

Xác định, phân tích đặc tính và đánh giá biểu hiện đặc thù của nhóm gene mã hóa nhân tố phiên mã Platz liên quan đến sinh trưởng và phát triển ở cây sắn (*Manihot esculenta*)

Hoàng Minh Chính<sup>1,2</sup>, Bùi Thị Thu Hương<sup>2</sup>, Phạm Phương Thu<sup>3</sup>,  
Lê Thị Ngọc Quỳnh<sup>4</sup>, Đồng Huy Giới<sup>2</sup>, Chu Đức Hà<sup>1\*</sup>

<sup>1</sup>Trường Đại học Công nghệ, Đại học Quốc gia Hà Nội

<sup>2</sup>Học viện Nông nghiệp Việt Nam

<sup>3</sup>Trường Đại học Sư phạm Hà Nội 2

<sup>4</sup>Trường Đại học Thủy lợi

Identification, characterization and analysis of expression patterns of genes encoding the Platz transcription factor related to the growth and development in cassava (*Manihot esculenta*)

Hoang Minh Chinh<sup>1,2</sup>, Bui Thi Thu Huong<sup>2</sup>, Pham Phuong Thu<sup>3</sup>,  
Le Thi Ngoc Quynh<sup>4</sup>, Dong Huy Gioi<sup>2</sup>, Chu Duc Ha<sup>1\*</sup>

<sup>1</sup>University of Engineering and Technology, Vietnam National University Hanoi

<sup>2</sup>Vietnam National University of Agriculture

<sup>3</sup>Hanoi Pedagogical University 2

<sup>4</sup>Thuyloi University

\*Corresponding author: cd.ha@vnu.edu.vn

<https://doi.org/10.55250/jo.vnuf.13.1.2024.003-011>

**TÓM TẮT**

Nhân tố phiên mã (transcription factor - TF) giàu AT và mang vùng bảo thủ đặc trưng cánh tay kềm đặc trưng ở thực vật (plant AT-rich sequences and zinc-binding, PLATZ) là một trong những nhóm protein điều hòa quan trọng liên quan đến quá trình sinh trưởng và phát triển ở cây trồng. Tuy nhiên, thông tin về nhóm TF PLATZ trên cây sắn (*Manihot esculenta*) vẫn chưa được ghi nhận, mặc dù genome của cây sắn đã được công bố gần đây. Trong nghiên cứu này, nhóm TF PLATZ ở cây sắn đã được xác định, định danh và phân tích bằng cách sử dụng các công cụ tin sinh học. Kết quả đã cho thấy, tổng số 20 thành viên của nhóm TF PLATZ đã được báo cáo trên dữ liệu của cây sắn. Phân tích trình tự protein đầy đủ cho thấy nhóm TF PLATZ ở cây sắn có kích thước từ 182 - 277 gốc amino acid, trọng lượng từ 21,36 - 30,75 kDa. Đa số thành viên trong nhóm TF PLATZ là các protein có tính base và không ổn định trong điều kiện ống nghiệm. Chỉ số độ ưa nước trung bình của các protein đều < 0, chứng tỏ nhóm TF PLATZ ở cây sắn có tính ưa nước. Tiếp theo, phân tích dữ liệu microarray cho thấy các gene mã hóa nhóm TF PLATZ có biểu hiện đa dạng ở các mẫu mô và cơ quan/bộ phận chính trên cây sắn. Tóm lại, kết quả của nghiên cứu này đã cung cấp những thông tin cơ bản về nhóm TF PLATZ ở cây sắn, đồng thời cung cấp một danh sách các gene ứng viên nhằm phân tích chức năng liên quan đến các quá trình sinh lý trong nghiên cứu tiếp theo.

**ABSTRACT**

Plant AT-rich sequences and zinc-binding (PLATZ) transcription factor (TF) have been regarded as one of the most important regulatory proteins that are related to the growth and development of crop species. Unfortunately, the TF PLATZ in cassava (*Manihot esculenta*) has not been reported, even though the genome of cassava plants was released recently. In this study, the TF PLATZ in cassava has been comprehensively identified, annotated and characterized by using various bioinformatics tools. As a result, a total of 20 members of the TF PLATZ have been found in the assembly of cassava. Analysis of full-length

**Thông tin chung:**

Ngày nhận bài: 23/10/2023

Ngày phản biện: 22/11/2023

Ngày quyết định đăng: 15/12/2023

**Từ khóa:**

biểu hiện, đặc tính, nhân tố phiên mã, PLATZ, sắn.

**Keywords:**

Cassava, characteristic, expression pattern, PLATZ, transcription factor.

protein sequences revealed that the TF PLATZ in cassava was varied from 182 - 277 amino acid residues in lengths and 21.36 - 30.75 kDa in masses. The majority of members of the TF PLATZ were basic proteins and unstable in the test tube. The scores of the grand average of hydropathy of all members were < 0, indicating that the TF PLATZ in cassava contained hydrophilic proteins. Next, microarray analysis indicated that genes encoding the TF PLATZ exhibited differential expression levels in various major tissues and organs in cassava plants. Taken together, our results could provide general information on the TF PLATZ in cassava, as well as a list of potential candidate genes for further functional characterization related to the physiological processes.

## 1. ĐẶT VẤN ĐỀ

Cây sắn (*Manihot esculenta*), một loài bản địa của Nam Mỹ [1, 2], đã trở thành một phần không thể thiếu của nền nông nghiệp ở nhiều quốc gia trên thế giới. Với nguồn gốc từ khu vực Đông Nam Brazil và phía Đông của dãy núi Andes [1, 2], cây sắn đã lan rộng khắp các lục địa từ châu Á, châu Phi đến các vùng nhiệt đới của châu Âu, nhờ khả năng thích nghi vượt trội với các điều kiện môi trường đa dạng [1, 3]. Sắn là nguồn cung cấp tinh bột lớn, đây là nguồn nguyên liệu quan trọng trong chế biến thực phẩm, và đồng thời là món ăn cơ bản cho hơn 800 triệu người trên toàn cầu [4, 5]. Bên cạnh đó, củ sắn cũng có thể được sử dụng trong chế biến thức ăn chăn nuôi và sản xuất nhiên liệu sinh học [5, 6]. Do đó, tìm hiểu về quá trình sinh trưởng và phát triển của cây sắn ở cấp độ phân tử được xem là chìa khóa để cải thiện kỹ thuật canh tác và nâng cao năng suất.

