

# PHÂN TÍCH HỌ GENE *KNOX* Ở CÂY DỨA (*ANANAS COMOSUS* L.) BẰNG PHƯƠNG PHÁP *IN SILICO*

## ANALYSIS OF *KNOX* GENE FAMILY IN PINEAPPLE (*ANANAS COMOSUS* L.) BY USING *IN SILICO* METHODS

Cao Phi Bằng

Trường Đại học Hùng Vương, Phú Thọ; phibang.cao@hvu.edu.vn

**Tóm tắt** - Nhờ sử dụng các phương pháp nghiên cứu *in silico*, tác giả đã xác định và phân tích 10 gene *KNOX* ở trong hệ gene của cây dứa (*Ananas comosus* L.). Căn cứ vào kết quả phân tích cấu trúc và cây phả hệ, các gene *KNOX* của cây dứa được phân chia thành hai phân nhóm là phân nhóm I (năm gene) và phân nhóm II (năm gene). Các gene *KNOX* của cây dứa có mức độ bảo tồn cao về cấu trúc exon/intron. Các *KNOX* của cây dứa đều có đủ bốn trình tự bảo thủ đặc trưng. Phân tích Gene Ontology cho thấy cả 10 gene *KNOX* của cây dứa đều có vai trò điều hòa sinh học. Phân tích dữ liệu RNA-seq cho thấy cả 10 gene *KNOX* của cây dứa đều biểu hiện ở tất cả các mô nghiên cứu. Trong đó, nhiều gene (6/10 gene) biểu hiện yếu nhất ở hoa khi so với các mô khác. Đặc biệt, hầu hết các gene biểu hiện mạnh trong quả ở các thời kì phát triển khác nhau.

**Từ khóa** - biểu hiện gene; cây dứa (*Ananas comosus* L.); cây phả hệ; *KNOX*; *in silico*

### 1. Đặt vấn đề

Các protein KNotted-like homeobox (KNOX) thuộc về họ protein có chứa vùng bảo thủ homeodomain. Các protein này có cấu trúc giống với protein được mã hóa bởi gene *knotted1* được phát hiện lần đầu ở cây ngô [14]. Dựa trên cấu trúc của gene, sự biểu hiện và các phân tích cây phả hệ, các *KNOX* ở thực vật được xếp vào ba phân nhóm [8, 12]. Phân nhóm I gồm các gene có vùng bảo thủ homeodomain gần giống với gene *KNI* (> 71%), các gene này biểu hiện trong các vùng giao nhau của các mô phân sinh đỉnh chồi nhưng biểu hiện thấp ở lá. Các protein phân nhóm I giữ vai trò quan trọng trong việc duy trì chức năng của mô phân sinh đỉnh chồi và trong quá trình phát sinh cơ quan. Cây *Arabidopsis thaliana* có bốn gene mã hóa protein phân nhóm I là *KNAT1* (*BREVIPEDICELLUS*), *KNAT2*, *KNAT6* và *STM* (*SHOOTMERISTEMLESS*). Trong đó, *STM* giữ vai trò chủ yếu trong việc hình thành và duy trì mô phân sinh đỉnh, hai gene *KNAT1* và *KNAT6* tham gia vào việc duy trì chức năng của mô phân sinh đỉnh chồi, đồng thời tham gia điều khiển sự phát triển của mầm hoa [5, 6]. Phân nhóm II gồm bốn gene *KNAT3*, *KNAT4*, *KNAT5* và *KNAT7* với đặc điểm chung là mang một intron trong vùng ELK, biểu hiện đa dạng [5]. Phân nhóm III chỉ gồm một gene *KNOXM* mới được xác định, mã hóa protein thiếu vùng bảo thủ homeodomain và biểu hiện ở các vùng bên của các tiền cơ quan và ở lớp bao của các cơ quan thành thực [9].

Ở mức độ hệ gene hoàn chỉnh, lần lượt tám, mười và mười lăm các gene *KNOX* đã được xác định trong hệ gene của các cây hai lá mầm như *A. thaliana* (chín gene), cây đào (10 gene) và cây sồi (15 gene), cây quýt (10 gene), cây đậu cove [1], cây ngô (12 gene) và cây lúa (13 gene) [2, 5, 6, 15].

