

NGHIÊN CỨU ĐẶC ĐIỂM CÁC GENE MÃ HÓA SWEET Ở CÂY ĐU ĐỦ (*Carica papaya* L.)

Lê Thị Mận^{1,2}, Nguyễn Phương Quý¹, Nguyễn Thị Huệ¹,
Nguyễn Thị Nguyệt Nga¹, Cao Phi Bằng^{1,2*}

¹Khoa Khoa học Tự nhiên, Trường Đại học Hùng Vương, Phú Thọ

²Nhóm nghiên cứu Công nghệ Sinh học, Trường Đại học Hùng Vương, Phú Thọ

Ngày nhận bài: 13/6/2021; Ngày chỉnh sửa: 16/6/2021; Ngày duyệt đăng: 18/6/2021

Tóm tắt

Đu đủ là loại cây ăn quả nhiệt đới quan trọng được trồng rộng rãi trên thế giới, trong đó có Việt Nam. Sự trao đổi và vận chuyển đường là vấn đề cần được quan tâm ở đu đủ. Nhờ sử dụng các phương pháp nghiên cứu tin sinh học, tổng số 12 gene mã hóa SWEET đã được xác định ở trong hệ gene của cây đu đủ (*Carica papaya* L.). Các gene SWEET của cây đu đủ có kích thước từ 1.203 đến 2.639 bp, trong đó hầu hết các gene có 5 intron, chỉ trừ gene CpSWEET12 có 2 intron. Các protein suy diễn có kích thước từ 234 tới 302 amino acid, với khối lượng phân tử 26,30 kDa tới 32,95 kDa. Các protein này có tính kiềm với giá trị pI dao động từ 7,69 đến 9,55. Cấu trúc không gian thể hiện các CpSWEET có 6 hoặc 7 xoắn xuyên màng. Căn cứ vào kết quả phân tích cây phá hệ, các SWEET của cây đu đủ được phân chia thành bốn nhóm I (ba gene), nhóm II (hai gene), nhóm III (năm gene) và nhóm IV (hai gene). Kết quả nghiên cứu này là tiền đề cho các nghiên cứu sâu hơn về tách dòng gene, phân tích chức năng của các gene trong họ SWEET và chọn giống ở cây đu đủ.

Từ khóa: Đặc điểm gene, đu đủ (*Carica papaya* L.), cây phá hệ, SWEET, tin sinh học

1. Đặt vấn đề

SWEET (Sugars Will Eventually Be Exported Transporter) là các protein có cấu trúc đặc trưng gồm 7 vùng xoắn xuyên màng. Chúng giữ chức năng quan trọng trong vận chuyển đường sucrose ở thực vật [1]. Đồng thời, nhóm protein này còn giữ nhiều chức năng trong sự phát triển các cơ quan sinh sản, vận chuyển gibberellin và điều hòa phản ứng với các stress vô sinh [1]. Do có nhiều vai trò quan trọng, họ gene mã hóa các SWEET

từ lâu đã được nghiên cứu ở nhiều loại thực vật, điển hình như *Arabidopsis thaliana* [2], lúa [3], sắn [4], ca cao [5] cũng như một số loài thực vật khác [6]. Tuy nhiên, chưa có nghiên cứu nào về họ gene SWEET ở cây đu đủ được thực hiện.

Cây đu đủ (*Carica papaya* L.) là loại cây ăn quả được trồng nhiều ở vùng nhiệt đới và cận nhiệt đới. Quả đu đủ khi chín có đặc điểm mềm, vị ngọt, giá trị dinh dưỡng cao, đặc biệt có nhiều tiền vitamin A, vitamin C

và các chất chống oxy hóa [7]. Do cây đu đủ sinh sống ở các vùng nhiệt đới hoặc cận nhiệt đới nên chịu tác động của nhiều nhân tố sinh thái như ánh sáng, nhiệt độ, chế độ nước, gió cũng như nhiều tác nhân stress sinh học [8].

Nghiên cứu này nhằm xác định và phân tích đặc điểm của các gene mã hóa SWEET ở cây đu đủ nhờ các phương pháp tin sinh học, vốn là phương pháp nghiên cứu hiện đại, được sử dụng phổ biến trong công nghệ sinh học hiện đại [9].