Ở thực vật, nhân tố phiên mã (transcription factor - TF) đóng vai trò trung tâm trong điều hòa các phản ứng sinh lý diễn ra trong suốt quá trình sinh trưởng và phát triển [7]. Thông qua việc bám trên vùng promoter, TF giúp điều hòa (kìm hãm hoặc thúc đẩy) sự biểu hiện của gene mục tiêu thông qua con đường tín hiệu phytohormone, giúp tế bào trả lời lại kích thích từ môi trường. Trong đó, cánh tay kẽm (zinc finger) là một trong những họ protein chức năng đặc trưng cho thực vật, gồm nhiều nhóm TF quan trọng, như nhân tố cảm ứng ethylene (ethylene responsive factor), nhân tố có vùng bảo thủ đặc trưng homeobox thực vật (plant homeobox domain), nhân tố có vùng bảo thủ WRKYGQK và nhân tố giàu AT và mang vùng bảo thủ đặc trưng cánh tay kẽm (plant AT-rich

sequences and zinc-binding, PLATZ). Các nghiên cứu gần đây đã ghi nhận vai trò của nhóm TF PLATZ tham gia vào các quá trình sinh học quan trọng trên các loài thực vật. Ví dụ, gene *ZmPLATZ12*, một thành viên trong nhóm TF PLATZ ở cây ngô (*Zea mays*), có biểu hiện đặc thù ở tế bào nội nhũ chứa tinh bột, được chứng minh là kiểm soát quá trình phát triển nội nhũ thông qua tương tác với phức hệ RNA polymerase III [8]. Một gene mã hóa TF PLATZ khác trên cây mô hình hai lá mầm *Arabidopsis thaliana* được ghi nhận tham gia vào cơ chế thúc đẩy sự biểu hiện của các gene liên quan đến kích thước và sự úa vàng ở lá [9]. Gần đây, nhóm TF PLATZ đã được nghiên cứu trên một số đối tượng cây trồng, như *Arabidopsis* [10], ngô [8], lúa gạo (*Oryza sativa*) [11], lúa mì (*Triticum aestivum*) [12] cải bắp Trung Quốc (*Brassica rapa*) [13], các loài thuộc họ Hoa hồng (Rosaceae), bao gồm táo tây (*Malus domestica*) thuộc giống Golden delicious [14], táo tây thuộc giống Hanfu [15], táo tây thuộc giống Gala [15], táo dại Tân Cương (*M. sieversii*) [15], táo dại Siberi (*M. baccata*) [15], táo dại châu Âu (*M. sylvestris*) [15], dâu tây dại (*Fragaria vesca*) [15], đào (*Prunus persica*) [15] và lê châu Âu (*Pyrus communis*) [15], bạch quả (*Ginkgo biloba*) [16].

Mục tiêu của nghiên cứu này nhằm sàng lọc và xác định nhóm TF PLATZ trong dữ liệu của cây sắn dựa trên các phân tích tin sinh học. Đầu tiên, các ứng viên của nhóm TF PLATZ được sàng lọc dựa trên việc đối chiếu trình tự tương đồng trên proteome của cây sắn. Các thông tin chú giải gene sau đó đã được xác định và kiểm chứng. Tiếp theo, nghiên cứu đã tiến hành phân tích các đặc điểm cơ bản của protein ở các

thành viên trong nhóm TF PLATZ ở cây sắn. Cuối cùng, mức độ biểu hiện của các gene mã hóa nhóm TF PLATZ ở cây sắn được khai thác dựa trên dữ liệu transcriptome ở các vị trí cơ quan/bộ phận chính trên cây sắn nhằm gợi mở các giả thuyết về chức năng của chúng trong quá trình sinh trưởng và phát triển.

## 2. PHƯƠNG PHÁP NGHIÊN CỨU

### 2.1. Dữ liệu nghiên cứu

Thông tin di truyền, bao gồm genome và proteome của giống sắn AM560-2 [17] được khai thác trên giao diện Phytozome (website: <https://phytozome-next.jgi.doe.gov/>) [18] và Trung tâm Quốc gia về Thông tin Công nghệ Sinh học của Hoa Kỳ (National Center for Biotechnology Information, NCBI, website: <https://www.ncbi.nlm.nih.gov/>) cho các phân tích tin sinh học trong nghiên cứu này.

Trình tự protein đầy đủ của TF PLATZ trên cây táo mô tả trong nghiên cứu trước đây [14] được khai thác để tiến hành tìm kiếm và sàng lọc trên dữ liệu của cây sắn.

Dữ liệu phiên mã transcriptome của cây sắn, bao gồm GSE82279 [19] được khai thác trên giao diện Dữ liệu biểu hiện gene (Gene Expression Omnibus, GEO, website: <https://www.ncbi.nlm.nih.gov/geo/>) của NCBI [20].

### 2.2. Phương pháp sàng lọc nhóm PLATZ

Sàng lọc và xác định nhóm PLATZ ở cây sắn được tiến hành dựa theo mô tả trong nghiên cứu trước đây [21]. Theo đó, trình tự protein đầy đủ của TF PLATZ trên cây táo, bao gồm MD00G1059100 và MD02G1017000 báo cáo gần đây [14] được đối chiếu trên proteome của cây sắn [17] trên Phytozome [18] bằng công cụ BlastP. Tất cả kết quả tìm kiếm có độ tin cậy E-value  $\leq 10^{-10}$  được lựa chọn để kiểm chứng sự có mặt của vùng bảo thủ đặc trưng cho TF PLATZ (mã định danh vùng bảo thủ: IPR006734) [22] bằng công cụ InterPro (website: <https://www.ebi.ac.uk/interpro/>) [23]. Trình tự protein đầy đủ của các kết quả tìm kiếm được khai thác cho các phân tích *in silico* tiếp theo.

