HNUE JOURNAL OF SCIENCE
Natural Sciences 2024, Volume 69, Issue 1, pp. 122-135
This paper is available online at http://stdb.hnue.edu.vn
DOI: 10.18173/2354-1059.2024-0012

PREDICTION OF DUPLICATION EVENTS IN THE PLATZ TRANSCRIPTION FACTOR IN CASSAVA (*Manihot esculenta*) SUGGESTS THE VARIATION OF THEIR FUNCTIONS IN DROUGHT CONDITION

Hoang Minh Chinh[1,2], Dong Huy Gioi[2], Chu Duc Ha[1], Tran Van Tien[3], La Viet Hong[4] and Tran Thi Thanh Huyen[5,*]

[1]*Faculty of Agricultural Technology, University of Engineering and Technology, Hanoi city, Vietnam*

[2]*Faculty of Biotechnology, Vietnam National University of Agriculture, Hanoi city, Vietnam*

[3]*Faculty of Rural Management, National Academy of Public Administration, Hanoi city, Vietnam*

[4]*Institute of Scientific Research and Application, Hanoi Pedagogical University 2, Hanoi city, Vietnam*

[5]*Faculty of Biology, Hanoi National University of Education, Hanoi city, Vietnam*

[*]Corresponding author: Tran Thi Thanh Huyen, e-mail: tranthanhhuyen@hnue.edu.vn

Received January 17, 2024. Revised March 16, 2024. Accepted March 23, 2024.

Abstract. The plant A/T-rich protein and zinc-binding protein (PLATZ) family has been regarded as one of the important plant-specific transcription factors that are involved in various biological processes during evolution. Unfortunately, the expansion of this gene family in cassava (*Manihot esculenta*) is hardly recognized. This recent work aims to explain the evolution of the *MePLATZ* gene family by using various bioinformatics tools. Based on the similarity, a total of eight duplicated *MePLATZ* genes, including seven duplicated pairs and one pair of three duplicated genes have been predicted in the *MePLATZ* gene family in cassava. Among them, segmental and tandem duplication events were noted to play a crucial role in the expansion of the *MePLATZ* gene family. We found that the majority of members of the *MePLATZ* genes contained three or four exons, while at least 10 conserved motifs have been found in the full-length protein sequences. Next, the MePLATZ family could be categorized into seven different groups similar to those described in the PLATZ family in other higher plant species. Interestingly, the expression levels of 17 duplicated *MePLATZ* genes in leaf samples under drought conditions suggested the hypothesis of the functional conservation, redundancy, and divergence that occurred in this family. Taken together, our study could provide a foundation to get insight into the *MePLATZ* gene family in cassava.

Keywords: PLATZ, categorization, structure, expression profile, duplication event, cassava.

1. Introduction

Cassava (*Manihot esculenta*), an essential staple crop, is critical to agriculture and food security, making it a valuable asset to worldwide societies [1], [2]. This drought-tolerant tube crop from South America has spread across continents to become a common sight in a variety of tropical and subtropical places throughout the world [2]. Cassava, known for its flexibility, is a critical source of nutrition for millions of people, giving vital carbohydrates and essential elements [3]. Its resistance to severe stress conditions and capacity to thrive in an array of soil types add to its agricultural significance. Furthermore, cassava's use extends beyond its nutritional value because it plays an important role in the industrial sector, functioning as an integral component in several food products and other commodities [3], [4]. Despite its major benefit to food security, cassava confronts hurdles in terms of post-harvest losses, demanding continued research and technological developments to fully realize its potential [2]. Cassava's complex nature and adaptability illustrate its tremendous impact on global agriculture and underline its continuous significance in addressing our times' growing concerns. Thus, understanding the growth and development of cassava plants under adverse environmental conditions at the molecular level will be important.

The plant A/T-rich protein and zinc-binding protein (PLATZ) transcription factors (TFs) represent a distinct group of regulatory proteins in the plant kingdom, notable for both their unique structural characteristics and their pivotal roles in plant development and stress responses [5], [6]. From a structural standpoint, these plant-specific TFs are defined by the presence of a zinc finger motif within their DNA-binding domain [7]. This motif facilitates specific binding to AT-rich sequences in the plant genome, a feature integral to the functional capacity of these proteins [7], [8]. Functionally, PLATZ proteins are implicated in a myriad of plant physiological processes. They play critical roles in regulating plant growth, orchestrating the development of various plant organs, and mediating plant responses to environmental stressors. This regulatory function is achieved through their capacity to modulate gene expression, either by activating or repressing the transcription of specific target genes [9]. Consequently, PLATZ TFs are instrumental in shaping plant developmental pathways and enabling adaptive responses to fluctuating environmental conditions [5], [9], [10]. In our recent work, a total of 20 members of the PLATZ TFs, namely MePLATZ, have been identified and annotated in the cassava assemblies [11]. The expression levels of this multiple-gene family in major organs/tissues during the growth and development of cassava plants have been well-characterized [11]. However, little is known about the evolution of the *MePLATZ* gene family in cassava. Recently, the shreds of evidence of the expansion of the *PLATZ* gene family in many plant species, such as apple (*Malus domestica*) [12] and several *Malus* spp. [13], Chinese cabbage (*Brassica rapa*) [14], ginkgo (*Ginkgo biloba*) [15], and wheat (*Triticum aestivum*) [16] were provided. Thus, it could be a great platform to get insight into the evolution of the *MePLATZ* gene family in cassava.

