcriptional regulation (TR) domain at C-terminal are the two main domain which showed the varied functions of *DOF* proteins (Yanagisawa, 1995; Yanagisawa and Schmidt, 1999, Negi *et al*., 2013).

PSTFs *DOF* show multiple functions that are specific to plants like nitrogen assimilation (Yanagisawa *et al*., 2004), controlling glucosinolate biosynthesis in Arabidopsis (Skirycz *et al*., 2006) development of functional stomata, (Negi *et al*., 2013) endosperm-specific and seed development in barley (Diaz *et al*., 2005) carbon metabolism and C4PEPC (C4 photosynthetic phosphoenolpyruvate carboxylase) gene expression in maiz, (Yanagisawa, 2000), development of spikes in *Eleusine coracana L* (Qu *et al*., 2016), plant defense against pest (Arnaiz *et al*., 2019) regulation of nitrogen metabolism in millet (Gupta *et al*., 2014), flowering and photoperiod response (Imaizumi, 2005;Fornara, 2009; Fornara *et al*., 2009), shoot and branch regulation and development (Papi *et al*., 2002; Zou *et al*., 2013; Bueso *et al*., 2016), regulation of stomata physiology (Negi *et al*., 2013), abiotic stress response (Corrales *et al*., 2014), protein movement inside the cell (Chen *et al*., 2013) and circadian cycle regulation (Yang *et al*., 2011;Corrales *et al*., 2014). *DOF* gene family members are diverse in different plant species which indicate their specific and diverse function in different crops .An effort has been made to predict and analyze the function of different *DOF* gene members in different plant species using modern molecular biology and bioinformatics tool. (Lijavetzky *et al*., 2003; Moreno-Risueno *et al*., 2007; Hernando-Amado *et al*., 2012;Cai *et al*., 2013; Ma *et al*., 2015a; Ma *et al*., 2015b; Malviya *et al*., 2015). Rice, Arabidopsis and popular has 30, 36 and 41 *DOF* gene family member that is predicted by genome-wide comparative phylogenetic analysis (Lijavetzky *et al*., 2003; Yang *et al*., 2006).

Asparagus is a perennial and dioecious plant (Harkess *et al*., 2017) belong to the family Asparagaceae. And grow up to 100–150 cm (40–60 in) height, with feathery foliage, branched and stout stems,. Asparagus has both male and female flowers on separate plants, but some type is hermaphrodite (Zhang *et al*., 2019) and is used as a vegetable owing to its definite flavor. Asparagus has medicinal properties and has a purported function as an aphrodisiac (Hartung *et al*., 1990). This plant has been depicted in ancient Egyptian literature dating to 3000 BC. Stem thickness indicates the age of the plant (Zohary and Spiegel-Roy, 1975).

This study is the first report on the identification and characterization of *DOF* gene family of asparagus using genome-wide and phylogenetic comparative analysis with Arabidopsis.

Material and methods

Database search and retrieval of sequence

A probe of 58 amino acids sequence related to the *DOF* domain of *A. thaliana (*Arabidopsis*)* (Accession no 175581) was separated using Pfam version 27.0 (PF02701) (http://pfam.xfam.org/) (Finn *et al*., 2014) and it was used to search *DOF* domains in the whole genome and proteome sequence of *A. officinalis L* . by using Blast (BLASTP) at Phytozome https://phytozome.jgi.doe.gov/pz/portal.html(Goodstein *et al*., 2014). *DOF* domain from thr retrieved amino acid sequences was subjected to searches at the protoparam (http://web.expasy.org/protoparam/)

(Schultz *et al.*, 1998; Schultz *et al.*, 2000; Letunic *et al.*, 2004;Letunic *et al.*, 2006;), and NCBI CCD (Conserved Domain Database) (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) (Marchler-Bauer *et al.*, 2007) to identify *DOF* domain. The proteins sequence were excluded that lack conserved *DOF* domain in their sequence (*PF02701*) (https://pfam.xfam.org/family/PF02701). By using Gene Structure Display Server (GSDS), gene structure (Extron intron) was visualized by (http://gsds.cbi.pku.edu.cn/)http://gsds.cbi.pku.edu.cn/) (Hu *et al.*, 2015) adding both coding sequences and genomic sequences of *AoDOF*. The protein length (amino acid residues), molecular weight(Mw), and theoretical pI (Isoelectric point) of *AoDOF* proteins were predicted using ProtParam tool (http://web.expasy.org/protparam/) (Gasteiger *et al.*, 2005).

The information for gene IDs, chromosomal position, the sequence of gene and protein, were retrieved from Phytozomehttps://phytozome.jgi.doe.gov/pz/portal.html (Goodstein *et al.*, 2014). These *AoDOF* genes were renamed according to the order of their physical position.

Comparative Phylogenetic analysis of AoDOF protein and Multiple sequence alignment

The amino acid sequences of *DOF* proteins were aligned using Clustal W version 2.1 (Thompson *et al.*, 2003; Thompson *et al.*, 1994) and the phylogenetic tree was constructed using MEGA v x.0 (http://www.megasoftware.net) program (Kumar *et al.*, 1994; Mello, 2018; S. *et al.*, 2018) with the maximum likelihood (Miyashima *et al.*, 2019) and bootstrapping at 1000 replications. There were 7 *AoDOF* and 35 *atDOF* protein (Yang and Tuskan, 2006) sequences were used for phylogenetic analysis. These Arabidopsis *DOF* sequences belong to 10 different subgroups (Lijavetzky *et al.*, 2003; Nasim *et al.*, 2016).

Subcellular localization analysis, promoter prediction and conserved motifs recognition

Subcellular localization of *AoDOF* was predicted by the online tool WoLF PSORT (https://wolfpsort.hgc.jp/) (Horton *et al.*, 2006). The nuclear localization signals (NLS) in olive DOF proteins were also predicted through an online server NLSdb (https://rostlab.org/services/nlsdb/) (Cokol *et al.*, 2000). For the analysis of the promoter region, a sequence of 1500-bp upstream was retrieved from the initiation codon of the putative AoDOF genes. Plant Care database (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) (Rombauts *et al.*, 1999) was then used to predict cis-regulatory elements in these sequences and validated in the PLACE databases (http://www.dna.affrc.go.jp/PLACE/) (Higo *et al.*, 1998; Higo *et al.*, 1999).

Meme analysis

MEME (Multiple EM for Motif Elicitation) programs (http://meme.nbcr.net/meme/) (Bailey *et al.*, 2015) were used to predict and analyze with the concluded protein sequences of the *AoDOF*s with a maximum number of motif set as 20. The minimum width of 06 and a maximum width of 50 amino acids were s*et al*ong with other factors as default values.

Distribution of chromosomes and gene duplication

The genome sequences files and annotation files of *A. officinalis* were downloaded from phytozome website (https://phytozome.jgi.doe.gov/pz/portal.html).

All *AoDOF* genes were mapped to *A. officinalis* chromosomes based on physical location information from the database of *A. officinalis* genome using Gene location visualize tool in tbtool software (https://github.com/CJ-Chen/TBtools). (Chen *et al.*, 2020) Non-synonymous (ka) and synonymous (ks) substitution of each duplicated *DOF* genes were calculated using Ka/Ks_Calculator in tbtools. The time of divergence of *A. officinalis DOF* gene family was estimated using Ks and Ka values. Protein sequences were aligned using Clustal W and the number of Ka and Ks substitution rates was conducted using the Nei-Gojobori model. The rate variation among sites was modeled with a gamma distribution (shape parameter = 1). The Ka/Ks ratios

were calculated. The Ka/Ks ratio was estimated to predict the rates of molecular evolution of each gene pair. Generally, Ka/Ks < 1 suggests purifying selection; Ka/Ks = 1 indicates neutral selection, and Ka/Ks > 1 predicts positive selection. The time of divergence was roughly estimated by computing Ks value in T = K s /2λ equation where λ represents the value of 6.05×10^{-9}.

Putative microRNA target sites analysis

The miRNA dataset of *A. officinalis* was obtained by Gene Expression Omnibus (GEO) NCBI (Konishi, 2007;Gardiner, 2010; Iwamoto, 2016; Miyashima, 2019) in an experiment related to the identification of miRNAs and its targets in male and female *A. officinalis*. The miRNA data were also downloaded fromhttp://www.mirbase.org/cgibin/mirna_summary.pl?org=aof. There, to find out the miRNAs which target the *AoDOF* genes, CDS sequences of all *AoDOF* genes were searched for the complementary sequences of miRNAs with the help of psRNATarget(https://plantgrn.noble.org/psRNATarget/analysis?function=3) (Samad, 2017) with default parameters.

Results

Identification of the DOF genes in A. officinalis

The sequence of the *DOF* domain was a blast to recognize the whole genome sequences of *A. officinalis* and non-redundant *DOF* genes were collected. Out of 8 identified 5 *DOF* protein sequences were non-redundant. The translated *DOF* proteins in *A. officinalis* possessed amino acid 85–461, molecular weight from 9.0-49.8 kD and pI ranges from 6.24-11.37 which is revealed by ExPASy server ProtParam (Table 1). Within the highly conserved sequences of *A. officinalis DOF* domain, 25 out of 54 amino acids were found to be 100 % conserved in all the *DOF* domain sequences. These sequences also included the highly conserved four cysteine residues that coordinate with zinc ion and are known to be a typical feature of *DOF* proteins (Fig. 1). Other conserved residues observed were Cys^{1}, Pro^{2}, Arg^{3}, Cys^{4}, Ser^{6}, Asn^{8},Thr^{9}, Lys^{10}, Phe^{11}, Cys^{12}, Tyr^{13}, Tyr^{14} , Asn^{15} , Asn^{16}, Tyr^{17} , Gln^{21} , Pro^{22}, Trp^{33}, Arg^{40}, Asn^{41}, Val^{42}, Pro^{43}, Val^{44}, Gly^{45}, Gly^{47} (Fig. 1) while the other 27 or 29 amino acids were discovered to be inconsistent in all the *AoDOF* proteins. While all others no one showed nuclear localization signal (NLS) as predicted using NLSdb (https://rostlab.org/services/nlsdb/) (Cokol *et al.*, 2000).