Cây dứa (*Ananas comosus* L.) là cây trồng có giá trị kinh tế quan trọng nhất trong nhóm cây mọng nước

**Abstract** - By using the *in silico* method, the author have identified and analyzed a total of 10 *KNOX* genes in the genome of Pineapple (*Ananas comosus* L.). Based on the protein structure and the phylogenetic analysis, the pineapple *KNOX* genes are divided into two subgroups, group I (five genes) and group II (five genes). The pineapple *KNOX* genes have a highly conserved level of exon/intron structure. All of ten *KNOX* proteins include four conserved domains. Gene Ontology analysis shows that all pineapple *KNOX* genes play an important role in biological regulation. RNA-seq analysis shows that all of ten *KNOX* genes are expressed in all study tissues. Many genes (6/10 genes) show the weakest expression level in flowers compared to other tissues. Especially, most of these genes have strong expression in fruits at different development stages.

**Key words** - gene expression; pineapple (*Ananas comosus* L.); phylogenetic tree; *KNOX*; *in silico*

(metabolism acid crassulacean, CAM) và là cây ăn trái nhiệt đới quan trọng thứ hai. Hệ gene của cây dứa được giải trình tự tạo cơ hội cho các nghiên cứu phân tích cấu trúc, chức năng của các gene, họ gene của loài cây này, cung cấp thông tin khoa học có giá trị trong công tác chọn tạo giống [10]. Nghiên cứu này hướng tới mục tiêu phân tích các đặc điểm của các gene *KNOX* ở cây dứa, một đại diện điển hình của thực vật CAM, bằng các phương pháp tin sinh học.

### 2. Giải quyết vấn đề

#### 2.1. Trình tự gene của cây dứa

Cơ sở dữ liệu về trình tự hệ gene (phiên bản v3) của cây dứa được lấy từ website phytozome (<http://www.phytozome.net/>) được giải trình tự bởi Ming và cộng sự [10, 11].

#### 2.2. Xác định các gene *KNOX* ở cây dứa

Các trình tự protein *KNOX* của cây lúa lấy từ Xiong và cộng sự [15] được dùng làm khuôn dò, chương trình TBLASTN được sử dụng để tìm kiếm các gene tương đồng trên toàn hệ gene của cây dứa.

#### 2.3. Xây dựng cây phả hệ

Trình tự protein suy diễn *KNOX* của cây dứa, cây lúa và *A. thaliana* được sắp dãy bằng MAFFT [7], cây phả hệ được xây dựng nhờ phần mềm MEGA7 [8].

#### 2.4. Phân tích tin sinh học các gene *KNOX* của cây dứa

Các đặc điểm vật lý, hóa học của các gene/protein nghiên cứu được khảo sát nhờ các công cụ ExpASY ProtParam (Expert Protein Analysis System) [4]. Vị trí khu trú dưới tế bào được phân tích nhờ ProtComp 9.0 (<http://linux1.softberry.com/berry.phtml?topic=protcomppl&group=help&subgroup=proloc>). Chức năng của các gene được phân tích nhờ công cụ agriGO (<http://bioinfo.cau.edu.cn/agriGO>) [13].

### 2.5. Phân tích sự biểu hiện gene

Sự biểu hiện của các gene được xác định qua phân tích dữ liệu RNA-seq đã được đặt trên ngân hàng Sequence read archive (SRA) (SRX1465561, SRX1465570, SRX1465592, SRX1465595, SRX1465596, SRX1465597, SRX1465632, SRX1465633, SRX1465636) của cây dứa, được lấy từ Ming và cộng sự [10]. Giá trị biểu hiện gene được tính bằng  $\log_{10}(\text{RPKM})$ ,  $\text{RPKM} = \text{Reads per kilo base per million mapped reads}$  (Số bản mã phiên (ARN)/(chiều dài đoạn gene mã hóa/1.000 nucleotide x tổng số mã phiên/1.000.000).

