APPLYING BIOINFORMATICS TO ANALYZE AUXIN-RELATED GENES IN ROBUSTA COFFEE GENOME (*Coffea canephora* L.)

Nguyen Dinh Sy[1]

Received Date: 14/06/2024; Revised Date: 26/06/2024; Accepted for Publication: 27/06/2024

ABSTRACT

Coffea canephora, which belongs to the Rubiaceae family, is one of the most popular cultivated coffea worldwide. In this study, we identified and analyzed candidate genes that involved in auxin-related gens in the *C. canephora* genome. The results showed that genome of *C. canephora* consists of 152 protein-coding genes related to auxin which are divided into 7 main groups depending on domain and motif: Auxin-induced protein; Auxin-binding protein; Auxin transporter-like protein; Auxin carrier component; Auxin response factor; Auxin-responsive protein; Auxin signaling protein. Using SMART software to analyze protein structure, the result indicated that there are some characteristic domains involved in auxin response such as EamA; AUX_IAA; Auxin inducible; Aldo_ket_red; Cupin; Aa_trans; B3, Auxin_resp; Mem_trans; B561; GH3; and LRR domain. The study on candidate protein-coding genes relating to auxin is important for elucidating protein functions involved in various cellular processes, growth, development and climate change adaptation of C. canephora.

Keywords: *Auxin, bioinformatic, C. canephora, domain protein, genome.*

1. INTRODUCTION

Although the Coffea genus includes more than 124 species, *C. canephora* (2n=2x=22) and *C. arabica* (2n=4x=44) are the most coffee bean productions with 40% and 60% in total production worldwide, respectively. FAS (2024) estimated that global coffee production in the 2024/2025 crop year will increase by 4.2% compared to the previous crop to 176 million bags, of which arabica production will increase by 4.4% to 99.86 million bags and robusta will increase by 3.9% to 76.38 million bags. Vietnam is the world's largest robusta coffee producer with export turnover in 2023 reaching 4.2 billion US Dollars.

Recently, several articles were published about genome sequencing (Pallavicini et al., 2005; Vieira et al., 2006; Denoeud et al., 2014 abiotic response genes (Nguyen Dinh et al., 2016; Dinh and Kang, 2017), genes for tolerance to disease (Barbosa et al., 2010; Albuquerque et al., 2015; Vadivelu, 2013) or caffeine biosynthetic pathway (Perrois *et al.,* 2015).

C. canephora genome sequence was published on Coffee Genome Hub (http://coffee-genome.org). Data available are the complete genome sequence of *C. canephora* along with gene structure, gene product information, metabolism, gene families, transcriptomics (ESTs, RNA-Seq), genetic markers and genetic maps. The hub provides also tools for easy querying, visualizing and downloading research data (Denoeud et al., 2014).

Diseases, pests and abiotic stresses are detrimental not only reducing yield and coffee quality, but also harmful for the economic and livelihood of coffee farmers who depend on it. Some research focused on several genes for tolerance and resistance. CaWRKY1 gene in *C. arabica* is a positive control against Rust fungus Hemileia vastatrix. α-amylase inhibitor-1 gene (α-AI1) was able to protect from coffee berry borer insect-pest by *Hypotheneumus hampei* for coffee plants (Barbosa et al., 2010; Albuquerque et al., 2015). CaNPR1 gene plays an important role in resistance against coffee leaf rust caused by *H. growatrix* in *C. arabica* and other plants (Vadivelu, 2013). Metallothionine gene expression studies, including CaMT4, CaMT15, CaMT3 and CaMT8 was elucidated the role of metallothionine in maintaining Cu and Zn homeostasis and in detoxifying these excess nutrients (Bulgarelli et al., 2016). The full-length *C. arabica* Protein Domain (CaBDP) gene sequence was extracted from the RNA of drought-tolerant *C. arabica* leaves. Genes have been cloned in *Arabidopsis* to characterize plant drought and salt tolerance (Nguyen Dinh et al., 2016; Dinh and Kang, 2017). Nguyen Dinh Sy et al., 2022 overview *C. canephora* L. genome and its function in stress response and caffeine biosynthesis, and analyzed candidate genes for dehydration stress response in *C. canephora* L.

[1]*Faculty of Natural Science and Technology, Tay Nguyen University;*
Corresponding author: Nguyen Dinh Sy; Tel: 0961367958; Email: ndsy@ttn.edu.vn.

Identification of the genes associated with the caffeine biosynthetic pathway in coffee provided the importance tool for regulating the caffeine biosynthesis to effectively help possibly produce more caffeine content and caffeine-free coffee for consumers in the future. Perrois *et al.,* 2015 demonstrated that the differential regulation of caffeine metabolism depends on the transcriptional activity that controls the differential expression of *XMT1* and *DXMT* genes in *C. arabica* and *C. canephora*. Recently, Raharimalala *et al.*, 2021 showed that *Coffea humblotiana*, a wild species from Comoro archipelago, which is lacks of caffeine synthase coding gene involved in the naturally decaffeinated status. Up to now, the evolution of NMT genes in *C. canephora* are NMT2; DXMT; XMT; MXMT; NMT3; MTL, which represent the methylation steps of the caffeine biosynthesis.