This current work aimed to explain the evolution of the *MePLATZ* gene family in cassava. We first predicted the gene duplication events that occurred in the *MePLATZ* gene family. Structural analysis, including gene organization and motif enrichment, was performed. Next, we constructed an unrooted phylogenetic tree to categorize the

MePLATZ proteins. Finally, the expression patterns of genes encoding several *MePLATZ* were re-analyzed.

2. Content

2.1. Materials and methods

2.1.1. Materials

The newest cassava assembly (NCBI RefSeq assembly: GCF_001659605.2) obtained from the previous work [17] was downloaded from the Phytozome [18] and NCBI sources.

Transcriptome atlas (GEO accession: GSE98537) obtained in leaf samples under drought conditions [19] was obtained in NCBI Gene Expression Omnibus [20].

Twenty well-characterized PLATZ proteins in cassava reported in the previous work [11] were explored to obtain the full-length protein sequences, coding DNA sequences (CDSs), and genomic DNA sequences (gDNAs) for *in silico* analysis.

2.1.2. Methods

Chromosomal distribution of genes: The annotation of each *MePLATZ* gene obtained in previous work [11] was used to retrieve its location. In particular, the gene identifier of each gene was searched against the cassava genome [17] in the Phytozome [18] and NCBI sources. The physical locations of the *MePLATZ* genes were then illustrated using the Adobe Illustrator software.

Prediction of the gene duplication: The duplicated *MePLATZ* genes were predicted as previously described [21]. In particular, all CDSs of 20 *MePLATZ* genes collected in the previous work [11] were used for multiple sequence alignment using the ClustalX software [22]. The similarity score was calculated by using the BioEDIT software [23]. Duplicated *MePLATZ* genes were defined as their corresponding CDSs may share a similarity of > 70% [21].

Estimation of the Ka/Ks value: The number of nonsynonymous substitutions per non-synonymous site (Ka) and the number of synonymous substitutions per synonymous site (Ks) of each duplicated *MePLATZ* pair were calculated as previously described [21]. Briefly, aligned CDSs of each duplicated *MePLATZ* gene pair were subjected to the DNASp software [24]. When the Ka/Ks ratio is larger than one, it shows positive selection; when it is less than one, it represents purifying or stabilizing selection; and when it is exactly one, it indicates neutral selection [21].

Construction of phylogenetic tree: The phylogenetic tree of the MePLATZ family has been generated as previously described [21]. In particular, full-length amino acid sequences of the MePLATZ proteins found in the previous work [11] were used to analyze in the Molecular Evolutionary Genetics Analysis software [25]. The Neighbor-Joining algorithm was applied to construct the phylogenetic tree with bootstrap values of 10,000. All results were then illustrated using the Adobe Illustrator software.

Analysis of gene structure: The exon/intron organizations of the *MePLATZ* genes were explored as previously described [21]. Briefly, the CDS and gDNA sequences of all *MePLATZ* genes obtained in the previous work [11] were analyzed using the Gene

Structure Display Server tool [26]. The arrangement of the *MePLATZ* genes was exactly followed by the order in the phylogenetic tree. All gene structures were then illustrated using the Adobe Illustrator software.

Analysis of conserved motifs: The conserved regions of the MePLATZ proteins were analyzed as previously described [21]. In particular, the full-length amino acid sequences of the MePLATZ proteins were subjected to the Multiple Em for Motif Elicitation tool [27]. The minimum width and maximum width of motifs were six to 50 amino acid residues and the cut-off value was < 1e-10 [21].

Analysis of gene expression: The expression profiles of the *MePLATZ* genes were analyzed using the NCBI Gene Expression Omnibus [20]. Based on the previous transcriptome database (GEO accession: GSE98537) [19], the expression levels of the *MePLATZ* genes in treated leaf samples were analyzed. The fold-change values of ≥ 2.00 and ≤ -2.00 indicated the up-regulated and down-regulated genes, respectively.

2.2. Results and Discussion

2.2.1. Chromosomal localization and prediction of duplication events of the PLATZ transcription factor family in cassava

To identify the physical distribution of the *MePLATZ* genes in the chromosomes of cassava plants, a gene identifier was searched against the cassava genome. As a result, Figure 1 illustrates the chromosomal localization of the *MePLATZ* genes. As expected, all *MePLATZ* genes were found to be randomly located in the whole 18 chromosomes of the cassava genome. In particular, chromosomes 01, 03, 05, 08, 09, and 15 contain two members of the *MePLATZ* family. It has been realized that chromosomes 06, 11, 14, 16, and 18 had only one *MePLATZ* gene each, including *Manes.06G043200*, *Manes.11G153600*, *Manes.14G120600*, *Manes.16G052100* and *Manes.18G000750*, respectively. Meanwhile, chromosome 17 contains the highest members of the *MePLATZ* family, including *Manes.17G083200*, *Manes.17G085975* and *Manes.17G011500*. Additionally, no *MePLATZ* genes have been found in chromosomes 02, 04, 07, 10, 12 and 13.

