IN SILICO ANALYSIS OF THE Dof TRANSCRIPTION FACTOR FAMILY IN *Coffea canephora*

Vinícius Garcia[1], Alessandra Ferreira Ribas[2], Luiz Gonzaga Esteves Vieira[2], Tiago Benedito dos Santos[1,2]

[1]Universidade do Oeste Paulista – UNOESTE, Especialização Lato Senso em Biotecnologia, Presidente Prudente, SP. [2]Universidade do Oeste Paulista – UNOESTE, Programa de Pós-Graduação em Agronomia, Presidente Prudente, SP. E-mail: dossantos@unoeste.com

ABSTRACT

The Dof family (DNA-binding with One Finger) is a group of transcription factors that are involved in a variety of functions of importance for different biological processes in plants, such as plant growth, development and response to biotic and abiotic stresses. The *Dof* genes have been identified and characterized in many plant species, but so far there is no information about these genes in coffee species. In the present study, we identified 24 Dof members in *Coffea canephora* using the Coffee Genome Hub database. Systematic bioinformatics analyses were performed to characterize all *CcDof* genes, including complete genome sequence, conserved protein domains, subcellular locations, phylogenetic relationships and gene expression profiles in different tissues. The results obtained here provide new insights into the *CcDof* gene family, allowing the design of future experiments for the molecular characterization of these genes in coffee plants.

Key words: Bioinformatics; *C. canephora*; Dof; transcription factor.

ANÁLISE *IN SILICO* DE FATORES DE TRANCRIÇÃO DA FAMÍLIA *Dof* EM *Coffea canephora*

RESUMO

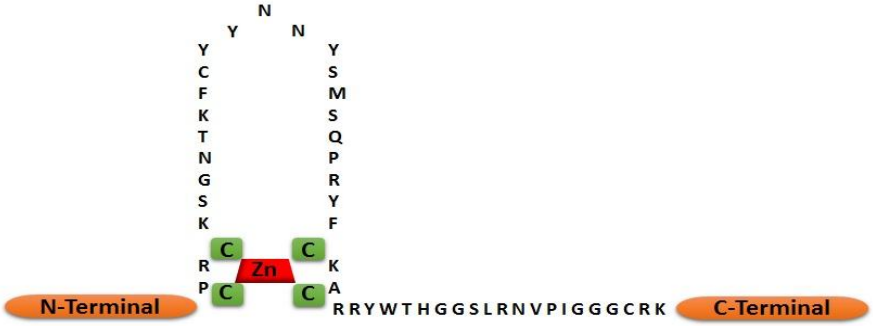
A família *Dof* (*DNA-binding with One Finger*) é um grupo de fatores de transcrição que desempenham papéis importantes no crescimento, desenvolvimento e na resposta das plantas aos estresses bióticos e abióticos. Os genes *Dof* foram identificados e caracterizados em várias espécies de plantas; entretanto até o presente momento não há informações sobre esses genes em café. No presente estudo foram identificados 24 membros da família *Dof* no genoma de *C. canephora* depositados no banco de dados *Coffee Genome Hub*. Análises sistemáticas de bioinformática foram realizadas para caracterizar os genes *Dof* em *C. canephora*, incluindo a análise de sequências genômicas, domínios proteicos conservados, localizações subcelulares, relações filogenéticas e perfis de expressão gênica em diferentes tecidos. Os resultados obtidos fornecem uma melhor compreensão sobre a família dos genes *CcDof* permitindo projetar experimentos futuros para caracterização molecular desses genes no cafeeiro.

Palavras-chave: Bioinformática; *C. canephora*; Dof; fator de transcrição.

INTRODUCTION

The Dof proteins are a family of plant-specific transcription factors (TFs) that play a variety of biological processes, such as biotic and abiotic stresses, seed germination, photosynthesis, secondary metabolic regulation, growth and development (GUPTA et al., 2015; YANG et al., 2018). The conserved Dof domain is formed by 52 amino acid residues is located at the N-terminal region and characterized by the presence of four cysteine residues (C: $CX_2CX_{21}CX_2C$) which bind covalently to a zinc atom (Zn^{2+}) forming a finger-like structure (DNA binding with One Finger) (Figure 1).

Figure 1. Representation of the Dof domain based on the amino acid sequences of *Eleusine coracana*. Figure adapted and reviewed by Gupta et al. (2015).



The Dof TF proteins, in addition to the DNA-binding domain, also contain a bipartite nuclear localization signal (NLS) that overlaps the conserved Dof DNA-binding domain (KREBS et al., 2010). The first report on the identification and characterization of the *Dof* gene was in *Zea mays* (YANAGISAWA; SHEEN, 1998).

Based on bioinformatics analysis generated by the large-scale sequencing genome projects, numerous studies were able to identify families of genes encoding the Dof transcription factors in different plant species: *Ricinus communis* (JIN; LIU, 2014), Chinese cabbage (MA et al., 2015), *Cajanus cajan* (MALVIYA et al., 2015), *Medicago truncatula* (SHU et al., 2015), *Sorghum bicolor* (GUPTA et al., 2016), *Capsicum annuum* (WU et al., 2016), *Phaseolus vulgaris* (ITO et al., 2017), *Setaria italica* (ZHANG et al., 2017a), *Prunus persica* (CHEN et al., 2017), *Jatropha curcas* (WANG et al., 2018), *Ananas comosus* (AZAM et al., 2018), *E. coracana* (GUPTA et al., 2018) and *Gossypium hirsutum* (LI et al., 2018). In this sense, the scope of this study was to identify and characterize the *C. canephora* Dof family based on information available at the Coffee Genome Hub, an integrated web-based database resource developed by IRD and Cirad (http://coffee-genome.org/). This study provides valuable information about candidate genes for future functional analysis for understanding the potential mechanisms of *CcDof* genes in this important coffee species.

MATERIAL AND METHODS

Identification and characterization of the *Dof* genes in *C. canephora*

From searches in the *C. canephora* database available at the Coffee Genome Hub website (http://coffee-genome.org/; DENOUED et al., 2014) were found 24 sequences referring to the keyword Dof. In order to verify if the identified sequences showed similarity with *Dof* genes of other organisms, each putative *CcDof* gene had its sequence individually confronted with the sequences deposited in the NCBI database (BlastX and BlastP; ALTSHUL et al., 1997). Additionally, the physico-chemical characteristics of Dof proteins, such as molecular weight (Mw) and isoelectric point (pI), were calculated using the ExPASy - Compute pI/Mw tool (https://web.expasy.org/compute_pi/). *In silico* analysis for the subcellular location, the Plant-mPLoc (http://www.csbio.sjtu.edu.cn/bioinf/plant-multi/; CHOU; SHEN, 2010) algorithms with default parameters were used. In addition, the hydrophilic/hydrophobic property index of all Dof protein sequences was performed through the GRAVY calculator program.

Chromosomal location and exon/intron structure of *CcDof* genes

All *CcDof* genes identified in the *C. canephora* genome were physically mapped in the chromosomes through the MapChart software (VORORIPS, 2002), according to their genome coordinates. The exon/intron structures of *CcDof* genes were generated by Gene Structure Display Server software (GSDS; http://gsds.cbi.pku.edu.cn/; HU et al., 2015) through a comparison of the predicted coding sequence (CDS) with their corresponding genomic DNA sequences (DENOUED et al., 2014).