2. Cơ sở dữ liệu và phương pháp nghiên cứu

2.1. Cơ sở dữ liệu hệ gene của cây đu đủ

Trình tự hệ gene của cây đu đủ (*Carica papaya* L.) đã được giải trình tự [10] và được đặt trên website phytozome V12 (<http://www.phytozome.net/>)

3. Kết quả nghiên cứu và thảo luận

3.1. Xác định họ gene SWEET ở cây đu đủ

2.2. Xác định các gene SWEET ở cây đu đủ

Các protein SWEET của cây *A. thaliana* [11] được sử dụng làm khuôn dò để tìm kiếm các gene tương đồng trên toàn hệ gene của cây đu đủ nhờ sử dụng chương trình TBLASTN [12].

2.3. Xây dựng cây phả hệ

Các protein SWEET của cây đu đủ, *A. thaliana* và cây lúa được sắp dãy bằng MAFFT [13], sau đó cây phả hệ được xây dựng nhờ phân mềm MEGA X [14].

2.4. Phân tích đặc điểm các gene SWEET của cây đu đủ

Công cụ ProtParam được sử dụng để phân tích các đặc điểm vật lý, hóa học của các gene/protein SWEET được phân tích trên server ExPASy (Expert Protein Analysis System) [15]. ProtComp 9.0 (<http://linux1.softberry.com/berry.phtml?Topic=protcomppl&group=help&subgroup=proloc>) được dùng để xác định vị trí khu trú dưới tế bào.