There are many genes, activators, and promoter genes in coffee plants that are continuously being discovered to elucidate their function, especially in a growth stimulation. Therefore, this research aims to screen and analyze auxin- related gene of coffee genome that select candidate genes for transgenic coffee plants to stimulate growth of coffee tree.

2. MATERIALS AND METHODS

2.1. Materials

- DNA sequencing genome of *C. canephora* that downloaded from website Coffee Genome Hub (coffee-genome.org/coffeacanephora).

2.2. Methods

- DNA sequencing genome of *C. canephora* was downloaded from website Coffee Genome

- To screening genes related to auxin response, the key words "auxin" was used.

- Using software SMART (http://smart.embl.de) to analyze protein structure (domain; motif).

3. RESULTS AND DISCUSSION

3.1. Identification and classification of auxin related genes

From a total of 25,574 genes in *C. canephora* genome (Denoeud et al., 2014) screening on protein-coding gene for auxin showed that there are total of 152 genes that anchored regularly in 11 chromosomes (chro.). The length of auxin-related genes is from 165 nucleotides (Cc11_g04780) to 3342 nucleotides (Cc00_g00210). Especially, Auxin transport protein BIG (Cc07_g04520) contains 15.333 nucleotides.

Among 152 protein-coding genes related to auxin, they are divided into 7 main groups depending on domain such as Auxin-induced protein; Auxin-binding protein; Auxin transporter-like protein; Auxin carrier component; Auxin response factor; Auxin-responsive protein; auxin signaling protein (table 1-3).

Table 1. The list of protein-coding-genes related to Auxin-induced protein

Gene	No. a.a	Gene	No. a.a	Gene	No. a.a	Gene	No. a.a
Cc01_g03080	858	Cc02_g16790	106	Cc07_g17560	372	Cc10_g01880	149
Cc01_g04600	331	Cc03_g04660	195	Cc07_g18590	361	Cc10_g02240	361
Cc01_g04610	92	Cc03_g05360	366	Cc07_g18600	378	Cc10_g10920	383
Cc01_g16330	251	Cc03_g08720	66	Cc07_g18610	405	Cc10_g12900	345
Cc01_g18860	97	Cc03_g14930	357	Cc07_g18620	377	Cc10_g14810	371
Cc02_g08320	380	Cc04_g03610	191	Cc07_g19260	190	Cc10_g14820	367
Cc02_g22010	390	Cc04_g06760	210	Cc07_g19990	377	Cc11_g04790	146
Cc02_g22030	412	Cc06_g03030	340	Cc08_g05620	394	Cc11_g04800	104
Cc02_g24230	179	Cc06_g12640	166	Cc08_g08390	156	Cc11_g10080	112
Cc02_g35440	368	Cc06_g14090	380	Cc08_g12360	321	Cc11_g10090	362
Cc02_g16700	99	Cc06_g14100	268	Cc08_g12980	184	Cc11_g14380	369
Cc02_g16710	107	Cc06_g20270	407	Cc08_g17100	112	Cc11_g14390	360
Cc02_g16730	106	Cc07_g02470	341	Cc09_g00760	405	Cc11_g14400	317
Cc02_g16740	103	Cc07_g17000	371	Cc09_g02950	392	Cc00_g13700	315
Cc02_g16750	95	Cc07_g17010	369	Cc09_g08800	376	Cc00_g15180	208
Cc02_g16760	97	Cc07_g17030	315	Cc10_g01860	106		

Table 2. The list of protein-coding-genes that related to auxin responsive protein/ Auxin response factor

Gene	No. a.a	Gene	No. a.a	Gene	No. a.a	Gene	No. a.a
Auxin responsive protein							
Cc01_g10550	144	Cc03_g06860	404	Cc06_g04040	277	Cc09_g00710	407
Cc01_g16320	401	Cc03_g09650	399	Cc06_g06020	129	Cc09_g07120	387
Cc01_g17790	326	Cc03_g13450	72	Cc06_g08150	216	Cc09_g10510	104
Cc02_g30730	246	Cc04_g00010	161	Cc06_g10040	338	Cc10_g08190	367
Cc02_g33360	483	Cc04_g02510	362	Cc06_g12650	122	Cc11_g04780	55
Cc02_g39040	189	Cc04_g02890	335	Cc06_g13230	180	Cc11_g09650	196
Cc02_g40000	183	Cc04_g03620	203	Cc07_g07780	375	Cc00_g04150	405
Cc03_g04670	240	Cc05_g14040	410	Cc07_g19210	103	Cc00_g26580	107
Cc03_g06400	266	Cc05_g16250	142	Cc08_g00560	173	Cc00_g29740	101
Auxin response factor							
Cc01_g11020	699	Cc02_g23580	222	Cc03_g13510	183	Cc07_g12410	846
Cc01_g11410	832	Cc02_g39520	697	Cc03_g13520	86	Cc08_g16330	694
Cc02_g11300	669	Cc03_g11270	221	Cc05_g00510	895	Cc09_g08740	907
Cc02_g14070	683	Cc03_g12730	216	Cc06_g03950	707	Cc10_g01900	950
Cc02_g23570	434	Cc03_g13500	94	Cc06_g12540	1079	Cc00_g00210	1114
						Cc00_g12260	863