Sequence alignment, phylogenetic analysis and classification of *CcDof* genes

Multiple sequence alignment of the identified CcDof protein sequences was performed using the CLC Main Workbench 8.0 program through the tool ClustalW, in order to identify the regions containing the characteristic domain of the Dof family (Figure 4). Next,

phylogenetic trees were generated with MEGA 7 software by using the Neighbor-Joining method, p-distance substitution model, pairwise deletion analysis. The reliability of the trees was assessed by the bootstrap method and branches of less than 50% bootstrap support (1000 replicates) were collapsed.

Conserved motif analysis

The prediction of the conserved Dof motifs of the putative proteins was performed by using the Multiple EM for Motif Elicitation (MEME; http://meme.ebi.edu.au/meme/intro.html; BAILEY et al., 2009), using the same parameters: motif length set to 6 - 100, motif sites set to 2 - 120, and maximum number of motifs set to 20, respectively.

Expression of the Dof gene in different tissues

The tissue-specific expression patterns of the different members of the *Dof* gene family in *C. canephora* was determined by employing the transcriptional profiles obtained by RNA sequencing (RNAseq) from root, stamen, pistil, leaf, perisperm and endosperm libraries (DENOUED et al., 2014). The RPKM (Reads Per Kilobase per Million mapped reads) values of each putative gene from each library were displayed in Heatmap format to infer the transcriptional profiles in the respective tissues of *C. canephora* plants.

RESULTS AND DISCUSSION

Dof (DNA-binding with one finger) domain proteins constitutes an important family of TFs that are associated to the regulatory network of various developmenta processes in plants (SILVA et al., 2016; KANG et al., 2016; CORRALES et al., 2017). As has been shown in the literature, the number of *Dof* genes described and characterized are variable among plant species: *S. tuberosum* (35), *Chrysanthemum morifolium* (20), *P. vulgaris* (36), *S. italica* (35), *A. comosus* L. (26), *Malus domestica* (60), *Boehmeria nivea* L. Gaud (19) (VENKATESH; PARK, 2015; SONG et al., 2016; ITO et al., 2017, ZHANG et al., 2017a; AZAM et al.; 2018, ZHANG et al., 2017b; XU et al., 2018), respectively. In this study, we were able to identify 24 *CcDof* genes within the *C. canephora* genome database (DENOUED et al., 2014), which were named (*CcDof1* through *CcDof24*) in descending order according to their position within the *C. canephora* chromossomes (Table 1).

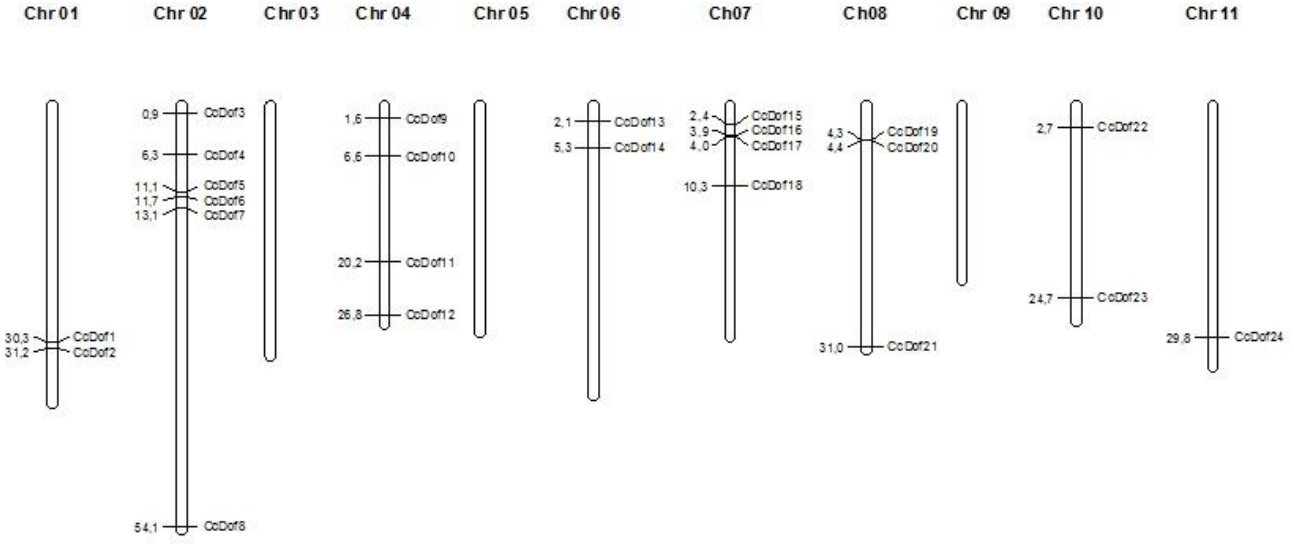
Detailed information on the physico-chemical characteristics of *CcDof* proteins is shown in Table 1. The deduced size of the amino acid sequences of the CcDof proteins ranged from 167 to 513 aa, the theoretical isoelectric point (pI) from 4.24 to 10.13 and the molecular weight (Mw) of from 14289.36 to 49223.78 Mw.

Table 1. List of the CcDof TFs identified in the genome of *C. canephora* and related information on gene name, locus ID, chormossome number, protein length (aa), molecular weight (Mw), isoelectric point (pI), hydropathy (GRAVY), subcellular localization and orthologous in *A. thaliana*