Previously, the *PLATZ* genes were also localized in the genome of higher plant species with uneven rates [12, 14, 15, 16]. For example, whole 17 *PLATZ* genes were found on 14 different chromosomes in the genome of an apple [12]. Among them, chromosomes 02, 06, and 16 have the most *PLATZ* genes, with two each, whereas chromosomes 01, 03, 05, 07, 10, 11, 12, 13, 14, 15, and 17 each had one [12]. Next, the *PLATZ* genes in Chinese cabbage were found on eight of the 10 chromosomes in an unequal distribution [14]. Specifically, chromosome 09 had six *PLATZ* genes, which were followed by chromosomes 07 (five *PLATZ* genes) and 08 (four *PLATZ* genes) [14]. Chromosomes 02, 03, 04, and 06 each contained two *PLATZ* genes, whereas chromosome 01 only had one *PLATZ* gene. No *PLATZ* genes were found in chromosomes 05 and 10 [14]. In ginkgo, 11 *PLATZ* genes were scattered irregularly across six (out of 12) chromosomes [15]. Number of *PLATZ* genes was greatest (three *PLATZ* genes) on chromosome 3, while chromosomes 2, 6, and 10 each contained two *PLATZ* genes [15]. Chromosomes 7 and 9 each had one *PLATZ* gene, while two remaining *PLATZ* genes were not annotated in the current assembly of ginkgo [15]. Except for chromosomes 4A,

4B, 4D, 5A, 5B, and 5D, all identified 62 *PLATZ* genes were discovered to be unequally distributed on 15 chromosomes in wheat [16]. A total of 40 (out of 62) *PLATZ* genes were found on chromosomes 2A, 2B, 2D, 6A, 6B, and 6D, while chromosomes 1A, 1B, and 1D each contained only two *PLATZ* genes [16].

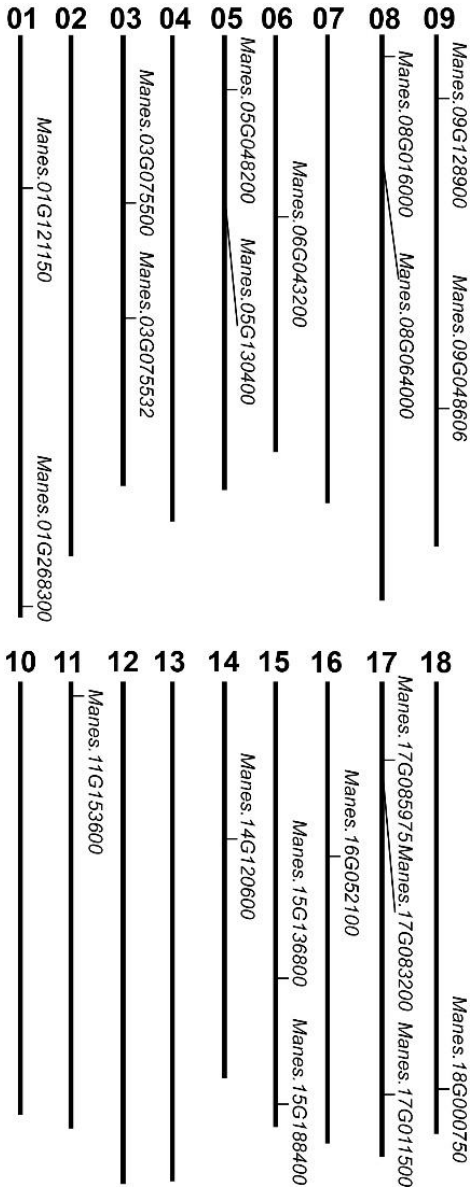

Figure 1. The physical location of the MePLATZ genes in the genome of cassava

Next, to explain the expansion of the *PLATZ* TFs, we performed a prediction of duplication events that occurred in this multiple-gene family. As a result, a total of eight duplication events, including seven gene pairs and one pair of three duplicated genes, have been found in the *PLATZ* family (Table 1). The similarities of these duplicated *MePLATZ* genes varied from 78.4 (*Manes.08G016000* and *Manes.09G048606*) to 98.4%

(*Manes.03G075500* and *Manes.03G075532*). Based on the chromosomal localization, it has been recognized that seven (out of eight) duplicated pairs have occurred from the segmental duplication. In particular, eight segmental duplication pairs included *Manes.05G130400* (on chromosome 05) and *Manes.18G000750* (on chromosome 18), *Manes.08G016000* (on chromosome 08), and *Manes.09G048606* (on chromosome 09), *Manes.08G064000* (on chromosome 08) and *Manes.09G128900* (on chromosome 09), *Manes.15G188400* (on chromosome 15) and *Manes.17G011500* (on chromosome 17), *Manes.01G268300* (on chromosome 01) and *Manes.05G048200* (on chromosome 05), *Manes.03G075500* (on chromosome 03) and *Manes.16G052100* (on chromosome 16), *Manes.03G075532* (on chromosome 03) and *Manes.16G052100* (on chromosome 16), *Manes.06G043200* (on chromosome 06) and *Manes.14G120600* (on chromosome 14), *Manes.15G136800* (on chromosome 15) and *Manes.17G085975* (on chromosome 17). Only one tandem duplication event was found to be *Manes.03G075500* and *Manes.03G075532* (on chromosome 03).

