Genome wide identification and functional assignments of C$_2$H$_2$ Zinc-finger family transcription factors in Dichanthelium oligosanthes

Manisha Mahapatra, Bijayalaxmi Mahanty, Raj Kumar Joshi*

Department of Biotechnology, Rama Devi Women’s University, Vidya Vihar, Bhubaneswar-751022, Odisha, INDIA; Dr. Raj Kumar Joshi - Email: rkJoshi@rdwu.ac.in; *Corresponding author

Received April 15, 2019; Revised October 16, 2019; Accepted October 16, 2019; Published October 16, 2019

DOI: 10.6026/97320630015689

Abstract:

Transcription factors (TFs) are biological regulators of gene function in response to various internal and external stimuli. C$_2$H$_2$ zinc finger proteins (C$_2$H$_2$-ZFPs) are a large family of TFs that play crucial roles in plant growth and development, hormone signalling and response to biotic and abiotic stresses. While C$_2$H$_2$-ZFPs have been well characterized in many model and crop plants, they are yet to be ascertained in the evolutionarily important C$_3$ plant Dichanthelium oligosanthes (Heller’s rosette grass). In the present study, we report 32 C$_2$H$_2$-ZF genes (DoZFs) belonging to three different classes-Q type, C-type and Z-type based on structural elucidation and phylogenetic analysis. Sequence comparisons revealed paralogs within the DoZFs and orthologs among with rice ZF genes. Motif assignment showed the presence of the distinctive C$_2$H$_2$-ZF conserved domain “QALGGH” in these proteins. Cis-element analysis indicated that majority of the predicted C$_2$H$_2$-ZFPs are associated with hormone signalling and abiotic stress responses. Further, their role in nucleic acid binding and transcriptional regulation was also observed using predicted functional assignment. Thus, we report an overview of the C$_2$H$_2$-ZF gene family in D. oligosanthes that could serve as the basis for future experimental studies on isolation and functional implication of these genes in different biological mechanism of C$_3$ plants.

Keywords: Zinc fingers, C$_2$H$_2$-ZFPs, transcription factors, phylogenetic analysis, Dichanthelium oligosanthes

Background:

Transcription factors (TFs) are regulatory proteins which play critical role in altering the expression of genes associated with multiple cellular pathways related to growth, development and stress responses [1]. Among the various TFs, the Zinc-finger proteins (ZFPs) are the largest group of transcription regulators in plants [2]. ZFPs constitute a two stranded antiparallel beta sheet and a helix stabilized by zinc finger domains consisting of zinc ion surrounded by cysteine and histidine residues. Since the discovery of the first ZFP from Petunia, several zinc-finger TFs have been identified from myraids of plants and their involvement in different biological processes including growth, development, reproduction, photosynthesis and stress responses have been reported [2].

Among all the ZFP types, C$_2$H$_2$-ZFPs are the most widely distributed transcription factors in eukaryotes. These are characterized by the presence of a conserved motif X$_2$-Cys-X$_3$-Cys-X$_2$-His-X$_3$-His, where X represents the amino acids that act as the spacer between the cysteine and the histidine residues [3]. Experimental analyses have shown that C$_2$H$_2$-ZFPs are represented by 3% of all genes in mammals, 2.3% of all the genes in Drosophila and 0.8% of all genes in yeast [4]. Compared to other eukaryotes, the
Identification and characteristics of \(C_{2}H_{2}\)-ZFPs

\(C_{2}H_{2}\)-ZFPs are characterized by the presence of highly conserved QALGGH motif in the zincfinger helices and have long spacers with variable length and sequence between the zinc finger domains [2, 4]. Extensive identification and characterization of \(C_{2}H_{2}\)-ZFPs have been reported in plants including 179 from \textit{Arabidopsis} [5], 189 in rice [6], 124 in foxtail millet [7], 109 in \textit{Populus trichocarpa} [8] and 122 in durum wheat [9]. Accumulating evidences indicate that \(C_{2}H_{2}\)-ZFPs are critically associated with transcriptional regulation, RNA metabolism and protein-protein interactions [10, 11]. A wide number of plant \(C_{2}H_{2}\)-ZFPs have been functionally implicated in multiple physiological processes including floral organogenesis [12], growth initiation [13], biogenesis of non-coding RNAs [14], abiotic stress responses [15, 16], pathogen defence [17].

