ZANCO Journal of Pure and Applied Sciences
The official scientific journal of Salahaddin University-Erbil
https://zancojournals.su.edu.krd/index.php/JPAS

ISSN (print):2218-0230**, ISSN (online):** 2412-3986**, DOI: http://dx.doi.org/10.21271/zjpas**

RESEARCH PAPER

Evolutionary Phylogenetic Analysis of Isochorismate Synthase in Photosynthetic Eukaryotes

Sevan O Majed1, Sardar Hussein Rasool2, Bushra K Amin3, Abdulkarim Y Karim3*

1-Biology Department, College of Education, Salahaddin University-Erbil, Kurdistan Region, Iraq
2- Ministry of Education-KRG, Erbil, Kurdistan Region, Iraq
3-Biology Department, College of Science, Salahaddin University-Erbil, Kurdistan Region, Iraq

A B S T R A C T

In the present research study sequence logo was used to display the frequency of conserved amino acids of multiple alignment of isochorismate synthase (ICS) sequences. The result revealed that residues in stacks with single amino acids (indicating conserved positions) were taller than those in stacks with multiple amino acids (where there is more change). The MSA result showed that different ICS sequences in different plants had a very close similarity to each other and also some amino acids in some aligned positions were highly conserved or identical. In addition, the ML tree displayed that these different ICS sequences in distinct plant species had an evolutionary relationship with each other because they shared from a single common ancestor and also have been evolved from each other through gene duplication event at the earliest stage of evolution. This event has played a key role in the evolution of different ICS enzymes in a single species called as paralogs. This event in different species has played a key role in the evolution of ICS2 from ICS1 because the variation in ICS2 happened more recently. Moreover, speciation event happened more recently and led to the evolution of different ICS enzymes in different species called as orthologs.

KEY WORDS: Vitamin K; Isochorismate synthase; Multiple sequence alignment
DOI: http://dx.doi.org/10.21271/ZJPAS.33.6.4
ZJPAS (2020) , 33(6);30-37 .

1.INTRODUCTION

Vitamin K, also called as phylloquinone, phytomenadione, or phytonadione, is produced by photosynthetic plants. It occurs at highest concentrations in green leaves and dark green leafy vegetables but it is also discovered in far smaller levels in other plant tissues, roots, fruits, tubers and seeds. The finding of phylloquinone arose up from the detection in the late 1920s and early 1930s that chicks reared on a reconstituted 'fat-free' diet developed a hemorrhagic illness identified by a severe impairment in blood coagulation (Almquist and Stokstad 1935; Dam 1935; Dam and Schönheyder 1934; Holst and Halbrook 1933; McFarlane et al. 1931; Schonheydee 1935).

* **Corresponding Author:**
Abdulkarim Y Karim
E-mail: abdulkarim.karim@su.edu.krd
Article History:
Received: 30/06/2021
Accepted: 21/11/2021
Published: 20/12 /2021

Phylloquinone is widely distributed in higher plants but it has been hardly found in lower plants like seaweeds, ferns and mosses (Bouga and Combet 2015; Majed and Karim 2017). In green plants, it plays an important role in photosynthesis: while in animalit involves in the carboxylation of glutamate amino acids in proteins to form gamma-carboxyglutamate (Gla) domains (Fu et al. 2017). The conserved amino acids are often (but not always) found in specific protein domains called Gla domains. Residues present in Gla domain usually participate calcium binding, and are crucial for the biological function of all characterised Gla domains (Furt et al. 2010; Wajih et al. 2007). They have been noticed to play key roles in the regulation of three physiological processes including blood coagulation, bone metabolism and vascular biology (Stafford 2005;

Vermeer and Schurgers 2000). The non-existence of fat in the food and diet as the cause of the disease was reasonably ruled out, because haemorrhages still appeared in chicks receiving a daily supplement of cholesterol and oil from cod-liver or flax seeds (Huttner et al. 2006; Hankey et al. 2014). It appears that the haemorrhagic diseases are caused by a lack of any of the known vitamins (Almquist and Stokstad 1935; Dam 1935; Dam and Schönheyder 1934).

The metabolism of phylloquinone production in photosynthetic plants is however composed of two separate metabolic routes: the naphthoquinone ring synthesis and the phytyl moiety synthesis. The later one is also used for the formation of chlorophylls and tocopherols. Hereafter, we would like to emphasis the enzymatic steps involved in the naphthoquinone ring biosynthesis, because the formation of the phytyl-diphosphate precursor from the methylerythritol-phosphate route in photosynthetic plants is not exact to phylloquinone (Lichtenthaler 1999; Rohmer 2003).