Gene Dof	Locus ID *C. canephora*	Chromosome number	Protein length (aa)	Molecular weight (Mw)	Theoretical (pI)	GRAVY	Subcellular localization	Ortholog in *Arabidopsis*
CcDof1	Cc01_g11670	chr1:30320517..30322291	290	24710.20	7.00	-0.670	Nucleus	AT2G28510.1
CcDof2	Cc01_g12730	chr1:31245653..31246860	307	25986.88	9.17	-0.559	Nucleus	AT3G55370.3
CcDof3	Cc02_g01210	chr2:972655..974150	311	27092.76	6.58	-0.530	Nucleus	AT4G24060.1
CcDof4	Cc02_g07960	chr2:6332990..6334233	336	29745.12	8.71	-0.530	Nucleus	AT3G55370.3
CcDof5	Cc02_g12890	chr2:11166886..11167800	288	24995.71	6.63	-0.501	Nucleus	AT1G07640.3
CcDof6	Cc02_g13590	chr2:11739465..11740752	332	29214.38	9.06	-0.503	Nucleus	AT1G28310.2
CcDof7	Cc02_g14960	chr2:13101644..13102165	174	12530.81	9.38	-0.761	Nucleus	AT1G29160.1
CcDof8	Cc02_g39630	chr2:54118424..54119805	290	24658.45	9.57	-0.520	Nucleus	AT3G61850.4
CcDof9	Cc04_g02150	chr4:1676203..1676925	241	18330.64	6.45	-0.778	Nucleus	AT1G51700.1
CcDof10	Cc04_g08250	chr4:6608932..6609990	353	30627.49	5.64	-0.554	Nucleus	AT5G60850.1
CcDof11	Cc04_g14150	chr4:20217841..20219685	307	27326.52	6.52	-0.645	Nucleus	AT5G62940.1
CcDof12	Cc04_g16510	chr4:26823179..26825416	502	47741.12	7.6	-0.762	Nucleus	AT3G47500.1
CcDof13	Cc06_g02740	chr6:2198639..2199961	356	31614.37	9.39	-0.462	Nucleus	AT2G37590.1
CcDof14	Cc06_g06640	chr6:5305481..5306557	340	30396.35	5.37	-0.527	Nucleus	AT3G52440.1
CcDof15	Cc07_g03470	chr7:2412880..2415816	513	49223.78	6.20	-0.802	Nucleus	AT5G39660.1
CcDof16	Cc07_g05550	chr7:3916861..3917589	243	18288.37	8.04	0.010	Nucleus	AT3G50410.1
CcDof17	Cc07_g05650	chr7:3979174..3980220	349	31143.82	9.01	-0.434	Nucleus	AT5G60200.1
CcDof18	Cc07_g13810	chr7:10399450..10400532	361	32196.76	8.93	-0.475	Nucleus	AT5G65590.1
CcDof19	Cc08_g03320	chr8:4306033..4308169	269	24871.87	9.85	-0.606	Nucleus	AT4G24060.1
CcDof20	Cc08_g03360	chr8:4335942..4338195	167	13459.31	10.13	-0.905	Nucleus	AT4G24060.1
CcDof21	Cc08_g16560	chr8:31049028..31050438	301	25693.09	7.17	-0.604	Nucleus	AT4G24060.1
CcDof22	Cc10_g03720	chr10:2787475..2788311	213	14289.36	6.45	-0.507	Nucleus	AT5G60200.1
CcDof23	Cc10_g14200	chr10:24787234..24788088	285	24832.08	4.24	-0.511	Nucleus	AT1G21340.1
CcDof24	Cc11_g12870	chr11:29855782..29858689	463	43776.71	8.61	-0.730	Nucleus	AT5G39660.1

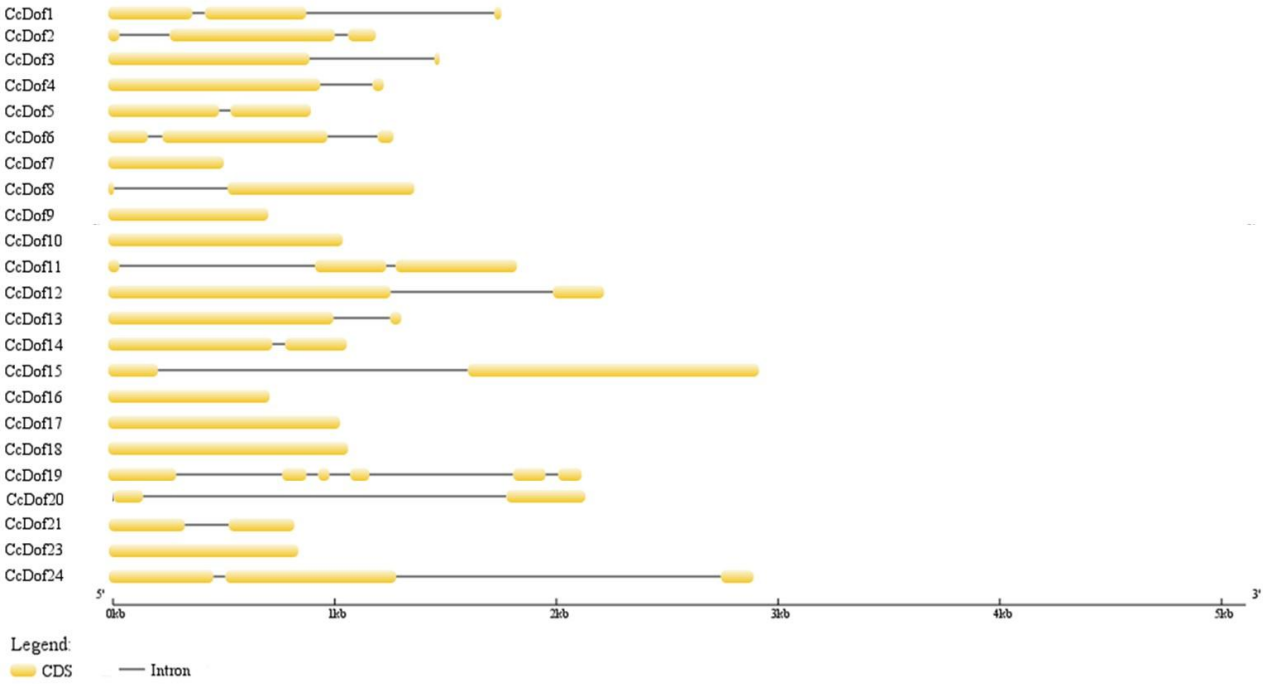
The 24 *CcDof* genes were unevenly distributed in almost all 11 chromosomes of *C. canephora*, with the exception of chromossomes 3, 5 and 9 (Figure 2). Chromosomes 2, the larger *C. canephora* chromosome (55 Mb), contained the highest number of *Dof* genes (6) (Figure 2). For comparison's sake, the 34 *Dofs* genes in *S. lycopersicum* were identified and distributed in 12 chromosomes, with the exception of chromosomes 7 and 12 (CAI et al., 2013), while in *A. comosus*, Azam et al. (2018) the identified 26 *Dofs* genes were mapped on all the 25 chromosomes of that species.

Figure 2. Distribution and chromosomal location of the *CcDof* genes in the *C. canephora* genome.



