

CHARACTERISTICS AND MOLECULAR EVOLUTION OF THE AP2 GENE FAMILY IN SOYBEAN (*Glycine max* (L.) Merr.)Nguyen Thu Giang^{1,3}, Nguyen Thi Hai Yen², Nguyen Huu Quan³, Chu Hoang Mau^{3*}¹TNU - University of Medicine and Pharmacy²TNU - University of Sciences, ³TNU - University of Education

ARTICLE INFO	ABSTRACT
Received: 08/4/2024 Revised: 20/5/2024 Published: 20/5/2024	Soybean [<i>Glycine max</i> (L.) Merrill] is a crop with nutritional and economic value and a soil improvement crop. Soybean belongs to a group of crops with poor resistance to adverse abiotic factors. Therefore, research to improve the ability of soybean plants to withstand abiotic stresses is essential in the context of climate change. This study aims to analyze the molecular evolution of the AP2 gene family, the DREB gene subfamily and the AP2 domain to select candidate genes for improving soybean resistance. Using BioEdit, BLAST and MEGA11 software to search, compare data and analyze phylogeny, the results show that the phylogenetic tree reflects the diversity and evolution of genes in the AP2 family and the DREB subfamily and AP2 domain in soybean. At the same time, 18 DREB genes were identified located on 17 soybean chromosomes. Among the DREB genes, the functions of some genes have yet to be thoroughly studied and need to be verified experimentally.

KEYWORDSAP2 gene family
AP2 domain
DREB gene subfamily
DNA binding sites
Soybean**ĐẶC ĐIỂM VÀ SỰ TIẾN HOÁ PHÂN TỬ CỦA HỌ GENE AP2 Ở ĐẬU TƯƠNG (*Glycine max* (L.) Merr.)**Nguyễn Thu Giang^{1,3}, Nguyễn Thị Hải Yến², Nguyễn Hữu Quân³, Chu Hoàng Mậu^{3*}¹Trường Đại học Y Dược - ĐH Thái Nguyên²Trường Đại học Khoa học - ĐH Thái Nguyên, ³Trường Đại học Sư phạm - ĐH Thái Nguyên

THÔNG TIN BÀI BÁO	TÓM TẮT
Ngày nhận bài: 08/4/2024 Ngày hoàn thiện: 20/5/2024 Ngày đăng: 20/5/2024	Đậu tương [<i>Glycine max</i> (L.) Merrill], loại cây trồng có giá trị dinh dưỡng, kinh tế và là cây cải tạo đất. Đậu tương thuộc nhóm cây trồng chống chịu kém các yếu tố bất lợi phi sinh học. Vì vậy nghiên cứu nâng cao khả năng chống chịu các stress phi sinh học của cây đậu tương là rất cần thiết trong bối cảnh biến đổi khí hậu hiện nay. Mục tiêu của nghiên cứu này nhằm phân tích tiến hoá phân tử của họ gene AP2, phân họ gene DREB và miền AP2 làm cơ sở lựa chọn gene ứng cử viên trong việc cải thiện khả năng chống chịu của cây đậu tương. Sử dụng phần mềm BioEdit, BLAST và MEGA11 trong việc tìm kiếm, so sánh dữ liệu và phân tích sự phát sinh chủng loại, kết quả cho thấy cây phát sinh chủng loại phản ánh sự đa dạng và tiến hoá của các gene trong họ AP2, phân họ DREB và miền AP2 ở đậu tương. Đồng thời, nghiên cứu xác định được 18 gene DREB có vị trí trên 17 nhiễm sắc thể của đậu tương. Trong số các gene DREB, chức năng của một số gene chưa được nghiên cứu đầy đủ cần được kiểm chứng bằng thực nghiệm.

TỪ KHÓAHọ gene AP2
Miền AP2
Phân họ gene DREB
Điểm liên kết DNA
Đậu tươngDOI: <https://doi.org/10.34238/tnu-jst.10058>

* Corresponding author. Email: chuhoangmau@tnu.edu.vn

1. Giới thiệu

Đậu tương [*Glycine max* (L.) Merrill] là loại cây trồng chiếm vị trí quan trọng trong cơ cấu cây nông nghiệp. Hạt đậu tương là nguồn cung cấp protein và dầu thực vật cho con người cũng như cho ngành công nghiệp thực phẩm chăn nuôi. Canh tác đậu tương không những mang lại giá trị dinh dưỡng mà còn có giá trị kinh tế và cải tạo đất. Biến đổi khí hậu tác động đến năng suất và chất lượng sản phẩm từ các loại cây trồng, trong đó có đậu tương. Cây đậu tương và các loại cây trồng khác luôn đối mặt các stress phi sinh học, trong đó có hạn và mặn. Tuy nhiên, đậu tương thuộc nhóm cây trồng chống chịu kém các yếu tố bất lợi phi sinh học [1], do vậy câu hỏi đặt ra là ứng dụng công nghệ nào để cải thiện khả năng chống chịu các nhân tố bất lợi phi sinh học của cây đậu tương.

Hạn hán, độ mặn cao, nhiệt độ thấp và sự tấn công của mầm bệnh là những stress phổ biến nhất ảnh hưởng đến sinh trưởng và phát triển của cây đậu tương. Hiện nay, ứng dụng kỹ thuật biểu hiện mạnh các gene mã hoá protein nhân tố phiên mã có khả năng đáp ứng các stress hạn, mặn và nhân tố bất lợi phi sinh học khác là cách tiếp cận đầy hứa hẹn trong chiến lược phát triển các giống đậu tương chống chịu, ứng phó với biến đổi khí hậu [2]. Tuy nhiên, tìm kiếm gene ứng cử viên trong kỹ thuật biến nạp tạo dòng đậu tương chống chịu giữ vai trò hết sức quan trọng.

