HNUE JOURNAL OF SCIENCE
Natural Sciences 2024, Volume 69, Issue 3, pp. 138-147
This paper is available online at http://hnuejs.edu.vn/ns
DOI: 10.18173/2354-1059.2024-0043

INVESTIGATION OF THE EVOLUTION OF SABATH GENE FAMILY IN CASSAVA (*Manihot esculenta*) REVEALS ITS POTENTIAL ROLE IN GROWTH AND DEVELOPMENT

Tong Van Hai[1], Luu Thi Bao Ngoc[1], Nguyen Quoc Trung[1], Dong Huy Gioi[1], Chu Duc Ha[2], Tran Van Tien[3], La Viet Hong[4], Le Thi Ngoc Quynh[5] and Tran Thi Thanh Huyen[6,*]

[1]*Faculty of Biotechnology, Vietnam National University of Agriculture, Hanoi city, Vietnam*
[2]*Faculty of Agricultural Technology, University of Engineering and Technology, 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, Vinh Phuc province, Vietnam*
[5]*Department of Biotechnology, Thuyloi University, Hanoi city, Vietnam*
[6]*Faculty of Biology, Hanoi National University of Education, Hanoi city, Vietnam*
[*]Corresponding author: Tran Thi Thanh Huyen, e-mail: tranthanhhuyen@hnue.edu.vn

Received August 12, 2024. Revised October 12, 2024. Accepted October 31, 2024.

Abstract. SABATH is one class of enzyme belonging to the class of methyltransferase, playing a crucial role in plant defense and stress response mechanisms. Despite its importance, no systematic analysis of the expansion of the *SABATH* gene family in cassava (*Manihot esculenta*) has been reported up to date. In this study, we investigated the *SABATH* gene family in cassava, revealing their random distribution across 18 chromosomes, with variable numbers of gene copies per chromosome. According to gene duplication analysis, five duplication events were identified, primarily segmental duplications, indicating their evolutionary significance. The Ka/Ks ratio analysis indicated that most duplicated genes are under negative selection, preserving their functions, while one pair showed signs of positive selection, suggesting adaptive benefits. Gene structure analysis showed diverse exon counts, primarily three or four. Expression profiling across 11 cassava tissues demonstrated tissue-specific expression patterns, with some genes highly or exclusively expressed in specific tissues such as root apical meristems, embryogenic calli, and fibrous roots, implying distinct functional roles in cassava growth and development. Overall, this study provides valuable insights into the evolution and functional diversity of the *SABATH* gene family in cassava and identifies candidate genes for further functional characterization.

Keywords: SABATH, gene duplication, gene structure, expression profile, cassava.

1. Introduction

Cassava (*Manihot esculenta*) is a crucial staple crop in many tropical and subtropical regions, known for its high carbohydrate content, mainly starch, a crucial energy source for millions of people [1]. Economically, cassava is significant due to its versatility being used for human food, animal feed, and industrial applications such as bioethanol production [2]. Cassava is notably stress-tolerant, thriving in poor soils with minimal inputs and displaying remarkable resistance to drought and high temperatures [3]. This resilience is attributed to its extensive root system, ability to reduce metabolic activity under stress, and efficient water usage [3], [4]. Thus, studying how cassava survives adverse environmental conditions is essential for improving food security, particularly in the face of climate change, by developing crops that can withstand harsh environments and ensuring stable yields in unpredictable climates.

SABATH is a group of enzyme belonging to the class of methyltransferase that plays a vital role in the plant's response to adverse environmental conditions [5]. Structurally, SABATH is a protein that catalyzes explicitly the methylation of carboxylic acids and nitrogen atoms [6]. Functionally, SABATH is essential in regulating various physiological processes, including growth, development, and stress responses [7]. The specialized methylated metabolites help plants adapt to environmental conditions by activating defense genes and enhancing resistance to herbivores, pathogens, and physical stressors like drought and salinity. This adaptation mechanism is critical for plant survival and productivity in changing environments. Interestingly, the SABATH family in higher plant species, such as rice (*Oryza sativa*) [6], tomato (*Solanum lycopersicum*) [8], *Hedychium coronarium* [9], *Neolamarckia cadamba* [10] and tea plants (*Camellia sinensis*) [11], has been reported to contain multiple genes. Thus, it would be interesting to study the expansion of the *SABATH* gene family and the role of their duplicated genes during the growth and development of cassava plants. Understanding the evolution and expansion of the *SABATH* gene family in cassava could show how these duplicated genes help the plant thrive under different environmental stresses.