The pathway of the naphthoquinone ring biosynthesis consists of nine enzymatic steps catalysed by isochorismate synthase (ICS), 2-Succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylic-acid (SEPHCHC) synthase, 2-S 2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate (SHCHC) synthase, O-Succinylbenzoate (OSB) synthase, 1,4-dihydroxy-2-naphthoyl-CoA (DHNA-CoA) synthase, 1,4-dihydroxy-2-naphthoyl-CoA (DHNA-CoA) thioesterase, 1,4-dihydroxy-2-naphthoate (DHNA) phytyltransferase (Almquist and Stokstad 1935; Dam 1935; Dam and Schönheyder 1934; McCoy et al. 2018; Emonds-Alt et al. 2017; Joyard et al. 2009). The present study focuses on the first and crucial enzyme, isochorismate synthase (ICS) of this pathway; because if this enzyme is not present in green plants, vitamin K cannot be synthesized. ICS catalyse the conversion of chorismate into isochorismate. Although most of photosynthetic plants possess single form of the ICS enzyme, several plants do possess two iso-forms, ICS1 and ICS2. Research studies reported that these two forms are encoded through two different genes. The *Arabidopsis* genome produces two different forms, ICS1 (At1g74710) and ICS2 (At1g18870), which share about 80% identity (Garcion et al. 2008; Strawn et al. 2007; Wildermuth et al. 2002). It is hypothesized that this enzyme has an evolutionary history and relationship within isoforms and different plants.

In present study, we plot the conservation and identity pattern of ICS enzyme isoforms sequence across different plants species using Logo and profile hidden Markov models. Furthermore, molecular phylogenetic analysis by maximum likelihood (ML) is used in order to construct evolution tree and understand the evolutionary events occurred during evolution of ICS1 and ICS2.

2. Materials and Methods

2.1 Retrieval of ICS enzyme sequences

The gene sequence of *Arabidopsis thaliana* Isochorismate synthase (ICS) was retrieved from NCBI's genomic sequence database and used to search for homologous sequences by BLAST-tool of NCBIhttp://blast.ncbi.nlm.nih.gov/Blast.cgi. After that, the whole collecting homologous ICS sequences were downloaded and then regulated into a fasta file. Then, the method of MUSCLE (Vesrion, 7) was used to construct the sequence alignment for the mutation detection. Then, the Maximum Likelihood tree tool (Version, 7.2) was applied to construct the phylogenetic tree for showing the phylogenetic relationship between the homologous ICS sequences.

2.2 Comparative sequence and phylogenetic analysis of ICS sequences

Firstly, in order to provide a compact graphical representation of the conservation pattern of a set of ICS sequences, HMMER 3.1 tool on the protein family (Pfam) seed alignment platform was used for constructing informative, interactive logos representing sequence alignments and profile hidden Markov models. These methods were applied to show the number of amino acids that play a key role in the ICS protein. Secondly, in order to reveal sequence similarity and identity. and also, molecular phylogenetic analysis among homologous ICS sequences, MEGA7 (Molecular Evolutionary Genetics Analysis version 7.0) was used to detect phylogenetic relationship.

3.Results and discussion

3.1 Sequence logo model of homologous sequences of Isochorismate synthase

Sequence logo also known as profile logo and HMM logo is one model of HMM logs. This was used for visualizing of letter (amino acid or residue) frequency of ICS sequences using HMM build from HMMER 3.1 on the pfam (Protein

family) seed alignment. Because the length of the sequence logo of ICS sequences was too much, the most important blocks were represented (fig. 1). The sequence logo represented each column with a stack of letters. The total height of each stack corresponds to a measure of the indifference of the column – typically, it is the information content of that position. The height of each letter within a stack depends on the frequency of that letter at that position. the profile logo was originally developed to show the extent of letter conservation in each column of an alignment. For example, the letter P (proline, blue colour) has the highest stack, meaning that is the highest conserved letter in ICS protein and also the letter Y (Tyrosine, brown) and (R) Arginine (dark red) are higher conserved than G (Glycine, yellow), E (Glutamic acid, red), I (Isoluesine, orange), F (Phenylalanine), K (Lycine, dark green) and V (Valine, orange). It can also be seen that the height of a given residue boosts with increasing frequency of the residue, and its height boost with increasing conservation of the aligned site; hence, residues within stacks with individual letters were taller than those in stacks with multiple letters.