Table 1. Duplication events of the MePLATZ genes in cassava

No.	Duplicated *MePLATZ* pairs	Duplication events	Similarity	Ka/Ks
1	*Manes.05G130400*	Segmental	85.6	0.97
	Manes.18G000750			
2	*Manes.08G016000*	Segmental	78.4	1.10
	Manes.09G048606			
3	*Manes.08G064000*	Segmental	84.7	0.65
	Manes.09G128900			
4	*Manes.15G188400*	Segmental	81.9	1.00
	Manes.17G011500			
5	*Manes.01G268300*	Segmental	88.2	1.38
	Manes.05G048200			
6	*Manes.03G075500*	Tandem	98.4	0.32
	Manes.03G075532	Segmental	85.1	0.58
	Manes.16G052100	Segmental	84.5	0.58
7	*Manes.06G043200*	Segmental	89.7	0.90
	Manes.14G120600			
8	*Manes.15G136800*	Segmental	79.5	0.70
	Manes.17G085975			

Previously, the segmental and tandem duplication events could be explained as the main reasons for the expansion of the *PLATZ* gene families of higher plant species. For example, a total of 20 duplicated pairs were recorded in the *PLATZ* gene family in Chinese cabbage [14]. All duplicated *PLATZ* genes were localized on the different chromosomes, clearly indicating that they were segmentally duplicated paralogous genes [14]. At least nine *PLATZ* pairs of segmental duplication events were found in the chromosomes of the *M. domestica* cv. Gala genome [13]. Similarly, eight segmental duplication pairs were found in the *PLATZ* gene families in *M. domestica* cv. Hanfu and

M. sieversii, whereas only one tandem duplication pair occurred in the *PLATZ* gene family in *M. sieversii* [13].

To understand the selection pressures acting on this gene family, we estimated the Ka/Ks ratio for 8 duplicated gene pairs that occurred in the *MePLATZ* gene family in cassava. As provided in Table 1, the Ka/Ks ratios of the duplicated *MePLATZ* gene pairs varied from 0.32 (*Manes.03G075500* and *Manes.03G075532*) to 1.38 (*Manes.01G268300* and *Manes.05G048200*). In particular, five (out of eight) duplication events, including four gene pairs and one pair of three duplicated genes, exhibited Ks/Ka values of less than 1.00. It indicated that these duplicated *MePLATZ* genes are most likely performing a critical function, alterations to its amino acid sequence (nonsynonymous modifications) are being rejected (negative selection). In contrast, the two remaining gene pairs, including *Manes.08G016000* and *Manes.09G048606*, and *Manes.01G268300* and *Manes.05G048200* had Ka/Ks ratios of more than 1.00, suggesting accelerated evolution with positive selection. This finding suggested that the nonsynonymous alterations that occurred in these duplicated *MePLATZ* genes are preferred they are more likely to provide an adaptive advantage to cassava plants during the evolution process. Additionally, the Ka/Ks rate of one gene pair (*Manes.15G188400* and *Manes.17G011500*) of 1 revealed that two *MePLATZ* genes are drifting neutrally (neutral selection). It signifies that the rate of nonsynonymous substitutions is equal to the rate of synonymous substitutions, implying that both copies of the *MePLATZ* gene are evolving without any functional limitations, or that the mutations that occur have no impact on the cassava's fitness.

2.2.1. Categorization and structural analysis of the PLATZ transcription factor family in cassava

To get insight into the evolution of the MePLATZ family in cassava, we constructed a Neighbor-Joining phylogenetic tree of all full-length amino acids. As a result, an unrooted phylogenetic tree of 20 members in the MePLATZ family has been generated and described in Figure 2. We found that the MePLATZ family could be clearly classified into seven groups, namely A, B, C, D, E, F and G. Specifically, groups A, B and F contained only two members of the MePLATZ family each, including Manes.01G268300 and Manes.05G048200, Manes.06G043200 and Manes.14G120600, and Manes.05G130400 and Manes.18G000750. Next, three (Manes.03G075532, Manes.03G075500, and Manes.16G052100) and three (Manes.08G016000, Manes.09G048606 and Manes.11G153600) members of the MePLATZ family were clustered into groups C and E, respectively. Interestingly, groups D and G shared the highest (four) members of the MePLATZ family, including Manes.01G121150, Manes.15G136800, Manes.17G083200 and Manes.17G085975, and 08G064000, Manes.09G128900, Manes.15G188400, Manes.17G011500, respectively.

Previously, the classification of the PLATZ families had been reported in other higher plant species. For example, a comprehensive phylogenetic tree of the PLATZ proteins in six *Malus* spp., woodland strawberry (*Fragaria vesca*), peach (*Prunus persica*), common pear (*Pyrus communis*), *Arabidopsis thaliana*, and rice (*Oryza sativa*) revealed that these proteins could be phylogenetically categorized into seven different

groups [13]. Another phylogenetic tree of the PLATZ proteins in Chinese cabbage, Arabidopsis, soybean, rice, wheat, maize, potato, and tomato revealed that these proteins could be categorized into seven different clades [16]. Similarly, this phenomenon was also recognized in the PLATZ families in apple [12], Chinese cabbage [14], and ginkgo [15]. Taken together, our study strongly suggested that the PLATZ family in higher plant species, perhaps in cassava, could be sorted into seven groups.