Các nhân tố phiên mã (Transcription Factor-TF) đóng vai trò quan trọng trong mạng lưới điều hòa biểu hiện gen để đối phó với các stress từ ngoại cảnh. TF kiểm soát các gene ở phía xuôi dòng bằng cách liên kết với yếu tố hoạt động *cis* trong vùng khởi động của gen mục tiêu. Các protein liên kết yếu tố đáp ứng APETALA2/ethylene (AP2/ERF) là một trong những họ TF lớn nhất và là nhân tố chính điều chỉnh phản ứng của thực vật trước các stress khác nhau. Họ AP2/ERF có thể được chia thành 5 phân họ, AP2, DREB, ERF, RAV và Soloist tùy theo loại và số lượng miền AP2 và B3 được bảo tồn [3]. Trong phân họ DREB (dehydration-responsive element binding), miền AP2 trong phân tử protein có khoảng 58-60 amino acid và là thành viên lớn nhất trong họ nhân tố phiên mã AP2/ERF. Protein DREB có thể điều chỉnh sự biểu hiện của các gen hạ lưu bằng cách liên kết với các yếu tố DRE/CRT tác động *cis* nằm trong vùng khởi động của các gen mục tiêu phản ứng với stress phi sinh học ở phía xuôi dòng [4], [5]. Đến nay, đã có một số nghiên cứu đề cập đến đặc điểm và phân tích sự biểu hiện của một số gene *DREB* trong mục đích xác định vai trò và tìm kiếm gene ứng cử viên cho cải thiện khả năng chống chịu của cây đậu tương [6]-[8]. Như vậy, có thể nói lựa chọn gene mã hoá nhân tố phiên mã thuộc họ AP2, trong đó có các gene *DREB* làm gene ứng cử viên trong mục đích cải thiện khả năng chống chịu các stress phi sinh học của cây đậu tương là hiệu quả hơn khi chọn một gene có tác động đơn lẻ.

Cơ sở của việc tìm kiếm gene ứng cử viên là sự tập hợp các trình tự gene từ các cơ sở dữ liệu, phân tích tiến hoá phân tử, xác định vị trí và mối quan hệ của gene lựa chọn trong cây phát sinh chủng loại. Nghiên cứu này trình bày kết quả phân tích đặc điểm và sự tiến hoá phân tử của họ gene AP2 và phân họ gene *DREB* ở cây đậu tương [*Glycine max* (L.) Merrill] trên cơ sở dữ liệu thu được từ các cơ sở dữ liệu sinh học.

2. Phương pháp nghiên cứu

2.1. Dữ liệu

Dữ liệu trình tự các gene thuộc họ AP2, các gene thuộc phân họ *DREB* và vùng AP2 của protein DREB được khai thác từ cơ sở dữ liệu National Center for Biotechnology Information (NCBI) [9], [10] và dữ liệu trên GenBank [11].

2.2. Phương pháp nghiên cứu

Dữ liệu về phân họ gene AP2, phân họ gene *DREB* và miền AP2 của protein DREB của đậu tương được sử dụng để phân tích đặc điểm và phát sinh chủng loại bằng phần mềm MEGA11 [12]. Lịch sử tiến hóa được suy luận bằng cách sử dụng phương pháp Maximum Likelihood (ML) và mô hình Tamura-Nei [13]. Cây đồng thuận bootstrap được suy ra từ 1000 lần lặp lại được thể hiện lịch sử tiến hóa của các đơn vị phân loại được phân tích [14]. So sánh miền AP2 bằng phần mềm BioEdit.

3. Kết quả và thảo luận

3.1. Sự phát sinh chủng loại của họ gen AP2 ở đậu tương

Từ cơ sở dữ liệu gene trên NCBI, 100 trình tự gene thuộc họ AP2 của đậu tương đã được khai thác và phân tích bằng BLAST. Kết quả phân tích BLAST thu được 100 trình tự gene thuộc họ gene AP2, vị trí của mỗi gene trên nhiễm sắc thể, các trình tự protein suy diễn có mã số trên GenBank. Các trình tự gen AP2 chia làm 16 loại, gồm AP2, AP2 like, AP2 -1, AP2-2, AP2-3, AP2-4, AP2-5, AP2-6, AP2-7, RAP2, AP2 -7 like, RAP2 -7, RAP2-7like, AP2-EREBP, SGR-AP2, Lf1 AP2) được phân bố trên 17 nhiễm sắc thể (1, 2, 3, 4, 5, 6, 8, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20). AP2 like có số lượng copy nhiều nhất (6 copy), AP2 có 4 bản copy, RAP2 -7 và AP2-ERF đều có 3 bản copy. Các gen còn lại trong họ gen AP2 chỉ gồm 1 đến 2 bản copy.

Kết quả phân tích phát sinh gene dựa trên 100 trình tự AP2 của đậu tương theo phương pháp ML và JTT maxtrix-based model trong MEGA11 cho thấy họ AP2 được phân thành 9 nhánh ký hiệu I, II, III, IV, V, VI, VII, VIII và IX (Hình 1). Nhánh VI và nhánh IX có số lượng trình tự gene AP2 lớn nhất, nhánh IX gồm 28 trình tự (AP2-7like, RAP2-7, AP2-7, AP2-8, AP2 like), nhánh VI với 22 trình tự (RAP2-7like, AP2-9, AP2-ERF. Các nhóm còn lại có từ 6 đến 11 trình tự. RAP2-7 XM006594938 là một nhánh phụ riêng của nhóm III, nhóm IV. Trình tự RAP2-7 phân bố ở cả nhánh II và IX.