The aim of this study was to investigate the duplication events within the SABATH gene family in cassava using computational tools. By analyzing the similarities between coding DNA sequences, we identified potential duplicated gene pairs and examined their structures. In addition, we assessed the expression profiles of SABATH genes across various cassava tissues during key growth and developmental stages. Our findings provide valuable insights into the specific roles of these duplicated genes, particularly in enhancing cassava's growth and stress tolerance. These insights can contribute to improving cassava's resilience to environmental challenges and supporting its productivity as a staple crop.

2. Content

2.1. Materials and methods

2.1.1. Data collection

The recent cassava reference genome (NCBI RefSeq assembly: GCF_001659605.2) [12] were obtained from the Phytozome v13 (https://phytozome-next.jgi.doe.gov/) [13] and NCBI (https://www.ncbi.nlm.nih.gov/) portals.

Transcriptome atlas (GEO accession number: GSE82279) of 11 samples under normal condition [14] was deposited in NCBI Gene Expression Omnibus (ncbi.nlm.nih.gov/geo/) [15].

Full-length protein sequences, coding DNA sequences, and genomic DNA sequences of 23 well-annotated members of the *SABATH* family in cassava obtained in the recent report were utilized for further computational analysis in this study.

2.1.2. Methods

Chromosomal localization of genes: The location of each *SABATH* gene was identified using its annotation. Gene identifiers were matched against the cassava genome [12] in Phytozome [13] and NCBI databases. Adobe Illustrator software was then used to visualize the physical locations of the *SABATH* genes.

Analysis of the gene duplication: Duplicated *SABATH* genes were identified following methods described previously [16]. Precisely, coding DNA sequences of all *SABATH* genes from earlier work were aligned using ClustalX v2.1 software [17]. BioEDIT v7.2.6 software [18] was utilized to calculate similarity scores. Genes sharing over 70% similarity were considered duplicates [19].

Estimation of Ka/Ks values: The Ka (nonsynonymous substitutions per nonsynonymous site) and Ks (synonymous substitutions per synonymous site) values for each duplicated *SABATH* pair were calculated as previously outlined [19]. Aligned coding DNA sequences were analyzed employing DNASp v6.12.03 tool [20]. A Ka/Ks ratio greater than 1.00, less than 1.00, and exactly 1.00 indicated positive selection, stabilizing selection, and neutral selection, respectively [19].

Analysis of gene organization: The exon/intron organization of *SABATH* genes was analyzed using the GSDS v2.0 website (https://gsds.gao-lab.org/) [21] as formerly reported [19]. The gene arrangements followed the phylogenetic tree order. Subsequently, Adobe Illustrator v28.0 software was utilized to illustrate all gene structures.

Analysis of microarray dataset: The expression profiles of the *SABATH* genes were analyzed using the recent transcriptome atlas [14] available from the NCBI Gene Expression Omnibus [15]. The FPKM (Fragments Per Kilobase of transcript per Million mapped reads) values were used to assess the tissue-specific expression of *SABATH* genes according to the previous report [22]. Eleven samples including friable embryogenic calli, fibrous root, lateral bud, somatic organized embryogenic structures, leaf, mid vein, petiole, root apical meristem, shoot apical meristem, stem, and storage root, were examined. An FPKM value less than 10.00 indicated that the gene was below the detection threshold, values between 10.00 and 50.00, between 50.00 and 100.00 or greater than 100.00 demonstrated gene expression, high expression, or exclusive expression, respectively [22]. A heatmap was thereby constructed using Python script.