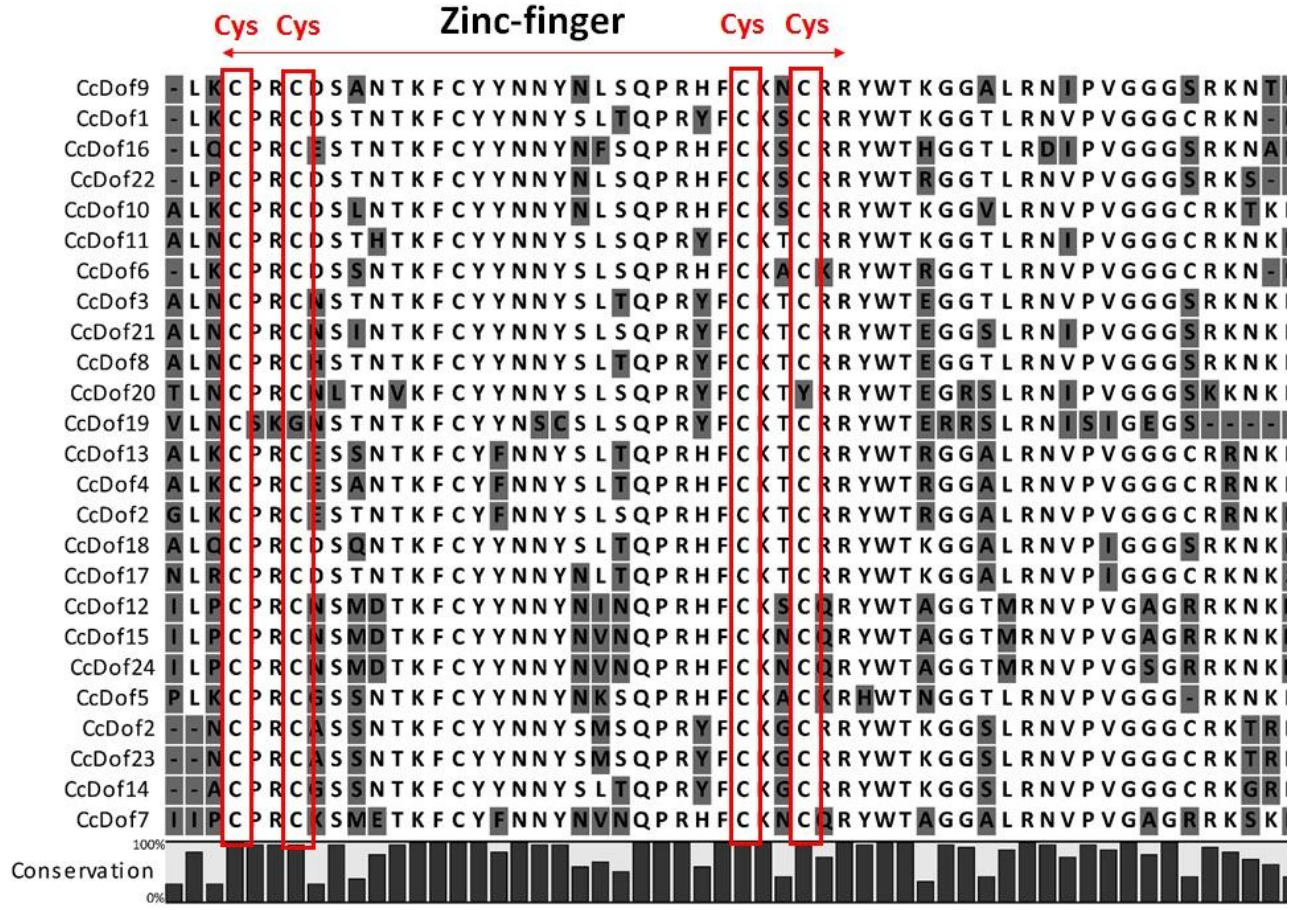
According to Koralewski and Krutovsky (2011), the structural and exon/intron organization can be used for examining the evolutionary relationships among genes or organisms. As commonly reported (KUSHWAHA et al., 2011; CHEN et al., 2017; ZHANG et al., 2017b), the distribution of intron in Dof genes is variable among plant species. In *C. canephora*, the exon/intron analysis depicted that the number of introns in the open reading frames (ORFs) varied from 0 to 5, with *CcDof19* presenting a maximum of 5 introns (Figure 3). The fact that *CcDof* genes presents on average less introns than other genes makes them more sensitive to transcriptional regulation, which suggests that they may trigger diverse biological processes in the plant (JIN et al., 2014).

Figure 3. Gene structure of the 24 *CcDof* genes in *C. canephora* genome. Yellow boxes indicate exon (CDS) and lines introns.



As mentioned earlier, the Dof TF family is characterized by having the conserved N-terminal DNA-binding domain ($CX_2CX_{21}CX_2C$) (YANAGISAWA, 1995; 1997). The multiple sequence alignment of the CcDof proteins revealed the typical presence of amino acid residues at the the Zinc-finger domain (Figure 4), similarly to that described for other plant species, such as banana (DONG et al., 2016), moso bamboo (WANG et al., 2016) and apple (ZHANG et al., 2017b).

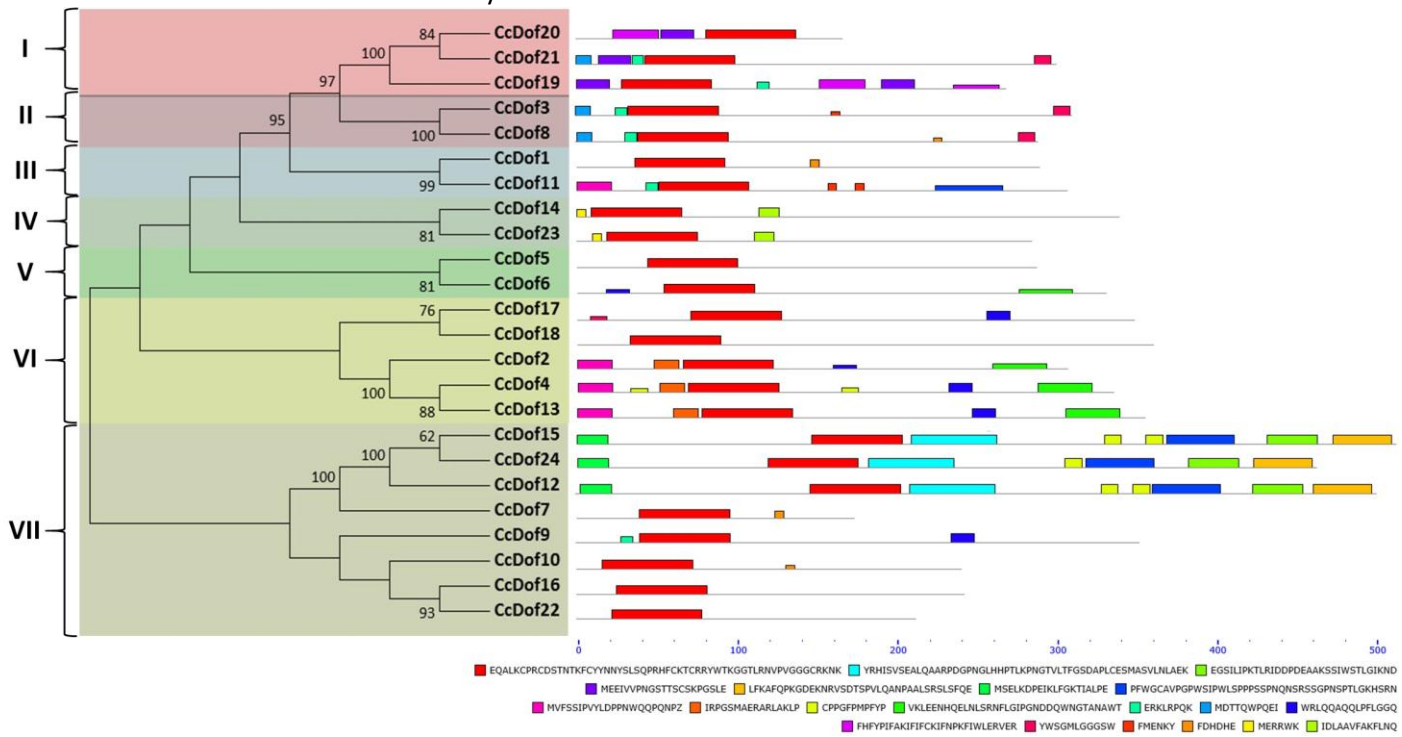
Figure 4. Multiple alignment of amino acid sequences of the DNA-binding domain of the *C. canephora* Dof transcription factors.