### 3. Kết quả nghiên cứu và bàn luận

**Bảng 1.** Các trình tự KNOX ở cây dứa

Gene	Phân nhóm	Tên locus	GS (bp)	PL (aa)	MW (kD)	pI	GRAVY	NST	I	SCL
<i>AcKNAT1</i>	I	Aco015871	7834	330	37,55	6,54	-0,74	1	4	N
<i>AcKNAT2</i>	I	Aco015873	7486	409	45,14	6,30	-0,75	1	4	N
<i>AcKNAT3</i>	II	Aco016892	9254	323	35,83	5,85	-0,73	2	4	N
<i>AcKNAT4</i>	I	Aco002548	7080	315	34,92	5,55	-0,66	4	4	N
<i>AcKNAT5</i>	I	Aco012803	10294	327	36,66	5,07	-0,69	5	4	N
<i>AcKNAT6</i>	I	Aco004983	5982	391	43,46	6,24	-0,73	7	4	N
<i>AcKNAT7</i>	II	Aco016636	3473	304	33,84	5,96	-0,53	8	4	N
<i>AcKNAT8</i>	II	Aco006448	5246	437	46,37	5,69	-0,50	14	4	N
<i>AcKNAT9</i>	II	Aco008183	4196	342	36,88	5,26	-0,57	19	5	N
<i>AcKNAT10</i>	II	Aco028389	7430	335	37,55	5,29	-0,77	2079*	5	N

Chú thích: GS = Kích thước gen, PL = Chiều dài phân tử protein, MW = Khối lượng phân tử protein, pI = Điểm đẳng điện, NST = Nhiễm sắc thể, I = Số lượng intron, SCL = Khu trú dưới tế bào, N = Nhân tế bào \* scaffold\_2079: chưa được lắp ráp vào nhiễm sắc thể nào

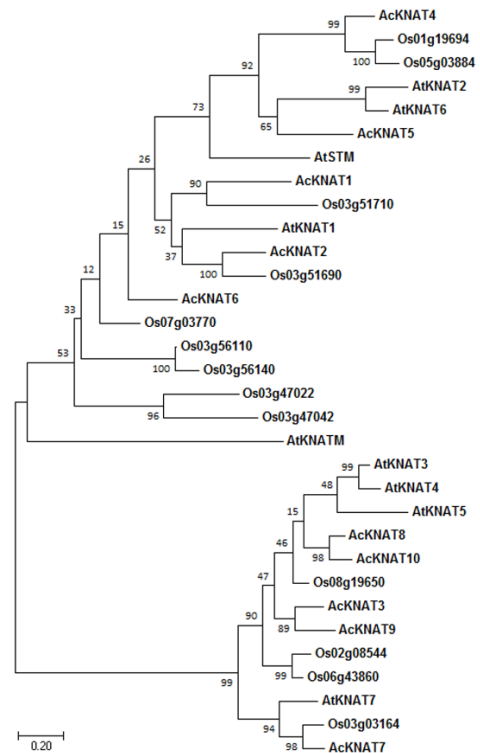
### 3.2. Phân tích đặc điểm và cây phả hệ của các KNOX ở cây dứa

Các gene KNOX của cây dứa có độ dài trình tự nucleotide khác nhau, gene dài nhất (*AcKNAT5*) chứa tới 10.294 nucleotide, trong khi gene ngắn nhất (*AcKNAT7*) chỉ có 3.473 nucleotide. Các gene này mã hóa cho các protein có độ dài từ 304 tới 437 amino acid, khối lượng phân tử của các protein suy diễn nằm trong khoảng 38,84 kDa tới 46,37 kDa. Các protein suy diễn có tính axit yếu hoặc trung tính, giá trị điểm đẳng điện (pI) lý thuyết của các KNAT ở cây dứa khá tương đồng, dao động từ 5,07 tới 6,54 (Bảng 1). Các đặc điểm này khá tương đồng với đặc điểm của các KNOX của cây quýt đường [2] và cây đậu cove [1].

Trình tự protein suy diễn của các KNOX của cây dứa được đóng hàng cùng với các trình tự tương đồng ở cây *A. thaliana* và cây lúa nhờ phần mềm MAFFT (<https://mafft.cbrc.jp/alignment/server/>) được sử dụng để tạo cây phả hệ (Hình 1). Cây phả hệ gồm có ba nhánh, trong đó, AtKNATM tạo thành một nhánh riêng, hai nhánh có đại diện của các KNOX ở cả ba loài. Phân tích cây phả hệ cho phép phân chia các KNOX của cây dứa thành hai phân nhóm. Phân nhóm I gồm có *AcKNAT1*, *AcKNAT2*, *AcKNAT4*, *AcKNAT5* và *AcKNAT6*, các protein này nằm cùng nhánh với AtSTM, AtKNAT1, AtKNAT2 và AtKNAT6 của cây *A. thaliana*. Phân nhóm 2 gồm năm protein còn lại, chúng nằm cùng nhánh với AtKNAT3, AtKNAT4, AtKNAT5 và AtKNAT7.