\textit{Dichanthelium oligosanthes}, also known as the Heller's rosette grass is a frost tolerant perennial wild penicoid grass species which utilizes the C3 pathway for carbon fixation and lacks Kranz anatomy [18]. Therefore, it can be used as a model species to understand the evolutionary developmental pattern of C4 photosynthesis when compared with important C4 relatives, including rice, wheat, and maize. The draft genome of \textit{D. oligosanthes} has been recently sequenced and a small suite of \(C_{2}H_{2}\)-ZFPs were identified [19]. While, extensive studies of \(C_{2}H_{2}\)-ZFPs and their association with biological and physiological mechanisms have been conducted in many plant species, no report is available from \textit{D. oligosanthes} so far. Therefore, it is important to perform a genome-wide identification and characterization of \(C_{2}H_{2}\)-ZF family of transcription factors to illuminate their molecular role in \textit{D. oligosanthes}. In the present study, we identified 32 \(C_{2}H_{2}\)-ZF genes from \textit{D. oligosanthes} utilizing varied bioinformatics tools. The structural organization of the identified genes including exon-intron arrangements, 5’/3’ untranslated regions (UTRs), conserved protein motifs and promoter \textit{cis}-elements were determined. Further, the identified proteins were analyzed for their phylogenetic relationship and orthology/paralogy within \textit{D. oligosanthes} as well as with other model plant species. Additionally, the functional characteristics of the identified \(C_{2}H_{2}\)-ZFPs were predicted using gene ontology (GO) analyses. These results will form the basis for future gene functional studies of \(C_{2}H_{2}\)-ZFPs in towards understanding physiological responses in \textit{D. oligosanthes}.

Materials & Methods

Identification and characteristics of \(C_{2}H_{2}\)-ZF gene family

The draft genome sequence of \textit{D. oligosanthes} (ASM163321v2) was downloaded from NCBI database (http://www.ncbi.nlm.nih.gov/). The hidden Markov model (HMM) profile of \(C_{2}H_{2}\)-ZF (PF00096) was downloaded from the Protein family (Pfam) database (http://pfam.xfam.org/) and subsequently used as a query in the HMMER database (https://www.ebi.ac.uk/Tools/hmmer) to search for \(C_{2}H_{2}\)-ZF proteins in \textit{D. oligosanthes}. The retrieved candidate protein sequences were further analyzed with the SMART (http://smart.embl-heidelberg.de/) database to confirm the presence of \(C_{2}H_{2}\)-ZF domain in the sequences. Specific properties of the deduced polypeptides including molecular weight, isoelectric points and hydropathy were calculated using the ExPaSy site (http://web.expasy.org/protparam/).

Sequence alignment and phylogenetic analysis

\(C_{2}H_{2}\)-ZF gene and protein sequences from model plant \textit{Arabidopsis} and rice were obtained from The \textit{Arabidopsis} Information Resource (TAIR, http://www.arabidopsis.org/index.jsp) and Rice Genome Browser (http://www.tigr.org/tigr-scripts/osa1web/gbrowse/rice) respectively. Multiple sequence alignment of the full length \(C_{2}H_{2}\)-ZF protein sequences from \textit{D. oligosanthes}, \textit{A. thaliana} and \textit{O. sativa} was performed using Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/) with default parameter and manually adjusted using BioEdit 7.1 software [20]. Phylogenetic analyses of the protein sequences were performed using Molecular Evolutionary Genetic Analysis (MEGA v 10.1) package [21]. A neighbourjoining (NJ) method with 1000 bootstrapping was performed to develop an unrooted phylogenetic tree.

Structural organization and identification of conserved motifs

The individual cDNA sequences of the \(C_{2}H_{2}\)-ZF genes and their corresponding genomic sequences were compared using the Gene Structures Display Server (GSDS 2.0; http://gds.sbi.pku.edu.cn/index.php) to generate the intron/exon organization. Motif structures of the predicted protein were analyzed using Multiple Expectation Maximization for motif Elicitation (MEME) tool [22] using the set parameters as follows: occurrence of motif repeats: any number, max number of motifs to be predicted: 20, and Min/Max motif width: 10/100.