It is well known that classificaton of genes and phylogeny are important to the functional analysis and to understand the evolutionary history of a particular gene family. The phylogenetic tree topology allowed to classify the CcDofs into 6 groups (groups I-VII). Each group was composed of a varied number of Dof genes (2 to 8 members) with a bootstrap value relatively high for each subgroup, indicating the relatedness of the Dof genes within each group (Figure 5). In *Ricinus communis* Dof FTs were clustered in four groups, which could be divided into seven subgroups (JIN et al., 2014), while Wei et al. (2018) divided the Dof TFs of eggplant in 4 groups and 9 subgroups.

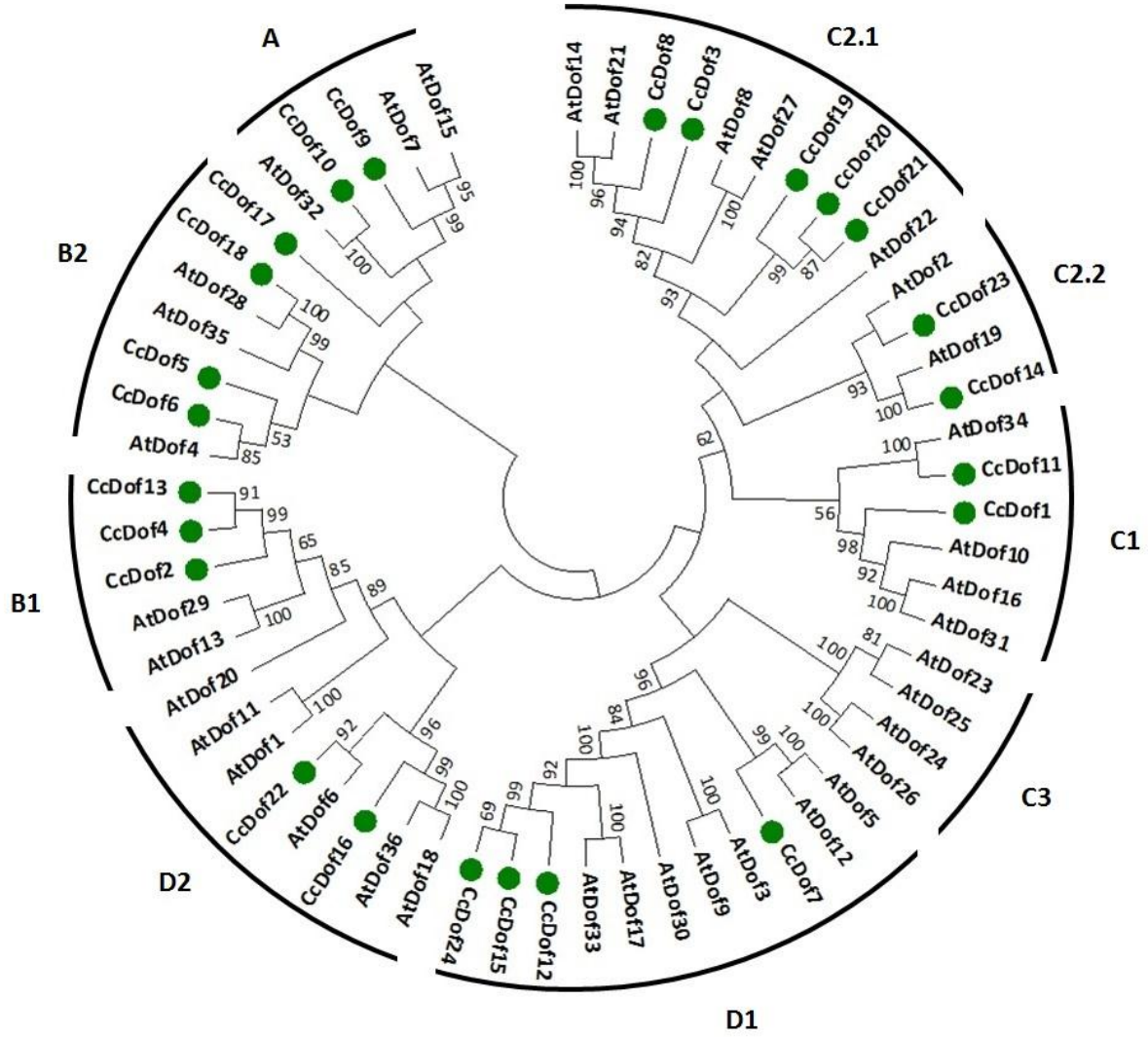
Using bioinformatics tools is possible to identify sequence Motifs, which represent highly conserved regions (functional regions) sharing the same function and, probably, having a common origen (GIBAS et al., 2001). A total of 24 conserved motifs have been identified in the CcDof proteins (Figure 5). The motif 1 (EQALKCPRCDSTNTKFCYYNNYSLQPRHFCKTCRRY WTKGGTLRNVPVGGGSRKNK) was confirmed as the conserved Dof domain in almost all sequences of the CcDof proteins (Figure 5).

Figure 5. Phylogenetic tree constructed from the alignment of the 24 CcDof proteins. The sequences were aligned using ClustalW at MEGA 7.0 software and the phylogenetic tree was constructed by Neighbor-Joining method. Different color boxes distinguish the groups. The motif analysis was performed using MEME tool and the motifs are indicated by different-colored boxes.



Another phylogenetic tree was constructed including 36 AtDof sequences from *Arabidopsis* aiming at clustering the CcDofs according the classification proposed by Lijavetzky et al. (2003) (Figure 6), where Dof proteins were grouped into four major clusters (A, B, C and D) and eight subclusters designated as B1, B2, C1, C2.1, C2.2, C3, D1 and D2.