### 3.1. Xác định họ gene KNOX ở cây dứa

Khi sử dụng các protein KNOX của cây lúa làm khuôn dò để chạy chương trình TBLASTN trên hệ gene của cây dứa, tác giả đã xác định được tổng số 10 gene có thể mã hóa cho các KNOX (Bảng 1). Kiểm tra Pfam [3] các protein suy diễn chỉ ra chúng có chứa các vùng bảo thủ KNOX1 (PF03790) và KNOX2 (PF03791), Homeodomain (Homeobox\_KN, PF05920) và ELK (PF03789). Khi so với một số loài cây khác, dứa có số lượng gene KNOX tương đương với các cây hai lá mầm thân gỗ như cây quýt đường [2], cây đào [6] nhưng ít hơn so với các cây một lá mầm như cây lúa (13 gene) và cây ngô (12 gene) [6].

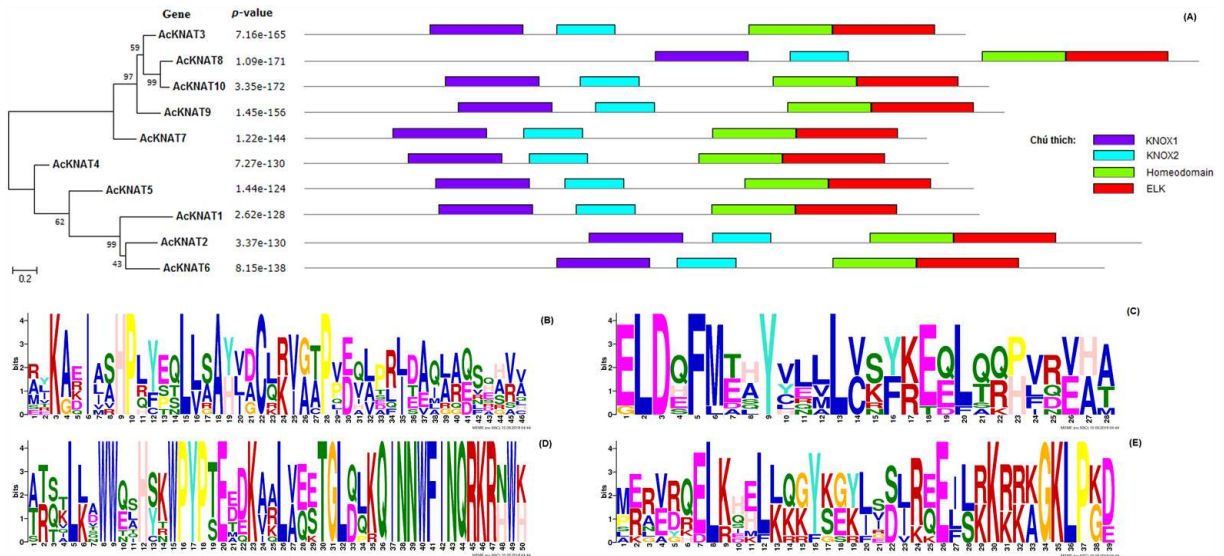


**Hình 1.** Cây phả hệ được thiết lập từ các KNOX ở ba loài *A. thaliana* (At), lúa (Os) và dứa (Ac) nhờ phần mềm MEGA7 [8]. Giá trị bootstraps được thể hiện ở gốc của nhánh. Tỷ lệ xích thể hiện xác suất amino acid bị biến đổi tại từng vị trí

Quan sát cây phả hệ cũng cho thấy có thể có hai sự kiện nhân gene trên quy mô hệ gene xảy ra với phân nhóm 2 của họ KNOX ở cây dứa. Sự kiện nhân gene thứ nhất hình thành nên hai cặp gene paralog *AcKNAT3* (nằm trên nhiễm sắc thể số 2) và *AcKNAT9* (nằm trên nhiễm sắc thể số 19). Trong khi đó, hai gene *AcKNAT8* (nằm trên nhiễm sắc thể số 14) và gene *AcKNAT10* (nằm trên scaffold\_2079) có thể là kết quả của sự kiện nhân gene thứ hai. Tuy nhiên, trường hợp cặp gene *AcKNAT8* và *AcKNAT10* còn cần được kiểm tra đối với các phiên bản hệ gene cây dứa sau, khắc phục được những lỗi ghép nối hệ gene hiện nay vì scaffold\_2079

chỉ có kích thước 16.983 nucleotide và chưa được ghép vào nhiễm sắc thể nào của cây dứa.