3.2Multiple sequence alignment for homologous sequences of Isochorismate synthase

Muscle tool from MEGA7 software was used to construct multiple sequence alignment (MSA) of ICS sequences. A small part of MSA was displayed, because our data includes 187 sequences. MSA result revealed that there is a very close similarity and identity among ICS sequences. This result is very similar to previous result. This confirms that our query sequence is a member of the chorismate binding family. It can be seen that many residues are shared among ICS sequences and also signed with an asterisk at the top of the alignment: including P (Proline), D (Aspartic acid), Q (Glutamine), N (Asparagine), A (Alanine) Y (Tyrosine) and R (Arginine), G (Glycine, yellow), E (Glutamic acid, red), I (Isoluesine), F (Phenylalanine), K (Lycine) and V (Valine) (Fig. 2). These conserved residues revealed that these different enzymes in different plants have the same function as well as similar structure. According to a research study conducted by Booth and Suttie (1998) from Tufts University in 1998 that the first step of the phylloquinone biosynthetic pathway in all photosynthetic plants is carried by ICS. The level of phylloquinone in vegetables increases with increasing the ICS level. Another research result reported that these enzymes have a specific domain called as *Chorismate_bind. The structure of this domain in different plants is similar* (Oostende et al. 2008; Li et al. 2011; Widhalm et al. 2012). Because ICS sequences had a very close similarity and identity with each other, they may have an evolutionary relationship with each other and may be descended from a single common ancestor via the events of evolution.

3.3 Molecular phylogenetic analysis for ICS enzyme

Biological activity of ICS enzyme is widely dispersed in photosynthetic eukaryotes. Many research studies reported that this enzyme in different plants has the same function as well as similar structure (Garcion et al. 2008; Gross et al. 2006; Maintinguer et al. 2011). In addition, the MSA result showed that there is a very high sequence similarity and identity among homologous sequences of ICS (fig. 2). It is predicted that this enzyme in a plant species is evolutionarily related to another and they may be descended from a common ancestor via the events of evolution. To prove this prediction, Molecular phylogenetic tree by maximum likelihood (ML) was constructed using 187 homologous ICS sequences. The ML tree shows that at the earliest stage of evolution the process of gene duplication played an important role in evolution of multiple homologous sequences into two big clades: Clade 1 and Clade 2 (fig. 3). It can be seen that the ICS enzyme is more widely evolved in eudicots than monocots. Within the tree, there was two distinct types of homology between different ICS proteins which were derived from a single gene in one ancestral species. Paralogs denotes more than one protein found in the same species that were descended from one gene. Paralogous ICS proteins were evolved by the process of gene duplication (Fig. 3, see red ball). This process could lead to the formation of multiple paralogs in a species. The other alternative is orthologs. Here, it denotes two different ICS proteins in two different species were evolved through the process of speciation events and also, they were descended from a single common ancestral gene (Fig. 3, see black ball). The values on the branches represent the value of bootstrap which give more reliability and confidence with ML tree. The bootstrap value in this tree was 1000 replications. Interestingly, two different types of ICS protein representing ICS1 and ICS2 can be seen in within many species. Both ICS1 and ICS2 in some species

were combined together in a clade. They are hypothesized to have an evolutionary relationship because they have the fewest amino acid change. The sequences of which are highly conserved; therefor, they were clustered together and shared a common ancestor.

3.4 Molecular evolutionary relationship between ICS1 and ICS2

The ICS enzyme which is observed in photosynthetic plant species has two types; ICS1 and ICS2. These two types are more likely to be produced from each other via gene duplication processes at the early stage of evolution. To prove this hypothesis, all sequences of ICS1 and ICS2 in *Arabidopsis thaliana* were isolated and then phylogenetic tree reconstructed for them. The MSA result showed that there is an extremely close similarity and identity between the sequences of ICS1 and ICS2 (Fig. 4A). This result emphasised that ICS1and ICS2 was evolved from each other at the earliest stage of evolution because they shared from a common ancestor (Fig. 4B). It appears that variation in ICS2 occurred more recently and rapidly than in ICS1. event of gene duplication has played a key role in the evolution of ICS2 from ICS1. Interestingly, all ICS1are clustered together in a clade because they have the fewest residual variation. Similarly, all ICS2 are combined together in another clade. This result is consistent with previous research results obtained by other researchers. According to some research studies, Arabidopsis genome produces in fact two different and catalytically functional ICS, ICS1(At1g74710) and ICS2 (At1g18870), that share about 80% identity(Strawn et al. 2007; Garcion et al. 2008; Wildermuth et al. 2002). They reported that these enzymes have the same function and may have similar structures. With decreasing of ICS1 and ICS2 levels in green plants, the concentration of phylloquinone is decreasing.

Conclusion:

The enzymes of ICS1 and ICS2, which are involved in the synthesis of vitamin k, were widely dispersed in plant classes. Gene duplation and speciation events played a key role in the distribution of these enzymes. These were structurally and functionally similar in the various plants species because numerous amino acids found in the ICS1 and ICS2 were conserved. Furthermore, phylogenetic tree analysis showed that these in the different classes of the plant were closely evolutionarily related to each other because they were derived from a common ancestor at the early stage of evolution.

Conflict of interest:

The authors declare that they have no conflict of interest.

Acknowledgement:

Authors acknowledge the financial and technical supports from Biology department, College of Science, Salahaddin University in this work.

