

PHÂN TÍCH HIỆN TƯỢNG LẬP Ở HỌ GENE MÃ HÓA PROTEIN VẬN CHUYỂN SUCROSE LIÊN QUAN ĐẾN ĐÁP ỨNG MẶN Ở LẠC (*Arachis hypogaea*)

**Chu Đức Hà¹, Hoàng Thị Hải Yến^{1,2}, Lê Thị Ngọc Quỳnh³, Phạm Thị Ngọc Bích^{1,4},
Hoàng Minh Chính^{1,4}, La Việt Hồng², Phạm Phương Thu², Nguyễn Thị Minh Nguyệt¹,
Lê Hùng Lĩnh¹, Lê Huy Hàm^{1,5}, Nguyễn Quốc Trung⁴, Phạm Xuân Hội¹**

¹ Viện Di truyền Nông nghiệp, Viện Khoa học Nông nghiệp Việt Nam

² Khoa Sinh - Kỹ thuật Nông nghiệp, Trường Đại học Sư phạm Hà Nội 2

³ Khoa Hóa và Môi trường, Trường Đại học Thủy lợi

⁴ Khoa Công nghệ Sinh học, Học viện Nông nghiệp Việt Nam

⁵ Khoa Công nghệ Nông nghiệp, Trường Đại học Công nghệ, Đại học Quốc gia Hà Nội

TÓM TẮT

Trong nghiên cứu này, tổng số 15 sự kiện lập (duplication event) đã được dự đoán trong họ gene *AhySWEET*. Tất cả các cặp gene lập đều được xuất phát từ hiện tượng lập đoạn xảy ra trên các nhiễm sắc thể (segmental duplication) trong genome của lạc. Tính toán trị số thay thế đồng nghĩa (Ks) và trái nghĩa (Ka) cho thấy chọn lọc tự nhiên đã kìm hãm các đột biến điểm xảy ra trong các cặp gene lập. Tiếp theo, nghiên cứu đã chỉ ra rằng họ protein AhySWEET chủ yếu gồm bảy cấu trúc xoắn cuộn xuyên màng (*trans*- membrane helices) và được phân bố chủ yếu tại các cơ quan tiết (secretory pathway) trong tế bào. Khai thác dữ liệu biểu hiện đã chứng minh các cặp gene lập cùng thể hiện xu hướng biểu hiện giống nhau tại mẫu rễ xử lý mặn. Cụ thể, các cặp gene *AhySWEET13/37* và *10/34* có mức độ phiên mã giảm, trong khi biểu hiện của hai gene *AhySWEET16/40* tăng ở rễ trong giai đoạn đầu của xử lý mặn. Nghiên cứu này sẽ được tiếp tục nhằm tìm hiểu vai trò của họ gene *AhySWEET* trong đáp ứng đa bất lợi ở lạc.

Từ khóa: Gene lập, lạc, mặn, SWEET, vận chuyển đường.

MỞ ĐẦU

Lạc (*Arachis hypogaea*) là một trong những loài cây họ Đậu quan trọng, được canh tác rộng rãi trên thế giới với nhiều mục đích sử dụng khác nhau, như thực phẩm, nguyên liệu cho sản xuất thức ăn chăn nuôi, nhiên liệu sinh học và cải tạo đất (Toomer, 2018). Tại Việt Nam, lạc nói riêng và các cây họ Đậu nói chung được đánh giá là loại cây công nghiệp ngắn ngày, trồng xen canh trên các diện tích canh tác hoa màu với tác dụng cải tạo đất, góp phần tăng thu nhập cho người nông dân. Tuy nhiên, năng suất và chất lượng lạc đang bị ảnh hưởng nghiêm trọng bởi tác động của các yếu tố bất lợi phi sinh học. Các nhà khoa học đã chỉ ra, khi cây trồng bị tác động bởi các bất lợi phi sinh học, hàm lượng đường sucrose tại vị trí tổn thương được gia tăng thông qua vai trò của nhóm protein vận chuyển SWEET (Sugars Will Eventually be Exported Transporter) (Chen, 2014; Patil *et al.*, 2015).