Các vùng bảo tồn của protein suy diễn họ KNOX của cây dứa được phân tích bằng chương trình MEME, kết quả cho thấy tất cả các KNOX của cây dứa đều mang 4 vùng bảo tồn: KNOX1, KNOX2, HOMEODOMAIN và ELK (Hình 2). Các motif bảo tồn của các vùng KNOX1, KNOX2, HOMEODOMAIN và ELK của các protein KNOX cũng đã được chỉ ra trong Hình 2.



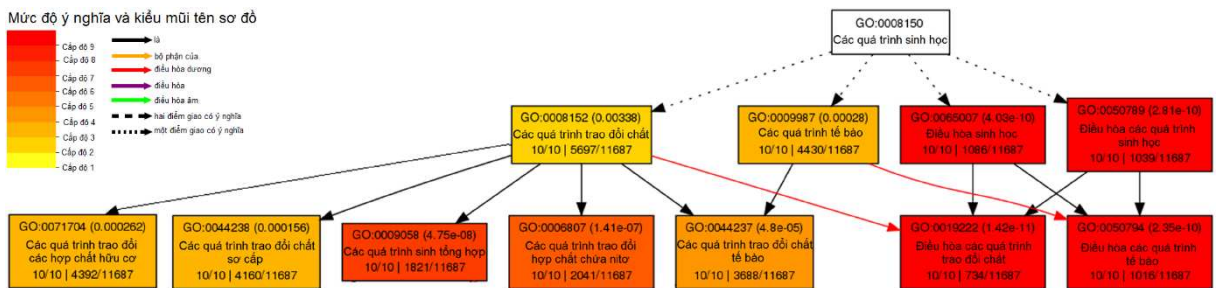
**Hình 2.** Các vùng bảo tồn trong phân tử protein KNOX của cây dứa, (A) vị trí tương đối các vùng bảo tồn, (B) motif vùng bảo tồn KNOX1, (C) motif vùng bảo tồn KNOX2, (D) motif vùng bảo tồn Homeodomain và (E) motif vùng bảo tồn ELK

**3.3. Vị trí khu trú dưới tế bào và phân tích chức năng của các gene**

Sử dụng công cụ ProtComp 9.0 để dự đoán vị trí khu trú lý thuyết của các protein KNOX của cây dứa, kết quả phân tích cho thấy tất cả các protein này đều có vị trí khu trú trong nhân tế bào. Kết quả này cũng phù hợp với kết quả phân tích

các KNOX của cây lúa và cây dương [15], đồng thời cũng phù hợp với các nghiên cứu trước đây chứng minh vai trò tác nhân điều hòa phiên mã của các KNOX [6].

Chức năng các KNOX của cây dứa được dự đoán nhờ phân tích ý nghĩa của gene (Gene Ontology) với công cụ agriGO v2.0 [13] (Hình 3).



**Hình 3.** Phân tích chức năng (GO) của các gene KNOX của cây dứa

Kết quả phân tích cho thấy cả 10 gene KNOX của cây dứa đều có chức năng trong các quá trình sinh học khác nhau như các quá trình trao đổi chất, các quá trình tế bào, điều hòa sinh học và điều hòa các quá trình sinh học. Cụ thể, cả 10 KNOX của cây dứa tham gia vào các quá trình trao đổi các hợp chất hữu cơ, các chất sơ cấp, các quá trình sinh tổng hợp, các hợp chất chứa nitơ và các quá trình trao đổi chất tế bào (cấp độ 4). Đặc biệt, các KNOX của cây dứa thể hiện chức năng điều hòa sinh học thông qua điều hòa các quá trình trao đổi chất và các quá trình tế bào ở cấp độ tin cậy cao (cấp độ 9). Kết quả nghiên cứu này khẳng định kết quả nghiên cứu về chức năng của các

KNOX ở thực vật [15].

