HNUE JOURNAL OF SCIENCE DOI: 10.18173/2354-1059.2023-0071
Natural Sciences 2023, Volume 68, Issue 3, pp. 154-161
This paper is available online at http://stdb.hnue.edu.vn

IDENTIFICATION, ANNOTATION AND CHARACTERIZATION OF THE TREHALOSE-6-PHOSPHATE SYNTHASE IN CASSAVA (*Manihot esculenta*)

Ha Duc Chu[1], Nguyen Duc Tien[2], Pham Phuong Thu[3], Le Thi Ngoc Quynh[4], Tran Van Tien[5] and Tran Thi Thanh Huyen[6]

[1]*Faculty of Agricultural Technology, University of Engineering and Technology*
[2]*Vietnam Institute of Agricultural Engineering and Post-Harvest Technology*
[3]*Faculty of Biology, Agricultural Technology, Hanoi Pedagogical University 2*
[4]*Department of Biotechnology, Thuyloi University*
[5]*Faculty of Social Management, National Academy of Public Administration*
[6]*Faculty of Biology, Hanoi National University of Education*

Abstract. The trehalose-6-phosphate synthase (TPS) proteins have been characterized to play an important role in numerous biological processes during the growth and development of plants. However, a comprehensive analysis of this important gene family is lacking for cassava (*Manihot esculenta*). The objective of this current study is to identify, annotate, and characterize the *TPS* genes (*MeTPS*) in cassava. A total of 12 *MeTPS* genes were identified in the cassava assembly. They were assigned to two different groups based on phylogeny. Our analyses indicated that these proteins shared slightly similar general features. Of our interest, the *MeTPS* genes exhibited divergent expression in major organs in cassava plants. Among them, *MeTPS10* was noted to exhibit a high expression in fibrous root, lateral bud, petiole, and stem, while *MeTPS11* was exclusively expressed in storage root tissues and strongly expressed in fibrous root and lateral bud tissues. Taken together, our study could provide a list of candidate *MeTPS* genes for further functional characterization.

Keywords: Trehalose-6-phosphate synthase, identification, characterization, expression profile, cassava.

1. Introduction

Cassava (*Manihot esculenta*), a vital staple crop, holds significant importance in both agriculture and food security, making it an indispensable asset to global communities [1, 2]. This drought-tolerant root vegetable, originating from South America, has traversed

Received August 3, 2023. Revised October 18, 2023. Accepted October 25, 2023.
Contact Tran Thi Thanh Huyen, e-mail address: tranthanhhuyen@hnue.edu.vn

continents to become a ubiquitous presence in various tropical and subtropical regions worldwide [2]. Renowned for its versatility, cassava serves as a fundamental source of sustenance, providing vital carbohydrates and essential nutrients to millions [3]. Its resilience against adverse environmental conditions and ability to thrive in diverse soil types further contribute to its agricultural prominence. Additionally, cassava's utilization extends beyond its nutritional value, as it plays a pivotal role in the industrial sector, serving as a key ingredient in various food products and other commodities [3-5]. However, despite its significant contributions to food security, cassava faces challenges in terms of post-harvest losses and susceptibility to pests and diseases, necessitating ongoing research and technological advancements to sustainably harness its potential [2]. Cassava's multifaceted nature and adaptive qualities highlight its profound impact on global agriculture and underscore its continued relevance in addressing the evolving challenges of our times. Thus, it would be significant to understand the growth and development of cassava plants under adverse environmental conditions at the molecular scale.

Trehalose, a remarkable non-reducing disaccharide, plays a crucial role in plants, acting as a protective agent against various environmental stresses [6, 7]. This natural sugar not only serves as a reserve carbohydrate, aiding plants in energy storage [8], but also exhibits remarkable abilities in mitigating detrimental effects caused by drought, extreme temperatures, and other abiotic stresses [7, 9, 10]. Of our interest, the synthesis of trehalose is facilitated by the trehalose-6-phosphate synthase (TPS) enzyme, which catalyzes the formation of trehalose-6-phosphate from glucose-6-phosphate and UDP-glucose [11]. TPS is an essential regulator of trehalose levels in plants, playing a pivotal role in various biological processes, including embryo development, senescence regulation, flower induction modulating responses to stressors, and facilitating adaptive mechanisms [11, 12]. Furthermore, recent research has unveiled the intricate involvement of trehalose and the TPS enzyme in enhancing plant tolerance to various stress conditions, holding promising implications for agricultural practices aimed at bolstering crop resilience and productivity in the face of an ever-changing climate [12, 13]. However, the information on the TPS family in cassava has not been fully reported. Recently, to explain the trehalose contents in cassava plants, biosynthesis pathway gene families, including TPS, trehalose-6-phosphate phosphatase, and trehalase genes have been recorded [14]. However, little information on the TPS family in cassava has been found, of which 12 members of the TPS family were identified. It would be very significant to get insight into this important gene family based on the newest cassava assemblies published in 2017 [15].