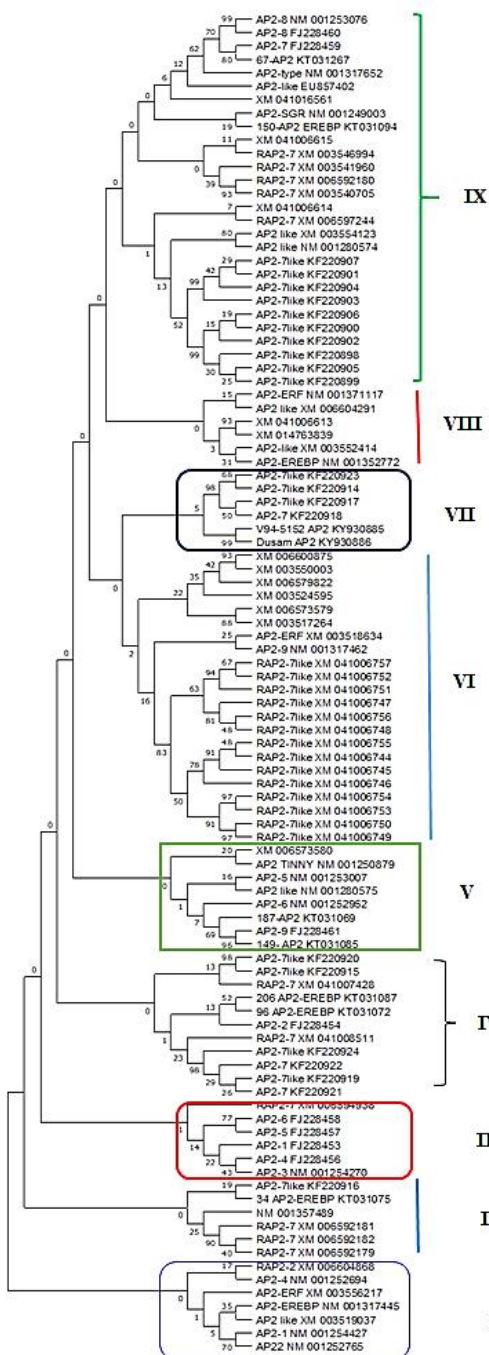
3.2. Sự phát sinh chủng loại của phân họ gen DREB ở đậu tương

Protein DREB, một nhóm nhân tố phiên mã có khả năng kích hoạt sự biểu hiện của các gene ở phía xuôi dòng phản ứng với stress phi sinh học. DREB liên kết với trình tự *cis* ở vùng khởi động ở những vị trí DNA binding của miền AP2. Sự liên kết này đã thúc đẩy hoạt động của enzyme phiên mã RNA polymerase làm tăng cường phiên mã của gene mục tiêu và tăng cường khả năng chống chịu các stress phi sinh học của ngoại cảnh ở nhiều loài thực vật. Các nhân tố phiên mã DREB1 và DREB2 đã được xác định trong cây *Arabidopsis* và chúng đã kích hoạt tăng biểu hiện phiên mã của các gene mục tiêu và nâng cao khả năng chống chịu các stress lạnh, hạn, mặn [15]-[17]. Đối với cây đậu tương, đã có một số gene *DREB* đã được phân tích và đề xuất là gene ứng cử viên cho việc cải thiện khả năng chịu hạn và mặn ở cây đậu tương [18], [19]. Tuy nhiên, nhiều gene *DREB* của đậu tương đến nay vẫn chưa rõ chức năng, cần tiếp tục xác định bằng khai thác các cơ sở dữ liệu và chứng minh bằng thực nghiệm.

Khai thác dữ liệu từ GenBank thu được 78 trình tự gen *DREB* của đậu tương chứa các thông tin về tên gene, gene ID, vị trí trên nhiễm sắc thể (NST), mã số gene trên GenBank và mã số protein suy diễn của mỗi gene. Phân tích 78 trình tự gene *DREB* đã xác định được 18 loại gene *DREB*, mỗi gene *DREB* có số bản copy khác nhau, dao động từ 1-8 copy (Bảng 1).

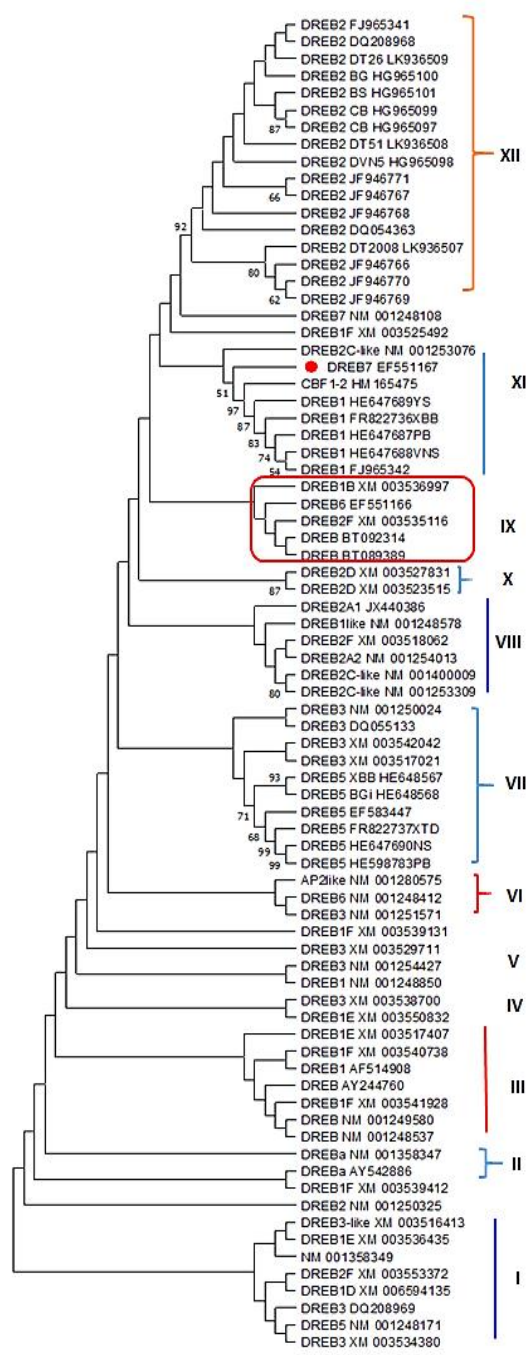
Bảng 1. Danh sách các gene *DREB*, số copy và vị trí trên nhiễm sắc thể của đậu tương