Trong nghiên cứu trước đây, họ protein SWEET đã được xem xét trên một số đối tượng cây trồng, như đậu tương (*Glycine max*) (Patil *et al.*, 2015) cải dầu (*Brassica napus*) (Jian *et al.*, 2016) và chuối (*Musa acuminata*) (Miao *et al.*, 2017). Đáng chú ý, các sự kiện lập (gene duplication) xảy ra trong họ gene *SWEET* ở những đối tượng này đã được kiểm chứng có mức độ phiên mã đáp ứng với bất lợi phi sinh học (Patil *et al.*, 2015; Jian *et al.*, 2016; Miao *et al.*, 2017). Đây được xem là cơ chế thúc đẩy sự nhân rộng của các họ gene (gene family expansion) ở thực vật trong quá trình tiến hóa nhằm đáp ứng với sự thay đổi của điều kiện ngoại cảnh. Tuy nhiên, nghiên cứu về hiện tượng lập gene liên quan đến đáp ứng bất lợi ở nhóm gene *AhySWEET* trên cây lạc cho đến nay vẫn chưa được tìm hiểu một cách đầy đủ (Ha *et al.*, 2020).

Trong nghiên cứu này, hiện tượng lập gene và vai trò của chọn lọc tự nhiên đã được dự đoán dựa trên các công cụ tin sinh học. Bên cạnh đó, trình tự protein của các AhySWEET khai thác từ nghiên cứu trước đây (Ha *et al.*, 2020) đã được sử dụng để phân tích cấu trúc vùng bảo thủ và vị trí cư trú nội bào. Quan trọng hơn cả, mức độ biểu hiện của các gene *AhySWEET* đã được đánh giá dựa trên khai thác dữ liệu RNA-Seq trong điều kiện xử lý mặn. Kết quả của nghiên cứu này có thể cung cấp những dữ liệu quan trọng cho việc định hướng phân tích chức năng gene nhằm cải thiện đặc tính chống chịu bất lợi ở cây lạc.

DỮ LIỆU VÀ PHƯƠNG PHÁP

Dữ liệu nghiên cứu

Dữ liệu genome và proteome của giống lạc Tifrunner (BioProject: PRJNA419393) trên Legume Information System và Phytosome (Bertioli *et al.*, 2019) được sử dụng làm hệ tham chiếu cho các phân tích *in silico*.

Trình tự gene (CDS và gDNA) của 43 *AhySWEET* được kế thừa từ nghiên cứu trước đây (Ha *et al.*, 2020).

Phương pháp nghiên cứu

Phân tích vị trí cư trú nội bào của *AhySWEET*: Trình tự của các protein *AhySWEET* (Ha *et al.*, 2020) được sử dụng để tìm kiếm các đoạn peptide tín hiệu (SP) đặc trưng cho các bào quan, bao gồm peptide chuyển vị của ty thể (mTP) và peptide chuyển vị của lục lạp (cTP) trên công cụ TargetP (Almagro *et al.*, 2019).

Xây dựng sơ đồ hình cây và vùng bảo thủ của *AhySWEET*: Sơ đồ hình cây của họ protein *AhySWEET* được xây dựng bằng thuật toán Neighbor-Joining (giá trị bootstrap = 1000) với trình tự của 43 *AhySWEET* (Ha *et al.*, 2020) trên MEGA 7.0 (Kumar *et al.*, 2016) theo các bước mô tả trong nghiên cứu trước đây (Chu *et al.*, 2018). Vùng bảo thủ đặc trưng của nhóm protein *AhySWEET* được phân tích trên TMHMM (*trans-* membrane by hidden Markov model) server 2.0 (Chaturvedi *et al.*, 2011).