This study aimed to identify and characterize the TPS family in cassava. We identified and annotated all putative members of the TPS family in the assembly of cassava. The common characteristics of the TPS proteins were then analyzed based on various bioinformatics tools. Of our interest, the expression profiles of genes encoding the TPS family were explored in various organs/tissues.

2. Content

2.1. Materials and methods

2.1.1. Materials

Recent assemblies of cassava reported in the previous study [14] were obtained from Phytozome [15] and NCBI servers.

Transcriptome atlas of cassava in various tissues/organs during the growth and development reported in the previous study [16] were collected from the NCBI GEO server [17].

Well-characterized TPS protein in rice [18], particularly HM050424 (Locus ID: LOC_Os05g44210) was downloaded from Phytozome [15] and NCBI servers for further analysis.

2.1.2. Methods

Survey of the TPS proteins in cassava: Well-characterized TPS protein in rice [18] was used as seed sequences for screening against the cassava assembly [14] as previously described [19]. Particularly, full-length HM050424 was BlastP-ed against the cassava proteome [14] available in the Phytozome server [15]. All results (E-value ≤ 1.00e-10) were validated by the Pfam tool [20]. Finally, the protein sequence of each putative TPS member was searched against the NCBI to obtain identifiers and gene sequences.

Estimation of features of the TPS proteins in cassava: The protein sequence of each member of the TPS family in cassava was used to calculate several parameters of the molecule as previously described [19]. Briefly, the full-length protein sequence of the TPS protein was subjected to the Expasy tool [21] to analyze the physical-chemical properties. Four general characteristics of the protein, including size, mass, theoretical isoelectric point (pI), and grand average of hydropathy (GRAVY) were consequently analyzed.

Phylogenetic analysis of the TPS proteins in cassava: All protein sequences of the TPS members in cassava were used to generate a phylogenetic tree as previously reported [19]. Particularly, MEGA software [22] was applied to construct an unrooted phylogenetic tree of TPS proteins by the Neighbor-Joining algorithm. The phylogenetic tree was then illustrated by the Adobe Illustrator software.

Expression analysis of the TPS genes in cassava: The expression profiles of the *TPS* genes in major organs were re-analyzed by the NCBI GEO [17]. Particularly, we used the GSE82279 dataset as previously provided [16] to investigate the expression levels of the *TPS* genes in the leaf blade, leaf midvein, petiole, stem, lateral bud, storage root, and fibrous root tissues obtained in 3-month-old TME 204 cassava plants. The fragments per kilobase of transcript per million reads mapped (FPKM) scores were counted to estimate the transcriptional changes of each *TPS* gene. An easy-to-use R package was applied to generate a heatmap of the expression levels of the *TPS* genes.

2.2. Results and discussion

2.2.1. Identification and annotation of the TPS family in cassava

To survey the TPS proteins in cassava, we used TPS protein in rice [18] to perform a comprehensive search against the cassava assembly [14] in Phytozome [15] and NCBI servers. After validation by the Pfam tool [20], 12 putative TPS proteins, namely MeTPS were found in the cassava assembly. According to the chromosomal distribution, 12 *MeTPS* genes were localized in the genome at an uneven rate (Figure 1). Particularly, chromosome 15 had the highest number of *MeTPS* genes, including *MeTPS06*, *MeTPS07, MeTPS08,* and *MeTPS09* (Figure 1). Chromosome 17 had two members of

the *MeTPS* gene family, like *MeTPS11* and *MeTPS12*, while chromosomes 01, 03, 05, 06, 14, and 16 had only one *MeTPS* gene each (Figure 1).