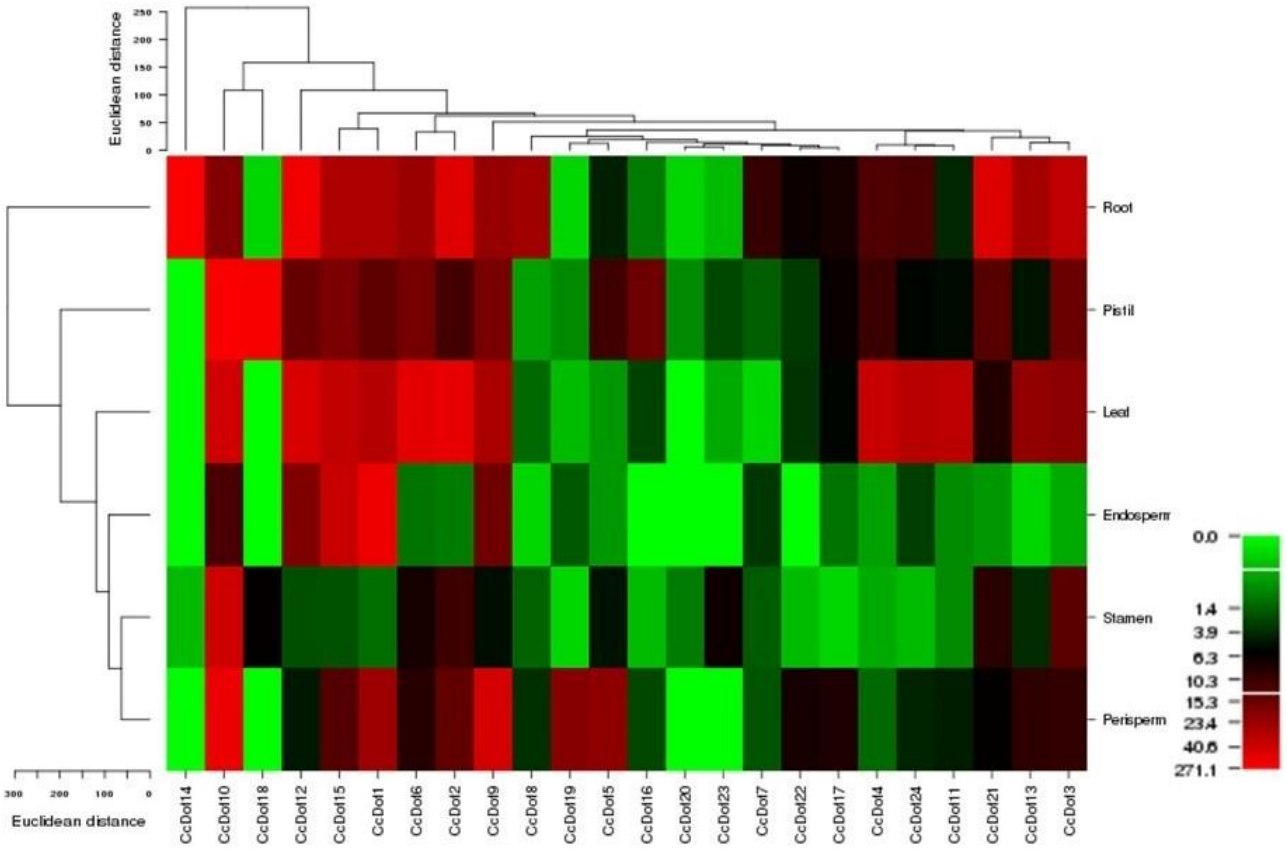
Figure 6. Phylogenetic tree of Dof proteins from *C. canephora* and *A. thaliana*. The 24 *C. canephora* and 36 *A. thaliana* protein sequences were aligned using ClustalW at MEGA 7.0 software and the phylogenetic tree was constructed by Neighbor-Joining method.



The groups B and C had the largest number of CcDofs, with 7 and 9 genes respectively, while group A represents the smallest clade with only 2 members (CcDof9 and 10) (Figure 6). The subgroup C3, which was originally identified in *Arabidopsis* (AtDofs 22, 23, 24, 25 and 26), do not possess any CcDof (Figure 6). Our results corroborate with earlier studies studies that demonstrated that this subgroup is found only in a few species, such as in cruciferous plants, with no apparent homologues observed in important crops as rice, tomato, chinese cabbage and banana (LIJAVETZKY; CARBONERO; VICENTE-CARBAJOSA; 2003; CAI et al., 2013; MA et al., 2015; FENG et al., 2016).

Recently, Denoued et al. (2014) provided the complete sequence of the genome of *C. canephora*, a perennial diploid species (2n = 2x = 22 chromosomes), also known as Robusta. In the scope of that same project, the RNA-seq data of various coffee cDNA libraries (roots, stamen, pistil, leaves, perisperm and endosperm) were made available at the Coffee Genome Hub platform (http://coffee-genome.org/). Using the RPKM data, we investigated the expression patterns of the 24 *CcDof* genes in those different tissues of *C. canephora* (Figure 7).

Figure 7. *In silico* expression profile of the *CcDof* genes of *C. canephora*. Heat map showing the expression pattern of 24 genes in different tissues. The expression of the genes is represented in the heat map in the colour scale of 0–272 in green-red colour scheme. The genes are represented in columns while tissues are shown in rows.



The heat map generated for *in silico* expression profiling showed differential transcript abundance of the *CcDof* genes in all libraries (Figure 7). For example, were able to identify 14 *CcDof* genes (*CcDof1, 2, 3, 4, 6, 8, 9, 10, 12, 13, 14, 15, 21* and *24)* abundantly expressed in roots. A large number of *CcDof* genes were also expressed in leaf tissues (*CcDof1*, *2*, *3*, *4*, *6*, *9*, *10*, *11*, *12*, *13*, *15* and 24). Similarly to our findings, Xu et al. (2018), analyzing the expression patterns of Dof genes in *Boehmeria nivea*, observed that many of these genes were highly expressed in roots and leaves. In potato, an expressive number of *StDof* genes (*StDof15a, StDof22, StDof24, StDof26, StDof29a, StDof32* and *StDof34*) showed maximum expression levels in the leaf tissues (VENKATESH; PARK, 2015). Also, specific expression of Dof genes in root, leaf and petiole tissues of *Daucus carota* was observed by Huang et al. (2016).

On the other hand, only 6 (*CcDof1*, *2*, *5*, *9, 10* and *19*) and 3 (*CcDof1*, *12* and *15*) *CcDofs* were highly expressed in the perisperm and in the endosperm, respectively. Particularly regarding the reproductive organs, the Dof family genes were more highly expressed in the pistils, as the cDNA library from stamen tissues showed lower number of transcripts of the majority of the identified *CcDof* genes (Figure 7). It is also worth mentioning that some *CcDof* genes (e.g. *CcDof1*, *2*, *10*, *12* and *15*) were ubiquitously expressed in the various coffee tissues. Finally, here we have also shown that the *CcDof* genes under investigations clustered together based on their expression in different tissues, which clearly indicates that they are involved in diverse physiological functions in *C. canephora*.

CONCLUSIONS

This study is the first to present data on *in silico* identification and analysis of *C. canephora* Dof family genes. A total of 24 *CcDof* members have been identified and were classified into seven phylogenetically related groups. Also, we conducted a detailed analysis regarding their gene structures, protein motifs and chromosome distribution. A heat map analysis of the *CcDof* genes revealed that they are primarily expressed in root and leaf tissues. Finally, further deep studies for evaluation of the

mechanism affected by the *CcDof* gene may help to devise new strategies and biotechnological applications for *C. canephora* plants.

REFERENCES

ALTSCHUL, S.F.; MADDEN, T.L.; SCHAFFER, A.A.; ZHANG, J.; ZHANG, Z.; MILLER, W.; LIPMAN, D. Gapped BLAST and PSIBLAST: a new generation of protein database search programs. **Nucleic Acids Research**, v.25, p.3389-3402, 1997. https://doi.org/10.1093/nar/25.17.3389

AZAM, S. M.; LIU, Y.; RAHMAN, Z. U.; ALI, H.; YAN, C.; WANG, L.; PRIYADARSHANI, S.V.G.N.; HU, B.; HUANG, X.; XIONG, J.; QIN, Y. Identification, Characterization and Expression Profiles of *Dof* Transcription Factors in Pineapple (*Ananas comosus* L.). **Tropical Plant Biology**, p. 1-16, 2018. https://doi.org/10.1007/s12042-018-9200-8.

CAI, X.; ZHANG, Y.; ZHANG, C.; ZHANG, T.; HU, T.; YE, J.; ZHANG, J.; WANG, T.; LI, H.; YE, Z. Genome-wide analysis of plant-specific *Dof* transcription factor family in tomato. **Journal of Integrative Plant Biology**, v.55, n.6, p.552-566, 2013. https://doi.org/10.1111/jipb.12043.