2.2. Results and discussion

2.2.1. Chromosomal distribution of the *SABATH* gene family in cassava

To get insight into the physical distribution of all 23 members of the *SABATH* gene family on cassava chromosomes, every SABATH gene was searched against the cassava genome. As described in Figure 1, all *SABATH* genes were found to be

randomly distributed across all 18 chromosomes of the cassava genome. Specifically, each of chromosomes 3, 4, 5, 13, and 18 contained only one *SABATH* gene, including *Manes.03G136100*, *Manes.04G074231*, *Manes.05G156400*, *Manes.13G061650* and *Manes.18G145282*, respectively. Chromosomes 2, 6, 10, and 17 each had two *SABATH* genes. Chromosomes 1 and 15 contained the highest number (five members) of *SABATH* genes. However, no *SABATH* genes were found on chromosomes 7, 8, 9, 11, 12, 14 and 16.

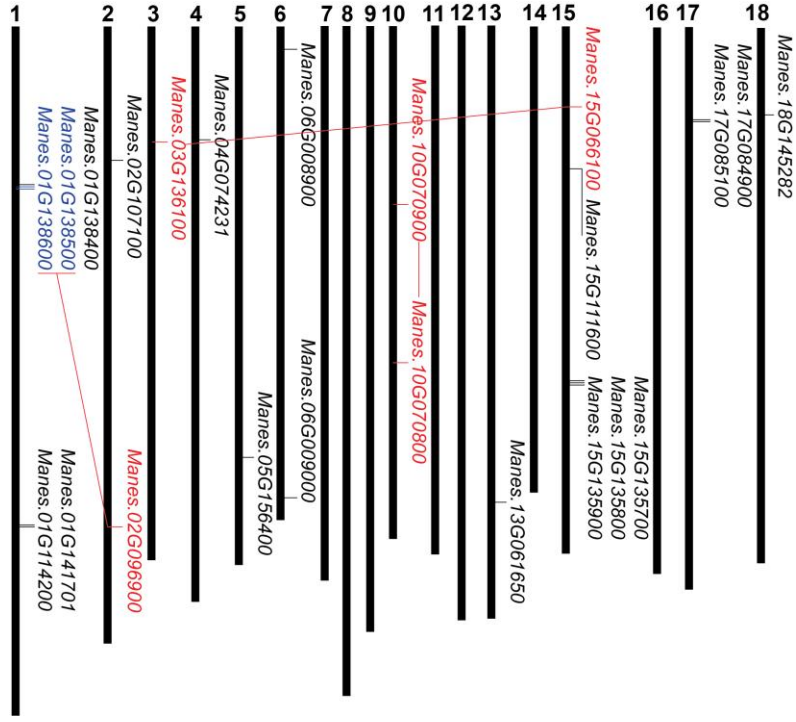

Figure 1. The physical location of the SABATH gene family in the genome of cassava. Red, blue, and black indicated segmental, tandem, and no duplication, respectively.

According to previous researches, the *SABATH* genes were also localized in the genome of higher plant species with uneven rates [8]-[10]. In tomato, all 20 members of the *SABATH* methyltransferase gene family, catalyzing the methylation of hormones, signal molecules, and other metabolites, were distributed unevenly across the 12 chromosomes [8]. Among them, chromosome 1 had the highest number, with seven genes, while chromosomes 9 and 10 each had three genes [8]. Chromosomes 2 and 4 each contained two genes, while each of chromosomes 7, 11, and 12 had only one gene [8]. No *SABATH* genes were present on chromosomes 3, 5, 6, and 8 [8]. In the case of *N. cadamba*, a total of 22 members of the *SABATH* gene family were reported to distribute unevenly across 12 chromosomes, with one member located on a scaffold [10]. Mainly, chromosome 19 had the highest concentration with four genes [10]. Chromosomes 9 and 13 had three genes in each one [10]. Chromosomes 10, 12, and 22 each contained two genes, while chromosomes 5, 6, 7, 14, 16, and 17 had only one gene each [10]. Recently, 11 (out of 12) *SABATH* genes were indicated to locate on three chromosomes of the *H. coronarium* genome and were unevenly distributed across these chromosomes [9].