### 2.3. Phương pháp định danh nhóm PLATZ

Các bước định danh nhóm TF PLATZ ở cây

sắn được tiến hành dựa theo báo cáo trước đây. Cụ thể, các trình tự protein đầy đủ lần lượt được kiểm tra các thông tin chú giải trên genome của cây sắn [17] cung cấp bởi NCBI. Các thông tin chú giải, bao gồm GeneID, ProteinID, mã định danh RNA và LocusID được thu thập. Trình tự đoạn mã hóa (coding DNA sequence, CDS), trình tự vùng gene (genomic DNA sequence, gDNA) và vị trí phân bố trên nhiễm sắc thể được khai thác cho các phân tích.

### 2.4. Phương pháp phân tích đặc tính nhóm PLATZ

Phân tích đặc tính protein của nhóm TF PLATZ ở cây sắn được tiến hành dựa theo báo cáo trước đây. Cụ thể, các trình tự protein đầy đủ của các thành viên của nhóm TF PLATZ được truy vấn trên giao diện ExPasy ProtParam (website: <https://web.expasy.org/cgi-bin/protparam/protparam>) như báo cáo của Gasteiger và cộng sự (2003) [24]. Đặc tính cơ bản, bao gồm kích thước protein (gốc amino acid), trọng lượng protein (kilo Dalton, kDa), điểm đẳng điện lý thuyết ( $< 7$  - tính acid,  $\sim 7$  - trung tính và  $> 7$  - tính base), độ bất ổn định ( $> 40$  - bất ổn định và  $< 40$  - ổn định), chỉ số béo và độ ưa nước trung bình ( $< 0$  - ưa nước và  $> 0$  - kỵ nước).

### 2.5. Phương pháp đánh giá biểu hiện gene của nhóm PLATZ

Phân tích mức độ biểu hiện của các gene mã hóa nhóm TF PLATZ ở cây sắn được tiến hành dựa trên các công cụ tin sinh học dữ liệu lớn như mô tả trong nghiên cứu trước đây. Cụ thể, dữ liệu GSE82279 [19] khai thác trên GEO NCBI [20] được sử dụng để phân tích mức độ biểu hiện của gene tại các mẫu mô và cơ quan/bộ phận chính trên cây sắn. Các mẫu được thu thập bao gồm mô sẹo phôi hóa (friable embryogenic callus, FEC), tổ chức phát sinh phôi cấu tạo soma (somatic organized embryogenic structures, OES), mô phân sinh chóp rễ (root apical meristem, RAM) và mô phân sinh đỉnh chồi (shoot apical meristem, SAM) [19]. Các vị trí cơ quan/bộ phận chính trên cây sắn được thu thập bao gồm lá, gân lá, cuống lá, rễ sợi, củ, chồi bên và thân [19].

## 2.6. Phương pháp xử lý số liệu

Dữ liệu biểu hiện của các gene được thể hiện bằng giá trị của đoạn mã hóa trên mỗi kb cho một đơn vị phiên mã trên mỗi triệu bản đọc (fragments per kilobase of transcript per million reads mapped, FPKM). Theo đó, biểu hiện của gene dưới ngưỡng xác định hoặc gene có biểu hiện yếu khi giá trị FPKM < 5, gene có xu hướng biểu hiện khi 5 < giá trị FPKM < 20. Trong khi gene có xu hướng biểu hiện mạnh và biểu hiện đặc thù khi giá trị FPKM > 50 và > 100. Các dữ liệu biểu hiện được xử lý và xếp hạng bằng thuật toán cluster trên ngôn ngữ lập trình

R [25].

## 3. KẾT QUẢ VÀ THẢO LUẬN

### 3.1. Xác định và định danh thông tin của nhóm TF PLATZ ở cây sắn

Để sàng lọc và xác định nhóm PLATZ ở cây sắn, trình tự của TF PLATZ trên cây táo [14] được khai thác để tiến hành so sánh trên dữ liệu của cây sắn [17]. Sau khi kiểm chứng trên InterPro [23], các kết quả đã được truy vấn để định danh thông tin chú giải. Kết quả định danh và chú giải thông tin của nhóm TF PLATZ trong dữ liệu của cây sắn được mô tả trong Bảng 1.

**Bảng 1. Định danh và chú giải của nhóm TF PLATZ từ dữ liệu của cây sắn**

STT	GeneID	ProteinID	Mã định danh RNA	LocusID
1	<i>Manes.01G121150</i>	-	-	-
2	<i>Manes.01G268300</i>	XP_021631323	XM_021775631.2	LOC110628818
3	<i>Manes.03G075532</i>	XP_021602542	XM_021746850.2	LOC110607704
4	<i>Manes.03G075500</i>	XP_021608834	XM_021753142.2	LOC110612369
5	<i>Manes.05G130400</i>	XP_021614340	XM_021758650.2	LOC110616275
6	<i>Manes.05G048200</i>	XP_021614143	XM_021758451.2	LOC110616128
7	<i>Manes.06G043200</i>	XP_021616867	XM_021761175.2	LOC110618121
8	<i>Manes.08G016000</i>	XP_021619379	XM_021763687.2	LOC110620106
9	<i>Manes.08G064000</i>	XP_021619561	XM_021763869.2	LOC110620225
10	<i>Manes.09G048606</i>	XP_021602886	XM_043959687.1	LOC110608000
11	<i>Manes.09G128900</i>	XP_021621856	XM_021766164.2	LOC110621886
12	<i>Manes.11G153600</i>	XP_021629155	XM_021773463.2	LOC110627192
13	<i>Manes.14G120600</i>	XP_021634141	XM_021778449.2	LOC110630842
14	<i>Manes.15G136800</i>	XP_021595879	XM_021740187.2	LOC110602620
15	<i>Manes.15G188400</i>	XP_021595641	XM_021739949.1	LOC110602427
16	<i>Manes.16G052100</i>	XP_021596936	XM_021741244.2	LOC110603496
17	<i>Manes.17G083200</i>	XP_021598899	XM_021743207.1	LOC110604921
18	<i>Manes.17G085975</i>	XP_021599857	XM_021744166.2	LOC110605544
19	<i>Manes.17G011500</i>	XP_021600027	XM_021744335.2	LOC110605688
20	<i>Manes.18G000750</i>	XP_021600277	XM_021744586.2	LOC110605905

Ghi chú: dấu "-" là "Không xác định".