Table 3. The list of protein-coding-genes that related to auxin-binding protein/ Auxin carrier protein/ Auxin transporter-like protein

Gene	No. a.a	Gene	No. a.a	Gene	No. a.a	Gene	No. a.a
Auxin-binding protein							
Cc01_g03720	140	Cc01_g05150	201	Cc06_g12080	208	Cc00_g05320	171
Auxin carrier protein							
Cc03_g13040	148	Cc07_g03020	360	Cc07_g12290	416	Cc10_g12950	345
Cc04_g06290	603	Cc07_g08300	458	Cc07_g12300	414	Cc10_g14830	363
Cc06_g00150	423	Cc07_g12020	423	Cc09_g03470	358	Cc11_g08680	666
Cc06_g12940	619	Cc07_g12270	411	Cc09_g03480	359	Cc11_g08940	451
Cc06_g19880	600	Cc07_g12280	412	Cc10_g00190	173		
Auxin transporter-like protein							
Cc02_g06770	475	Cc05_g00830	477	Cc07_g04520	5111		
Cc02_g16390	502	Cc06_g01510	486	Cc10_g11120	398		
Auxin signaling							
Cc02_g13650	126	Cc04_g00930	144	Cc07_g01170	464		

3.2. Motif and domain structure

SMART software (http://smart.embl.de) was applied to analyze domain and motif protein structure of 7 groups: Auxin-induced protein; Auxin-binding protein; Auxin transporter-like protein; Auxin carrier protein; Auxin response factor; Auxin responsive protein; Auxin signaling protein.

Group 1: Auxin-induced protein

Auxin-induced protein group contains EamA domain; or AUX_IAA domain; or Auxin inducible domain (Figure 1).

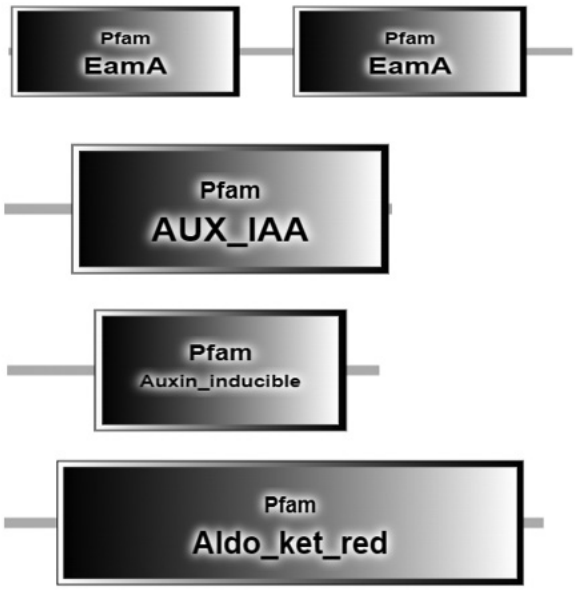

Figure 1. EamA, AUX_IAA, Auxin inducible, and Aldo_ket_red domain structure of Auxin-induced protein

The EamA domain, named after the O-acetyl-serine/cysteine export gene in *E. coli,* can be found in various proteins. One example is the PecM protein in *Erwinia chrysanthemi* regulates pectinase, cellulase, and blue pigment. Another example is the PagO protein in *Salmonella typhimurium*, although its function is unknown. Additionally, some members of the solute carrier family group 35 (SLC35) nucleoside-sugar transporters also possess this domain. Many proteins in this family are classified as drug/metabolite transporters, yet their function remains unidentified. These proteins are anticipated to be integral membrane proteins, and it's worth noting that many of them contain two copies of the EamA domain (Jack et al., 2001).

Transcription of the AUX/IAA genes occurs quickly in response to the plant hormone auxin (Abel et al., 1995). Certain members of this gene family are longer and possess a DNA binding domain at the beginning (like O64965). The inclusion here signifies the C-terminal portion of the AUX/IAA proteins. However, the specific function of this region remains uncertain.