2.2.2. Gene duplication events in the *SABATH* gene family in cassava

The *SABATH* gene families in higher plant species have been demonstrated to contain multiple genes [8]-[10]. Thus, describing the evolution of the *SABATH* gene family in cassava would be significant. As a result, the duplication events in the *SABATH* gene family have been provided in Figure 1 and Table 1. We predicted that at least five duplication events have been recorded in the *SABATH* gene family in cassava. Notably, the similarities of these duplicated *SABATH* genes ranged from 79.60 (*Manes.01G138600* and *Manes.02G096900*) to 87.50% (*Manes.01G138600* and *Manes.01G138500*). According to the chromosomal distribution, it has been recognized that four (out of five) duplication events, including a pair of *Manes.03G136100* and *Manes.15G066100*, *Manes.10G070900* and *Manes.10G070800*, *Manes.02G096900* and *Manes.01G138500*, and *Manes.01G138600* and *Manes.02G096900*, have occurred as a result of the segmental duplication, whereas only one tandem duplication event was found to be *Manes.01G138600* and *Manes.01G138500*.

Table 1. Summary of the duplication events occurring in the SABATH gene family of cassava

#	Duplicated pairs	Duplication events	Similarity	Ka	Ks	Ka/Ks
1	*Manes.03G136100* *Manes.15G066100*	Segmental duplication	83.40	0.17	0.16	1.06
2	*Manes.10G070900* *Manes.10G070800*	Segmental duplication	81.10	0.09	0.14	0.64
3	*Manes.01G138600* *Manes.01G138500*	Tandem duplication	87.50	0.10	0.15	0.67
4	*Manes.02G096900* *Manes.01G138500*	Segmental duplication	81.10	0.2	0.26	0.77
5	*Manes.01G138600* *Manes.02G096900*	Segmental duplication	79.60	0.19	0.27	0.70

Note: Ka - Nonsynonymous substitutions per nonsynonymous site, Ks - Synonymous substitutions per synonymous site.

To detect the selection pressures affecting the *SABATH* gene family in cassava, the Ka/Ks ratios for five duplicated gene pairs were estimated. As shown in Table 1, the Ka/Ks ratios of these duplicated genes ranged from 0.64 (*Manes.10G070900* and *Manes.10G070800*) to 1.06 (*Manes.03G136100* and *Manes.15G066100*). Notably, four duplication events had Ka/Ks values less than 1.00. This indicated that these duplicated genes were likely driven by negative selection, where alterations to their amino acid sequences (nonsynonymous changes) are being rejected to maintain their critical functions. In contrast, only one duplicated pair showed Ka/Ks ratios greater than 1.00, suggesting that this duplication event was under positive selection, with nonsynonymous changes being favored for providing adaptive advantages to cassava plants.Recently, the expansion of the *SABATH* gene families in higher plant species has been investigated. To understand the evolutionary patterns of the *SABATH* gene family in tomatoes, a total of seven tandem duplication and two segmental duplication events were predicted using a Blastall tool, respectively [8]. Furthermore, all tandem duplicated genes had

Ka/Ks ratios less than 1.00, indicating they were under purifying selection [8]. Thus, positive selection did not drive gene divergence after *SABATH* tandem duplication [8]. In *N. cadamba*, the collinearity analysis revealed 12 duplicated gene pairs within the *SABATH* gene family [10]. Particularly, only two tandem duplicated gene pairs were predicted, located on chromosomes 9 and 12 of the *N. cadamba* genome [10]. Additionally, 69.57% (16 out of 23) of *SABATH* genes in *N. cadamba* underwent segmental duplication, forming ten segmental duplicated gene pairs distributed across 9 of the 22 chromosomes of the *N. cadamba* genome [10]. This indicated that segmental duplication likely played a significant role in the expansion of the *SABATH* gene family in *N. cadamba* [10]. All Ka/Ks ratios of these 12 duplication events were less than 1.00, suggesting that the duplicated genes were driven by purifying selection [10].

2.2.3. Gene structure of the *SABATH* gene family in cassava

To investigate the gene structure of *SABATH* genes in cassava, the GSDS website [21] was applied to analyze the coding DNA sequences and genomic DNA sequences. The exon/intron organization of the *SABATH* genes in cassava is provided in Figure 2. It is shown that 12 members of the *SABATH* gene family in cassava contained four exons. Additionally, nine *SABATH* genes exhibited three exons. Only *Manes.01G141701* and *Manes.15G135800* had five and two exons, respectively.