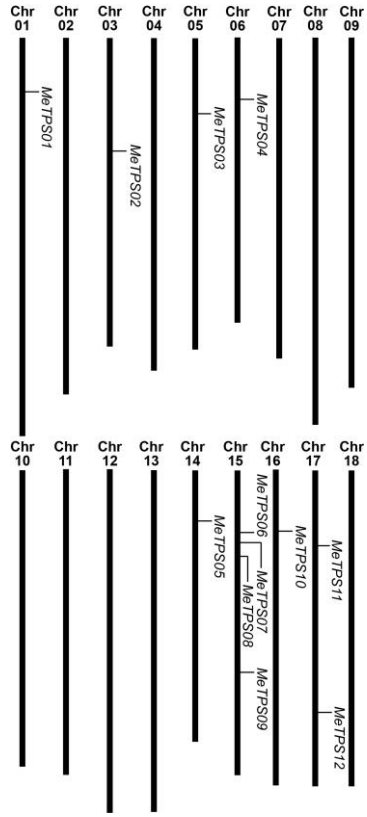

Figure 1. Physical distribution of the MeTPS genes in the cassava genome

Previously, great efforts have been made to identify the TPS families in plant species. Specifically 13 *TPS* genes have been found in the *Populus tomentosa* genome [23], while 7 and 8 members of the *TPS* gene family were recorded in cucumber (*Cucumis sativus*) and sweet orange (*Citrus sisnensis*) [24, 25]. Additionally, 11 members of the TPS gene family have been identified in rice (*Oryza sativa*) [18]. Recently, 11 *TPS* genes have been annotated in *Medicago truncatula*, whereas 26 putative genes have been well-annotated in *Brassica napus* [26, 27]. Taken together, our study revealed that the members of the TPS family in plant species were greatly variable.

2.2.2. Analysis of characteristics categorization of the TPS family in cassava

To investigate the features of the MeTPS proteins in cassava, the full-length protein sequence of each member was used to analyze the size, mass, pI, and GRAVY scores. As a result, four typical characteristics were well-described in Figures 2A and 2B. Firstly, we found that the size and mass of the MeTPS proteins in cassava ranged from 853 (MeTPS10) to 928 residues (MeTPS07) and 96.04 (MeTPS10) to 104.71 kDa (MeTPS07), respectively (Figure 2A). According to the pI scores, all MeTPS proteins were acidic (pI less than 7.0), ranging from 5.43 (MeTPS06) to 6.73 (MeTPS07) (Figure 2B). Finally, the GRAVY scores of all MeTPS proteins were less than 0, ranging from -0.17 (MeTPS01 and MeTPS11) to -0.40 (MeTPS07) (Figure 2B). It has been realized that the MeTPS proteins were hydrophilic molecules.

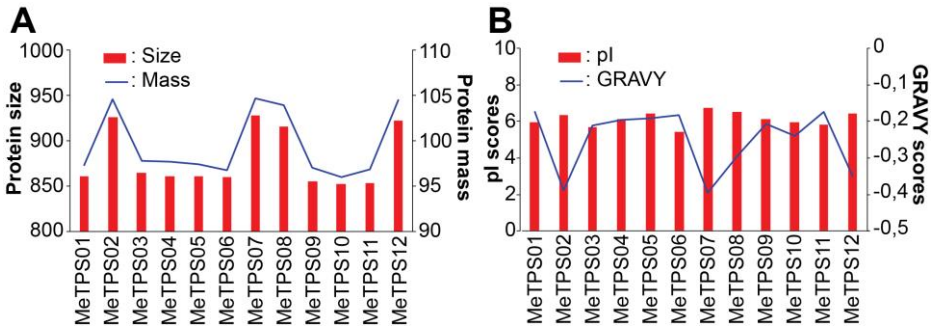

Figure 2. Physic-chemical properties, including (A) size and mass, (B) pI and GRAVY scores of the MeTPS proteins in cassava

In recent studies, the common features of the TPS families have been analyzed in several plant species. For example, the TPS proteins in *Populus* spp. varied from 846 to 922 residues, while all proteins were relatively hydrophilic [23]. Recently, the TPS proteins in sugarcane (*Saccharum* spp. hybrids) varied from 818 to 976 amino acid residues in size and from 91.05 to 108.39 kDa in weight [28]. The pI scores of these TPS proteins were acidic (pH less than 7.0) and hydrophilic (GRAVY less than 0) [28]. The length of cucumber TPS proteins was between 850 and 971 residues, and their mW was between 96.31 and 110.33 kDa [25]. Most of the cucumber TPS proteins were acidic [25]. Taken together, our comparisons suggested that the TPS proteins in plant species shared slightly variable characteristics.