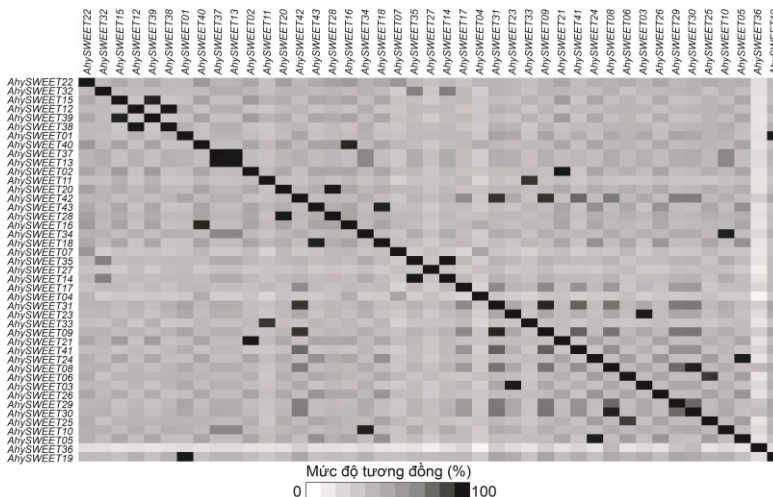
Phân tích sự tiến hóa của gene *AhySWEET*: Mức độ tương đồng (%) của các gene *AhySWEET* được phân tích dựa trên phân tích căn trình tự tương đồng đoạn CDS. Cặp gene lập được định nghĩa là các gene *AhySWEET* có mức độ tương đồng > 90%. Giá trị thay thế đồng nghĩa (Ks) và trái nghĩa (Ka) của sự kiện lập (duplication event) được tính toán trên DNAsp 6.0 (Rozas *et al.*, 2017).

Khai thác dữ liệu biểu hiện RNA-Seq: Dữ liệu phiên mã RNA-Seq ở lạc trong điều kiện xử lý mặn (Accession: GSE75625) được khai thác từ nghiên cứu trước đây (Chen *et al.*, 2016). Trong đó, mẫu rễ ở giai đoạn cây non của giống lạc Huayu 19 được xử lý trong dung dịch NaCl 200 mM trong 0 (đối chứng), 3 và 48h. RNA tổng số được sử dụng làm khuôn cho phân tích oligonucleotide microarray (Chen *et al.*, 2016). Dữ liệu biểu hiện của các gene *AhySWEET* được phân tích mô hình hóa bằng bảng nhiệt màu thông qua ngôn ngữ R và gplots (Liao *et al.*, 2019).

KẾT QUẢ VÀ THẢO LUẬN

Phân tích sự nhân rộng của họ gene *AhySWEET* ở lạc

Tìm hiểu về cơ chế phát tán của họ gene *SWEET* ở thực vật cho phép đưa ra các giả thuyết về chức năng của các gene lập trong quá trình tiến hóa của loài (Patil *et al.*, 2015). Trong nghiên cứu này, hiện tượng lập ở họ gene *AhySWEET* đã được dự đoán bằng cách xem xét mức độ tương đồng (> 90%). Kết quả đã xác định được tổng số 15 sự kiện lập gen, với mức độ tương đồng dao động từ 92,4 (*AhySWEET06/25*) đến 99,2% (*AhySWEET12/38*) (Bảng 1, Hình 1). Đáng chú ý, tất cả các cặp gene lập đều xuất phát từ những nhiễm sắc thể (NST) khác nhau, trong khi không có sự kiện lập gene nào ở họ *AhySWEET* diễn ra trên cùng NST được ghi nhận (Bảng 1). Kết quả này đã chỉ ra rằng, hiện tượng lập trên các đoạn NST (segmental duplication) là cơ chế chính giải thích cho quá trình nhân rộng của họ gene *AhySWEET* ở lạc. Trước đó, 56 gene lập trong cùng một NST (tandemly duplication) và 95 sự kiện lập trên các đoạn NST trên tổng số 411 gene *SWEET* ở 23 loài thực vật (Patil *et al.*, 2015). Bên cạnh đó, hiện tượng lập trên các đoạn NST là nguyên nhân chính giải thích cho sự nhân rộng của họ gene *BnSWEET* ở cải dầu (21/23 sự kiện lập) (Jian *et al.*, 2016) và *BoSWEET* ở cải bắp (toàn bộ bảy sự kiện lập) (Zhang *et al.*, 2019). Những kết quả này đã thống nhất rằng hiện tượng lập gene ở họ *SWEET* ở lạc nói riêng và thực vật nói chung chủ yếu xuất phát từ cơ chế lập trên các đoạn NST.