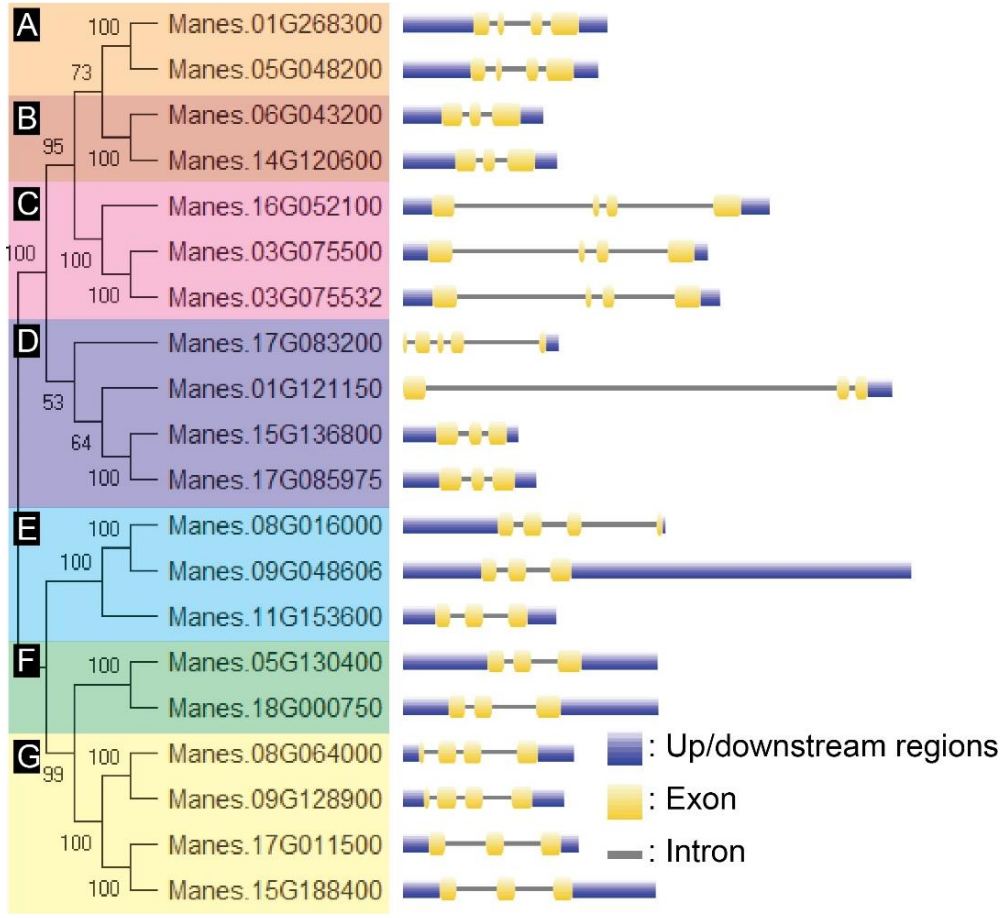

Figure 2. Categorization and gene structure of the MePLATZ family in cassava

Next, we analyzed the gene structure of all members of the MePLATZ family in cassava. As provided in Figure 2, the numbers of exons of the *MePLATZ* genes ranged from three to five. In particular, 11 (out of 20) members of the *MePLATZ* genes had three exons, while eight and one (out of 20) *MePLATZ* genes contained four and five exons, respectively. Interestingly, we found that the *MePLATZ* genes in the same clade should share similar structural patterns. For example, two *MePLATZ* genes in group A, including *Manes.01G268300* and *Manes.05G048200*, shared the same gene sizes and amounts of the exon. This finding was also recorded in the *MePLATZ* genes belonging to groups B, C, F, and G.