The subcellular localization of *AoDOF* protein showed nucleus, cytoplasm, chloroplast, mitochondrion, vacuole, cytoskeleton and extracellular spaces localization as predicted using the online tool WoLF PSORT (https://wolfpsort.hgc.jp/) (Horton *et al.*, 2006) (Table 3).

Table 1. Information about 7 *DOF* genes discovered from the genome of *A. officinalis* amino acid sequence length; MW, molecular weight; Pi value.

DOF gene	Source accession no	Chromosome no	Chromosome location	No of introns	mRNA length (bp)	Amino acid ength	sequence	pI value	Mw(kd)
Ao*DOF*1	AsparagusV1_08.412	08	7290278..7291576	1	702	261	1299	cannot be computed	undefined
Ao*DOF*2	AsparagusV1_08.3278	08	124421297..124433633	0	1386	461	3278	6.24	49892.69
Ao*DOF*3	AsparagusV1_05.2213	05	94112838..94115343	0	1221	309	2506	8.89	32903.18
Ao*DOF*4	AsparagusV1_03.1160	03	25249244..25250099	1	780	241	856	8.77	26016.75
Ao*DOF*5	AsparagusV1_08.3510	08	129286205..129305819	2	570	189	720	11.37	20361.89
Ao*DOF*6	AsparagusV1_07.3114	07	130311228..130333076	1	258	85	258	11.19	9027.19
Ao*DOF*8	AsparagusV1_04.3405	04	140213750..140216572	1	657	218	657	11.24	24046.44

Comparative phylogenetic relatedness of AoDOF gene family with Arabidopsis

To investigate the evolutionary relationships between *AoDOF* TFs and *Arabidopsis thaliana* a neighbor-joining (NJ) phylogenetic tree was constructed by aligning their full-length protein sequences. The results depicted that 7 *AoDOF* proteins were distributed among 4 subgroups named as D1, C2.1, B2 and A (Table 2 and Fig. 1). Subgroup D1 consisted of a total of 1 *DOF*-like Protein in which 7 are

arabidopsis *DOF*-like proteins, AT2G34140, AT1G29160, AT1G69570, AT3G47500, AT5G39660, AT1G26790, AT5G62430, while the remaining belongs to asparagus Ao*DOF*2. C3 group consisted of 4 *DOF*-like proteins that are all of Arabidopsis AT4G21030, AT4G21080, AT4G21050, AT4G21040. None of the *DOF*-like proteins in this clade belongs to *A officinalis*. B2 Group contained 1 *DOF*-like protein in which only 3 are of Arabidopsis, AT4G38000, AT5G65590, AT1G28310, while 1 belongs to *A officinalis* Ao*DOF*5. C1 contained 4 are of Arabidopsis AT5G62940, AT2G28510, AT3G45610 and AT5G60200 and no one is of *A officinalis*. Group C2.1 had 5 of of Arabidopsis AT4G00940, AT3G61850, AT1G64620, AT2G46590 and AT4G24060, and only one belongs to *AoDOF 1*. B1 consisted of 5 Arabidopsis, AT1G07640, AT2G28810, AT5G02460, AT3G55370, AT2G37590 and no one from *A officinalis*. Group A consisted of 3 from Arabidopsis AT5G60850, AT3G21270, AT1G51700 and no one from *A officinalis*. D2 consisted of 5 *DOF*-like Proteins, 2 of Arabidopsis, AT3G50410, AT5G66940, and 3 from *A officinalis* Ao*DOF 3, AoDOF 6* and *AoDOF 8*. Proteins of common clades usually seem to show similarity in structure and functioning (Fig. 5).

So, all the *DOF*-like proteins of similar Clades may have a similar structure as well as functions. Among the *DOF* domain sequence of olive and Arabidopsis, 3 amino acids were found to exist on the same location.

Table 2. *A. officinal is DOF* gene family distribution among groups based on phylogenetic analysis with Arabidopsis *DOF* member.

Group	Number of *DOF* gene	Gene id
D2	3	*AoDOF3, AoDOF6, AoDOF8*
D1	1	*AoDOF2*
E	1	*AoDOF4*
B2	1	*AoDOF5*
C2.1	1	*AoDOF1*

Table 3. Prediction of the subcellular localization of *DOF* proteins.

DOF gene	Source accession no	Nucleus	Cytoplasm	Mitochondria	Extra	Cytsk	Vacoule	Chloroplast
AoDOF1	AsparagusV1_08.412	8	1	3	1	1	0	0
AoDOF2	AsparagusV1_08.3278	13	0	0	0	0	1	0
AoDOF3	AsparagusV1_05.2213	14	0	0	0	0	0	0
AoDOF4	AsparagusV1_03.1160	2	0	1	0	0	0	11
AoDOF5	AsparagusV1_08.3510	8	0	2	0	0	0	4
AoDOF6	AsparagusV1_07.3114	5	0	2	2	0	0	5
AoDOF8	AsparagusV1_04.3405	11	0	1	0	0	0	2

Gene structures and recognition of conserved motifs and domain

The organization of exon and intron provide are the backbones of genes and help in assisting verification for the study of the evolutionary relationship between genes or organisms (Koralewski and Krutovsky, 2011). Their numbers and distribution patterns are an evolutionary mark for a gene family. A comprehensive demonstration of the exon-intron structures of asparagus *DOF* genes along with phylogenetic revealed that the gene structure pattern was consistent with the phylogenetic analysis. The number of introns varied from zero to two in *A officinalis* (Fig.4, Table 6). There are two Ao*DOF* genes without intron (28.5%), four *AoDOF* genes with one intron (57.1), one *AoDOF* genes with two introns (14.2%) (Table 6, and Fig. 4.

All of the *AoDOF* genes in subfamily D2 varied from zero to one intron, while the number of introns of the *AoDOF* gene in subfamily C2.1 contained one intron (Table 6). Similar to the *DOF* genes studied in various species, some *DOF* genes in olive possess no intron while other *DOF* genes possess multiple introns, up to two (Table 6 and Fiq. 4).

Table 4. Identif ication of *AoDOF* miRNAs in male and female *A. officinalis*.

miRNA_Acc.	Target_Acc.	miRNA_start	miRNA_end	Target_start	Target_end	miRNA_aligned_fragment
aof-miR168c	Ao*DOF*8	1	20	557	576	CCGCCUUGCACCAACUGAAU
aof-miR477i	Ao*DOF*8	1	21	421	441	ACUCUCCCUCAAGGGCUUCCG
aof-miR166i-5	Ao*DOF*5	1	21	80	100	UCGGACCCGGCUUCAUUCCCC
aof-miRn34	Ao*DOF*2	1	21	150	170	UGGUCGAUUGUUUUUGGGAUG
aof-miRn34	Ao*DOF*3	1	21	560	580	UGGUCGAUUGUUUUUGGGAUG
aof-miRn35	Ao*DOF*2	1	22	149	170	UGGUCGAUUGUUUUUGGGAUGC
aof-miRn35	Ao*DOF*3	1	22	559	580	UGGUCGAUUGUUUUUGGGAUGC

Table 5. miRNA targets prediction of *AoDOF*. The miRNA data was downloaded from plant micro RNA http://plantgrn.noble.org/psRNATarget/analysis.

Target_Acc.	*miRNA length*	*Target*	*miRNA_aligned_fragment*
AoDOF8	*22*	*423-444*	*AUCUCUCUCCCUCAAAGGCUCU*
AoDOF8	*21*	*421-441*	*ACUCUCCCUCAAGGGCUUCCG*
AoDOF8	*21*	*75-95*	*CGAGUUUUCACGUUUGGGCGA*
AoDOF2	*22*	*1279-1300*	*UUUGAUUAUUGGAUUGUUGCCU*

The identification and distributions of 20 motifs within all the *A. officinalis DOF* proteins were studied using the MEME program (Fig. 3a, Fig. 9). The presence of *DOF* domain was consistent among 7 out of 8 *AoDOF* proteins (Fig. 3b). It was witnessed that the *DOF* genes present in the same group contain motifs that are alike which propose that these conserved motifs take an essential part in activities that are specific in a group or subgroup. The distribution of similar motifs among various *DOF* genes suggests that such genes might come into existence as a result of gene expansion.

Table 6. In *A.officinalis* predicted no of intron and exon in *AoDOF* genes.

Group	*DOF* gene	Source accession	Intron	Exon
Name	Name	No.	No.	No.
D2	*AoDOF3*	AsparagusV1_05.2213	0	2
	AoDOF6	AsparagusV1_07.3114	1	1
	AoDOF8	AsparagusV1_04.3405	1	1
D1	*AoDOF2*	AsparagusV1_08.3278	0	2
B2	*AoDOF5*	AsparagusV1_08.3510	2	3
C2.1	*AoDOF1*	AsparagusV1_08.412	1	2
E	*AoDOF4*	AsparagusV1_03.1160	1	2

Table 7. The detailed information of the motifs in *A. officinalis DOF* proteins.