Hình 1. Mức độ tương đồng trong họ gene *AhySWEET* ở lạc

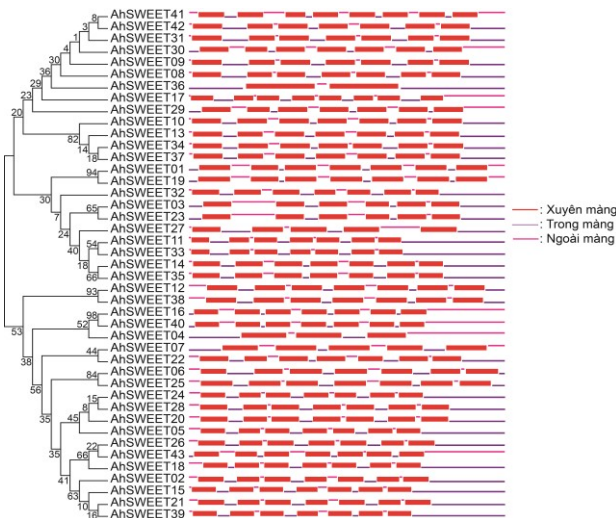
Để đánh giá áp lực của chọn lọc tự nhiên đến các sự kiện lặp ở họ gene *AhySWEET* ở lạc, giá trị K_a và K_s đã được tính toán bằng DNAsp (Rozas *et al.*, 2017). Kết quả cho thấy hầu hết tỷ lệ K_a/K_s ở các sự kiện lặp (11/15) đều < 1 (Bảng 1), đặt ra giả thuyết rằng, các gene lặp *AhySWEET* được bảo tồn về mặt cấu trúc (do các giá trị thay thế trái nghĩa K_a nhỏ hơn giá trị thay thế đồng nghĩa K_s). Điều này được giải thích có thể do áp lực của chọn lọc tự nhiên đã kìm hãm hiện tượng đột biến điểm xảy ra, dẫn đến cấu trúc của các gene lặp được bảo tồn để đảm bảo chức năng của gene. Trong khi đó, bốn cặp gene lặp, *AhySWEET06/25*, *08/30*, *14/35*, *18/43* có tỷ lệ $K_a/K_s > 1$ (Bảng 1), gợi ý rằng chọn lọc tự nhiên thúc đẩy đột biến xảy ra, các gene lặp có sự sai khác về mặt cấu trúc, dẫn đến thay đổi về mặt chức năng để đáp ứng với sự thay đổi của ngoại cảnh. Những dự đoán này cũng đồng thuận với những ghi nhận trước đây trên các loài thực vật, ví dụ như đậu tương (Patil *et al.*, 2015), cải dầu (Jian *et al.*, 2016), chuối (Miao *et al.*, 2017) và cải bắp (Zhang *et al.*, 2019).

Bảng 1. Các sự kiện lặp xảy ra trong họ gene *AhySWEET* ở lạc

TT	Cặp gene lặp	Chr	K_a	K_s	K_a/K_s
1	<i>AhySWEET01/19</i>	01/11	0,004	0,021	0,183
2	<i>AhySWEET03/23</i>	03/13	0,009	0,034	0,265
3	<i>AhySWEET05/24</i>	03/13	0,016	0,023	0,700
4	<i>AhySWEET06/25</i>	03/13	0,063	0,036	1,731
5	<i>AhySWEET08/30</i>	05/15	0,022	0,018	1,200
6	<i>AhySWEET09/31/42</i>	05/15/20	0,017/0,026	0,058/0,058	0,291/0,455
7	<i>AhySWEET10/34</i>	06/16	0,018	0,027	0,680
8	<i>AhySWEET11/33</i>	06/16	0,052	0,055	0,935
9	<i>AhySWEET12/38</i>	07/18	0,004	0,020	0,188
10	<i>AhySWEET13/37</i>	08/17	0,007	0,020	0,351
11	<i>AhySWEET14/35</i>	08/17	0,015	0,006	2,475
12	<i>AhySWEET15/39</i>	08/18	0,023	0,026	0,896
13	<i>AhySWEET16/40</i>	08/18	0,027	0,031	0,892
14	<i>AhySWEET18/43</i>	10/20	0,014	0,010	1,314
15	<i>AhySWEET20/28</i>	13/14	0,010	0,025	0,390