This entry represents a group of plant proteins that respond to auxin, known as small auxin-up RNA (SAUR) (Gil and Green, 1997). The first SAUR gene was initially discovered in soybean hypocotyls (McClure and Guilfoyle, 1987). SAUR genes are primarily active in growing hypocotyls or other tissues involved in elongation. This suggests that they play a role in regulating cell elongation SAUR proteins might serve as a connection between auxin and plasma membrane H+-ATPases (PM H+-ATPases) in Arabidopsis thaliana (Spartz et al., 2014).

The aldo-keto reductase family consists of various related enzymes that are monomeric and depend on NADPH to carry out oxidoreduction reactions. Some examples of these enzymes are aldehyde reductase, aldose reductase, prostaglandin F synthase, xylose reductase, rho crystallin, and others (Bohren et al, 1989). They all share a similar structure characterized by a beta-alpha-beta fold, which is typical of proteins that bind nucleotides (Schade et al., 1990). This fold comprises a barrel shape with parallel beta-strands and alpha helices, containing a unique motif that binds NADP. The binding site is situated in a large, deep, elliptical pocket at the C-terminal end of the beta- sheet, where the substrate binds in an extended form. The pocket's hydrophobic nature means it favors aromatic and non-polar substrates, rather than highly polar ones (Wilson et al., 1992). When the NADPH coenzyme binds, it induces a significant conformational change that repositions a loop, effectively securing the coenzyme in place. This binding is more akin to FAD-binding oxidoreductases than NAD(P)-binding ones (Borhani et al., 1992). In some proteins within this category, there is an additional domain called the K+ ion channel beta chain regulatory domain, which has been shown to possess oxidoreductase activity (Gulbis et al., 2000). This entry represents the domain found in these proteins responsible for NADP-dependent oxidoreductase activity.

Group 2: Auxin-binding protein

This family represents the conserved barrel domain of the 'cupin' superfamily (Figure 2). This family contains 11S and 7S plant seed storage proteins, and germins. Plant seed storage proteins provide the major nitrogen source for the developing plant (Dunwell, 1998).

Figure 2. Cupin domain structure of Auxin-binding protein

Red: signal peptide.

Group 3: Auxin transporter-like protein

This transmembrane domain is present in various amino acid transporters, such as P34579 (UNC-47) and P40501 (MTR). UNC-47 encodes a vesicular amino butyric acid (GABA) transporter (VGAT), and it is predicted to consist of 10 transmembrane domains (UNC47_CAEEL) (McIntire et al., 1997). MTR is a protein associated with the N system amino acid transporter system, which is involved in methyltryptophan resistance (MTR_NEUCR). Other proteins possess this domain, including proline transporters and amino acid transporters with unidentified specificities.

Figure 3. Aa_trans domain structure of Auxin transporter-like protein

Pink: low complexity region.

Group 4: Auxin carrier protein

This entry represents a family of membrane transport proteins that have not been fully characterized yet. These proteins are found in eukaryotes, bacteria, and archaea. The most well-studied members of this family are the PIN components of auxin efflux systems in plants. These carriers are specific to auxin, meaning they only transport auxin molecules, and they are found at the basal ends of cells that can transport auxin (Blakeslee et al., 2005; Kramer, 2004).

Figure 4. Mem_trans domain structure of Auxin carrier protein

Plants usually have multiple proteins from this family, each with a unique pattern of expression in specific tissues. They are present in various plant tissues, including vascular tissues and roots. These proteins play a role in several processes, such as establishing embryonic polarity, promoting plant growth, forming apical hooks in seedlings, and influencing responses to light and gravity. On average, these plant proteins are made up of 600-700 amino acids and contain 8-12 segments that cross the cell membrane.

Group 5: Auxin response factor

B3 DNA Binding Domain:

RAV1 and RAV2, two DNA- binding proteins found in Arabidopsis thaliana, possess unique amino acid sequence domains exclusive to higher plant species. The N-terminal regions of RAV1 and RAV2 share similarities with the AP2 DNA-binding domain, which belongs to a family of transcription factors. On the other hand, the C-terminal region of RAV1 and RAV2 shows similarities with the highly conserved C-terminal domain, known as B3, of VP1/ABI3 transcription factors (Kagaya et al., 1999).

In the case of RAV1, its AP2 and B3-like domains independently bind to the CAACA and CACCTG motifs, respectively. When these two domains work together, they achieve a strong affinity and specificity for binding. Interestingly, there is a suggestion that a highly flexible structure connects the AP2 and B3-like domains of RAV1. This allows the two domains to bind to the CAACA and CACCTG motifs in various spacings and orientations (Kagaya et al., 1999).

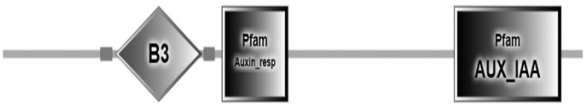

Figure 5. B3, Auxin_resp, and AUX_IAA domain structure of Auxin response protein

Pink: low complexity region.