STT	Tên gen	Số bản copy	Vị trí trên NST số
1	<i>DREB</i>	4	8, 12, 13, 14
2	<i>DREB1</i>	1	9
3	<i>DREB1like</i>	1	9
4	<i>DREBa</i>	1	12
5	<i>DREB1B</i>	1	10
6	<i>DREB1D</i>	1	13
7	<i>DREB1E</i>	3	1, 10, 17
8	<i>DREB1F</i>	4	5, 11, 12, 13
9	<i>DREB2</i>	3	4, 6, 8
10	<i>DREB2A2</i>	1	14
11	<i>DREB2Clike</i>	2	2, 6
12	<i>DREB2D</i>	2	4, 6
13	<i>DREB2F</i>	2	10, 19
14	<i>DREB3</i>	8	1, 3, 4, 7, 9, 11, 13, 17
15	<i>DREB3 like</i>	2	1, 6
16	<i>DREB5</i>	2	12, 13
17	<i>DREB6</i>	1	5
18	<i>DREB7</i>	1	20



Hình 1. Cây phát sinh và quan hệ tiến hóa phân tử của các thành viên của họ gene AP2 ở đậu tương theo phương pháp ML và JTT maxtrix-based model trong MEGA11

Thông tin từ bảng 1 cho thấy hai gene *DREB* và *DREB1F* đều có 4 copy và phân bố ở 4 nhiễm sắc thể, gene *DREB3* có 8 copy và phân bố ở 8 nhiễm sắc thể, *DREB1E* và *DREB2* có 3 copy, còn các gene *DREB* khác phân bố trên 1 hoặc 2 nhiễm sắc thể. Hiện tượng này cho thấy các gene *DREB* có sự khác nhau về cấu trúc và các gene có nhiều copy hiện đang trong quá trình tiến hoá mạnh.



Hình 2. Cây phát sinh và quan hệ tiến hóa phân tử của các thành viên của phân họ gene *DREB* ở đậu tương theo phương pháp ML và JTT maxtrix-based model trong MEGA11

Kết quả phân tích sự phát sinh chủng loại phân họ gene *DREB* của đậu tương cho thấy 78 trình tự gen *DREB* được chia thành hai nhánh với 12 nhóm lớn ký hiệu từ I-XII (Hình 2). Sơ đồ phát sinh gene cho thấy 17 trình tự *DREB2* tạo nên nhóm lớn nhất (XII) và 6 trình tự tạo thành các nhóm riêng là *DREB7* (NM001248108), *DREB1F* (XM003525492), *DREB1F* (XM00353913), *DREB3* (XM 003529711), *DREBa* (MN001358347), *DREB2* (NM 001250325). Các nhóm gồm hai gen *DREB*, đó là nhóm II (*DREBa* và *DREB1F*), nhóm IV (*DREB3* và *DREB1E*), nhóm V (*DREB3* và *DREB1*), nhóm IX (*DREB2D*). Đáng chú ý, một số trình tự *DREB3* phân bố ở những nhánh riêng, hai gen *DREB3* (DQ055133 và NM_001250024) tách thành một nhánh trong nhóm VII, *DREB3* (XM 003529711) là một nhóm phụ của nhánh thứ hai. Hai nhóm (IV, V) đều có trình tự *DREB3*. Trình tự *DREB7* (EF551167) có vị trí ở nhóm XI, trong khi đó *DREB7* (NM 001248108) lại trong nhóm XII. Kết quả phân tích có thể thấy mức độ đột biến của các gene *DREB* là khác nhau.

Các nhân tố phiên mã DREB điều hoà sự biểu hiện của các gen phản ứng với stress và do đó đóng vai trò quan trọng trong phản ứng của thực vật với stress phi sinh học. Zhou và cộng sự (2020) đã xác định 73 gen thuộc họ DREB từ hệ gen đậu tương. Các gen DREB này được chia thành sáu nhóm nhỏ dựa trên phân tích phát sinh gen. Phân tích cấu trúc gen cho thấy hầu hết các gen DREB đều có một exon duy nhất không có intron [20]. Tuy nhiên, đến nay chúng tôi đã xác định được 78 trình tự gene DREB của đậu tương từ các cơ sở dữ liệu NCBI và GenBank.

3.3. Đặc điểm và sự phát sinh chủng loại của miền AP2 trong protein DREB của đậu tương

Protein nhân tố phiên mã DREB chứa miền bảo thủ AP2 chứa các điểm amino acid liên kết với trình tự *cis* trong promoter của gene mục tiêu. Khi nhân tố phiên mã DREB bám vào và liên kết với promoter, enzyme phiên mã RNA polymerase được kích hoạt và tăng tốc độ phiên mã của gene mục tiêu. Miền AP2 của 18 protein DREB suy diễn của đậu tương có sự khác nhau về số lượng và thành phần amino acid, nhưng phổ biến từ 58-60. Tuy nhiên, một số miền AP2 của một số DREB có số lượng amino acid ít hơn, như *DREB2_DQ208968* và *DREB2A2* có 46 amino acid hoặc chỉ có 43 amino acid (*DREB 2C like*) hay 50 amino acid (*DREB6_NM001248412*). Một số trình tự DREB có hơn 60 amino acid, như *DREB3* (XM003542042 và XM003538700) có 70 amino acid, các gene *DREB*, *DREBa*, *DREB2C*, *DREB2A2* (Hình 3).