Phân tích cấu trúc vùng bảo thủ, mô hình của *AhySWEET* ở lạc

Thông thường, họ protein SWEET ở thực vật thường chứa bảy cấu trúc xoắn cuộn xuyên màng (*trans*-membrane helices, TMH) (Patil *et al.*, 2015) nhằm đảm bảo chức năng vận chuyển sucrose trong tế bào (Chen, 2014). Trong nghiên cứu này, vùng bảo thủ của họ protein *AhySWEET* ở lạc được xác định bằng TMHMM (Chaturvedi *et al.*, 2011). Kết quả cho thấy, hầu hết các *AhySWEET* (35/43) đều có 7 đoạn TMH đặc trưng. Cấu trúc đặc trưng này bao gồm hai lần lặp của ba nhóm TM ('TM1-TM2-TM3' và 'TM5-TM6-TM7') được chia tách bởi một cấu trúc TM đơn lẻ (Hình 2), tương tự như trên các loài thực vật khác. Trước đây, các nghiên cứu về protein GmSWEET trên đậu tương (Patil *et al.*, 2015) và BoSWEET ở cải bắp (Zhang *et al.*, 2019) đã chỉ ra rằng, tất cả các phân tử này có cấu trúc gồm bảy vùng TMH đều phân bố trên các màng với chức năng như những thụ thể trên màng. Các dữ liệu này cho thấy đa số thành viên trong họ protein *AhySWEET* ở lạc có thể cư trú trên các màng bào quan để thực hiện quá trình vận chuyển sucrose qua màng.



Hình 2. Cấu trúc xoắn cuộn xuyên màng của họ *AhySWEET* ở lạc

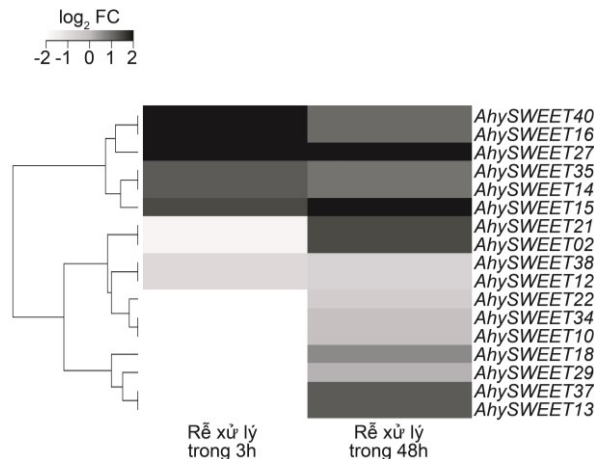
Để xem xét rõ hơn về vị trí cư trú nội bào của họ protein AhySWEET ở lạc, trình tự protein lần lượt được truy vấn trên TargetP (Almagro *et al.*, 2019). Kết quả dự đoán vị trí cư trú nội bào của phân tử AhySWEET dựa trên sự có mặt của các đoạn cTP (đặc trưng cho lục lạp), mTP (đặc trưng cho ty thể), SP (đặc trưng cho các bào quan tiết như mạng lưới nội chất và thể Golgi) được thể hiện ở bảng 2. Hầu hết các phân tử AhySWEET (38/43) cư trú ở hệ thống bao gói, ngoại trừ protein AhySWEET04 phân bố ở ty thể (Bảng 2). Trong khi đó, protein AhySWEET07, 12, 38 và 42 chưa rõ vị trí phân bố trong tế bào. Kết quả này cũng phù hợp với phân tích về vùng TMH của họ protein AhySWEET ở lạc.