Auxin, a plant hormone also known as indole-3-acetic acid can control the gene expression of various families, such as Aux/IAA, GH3, and SAUR. Among these families, there are two closely related groups of proteins, namely Aux/IAA proteins (IPR003311) and auxin response factors (ARF), which play a crucial role in regulating the gene expression influenced by auxin (Liscum and Reed, 2002). Multiple ARF proteins exist, with some activating and others repressing transcription. ARF proteins bind to specific promoter elements named auxin-responsive cis-acting promoter elements (AuxREs) using a DNA-binding domain located at their N-terminal. It is believed that Aux/IAA proteins activate transcription by modifying the

activity of ARF proteins through the protein-protein interaction domains present at the C-terminal region of both Aux/IAA and ARF proteins.

Group 6: Auxin responsive protein

Signal peptide/ trange membrance

Cytochromes b561 constitute a class of intrinsic membrane proteins containing two haem molecules involved in ascorbate (vitamin C) regeneration. They have been suggested to function as electron transporters, shuttling electrons across membranes from ascorbate to an acceptor molecule. The one-electron oxidation product of ascorbate, monodehydro-ascorbate (MDHA) has been shown to function as an election acceptor for mammalian and plant cytochromes b561. The cytochrome b561-catalyzed reduction of MDHA results in the regeneration of the fully reduced ascorbate molecule. Cytochromes b561 have been identified in a large number of phylogenetically distant species, but are absent in prokaryotes. Most species contain three or four cytochrome b561 paralogous proteins (Verelst and Asard, 2003).

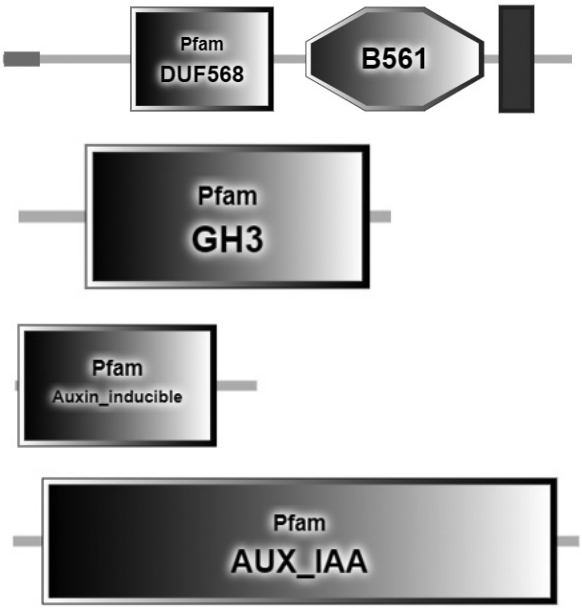

Figure 6. B561, GH3, Auxin_inducible, and AUX_IAA domain structure of Auxin response protein

Red: signal peptide; Blue: transmembrane helix region.

GH3 protein was initially discovered in soybeans (Glycine max) as a gene that responds to auxin at an early stage (Hagen and Guilfoyle, 1985). Subsequently, various GH3 proteins from plants were identified and classified into three groups: group I proteins synthesize JA-amino acid conjugates (Wakuta et al., 2011), group II proteins produce indole-3-acetic acid (IAA) conjugates (Staswick et al., 2005), and group III proteins are involved in amino acid conjugation to 4-substituted benzoate (Okrent et al., 2009). It is worth noting that this entry also encompasses proteins from bacteria, fungi, and animals.

Group 7: Auxin signaling protein

This family includes various proteins associated with plant dormancy and proteins repressed by auxin (Stafstrom et al., 1998). The expression of DRM/ARP family members might be influenced by stress and environmental factors (Rae et al., 2014).

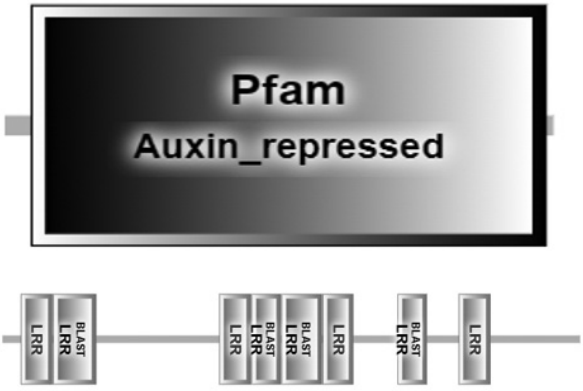

Figure 7. Auxin_repressed, and LRR domain structure of Auxin signaling protein

Leucine-rich repeats: Secreted by many different cell types, leucine-rich repeats (LRRs) play a crucial role in protein-protein interactions. These protein motifs possess a unique structural fold, enabling them to participate in diverse cellular signaling pathways. LRRs function as key mediators in recognition systems, aiding cell adhesion, immune responses, and pathogen detection. Their modular nature allows for remarkable versatility and adaptability. Through their involvement in various biological processes, LRRs have demonstrated their indispensability as essential molecular players, shedding light on potential therapeutic targets. As we delve further into their intricate mechanisms, we uncover the intricate tapestry of LRRs in shaping the functionality of living organisms.