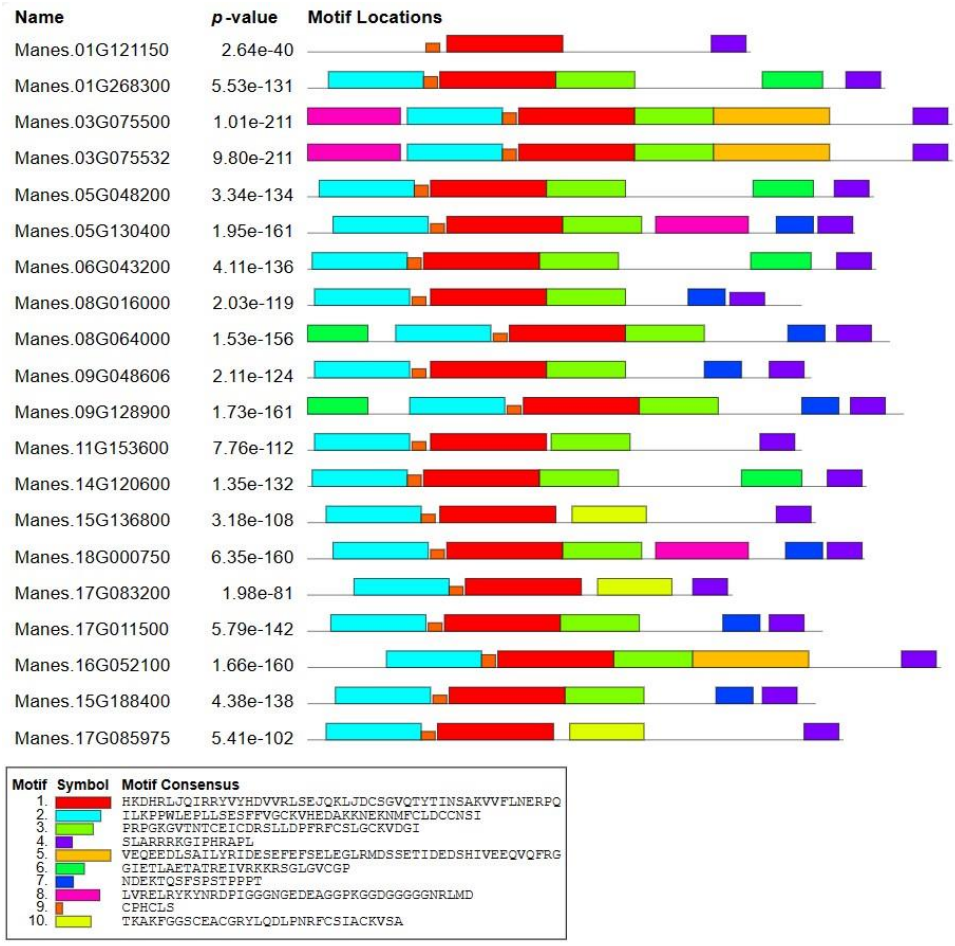

Figure 3. Conserved motifs of the MePLATZ transcription factor family in cassava

It is well understood that structural variation among genes drives the expansion of multigene families. Here, we discussed the gene structure of the *PLATZ* gene families in other higher plant species [15]. In ginkgo, a majority of the *PLATZ* genes contained 2 - 4 introns, whereas the remaining *PLATZ* genes were intronless [15]. The amounts of exon in the *PLATZ* gene family in Chinese cabbage ranged from two to five [14]. Among them, 14 (out of 24) *PLATZ* genes had four exons, whereas seven, one, and two *PLATZ* genes contained three, two, and five exons, respectively [14].

As a major part of this current work, the typical motifs of the MePLATZ proteins have been screened. Based on the *p*-value, a total of 10 specific motifs have been found in the MePLATZ proteins. As illustrated in Figure 3, motifs 1, 4, and 9 occurred in the whole members of the MePLATZ proteins. Motif 2 was found to be localized in the majority (19 out of 20) members of the MePLATZ proteins, excluding Manes.01G121150. Numerous of the MePLATZ proteins contain motif 3, excluding Manes.01G121150, Manes.15G136800, Manes.17G083200 and Manes.17G085975. The remaining motifs were found in the similar MePLATZ proteins. In previous works, the presence of 10 common motifs was also recorded in the PLATZ families in higher plant species [12]-[16].

2.2.3. Expression analysis of the PLATZ transcription factor family under the drought condition in cassava

Cassava has been well-reported to be a drought-tolerant crop. Thus, to understand how the duplicated *MePLATZ* genes are involved in drought resistance in cassava, we investigated the expression patterns of 17 duplicated genes in treated leaf samples. Based on the previous RNA-Seq dataset [19], the fold-change values of eight duplicated pairs were re-analyzed and provided in Figure 4.

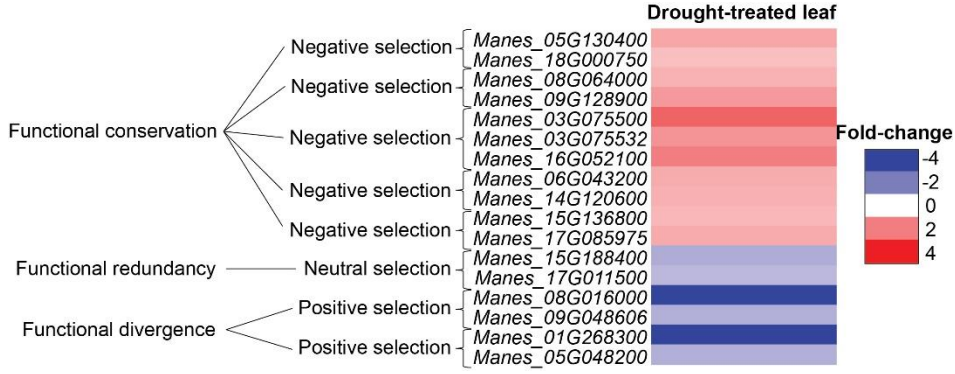

Figure 4. Expression profiles of the duplicated MePLATZ genes in drought-treated leaf samples in cassava

Our study revealed that the duplicated *MePLATZ* genes exhibited variable expression levels in leaf samples under drought conditions. In particular, four duplicated *MePLATZ* genes that occurred by negative selection, including *Manes_05G130400* and *Manes_18G000750*, *Manes_08G064000* and *Manes_09G128900*, *Manes_06G043200* and *Manes_14G120600*, and *Manes_15G136800* and *Manes_17G085975*, were not significantly expressed ($|\text{fold-change}| \leq 2$) in leaves under drought condition. Similarly, we also found that the expression levels of one duplicated *MePLATZ* gene pair that occurred by neutral selection were not highly altered in leaf samples under drought conditions. Their non-differential expression under drought conditions implies these duplicated genes might not be directly involved in the drought response, or this response is mediated at a post-transcriptional level or by other genes/pathways. This constancy in expression could also signify robustness to environmental stresses, ensuring the stability of vital processes under varying conditions. Additionally, the presence of these duplicated *MePLATZ* genes might contribute to functional redundancy, allowing the cassava plant to maintain its basic functions even when one gene copy undergoes mutations or changes.

Of our interest, we found that two duplicated *MePLATZ* gene pairs exhibited different expression levels in leaf samples under drought conditions. In particular, *Manes_01G268300* and *Manes_08G016000* were noted to be highly reduced in drought-treated leaves, whereas *Manes_05G048200* and *Manes_09G048606* were not significantly expressed in the tested tissues. The different expression patterns under drought conditions could be a manifestation of this divergence, where two genes, including *Manes_01G268300* and *Manes_08G016000* might have evolved a specific role in the drought response while the remaining genes, like *Manes_05G048200* and *Manes_09G048606* maintained their original function or has developed a different role. It could be explained that positive selection is a key driver of adaptive evolution, leading

to functional divergence between gene copies. In particular, mutations that occurred in these duplicated *MePLATZ* genes have been beneficial for cassava, providing some sort of evolutionary advantage, perhaps in response to environmental stresses like drought. Over time, these advantageous mutations are retained and accumulated, leading to the evolution of specialized functions. This hypothesis could explain why one *MePLATZ* gene's expression is significantly reduced under drought conditions while the other remains unchanged.