Bảng 2. Vị trí phân bố nội bào của họ protein AhySWEET ở lạc

Tên protein	cTP	mTP	SP	LC	RC	Tên protein	cTP	mTP	SP	LC	RC
AhySWEET01	0,009	0,015	0,794	S	4	AhySWEET23	0,150	0,014	0,639	S	4
AhySWEET02	0,004	0,013	0,921	S	3	AhySWEET24	0,006	0,016	0,938	S	2
AhySWEET03	0,155	0,014	0,629	S	4	AhySWEET25	0,005	0,021	0,855	S	3
AhySWEET04	0,005	0,674	0,150	M	5	AhySWEET26	0,006	0,023	0,791	S	3
AhySWEET05	0,005	0,017	0,930	S	2	AhySWEET27	0,012	0,071	0,618	S	5
AhySWEET06	0,010	0,024	0,755	S	3	AhySWEET28	0,002	0,017	0,961	S	2
AhySWEET07	0,004	0,268	0,259	-	2	AhySWEET29	0,013	0,025	0,838	S	3
AhySWEET08	0,026	0,099	0,452	S	5	AhySWEET30	0,026	0,098	0,459	S	5
AhySWEET09	0,019	0,046	0,572	S	5	AhySWEET31	0,020	0,045	0,611	S	5
AhySWEET10	0,067	0,029	0,397	S	5	AhySWEET32	0,027	0,084	0,608	S	5
AhySWEET11	0,004	0,037	0,797	S	4	AhySWEET33	0,004	0,037	0,797	S	4
AhySWEET12	0,062	0,031	0,468	-	5	AhySWEET34	0,068	0,031	0,353	S	5
AhySWEET13	0,011	0,009	0,888	S	2	AhySWEET35	0,014	0,117	0,617	S	5
AhySWEET14	0,014	0,117	0,617	S	5	AhySWEET36	0,006	0,022	0,931	S	3
AhySWEET15	0,010	0,006	0,935	S	2	AhySWEET37	0,016	0,008	0,847	S	3
AhySWEET16	0,010	0,029	0,836	S	3	AhySWEET38	0,062	0,031	0,468	-	5
AhySWEET17	0,013	0,046	0,629	S	5	AhySWEET39	0,011	0,009	0,852	S	3
AhySWEET18	0,004	0,021	0,869	S	3	AhySWEET40	0,010	0,031	0,835	S	3
AhySWEET19	0,009	0,015	0,794	S	4	AhySWEET41	0,049	0,007	0,883	S	2
AhySWEET20	0,002	0,018	0,986	S	2	AhySWEET42	0,014	0,053	0,575	-	5
AhySWEET21	0,004	0,013	0,921	S	3	AhySWEET43	0,003	0,020	0,875	S	3
AhySWEET22	0,003	0,020	0,871	S	3	-	-	-	-	-	-

Đánh giá dữ liệu biểu hiện của họ gene AhySWEET ở lạc trong điều kiện xử lý mặn

Để đánh giá biểu hiện của họ gene AhySWEET, nghiên cứu này đã khai thác và phân tích dữ liệu RNA-Seq ở mẫu rễ lạc trong điều kiện mặn (Chen *et al.*, 2016). Kết quả đã cho thấy, 17 (trên tổng số 43) gene AhySWEET có thông tin về mức độ biểu hiện trong điều kiện xử lý mặn (Hình 3). Trong đó, họ gene AhySWEET có hai xu hướng đáp ứng trong giai đoạn đầu của xử lý mặn (3h sau xử lý). Cụ thể, chín gene, AhySWEET02/10/13/18/21/22/29/34 và 37 giảm mức độ phiên mã ở rễ xử lý mặn ở thời điểm 3h, trong khi ba gene, AhySWEET16/27 và 40 tăng cường biểu hiện trong giai đoạn này (Hình 3). Tiếp theo, các gene AhySWEET đều có xu hướng tăng cường biểu hiện ở rễ ở thời điểm 48h xử lý mặn, ngoại trừ hai gene, AhySWEET16 và 40, giảm biểu hiện (Hình 3).