Kết quả cho thấy, đa số (19 trên tổng số 20) thành viên của nhóm TF PLATZ ở cây sắn, ngoại trừ gene *Manes.01G121150* chưa có thông tin chú giải. Kết quả này được giải thích do bản tham chiếu mới nhất của cây sắn (mã số dự án giải mã genome trên NCBI: GCF\_001659605.2) mới chỉ bao phủ khoảng 95,7% số lượng gene mã hóa protein thực tế [17]. Cụ thể, Bredeson và cộng sự (2016) ghi nhận dữ liệu genome hiện tại của cây sắn có kích thước khoảng 639,6 Mb [17], trong khi genome thực tế của cây sắn được dự đoán có kích thước xấp xỉ 772 Mb [26].

Vì vậy, thông tin chú giải của gene *Manes.01G121150*, bao gồm ProteinID, mã định danh RNA và LocusID được dự đoán có thể mô tả đầy đủ trong phiên bản cập nhật tiếp theo của genome cây sắn.

Gần đây, nhóm TF PLATZ đã được mô tả trong các loài cây trồng khác nhau [10, 13, 14]. Cụ thể, tổng số 17 thành viên của TF PLATZ, đặt tên là MdPLATZ, đã được báo cáo trên hệ tham chiếu của loài [14]. Năm 2020, tổng số 24 gene thành viên thuộc nhóm TF PLATZ, từ *BrPLATZ1* đến *BrPLATZ24* đã được báo cáo trên đối tượng

cây bắp cải Trung Quốc [13]. Nhóm TF PLATZ ở Arabidopsis gồm 12 thành viên cũng đã được xác định [10], trong khi số lượng thành viên trong nhóm TF PLATZ ở các loài ngô, lúa gạo và lúa mì đạt lần lượt là 17; 15 và 62 [8, 12]. Ở một số loài trong họ hoa hồng, nhóm TF PLATZ có số lượng thành viên lần lượt từ 14; 17; 17; 18; 19 và 19 ở loài táo tây thuộc giống Hanfu [15], táo tây thuộc giống Golden delicious [14], táo tây thuộc giống Gala [15], táo dại Tân Cương [15], táo dại châu Âu [15] và táo dại Siberi [15]. Các so sánh này cho thấy nhóm TF PLATZ ở các loài thực vật có số lượng thành

viên tương đối đa dạng, không phụ thuộc vào kích thước genome của từng loài.

### 3.2. Phân tích đặc tính cơ bản của nhóm TF PLATZ ở cây sắn

Để phân tích đặc tính protein của nhóm TF PLATZ ở cây sắn, trình tự protein đầy đủ của từng thành viên được truy vấn trên công cụ ExPasy ProtParam [24]. Theo đó, sáu tính chất cơ bản của protein, bao gồm kích thước phân tử, trọng lượng phân tử, điểm đẳng điện, độ bất ổn định, chỉ số béo và độ ưa nước trung bình được tính toán và thể hiện ở Bảng 2.

**Bảng 2. Đặc tính cơ bản của nhóm TF PLATZ của cây sắn**

STT	GeneID	Kích thước	Trọng lượng	Điểm đẳng điện	Độ bất ổn định	Chỉ số béo	Độ ưa nước trung bình
1	Manes.01G121150	190	22,02	9,66	67,54	65,68	-0,69
2	Manes.01G268300	248	27,61	9,08	56,40	84,11	-0,13
3	Manes.03G075532	277	30,72	5,94	60,29	75,63	-0,40
4	Manes.03G075500	277	30,75	5,94	59,62	75,27	-0,41
5	Manes.05G130400	235	26,48	9,27	56,27	69,23	-0,49
6	Manes.05G048200	243	27,17	8,85	56,23	80,62	-0,14
7	Manes.06G043200	244	27,70	8,62	56,23	70,20	-0,45
8	Manes.08G016000	212	24,04	8,33	65,98	82,31	-0,38
9	Manes.08G064000	250	28,04	9,10	58,84	64,76	-0,60
10	Manes.09G048606	216	24,64	9,47	78,96	72,64	-0,55
11	Manes.09G128900	256	28,93	9,04	47,13	65,47	-0,70
12	Manes.11G153600	212	24,45	8,68	80,37	67,50	-0,76
13	Manes.14G120600	240	27,27	8,71	64,75	70,62	-0,43
14	Manes.15G136800	218	25,33	9,19	58,28	67,98	-0,78
15	Manes.15G188400	218	24,84	9,22	60,49	61,70	-0,74
16	Manes.16G052100	272	30,15	6,75	59,76	74,52	-0,38
17	Manes.17G083200	182	21,36	9,91	39,80	72,86	-0,61
18	Manes.17G085975	230	26,58	9,38	66,98	64,00	-0,69
19	Manes.17G011500	221	24,90	9,29	55,57	66,11	-0,61
20	Manes.18G000750	239	26,96	8,92	54,59	68,49	-0,52