ID	Motif	Length (AA)
1	TKFCYYNNYSLSQPRHFC	18
2	TCPRYWT	7
3	NHNHDL	6
4	FMPMAPEPGPEYGSGFGLNEF	21
5	LRNVPVGAGSRKNK	14
6	LPCPRCNSTN	10
7	QVDHAVSIQT	10
8	MAADMGEPHDEI	12
9	LEPHKDQ	7
10	WDDGRRWRWR	10
11	VKDRKADVNERLGQDFEC	18
12	YWSGANWG	8
13	QHCFLN	6
14	MIGAKRGR	8
15	RJPQPQQL	8
16	FPFEDL	6
17	FWPCTN	6
18	AMELLR	6
19	NCPPLC	6
20	ZWNNGAGMAAVNCSF	15

Location of chromosomes and assessment of gene duplication of A.officinalis DOF genes

Distribution of chromosomes of the analyzed *A officinal is DOF* genes demonstrated that *AoDOF* genes were present on five out of ten chromosomes. The maximum number, as three, of *AoDOF* genes were located on chromosomes eight. Likewise, chromosomes three, four, five and seven contain one *DOF* genes. (Fig. 5). *AoDOF 1, AoDOF5* and *AoDOF 2* are present on chromosome eight (Fig. 5). The date of duplication of the gene was also roughly estimated through MEGA-X using pairwise alignment that provided Ks and Ka values and then Ka/Ks was calculated manually (Fig. 4). Ks depicts the number of synonymous substitutions per synonymous site whereas Ka shows the number of nonsynonymous substitutions per nonsynonymous site and the ratio of nonsynonymous (Ka) versus synonymous (Ks) mutation was represented by Ka/Ks.

Table 8. The Arabidopsis *DOF* genes GO annotations and its involvement in biological function.

Group	Ca Gene IZ	Ortholgue		GO Annotations : involved in Biological Process	Refere
		Accession	Gene IDS		
D2	Ca*DOF*1	AT5G66940		auxin and salt signals to regulate *Arabidopsis* thaliana seed germination.	
		AT5G62940.1		phloem or xylem histogenesis,	
		AT3G50410.1		cell wall modification, positive regulation of cell cycle, positive regulation of transcription, response to auxin, response to salicylic acid	
B2		AT5G65590.1		guard cell differentiation,	
		AT1G28310.2		regulation of transcription,	
		AT4G38000		floral organ abscission,	
B1		AT2G28810.1		regulation of transcription, transcription, DNA-templated	
		AT2G37590.1		regulation of transcription, transcription, DNA-templated	
		AT3G55370.3		photomorphogenesis, positive regulation of transcription, DNA-templated	
		AT5G02460		regulation of transcription,	
A		AT1G51700		regulation of transcription,	
		AT3G21270.1		regulation of transcription, transcription, DNA-templated	
		AT5G60850.1		regulation of transcription,	
C1		AT1G47655.1		regulation of transcription, transcription, DNA-templated	
		AT1G07640.3		regulation of glucosinolate biosynthetic process, transcription, response to insect, response to jasmonic acid, response to wounding,	
		AT2G28510.1		regulation of transcription, transcription, DNA-templated	
		AT5G60200		regulation of transcription, DNA-templated, root development, Transcription	
		AT3G45610		regulation of transcription,	
C2		AT1G21340.1		regulation of transcription,	
C2.2		AT3G52440.1		regulation of transcription, transcription, DNA-templated	
C2.1		AT3G61850.4		negative regulation of transcription, regulation of transcription, DNA-templated, response to red or far red light, seed germination	
		AT4G24060.1		regulation of transcription, transcription, DNA-templated	
		AT1G64620		regulation of transcription, transcription, DNA-templated	
		AT2G46590		cellular response to red light, cellular response to water stimulus, positive regulation of gibberellin biosynthetic process, positive regulation of seed germination, red light signaling pathway, regulation of transcription, response to cold, response to light stimulus, seed germination, transcription,	

		DNA-templated
	AT4G00940	regulation of transcription, transcription, DNA-templated
C3	AT4G21030	cotyledon development, mucilage metabolic process involved in seed coat development, positive regulation of transcription, regulation of secondary shoot formation, regulation of transcription, seed coat development, transcription, DNA-templated
	AT4G21040	regulation of transcription, transcription, DNA-templated
	AT4G21050	fruit development, positive regulation of transcription, regulation of secondary shoot formation, regulation of transcription, transcription, DNA-templated
	AT4G21080	regulation of transcription, transcription, DNA-templated
D1	AT1G69570	flower development, negative regulation of long-day photoperiodism, negative regulation of short-day photoperiodism, regulation of transcription, DNA-templated, Transcription
	AT2G34140	flower development, negative regulation of stem cell population maintenance, negative regulation of transcription, Transcription DNA-templated
	AT1G26790	regulation of transcription, transcription, DNA-templated
	AT5G62430.1	flower development, regulation of timing of transition from vegetative to reproductive phase, vegetative to reproductive phase transition of meristem
	AT3G47500.1	flower development, regulation of transcription, DNA-templated, transcription,
	AT5G39660.1	flower development, regulation of transcription, transcription, DNA-templated
	AT1G29160.1	regulation of transcription, seed coat development, transcription, DNA-templated

Table 9. Putative functions of Cis-regulatory elements identified in *A.officinalis DOF* promoter region.

Cis-element	Function
3-AF1 binding site	light responsive element
A-BOX	cis-acting regulatory element
ABRE	cis-acting element involved in the abscisic acid responsiveness
ABRE3a	cis-acting element involved in the abscisic acid responsiveness
ABRE4	cis-acting element involved in the abscisic acid responsiveness
AE-BOX	part of a module for light response
ARE	cis-acting regulatory element essential for the anaerobic induction
as-1	cis-acting regulatory element involved in the root-specific expression
AUXRR-CORE	cis-acting regulatory element involved in auxin responsiveness
Box 4	part of a conserved DNA module involved in light responsiveness
Box II	part of a light responsive element
CAAT -BOX	Common cis-acting element in promoter and enhancer regions
CAT-BOX	cis-acting regulatory element related to meristem expression
CCAAT-BOX	MYBHv1 binding site
CCGTCC-BOX	cis-acting regulatory element related to meristem specific activation
chs-CMA2A	part of a light responsive element
circadian	cis-acting regulatory element involved in circadian control
ERE	ethylene responsive element
GA-motif	part of a light responsive element
GARE-motif	gibberellin-responsive element
GATA-motif	part of a light responsive element
G-Box	Cis-acting regulatory element involved in light responsiveness

G-box	cis-acting regulatory element involved in light responsiveness
GC-motif	enhancer-like element involved in anoxic specific inducibility
GCN4-motif	cis-regulatory element involved in endosperm expression
GT 1-motif	light responsive element
HD-zip 1	element involved in differentiation of the palisade mesophyll cells
I-BOX	part of a light responsive element
LAMP-element	part of a light responsive element
LTR	cis-acting element involved in low-temperature responsiveness
MBS	MYB binding site involved in drought-inducibility
MRE	MYB binding site involved in light responsiveness
MSA-like	cis-acting element involved in cell cycle regulation
MYB	responses to biotic and abiotic stresses, development, differentiation, metabolism, defense
Myb	Myb Transcription factor binding site
MYB recognition site	
Myb-binding site	
MYB-like sequence	
MYC	
Myc	
O2-SITE	cis-acting regulatory element involved in zein metabolism regulation
p-BOX	gibberellin-responsive element
RY-element	cis-acting regulatory element involved in seed-specific regulation
Sp1	light responsive element
TATA-BOX	Core promoter element around -30 of transcription start
TATC-box	GA-responsive element
TCA-elements	cis-acting element involved in salicylic acid responsiveness
TCCC-motif	part of a light responsive element
TC-rich repeats	cis-acting element involved in defense and stress responsiveness
TCT-motif	part of a light responsive element
TGACG-motif	cis-acting regulatory element involved in the MeJA-responsiveness
TGA-element	auxin-responsive element
W box	WRKY DNA binding proteins
WUN-motif	wound-responsive element

The speculative date for tandem gene duplication of the paralogous group *AoDOF 08/AoDOF 01* was calculated to be 2.76 Mya while for the remaining 3 paralogous pairs the segmental duplication date was estimated in the range from 1.98 Mya for paralogous pair 2 to 2.48 Mya for paralogous pair 1. All 4 paralogous group pairs in *A. officinalis* had Ka/Ks ratio greater than 0.3 which suggests probability of considerable functional divergence after the occurrence of the duplication process.

Table 10. Duplicated *DOF*-like Genes and their dates of duplication in *A.officinali.s*.

Gene 1	Gene 2	Ks	Ka	Ka/Ks	Date (Mya)
Ao*DOF*3	Ao*DOF*1	0.745	0.742	0.996	2.48
Ao*DOF*6	Ao*DOF*4	0.596	0.434	0.729	1.98
Ao*DOF*8	Ao*DOF*1	0.828	0.489	0.590	2.76
Ao*DOF*6	Ao*DOF*5	0.72	0.468	0.650	2.4

Regulation of AoDOF by miRNAs

To investigate how miRNAs interact with *AoDOF* gene, we predicted the target sites for the *A. officinalis* miRNAs using all *DOF* full-length isoforms. In *A. officinalis*, 4 Ao*DOF* (Ao*DOF* 2, Ao*DOF*3, Ao*DOF* 5 and Ao*DOF* 8) from all 7 *AoDOF* genes were identified to be putative targets of 11 Aof-miRNAs, and 3 *DOF* mRNA transcribed from *AoDOF 1, AoDOF 4 and AoDOF 6* was not targeted by the miRNAs (Table 4 and Table 5). *AoDOF8* gene mRNA was predicted to be targeted by aof-miR168c, aof-miR477i, aof-miR477a, aof-miR477b and aof-miR12161 respectively (Table 4 and Table 5) while Ao*DOF* 5 mRNA was predicted to be targeted by aof-miR166i-5. Ao*DOF* 2 and Ao*DOF* 3 both were targeted by aof-miRn34 and aof-miRn35.