Hình 3. Biểu hiện của gene AhySWEET ở rễ trong điều kiện xử lý mặn

KẾT LUẬN VÀ ĐỀ NGHỊ

Đã xác định được 15 sự kiện lặp gene xuất phát từ hiện tượng lặp đoạn trên các nhiễm sắc thể xảy ra trong họ gene *AhySWEET* ở lạc. Phân tích cấu trúc protein cho thấy 35/43 protein *AhySWEET* đều chứa bảy vùng xoắn cuộn xuyên màng đặc trưng để đảm bảo cho chức năng vận chuyển sucrose nội bào. Đáng chú ý, 38/43 *AhySWEET* nằm ở hệ thống tiết nội bào. Đánh giá dữ liệu RNA-Seq cho thấy *AhySWEET13/37* và *10/34* có đáp ứng giảm, trong khi *AhySWEET16/40* có đáp ứng tăng ở rễ trong giai đoạn đầu của xử lý mặn.

Lời cảm ơn: Nghiên cứu này được thực hiện từ kinh phí của đề tài nghiên cứu cơ bản mã số 08/HĐĐT-KHCN do Trường Đại học Sư phạm Hà Nội 2 tài trợ.

TÀI LIỆU THAM KHẢO

- Almagro JJ, Salvatore M, Emanuelsson O, Winther O, von Heijne G, Elofsson A, Nielsen H (2019). Detecting sequence signals in targeting peptides using deep learning. *Life Sci Alliance* 2(5): e201900429.
- Bertioli DJ, Jenkins J, Clevenger J, Dudchenko O, Gao D, Seijo G, Leal-Bertioli SCM, Ren L, Farmer AD, Pandey MK, Samoluk SS, Abernathy B, Agarwal G, Ballén-Taborda C, Cameron C, Campbell J, Chavarro C, Chitikineni A, Chu Y, Dash S, El Baidouri M, Guo B, Huang W, Kim KD, Korani W, Lanciano S, Lui CG, Mirouze M, Moretzsohn MC, Pham M, Shin JH, Shirasawa K, Sinharoy S, Sreedasyam A, Weeks NT, Zhang X, Zheng Z, Sun Z, Froenicke L, Aiden EL, Michelmore R, Varshney RK, Holbrook CC, Cannon EKS, Scheffler BE, Grimwood J, Ozias-Akins P, Cannon SB, Jackson SA, Schmutz J (2019). The genome sequence of segmental allotetraploid peanut *Arachis hypogaea*. *Nat Genet* 51(5): 877-884.
- Chaturvedi N, Shanker S, Singh VK, Sinha D, Pandey PN (2011). Hidden markov model for the prediction of transmembrane proteins using MATLAB. *Bioinfo* 7(8): 418-421.
- Chen LQ (2014). SWEET sugar transporters for phloem transport and pathogen nutrition. *New Phytol* 201(4): 1150-1155.
- Chen N, Su M, Chi X, Zhang Z, Pan L, Chen M, Wang T, Wang M, Study L, Yu S (2016). Transcriptome analysis reveals salt-stress-regulated biological processes and key pathways in roots of peanut (*Arachis hypogaea* L.). *Gene Genet* 38: 493-507.
- Chu HD, Nguyen KH, Watanabe Y, Le DT, Pham TLT, Mochida K, Tran LP (2018). Identification, structural characterization and gene expression analysis of members of the Nuclear factor-Y family in chickpea (*Cicer arietinum* L.) under dehydration and abscisic acid treatments. *Int J Mol Sci* 19(11): E3290.
- Ha DC, Quynh TNL, Yen HTH, Thu PP, Hong VL, Nguyet MTN, Anh XD, Thao DL, Linh HL (2020). Identification, structural characterization and *in silico* expression analysis of the sucrose transporter 'SWEET' gene family in peanut (*Arachis hypogaea*). *Vietnam J Sci Technol Engin* 62(2): 5-10.
- Jian H, Lu K, Yang B, Wang, T, Zhang L, Zhang A, Wang J, Liu L, Qu, C, Li J (2016). Genome-wide analysis and expression profiling of the SUC and SWEET gene families of sucrose transporters in oilseed rape (*Brassica napus* L.). *Front Plant Sci* 7: 1464.
- Kumar S, Stecher G, Tamura K (2016). MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol* 33(7): 1870-1874.
- Liao Y, Smyth GK, Shi W (2019). The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Res* 47(8): e47.
- Miao H, Sun P, Liu Q, Miao Y, Liu J, Zhang K, Hu W, Zhang J, Wang J, Wang Z, Jia C, Xu B, Jin Z. (2017). Genome-wide analyses of SWEET family proteins reveal involvement in fruit development and abiotic/biotic stress responses in banana. *Sci Reps* 7(1): 3536.
- Patil G, Valliyodan B, Deshmukh R, Prince S, Nicander B, Zhao M, Sonah H, Song L, Lin L, Chaudhary J, Liu Y, Joshi T, Xu D, Nguyen HT (2015). Soybean (*Glycine max*) SWEET gene family: Insights through comparative genomics, transcriptome profiling and whole genome re-sequencing analysis. *BMC Genomics* 16: 520.
- Rozas J, Ferrer-Mata A, Sanchez-DelBarrio C, Guirao-Rico S, Librado P, Ramos-Onsins SE, Sanchez-Gracia A. (2017). DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Mol Biol Evol* 34(12): 3299-3302.
- Toomer OT (2018). Nutritional chemistry of the peanut (*Arachis hypogaea*). *Crit Rev Food Sci Nutr* 58(17): 3042-3053.
- Zhang W, Wang S, Yu F, Tang J, Shan X, Bao K, Yu L, Wang H, Fei Z, Li J (2019). Genome-wide characterization and expression profiling of SWEET genes in cabbage (*Brassica oleracea* var. *capitata* L.) reveal their roles in chilling and clubroot disease responses. *BMC Genomics* 20(1): 93.