Previously, *Manes_01G268300*, *Manes_09G048606* and *Manes_08G016000* were not expressed in any major organs, including root, lateral bud, leaf, mid vein, stem, and petiole tissues under normal conditions [11]. Additionally, *Manes_05G048200* was expressed in lateral bud tissues under normal conditions [11]. This finding suggested that *Manes_01G268300*, *Manes_09G048606*, and *Manes_08G016000* could be house-keeping genes, while *Manes_05G048200* might regulate the biological processes related to the lateral bud tissues during the growth and development of cassava plants. Recently, several duplicated *PLATZ* genes in apples, including *MdPLATZ1*, *MdPLATZ6*, *MdPLATZ7*, and *MdPLATZ16* were ubiquitously expressed across all examined organs with the exception of pollen, suggesting their pivotal role in supporting apple growth and development [12]. Conversely, the expression of *MdPLATZ15* demonstrated a progressive decline concurrent with the ripening of fruit, potentially implicating its involvement in the modulation of fruit firmness [12]. Taken together, the expression profiles of these duplicated *MePLATZ* genes under various environmental conditions should be validated to confirm this hypothesis.

3. Conclusions

To get insight into the evolution of the *MePLATZ* gene family in cassava, we provided a comprehensive analysis of the duplication event, gene structure, classification, conserved motifs, and expression profiles of the *MePLATZ* genes by using various bioinformatics approaches. Structural analysis indicated that three and four exons are the most common gene organizations of the *MePLATZ* genes, while at least 10 conserved motifs were found in the structure of the corresponding proteins. Our phylogenetic tree revealed that the MePLATZ proteins could be grouped into seven clades. The segmental and tandem duplication events were predicted as the main reasons for the expansion of the *MePLATZ* gene family in cassava. Among them, all 17 duplicated *MePLATZ* genes exhibited different expression levels in leaf tissues under drought conditions. To sum up, this current work could provide solid evidence for the expansion of this multiple-gene family in cassava.

REFERENCES

[1] Chavarriaga-Aguirre P, Brand A, Medina A, Prias M, Escobar R, Martinez J, Diaz P, Lopez C, Roca WM & Tohme J, (2016). The potential of using biotechnology to improve cassava: a review. *In Vitro Cellular & Developmental Biology - Plant*, 52(5), 461-478. 10.1007/s11627-016-9776-3.

[2] Malik AI, Kongsil P, Nguyen VA, Ou W, Sholihin, Srean P, Sheela MN, Becerra Lopez-Lavalle LA, Utsumi Y, Lu C, Kittipadakul P, Nguyen HH, Ceballos H, Nguyen TH, Selvaraj Gomez M, Aiemnaka P, Labarta R, Chen S, Amawan S, Sok S, Youabee L, Seki M, Tokunaga H, Wang W, Li K, Nguyen HA, Nguyen VD, Ham LH & Ishitani M, (2020). Cassava breeding and agronomy in Asia: 50 years of history and future directions. *Breed Science,* 70(2), 145-166. 10.1270/jsbbs.18180.

[3] Li S, Cui Y, Zhou Y, Luo Z, Liu J & Zhao M, (2017). The industrial applications of cassava: current status, opportunities, and prospects. *Journal of the Science of Food and Agriculture,* 97(8), 2282-2290. 10.1002/jsfa.8287.

[4] Nilusha RAT, Jayasinghe J, Perera O, Perera PIP & Jayasinghe CVL, (2021). Proximate composition, physicochemical, functional, and antioxidant properties of flours from selected cassava (*Manihot esculenta* Crantz) varieties. *International Journal of Food Science,* 6064545. 10.1155/2021/6064545.

[5] Wang A, Hou Q, Si L, Huang X, Luo J, Lu D, Zhu J, Shangguan Y, Miao J, Xie Y, Wang Y, Zhao Q, Feng Q, Zhou C, Li Y, Fan D, Lu Y, Tian Q, Wang Z & Han B, (2019). The PLATZ transcription factor GL6 affects grain length and number in rice. *Plant Physiol,* 180(4), 2077-2090. 10.1104/pp.18.01574.

[6] Kim JH, Kim J, Jun SE, Park S, Timilsina R, Kwon DS, Kim Y, Park S-J, Hwang JY, Nam HG, Kim G-T & Woo HR, (2018). ORESARA15, a PLATZ transcription factor, mediates leaf growth and senescence in *Arabidopsis*. *New Phytologist,* 220(2), 609-623. https://doi.org/10.1111/nph.15291.

[7] Nagano Y, Furuhashi H, Inaba T & Sasaki Y, (2001). A novel class of plant-specific zinc-dependent DNA-binding protein that binds to A/T-rich DNA sequences. *Nucleic Acids Research,* 29(20), 4097-4105. 10.1093/nar/29.20.4097.

[8] Sandhu JS, Webster CI & Gray JC, (1998). A/T-rich sequences act as quantitative enhancers of gene expression in transgenic tobacco and potato plants. *Plant Molecular Biology,* 37(5), 885-896. 10.1023/a:1006051832213.

[9] Zhang J, Li C, Zhang W, Zhang X, Mo Y, Tranquilli GE, Vanzetti LS & Dubcovsky J, (2023). Wheat plant height locus RHT25 encodes a PLATZ transcription factor that interacts with DELLA (RHT1). *Proceedings of the National Academy of Sciences,* 120(19), e2300203120. 10.1073/pnas.2300203120.