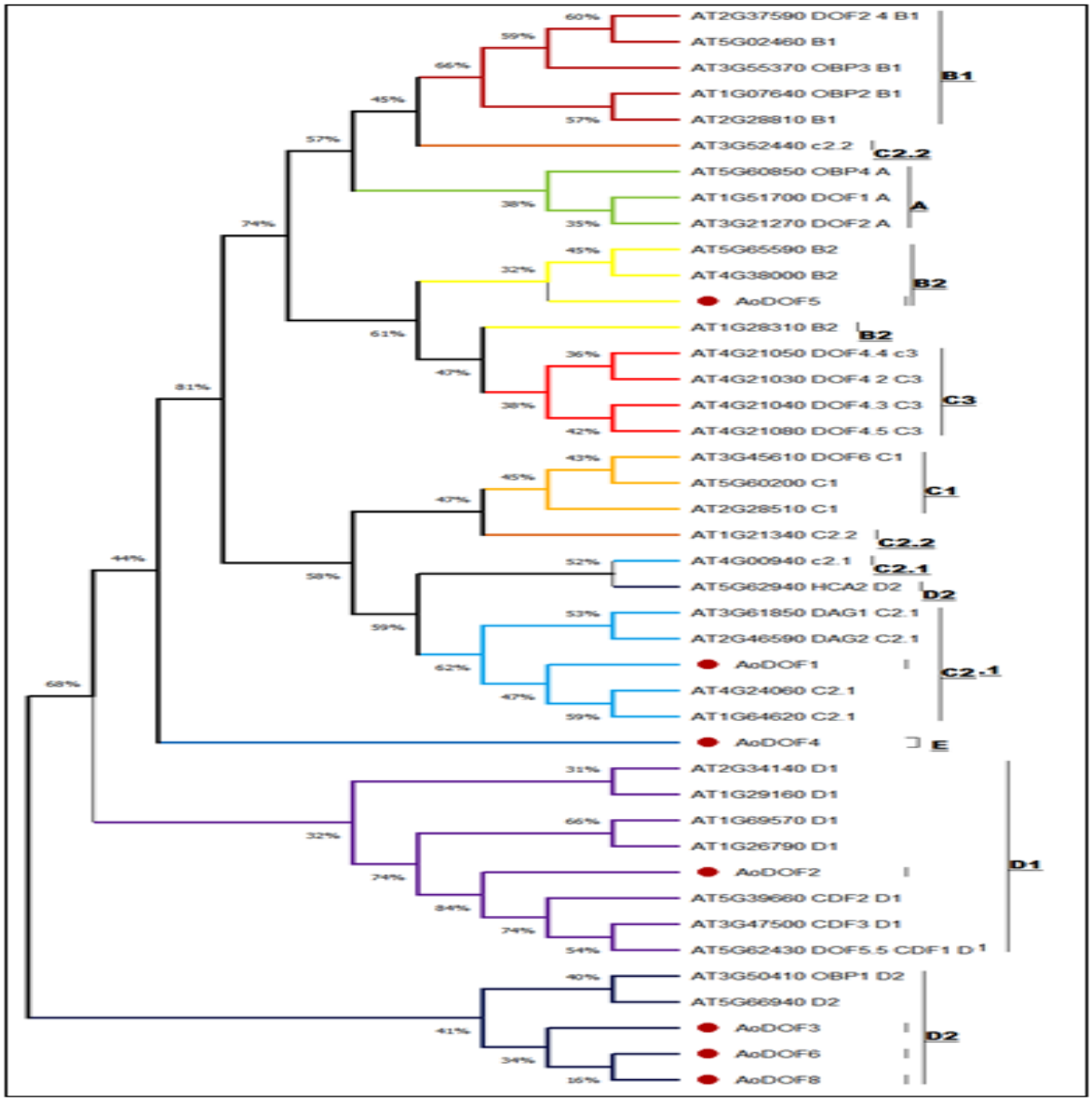


Fig. 1. Phylogenetic Relationship between *AoDOF* and *AtDOF* Proteins where *AoDOF* proteins are marked with red circle. The evolutionary history was inferred by using the Maximum Likelihood method and JTT matrix-based model. The tree with the highest log likelihood (-26991.60) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the JTT model, and then selecting the topology with superior log likelihood value. The proportion of sites where at least 1 unambiguous base is present in at least 1 sequence for each descendent clade is shown next to each internal node in the tree. This analysis involved 42 amino acid sequences. There were a total of 546 positions in the final dataset. Evolutionary analyses were conducted in MEGA X (P.H.A. and R.R., 1973); (J., 1985); (E. and L., 1965); (S. *et al.*, 2018).

The data about the expression of miRNA34, miRNA35, aof-miR477i and aof-miR166i-5 were also downloaded (NCBI GEO Accession GSE72594) and analysed. The data shows that there is no significant difference for expression of miRNA34, miRNA35, and aof-miR166i-5 while the expression of aof-miR477i is higher in male than female's *A. officinalis* which shows their involvement in the development of male *A. officinalis*.

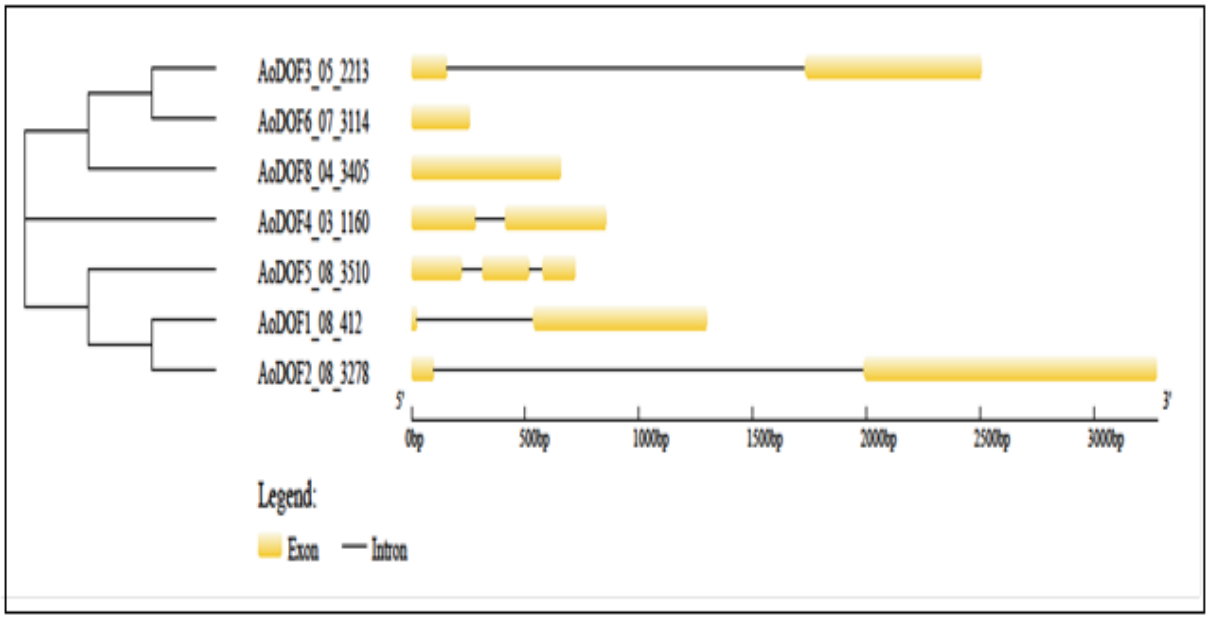


Fig. 2. Phylogenetic relationships and gene structures of the *DOF* genes from *A. officinalis*. The phylogenetic tree was constructed based on the full-length sequences of *A. officinalis DOF* proteins. b Exon-intron structures of the *A. officinalis DOF* genes. Yellow boxes indicate exons; and black lines indicate introns.

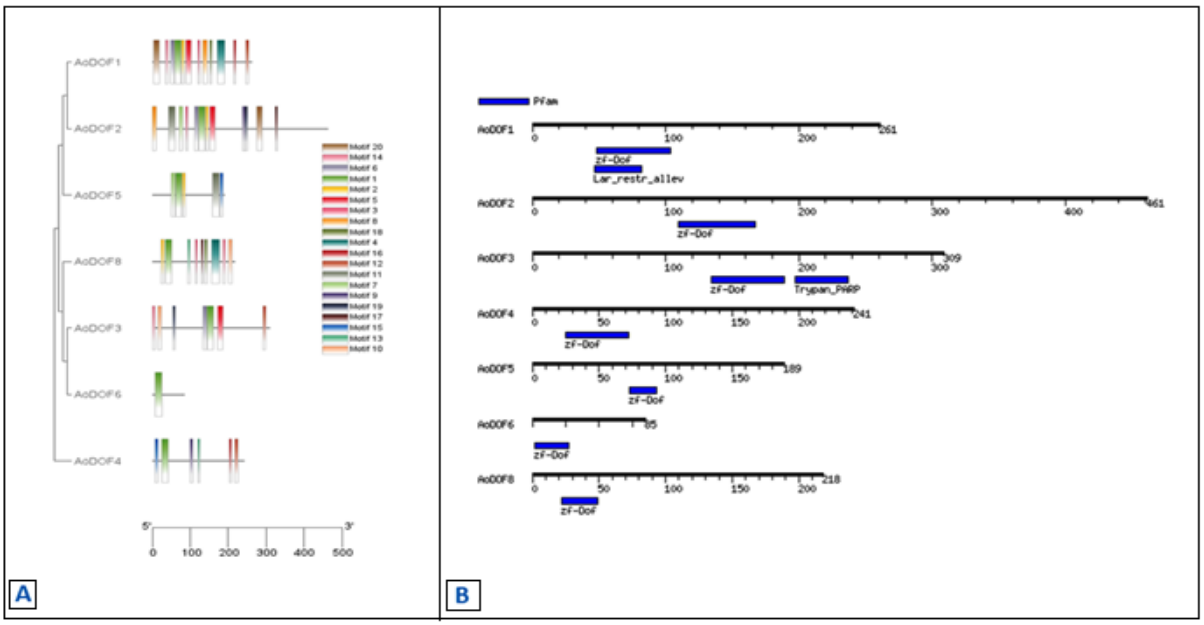


Fig. 3. *The AoDOF* protein domain and motif analysis. (A) The distribution of 20 motifs present in 7 *DOF* proteins of *A. officinalis* by using MEME version 4.9.0 and interlinking it phylogenetic tree to develop a good understanding of their association. (B) The Domain structure and distribution from the protein sequences of *AoDOF* gene. The full-length amino acid sequences of *AoDOF* was uploaded in https://www.genome.jp/tools/motif/pfam-hmmer.htm#Score and structure were retrieved.