POTENTIAL ROLES OF THE DUPLICATION EVENTS IN THE GENES ENCODING SUCROSE TRANSPORTER IN THE SALT STRESS RESPONSES IN PEANUT (*Arachis hypogaea*)

Chu Duc Ha¹, Hoang Thi Hai Yen^{1,2}, Le Thi Ngoc Quynh³, Pham Thi Ngoc Bich^{1,4}, Hoang Minh Chinh^{1,4}, La Viet Hong², Pham Phuong Thu², Nguyen Thi Minh Nguyet¹, Le Hung Linh¹, Le Huy Ham^{1,5}, Nguyen Quoc Trung⁴, Pham Xuan Hoi¹

¹ Agricultural Genetics Institute, Vietnam Academy of Agricultural Sciences

² Faculty of Biology - Agricultural engineering, Hanoi Pedagogical University 2

³ Faculty of Chemistry and Environment, Thuyloi University

⁴ Faculty of Biotechnology, Vietnam National University of Agriculture

⁵ Faculty of Agricultural technology, University of Engineering and Technology, Vietnam National University

SUMMARY

In this study, a total of 15 duplication events has been found and analyzed in the *AhySWEET* gene family in peanut. All duplicated genes were noted to occur from the segmental duplication events in the peanut's genome. The synonymous (Ks) and non-synonymous (Ka) mutation rate of these events indicated the strongly purifying selection occurring in the evolution of these *AhySWEET* gene families. Next, we also found that the most of *AhySWEET* proteins contained seven *trans*- membrane helices and were mapped on the secretory pathway. According to the transcriptome atlas, we found that the duplicated genes shared the similar expression patterns in roots under the salt stress treatments. Particularly, the expression of *AhySWEET13/37* and *10/34* were reduced, whereas, that for *AhySWEET16/40* were strongly increased in roots at the early stage of the treatments. Taken together, our study would provide the fundamental information for further functional characterization of *AhySWEET* genes of peanut in multi-stress response.

Keywords: Gene duplication, peanut, salt, SWEET, sucrose transporter.