Promoter \textit{cis}-element analysis and identification of paralogs and orthologs

Promoter sequences about 2Kb upstream of the translation start site for all the \(C_{2}H_{2}\)-ZF genes were obtained from the NCBI database. The \textit{cis}-acting regulatory elements were located and predicted from the putative \(C_{2}H_{2}\)-ZF promoter regions by using Plant-CARE [23]. All the cDNA sequences of the \(C_{2}H_{2}\)-ZF genes were compared amongst themselves (all-against-all) by performing BLASTn to identify the paralogous ZFs in \textit{D. oligosanthes}. After each round of
BLASTn, sequences showing ≥ 40% sequence similarity with at least 300bp sequence alignment were considered to be paralogous [24]. To predict the orthologs in rice, each of the rice C2H2-ZF sequences was used as a query to search against all DoZF sequences by using BLASTn. The BLASTn results showing the best hits with at least 300 bp region of alignment with a DoZF was considered to be an ortholog [24].

Sub-cellular localization and gene ontology (GO) analysis
The subcellular localization of C2H2-ZF proteins was predicted using the mGOASVM (Plant V2) server [25]. The functional grouping of C2H2-ZF sequences from D. oligosanthes and the annotation data were computed using the Blast2GO v3.0 [26] and cross verified using the DeepGO protein function prediction tool with the protein GO classes [27]. Blast2GO annotation associates genes or transcripts with GO terms classified into three categories: biological processes, molecular functions and cellular components.

Results & Discussion
The HMM profile of the C2H2-ZF domain (PF00096) was used as a query to search for C2H2-ZF genes of D. oligosanthes within the protein databases using HMMER software. A total of 57 C2H2-ZF genes were obtained. A recent study using similar approach identified 14 Squamosa promoter-binding protein-like (SPL) TFs in D. oligosanthes [28]. The candidate sequences thus obtained were analysed using the Simple Modular Architecture Research Tool (SMART; SM000355) and the Conserved Domain Database (CDD) to validate the presence of C2H2-ZFs. Finally, 32 C2H2-ZF genes were identified and names as DoZFP1 to DoZFP32 (C2H2ZFPs of D. oligosanthes). This number is quite less than those found in Arabidopsis, rice, foxtail millet and Populus [5-8]. Analysis of the peptide properties showed that DoZFPs had molecular masses ranging from 21133.18 Da (DoZF2) to 166234.58 Da (DoZF13). Likewise, the length of the amino acids in the encoded proteins of DoZFPs greatly varied between 196aa (DoZF3) to 15103aa (DoZF13). Additionally, the hydropathy plot obtained from Expasy protscale revealed that majority of the identified DoZFPs were basic in nature (data not shown). Also, 26 DoZFs were basic while the remaining 6 predicted proteins were found
acidic in nature. The details of the properties of the DoZFP nucleic acid and protein sequences are represented in Table 1.

Figure 3: Unrooted phylogenetic tree representing the relationship among C2H2-ZFPs of D. oligosanthes, rice and Arabidopsis. The protein sequences of C2H2-ZFPs were aligned with Clustal Omega and phylogenetics tree was constructed using the neighbor-joining method in MEGA 10.0. The Bootstrap value was 1,000 replicates.

Diversity of the gene structure, cis-regulatory elements and conservation of protein motifs is possible instrument for the evolution of gene families in plants [29]. The intron/exon organization of the DoZFPs was determined by comparing the coding sequence with their corresponding genomic DNA sequences using GSDS software. The number of exons varied from 1 (DoZF3, DoZF6, DoZF7, DoZF11, DoZF12, DoZF19, DoZF20, DoZF25, DoZF29, DoZF30) to 7 (DoZF13) with 13 DoZFPs composed of three or more exons (Figure 1). In contrast, 14 DoZFPs had two or more than two introns while 10 DoZFPs had no introns. Similar organization of introns/exon organization has been reported for C2H2-ZFPs in Populus and rice [6, 8]. Cis regulatory elements are key factors in controlling the transcriptional regulation of genes [30]. Therefore, the interaction between key transcription factors and specific cis-element is crucial in plants’ response to phyto hormones as well as biotic and abiotic stresses [31]. Promoter sequence 2000 bp upstream of the translation initiation site in the 32 DoZFP genes were examined for the presence of cis-element using the PlantCARE database. Results revealed that 1 to 11 TATA box element and 1 to 8 CAAT box elements were found in the promoter regions of 32 DoZF genes. In addition, DoZFP gene promoters contains multiple cis regulatory elements responsive to phyto hormone and stress signalling, including ABRE (Abscisic acid responsive element), TCA (Salicylic acid responsive element), MYB and MYC regions, CGTCA (Methyl jasmonate responsive element), ERE (ethylene responsive element), G-box (light responsive element), and W-box (WRKY binding draught responsive element). Similar cis-elements have been reported in the promoters of C2H2-ZFPs in Arabidopsis thaliana [5] and further in-depth analysis of these regulatory regions would be needed to validate their roles in stress responsiveness of D. oligosanthes.