Analysis of Cis-regulatory elements

The time and space-related transcriptomic expression of genes is affected by the presence and organization of various *cis*-regulatory elements (CREs) at the binding site of transcription factors on the promoter region (Jones and Vandepoele, 2020). Analysis of various cis-regulatory elements perhaps can be employed to evaluate the putative functions of genes based on bioinformatics (Bulow and Hehl, 2016; Jones and Vandepoele, 2020).

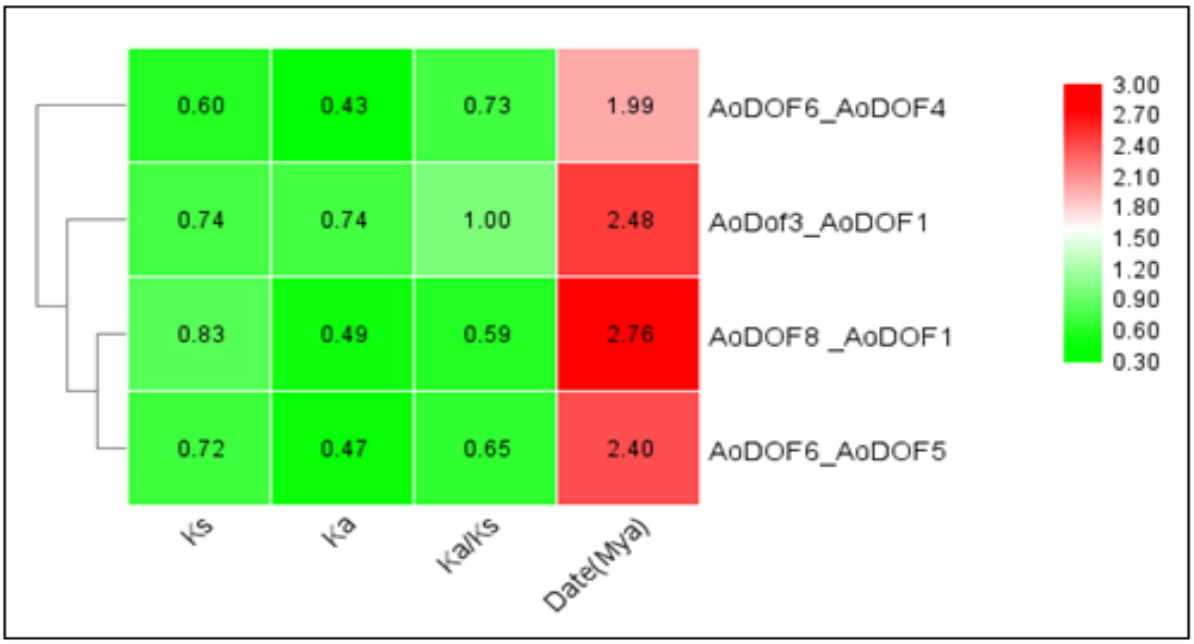


Fig. 4. Time of gene duplication estimated for different paralogous pairs of AoDOF genes on the basis of Ks and Ka values. Analyses were conducted using the Nei-Gojobori model.

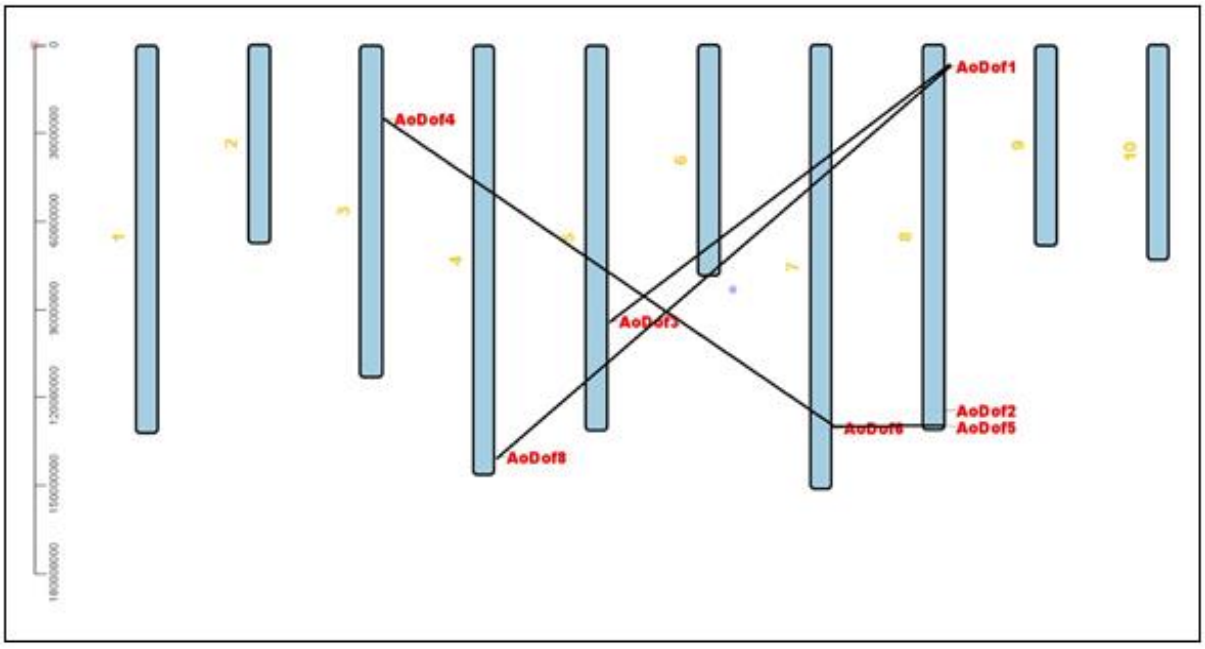


Fig. 5. Distribution of *AoDOF* genes on *A. officinalis* chromosomes. The scale represents a bp (base pair) chromosomal distance. Chromosome map was generated using tbtool. The black lines link the segmented duplicate gene. Genome-wide chromosomal analysis of *A. officinalis DOF* genes showing the dominance of segmental duplication.

The CREs evaluation of recognized *A. officinalis DOF* genes was performed by recovering 1500 base pairs sequence upstream from the starting site of *A. officinalis* genomic sequences using Phytozome database. An ample number of cis-regulatory elements were analyzed in all the 7 *A. officinalis DOF* genes by employing PlantCare database. CREs related to principle physiological processes such as response to light, circadian, wound, seed regulation and endosperm specific, hormone specific (Auxin MeJA SA and ABA), meristem specific and stress were observed (Fig. 6 and Table 9).

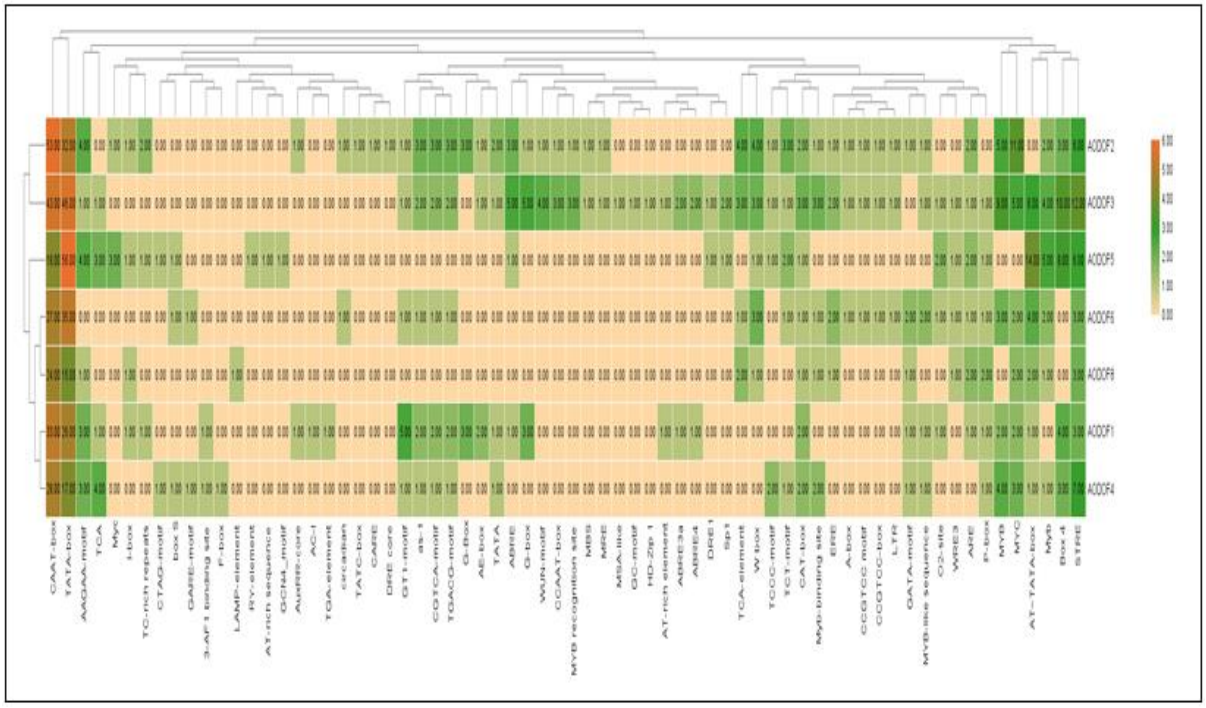


Fig. 6. Different *cis*-acting elements in putative *AoDOF* promoters which are associated with abiotic stresses, hormone responses, growth and development.