Next, to focus on the relationship of the MeTPS family in cassava, we constructed a Neighbor-Joining-based phylogenetic tree of whole members. As provided in Figure 3, the MeTPS family in cassava could be classified into two groups, groups 1 and 2. Particularly, group 1 contained two sub-groups, including clade 1A (MeTPS09, MeTPS10, MeTPS11 and MeTPS16) and clade 1B (MeTPS01, MeTPS03, MeTPS04 and MeTPS05) (Figure 3). There are two sub-groups found in group 2, namely clade 2A (MeTPS02 and MeTPS07) and clade 2B (MeTPS08 and MeTPS12) (Figure 3). The categorization of members of the MeTPS family in cassava was confirmed by previous studies. Recently, the TPS proteins in five species, including *Populus*, rice, *Arabidopsis*, soybean and Medicago have been generated and revealed a categorization similar to the MeTPS family in cassava [23].

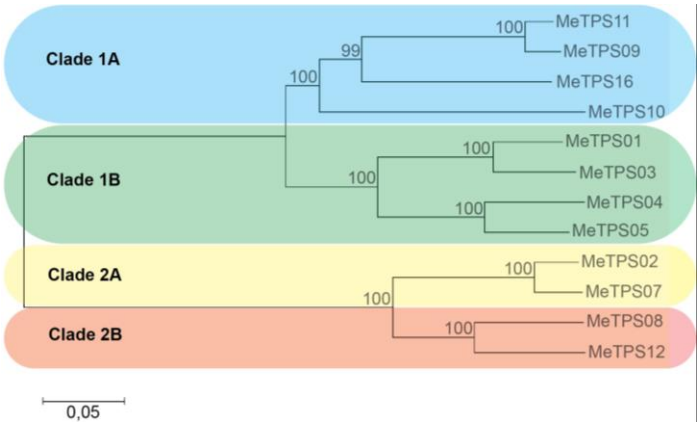

Figure 3. Phylogenetic tree of the MeTPS proteins in cassava

2.2.3. Expression analysis of the TPS family in cassava

As a main part of this study, the expression patterns of the *MeTPS* genes in various organs during the growth and development of cassava plants have been investigated. According to the GSE82279 dataset [16], the expression levels of the *MeTPS* genes were differentially expressed in seven major organs (Figure 4).

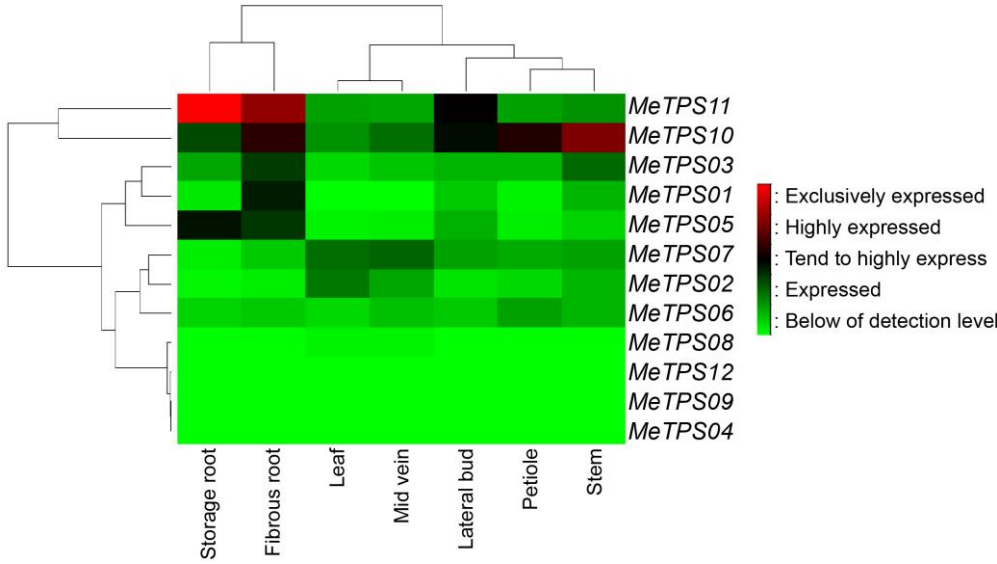

Figure 4. Expression patterns of the MeTPS genes in various major organs in cassava

Two genes, including *MeTPS01* and *MeTPS03,* tended to be highly expressed in fibrous root samples (Figure 4). Particularly, one gene, *MeTPS05* was found to be differentially expressed in storage roots and fibrous roots (Figure 4). Interestingly, we found that two genes *MeTPS10* and *MeTPS11* exhibited significant expression levels in various organs (Figure 4). Particularly, *MeTPS10* was thought to exhibit a strong expression in four major organs, including fibrous root, lateral bud, petiole, and stem (Figure 4). Moreover, another gene, *MeTPS11* was exclusively expressed in storage root tissues and highly expressed in fibrous root and lateral bud samples (Figure 4).