CHEN, M.; LIU, X.; HUAN, L.; SUN, M.; LIU, L.; CHEN, X.; GAO, D.; LI, L. Genome-wide analysis of Dof family genes and their expression during bud dormancy in peach (*Prunus persica*). **Scientia Horticulturae**, v.214, p.18-26, 2017. https://doi.org/10.1016/j.scienta.2016.11.014.

CORRALES, A.R.; CARRILLO, L.; LASIERRA, P.; NEBAUER, S. G.; DOMÍNGUEZ-FIGUEROA, J.; RENAU-MORATA, B.; POLLMANN, S.; GRANELL, A.; MOLINA, R. V.; VICENTE-CARBAJOSA, J.; MEDINA, J. Multifaceted role of cycling Dof Factor 3 (CDF3) in the regulation of flowering time and abiotic stress responses in Arabidopsis. **Plant Cell Environment**, v.40, p.748–764, 2017. https://doi.org/10.1111/pce.12894.

DENOEUD, F.; CARRETERO-PAULET, L.; DEREEPER, A.; DROC, G.; GUYOT, R.; PIETRELLA, M.; ZHENG, C.; ALBERTI, A.; ANTHONY, F.; APREA, G. et al. The coffee genome provides insight into the convergent evolution of caffeine biosynthesis. **Science**, v.345, n.6201, p.1181-1184, 2014. https://doi.org/10.1126/science.1255274.

FENG, B.H.; HAN, Y.-C, XIAO, Y.-Y. ; KUANG, J.-F.; FAN, Z.-Q.; CHEN, J.-Y.; LU, WJ. The banana fruit Dof transcription factor MaDof23 acts as a repressor and interacts with MaERF9 in regulating ripening-related genes. **Journal of Experimental Botany**, v. 67, p. 2263-2275, 2016. https://doi.org/10.1093/jxb/erw032.

GIBAS, C.; JAMBECK, P. **Developing Bioinformatics Computer Skills**. 1. ed. Lorrie Lejeune: O'Reilly Media, 2001 p. 446.

HU, BO.; JIN, J.; GUO, A-Y.; ZHANG, HE.; LUO, J.; GAO, G. GSDS 2.0: an upgraded gene feature visualization server. **Bioinformatics,** v.31, n.8, p.1296-1297, 2015. https://doi.org/10.1093/bioinformatics/btu817.

GUPTA, S.; ARYA, G. C.; MALVIYA, N.; BISHT, N.C.; YADAV, D. Molecular cloning and expression profiling of multiple Dof genes of *Sorghum bicolor* (L) Moench. **Molecular Biology Reports**, v.43, n.3, p.767-774, 2016. https://doi.org/10.1007/s11033-016-4019-6.

GUPTA, S.; MALVIYA, N.; KUSHWAHA, H.; NASIM, J.; BISHT, N. C.; SINGH, V. K.; YADAV, D. Insights into structural and functional diversity of Dof (DNA binding with one finger) transcription factor. **Planta**, v. 241, n. 3, p. 549-562, 2015. https://doi.org/10.1007/s00425-014-2239-3.

GUPTA, S.; PATHAK, R. K.; GUPTA, S. M.; GAUR, V. S.; SINGH, N. K.; KUMAR, A. Identification and molecular characterization of Dof transcription factor gene family preferentially expressed in developing spikes of *Eleusine coracana* L. **3 Biotech**, v.8, n.2, p.82, 2018. https://doi.org/10.1007/s13205-017-1068-z.

HUANG, W.; HUANG, Y.; LI, M.Y.; WANG, F.; XU, Z.S.; XIONG, A.S. 2016. Dof transcriptionfactors in carrot: genome-wide analysis and their response to abiotic stress. **Biotechnology Letters,** v. 38, p. 145–155, 2016. https://doi.org/10.1007/s10529-015-1966-2.

HOU, Y.; WU, A. L.; HE, Y. Y.; LI, F. D.; WEI, C. L. Genome-wide characterization of the basic leucine zipper transcription factors in *Camellia sinensis*. **Tree Genetics & Genomes**, v. 14, p. 27, 2018. https://doi.org/10.1007/s11295- 018-1242-4.

ITO, T.M.; TREVIZAN, C.B.; DOS SANTOS, T.B.; DE SOUZA, S.G.H. Genome-Wide Identification and

Characterization of the Dof Transcription Factor Gene Family in *Phaseolus vulgaris* L. **American Journal of Plant Sciences**, v. 8, p. 3233-3257, 2017. https://doi.org/10.4236/ajps.2017.812218.

JIN, Z.; CHANDRASEKARAN, U.; LIU, A. Genome-wide analysis of the Dof transcription factors in castor bean (*Ricinus communis* L.). **Genes Genomics,** v.36, p.527-537, 2014. https://doi.org/10.1007/s13258-014-0189-6.

KANG, W.-H.; KIM, S.; LEE, H.-A.; CHOI, D; YEOM, S.-I. Genome-wide Analysis of Dof Transcription Factors Reveals Functional Characteristics during Development and Response to Biotic Stresses in Pepper. **Scientific Reports**, v.6, n.33332, 2016. https://doi.org/10.1038/srep33332.

KORALEWSKI, T.E.; KRUTOVSKY, K. V. Evolution of exon-intron structure and alternative splicing. **PLoS One,** v.6, e18055. 2011 https://doi.org/10.1371/journal.pone.0018055.

KREBS, J.; MUELLER-ROEBER, B.; RUZICIC, S. A novel bipartite nuclear localization signal with an atypically long linker in DOF transcription factors. **Journal of Plant Physiology**, v.167, n.7, p.583-586, 2010. https://doi.org/10.1016/j.jplph.2009.11.016.

KUSHWAHA, H.; GUPTA, S.; SINGH, V.K.; RASTOGI, S.; YADAV, D. Genome wide identification of Dof transcription factor gene family in sorghum and its comparative phylogenetic analysis with rice and Arabidopsis. **Molecular Biology Reports**, v. 38, p. 5037-5053, 2011. https://doi.org/10.1007/s11033-010-0650-9.

LI, DH.; LIU, P.; YU, J. Y.; WANG, L. H.; DOSSA, K.; ZHANG. Y. X.; ZHOU, R.; WEI, X.; ZHANG, X. R. Genome-wide analysis of WRKY gene family in the sesame genome and identification of the WRKY genes involved in responses to abiotic stresses. **BMC Plant Biology** v. 17, n. 1, p. 152, 2017. https://doi.org/10.1186/s12870-017-1099-y.

LIJAVETZKY, D.; CARBONERO, P.; VICENTE-CARBAJOSA, J. Genome-wide comparative phylogenetic analysis of the rice and Arabidopsis Dof gene families. **BMC Evolutionary Bioogy**. v. 3, n.17, 2003. https://doi.org/10.1186/1471-2148-3-17.

MA, J.; LI, M-Y.; WANG, F.; TANG, J.; XIONG, A. S. Genome-wide analysis of Dof family transcription factors and their responses to abiotic stresses in Chinese cabbage. **BMC Genomics** v. 16, p. 33, 2015. https://doi.org/10.1186/s12864-015-1242-9.