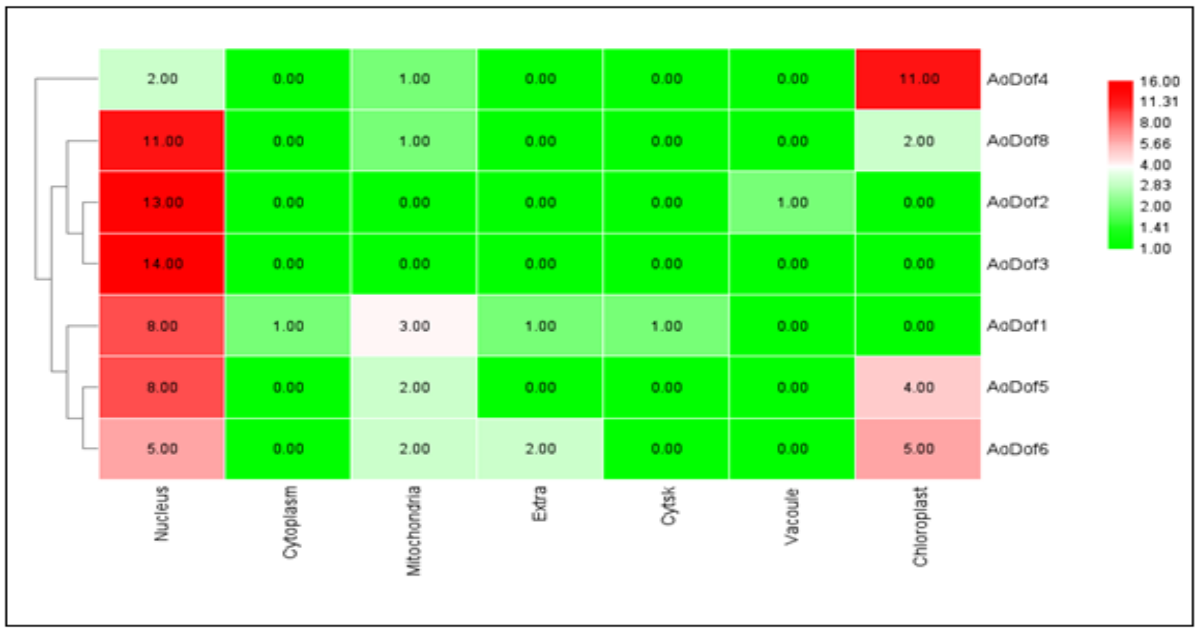


Fig. 7. Subcellular localization in putative *AoDOF* gene and showed the presence of more nuclear localization in all genes except in Ao*DOF* 4 which shows more chloroplast localization.

Mainly, Ao*DOF*3 genes contain elements essential for stress response (MYb and ABA response). All Ao*DOF* genes possess element involved in light responsiveness, 04 *AoDOF* genes possess fragment of a conserved DNA module that takes part in light responsiveness and also element involved in the abscisic acid response, 05 *AoDOF* genes possess elements responsive to methyl Jasmonic acid, 04 *AoDOF* genes possess elements involved in salicylic acid responsiveness, 02 *AoDOF* genes showed wound-responsive elements while 06 *AoDOF* genes showed elements that show response in defense and stress and are related to meristem expression and 02 *AoDOF* genes are related to drought-inducibility, 03

AoDOF genes possess elements involved in low-temperature responsiveness, 02 *AoDOF* genes possess Auxin-responsive element, 01 *AoDOF* genes showed element specific to anoxic inducibility, 01 *AoDOF* genes possess elements specific to seed regulation, 01 *AoDOF* genes possess elements involved in endosperm expression, all *AoDOF* genes possess elements involved in element involved in circadian control and gibberellin response and 01 *AoDOF* genes showed element involved in cell cycle regulation. The cis-regulatory elements identified among 07 *DOF* genes of asparagus along with their functional annotation are shown in (Fig. 6 and Fig. S8).

Fig. 8. The sequence logos of 20 conserved motif present in all *AoDOF* proteins.

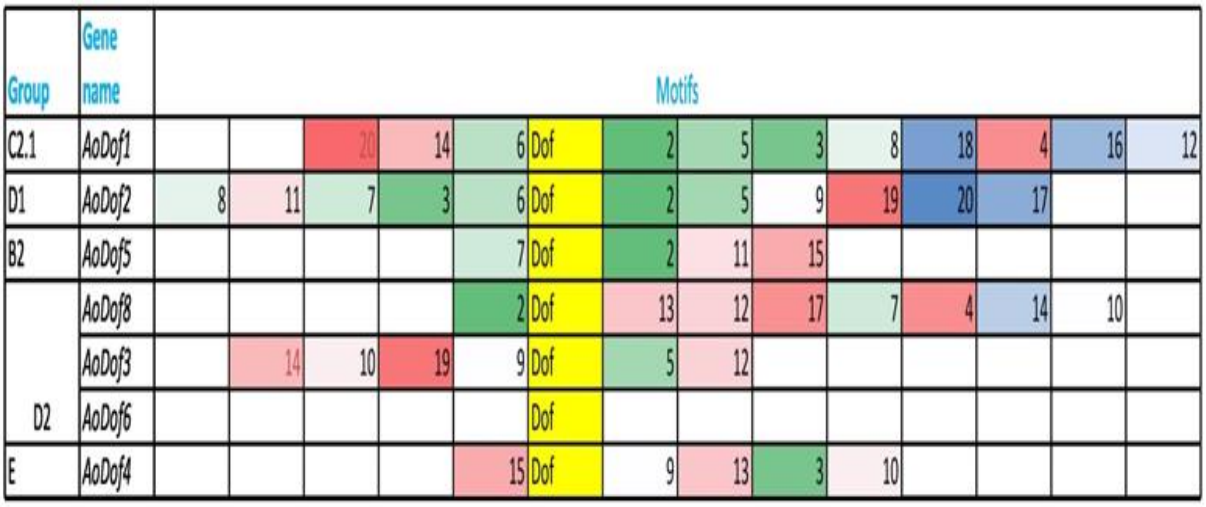
Discussion

Transcription factors (TFs) are important regulatory molecules and has a main role in the regulation of gene transcription and networking. Characterization and identification of TFs provide a better understanding of plant growth and development under environmental stimuli (Yanagisawa and Schmidt, 1999; Wen *et al.*, 2016; Jones and Vandepoele, 2020).

According to the phylogenetic and domain analysis of *A. thaliana* citrus (Wu *et al.*, 2016) and eggplant (Wei *et al.*, 2018), *DOF* transcription factors (TFs) were classified into nine subfamilies (Group A, B1, B2, C1, C2.1, C2.2, C3, D1 and D2). In this study, we used recently released *A.officinalis* genome database (https://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org_Aofficinalis_er) to identify 07 *DOF* gene (Table 1) at the genome wide level. Seven *A.officinalis DOF* gene were classified into five subfamilies (Group B2, E, C2.1, D1 and D2) using the phylogenetic analysis (Fig. 1, Table 2) while A,B1,C2.2,B2,C3 and C1 subfamily (Zou *et al.*, 2013) At*DOF*5 (AT5G21970), At*DOF*6 (AT1G06870), (AT2G30440) and At*DOF*8 are missing in *A.officinalis* genome. At*DOF*1 (AT4G24060) helps in the regulation of transcription, transcription, DNA-templated.

The member *DOF* *gene family* *A.officinalis* was lower as compared to other plant species i.e. banana (74 Ma*DOF*) (Dong *et al.*, 2016), Chinese cabbage (76 BrAT*DOF*) (Ma *et al.*, 2015b) rice (30 Os*DOF*) (Yang and Tuskan, 2006), Arabidopsis (36 At*DOF*) (Yang and Tuskan, 2006) and tomato (34 Si*DOF*) (Cai *et al.*, 2013). For understanding the evolutionary relationships, the exon-intron structure used in different genes or organisms (Koralewski and Krutovsky, 2011;Bondarenko and Gelfand, 2016).

In general, the *A.officinalis* (*AoDOF)* genes in the same subfamily share two similar exon-intron structures but one is different (Fig. 2), further differences are present in different subfamilies. The distribution of motifs among *DOF* proteins (Fig. 3) is indicative of the evolutionary relationship as deduced by a phylogenetic tree (Gupta *et al.*, 2015; Malviya *et al.*, 2015). The motif data analysis by MEME (Fig. 3a), and domain analysis using NCBI CDD (Fig.3b).



Group	Gene name	Motifs													
C2.1	AoDof1			20	14	6	Dof	2	5	3	8	18	4	16	12
D1	AoDof2	8	11	7	3	6	Dof	2	5	9	19	20	17		
B2	AoDof5					7	Dof	2	11	15					
	AoDof8					2	Dof	13	12	17	7	4	14	10	
	AoDof3		14	10	19	9	Dof	5	12						
D2	AoDof6						Dof								
E	AoDof4					15	Dof	9	13	3	10				

Fig. 9. Schematic distributions of the conserved motifs among defined *AoDOF* gene groups. Motifs were identified by means of MEME software using the deduced amino-acid sequences of the 7 *AoDOFs*. The relative position of each identified motif in all *DOF* proteins is shown. Multilevel consensus sequences for the MEME defined motifs are listed in Fig. 2 and Table 7.

The transcription factors of *DOF* gene family have been evolutionarily conserved in plants. Apart from the *DOF* domain, twenty distinct motifs were identified that were differentially distributed among *AoDOF* s (Fig. 3). Meanwhile, at least one conservative motif types and spatial distribution in *AoDOF* s are present in each subfamily while some differences were present between the member of different, implying certain functional similarities of *AoDOF* members within the same subfamily.

Besides, *AoDOF* genes showed structural conservation in subfamilies and was consistent with other plants such as *arabidopsis* banana rice and chickpea (Lijavetzky *et al.*, 2003; Yang and Tuskan, 2006; Nasim *et al.*, 2016; Dong *et al.*, 2016).