To further reveal the diversification of C2H2-ZFPs in D. oligosanthes, conserved protein motif sequences were predicted using MEME web server [22]. A total of 15 distinct structural motifs were predicted (Figure 2; Table 2). Motif 1, 2, 7 and 11 represented distinctive conserved regions of the C2H2-ZFPs. Motif 7 and 11 constituted the plant specific conserved domain “QALGGH” and were found in 11DoZFPs that were identified as Q-type. Among the Q-types, DoZF29 have a modified conserved sequence “ALGGH” and classified as M-typeC2H2-ZFP. Likewise, 15DoZFPs consisted of Motif 1 with conserved sequence “CGKGFQRDQNLQLHRRGH” and motif 2 with conserved sequence “CGKGFKRDNLMHLRRGH”, the characteristic features of the Z-type C2H2-ZFPs. The remaining 6 DoZFPs (DoZF4, DoZF9, DoZF13, DoZF15, DoZF25 and DoZF32) did not contain any known conserved motif in the ZF region and were categorized as C-type C2H2-ZFPs. Additionally, 11 unidentified conserved motifs were also identified that were randomly placed across all the DoZFPs. Taken together, our results suggest that functionally divergent group of C2H2-ZFPs are associated in numerous plant developmental and physiological processes of D. oligosanthes.
To explore the evolutionary association of the identified DoZFPs, full length protein sequences of 32 DoZFPs, 15 AtZFPs and 29 OsZFPs were used to construct a neighbor-joining tree (Figure 3). The resulting tree clustered all the C2H2-ZFPs into two groups- I and II similar to previous grouping of C2H2-ZFPs reported in rice [6] and Arabidopsis [5]. Group I consisted of 40 proteins including 15 Q-type DoZFPs and 2 C-type DoZFPs. Likewise, group II categorized 36 proteins including 15Z-type DoZFPs. Previous reports have shown that C-type ZFs are grouped with Z-type as well as Q-type ZFs [8]. Nevertheless, our results support the hypothesis that Q-type plant specific ZFs have evolved from C-type ZFs through conservation of the “QALGGH” sequence [6]. Further, assessment of paralogy among DoZFs and orthology of DoZFs with OsZFs revealed that 12 DoZFPs were paralogous with an average of 90% similarity while 21 were orthologous (68% similarity) with OsZFs (Table 3). The genomic expansion and evolutionary divergence of a species depends on genetic duplication of functional traits [32]. Similar to C2H2-ZFPs, several TFs in different plants including NAC, WRKY and HD-Zip exhibit gene duplication as an adaptive mechanism towards dynamic environmental conditions [33, 34].