**3.4. Phân tích sự biểu hiện gene**

Sự biểu hiện gene KNOX của cây dứa được phân tích thông qua dữ liệu RNA-seq đã được đặt trên hệ dữ liệu SRA của cây dứa với các mô rễ, lá, hoa và quả ở các giai đoạn phát triển khác nhau (Hình 4). Tất cả các gene KNOX của cây dứa đều biểu hiện ở tất cả các mô phân tích trong ngân hàng dữ liệu SRA được xây dựng bởi Ming và cộng sự [10]. Đáng chú ý, khi so sánh sự biểu hiện của cùng gene trong các loại mô khác nhau, có tới 6/10 gene biểu hiện

thấp nhất ở hoa, hai gene biểu hiện thấp nhất ở rễ (*AcKNAT2* và *AcKNAT3*), hai gene biểu hiện thấp nhất ở quả giai đoạn S1 (*AcKNAT8* và *AcKNAT10*). Chỉ duy nhất gene *AcKNAT3* biểu hiện mạnh nhất ở hoa trong số các mô được phân tích, các gene còn lại biểu hiện mạnh nhất ở quả (một trong các giai đoạn phát triển). Lần đầu tiên sự biểu hiện của các gene *KNOX* ở quả được báo cáo khi phân tích dữ liệu RNA-seq. Trong những nghiên cứu trước đây, sự biểu hiện của các gene *KNOX* được phân tích chủ yếu ở đỉnh sinh trưởng như ở *A. thaliana* [5], ở cây đậu cove [1], cây quýt đường [2]. Gần đây, sự biểu hiện của các gene *KNOX* ở các mô sinh sản như hoa cũng đã được xác định như ở cây đậu cove [1], cây lúa [15]. Sự biểu hiện của các gene *KNOX* ở tất cả các mô nghiên cứu của cây dứa gợi ý rằng các gene này có vai trò trong sự phát triển của cây dứa ở các giai đoạn phát triển khác nhau.



**Hình 4.** Sự biểu hiện của các gene *KNOX* trong các cơ quan của cây dứa (đơn vị tính  $\log_{10}$  (RPKM), Số bản mã phiên (ARN)/(chiều dài đoạn mã hóa/1.000 nucleotide \* tổng số mã phiên/1.000.000)

Chú thích: R = Rễ; L = Lá; F = Hoa; S1, S2, S3, S4, S5 và S7 = Quả ở các giai đoạn 1, 2, 3, 4, 5 và 7.

#### 4. Kết luận

Họ gene *KNOX* ở cây dứa đã được xác định và phân tích bằng phương pháp *tin sinh học*. Trong bộ gene của cây dứa có tổng số 10 gene mã hóa *KNOX* được xác định. Cấu trúc gene và các đặc điểm của các protein *KNOX* cũng được phân tích. Dựa trên cây phả hệ, có thể phân loại các gene *KNOX* của cây dứa thành hai phân nhóm I và II. Phân tích Gene Ontology cho thấy các *KNOX* của cây dứa có vai trò đa dạng, nhất là chức năng điều hòa sinh học. Sự biểu hiện gene của các gene *KNOX* ở rễ, lá, hoa và quả dứa ở các giai đoạn khác nhau đã được phân tích nhờ hệ dữ liệu RNA-seq đặt trên hệ dữ liệu SRA vốn được xây dựng bởi Ming và cộng sự. (2015). Tất cả các gene *KNOX* của cây dứa đều biểu hiện ở trong các mô nghiên cứu, tuy nhiên với mức độ khác nhau tùy gene và tùy mô. Kết quả này có ý nghĩa lớn, mở đường cho việc tách dòng gene và phân tích chức năng của các gene trong họ *KNOX* ở các cây mọng nước một lá mầm khác.

**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.

(BBT nhận bài: 05/7/2018, hoàn tất thủ tục phản biện: 08/8/2018)