Figure legends

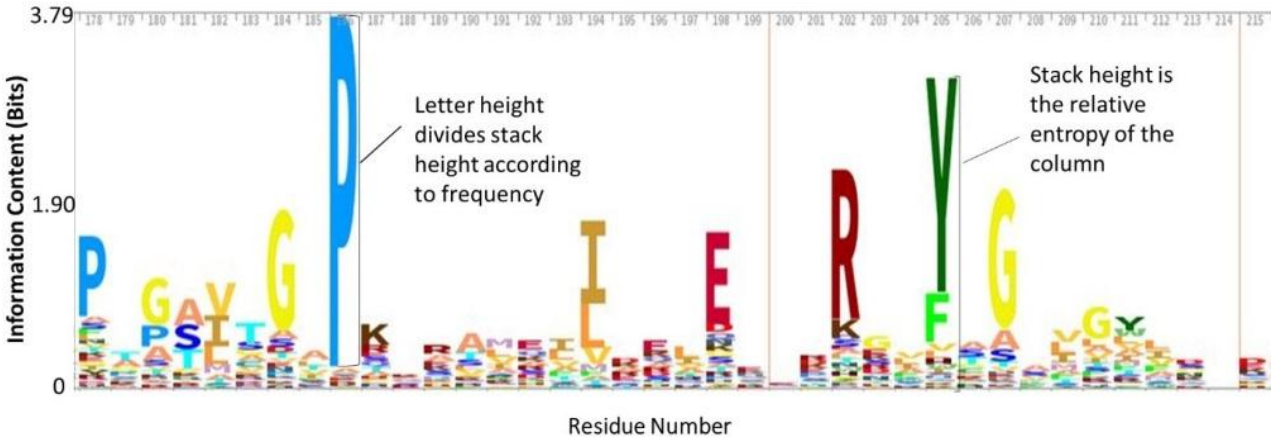

Figure 1 A partial sequence logo for Isochorismate synthase, it shows positions 178 to 215 of the Isochorismate synthase HMM logo from pfam (family: *Chorismate_bind* (pfam: PF00425)) produced using Skylign. This logo consists of colour coded stacks of letters representing amino acids at significant positions. This part is the highest conserved position in ICS enzyme.