Bảng 2 trình bày thành phần amino acid của những điểm liên kết trình tự *cis* của promoter (DNA binding site) cho thấy trong tất cả 18 loại protein DREB, miền AP2 đều có 11 điểm amino acid liên kết, nhưng khác nhau ở thành phần và biểu hiện ở 11 dạng. Dạng RGRRWKERRWT có ở 8 protein DREB (*DREB*, *DREB1*, *DREB2A2*, *DREB2D*, *DREB3*, *DREB3-like*, *DREB5*, *DREB6*), dạng RGRRSKERRWT (*DREB1*, *DREB-like*, *DREB7*), dạng KGRRNKERRWT (*DREB1B* và *DREB1D*), dạng KGRRWKERRWT (*DREB2D* và *DREB2F*), các dạng còn lại chỉ có ở một gene.

Bảng 2. Các dạng liên kết giữa AP2 của protein DREB với trình tự *cis* của promoter

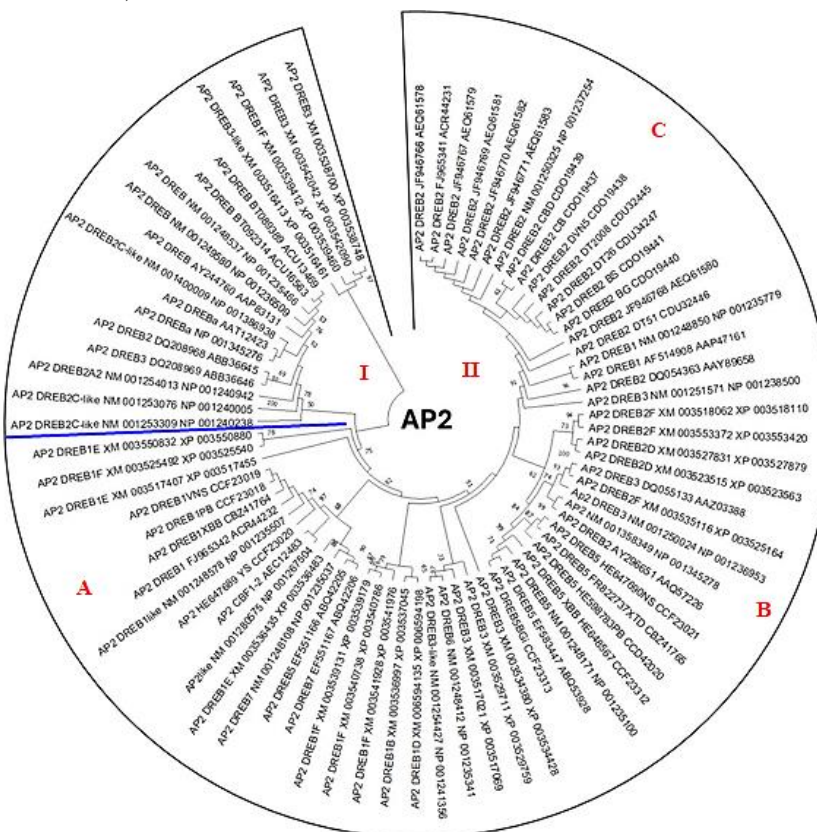
Dạng liên kết	Gene	Dạng liên kết	Gene
RGRRWKERRWT	<i>DREB</i> , <i>DREB1</i> , <i>DREB2A2</i> , <i>DREB2D</i> , <i>DREB3</i> , <i>DREB3-like</i> , <i>DREB5</i> , <i>DREB6</i>	KGRRNKERRWT	<i>DREB1D</i> , <i>DREB1B</i>
RGRRSKERRWT	<i>DREB1</i> , <i>DREB-like</i> , <i>DREB7</i>	RGRRWKERRWS	<i>DREB2</i> , <i>DREB2F</i>
RGRRNKERRWT		RGRRNKERRWT	<i>DREB E</i>
RGRRRKERRWT	<i>DREB1F</i>	RGRRKKERRWT	
RGRRGKERRWT		KGRRWKERRWS	<i>DREB3</i>
		HGRRWKERRWT	