Previously, functions of the *TPS* genes in plants during the growth and development processes were recorded. Trehalose-6-phosphate has been demonstrated to influence plant starch metabolism [11]. Exogenous trehalose causes starch accumulation in *Arabidopsis* by boosting the activity of ADP-glucose pyrophosphorylase, a key enzyme in starch synthesis [29]. Avonce *et al.* (2004) demonstrated that the *AtTPS1* gene controls glucose content, abscisic acid, and stress signaling in *Arabidopsis* [30]. Furthermore, *OsTPS1 gene* overexpression improved transgenic rice tolerance to stressors such as salinity, drought, and cold by increasing trehalose and proline content and modulating the expression of stress-related genes [31]. Meanwhile, the *TPS* genes in sweet orange exhibited different expression levels in various tissues during growth and development, suggesting that they may be involved in the development of these organs [24, 25]. Here, we hypothesized that two *TPS* genes, like *MeTPS10* and *MeTPS11,* may play a role in the root tissues related to the starch metabolism in Vietnamese cassava cultivars.

3. Conclusions

This study identified, annotated, and characterized the *MeTPS* genes from the cassava genome. As a result, 12 members of the *MeTPS* genes were found in cassava, which were categorized into two distinct groups. Based on various tools, we found that these MeTPS proteins shared high similarity in general physic-chemical properties. According to the previous transcriptome atlas, the *MeTPS* genes were significantly expressed in seven major organs in cassava plants during the growth and development processes. In summary, this study could provide basic information for the further understanding of this important enzyme family in cassava.

REFERENCES

[1] P. Chavarriaga-Aguirre, A. Brand, A. Medina, M. Prías, R. Escobar, J. Martinez, P. Díaz, C. López, M. Roca, J. Tohme, 2016. The potential of using biotechnology to improve cassava: A review. *In Vitro Cell Dev. Biol. Plant*, 52(5), pp. 461-478.

[2] A. Malik, P. Kongsil, V. Nguyen, W. Ou, P. Srean, M. Sheela, L. Becerra, Y. Utsumi, C. Lu, P. Kittipadakul, H. Nguyen, H. Ceballos, T. Nguyen, M. Selvaraj Gomez, P. Aiemnaka, R. Labarta, S. Chen, S. Amawan, S. Sok, L. Youabee, M. Seki, H. Tokunaga, W. Wang, K. Li, H. Nguyen, V. Nguyen, H. Ham, M. Ishitani, 2020. Cassava breeding and agronomy in Asia: 50 years of history and future directions. *Breed Sci.*, 70(2), pp. 145-166.

[3] S. Li, Y. Cui, Y. Zhou, Z. Luo, J. Liu, M. Zhao, 2017. The industrial applications of cassava: current status, opportunities, and prospects. *J. Sci. Food Agric.*, 97(8), pp. 2282-2290.

[4] R. J. Hillocks, J. M. Thresh, 2002. Cassava: biology, production, and utilization. Wallingford, UK; New York: CABI Pub, 332 pages.

[5] R. Nilusha, J. Jayasinghe, O. Perera, P. Perera, C. Jayasinghe, 2021. Proximate composition, physicochemical, functional, and antioxidant properties of flours from selected cassava (*Manihot esculenta* Crantz) varieties. *Int J. Food Sci.*, 6064545, pp. 1-13.

[6] J. Cervantes-Chávez, L. Valdés-Santiago, G. Bakkeren, E. Hurtado-Santiago, C. León-Ramírez, E. Esquivel-Naranjo, F. Landeros-Jaime, Y. Rodríguez-Aza, J. Ruiz-Herrera, 2016. Trehalose is required for stress resistance and virulence of the *Basidiomycota* plant pathogen *Ustilago maydis*. *Microbiol*, 162(6), pp. 1009-1022.

[7] B. Williams, 2015. Trehalose accumulation triggers autophagy during plant desiccation. *PLoS Genet*, 11(12), pp. e1005705.

[8] L. O'Hara, et al., 2013. How do sugars regulate plant growth and development? New insight into the role of trehalose-6-phosphate. *Mol. Plant*, 6(2), pp. 261-274.