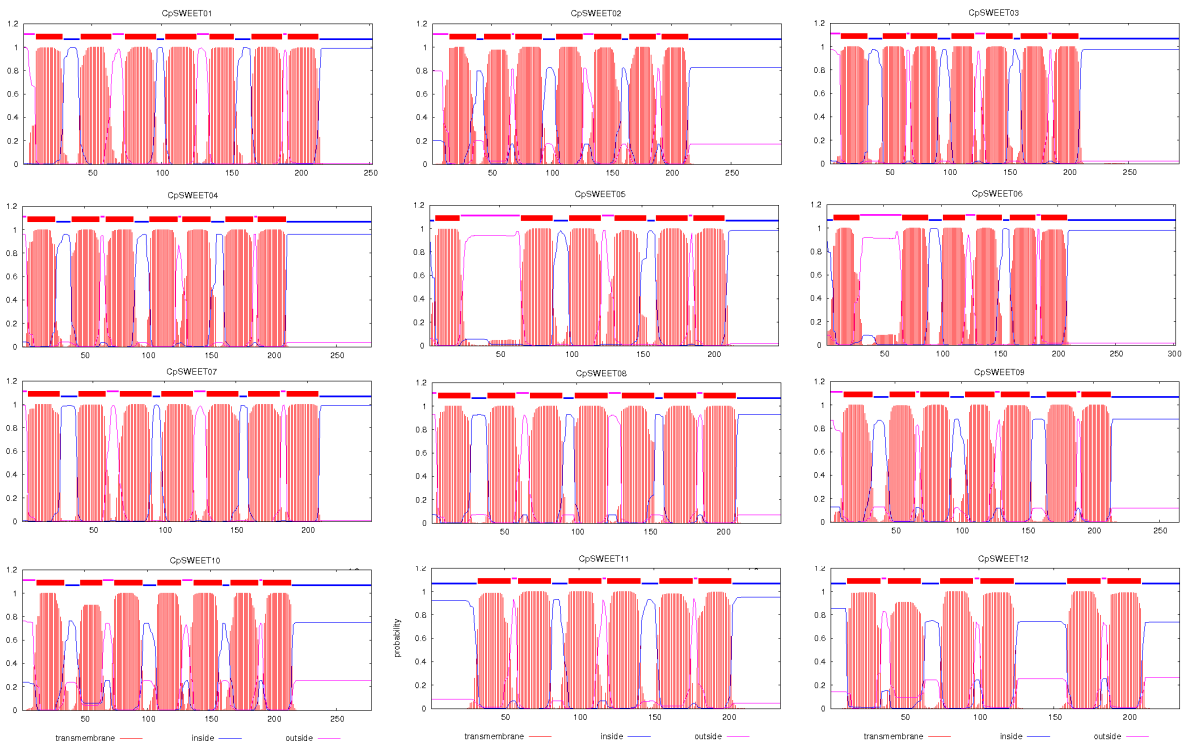
Bảng 1. Các trình tự SWEET ở cây đu đủ

Gene	Tên locus	Phân nhóm	GS (bp)	PL (aa)	MW (kD)	pI	GRAVY	In	SCL	THM
CpSWEET01	evm.TU.supercontig_3.56	I	1499	252	28,16	9,53	0,62	5	PM	7
CpSWEET02	evm.TU.supercontig_4.141	III	2015	292	32,82	8,64	0,66	5	PM	7
CpSWEET03	evm.TU.supercontig_4.143	III	1584	292	32,67	8,28	0,68	5	PM	7
CpSWEET04	evm.TU.supercontig_9.173	III	1400	278	31,30	8,28	0,48	5	PM	7
CpSWEET05	evm.TU.supercontig_48.105	IV	1523	246	27,32	9,16	0,66	5	PM	6
CpSWEET06	evm.TU.supercontig_48.106	IV	1638	302	32,95	8,92	0,44	5	PM	6
CpSWEET07	evm.TU.supercontig_49.77	I	1203	245	26,76	9,36	0,75	5	PM	7
CpSWEET08	evm.TU.supercontig_49.78	I	2112	240	26,61	9,39	0,77	5	PM	7
CpSWEET09	evm.TU.supercontig_51.71	III	1536	265	29,74	9,39	0,62	5	PM	7
CpSWEET10	evm.TU.supercontig_99.60	III	1869	278	31,33	7,69	0,50	5	PM	7
CpSWEET11	evm.TU.contig_35275	II	1209	235	26,30	9,55	0,64	5	PM	6
CpSWEET12	evm.TU.contig_36848	II	2639	234	26,60	9,17	0,23	2	PM	6

Chú thích: GS = Kích thước gen, PL = Chiều dài phân tử protein, MW = Khối lượng phân tử protein, pI = điểm đẳng điện, In = Số lượng intron, SCL = Khu trú dưới tế bào, Chlo: lục lạp, Mito: ti thể, Nucl: nhân tế bào

Tổng số 12 gene SWEET đã được xác định ở trong hệ gene cây đu đủ (bảng 1). Khi xét trên quy mô hệ gene, họ SWEET của cây đu đủ có ít gene hơn so với cây *A. thaliana* (17 gene) [2], cây lúa và cây ca cao (cùng 21 gene) [5], cũng như cây sắn (28 gene) [4]. Các CpSWEET đều mang một hoặc vùng bảo tồn đặc trưng (*MtN3_slv* (PF03083) cho họ SWEET [2]. Các gene mã hóa SWEET ở cây đu đủ có chiều dài dao động từ 1203 đến 2639 nucleotide (bảng 1). Tất cả 12 gene CpSWEET đều mã hóa không liên tục, 11

trong tổng số 12 gene có 5 intron, chỉ riêng gene CpSWEET12 có 2 intron (bảng 1). Kích thước các phân tử protein suy diễn dao động trong khoảng từ 234 (CpSWEET12) tới 302 (CpSWEET06) amino acid, khối lượng phân tử từ 26,3 kDa (CpSWEET11) tới 32,95 kDa (CpSWEET06). Các CpSWEET có tính base với giá trị pI lý thuyết nằm trong khoảng từ 7,69 tới 9,55. Như vậy, các SWEET của cây đu đủ có đặc điểm lý-hóa khá tương đồng với SWEET của cây ca cao [5].



Hình 1. Mô hình cấu trúc không gian của một số AhSWEET được xây dựng nhờ TMHMM Server v.2.0 (<https://services.healthtech.dtu.dk/service.php?TMHMM-2.0>).