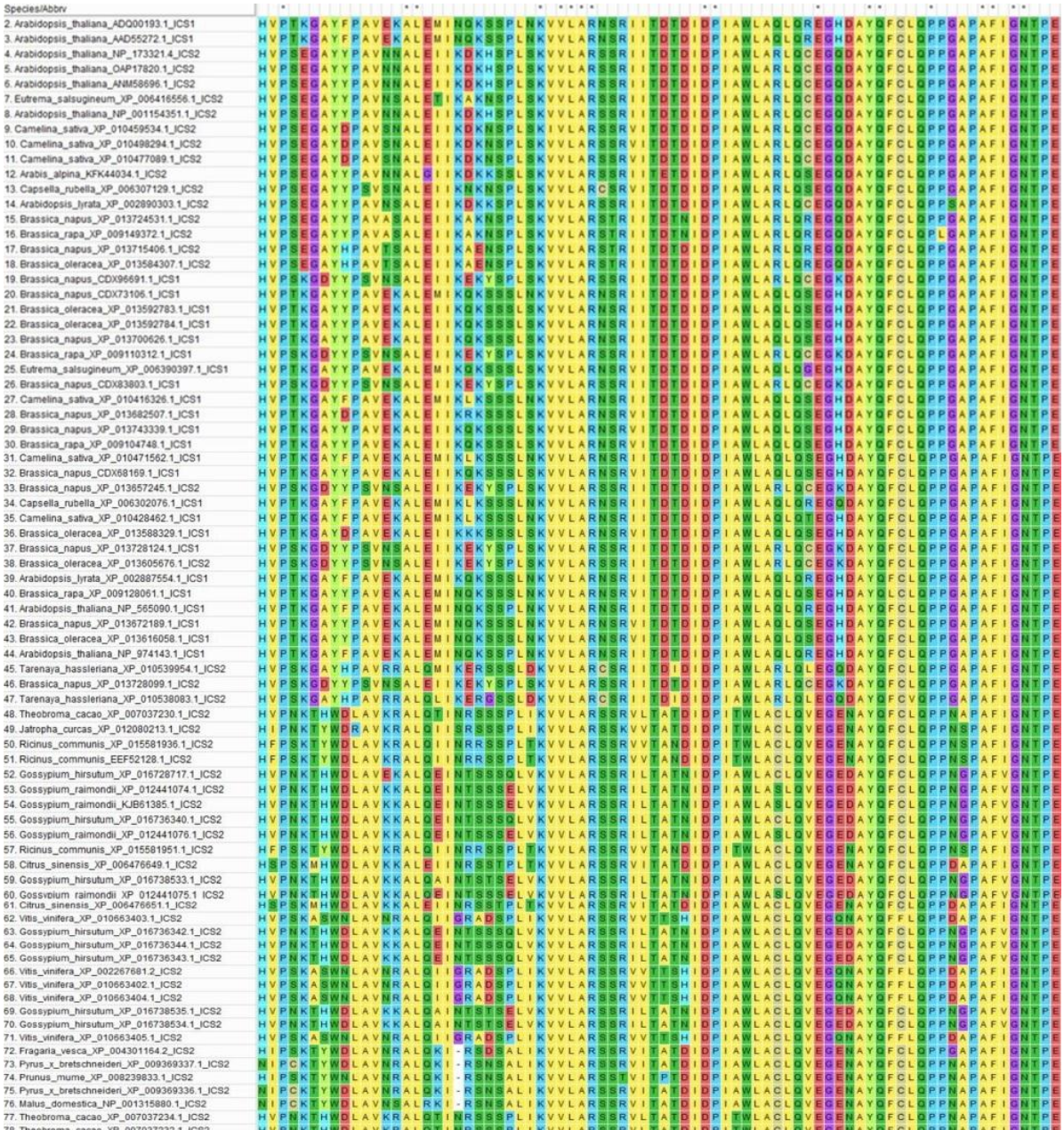

Figure 2 A part of multiple sequence alignements (MSAs) of ICS was indicated as an example. Representation of an example of MSAs of ICS constructed with MUSCLE algorithm. 187 homologous sequences from protein database are collected from (http://blast.ncbi.nlm.nih.gov/Blast.cgi) using *Arabidopsis thaliana* ICS as a query sequence. The colours correspond to amino acid conservations according to the presence of residual frequencies in each column and the features. Notice that more than one completely conserved amino acid signed with an asterisk at the top of the alignment.