Kết quả phân tích kích thước protein của nhóm TF PLATZ ở cây sắn cho thấy, các thành viên đều có chiều dài khá tương đồng, với Manes.17G083200 là protein có kích thước bé nhất (182 gốc amino acid) và hai protein, bao gồm Manes.03G075532 và Manes.03G075500, chia sẻ kích thước lớn nhất (277 gốc amino acid). Tương tự, trọng lượng phân tử của các thành viên trong nhóm TF PLATZ ở cây sắn dao động từ 21,36 (Manes.17G083200) -30,75 kDa (Manes.03G075500), với giá trị trung bình đạt khoảng 26,50 kDa. Phần lớn (17 trên tổng số 20)

các thành viên của nhóm TF PLATZ ở cây sắn đều là các protein có tính base, với giá trị điểm đẳng điện dao động từ 8,33 (Manes.08G016000) - 9,91 (Manes.17G083200), ngoại trừ ba thành viên còn lại là phân tử protein có tính acid, bao gồm Manes.03G075532 (điểm đẳng điện đạt 5,94), Manes.03G075500 (điểm đẳng điện đạt 5,94) và Manes.16G052100 (điểm đẳng điện đạt 6,75). Tiếp theo, giá trị bất ổn định của các thành viên trong nhóm TF PLATZ ở cây sắn phần lớn đều > 40,00, ngoại trừ Manes.17G083200 (độ bất ổn định đạt 39,80), chứng tỏ đa số các

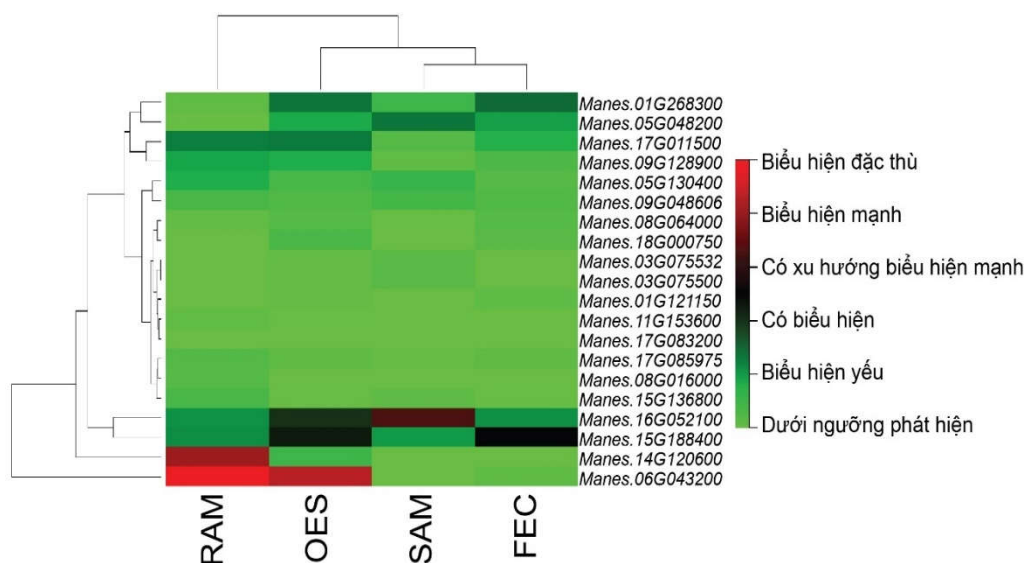
thành viên trong nhóm TF PLATZ ở cây sắn có tính bất ổn định trong ống nghiệm. Phân tích cho thấy chỉ số béo của các protein thuộc nhóm TF PLATZ ở cây sắn có giá trị dao động từ 61,70 (Manes.15G188400) - 84,11 (Manes.01G268300), với chỉ số béo trung bình đạt 74,59. Cuối cùng, kết quả phân tích độ ưa nước trung bình của các thành viên trong nhóm TF PLATZ ở cây sắn cho thấy tất cả các protein này đều có tính ưa nước, với giá trị độ ưa nước < 0, dao động từ -0,78 (Manes.15G136800) - -0,13 (Manes.01G268300).

Có thể thấy rằng, kết quả phân tích đặc tính cơ bản về protein của các thành viên trong nhóm TF PLATZ ở cây sắn cũng tương đồng với những báo cáo trước đây. Cụ thể, nhóm TF PLATZ ở cây cải bắp Trung Quốc có kích thước phân tử dao động từ 165 - 325 gốc amino acid, tương ứng với trọng lượng phân tử đạt từ 19,158 - 36,917 kDa [13]. Trong đó, chỉ có hai (trên tổng số 24) thành viên trong nhóm TF PLATZ ở cây cải bắp Trung Quốc có tính acid (điểm đẳng điện < 7), trong khi tất cả các protein còn lại đều thể hiện tính base, với điểm đẳng điện > 7 [13]. Nghiên cứu này cũng cho thấy, giá trị độ ưa nước trung bình của các thành viên trong nhóm TF PLATZ ở cây cải bắp Trung Quốc đều < 0, dao động từ -0,202 - -0,819, chứng tỏ chúng cũng có tính ưa nước [13], tương tự như ghi nhận trong nghiên cứu này. Ở lúa mì, các thành viên của nhóm TF PLATZ có kích thước phân tử dao động từ 145 - 275 gốc amino acid

(giá trị trung bình đạt 221 gốc amino acid), trọng lượng phân tử đạt từ 16,67 - 29,76 kDa [12], điểm đẳng điện có giá trị từ 5,21 - 9,68 [12]. Gần đây, nhóm TF PLATZ ở cây bạch quả cũng đã được phân tích đặc tính lý hóa [16]. Theo đó, chiều dài trung bình của TF PLATZ ở cây bạch quả có giá trị đạt khoảng 270,27 gốc amino acid, với GbPLATZ1 có kích thước dài nhất, đạt 348 gốc amino acid [16]. Trọng lượng phân tử của TF PLATZ ở cây bạch quả đạt từ 25,84 - 38,74 kDa, với điểm đẳng điện dao động trong khoảng base, từ 8,02 - 9,26 [16]. Ở các loài trong chi *Malus* spp., nhóm TF PLATZ có kích thước đa dạng, từ 133 - 893 gốc amino acid, với trọng lượng phân tử ở ngưỡng từ 15,31 - 99,11 kDa [15]. Trong khi đó, giá trị điểm đẳng điện của nhóm TF PLATZ ở chi *Malus* dao động từ 6,10 - 9,94 [15].