4. CONCLUSION

C. canephora genome consists of 152 protein-coding genes related to auxin response which are divided into 7 main groups depending on domain such as Auxin-induced protein; Auxin-binding protein; Auxin transporter-like protein; Auxin carrier component; Auxin response factor; Auxin-

responsive protein; Auxin signaling protein.

There are some characteristic domains involved in auxin response such as EamA; AUX_IAA; Auxin inducible; Aldo_ket_red; Cupin; Aa_trans; B3, Auxin_resp; Mem_trans; B561; GH3; and LRR domain.

ỨNG DỤNG CÔNG NGHỆ SINH TIN HỌC TRONG PHÂN TÍCH NHỮNG GEN ỨNG VIÊN LIÊN QUAN ĐẾN AUXIN TRONG BỘ GEN CÀ PHÊ VỐI (*Coffea canephora* L.)

Nguyễn Đình Sỹ[1]

Ngày nhận bài: 14/06/2024; Ngày phản biện thông qua: 26/06/2024; Ngày duyệt đăng: 27/06/2024

TÓM TẮT

Cà phê vối (Coffea canephora) thuộc họ Rubiaceae, là một trong 2 loài cà phê được trồng nhiều nhất trên thế giới. Trong nghiên cứu này, chúng tôi sàng lọc và phân tích những gen ứng viên có liên quan đến chất điều hòa sinh trưởng auxin. Kết quả thu được cho thấy trong bộ gen của C. canephora có chứa 152 gen mã hóa protein có khả năng đáp ứng auxin. Dựa vào phân tích cấu trúc domain thì chúng được chia thành 7 nhóm chính: Protein được cảm ứng bởi auxin (Auxin-induced protein); Protein liên kết với auxin (Auxin-binding protein); Protein vận chuyển auxin (Auxin transporter-like protein); Protein là thành phần chất mang auxin (Auxin carrier component); Protein là nhân tố phản ứng auxin (Auxin response factor); Protein phản ứng với auxin (Auxin-responsive protein); Protein tín hiệu của auxin (Auxin signaling protein). Kết quả phân tích cũng đã chỉ ra rằng một số domain đặc trưng đã được tìm thấy trong cấu trúc của những protein liên quan đến auxin như EamA; AUX_IAA; Auxin inducible; Aldo_ket_red; Cupin; Aa_trans; B3, Auxin_resp; Mem_trans; B561; GH3; và LRR domain. Kết quả từ nghiên cứu này sẽ tạo nguồn dữ liệu tin cậy nhằm giúp các nhà khoa học định hướng chọn lựa những gen ứng viên tiềm năng nhất để nghiên cứu sâu chức năng của gen, từ đó ứng dụng vào chọn và tạo giống cà phê có khả năng sinh trưởng tốt hơn và thích ứng với biến đổi khí hậu trong tương lai.

Từ khóa: *Auxin, Bộ gen, C. canephora, Domain, sinh tin học.*

[1]*Khoa Khoa học Tự nhiên và Công nghệ, Trường Đại học Tây Nguyên;*
Tác giả liên hệ: Nguyễn Đình Sỹ; Tel: 0961367958; Email: ndsy@ttn.edu.vn.

REFERENCES

Abel S, Nguyen MD, Theologis A. The PS-IAA4/5-like family of early auxin-inducible mRNAs in Arabidopsis thaliana. J Mol Biol. 1995 Aug 25;251(4):533-49. doi: 10.1006/jmbi.1995.0454. PMID: 7658471.

Albuquerque, É. V. S. et al. (2015) 'Seed-Specific Stable Expression of the α-AI1 Inhibitor in Coffee Grains and the In Vivo Implications for the Development of the Coffee Berry Borer', Tropical Plant Biology, 8(3–4), pp. 98–107. doi: 10.1007/s12042-015-9153-0.

Barbosa, A. E. A. D. et al. (2010) 'α-Amylase inhibitor-1 gene from Phaseolus vulgaris expressed in Coffea arabica plants inhibits α-amylases from the coffee berry borer pest', BMC Biotechnology, 10(May 2014). doi: 10.1186/1472-6750-10-44.

Blakeslee JJ, Peer WA, Murphy AS. Auxin transport. Curr Opin Plant Biol. 2005 Oct;8(5):494-500. doi: 10.1016/j.pbi.2005.07.014. PMID: 16054428.

Bohren KM, Bullock B, Wermuth B, Gabbay KH. The aldo-keto reductase superfamily. cDNAs and deduced amino acid sequences of human aldehyde and aldose reductases. J Biol Chem. 1989 Jun 5;264(16):9547-51. PMID: 2498333.