Also, as predicted by *in silico* analyses, no *AoDOF* deduced harbored NLSs to help localize them to the nucleus but subcellular localization analysis using online tool WoLF PSORT (https://wolfpsort.hgc.jp/), deduced nucleus localization in all Ao*DOF* protein. MicroRNAs are very important regulators of plants that regulate almost every biological process ranging from growth and development to combating pathogens and maintaining proper internal conditions (Terzi, 2008;Spanudakis, 2014;Samad, 2017; Carbone *et al.*, 2019). miRNAs are highly conserved among different species, meaning that each microRNA performs a specific function, regardless of the type of species in which they were observed. Cyclic *DOF* (CDF2) in arabidopsis binds directly to the promoters of some miRNAs and works as a transcription activator or repressor for miRNA genes (Sun *et al.*, 2015). Growing evidence reveals that miRNA-guided regulation of *DOF* genes at the post-transcriptional level is essential for normal growth and development (Sun *et al.*, 2015; Song *et al.*, 2016; Wen *et al.*, 2016) and help in the regulation of flowering timing and radial and growth of pear (Liu *et al.*, 2017; Miyashima *et al.*, 2019). Many miRNAs target Tfs (transcription factors) play a key role in regulatory pathways (Jones-Rhoades *et al.*, 2006). We predicted that 7 *AoDOFs* could be targeted by 11miRNA families, respectively (Table 4 and Table 5). Aof-miR168c, aof-miR477i, aof-miR477a, aof-miR477b and aof-miR12161 was predicted to be targeted of Ao*DOF* 8 (Table 4 and Table 5). miR168-3p along with miR159a.1, miR164a is the putative ER protein processing and exposed lower expression levels in flowering samples as compared to non-flowering leaves (Qu *et al.*, 2016) showed that miR168 expression has differed between monocots and dicots. Heatmap of microRNA expression profiles in the male and female *A. officinalis* plant targeting *DOF* genes (Fig.3b) revealed that expression of aof-

miR477i is higher in male than female *A. officinalis* which shows their involvement in the development of male *A. officinalis*. Aof-miR477i were also Identified as miRNA-mediated browning regulatory networks in *Luffa cylindrica* (Xu *et al*., 2018).

Conclusion

In this study, *AoDOF* transcription factors were analyzed and supported by genes in the asparagus genome and were classified into five subgroups that characterize the structural and functional properties of each *AoDOF* member. Many of the *AoDOF* genes were involved in flower and stem development. miRNA data targeting *AoDOF* gene revealed that Ao*DOF* 8 was targeted during drupe development in olive suggests their role in fruit growth and development. The detailed computational inspection of olive *DOF* proteins revealed in the current study might be selected for cloning purposes at the molecular level, portraying gene expression and study their interaction with different transcription factors. The presence of an almost similar number of *DOF* genes in some plants such as 33 in tomato, 34 in pepper and 35 in potato genome and the relatively higher number of *DOF* genes in other plants like 78 in soybean and 51 in olive suggests that these genes might be the result of duplications which leads to the expansion of *DOF* gene family.

Acknowledgments

This work was supported by the University of the Punjab, Lahore, Pakistan.

Conflict of Interest

The authors have no potential conflict of interest.

References

Arnaiz A, Martinez M, Gonzalez-Melendi P, Grbic V, Diaz I, Santamaria ME. 2019. Plant Defenses Against Pests Driven by a Bidirectional Promoter. Frontiers in Plant Sciences **10,** 930.

Bailey TL, Johnson J, Grant CE, Noble WS. 2015. The MEME suite**.** Nucleic acids research **43,** W39-W49.

Bondarenko VS, Gelfand MS. 2016. Evolution of the Exon-Intron Structure in Ciliate Genomes. PLoS One **11,** e0161476.

Bueso E, Munoz-Bertomeu J, Campos F, Martinez C, Tello C, Martinez-Almonacid I, Ballester P, Simon-Moya M, Brunaud V, Yenush L, Ferrandiz C, Serrano R. 2016. Arabidopsis COGWHEEL1 links light perception and gibberellins with seed tolerance to deterioration. The Plant Journal **87,** 583-96.

Bulow L, Hehl R. 2016. Bioinformatic Identification of Conserved Cis-Sequences in Coregulated Genes. Methods in Molecular Biology **1482,** 233-45.

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. Journal of Integrative Plant Biology **55,** 552-566.

Carbone F, Bruno L, Perrotta G, Bitonti MB, Muzzalupo I, Chiappetta A. 2019. Identification of miRNAs involved in fruit ripening by deep sequencing of Olea europaea L. transcriptome. PLoS One **14,** e0221460.

Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, Xia R. 2020. TBtools - an integrative toolkit developed for interactive analyses of big biological data. Molecular Plant.

Chen H, Ahmad M, Rim Y, Lucas WJ, Kim J. Y. 2013. Evolutionary and molecular analysis of D of transcription factors identified a conserved motif for intercellular protein trafficking. New Phytologist **198,** 1250-1260.

Chen R, Ni Z, Qin Y, Nie X. 2005. Isolation and characterization of Ta*DOF*1 transcription factor in wheat (Triticum. aestivum. L) Full Length Research Paper. DNA Sequence **16,** 358-363.

Cokol M, Nair R, Rost B. 2000. Finding nuclear localization signals. EMBO reports **1,** 411-415.

Corrales AR, Nebauer SG, Carrillo L, Fernandez-Nohales P, Marques 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. Journal of Experimental Botany **65,** 995-1012.

Diaz I, Martinez M, Isabel-LaMoneda I, Rubio-Somoza I, Carbonero P. 2005. The *DOF* protein, SAD, interacts with GAMYB in plant nuclei and activates transcription of endosperm-specific genes during barley seed development. The Plant Journal **42,** 652-662.

Dong C, Hu H, Xie J. 2016. Genome-wide analysis of the DNA-binding with one zinc finger (*DOF*) transcription factor family in bananas. Genome **59,** 1085-1100.

Dong G, Ni Z, Yao Y, Nie X, Sun Q. 2007. Wheat *DOF* transcription factor WPBF interacts with TaQM and activates transcription of an alpha-gliadin gene during wheat seed development. Plant Molecular Biology **63,** 73-84.

Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, Heger A, Hetherington K, Holm L, Mistry J. 2014. Pfam: the protein families database. Nucleic acids research **42,** D222-D230.

Fornara F. 2009. Arabidopsis *DOF* transcription factors act redundantly to reduce CONSTANS expression and are essential for a photoperiodic flowering response. Europe PMC.

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. Developmental cell **17,** 75-86.

Gardiner J. 2010. Expression of *DOF* genes identifies early stages of vascular development in Arabidopsis leaves. Europe PMC.

Gasteiger E, Hoogland C, Gattiker A, Wilkins MR, Appel RD, Bairoch A. 2005. Protein identification and analysis tools on the ExPASy server. *In* "The proteomics protocols handbook", pp. 571-607. Springer.

Goodstein D, Batra S, Carlson J, Hayes R, Phillips J, Shu S, Schmutz J, Rokhsar D. 2014. Phytozome Comparative Plant Genomics Portal.

Gupta S, Gupta SM, Gupta AK, Gaur VS, Kumar A. 2014. Fluctuation of *DOF*1/*DOF*2 expression ratio under the influence of varying nitrogen and light conditions: involvement in differential regulation of nitrogen metabolism in two genotypes of finger millet (Eleusine coracana L.). Gene **546,** 327-35.

Gupta S, Malviya N, Kushwaha H, Nasim J, Bisht NC, Singh V, Yadav D. 2015. Insights into structural and functional diversity of *DOF* (DNA binding with one finger) transcription factor. Planta **241,** 549-562.

Harkess A, Zhou J, Xu C, Bowers JE, Van der Hulst R, Ayyampalayam S, Mercati F, Riccardi P, McKain MR, Kakrana A, Tang, H, Ray J, Groenendijk J, Arikit S, Mathioni SM, Nakano M, Shan H, Telgmann-Rauber A, Kanno A, Yue Z, Chen H, Li W, Chen Y, Xu X, Zhang Y, Luo S, Gao J, Mao Z, Pires JC, Luo M, Kudrna, D, Wing RA, Meyers BC, Yi K, Kong H, Lavrijsen P, Sunseri F, Falavigna A, Ye Y, Leebens-Mack JH, Chen G. 2017. The asparagus genome sheds light on the origin and evolution of a young Y chromosome. Nature Communications **8,** 1279.

Hartung AC, Nair MG, Putnam AR. 1990. Isolation and characterization of phytotoxic compounds from asparagus (*Asparagus officinalis L* .) roots. Journal of Chemical Ecology **16,** 1707-18.

Hernando-Amado S, Gonzalez-Calle V, Carbonero P, Barrero-Sicilia C. 2012. The family of *DOF* transcription factors in Brachypodium distachyon: phylogenetic comparison with rice and barley *DOF*s and expression profiling. BMC Plant Biology **12**, 202.

Higo K, Ugawa Y, Iwamoto M, Higo H. 1998. PLACE: a database of plant cis-acting regulatory DNA elements. Nucleic Acids Research **26,** 358-9.

Higo K, Ugawa Y, Iwamoto M, Korenaga T. 1999. Plant cis-acting regulatory DNA elements (PLACE) database: 1999. Nucleic Acids Research **27,** 297-300.

Horton P, Park KJ, Obayashi T, Nakai K. 2006. Protein subcellular localization prediction with WoLF PSORT. *In* "Proceedings of the 4th Asia-Pacific bioinformatics conference", p 39-48. World Scientific.

Hu B, Jin J, Guo AY, Zhang H, Luo J, Gao G. 2015. GSDS 2.0: an upgraded gene feature visualization server. Bioinformatics **31,** 1296-7.

Imaizumi T. 2005. FKF1 F-box protein mediates cyclic degradation of a repressor of CONSTANS in Arabidopsis. Europe PMC.

Iwamoto M. 2016. MicroRNA-targeted transcription factor gene RDD 1 promotes nutrient ion uptake and accumulation in rice. The Plant Journal.

Jones DM, Vandepoele K. 2020. Identification and evolution of gene regulatory networks: insights from comparative studies in plants. Current Opinion Plant Biology **54,** 42-48.

Konishi M. 2007. Sequential activation of two *DOF* transcription factor gene promoters during vascular development in Arabidopsis thaliana. Europe PMC.