### 3.3. Đánh giá mức độ biểu hiện của nhóm TF PLATZ trong giai đoạn sinh trưởng và phát triển của cây sắn

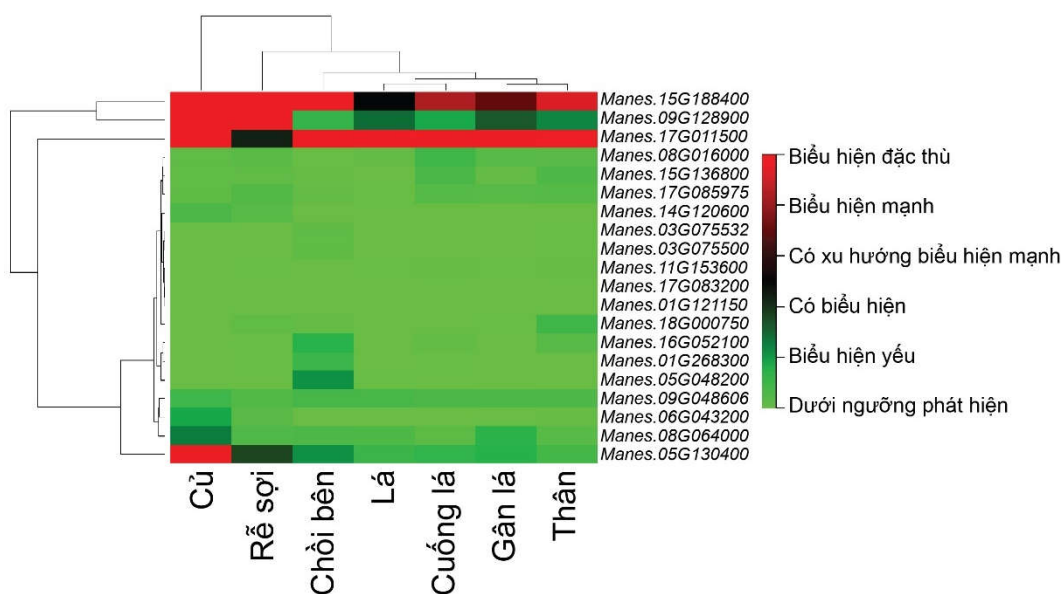
Để đánh giá biểu hiện đặc thù của các gene mã hóa nhóm TF PLATZ ở cây sắn, dữ liệu GSE82279 [19] trên GEO NCBI [20] đã được khai thác để phân tích giá trị FPKM của từng gene tại mỗi mẫu mô cơ quan chính. Kết quả phân tích dữ liệu biểu hiện tại bốn mẫu mô, bao gồm FEC, OES, RAM và SAM được thể hiện ở Hình 1 và tại 07 cơ quan/bộ phận chính, bao gồm lá, gân lá, cuống lá, rễ sợi, củ, chồi bên và thân được minh họa ở Hình 1.



Hình 1. Mức độ biểu hiện của các gene mã hóa nhóm TF PLATZ ở bốn mẫu mô trên cây sắn

Theo đó, kết quả phân tích cho thấy tám (trên tổng số 20) gene mã hóa nhóm TF PLATZ ở cây sắn, bao gồm *Manes.01G121150*, *Manes.03G075532*, *Manes.03G075500*, *Manes.08G016000*, *Manes.08G064000*, *Manes.11G153600*, *Manes.17G083200* và *Manes.17G085975* có giá trị FPKM ở dưới ngưỡng phát hiện hoặc có biểu hiện yếu. Trong khi đó, 12 (trên tổng số 20) gene còn lại có biểu hiện tại ít nhất một mẫu mô trên cây sắn. Cụ thể, gene *Manes.01G268300* có xu hướng biểu

hiện tại FEC và OES, trong khi *Manes.05G048200* là gene có biểu hiện tại SAM. Nghiên cứu cũng đã chỉ ra rằng, gene *Manes.17G011500* có biểu hiện tại OES và RAM. Đáng chú ý, *Manes.15G188400* được xác định là gene có xu hướng biểu hiện mạnh tại FEC và OES, trong khi *Manes.16G052100* có biểu hiện mạnh tại SAM và OES. Đặc biệt, *Manes.06G043200* là gene có biểu hiện đặc thù tại OES và RAM, và *Manes.14G120600* là gene có biểu hiện mạnh tại RAM.



Hình 2. Mức độ biểu hiện của các gene mã hóa nhóm TF PLATZ ở bảy cơ quan/bộ phận chính trên cây sắn

Kết quả phân tích mức độ biểu hiện của các gene mã hóa nhóm TF PLATZ ở bảy vị trí chính trên cây sắn cho thấy, 11 (trên tổng số 20) gene, bao gồm *Manes.01G121150*, *Manes.03G075532*, *Manes.03G075500*, *Manes.08G016000*, *Manes.09G048606*, *Manes.11G153600*, *Manes.14G120600*, *Manes.15G136800*, *Manes.17G083200*, *Manes.17G085975* và *Manes.18G000750* có mức độ biểu hiện dưới ngưỡng phát hiện hoặc biểu hiện yếu ở tất cả các cơ quan/bộ phận chính. Các gene còn lại được xác định có biểu hiện ở ít nhất một cơ quan/bộ phận chính trên cây sắn. Cụ thể, *Manes.05G130400* là gene có biểu hiện đặc thù ở củ sắn, trong khi *Manes.09G128900* có biểu hiện mạnh tại củ và

rễ sợi, đặt ra giả thuyết rằng hai gene này có chức năng liên quan đến các quá trình sinh lý diễn ra ở các bộ phận dưới mặt đất. Trong khi đó, *Manes.15G188400* và *Manes.17G011500* có biểu hiện mạnh ở tất cả 07 cơ quan/bộ phận chính trên cây sắn, ngoại trừ tại lá và rễ sợi, gợi ý rằng 02 gene này có khả năng tham gia vào điều hòa quá trình sinh trưởng và phát triển tại các cơ quan này trên cây sắn.