#### TÀI LIỆU THAM KHẢO

- [1] Cao Phi Bằng, “Phân tích in silico họ gen *KNOX* ở cây đậu cove (*Phaseolus vulgaris* L.)”, *Chuyên san CNSH, Tạp chí Di truyền và ứng dụng*, Số 10, 2015, trang 4-11.
- [2] Cao Phi Bằng, Dunand C., “Xác định và phân tích in silico họ gen *KNOX* ở cây quýt đường (*Citrus clementina*)”, *Tạp chí Công nghệ Sinh học*, 12(1), 2014, trang 113-123.
- [3] Finn R. D., Bateman A., Clements J., Coggill P., Eberhardt R. Y., Eddy S. R., Heeger A., Hetherington K., Holm L., Mistry J., Sonnhammer E. L., Tate J., & Punta M., “Pfam: The Protein Families Database”, *Nucleic Acids Res*, 42(Database issue), 2014, D222-230. doi:10.1093/nar/gkt1223.
- [4] Gasteiger E., Hoogland C., Gattiker A., Wilkins M. R., Appel R. D., & Bairoch A., *Protein Identification and Analysis Tools on the ExPASy Server*, The Proteomics Protocols Handbook, Springer, 2005, pp. 571-607.
- [5] Hake S., Smith H. M., Holtan H., Magnani E., Mele G., & Ramirez J., “The Role of *KNOX* Genes in Plant Development”, *Annu Rev Cell Dev Biol*, 20, 2004, pp. 125-151.
- [6] Hay A., & Tsiantis M., “*KNOX* Genes: Versatile Regulators of Plant Development and Diversity”, *Development*, 137(19), 2010, pp. 3153-3165.
- [7] Katoh K., & Standley D. M., “MAFFT Multiple Sequence Alignment software Version 7: Improvements in Performance and Usability”, *Mol Biol Evol*, 30(4), 2013, pp. 772-780.
- [8] Kumar S., Stecher G., & Tamura K., “MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets”, *Mol Biol Evol*, 33(7), 2016, pp. 1870-1874.
- [9] Magnani E., & Hake S., “*KNOX* Lost the OX: the Arabidopsis *KNATM* Gene Defines a Novel Class of *KNOX* Transcriptional Regulators Missing the Homeodomain”, *Plant Cell*, 20(4), 2008, pp. 875-887.
- [10] Ming R., VanBuren R., Wai C. M., Tang H., Schatz M. C., Bowers J. E., Lyons E., Wang M. L., Chen J., Biggers E., Zhang J., Huang L., Zhang L., Miao W., Zhang J., Ye Z., Miao C., Lin Z., Wang H., Zhou H., Yin W. C., Priest H. D., Zheng C., Woodhouse M., Edger P. P., Guyot R., Guo H. B., Guo H., Zheng G., Singh R., Sharma A., Min X., Zheng Y., Lee H., Gurtowski J., Sedlaczek F. J., Harkess A., McKain M. R., Liao Z., Fang J., Liu J., Zhang X., Zhang Q., Hu W., Qin Y., Wang K., Chen L. Y., Shirley N., Lin Y. R., Liu L. Y., Hernandez A. G., Wright C. L., Bulone V., Tuskan G. A., Heath K., Zee F., Moore P. H., Sunkar R., Leebens-Mack J. H., Mockler T., Bennetzen J. L., Freeling M., Sankoff D., Paterson A. H., Zhu X., Yang X., Smith J. A., Cushman J. C., Paull R. E., & Yu Q., “The Pineapple Genome and The Evolution of CAM Photosynthesis”, *Nature Genetics*, 47, 2015, pp. 1435-1442.
- [11] Ming R., Wai C. M., & Guyot R., “Pineapple Genome: A Reference for Monocots and CAM Photosynthesis”, *Trends Genet*, 32(11), 2016, pp. 690-696.
- [12] Mukherjee K., Brocchieri L., & Burglin T. R., “A Comprehensive Classification and Evolutionary Analysis of Plant Homeobox Genes”, *Mol Biol Evol*, 26(12), 2009, pp. 2775-2794.
- [13] Tian T., Liu Y., Yan H., You Q., Yi X., Du Z., Xu W., & Su Z., “agriGO v2.0: a GO Analysis Toolkit for The Agricultural Community, 2017 Update”, *Nucleic Acids Res*, 45(W1), 2017, pp. W122-W129.
- [14] Vollbrecht E., Veit B., Sinha N., & Hake S., “The Developmental Gene *Knotted-1* is a Member of a Maize Homeobox Gene Family”, *Nature*, 350(6315), 1991, pp. 241-243.
- [15] Xiong H., Shi A., Wu D., Weng Y., Qin J., Shu X., & Zhou W., “Genome-Wide Identification, Classification and Evolutionary Expansion of *KNOX* Gene Family in Rice (*Oryza sativa*) and Populus (*Populus trichocarpa*)”, *Am J Plant Sci*, 9(06), 2018, pp. 1071-1092.