	10	20	30	40	50	60
AP2_CBF1-2_AEC12483	YRQVRRRNTD	KWVSEVREPN	KKT-RIWLG	FPTPEMAARA	HDVAAMALRG	RYACLNFADS
AP2_DREB_AY244760 AAP8313	Q.TWG	A.I	RGN-L	AIG-L	Y.E.Q.MY	SC.R..PN-
AP2_DREB_BT089389 ACU1346	Q.TWG	A.I	RGN-L	AIG-L	Y.E.R.MY	SC.R..PN-
AP2_DREB_BT092314 ACU1656	Q.TWG	A.I	RGN-L	AIG-L	Y.E.R.MY	SC.R..PN-
AP2_DREB_NM_001248537 NP_0	Q.TWG	A.I	RGN-L	AIG-L	Y.E.R.MY	SC.R..PN-
AP2_DREB_NM_001249580 NP_0	Q.TWG	A.I	RGN-L	AIG-L	Y.E.R.MY	SC.R..PN-
AP2_DREB1_VNS_CCF23019	S	C				
AP2_DREB1_PB_CCF23018	S	C				
AP2_DREB1_XBB_CBZ41764	S	C				
AP2_DREB1_AF514908 AAP471	I.M.KWG	A.I	RS-S	YT.VA	Y.T.VFY	PT.R..PEL
AP2_DREB1_FJ965342 ACR4423	S	C				
AP2_DREB1_NM_001248850 NP	I.M.KWG	A.I	RS-S	YT.VA	Y.T.VFY	PT.R..PEL
AP2_DREB1B_XM_003536997 XP	K.Q.-G	L.Q	N.NA.V	H.D.I	Y.L.FK	DS.S..PN-
AP2_DREB1D_XM_006594135 XP	K.Q.-G	C.L.Q	N.NA.V	TH.D.I	Y.L.FK	DN.S..PHA
AP2_DREB1E_XM_003517407 XP	KN	C.M.V	NNS-	Y	L	KS
AP2_DREB1E_XM_003536435 XP	DSG	C	S	A	L	S
AP2_DREB1E_XM_003550832 XP	NN	C.V	D.ST	Y	LS	KS
AP2_DREB1F_XM_003539412 X	S	C				
AP2_DREB1F_XM_003525492 XP	NN	C.V	D.ST	Y.V	L	KS
AP2_DREB1F_XM_003539131 XP	Q.GN	C.I	S-V.V	Y	VL.S	TS.KF.P
AP2_DREB1F_XM_003540738 XP	Q.RN	C.I	S-V	Y	VL.S	TS.NF.P
AP2_DREB1F_XM_003541928 XP	Q.GN	C.I	S	Y.S	VL.K	TS.VF.P
AP2_DREB1like_NM_001248578	S	C				
AP2_DREB2_AY296651 AAQ572	Q.HWG	A.I.L.K	NR-L	D.A.E.L	Y.N.FK	EF.R..PHL
AP2_DREB2_JF946768 AEQ615	I.M.KWG	A.I	SS-	YA.VA	Y.T.VFH	PS.R..PEL
AP2_DREB2_BG_CDO19440	I.M.KWG	A.I	RS-	YA.VA	Y.T.VFH	PS.R..PEL
AP2_DREB2_BS_CDO19441	I.M.KWG	A.I	RS-	YA.VA	Y.T.VFH	PS.R..PEL
AP2_DREB2_CB_CDO19437	I.M.KWG	A.I	RS-	YA.VA	Y.T.VFH	PS.R..EL
AP2_DREB2_CBD_CDO19439	I.M.KWG	A.I	RS-	YA.VA	Y.T.VFH	PS.R..EL
AP2_DREB2_DQ054363 AAY8965	I.M.KWG	A.I	RS-	YA.VT	Y.T.VFH	PS.R..PEL
AP2_DREB2_DQ208968 ABB3664	K.K.KWG	I.L	SRQ-	YD.K	F.A.MFC	N.KF.P-
AP2_DREB2_DT26_CDU34247	I.M.KWG	A.I	RS-	YA.VA	Y.T.VFH	PS.R..PEL
AP2_DREB2_DT51_CDU32446	I.M.KWG	A.I	RS-	YA.VA	Y.T.VLH	PS.R..PEL
AP2_DREB2_DT2008_CDU32445	I.M.KWG	A.I	RS-	YA.VA	Y.T.VFH	PS.R..PEL
AP2_DREB2_DVN5_CDO19438	I.M.KWG	A.I	RS-	YA.VA	Y.T.VFH	PS.R..PEL
AP2_DREB2_FJ965341 ACR4423	I.M.KWG	A.I	RS-	YA.VA	Y.T.VFH	PS.R..PEL
AP2_DREB2_JF946766 AEQ6157	I.M.KWG	A.I	RS-	YA.VA	Y.T.VFH	PS.R..PEL
AP2_DREB2_JF946767 AEQ6157	I.M.KWG	A.I	RS-	YA.VA	Y.T.VFH	PS.R..PEL
AP2_DREB2_JF946769 AEQ6158	I.M.KWG	A.I	RS-	YA.VA	Y.T.VFH	PS.R..PEL
AP2_DREB2_JF946770 AEQ6158	I.M.KWG	A.I	RS-	YA.VA	Y.T.VFH	PS.R..PEL
AP2_DREB2_JF946771 AEQ6158	I.M.KWG	A.I	RS-	YA.VA	Y.T.VFH	PS.R..PEL
AP2_DREB2_NM_001250325 NP	I.M.KWG	A.I	RS-	YA.VA	Y.T.VFH	PS.R..PEL
AP2_DREB2A2_NM_001254013 N	Q.TWG	G.I	RGS-L	SSAQE.L	Y.E.R.MY	-P.R..P-
AP2_DREB2C-like_NM_0012530	Q.TWG	G.I	RGS-L	SSAQE.L	Y.E	
AP2_DREB2C-like_NM_0012533	Q.TWG	A.I	RGS-L	AIS.L	Y.E.R.MY	SC.R..PNV
AP2_DREB2C-like_NM_0014000	Q.TWG	A.I	RGS-L	AIS.L	Y.E.R.MY	SC.R..PNV
AP2_DREB2D_XM_003523515 XP	K.Q.TWG	A.I	RGA-L	E.SHE.L	Y.A.RK.Y	SD.K.LPEL
AP2_DREB2D_XM_003527831 XP	K.Q.TWG	A.I	RGA-L	E.SHE.L	Y.A.RK.Y	SD.K.LPEL
AP2_DREB2F_XM_003553372 X	Q.TWG	A.I.K	R-L	S.A.A.E.L	Y.E.RR.Y	PD.Y.LPHL
AP2_DREB2F_XM_003518062 XP	Q.TWG	A.I.K	R-L	S.A.A.E.L	Y.E.RR.Y	PD.Y.LPHL
AP2_DREB2F_XM_003535116 XP	Q.HWG	A.I.L.K	NR-L	D.A.E.L	Y.K.YK	DF.R..PNL
AP2_DREB3_DQ055133 AAZ033	Q.HWG	A.I.L.K	NR-L	D.A.E.L	Y.K.YK	DF.R..PNL
AP2_DREB3_DQ208969 ABB3664	K.K.KWG	I.L	SRQ-	YD.K	F.A.MFC	N.KF.P-
AP2_DREB3_NM_001250024 NP	Q.HWG	A.I.L.K	NR-L	D.A.E.L	Y.K.YK	DF.R..PNL
AP2_DREB3_NM_001251571 NP	K.K.KWG	I.L	SRQ-	YD.K	F.A.MFC	N.KF.P.N
AP2_DREB3_XM_003517021 XP	M.WG	I.R	S-	A	LTIK	SS.I..PEL
AP2_DREB3_XM_003529711 XP	H.K.WG	I.R	S-	S	LTIK	QS.I..PE-
AP2_DREB3_XM_003534380 XP	F.K.SWG	R.Y	I.L.G	Q.	S	GS
AP2_DREB3_XM_003538700 XP	M.KWG	I.L.K	S-	S	LTIK	TS.F..PEL
AP2_DREB3_XM_003542042 XP	RNRDPTKHSD	H.M.WG	I.R	S	A	LSIK
AP2_DREB3-like_NM_00125442	M.WG	I.R	S-	S	LSIK	SA.I..PH-
AP2_DREB3-like_XM_00351641	M.WG	I.W	S-	S	VW	N.LSIK
AP2_DREB5_BG1_CCF23313	Q.HWG	A.I.L.R	NR-L	D.A.D.M	Y.RE.FK	EN.R..PEL
AP2_DREB5_FR822737;XTD_CB	Q.HWG	A.I.L.R	NR-L	D.A.D.M	Y.RE.FKQ	EN.R..PEL
AP2_DREB5_EF551166 ABQ4220	DSG	C	S	S.A	I	S
AP2_DREB5_EF583447 ABQ5392	Q.HWG	A.I.L.R	NR-L	D.A.D.M	Y.RE.FK	EN.R..PEL
AP2_DREB5_HE598783;PB_CCD4	Q.HWG	A.I.L.R	NR-L	D.A.D.M	Y.RE.FKQ	EN.R..PEL
AP2_DREB5_HE647690;NS_CCF2	WQ.HWG	A.I.L.R	NR-L	D.A.D.M	Y.RE.FKQ	EN.R..PEL
AP2_DREB5_NM_001248171 NP	Q.HWG	A.I.L.R	NR-L	D.A.D.M	Y.RE.FK	EN.R..PEL
AP2_DREB5_XBB_HE648567_CCF	Q.HWG	A.I.L.R	NR-L	D.A.D.M	Y.RE.FK	EN.R..PEL
AP2_DREB6_NM_001248412 NP	M.QWG	I.R	S-	D	LTIK	
AP2_DREB7_EF551167 ABQ4220	DSG	C	S	S.A	I	S
AP2_DREB7_NM_001248108 NP	DSG	C	S	A	I	S
AP2_DREBa_AAT12423	Q.TWG	A.I	RGS-L	AIS.L	Y.E.MY	FC.R..PNV
AP2_DREBa NP_001345276	Q.TWG	A.I	RGS-L	AIS.L	Y.E.MY	FC.R..PNV
AP2_HE647689_Ys_CCF23020	S	C				
AP2_NM_001358349 NP_001345	Q.HWG	A.I.L.K	NR-L	D.A.E.L	Y.N.FK	EF.R..PHL
AP2like_NM_001280575 NP_00						