Borhani DW, Harter TM, Petrash JM. The crystal structure of the aldose reductase.NADPH binary complex. J Biol Chem. 1992 Dec 5;267(34):24841-7. doi: 10.2210/pdb1abn/pdb. PMID: 1447221).

Bulgarelli, R. G. et al. (2016) 'Expression of metallothionein genes in coffee leaves in response to the absence or excess of Cu and Zn', Theoretical and Experimental Plant Physiology. Springer International Publishing, 28(4), pp. 371–383. doi: 10.1007/s40626-016-0075-5.

Coffee-genome hub database : http://coffee-genome.org

Denoeud, F. et al. (2014) 'The coffee genome provides insight into the convergent evolution of caffeine biosynthesis.', Science (New York, N.Y.), 345(6201), pp. 1181–1184. doi: 10.1126/science.1255274.

Dinh, S. N. and Kang, H. (2017) 'An endoplasmic reticulum-localized Coffea arabica BURP domain-containing protein affects the response of transgenic Arabidopsis plants to diverse abiotic stresses', Plant Cell Reports, 36(11), pp. 1829–1839. doi: 10.1007/s00299-017-2197-x.

Dunwell JM. Cupins: a new superfamily of functionally diverse proteins that include germins and plant storage proteins. Biotechnol Genet Eng Rev. 1998;15:1-32. doi: 10.1080/02648725.1998.10647950. PMID: 9573603.

Esteves Vieira, L. G. et al. (2006) 'Brazilian coffee genome project: An EST-based genomic resource', Brazilian Journal of Plant Physiology, 18(1), pp. 95–108. doi: 10.1590/S1677-04202006000100008.

Foreign Agricultural Service (FAS) (2024). https://fas.usda.gov/

Gil P, Green PJ. Regulatory activity exerted by the SAUR-AC1 promoter region in transgenic plants. Plant Mol Biol. 1997 Jul;34(5):803-8. doi: 10.1023/a:1005875300606. PMID: 9278170.

Gulbis JM, Zhou M, Mann S, MacKinnon R. Structure of the cytoplasmic beta subunit-T1 assembly of voltage-dependent K+ channels. Science. 2000 Jul 7;289(5476):123-7. doi: 10.1126/science.289.5476.123. PMID: 10884227.

Hagen G, Guilfoyle TJ. Rapid induction of selective transcription by auxins. Mol Cell Biol. 1985 Jun;5(6):1197-203. doi: 10.1128/mcb.5.6.1197-1203.1985. PMID: 4041007; PMCID: PMC366846.

Jack DL, Yang NM, Saier MH Jr. The drug/metabolite transporter superfamily. Eur J Biochem. 2001 Jul; 268(13):3620-39. doi: 10.1046/j.1432-1327.2001.02265.x. PMID: 11432728

Kagaya Y, Ohmiya K, Hattori T. RAV1, a novel DNA-binding protein, binds to bipartite recognition sequence through two distinct DNA-binding domains uniquely found in higher plants. Nucleic Acids Res. 1999 Jan 15;27(2):470-8. doi: 10.1093/nar/27.2.470. PMID: 9862967; PMCID: PMC148202.).

Kramer EM. PIN and AUX/LAX proteins: their role in auxin accumulation. Trends Plant Sci. 2004 Dec;9(12):578-82. doi: 10.1016/j.tplants.2004.10.010. PMID: 15564124.).

Liscum E, Reed JW. Genetics of Aux/IAA and ARF action in plant growth and development. Plant Mol Biol. 2002 Jun-Jul;49(3-4):387-400. PMID: 12036262.

McCarthy, A. A. et al. (2007) 'Cloning, expression, crystallization and preliminary X-ray analysis of the XMT and DXMT N-methyltransferases from Coffea canephora (robusta)', Acta Crystallographica Section F: Structural Biology and Crystallization Communications. International Union of Crystallography, 63(4), pp. 304–307. doi: 10.1107/S1744309107009268.

McClure BA, Guilfoyle T. Characterization of a class of small auxin-inducible soybean polyadenylated RNAs. Plant Mol Biol. 1987 Nov;9(6):611-23. doi: 10.1007/BF00020537. PMID: 24277197

McIntire SL, Reimer RJ, Schuske K, Edwards RH, Jorgensen EM. Identification and characterization of the vesicular GABA transporter. Nature. 1997 Oct 23;389(6653):870-6. doi: 10.1038/39908. PMID: 9349821.

Mishra, M. K. and Slater, A. (2012) 'Recent Advances in the Genetic Transformation of Coffee', Biotechnology Research International, 2012, pp. 1–17. doi: 10.1155/2012/580857.