Trong nghiên cứu trước đây, các gene mã hóa nhóm TF PLATZ ở các loài thực vật cũng thể hiện mức độ biểu hiện khác biệt giữa các mẫu mô, chứng tỏ rằng nhóm protein điều hòa này có thể liên quan đến nhiều vai trò khác nhau. Ví dụ, gene *AtPLATZ1* được tăng cường biểu hiện ở mẫu rễ trong điều kiện xử lý với abscisic acid

và điều khiển sự phát triển của rễ chính trên cây *Arabidopsis* [27], trong khi gene *AtPLATZ7* được chứng minh là kiểm soát kích thước của mô phân sinh rễ thông qua kích thích của các gốc ôxi nguyên tử hoạt động [28]. Một gene khác trong nhóm TF PLATZ ở cây *Arabidopsis* là *AtPLATZ3* có chức năng tăng cường sự phát triển của lá thông qua cải thiện tỷ lệ và độ dài của quá trình biệt hóa tế bào xảy ra ở lá [29]. Ở cây bạch quả, gene *GbPLATZ9* biểu hiện mạnh ở rễ và hạt trong quá trình phát triển của hạt [16]. Trên cây lúa mì, tám (trên tổng số 62) gene thành viên trong nhóm TF PLATZ có biểu hiện mạnh ở rễ, thân, lá và hạt ở một số giai đoạn phát triển [12]. Ví dụ, ba gene thành viên, bao gồm *TaPLATZ26*, *TaPLATZ31* và *TaPLATZ32* có biểu hiện mạnh ở hạt giai đoạn 5 và 10 ngày sau khi hình thành, trong khi gene *TaPLATZ30* biểu hiện mạnh ở tất cả các giai đoạn phát triển của hạt [12]. Ba gene khác, bao gồm *TaPLATZ6*, *TaPLATZ2* và *TaPLATZ4* có biểu hiện đặc thù ở rễ và lá của cây lúa mì [12]. Mặt khác, mức độ biểu hiện của gene *GhPLATZ1* ở loài *Gossypium hirsutum* được tăng cường trong điều kiện bất lợi và xử lý phytohormone ngoại sinh ở mẫu cây non 20 ngày tuổi [30]. Các kết quả này đã gợi ý về chức năng của các gene mã hóa nhóm TF PLATZ ở cây sắn tại các vị trí cơ quan/bộ phận chính mà chúng có biểu hiện đặc thù.

#### 4. KẾT LUẬN

- Tổng số 20 thành viên của nhóm TF PLATZ đã được sàng lọc và xác định trên cây sắn. Trong đó, 19 trên tổng số 20 thành viên của nhóm TF PLATZ, ngoại trừ *Manes.01G121150*, đã được chú giải và định danh.

- Phân tích đặc tính cho thấy các thành viên của nhóm TF PLATZ ở cây sắn có kích thước và trọng lượng phân tử tương đồng, có tính ưa nước, với đa số protein thể hiện tính base và có cấu trúc không ổn định trong điều kiện ống nghiệm, tương tự như trên các loài cây trồng khác.

- Khai thác dữ liệu RNA-Seq cho thấy 06 gene, bao gồm *Manes.05G130400*, *Manes.06G043200*, *Manes.09G128900*, *Manes.14G120600*, *Manes.15G188400*,

*Manes.17G011500* có biểu hiện đặc thù tại ít nhất một mẫu mô hoặc cơ quan/bộ phận chính trên cây sắn.

#### TÀI LIỆU THAM KHẢO

- [1]. Guira, F., Some K., Kabore D., Sawadogo-Lingani H., Traore Y. & Savadogo A. (2017). Origins, production, and utilization of cassava in Burkina Faso, a contribution of a neglected crop to household food security. *Food Science & Nutrition*. 3: 415-423.
- [2]. Olsen M. & Schaal A. (1999). Evidence on the origin of cassava: Phylogeography of *Manihot esculenta*. *Proceedings of the National Academy of Sciences*. 10: 5586-5591.
- [3]. Hillocks R. J., Thresh J. M. & Bellotti A. (2002). *Cassava: Biology, production and utilization*. CABI Pub., Wallingford, UK; New York.
- [4]. De Souza P., Massenbun N., Jaiswal D., Cheng S., Shekar R. & Long, P. (2017). Rooting for cassava: insights into photosynthesis and associated physiology as a route to improve yield potential. *New Phytologist*. 1: 50-65.
- [5]. Chavarriaga-Aguirre P., Brand A., Medina A., Prias M., Escobar R., Martinez J., Diaz P., Lopez C., Roca W. M. & Tohme J. (2016). The potential of using biotechnology to improve cassava: A review. *In Vitro Cellular & Developmental Biology*. 5: 461-478.
- [6]. Nilusha R. A. T., Jayasinghe J., Perera O., Perera P. I. P. & Jayasinghe C. V. L. (2021). Proximate composition, physicochemical, functional, and antioxidant properties of flours from selected cassava (*Manihot esculenta* Crantz) varieties. *International Journal of Food Science*. 6064545.
- [7]. Liu Y., Huang Y., Li Z., Feng M., Ge W., Zhong C. & Xue R. (2023). Genome-wide identification of the TGA genes in common bean (*Phaseolus vulgaris*) and revealing their functions in response to *Fusarium oxysporum* f. sp. *phaseoli* infection. *Frontiers in Genetics*. 14: 1137634.
- [8]. Wang J., Ji C., Li Q., Zhou Y. & Wu Y. (2018). Genome-wide analysis of the plant-specific PLATZ proteins in maize and identification of their general role in interaction with RNA polymerase III complex. *BMC Plant Biology*. 1: 221.
- [9]. Kim J. H., Kim J., Jun S. E., Park S., Timilsina R., Kwon D. S., Kim Y., Park S., Hwang J. Y., Nam H. G., Kim G. T. & Woo H. R. (2018). ORESARA15, a PLATZ transcription factor, mediates leaf growth and senescence in *Arabidopsis*. *New Phytologist*. 2: 609-623.
- [10]. González-Morales S. I., Chávez-Montes R. A., Hayano-Kanashiro C., Alejo-Jacuinde G., Rico-Cambron T. Y., de Folter S. & Herrera-Estrella L. (2016). Regulatory network analysis reveals novel regulators of seed desiccation tolerance in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences of the*