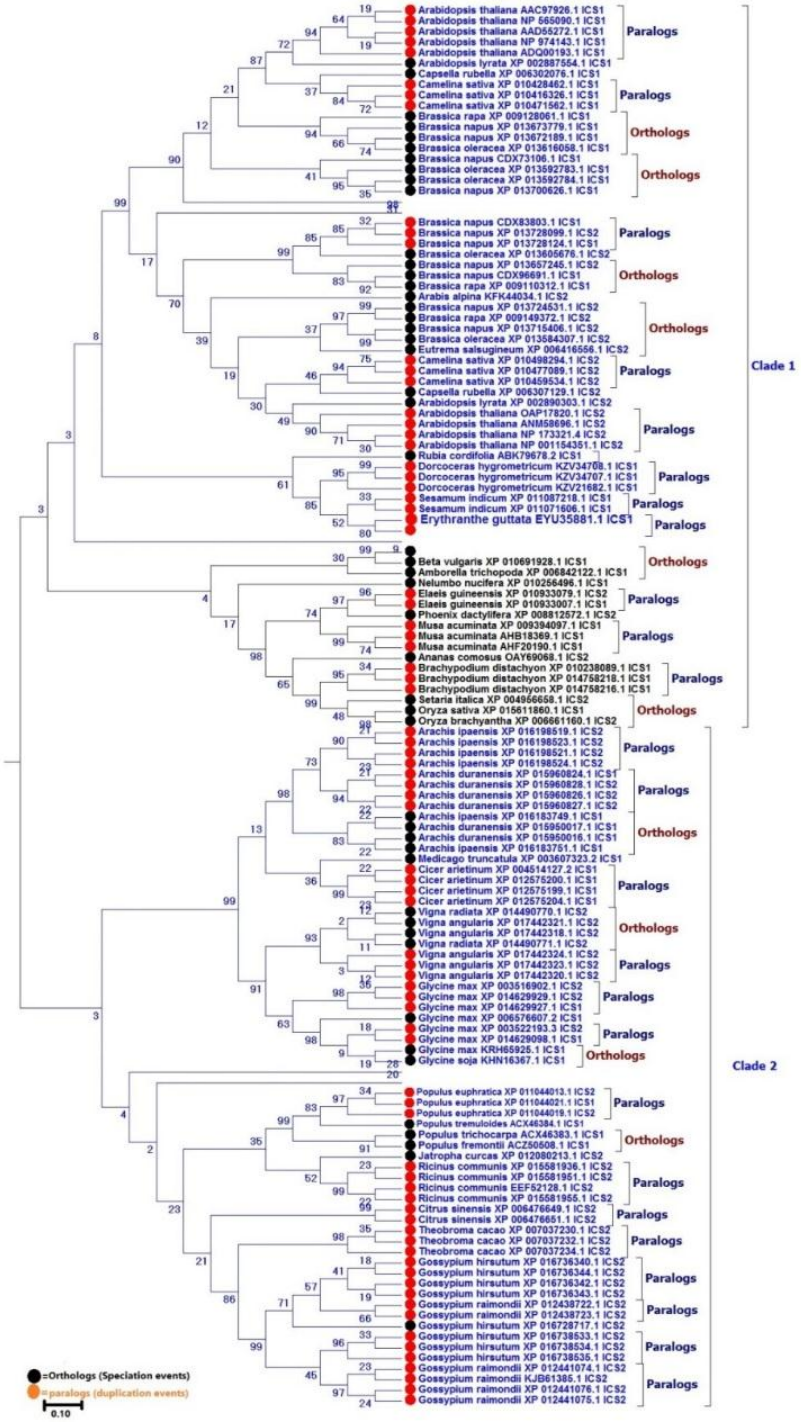

Figure 3 phylogenetic tree generated by maximum likelihood (ML) method, 187 sequences of ICS is compared. The ML tree represents two major homologous groups shared single common ancestor. Gene duplication and speciation events have played an important role in the distribution of the paralogs and orthologs, respectively. It appears that a close evolutionary relationship between the ICS sequences found in eudicots and monocots; because, they were clustered together in the clades. The numbers at the branches are bootstrap values (1000 replication).