| Name   | Accession no. | Gene Length (bp) | Protein length (aa) | pf | Mw  | No. of Exons | Nature | Location | Molecular Function | Biological Process | Cellular Component |
|--------|---------------|------------------|---------------------|-----|-----|-------------|--------|----------|-------------------|-------------------|-------------------|
| DoZF1  | A0A1E5Y3C44   | 1905             | 778                 | 4.75| 925 | 2           | Basic  | Nucleus  | DNA binding       | Regulation of DNA transcription; RNA biosynthesis | cell part          |
| DoZF2  | A0A1E5Y3C44   | 1905             | 778                 | 4.75| 925 | 2           | Basic  | Nucleus  | DNA binding       | RNA biosynthesis   | cell part          |
| DoZF3  | A0A1E5Y3C44   | 1905             | 778                 | 4.75| 925 | 2           | Basic  | Nucleus  | DNA binding       | RNA biosynthesis   | cell part          |
| DoZF4  | A0A1E5Y3C44   | 1905             | 778                 | 4.75| 925 | 2           | Basic  | Nucleus  | DNA binding       | RNA biosynthesis   | cell part          |
| DoZF5  | A0A1E5Y3C44   | 1905             | 778                 | 4.75| 925 | 2           | Basic  | Nucleus  | DNA binding       | RNA biosynthesis   | cell part          |
| DoZF6  | A0A1E5Y3C44   | 1905             | 778                 | 4.75| 925 | 2           | Basic  | Nucleus  | DNA binding       | RNA biosynthesis   | cell part          |
| DoZF7  | A0A1E5Y3C44   | 1905             | 778                 | 4.75| 925 | 2           | Basic  | Nucleus  | DNA binding       | RNA biosynthesis   | cell part          |
| DoZF8  | A0A1E5Y3C44   | 1905             | 778                 | 4.75| 925 | 2           | Basic  | Nucleus  | DNA binding       | RNA biosynthesis   | cell part          |
| DoZF9  | A0A1E5Y3C44   | 1905             | 778                 | 4.75| 925 | 2           | Basic  | Nucleus  | DNA binding       | RNA biosynthesis   | cell part          |
| DoZF10 | A0A1E5Y3C44   | 1905             | 778                 | 4.75| 925 | 2           | Basic  | Nucleus  | DNA binding       | RNA biosynthesis   | cell part          |
| DoZF11 | A0A1E5Y3C44   | 1905             | 778                 | 4.75| 925 | 2           | Basic  | Nucleus  | DNA binding       | RNA biosynthesis   | cell part          |
| DoZF12 | A0A1E5Y3C44   | 1905             | 778                 | 4.75| 925 | 2           | Basic  | Nucleus  | DNA binding       | RNA biosynthesis   | cell part          |
| DoZF13 | A0A1E5Y3C44   | 1905             | 778                 | 4.75| 925 | 2           | Basic  | Nucleus  | DNA binding       | RNA biosynthesis   | cell part          |
| DoZF14 | A0A1E5Y3C44   | 1905             | 778                 | 4.75| 925 | 2           | Basic  | Nucleus  | DNA binding       | RNA biosynthesis   | cell part          |
| DoZF15 | A0A1E5Y3C44   | 1905             | 778                 | 4.75| 925 | 2           | Basic  | Nucleus  | DNA binding       | RNA biosynthesis   | cell part          |
| DoZF16 | A0A1E5Y3C44   | 1905             | 778                 | 4.75| 925 | 2           | Basic  | Nucleus  | DNA binding       | RNA biosynthesis   | cell part          |
| DoZF17 | A0A1E5Y3C44   | 1905             | 778                 | 4.75| 925 | 2           | Basic  | Nucleus  | DNA binding       | RNA biosynthesis   | cell part          |
| DoZF18 | A0A1E5Y3C44   | 1905             | 778                 | 4.75| 925 | 2           | Basic  | Nucleus  | DNA binding       | RNA biosynthesis   | cell part          |
| DoZF19 | A0A1E5Y3C44   | 1905             | 778                 | 4.75| 925 | 2           | Basic  | Nucleus  | DNA binding       | RNA biosynthesis   | cell part          |
| DoZF20 | A0A1E5Y3C44   | 1905             | 778                 | 4.75| 925 | 2           | Basic  | Nucleus  | DNA binding       | RNA biosynthesis   | cell part          |
| DoZF21 | A0A1E5Y3C44   | 1905             | 778                 | 4.75| 925 | 2           | Basic  | Nucleus  | DNA binding       | RNA biosynthesis   | cell part          |
| DoZF22 | A0A1E5Y3C44   | 1905             | 778                 | 4.75| 925 | 2           | Basic  | Nucleus  | DNA binding       | RNA biosynthesis   | cell part          |
| DoZF23 | A0A1E5Y3C44   | 1905             | 778                 | 4.75| 925 | 2           | Basic  | Nucleus  | DNA binding       | RNA biosynthesis   | cell part          |
| DoZF24 | A0A1E5Y3C44   | 1905             | 778                 | 4.75| 925 | 2           | Basic  | Nucleus  | DNA binding       | RNA biosynthesis   | cell part          |
| DoZF25 | A0A1E5Y3C44   | 1905             | 778                 | 4.75| 925 | 2           | Basic  | Nucleus  | DNA binding       | RNA biosynthesis   | cell part          |
| DoZF26 | A0A1E5Y3C44   | 1905             | 778                 | 4.75| 925 | 2           | Basic  | Nucleus  | DNA binding       | RNA biosynthesis   | cell part          |
| DoZF27 | A0A1E5Y3C44   | 1905             | 778                 | 4.75| 925 | 2           | Basic  | Nucleus  | DNA binding       | RNA biosynthesis   | cell part          |
| DoZF28 | A0A1E5Y3C44   | 1905             | 778                 | 4.75| 925 | 2           | Basic  | Nucleus  | DNA binding       | RNA biosynthesis   | cell part          |
| DoZF29 | A0A1E5Y3C44   | 1905             | 778                 | 4.75| 925 | 2           | Basic  | Nucleus  | DNA binding       | RNA biosynthesis   | cell part          |
| DoZF30 | A0A1E5Y3C44   | 1905             | 778                 | 4.75| 925 | 2           | Basic  | Nucleus  | DNA binding       | RNA biosynthesis   | cell part          |
| DoZF31 | A0A1E5Y3C44   | 1905             | 778                 | 4.75| 925 | 2           | Basic  | Nucleus  | DNA binding       | RNA biosynthesis   | cell part          |
| DoZF32 | A0A1E5Y3C44   | 1905             | 778                 | 4.75| 925 | 2           | Basic  | Nucleus  | DNA binding       | RNA biosynthesis   | cell part          |