MALVIYA, N.; GUPTA, S.; SINGH, V.K.; YADAV, M.K.; BISHT, N.C.; SARANGI, B.K.; YADAV, D. Genome wide *in silico* characterization of *Dof* gene families of pigeonpea (*Cajanus cajan* (L) Millsp.). **Molecular Biology Reports**, v. 42, n. 2, p. 535-552, 2015. https://doi.org/10.1007/s11033-014-3797-y.

MORENO-RISUENO, M.Á.; DÍAZ, I.; CARRILLO, L.; FUENTES, R.; CARBONERO, P. The HvDOF19 transcription factor mediates the abscisic acid-dependent repression of hydrolase genes in germinating barley aleurone. **Plant Journal,** v. 51, p.352–365, 2007. https://doi.org/10.1111/j.1365-313X.2007.03146.x.

PROSDOCIMI, F.; CERQUEIRA G. C.; BINNECK, E.; SILVA, A. F.; NEVES, A. ET AL. Bioinformática: manual do usuário. **Biotecnologia Ciência & Desenvolvimento**, v. 29, p. 12-25, 2002.

ROY, S.W.; GILBERT, W. The evolution of spliceosomal introns: patterns, puzzles and progress. **Nature Reviews Genetics**, v. 7, p. 211–221, 2006. https://doi.org/10.1038/nrg1807.

SHU, Y.J.; SONG, L.L.; ZHANG, J.; LIU, Y.; GUO, C. H. Genome-wide identification and characterization of the *Dof* gene family in *Medicago truncatula*. **Genetics and Molecular Research**, v.14, n.3, p.10645-10657, 2015. https://doi.org/10.4238/2015.September.9.5.

SILVA, D.C.; FALAVIGNA, V.S.; FASOLI, M.; BUFFON, V.; PORTO, D.D.; PAPPAS JR, G.J.; PEZZOTTI, M.; PASQUALI, G.; REVERS, L.F. Transcriptome Analyses of the Dof-Like Gene Family in Grapevine Reveal Its Involvement in Berry, Flower and Seed Development. **Horticulture Research**, v.3, n.16042, 2016. https://doi.org/10.1038/hortres.2016.422.

SONG, A.; GAO, T.; LI, P.; CHEN, S., WU, D.; XIN, J.; FAN, Q.; ZHAO, K.; CHEN, F. Transcriptome-Wide Identification and Expression Profiling of the DOF Transcription Factor Gene Family in *Chrysanthemum morifolium*. **Frontiers in Plant**

Science v.7, n.199, 2016. https://doi.org/10.3389/fpls.2016.00199.

VENKATESH, J.; PARK, S.W. Genome-wide analysis and expression profiling of DNA-binding with one zinc finger (Dof) transcription factor family in potato. **Plant Physiology and Biochemistry,** v.94, p.73-85, 2015. https://doi.org/10.1016/j.plaphy.2015.05.010.

VOORRIPS, R.E. MapChart: Software for the Graphical Presentation of Linkage Maps and QTLs. **Journal of Heredity**, v.93, p.77-78, 2002. https://doi.org/10.1093/jhered/93.1.77

WANG, T.; YUE, J.-J.; WANG, X.-J.; XU, L.; LI, L.-B.; GU, X.-P. Genome-Wide Identification and Characterization of the Dof Gene Family in Moso Bamboo (*Phyllostachys heterocycla* var. *pubescens*). **Genes & Genomics**, v.38, n.733, 2016. https://doi.org/10.1007/s13258-016-0418-2.

WANG, P.; LI, J.; GAO, X.; ZHANG, D.; LI, A.; LIU, C. Genome-Wide Screening and Characterization of the Dof Gene Family in Physic Nut (*Jatropha curcas* L.). **International Journal of Molecular sciences**, v.19, n.6, p.1598, 2018. https://doi.org/10.3390/ijms19061598

WEI, Q.; WANG, W.; HU, T.; HU, H.; MAO, W.; ZHU, Q.; BAO, C. Genome-wide identification and characterization of Dof transcription factors in eggplant (*Solanum melongena* L.). **PeerJ**, n. 6e4481, 2018. https://doi.org/10.7717/peerj.4481/supp-1.

WEN, C. L.; CHENG, Q.; ZHAO, L.; MAO, A.; YANG, J.; YU, S.; WENG, Y.; XU, Y. Identification and characterisation of Dof transcription factors in the cucumber genome. **Scientific Reports,** v. 6, n.23072, 2016. https://doi.org/10.1038/srep23072.

WU, Z.; CHENG, J.; CUI, J.; XU, X.; LIANG, G.; LUO, X.; CHEN, X.; TANG, X.; HU, K.; QIN, C. Genome-wide identification and expression profile of dof transcription factor gene family in pepper (*Capsicum annuum* L.). **Frontiers in Plant Science**, v.7, n.574, 2016. https://doi.org/10.3389/fpls.2016.00574.

XU, X.; LI, F.; WANG, Y.; TANG, S.; DAI, Q.; ZHU, S.; LIU, T. Identification of Dof transcription factors in ramie (*Boehmeria nivea* L. Gaud) and their expression in response to different nitrogen treatments. **3 Biotechonoly**, v. 8, n. 12, p. 496, 2018 https://doi.org/10.1007/s13205-018-1512-8.

YANAGISAWA S. A novel DNA-binding domain that may form a single zinc finger motif. **Nucleic Acids Research** v. 23, p. 3403–3410, 1995 https://doi.org/10.1093/nar/23.17.3403.

YANAGISAWA S. Dof DNA-binding domains of plant transcription factors contribute to multiple protein-protein interactions. **European Journal of Biochemistry**, v. 250, p. 403–410, 1997. https://doi.org/10.1111/j.1432-1033.1997.0403a.x.

YANAGISAWA, S.; SHEEN J. Involvement of maize Dof zinc finger proteins in tissue specific and light-regulated gene expression. **The Plant Cell** v.10, n.1, p.75-89, 1998. https://doi.org/10.1105/tpc.10.1.75

YANG, Q.; CHEN, Q.; ZHU, Y.; LI, T. Identification of *MdDof* genes in apple and analysis of their response to biotic or abiotic stress. **Functional Plant Biology**, v. 45, n. 5, p. 528-541, 2018. https://doi.org/10.1071/FP17288.

ZHANG, L.; LIU, B.; ZHENG, G.; ZHANG, A.; LI, R. Genome-wide characterization of the *SiDof* gene family in foxtail millet (*Setaria italica*). **Biosystems**, v. 151, p. 27-33, 2017a. https://doi.org/10.1016/j.biosystems.2016.11.007.

ZHANG, Z.; YUAN, L.; LIU, X.; CHEN, X.; WANG, X. Evolution analysis of Dof transcription factor family and their expression in response to multiple abiotic stresses in *Malus domestica*. **Gene,** v.639, p.137-148, 2017b. https://doi.org/10.1016/j.gene.2017.09.039.

ZHANG, C.; MA, R.; XU, J.; YAN, J.; GUO, L.; SONG, J.; FENG, R.; YU, M. Genome-wide identification and classification of MYB superfamily genes in peach. **PLoS One** n. 13:e0199192, 2018. https://doi.org/10.1371/journal.pone.0199192.

Recebido para publicação em 17/12/2018
Revisado em 14/02/2019
Aceito em 14/02/2019