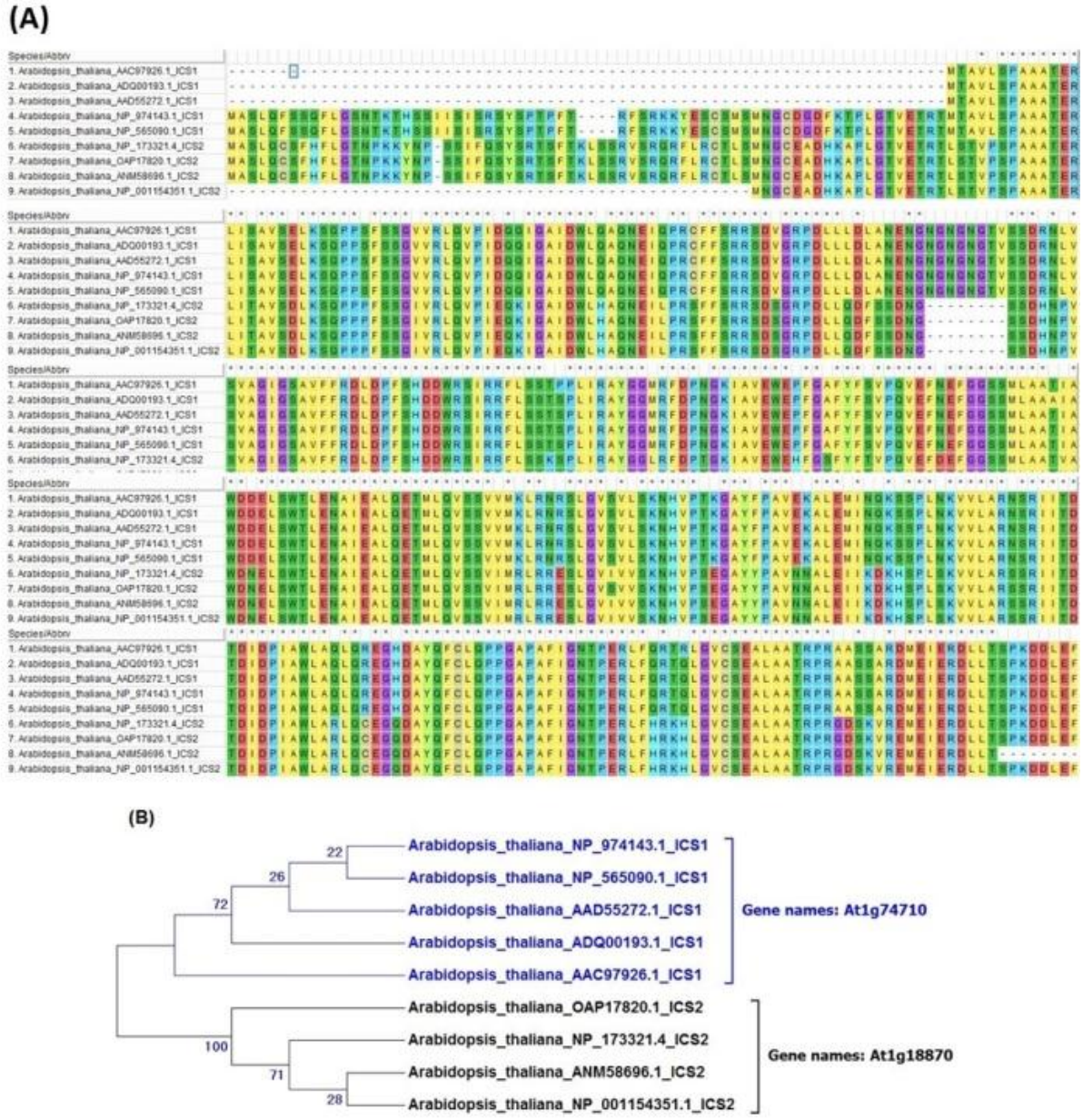

Figure 4 (A) Showing multiple sequence alignment (MSA) for ICS1 and ICS2 in *Arabidopsis thaliana.* The highest sequences similarity and identity was seen between these two forms of ICS; ICS1 and ICS2. **(B)** Evolutionary relationship among ICS1 and ICS2 was displayed**.** The numbers at the branches are bootstrap values (1000 replication).

Reference:

Almquist, H. J., & Stokstad, E. L. (1935). Hemorrhagic chick disease of dietary origin. Journal of Biological Chemistry. *Nature.* 111.p.105-113

Booth, S. L., & Suttie, J. W. (1998). Dietary Intake and Adequacy of Vitamin K. *The Journal of Nutrition* 128 (5). p.785-788. doi:10.1093/jn/128.5.785

Bouga, M., & Combet, E. (2015) Emergence of seaweed and seaweed-containing foods in the UK: focus on labeling, iodine content, toxicity and nutrition. *Foods.* 4 (2). p.240-253

Dam, H., (1935). The antihaemorrhagic vitamin of the chick. *The Biochemical journal*. 29 (6). p.1273-1285

Dam, H., & Schönheyder, F. (1934). A deficiency disease in chicks resembling scurvy. *J Biochemical Journal*, 28 (4). p.1355-1359. doi:10.1042/bj0281355 %

Emonds-Alt, B., Coosemans, N., Gerards, T., Remacle, C., & Cardol, P. (2017) Isolation and characterization of mutants corresponding to the MENA, MENB, MENC and MENE enzymatic steps of 5′-monohydroxyphylloquinone biosynthesis in Chlamydomonas reinhardtii. *The Plant Journal* 89 (1). P.141-154

Fu, X., Harshman, SG., Shen, X., Haytowitz ,D. B., Karl, J. P., Wolfe, BE., & Booth, S. L. (2017). Multiple Vitamin K Forms Exist in Dairy Foods. Current *Developments in Nutrition.* 1 (6). p. 12-26. doi:10.3945/cdn.117.000638

Furt, F., Oostende, C., Widhalm, J. R., Dale, M. A., Wertz, J., & Basset, G. J. (2010). A bimodular oxidoreductase mediates the specific reduction of phylloquinone (vitamin K1) in chloroplasts. *The Plant Journal.* 64 (1). p.38-46

Garcion, C., Lohmann, A., Lamodière, E., Catinot, J., Buchala, A., Doermann, P., & Métraux, J-P. (2008). Characterization and Biological Function of the ISOCHORISMATE SYNTHASE2 Gene of Arabidopsis. *Plant Physiology*. 147 (3). p.1279-1287. doi:10.1104/pp.108.119420 %J Plant Physiology

Gross, J., Cho, W, K., Lezhneva, L., Falk, J., Krupinska, K., Shinozaki, K., Seki, M., Herrmann, R. G., & Meurer, J., (2006). A Plant Locus Essential for Phylloquinone (Vitamin K1) Biosynthesis Originated from a Fusion of Four Eubacterial Genes. *Plant physiology*. 281 (25). p.17189-17196. doi:10.1074/jbc.M601754200

Hankey, G. J., Stevens, S. R., Piccini, J. P., Lokhnygina ,Y., Mahaffey, K. W., Halperin, J. L., Patel, M. R., Breithardt, G., Singer, D. E., & Becker, R. C. (2014). Intracranial hemorrhage among patients with atrial fibrillation anticoagulated with warfarin or rivaroxaban: the rivaroxaban once daily, oral, direct factor Xa inhibition compared with vitamin K antagonism for prevention of stroke and embolism trial in atrial fibrillation. Stroke. *Nature.* 45 (5). p.1304-1312

Holst, W. F., & Halbrook E. R. (1933). A "SCURVY-LIKE" DISEASE IN CHICKS. *J Science.* 77 (1997). p.354-354. doi:10.1126/science.77.1997.354 %J Science

Huttner, H. B., Schellinger P., Hartmann, M., Köhrmann, M, Juettler, E., Wikner J, Mueller, S., Meyding-Lamade, U., Strobl, R., & Mansmann, U. (2006). Hematoma growth and outcome in treated neurocritical care patients with intracerebral hemorrhage related to oral anticoagulant therapy: comparison of acute treatment strategies using vitamin K, fresh frozen plasma, and prothrombin complex concentrates. *Stroke.* 37 (6). p.1465-1470