Gene ontology (GO) term analyses of the predicted proteins using Blast2GO v3.0 categorized them into cellular components, molecular functions and biological processes (Table 1). Under the biological process categories, all the DoZFPs represented regulation of DNA transcription (GO: 190306) and RNA biosynthesis (GO: 2001141). Similarly, cellular component prediction showed that 31 DoZFPs were represented by ‘cell part’ (GO: 004464) while only DoZF4 was represented as ‘intracellular part’ (GO: 0044424). Within the ‘molecular function category’, 31DoZFPs were represented by GO terms ‘DNA binding (GO: 0003677)’ and nucleic acid binding...
(GO: 0003676)‘ suggesting their primary molecular role as interaction modules that binds to DNA, RNA and proteins [35]. In addition, DoZF4 represented transporter activity (GO: 0022891).

Table 3: Paralogs and orthologs of C_H-ZF gene pairs in D. oligosanthes and Oryza sativa

| PARALOGS within DoZFs | ORTHOLOGS of DoZFs in Oryza sativa |
|-----------------------|-----------------------------------|
| DoZF4/DoZF5           | DoZF11/ LOC_Os10g28330            |
| DoZF23/DoZF22         | DoZF12/ LOC_Os04g0290, LOC_Os10g2220 |
| DoZF18/DoZF22         | DoZF12/ LOC_Os04g0290, LOC_Os10g2220 |
| DoZF14/DoZF21         | DoZF13/ LOC_Os06g0150            |
| DoZF16/DoZF23         | DoZF16/ LOC_Os04g0290, LOC_Os10g2220 |
| DoZF10/DoZF12         | DoZF17/ LOC_Os04g0290, LOC_Os10g2220 |
| DoZF11/DoZF12         | DoZF18/ LOC_Os04g0290, LOC_Os10g2220 |
| DoZF13/DoZF14         | DoZF19/ LOC_Os04g0290, LOC_Os10g2220 |
| DoZF14/DoZF21         | DoZF20/ LOC_Os04g0290, LOC_Os10g2220 |
| DoZF11/DoZF12         | DoZF21/ LOC_Os04g0290, LOC_Os10g2220 |
| DoZF13/DoZF14         | DoZF22/ LOC_Os04g0290, LOC_Os10g2220 |
| DoZF14/DoZF21         | DoZF23/ LOC_Os04g0290, LOC_Os10g2220 |
| DoZF11/DoZF12         | DoZF24/ LOC_Os04g0290, LOC_Os10g2220 |
| DoZF13/DoZF14         | DoZF25/ LOC_Os04g0290, LOC_Os10g2220 |
| DoZF14/DoZF21         | DoZF26/ LOC_Os04g0290, LOC_Os10g2220 |

**Conclusion:**

A comprehensive genome wide analysis including phylogenetic relationships, structural prediction, conserved motif analysis and gene functions of the C_H-ZF gene family in _D. oligosanthes_ were performed. Our analysis identified 32 C_H-ZF genes in _D. oligosanthes_. Phylogenetic analysis grouped the DoZFs into three clusters similar to their orthologs in _Arabidopsis_ and rice. Structural and motif elucidation demonstrated the presence of multiple conserved domains “QALGQH” suggesting their implication in DNA binding and transcription factor activity. Further, the cis-element analysis of the DoZFs showed their involvement in hormone signalling and stress responses. These data form the basis for functional characterization of suitable candidate genes to untangle their different roles in biological regulation.