[9] M. Paul, et al., 2018. The role of trehalose 6-phosphate in crop yield and resilience. *Plant Physiol*, 177(1), pp. 12-23.

[10] J. Vilchez, et al., 2016. Plant drought tolerance enhancement by trehalose production of desiccation-tolerant microorganisms. *Front. Microbiol*, 7, pp. 1577.

[11] J. Ponnu, et al., 2011. Trehalose-6-phosphate: connecting plant metabolism and development. *Front. Plant Sci.*, 2, pp. 70.

[12] A. Grennan, 2007. The role of trehalose biosynthesis in plants. *Plant Physiol*, 144(1), pp. 3-5.

[13] M. Iordachescu, et al., 2008. Trehalose biosynthesis in response to abiotic stresses. *J. Integr. Plant Biol.*, 50(10), pp. 1223-1229.

[14] J. Bredeson, et al., 2016. Sequencing wild and cultivated cassava and related species reveals extensive interspecific hybridization and genetic diversity. *Nat. Biotechnol*, 34(5), pp. 562-570.

[15] D. Goodstein, et al., 2012. Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res.*, 40, pp. D1178-D1186.

[16] M. Wilson, et al., 2017. Gene expression atlas for the food security crop cassava. *New Phytol.*, 213(4), pp. 1632-1641.

[17] T. Barrett, et al., 2013. NCBI GEO: archive for functional genomics data sets - Update. *Nucleic Acids Res.*, 41, pp. D991-D995.

[18] B. Zang, et al., 2011. Analysis of the trehalose-6-phosphate synthase (TPS) gene family suggests the formation of TPS complexes in rice. *Plant Mol. Biol.*, 76(6), pp. 507-522.

[19] H. La, et al., 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, pp. 146210.

[20] J. Mistry, et al., 2021. Pfam: The protein families database in 2021. *Nucleic Acids Res.*, 49(D1), pp. D412-D419.

[21] E. Gasteiger, et al., 2005. Protein identification and analysis tools on the ExPASy server, in *The proteomics protocols handbook*, Springer, pp. 571-607.

[22] K. Tamura, et al., 2021. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Mol. Biol. Evol.*, 38(7), pp. 3022-3027.

[23] Y. Gao, et al., 2021. Characterization and expression pattern of the trehalose-6-phosphate synthase and trehalose-6-phosphate phosphatase gene families in *Populus*. *Int. J. Biol. Macromol*, 187, pp. 9-23.

[24] K. Liu, et al., 2022. Genome-wide identification of the trehalose-6-phosphate synthase gene family in sweet orange (*Citrus sinensis*) and expression analysis in response to phytohormones and abiotic stresses. *Peer J*, 10, pp. e13934.

[25] Y. Dan, et al., 2021. Genome-wide identification and expression analysis of the trehalose-6-phosphate synthase (TPS) gene family in cucumber (*Cucumis sativus* L.). *Peer J*, 9, pp. e11398.

[26] M. Hu, et al., 2022. Genome-wide characterization of trehalose-6-phosphate synthase gene family of *Brassica napus* and potential links with agronomic traits. *Int. J. Mol. Sci.*, pp. 23.

[27] J. Song, et al., 2021. Identification of the trehalose-6-phosphate synthase gene family in *Medicago truncatula* and expression analysis under abiotic stresses. *Gene*, 787, pp. 145641.

[28] X. Hu, et al., 2020. Genome-wide analysis of the trehalose-6-phosphate synthase (TPS) gene family and expression profiling of ScTPS genes in sugarcane. *Agronomy*, 10(7), pp. 969.

[29] A. Wingler, et al., 2000. Trehalose induces the ADP-glucose pyrophosphorylase gene, ApL3, and starch synthesis in *Arabidopsis*. *Plant Physiol.*, 124(1), pp. 105-114.

[30] N. Avonce, B. Leyman, J. Mascorro-Gallardo, P. Dijck, J. Thevelein, G. Iturriaga, 2004. The Arabidopsis trehalose-6-P synthase AtTPS1 gene is a regulator of glucose, abscisic acid, and stress signaling. *Plant Physiol.*, 136, pp. 3649-3659.\

[31] H. Li, B. Zang, X. Deng, X. Wang, 2011. Overexpression of the trehalose-6-phosphate synthase gene OsTPS1 enhances abiotic stress tolerance in rice. *Planta*, 234, pp. 1007-1018.