Hình 3. Đặc điểm miền AP2 của 78 protein DREB ở đậu tương

Phân tích tiên hoá phân tử dựa trên thành phần và trình tự amino acid miền AP2 của 78 protein DREB ở đậu tương bằng MEGA11, kết quả cho thấy sơ đồ cây phát sinh gồm 2 nhánh được thể hiện ở hình 4. Nhánh I chia thành 2 nhóm con, một nhóm gồm 4 trình tự, 13 trình tự ở

nhóm còn lại. Nhánh II chia thành 8 nhóm con, trong đó có 3 nhóm gồm nhiều trình tự nhóm A (12 trình tự), nhóm B (16 trình tự), nhóm C (20 trình tự). Ngoài ra còn có 1 nhóm gồm 2 trình tự AP2 của DREB1 (E, F); 2 nhóm chỉ gồm 1 trình tự là AP2 (DREB1E và DREB3_XM003534380).



Hình 4. Sơ đồ phát sinh miền AP2 của phân họ DREB ở đậu tương được thiết lập dựa trên 78 trình tự amino acid miền AP2 theo phương pháp Maximum likelihood trong MEGA11 và JTT maxtric-based model với bootstrap được lập lại 1000 lần

4. Kết luận

Trong nghiên cứu này, dựa trên cơ sở dữ liệu họ gene AP2 và phân họ gene DREB đã xác định được 18 gene DREB có vị trí trên 17 nhiễm sắc thể của đậu tương. Sự đa dạng của phân họ gene DREB thể hiện ở số bản copy của gene, trình tự nucleotide, trình tự amino acid của miền AP2 và các motif điểm liên kết amino acid với trình tự cis của promoter. Cây phát sinh chủng loài đã thể hiện sự tiến hoá phân tử của họ gene AP2, phân họ gene DREB và miền AP2 trong protein DREB. Trong các trình tự gene DREB có nhiều gene giả định, gene tham chiếu và chức năng của một số gene như DREB5, DREB7, RAP2-4... chưa nghiên cứu đầy đủ cần được kiểm chứng bằng thực nghiệm.

Lời cảm ơn

Nghiên cứu này được tài trợ bởi Bộ Giáo dục và Đào tạo thông qua đề tài cấp Bộ, mã số B2023-TNA-26. Các tác giả chân thành cảm ơn sự hỗ trợ này.