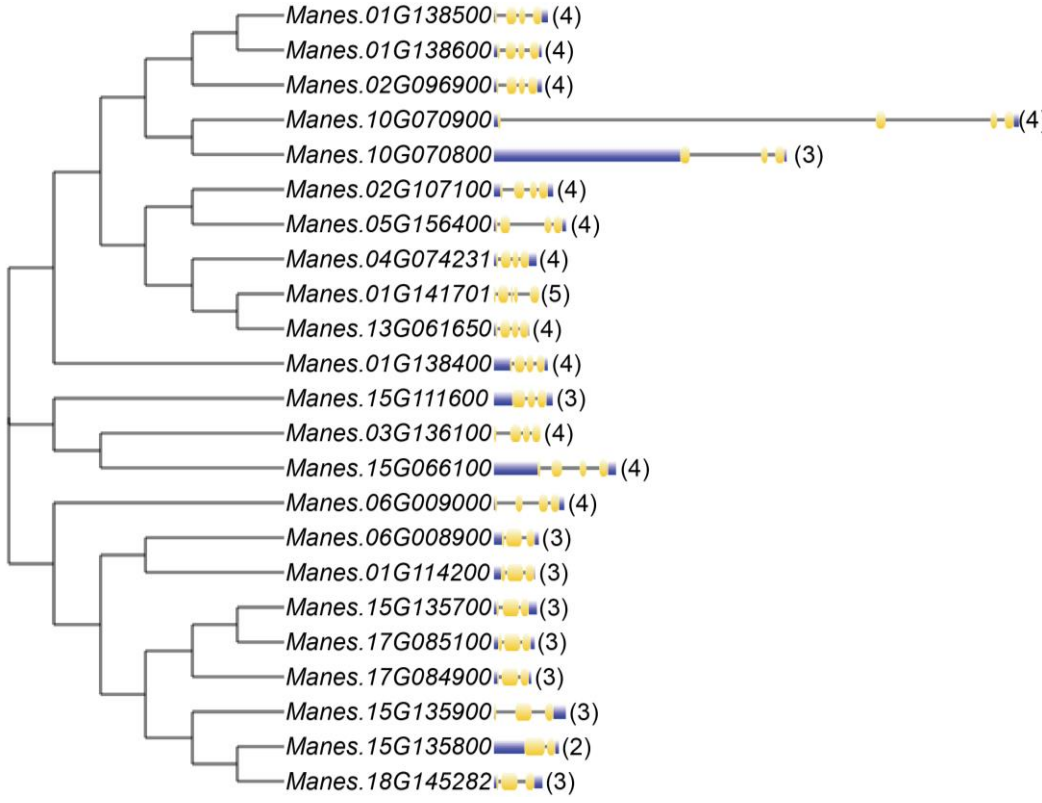

Figure 2. Gene structure of the SABATH gene family in cassava. Blue and yellow boxes indicated upstream/downstream regions and exons respectively, while the black line indicated introns

In recent studies, the gene structure of the *SABATH* gene family in higher plant species was focused [8], [10]. For example, nearly all coding sequences of *SABATH* genes in tomatoes were interrupted by one or more introns [8]. Among them, the number of exons ranged from two to five [8]. According to the structural features of each member of the *SABATH* gene family, 26 exons were symmetric with phase 0 introns, representing 38% of the 68 total exons, while no exons were symmetric with phase 1 or 2 introns [8]. To explore the structural diversity of *SABATH* genes in *N. cadamba*, the intron-exon structure of each gene was analyzed by examining its genomic DNA sequence and coding DNA sequence [10]. This analysis showed that all coding sequences of the *SABATH* genes contained one or more introns, ranging from one to seven, while the number of exons varied from two to eight [10].