Về cấu trúc không gian, các CpSWEET có cấu trúc gồm nhiều xoắn xuyên màng đặc trưng. Trong đó, có tới 8 trong tổng số 12 CpSWEET có 7 vùng xoắn xuyên màng (hình 1), bốn phân tử còn lại chỉ có 6 xoắn xuyên

màng (bảng 1). Cấu trúc xoắn xuyên màng đặc trưng của các CpSWEET tương đồng với các SWEET của nhiều loài loài đã biết như *A. thaliana*, lúa [2], sắn [4], ca cao [5]. Cấu trúc xoắn xuyên màng giải thích cho kết quả

Kết quả phân tích cây phả hệ (hình 2) thể hiện các SWEET của cây đu đủ được xếp thành bốn nhóm I-IV, tương tự như của các loài khác [5, 6]. Nhóm I (3 gene), nhóm II (2 gene), nhóm III (5 gene) và nhóm IV (2 gene). Nhóm III có số lượng nhiều gene nhất trong họ SWEET, tương tự ở cây ca cao [5], cây *A. thaliana* [6].

4. Kết luận

Tổng số 12 gene SWEET được phát hiện và phân tích ở cây đu đủ bằng phương pháp tin sinh học. Cấu trúc gene và các đặc điểm lý - hóa của các gene/protein SWEET này cũng được nghiên cứu. Đặc điểm nổi bật là các gene này mã hóa không liên tục với 11 gene có 5 intron, 1 gen có 2 intron. Các protein đều có các xoắn xuyên màng đặc trưng với 8 phân tử có 7 xoắn, 5 phân tử có 6 xoắn. Kết quả phân tích cây phả hệ cho thấy các SWEET của cây đu đủ được xếp vào bốn nhóm, nhóm I (3 gene), II (2 gene), III (5 gene) và IV (2 gene). Kết quả này có ý nghĩa lớn, mở đường cho việc tách dòng gene và phân tích chức năng của các gene trong họ SWEET ở cây đu đủ trong đáp ứng với các điều kiện môi trường hoặc sự phát triển.

Lời cảm ơn

Công trình này được hoàn thành với sự hỗ trợ kinh phí từ chương trình nghiên cứu khoa học cơ bản của Trường Đại học Hùng Vương (Đề tài mã số: 24/2020/HĐKH.HV20-24).

Tài liệu tham khảo

- [1] Jeena G. S., Kumar S., & Shukla R. K. (2019). Structure, evolution and diverse physiological roles of SWEET sugar transporters in plants. *Plant Mol Biol*, 100(4-5) 351-365, doi: 10.1007/s11103-019-00872-4.
- [2] Chen L. Q., Hou B-H., Sylvie Lalonde S., Takanaga H., Hartung M. L., Qu X-Q., Guo W-J., Kim J-G., Underwood W., Chaudhuri B., Chermak D., Antony G., White F. F., Somerville S. C., Mudgett M. B. & Frommer W. B. (2010). Sugar transporters for intercellular exchange and nutrition of pathogens. *Nature*, 468(7323), 527-32, doi: 10.1038/nature09606.
- [3] Yuan M. & Wang S. (2013). Rice MtN₃/saliva/SWEET family genes and their homologs in cellular organisms. *Mol Plant*, 6(3) 665-74, doi: 10.1093/mp/sst035.
- [4] Chu Đức Hà, Phan Thị Quỳnh, Phạm Thị Lý Thu, Nguyễn Văn Cường & Lê Tiến Dũng (2018). Xác định họ gen mã hóa protein vận chuyển Sweet trên cây sắn (*Manihot esculenta* Crantz), *Tạp chí Khoa học Trường Đại học Sư phạm Hà Nội*, 63(3). 140-149.
- [5] Cao Phi Bằng, Nguyễn Văn Đỉnh, Trần Thị Thanh Huyền, Lê Thị Mận & Vũ Xuân Dương (2020). In silico characterisation of genes encoding SWEET protein in cocoa (*Theobroma cacao* L.). *Báo cáo Hội nghị Quốc gia lần thứ 4 về nghiên cứu và giảng dạy Sinh học ở Việt Nam*, Vinh Phuc, Vietnam, 2020.
- [6] Li X., Si W., Qin Q., Wu H. & Jiang H. (2018). Deciphering evolutionary dynamics of SWEET genes in diverse plant lineages. *Sci. Rep.*, 8(1), 13440, doi: 10.1038/s41598-018-31589-x.
- [7] Ming R. Qingyi Yu Q., Moore P. H., Paull R. E., Chen N. J., Wang M-L., Zhu Y. J., Schuler M. A., Jiang J. & Paterson A. H. (2012). Genome of papaya, a fast growing tropical fruit tree. *Tree Genet. Genom.*, 8(3), 445-462, doi: 10.1007/s11295-012-0490-y.
- [8] Campostrini E. & Glenn D. M. (2007). Ecophysiology of papaya: a review. *Brazilian Journal of Plant Physiology*, 19(4), 413-424.
- [9] Cao Phi Bằng (2015). Xác định và phân tích in silico các gen DREB₂ ở cây quýt (*Citrus clementina*). *Tạp chí Khoa học Đại học Sư phạm Hà Nội*, 60(4), 127-131, doi: 10.18173/2354-1059.2015-00018.
- [10] Ming R., Hou S., Feng Y. & et al. (2008). The draft genome of the transgenic tropical fruit tree papaya (*Carica papaya* Linnaeus). *Nature*, 452: 991-996, 10.1038/nature06856