Joyard, J., Ferro, M., Masselon, C., Seigneurin-Berny, D., Salvi, D., Garin, J., & Rolland, N. (2009). Chloroplast proteomics and the compartmentation of plastidial isoprenoid biosynthetic pathways. *Molecular plant.* 2 (6). p.1154-1180

Li, Q., Mavrodi, D. V., Thomashow, L. S., Roessle, M., & Blankenfeldt, W. (2011). Ligand binding induces an ammonia channel in 2-amino-2-desoxyisochorismate (ADIC) synthase PhzE. *Journal of Biological Chemistry*. 286 (20). p.18213-18221

Lichtenthaler, H. K. (1999). THE 1-DEOXY-D-XYLULOSE-5-PHOSPHATE PATHWAY OF ISOPRENOID BIOSYNTHESIS IN PLANTS. *Molecular plant.* 50 (1). p.47-65. doi:10.1146/annurev.*arplant*.50.1.47

Maintinguer, S. I, Fernandes, B. S, Duarte, I. C., Saavedra, N. K., Adorno, M. T., & Varesche, M. A. (2011). Fermentative hydrogen production with xylose by Clostridium and Klebsiella species in anaerobic batch reactors. *International Journal of Hydrogen Energy.* 36 (21). p.13508-13517

Majed, S. O, & Karim, A. Y. (2017). Phylogenetic Relationship among Phosphomannose Isomerase (PMI), Phosphomannose Mutase (PMM) and GDP-D-Mannose Pyrophosphorylase (GMP) in Photosynthetic Eukaryotes. *ZANCO Journal of Pure and Applied Sciences.* 28 (6). p.27-38

McCoy, R. M., Utturkar S. M. , Crook, J. W., Thimmapuram. J., & Widhalm, J. R. (2018). The origin and biosynthesis of the naphthalenoid moiety of juglone in black walnut. *Horticulture research*. 5 (1). p.50-67

McFarlane, W. D., Graham, J. R., & Hall, G. E. (1931). Studies in Protein Nutrition of the ChickI. The Influence of Different Protein Concentrates on the Growth of Baby Chicks, when Fed as the Source of Protein in Various Simplified Diets. *The Journal of Nutrition*. 4 (3). p.331-349. doi:10.1093/jn/4.3.331

Oostende, C., Widhalm, J. R, & Basset G. J. (2008). Detection and quantification of vitamin K(1) quinol in leaf tissues. *Phytochemistry*. 69 (13). p.2457-2462. doi:10.1016/j.phytochem.2008.07.006

Rohmer, M. (2003). Mevalonate-independent methylerythritol phosphate pathway for isoprenoid biosynthesis. Elucidation and distribution. *Pure and Applied Chemistry*. vol 75. p.1-19. doi:10.1351/pac200375020375

Schonheydee, F. (1935). The antihaemorrhagic vitamin of the chick. Measurement and biological action. *Nature*. 135. p.640-653. doi:10.1038/135653a0

Stafford, D. (2005). The vitamin K cycle. *Journal of Thrombosis and Haemostasis.* 3 (8). p1873-1878

Strawn, M. A., Marr, S. K., Inoue, K., Inada, N., Zubieta, C., & Wildermuth B. C. (2007). Arabidopsis isochorismate synthase functional in pathogen-induced salicylate biosynthesis exhibits properties consistent with a role in diverse stress responses. *Molecular Plant*. 282 (8). p.5919-5933

Vermeer, C., & Schurgers, L. J., (2000). A comprehensive review of vitamin K and vitamin K antagonists. *Hematology/oncology clinics of North America.* 14 (2). p.339-353

Wajih, N., Hutson, S. M., & Wallin, R. (2007). Disulfide-dependent Protein Folding Is Linked to Operation of the Vitamin K Cycle in the Endoplasmic Reticulum A PROTEIN DISULFIDE ISOMERASE-VKORC1 REDOX ENZYME COMPLEX APPEARS TO BE RESPONSIBLE FOR VITAMIN K1 2, 3-EPOXIDE REDUCTION. *Journal of Biological Chemistry.* 282 (4). p.2626-2635

Widhalm, J. R., Ducluzeau, A. L., Buller, N. E., Elowsky, C. G., Olsen, L. J., & Basset, G. J. (2012) Phylloquinone (vitamin K1) biosynthesis in plants: two peroxisomal thioesterases of lactobacillales origin hydrolyze 1, 4-dihydroxy-2-naphthoyl-coa. The plant journal 71 (2):205-215

Wildermuth, M. C., Dewdney, J., Wu, G., & Ausubel, F. M. (2002) corrigendum: Isochorismate synthase is required to synthesize salicylic acid for plant defence. Nature 417:571. doi:10.1038/417571a.