2.2.4. Expression analysis of the *SABATH* gene family in major organs during the growth and development of cassava

To understand the potential function of the *SABATH* genes during the growth and development of cassava plants, the expression pattern of each *SABATH* gene was explored by re-analyzing the recent transcriptome database. As a result, the expression profiles of the *SABATH* genes in 11 major tissue samples are provided in Figure 3. We found that *Manes.01G141701* was exclusively expressed in root apical meristems, while the expression of *Manes.13G061650* was specific in friable embryogenic calli. Two *SABATH* genes, including *Manes.06G009000* and *Manes.15G135800*, were highly expressed in fibrous root tissues. Manes.02G096900 was exclusively expressed in the lateral bud, leaf, and petiole tissues. Finally, *Manes.04G074231* was specifically expressed in fibrous roots and root apical meristems.

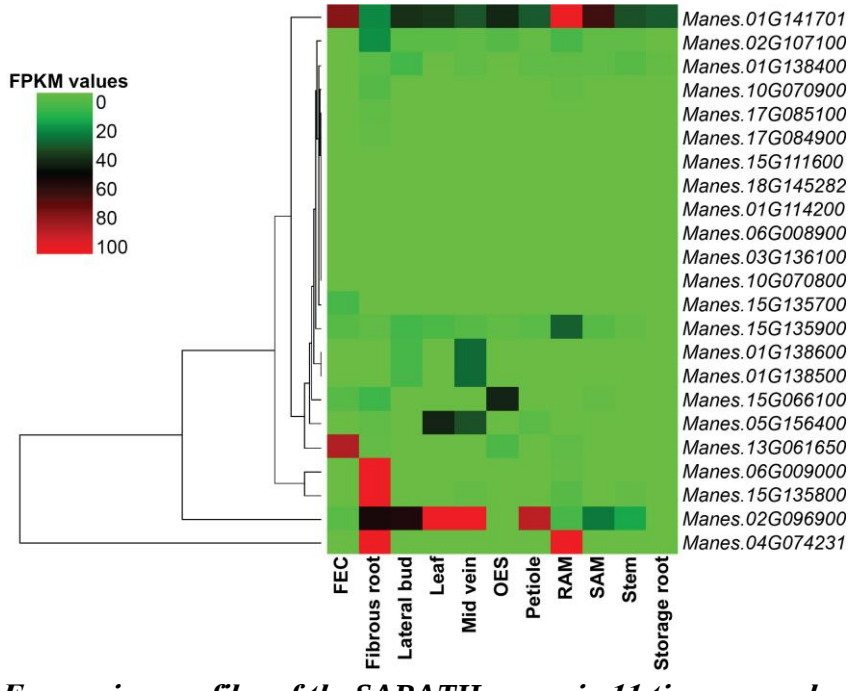

Figure 3. Expression profiles of the SABATH genes in 11 tissue samples in cassava

The expression levels of the *SABATH* in major organs of higher plant species were also explored to suggest their putative roles in various biological processes. For example, to identify the highly expressed *SABATH* genes in the flowers of *H. coronarium*, the expression profiles of *SABATH* genes in different tissues were analyzed using transcriptomic data [9]. The results showed that *SABATH* gene expression was tissue-specific, clustering into two main categories. Category I genes exhibited the highest expression in petals. In contrast, category II genes were most highly expressed in leaves [9]. In the case of *N. cadamba*, 22 members of the *SABATH* gene family were expressed in at least one of the 16 tissues examined, and six of the genes exhibited very low expression levels across all tissues [10]. For instance, three genes were primarily expressed in buds and young leaves, while two genes had high expression in the xylem [10]. Three genes were predominantly expressed in fruit and old leaves, while two genes showed high expression in the cambium [10]. Notably, four genes were highly expressed in most tested tissues, particularly in bark, bud, cambium, and phloem [10].

3. Conclusions

In summary, our study comprehensively investigated the evolution of the *SABATH* gene family in cassava applying various computational tools. All members of the *SABATH* gene family exhibited a random distribution across 18 chromosomes of the cassava genome, with some chromosomes containing multiple genes and others none at all. Analysis of gene duplication events revealed at least five duplication occurrences, with four segmental and one tandem duplication, indicating that segmental duplication has played a significant role in the evolution of the *SABATH* gene family in cassava. The Ka/Ks ratio analysis showed that most duplicated genes are under negative selection, maintaining their critical functions, while only one pair is under positive selection, suggesting adaptive advantages. Gene structure analysis indicated a variable number of exons among *SABATH* genes in cassava, with the majority having three or four exons. Expression analysis across 11 major cassava tissues highlighted tissue-specific expression patterns, with certain genes showing exclusive or high expression in specific tissues, such as root apical meristems, embryogenic calli, and fibrous roots, indicating their potential functional roles in cassava growth and development.