Nguyen Dinh, S. et al. (2016) 'Abiotic stresses affect differently the intron splicing and expression of chloroplast genes in coffee plants (Coffea arabica) and rice (Oryza sativa)', Journal of Plant Physiology. Elsevier GmbH., 201, pp. 85–94. doi: 10.1016/j.jplph.2016.07.004.

Okrent RA, Brooks MD, Wildermuth MC. Arabidopsis GH3.12 (PBS3) conjugates amino acids to 4-substituted benzoates and is inhibited by salicylate. J Biol Chem. 2009 Apr 10;284(15):9742-54. doi: 10.1074/jbc.M806662200. Epub 2009 Feb 2. PMID: 19189963; PMCID: PMC2665095.

Pallavicini, A. et al. (2005) Transcriptomics of resistance response in Coffea arabica L.

Perrois, C. et al. (2015) 'Differential regulation of caffeine metabolism in Coffea arabica (Arabica) and Coffea canephora (Robusta)', Planta, 241(1), pp. 179–191. doi: 10.1007/s00425-014-2170-7.

Privat, I. et al. (2011) 'The "PUCE CAFE" Project: The First 15K Coffee Microarray, a New Tool for Discovering Candidate Genes correlated to Agronomic and Quality Traits', BMC Genomics, 12. doi: 10.1186/1471-2164-12-5.

Rae GM, Uversky VN, David K, Wood M. DRM1 and DRM2 expression regulation: potential role of splice variants in response to stress and environmental factors in Arabidopsis. Mol Genet Genomics. 2014 Jun;289(3):317-32. doi: 10.1007/s00438-013-0804-2. Epub 2014 Jan 18. PMID: 24442277.

Raharimalala, N. et al. (2021) 'The absence of the caffeine synthase gene is involved in the naturally decaffeinated status of Coffea humblotiana, a wild species from Comoro archipelago', Scientific Reports. Nature Publishing Group UK, 11(1), pp. 1–14. doi: 10.1038/s41598-021-87419-0.

Schade SZ, Early SL, Williams TR, Kézdy FJ, Heinrikson RL, Grimshaw CE, Doughty CC. Sequence analysis of bovine lens aldose reductase. J Biol Chem. 1990 Mar 5;265(7):3628-35. PMID: 2105951.

Spartz AK, Ren H, Park MY, Grandt KN, Lee SH, Murphy AS, Sussman MR, Overvoorde PJ, Gray WM. SAUR Inhibition of PP2C-D Phosphatases Activates Plasma Membrane H+-ATPases to Promote Cell Expansion in Arabidopsis. Plant Cell. 2014 May;26(5):2129-2142. doi: 10.1105/tpc.114.126037. Epub 2014 May 23. PMID: 24858935; PMCID: PMC4079373.

Stafstrom JP, Ripley BD, Devitt ML, Drake B. Dormancy-associated gene expression in pea axillary buds. Cloning and expression of PsDRM1 and PsDRM2. Planta. 1998 Aug;205(4):547-52. doi: 10.1007/s004250050354. PMID: 9684359.

Staswick PE, Serban B, Rowe M, Tiryaki I, Maldonado MT, Maldonado MC, Suza W. Characterization of an Arabidopsis enzyme family that conjugates amino acids to indole-3-acetic acid. Plant Cell. 2005 Feb;17(2):616-27. doi: 10.1105/tpc.104.026690. Epub 2005 Jan 19. PMID: 15659623; PMCID: PMC548830.

Software SMART (http://smart.embl.de).

Torres, L. F. et al. (2019) 'Expression of DREB-Like Genes in Coffea canephora and C. arabica Subjected to Various Types of Abiotic Stress', Tropical Plant Biology, 12(2), pp. 98–116. doi: 10.1007/s12042-019-09223-5.

Vadivelu, J. (2013) 'Microbial Pathogens and Strategies for Combating them: Science, Technology and Education', in.

Verelst W, Asard H. A phylogenetic study of cytochrome b561 proteins. Genome Biol. 2003;4(6):R38.

doi: 10.1186/gb-2003-4-6-r38. Epub 2003 May 28. PMID: 12801412; PMCID: PMC193617.

Wakuta S, Suzuki E, Saburi W, Matsuura H, Nabeta K, Imai R, Matsui H. OsJAR1 and OsJAR2 are jasmonyl-L-isoleucine synthases involved in wound- and pathogen-induced jasmonic acid signalling. Biochem Biophys Res Commun. 2011 Jun 17;409(4):634-9. doi: 10.1016/j.bbrc.2011.05.055. Epub 2011 May 17. PMID: 21619871.

Wilson DK, Bohren KM, Gabbay KH, Quiocho FA. An unlikely sugar substrate site in the 1.65 A structure of the human aldose reductase holoenzyme implicated in diabetic complications. Science. 1992 Jul 3;257(5066):81-4. doi: 10.1126/science.1621098. PMID: 1621098

WolF PSORT (https://wolfpsort.hgc.jp/)