- [11] Chen L. Q., Hou B. H., Lalonde S., Takanaga H., Hartung M. L., Qu X. Q., Woei-Jiun Guo W. J., Kim J. G., Underwood W., Chaudhuri B., Chermak D., Antony G., White F. F., Somerville S. C., Mudget M. B. & Frommer W. B. (2010). Sugar transporters for intercellular exchange and nutrition of pathogens. *Nature*, 468(7323), 527-532. doi:10.1038/nature09606
- [12] Gertz E. M., Yu Y. K., Agarwala R., Schaffer A. A. & Altschul S. F. (2006) Composition-based statistics and translated nucleotide searches: Improving the TBLASTN module of BLAST. *BMC Biol*, 4, 41, doi: 10.1186/1741-7007-4-41.
- [13] Katoh K. & Standley D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*, 30(4), 772-80, doi: 10.1093/molbev/mst010.
- [14] Kumar S., Stecher G., Li M., Knyaz C., & Tamura K. (2018) MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol Biol Evol*, 35(6), 1547-1549, doi: 10.1093/molbev/msy096.
- [15] Gasteiger E., Hoogland C., Gattiker A., Wilkins M. R., Appel R. D., & Bairoch A. (2005). Protein identification and analysis tools on the ExPASy server. In the proteomics protocols handbook: Springer, 571-607.

STUDY ON CHARACTERISTICS OF SWEET ENCODING GENES IN PAPAYA (*Carica papaya* L.)

Le Thi Man^{1,2}, Nguyen Phuong Quy¹, Nguyen Thi Hue¹,
Nguyen Thi Nguyet Nga¹, Cao Phi Bang^{1,2}

¹Faculty of Natural Sciences, Hung Vuong University, Phu Tho

²Biotechnology Research Group, Hung Vuong University, Phu Tho

Abstract

Papaya (*Carica papaya* L.) is an important tropical fruit crop which is widely cultivated in the world, including Vietnam. Sugar metabolism and transport is an important matter of concern. By using the bioinformatics methods, total of 12 SWEET encoding genes have been identified in the genome of papaya (*Carica papaya* L.). The sizes of papaya SWEET genes were ranging from 1203 to 2639 bp. Among which, all of these genes had five introns, except CpSWEET12 including two introns. The predicted protein sequences contained from 234 to 302 amino acids, according to the molecular weight ranged from 26.30 to 32.95 kDa. These proteins were alkaline with a pI value ranging from 7.69 to 9.55. The secondary structure showed that the CpSWEET included six or seven transmembrane helices. Based on the phylogenetic analysis, the papaya SWEET were classified into four groups, I (three members), II (two members), III (five members) and IV (two members). The results of this study have an important significance and are the basis of the further research on gene cloning, functional analysis of SWEET genes.

Keywords: Gene characterization, *in silico*, papaya (*Carica papaya* L.), phylogenetic tree, SWEET