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.

[2] 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.

[3] 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.

[4] Zhao P, Liu P, Shao J, Li C, Wang B, Guo X, Yan B, Xia Y & Peng M, (2015). Analysis of different strategies adapted by two cassava cultivars in response to drought stress: ensuring survival or continuing growth. *Journal of Experimental Botany*, 66(5), 1477-1488.

[5] Seo HS, Song JT, Cheong JJ, Lee YH, Lee YW, Hwang I, Lee JS & Choi YD, (2001). Jasmonic acid carboxyl methyltransferase: a key enzyme for jasmonate-regulated plant responses. *Proceedings of the National Academy of Sciences of the United States of America*, 98(8), 4788-4793.

[6] Qi J, Li J, Han X, Li R, Wu J, Yu H, Hu L, Xiao Y, Lu J & Lou Y, (2016). Jasmonic acid carboxyl methyltransferase regulates development and herbivory-induced defense response in rice. *Journal of Integrative Plant Biology*, 58(6), 564-576.

[7] Cheong JJ & Choi YD, (2003). Methyl jasmonate as a vital substance in plants. *Trend in Genetics*, 19(7), 409-413.

[8] Wei X, Tao K, Zhang J, Lu S, Chen S & Liao J, (2021). Identification of SABATH family members in *Solanum lycopersicum* and their expression patterns under abiotic/biotic stresses. *Plant Molecular Biology Reporter*, 39(2), 403-418.

[9] Yue Y, Zhang X, Wang L, He J, Yang S, Li X, Yu Y, Yu R & Fan Y, (2024). Identification and characterization of jasmonic acid methyltransferase involved in the formation of floral methyl jasmonate in *Hedychium coronarium*. *Plants*, 13(1), 8.

[10] Ren R, Zhang S, Guo T, Long J & Peng C, (2023). Genome-wide identification and expression pattern analysis of the SABATH gene family in *Neolamarckia cadamba*. *Journal of Forest Research*, 3(1), 1-13.

[11] Guo Y, Qiao D, Yang C, Chen J, Li Y, Liang S, Lin K & Chen Z, (2020). Genome-wide identification and expression analysis of SABATH methyltransferases in tea plant (*Camellia sinensis*): insights into their roles in plant defense responses. *Plant Signaling & Behavior*, 15(10), 1804684.

[12] 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.

[13] 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.

[14] Wilson MC, Mutka AM, Hummel AW, Berry J, Chauhan RD, Vijayaraghavan A, Taylor NJ, Voytas DF, Chitwood DH & Bart RS, (2017). Gene expression atlas for the food security crop cassava. *New Phytologists*, 213(4), 1632-1641.

[15] 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.

[16] Hoang MC, Pham PT, Le TNQ, 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*). *Forestry Science and Technology Journal,* 13(1), 13-20.

[17] Thompson JD, Gibson TJ & Higgins DG, (2002). Multiple sequence alignment using ClustalW and ClustalX. *Curr Protoc Bioinformatics,* Chapter 2(Unit 2 3,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.

[18] 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.

[19] La HV, Chu DH, Tran CD, Nguyen KH, Le QTN, Hoang CM, Cao BP, Pham ATC, Nguyen BD, Nguyen TQ, Nguyen VL, 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.

[20] 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.

[21] 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.

[22] Tran VT, La VH, Nguyen QT, Phi CT, Bui TTH, Nguyen VL, Dong HG, Le TNQ, Tran TTH, Chu DH & Cao PB, (2024). Genome-wide identification and characterization of the GATA transcription factor family suggests functional expression patterns against various environmental conditions in cassava (*Manihot esculenta*). *Journal of Animal and Plant Sciences,* 34(2), 1-10.