TÀI LIỆU THAM KHẢO/ REFERENCES

- [1] H. M. Chu, T. T. H. Nguyen, V. T. T. Nguyen, and H. H. Chu, *Gene and resistance characteristics of soybean plants*. Vietnam National University Pres, Hanoi, 2011.
- [2] C. Lata and M. Prasad, "Role of DREBs in regulation of abiotic stress responses in plants," *Journal of Experimental Botany*, vol. 62, pp. 4731-4748, 2011.

- [3] Y. Sakuma, Q. Liu, J. G. Dubouzet, H. Abe, K. Shinozaki, and K. Yamaguchi-Shinozaki, "DNA-binding specificity of the ERF/AP2 domain of *Arabidopsis* DREBs, transcription factors involved in dehydration-and cold-inducible gene expression," *Biochem. Biophys. Res. Commun.*, vol. 290, pp. 998-1009, 2002.
- [4] D. Kizis, V. Lumberras, and M. pages, "Role of AP2/EREBP transcription factors in gene regulation during abiotic stress," *FEBS Letters*, vol. 498, pp. 187-189, 2001.
- [5] M. Tang, M. Tang, J. Sun, Y. Liu, F. Chen, and S. Shen, "Isolation and functional characterization of the JcERF gene, a putative AP2/EREBP domain-containing transcription factor, in the woody oil plant *Jatropha curcas*," *Plant Mol. Bio.*, vol. 63, pp. 419-428, 2007.
- [6] T. H. Phang, G. Shao, and H. M. Lam, "Salt tolerance in soybean," *J Integr Plant Biol.*, vol. 50, pp. 1196-1212, 2008.
- [7] X. T. Dao, M. T. Ho, T. T. T. Vu, V. S. Le, and H. M. Chu, "Cloning and overexpression of GmDREB2 gene from a Vietnamesedrought-resistant Soybean variety," *Braz. Arch. Biol. Technol.*, vol. 58, pp. 651-657, 2015.
- [8] T. Q. Tu, P. Vacixaxa, T. T. M. Lo, N. H. Nguyen, N. T. T. Pham, Q. H. Nguyen, P. T. Do, L. T. N. Nguyen, Y. T. H. Nguyen, and M. H. Chu, "GmDREB6, a soybean transcription factor, notably affects the transcription of the *NtP5CS* and *NtCLC* genes in transgenic tobacco under saltstress conditions," *Saudi Journal of Biological Sciences*, vol. 28, 12, pp. 7175-7181, 2021, doi 10.1016/j.sjbs.2021.08.018.
- [9] NCBI, "AP2 in Glycine max," *Gene*. [Online]. Available: <https://www.ncbi.nlm.nih.gov/gene/?term=AP2+in+Glycine+max>. [Accessed Feb. 20, 2024].
- [10] NCBI, "Glycine max gene for dehydration-responsive element-binding protein," *Gene*. [Online]. Available: <https://www.ncbi.nlm.nih.gov/gene/?term=Glycine+max+gene+for+dehydration-responsive+element-binding+protein>. [Accessed Feb. 20, 2024]
- [11] National Center for Biotechnology Information, *GenBank*. [Online]. Available: <https://www.ncbi.nlm.nih.gov/nuccore>. [Accessed Feb. 20, 2024]
- [12] K. Tamura, G. Stecher, and S. Kumar, "MEGA11: Molecular Evolutionary Genetics Analysis Version 11," *Mol Biol Evol.*, vol. 38, pp. 3022-3027, 2021.
- [13] K. Tamura and M. Nei, "Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees," *Mol Biol Evol.*, vol. 10, pp. 512-526, 1993.
- [14] J. Felsenstein, "Confidence limits on phylogenies: an approach using the bootstrap," *Evolution*, vol. 39, pp. 783-791, 1985.
- [15] E. J. Stockinger, S. J. Gilmour, and M. F. Thomashow, "Arabidopsis thaliana CBF1 encodes an AP2 domain-containing transcriptional activator that binds to the C-repeat/DRE, a cis-acting DNA regulatory element that stimulates transcription in response to low temperature and water deficit," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 94, pp. 1035-1040, 1997.
- [16] Q. Liu, M. Kasuga, Y. Sakuma, H. Abe, S. Miura, K. Yamaguchi-Shinozaki, and K. Shinozaki, "Two transcription factors, DREB1 and DREB2, with an EREBP/AP2 DNA binding domain separate two cellular signal transduction pathways in droughtand low-temperature-responsive gene expression, respectively, in Arabidopsis," *Plant Cell*, vol. 10, pp. 1391-1406, 1998.
- [17] K. R. Jaglo-Ottosen, S. J. Gilmour, D. G. Zarka, O. Schabenberger, and M. F. Thomashow, "Arabidopsis CBF1 overexpression induces COR genes and enhances freezing tolerance," *Science*, vol. 280, pp. 104-106, 1998.
- [18] Q. H. Nguyen, L. T. K. Vu, L. T. N. Nguyen, N. T. T. Pham, Y. T. H. Nguyen, S. V. Le, and M. H. Chu, "Overexpression of the *GmDREB6* gene enhances proline accumulation and salt tolerance in genetically modified soybean plants," *Sci Rep.*, vol. 9, 2019, Art. no. 19663.
- [19] T. T. N. Pham, H. Q. Nguyen, T. N. L. Nguyen, X. T. Dao, D.T. Sy, V. S. Le, and H. M. Chu, "Overexpression of the *GmDREB2* gene increases proline accumulation and tolerance to drought stress in soybean plants," *AJCS*, vol. 14, pp. 495-503, 2020.
- [20] Y. Zhou, W. Zhou, H. Liu, P. Liu, and Z. Li, "Genome-wide analysis of the soybean DREB gene family: identification, genomic organization and expression profiles in response to drought stress," *Plant Breeding*, vol. 139, pp. 1158-1167, 2020.