United States of America. 35: E5232-E5241.

[11]. Wang A. & Hou Q. (2019). The PLATZ transcription factor GL6 affects grain length and number in rice. *Plant Physiology*. 4: 2077-2090.

[12]. 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*. 23: 8934.

[13]. Azim J. B., Khan F., Hassan L. & Robin A. H. (2020). Genome-wide characterization and expression profiling of plant-specific PLATZ transcription factor family genes in *Brassica rapa* L. *Plant Breeding and Biotechnology*. 1: 28-45.

[14]. 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*. 1: 329.

[15]. 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.

[16]. Han X., Rong H., Tian Y., Qu Y., Xu M. & Xu L. (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.

[17]. Bredeson J. V., Lyons J. B., Prochnik S. E., Wu G. A., Ha C. M., Edsinger-Gonzales E., Grimwood J., Schmutz J., Rabbi I. Y., Egesi C., Nauluvula P., Lebot V., Ndunguru J., Mkamilo G., Bart R. S., Setter T. L., Gleadow R. M., Kulakow P., Ferguson M. E., Rounsley S. & Rokhsar D. S. (2016). Sequencing wild and cultivated cassava and related species reveals extensive interspecific hybridization and genetic diversity. *Nature Biotechnology*. 5: 562-570.

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

[19]. Wilson M. C., Mutka A. M., Hummel A. W., Berry J., Chauhan R. D., Vijayaraghavan A., Taylor N. J., Voytas D. F., Chitwood D. H. & Bart R. S. (2017). Gene expression atlas for the food security crop cassava. *New Phytologist*. 4: 1632-1641.

[20]. Barrett T., Wilhite S. E., Ledoux P., Evangelista C., Kim I. F., Tomashevsky M., Marshall K. A., Phillippy K. H., Sherman P. M., Holko M., Yefanov A., Lee H., Zhang N., Robertson C. L., Serova N., Davis S. & Soboleva A. (2013). NCBI GEO: archive for functional genomics data

sets - Update. *Nucleic Acids Research*. 41: D991-D995.

[21]. La Viet Hong, Chu Duc Ha, Tran Duy Cuong, Nguyen Huu Kien, Le Thi Ngoc Quynh, Hoang Minh Chinh, Cao Phi Bang, Pham Cong Anh Tuyen, Nguyen Quoc Trung, Nguyen Van Loc, Ha Van Chien, Le Huy Ham, Le Tien Dung & Tran Phan Lam Son. (2022). Insights into the gene and protein structures of the CaSWEET family members in chickpea (*Cicer arietinum*), and their gene expression patterns in different organs under various stress and abscisic acid treatments. *Gene*. 819: 146210.

[22]. 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.

[23]. Paysan-Lafosse T., Blum M., Chuguransky S., Grego T., Pinto B. Lázaro S., Gustavo A., Bileschi M., Bork P., Bridge A., Colwell L., Gough J., Haft D., Letunić I., Marchler-Bauer A., Mi H., Natale D., Orengo C., Pandurangan A., Rivoire C., Sigrist C., Sillitoe I., Thanki N., Thomas P., Silvio C. E., Cathy H. & Bateman A. (2023). InterPro in 2022. *Nucleic Acids Research*. (D1): D418-D427.

[24]. Gasteiger E., Gattiker A., Hoogland C., Ivanyi I., Appel R. D. & Bairoch A. (2003). ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Research*. 13: 3784-3788.

[25]. Marini F., Linke J. & Binder H. (2020). ideal: an R/Bioconductor package for interactive differential expression analysis. *BMC Bioinformatics*. 1: 565.

[26]. Awoleye F., van Duren M., Dolezel J. & Novak F. J. (1994). Nuclear DNA content and in vitro induced somatic polyploidization cassava (*Manihot esculenta* Crantz) breeding. *Euphytica*. 76 (3): 195-202.

[27]. Dong T., Yin X., Wang H., Lu P., Liu X., Gong C. & Wu Y. (2021). ABA-INDUCED expression 1 is involved in ABA-inhibited primary root elongation via modulating ROS homeostasis in *Arabidopsis*. *Plant Science*. 304: 110821.

[28]. Yamada M., Han X. & Benfey P. N. (2020). RGF1 controls root meristem size through ROS signalling. *Nature*. 7788: 85-88.

[29]. Jun S. E., Kim J. H., Hwang Ji. Y., Thien Tu Huynh Le & Kim G. (2020). ORESARA15 acts synergistically with ANGUSTIFOLIA3 and separately from AINTEGUMENTA to promote cell proliferation during leaf growth. *International Journal of Molecular Sciences*. 1: 241.

[30]. Zhang S., Yang R., Huo Y., Liu S., Yang G., Huang J., Zheng C. & Wu C. (2018). Expression of cotton PLATZ1 in transgenic *Arabidopsis* reduces sensitivity to osmotic and salt stress for germination and seedling establishment associated with modification of the abscisic acid, gibberellin, and ethylene signalling pathways. *BMC Plant Biology*. 1: 218.