Koralewski TE, Krutovsky KV. 2011. Evolution of exon-intron structure and alternative splicing. PloS one **6,** e18055.

Kumar S, Tamura K, Nei M. 1994. MEGA: Molecular Evolutionary Genetics Analysis software for microcomputers. Computer Applications in Biosciences **10,** 189-91.

Letunic I, Copley RR, Pils B, Pinkert S, Schultz J, Bork P. 2006. SMART 5: domains in the context of genomes and networks. Nucleic Acids Research **34,** D257-60.

Letunic I, Copley RR, Schmidt S, Ciccarelli FD, Doerks T, Schultz J, Ponting CP, Bork P. 2004. SMART 4.0: towards genomic data integration. Nucleic Acids Research **32,** D142-4.

Lijavetzky D, Carbonero P, Vicente-Carbajosa J. 2003. Genome-wide comparative phylogenetic analysis of the rice and Arabidopsis *DOF* gene families. BMC evolutionary biology **3,** 17.

Liu J, Cheng X, Liu P, Li D, Chen T, Gu X, Sun J. 2017. MicroRNA319-regulated TCPs interact with FBHs and PFT1 to activate CO transcription and control flowering time in Arabidopsis. PLoS Genetics **13,** e1006833.

Ma J, Li MY, Wang F, Tang J, Xiong AS. 2015a. Genome-wide analysis of *DOF* family transcription factors and their responses to abiotic stresses in Chinese cabbage. BMC Genomics **16,** 33.

Ma J, Li MY, Wang F, Tang J, Xiong AS. 2015b. Genome-wide analysis of *DOF* family transcription factors and their responses to abiotic stresses in Chinese cabbage. BMC Genomics **16,** 33.

Malviya N, Gupta S, Singh V, Yadav M, Bisht N, Sarangi B, Yadav D. 2015. Genome wide in

silico characterization of *DOF* gene families of pigeonpea (Cajanus cajan (L) Millsp.). Molecular biology reports **42,** 535-552.

Marchler-Bauer A, Anderson JB, Derbyshire MK, DeWeese-Scott C, Gonzales NR, Gwadz M, Hao, L, He S, Hurwitz DI, Jackson JD, Ke Z, Krylov D, Lanczycki CJ, Liebert CA, Liu C, Lu F, Lu S, Marchler GH, Mullokandov M, Song JS, Thanki N, Yamashita RA, Yin JJ, Zhang D, Bryant SH. 2007. CDD: a conserved domain database for interactive domain family analysis. Nucleic Acids Research **35,** D237-40.

Mello B. 2018. Estimating timetrees with MEGA and the TimeTree resource. Molecular biology and evolution **35,** 2334-2342.

Mena M, Cejudo FJ, Isabel-Lamoneda I, Carbonero P. 2002. A role for the *DOF* transcription factor BPBF in the regulation of gibberellin-responsive genes in barley aleurone. Plant Physiology **130,** 111-9.

Miyashima S. 2019. Mobile PEAR transcription factors integrate positional cues to prime cambial growth. Europe PMC.

Miyashima S, Roszak P, Sevilem I, Toyokura K, Blob B, Heo JO, Mellor N, Help-Rinta-Rahko H, Otero S, Smet W, Boekschoten M, Hooiveld G, Hashimoto K, Smetana O, Siligato R, Wallner ES, Mahonen AP, Kondo Y, Melnyk CW, Greb T, Nakajima K, Sozzani R, Bishopp A, De Rybel B, Helariutta Y. 2019. Mobile PEAR transcription factors integrate positional cues to prime cambial growth. Nature **565,** 490-494.

Moreno-Risueno MA, Martinez M, Vicente-Carbajosa J, Carbonero P. 2007. The family of *DOF* transcription factors: from green unicellular algae to vascular plants. Molecular Genetics and Genomics **277,** 379-90.

Nasim J, Malviya N, Kumar R, Yadav D. 2016. Genome-wide bioinformatics analysis of *DOF* transcription factor gene family of chickpea and its comparative phylogenetic assessment with Arabidopsis and rice. Plant Systematics and Evolution **302,** 1009-1026.

Negi J, Moriwaki K, Konishi M, Yokoyama R, Nakano T, Kusumi K, Hashimoto-Sugimoto, M, Schroeder JI, Nishitani K, Yanagisawa S. 2013. A *DOF* transcription factor, SCAP1, is essential for the development of functional stomata in Arabidopsis. Current Biology **23,** 479-484.

Papi M, Sabatini S, Altamura MM, Hennig L, Schafer E, Costantino P, Vittorioso P. 2002. Inactivation of the phloem-specific *DOF* zinc finger gene DAG1 affects response to light and integrity of the testa of Arabidopsis seeds. Plant Physiology **128,** 411-7.

Qu D, Yan F, Li MA, Varotto C, Zhao ZY. 2016. Comparative analysis of MIR168 promoters in three plant species. Genet Mol Res **15**.

Rombauts S, Déhais P, Van Montagu M, Rouzé P. 1999. PlantCARE, a plant cis-acting regulatory element database. Nucleic acids research **27,** 295-296.

SKGSMLCKKT. 2018. MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. Molecular biology and evolution, 1547-1549.

Samad AFA. 2017. MicroRNA and Transcription Factor: Key Players in Plant Regulatory Network. Frontiers in Plant Science.

Schultz J, Copley RR, Doerks T, Ponting CP, Bork P. 2000. SMART: a web-based tool for the study of genetically mobile domains. Nucleic Acids Research **28,** 231-4.

Schultz J, Milpetz F, Bork P, Ponting CP. 1998. SMART, a simple modular architecture research tool:

identification of signaling domains. Proceedings of the Natlional Academy Sciences of United States of America **95,** 5857-64.

Skirycz A, Reichelt M, Burow M, Birkemeyer C, Rolcik J, Kopka J, Zanor MI, Gershenzon J, Strnad M, Szopa J, Mueller-Roeber B, Witt I. 2006. *DOF* transcription factor At*DOF*1.1 (OBP2) is part of a regulatory network controlling glucosinolate biosynthesis in Arabidopsis. Plant Journal **47,** 10-24.

Song A, Gao T, Li P, Chen S, Guan Z, Wu D, Xin J, Fan Q, Zhao K, Chen F. 2016 Transcriptome-Wide Identification and Expression Profiling of the *DOF* Transcription Factor Gene Family in Chrysanthemum morifolium. Frontiers in Plant Sciences **7**, 199.

Spanudakis E. 2014. The role of microRNAs in the control of flowering time. Journal of Experimental Botany.

Sun Z, Guo T, Liu Y, Liu Q, Fang Y. 2015. The Roles of Arabidopsis CDF2 in Transcriptional and Posttranscriptional Regulation of Primary Micro RNAs. PLoS Genetics **11,** e1005598.

Terzi C. 2008. Regulation of Flowering Time by RNA Processing. *Springer Link*.

Thompson JD, Gibson TJ, Higgins DG. 2003. Multiple sequence alignment using ClustalW and ClustalX. Current protocols in bioinformatics, 2.3. 1-2.3. 22.

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 research **22,** 4673-4680.

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.). Peer-reviewed Journal **6,** e4481.

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. Scientific Reports **6,** 23072.

Wu J, Fu L, Yi H. 2016. Genome-Wide Identification of the Transcription Factors Involved in Citrus Fruit Ripening from the Transcriptomes of a Late-Ripening Sweet Orange Mutant and Its Wild Type. PLoS One **11,** e0154330.

Xu Y, Liu Z, Lou L, Su X. 2018. Identification of browning-related microRNAs and their targets reveals complex miRNA-mediated browning regulatory networks in Luffa cylindrica. Scientific Reports **8,** 16242.

Yanagisawa S. 1995. A novel DNA-binding domain that may form a single zinc finger motif. Nucleic acids research **23,** 3403-3410.

Yanagisawa S. 2000. *DOF*1 and *DOF*2 transcription factors are associated with expression of multiple genes involved in carbon metabolism in maize. Plant Journal **21,** 281-8.

Yanagisawa S. 2002. The *DOF* family of plant transcription factors. Trends in plant science **7,** 555-560.

Yanagisawa S, Akiyama A, Kisaka H, Uchimiya H, Miwa T. 2004. Metabolic engineering with *DOF*1 transcription factor in plants: improved nitrogen assimilation and growth under low-nitrogen conditions. Proceedings of the National Academy of Sciences **101,** 7833-7838.

Yanagisawa S, Schmidt RJ. 1999. Diversity and similarity among recognition sequences of *DOF* transcription factors. The Plant Journal **17,** 209-214.

Yang J, Yang MF, Zhang WP, Chen F, Shen SH. 2011. A putative flowering-time-related *DOF*

transcription factor gene, Jc*DOF*3, is controlled by the circadian clock in Jatropha curcas. Plant Science **181,** 667-674.

Yang X, Tuskan GA. 2006. Divergence of the *DOF* gene families in poplar, Arabidopsis, and rice suggests multiple modes of gene evolution after duplication. Plant physiology **142,** 820-830.

Zhang H, Birch J, Pei J, Mohamed Ahmed IA, Yang H, Dias G, Abd El-Aty AM, Bekhit AE. 2019. Identification of Six Phytochemical Compounds from *Asparagus officinalis L* . Root Cultivars from New Zealand and China Using UAE-SPE-UPLC-MS/MS: Effects of Extracts on H(2)O(2)-Induced Oxidative Stress. Nutrients **11**.

Zohary D, Spiegel-Roy P. 1975. Beginnings of fruit growing in the old world. Science **187,** 319-27.

Zou HF, Zhang YQ, Wei W, Chen HW, Song QX, Liu YF, Zhao MY, Wang F, Zhang BC, Lin Q, Zhang WK, Ma B, Zhou YH, Zhang JS, Chen SY. 2013. The transcription factor At*DOF*4.2 regulates shoot branching and seed coat formation in Arabidopsis. Biochemical Journal **449,** 373-88.