**Acknowledgment:**

BM is grateful to Dept. of Biotechnology (DBT), Govt. of India for financial support in form of Junior Research Fellowship. The authors are thankful to the Head, Dept. of Biotechnology, Rama Devi Women’s University, Vidya Vihar, Bhubaneswar for her guidance and support. The open access for the article is partly sponsored by Biomedical Informatics (P) Ltd, India.

**References:**

[1] Qin Z et al. _BMC Genomics_. 2018 19:193. [PMID: 29703141]
[2] Ciftci-Yilmaz S & Mittler R. _Cell Mol Life Sci_. 2008 65:1150 [PMID: 18193167]
[3] Huang J et al. _Yi Chuan_. 2004 26:414. [PMID: 15640031]
[4] Chung HR et al. _EMBO Rep_. 2002 3:1158. [PMID: 12446571]
[5] Engblight CC et al. _BMC Genomics_. 2004 5:39. [PMID: 15236668]
[6] Agrawal et al. _Plant Mol Biol_. 2007 65:467. [PMID: 17610133]
[7] Muthumalarasan M et al. _Func. Int. Genom_. 2014 14:431. [PMID: 24915771]
[8] Liu Q et al. _PLos ONE_. 2015 10: e0134753. [PMID: 26237514]
[9] Faraji S et al. _Biometals_. 2018 31:1019. [PMID: 30288657]
[10] Gamsjaeger R et al. _Trends Biochem Sci_. 2007 32:63. [PMID: 17210253]
[11] Brayer KJ & Segal DJ. _Cell Biochem Biophys_. 2008 50:111. [PMID: 18253864]
[12] Dinneny JR et al. _Development_. 2006 133:2285. [PMID: 16554365]
[13] Grigg SP et al. _Nature_. 2005 437:1022. [PMID: 16222298]
[14] Yang L et al. _Plant J_. 2006 47:841. [PMID: 16889646]
[15] Huang J et al. _Biochem Biophs Res Co_. 2009 389:556. [PMID: 19751706]
[16] Wang K et al. _Physiol Plant_. 2019 165:690. [PMID: 29572849]
[17] Kim SH et al. _Plant Mol Biol_. 2004 55:883. [PMID: 15604723]
[18] Brautigam A et al. _J Exp Bot_. 2014 65:3579. [PMID: 24642845]
[19] Studer AJ et al. _Genome Biol_. 2016 17:223. [PMID: 27793170]
[20] Hall TA. _Nucleic Acids Symposium Series_. 1999 41:95.
[21] Kumar S et al. _Mol Biol Evol_. 2018 35:1547. [PMID: 29722887]
[22] Bailey TL et al. _Nucleic Acids Res_. 2006 34:369. [PMID: 16854028]
[23] Lescoat M et al. _Nucleic Acids Res_. 2002 30:325. [PMID: 11752327]
[24] Blanc G and Wolfe KH. _Plant Cell_. 2004 16:1667. [PMID: 15208399]
[25] Wan S et al. _BMC Bioinformatics_. 2012, 13: 290. [PMID: 23130999]
[26] Conesa A et al. _Int J Plant Genomics_. 2008 2008:619832. [PMID: 18483572]
[27] Kulmanov M et al. _Bioinformatics_. 2018 34:660. [PMID: 29028931]
[28] Nanda S & Hussain S. _Bioinformation_. 2019 15:165. [PMID: 31354191]
[29] Hu R et al. _Plant Biol_. 2010 16:145. [PMID: 20630103]
[30] Nakashima et al. _Front Plant Sci_. 2014 5:170. [PMID: 24904597]
[31] Santrner A & Estelle M. _Nature_. 2009 459:1071. [PMID: 19553990]
[32] Vision TJ et al. _Science_. 2000 290:2114. [PMID: 11118139]
[33] Li W et al. _Gene_. 2019 697:103. [PMID: 30776460]
[34] Liu Y et al. _Plant Biotech J_. 2019 17:1770. [PMID: 30801865]
[35] Klug A. _Ann. Rev. Biochem_. 2010 79:213. [PMID: 20192761]