[10] Zhao J, Zheng L, Wei J, Wang Y, Chen J, Zhou Y, Chen M, Wang F, Ma Y & Xu Z-S, (2022). The soybean PLATZ transcription factor GmPLATZ17 suppresses drought tolerance by interfering with stress-associated gene regulation of GmDREB5. *Crop Journal,* 10(4), 1014-1025. https://doi.org/10.1016/j.cj.2022.03.009.

[11] Hoang MC, Bui TH, Pham PT, Le NQ, Dong HG & Chu DH, (2024). Identification, characterization, and analysis of expression patterns of genes encoding the Platz transcription factor related to the growth and development of cassava (*Manihot esculenta*). *Journal* of *Forestry Science and Technology,* 13(1), 13-20.

[12] Li J, Zhao Y, Zhang Y, Ye F, Hou Z, Zhang Y, Hao L, Li G, Shao J & Tan M, (2023). Genome-wide analysis of MdPLATZ genes and their expression during axillary bud outgrowth in apple (*Malus domestica* Borkh.). *BMC Genomics,* 24(1), 329. 10.1186/s12864-023-09399-x.

[13] Sun Y, Liu Y, Liang J, Luo J, Yang F, Feng P, Wang H, Guo B, Ma F & Zhao T, (2023). Identification of PLATZ genes in *Malus* and expression characteristics of MdPLATZs in response to drought and ABA stresses. *Frontiers in Plant Science*, 13, 1109784. 10.3389/fpls.2022.1109784.

[14] Azim JB, Khan MFH, Hassan L & Robin AHK, (2020). Genome-wide characterization and expression profiling of plant-specific PLATZ transcription factor family genes in *Brassica rapa* L. *Plant Breeding and Biotechnology*, 8(1), 28-45. 10.9787/PBB.2020.8.1.28.

[15] Han X, Rong H, Tian Y, Qu Y, Xu M & Xu L-a, (2022). Genome-wide identification of PLATZ transcription factors in *Ginkgo biloba* L. and their expression characteristics during seed development. *Frontiers in Plant Science*, 13, 946194. 10.3389/fpls.2022.946194.

[16] Fu Y, Cheng M, Li M, Guo X, Wu Y & Wang J, (2020). Identification and characterization of PLATZ transcription factors in wheat. *International Journal of Molecular Sciences*, 21(23), 8934. 10.3390/ijms21238934.

[17] Bredeson JV, Lyons JB, Prochnik SE, Wu GA, Ha CM, Edsinger-Gonzales E, Grimwood J, Schmutz J, Rabbi IY, Egesi C, Nauluvula P, Lebot V, Ndunguru J, Mkamilo G, Bart RS, Setter TL, Gleadow RM, Kulakow P, Ferguson ME, Rounsley S & Rokhsar DS, (2016). Sequencing wild and cultivated cassava and related species reveals extensive interspecific hybridization and genetic diversity. *Nature Biotechnology*, 34(5), 562-570. 10.1038/nbt.3535.

[18] Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, Mitros T, Dirks W, Hellsten U, Putnam N & Rokhsar DS, (2012). Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Research*, 40(Database issue), D1178-D1186. 10.1093/nar/gkr944.

[19] Zhu Y, Luo X, Wei M, Khan A, Munsif F, Huang T, Pan X & Shan Z, (2020). Antioxidant enzymatic activity and its related gene expression in cassava leaves at different growth stages play key roles in sustaining yield and drought tolerance under moisture stress. *Journal of Plant Growth Regulation*, 39(2), 594-607. 10.1007/s00344-019-10003-4.

[20] Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, Marshall KA, Phillippy KH, Sherman PM, Holko M, Yefanov A, Lee H, Zhang N, Robertson CL, Serova N, Davis S & Soboleva A, (2013). NCBI GEO: archive for functional genomics data sets--update. *Nucleic Acids Research*, 41(Database issue), D991-D995. 10.1093/nar/gks1193.

[21] La HV, Chu HD, Tran CD, Nguyen KH, Le QTN, Hoang CM, Cao BP, Pham ATC, Nguyen BD, Nguyen TQ, Van Nguyen L, Ha CV, Le HT, Le HH, Le TD & Tran LP, (2022). Insights into the gene and protein structures of the CaSWEET family members in chickpeas (*Cicer arietinum*), and their gene expression patterns in different organs under various stress and abscisic acid treatments. *Gene*, 819, 146210. 10.1016/j.gene.2022.146210.

[22] Thompson J, Gibson T, Plewniak F, Jeanmougin F & Higgins D, (1997). The ClustalX Windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research,* 25, 4876-4882.

[23] Hall TA, (1999). BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series,* 41, 95-98.

[24] Rozas J, Ferrer-Mata A, Sanchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE & Sanchez-Gracia A, (2017). DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets. *Molecular Biology and Evolution,* 34(12), 3299-3302. 10.1093/molbev/msx248.

[25] Kumar S, Stecher G & Tamura K, (2016). MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution,* 33(7), 1870-1874. 10.1093/molbev/msw054.

[26] Hu B, Jin J, Guo AY, Zhang H, Luo J & Gao G, (2015). GSDS 2.0: an upgraded gene feature visualization server. *Bioinformatics,* 31(8), 1296-1297. 10.1093/bioinformatics/btu817.

[27] Bailey TL, Johnson J, Grant CE & Noble WS, (2015). The MEME Suite. *Nucleic Acids Research,* 43(W1), W39-W49. 10.1093/nar/gkv416.