IJMS 2023 10 (1): 134 - 150

International Journal of Multidisciplinary Studies (IJMS)

Volume 10, Issue I, 2023

Comparative bioinformatics analysis of Pyrroline-5-carboxylate synthase (*P5CS*), Betaine aldehyde dehydrogenase (*BADH*) and Ferritin encoding genes in several crop species

Kaluthanthri D.V.S.[1], Dasanayaka P.N.[1*] and Perera S.A.C.N.[2]

[1] Department of Botany, Faculty of Applied Sciences, University of Sri Jayewardenepura, Sri Lanka
[2] Department of Agricultural Biology, Faculty of Agriculture, University of Peradeniya, Sri Lanka

ABSTRACT

*Improving abiotic stress tolerance in crops using both conventional breeding and transgenic techniques has become highly important. Of them, the transgenic technique provides means for the introgression of novel genes across species. Prior knowledge of conserved regions of a gene throughout the evolutionary process and the evolutionary divergence of genes among different plant species and families for coding sequences is vitally important in the genetic engineering processes. The genes***:** *Pyrroline-5-carboxylate synthase (P5CS), Betaine aldehyde dehydrogenase (BADH) and Ferritin have been reported to be involved in delivering tolerance to abiotic stresses in crops. In this study, bioinformatics analysis was used to study coding sequences of the three genes encoding P5CS, BADH and Ferritin in nine different crop species. Coding sequences were retrieved from National Center for Biotechnology Information (NCBI) and ClustalW multiple sequence alignment was performed using MEGA$_5$ software. A phylogenetic tree following the maximum likelihood approach with 100 bootstrap analysis and pairwise distances was obtained. DnaSP$_5$ software was used to analyze the conserved regions. P5CS, BADH and Ferritin genes recorded sixteen, seven and six conserved regions respectively with significant ($P\leq 0.05$) conservation and homozygosity values. Phylogenetic trees of P5CS and BADH showed three distinct clusters whereas only two clusters were observed for the Ferritin gene. There was a significant evolutionary divergence among CDS of P5CS (0.011 - 0.458), BADH (0.017 - 0.406) and Ferritin (0.009 - 0.509) genes. The results revealed the three genes to* ***possess*** *several regions conserved throughout the evolutionary process. Phylogeny reconstruction of the genes revealed the existence of two groups, separating monocot from dicot plants. The information derived will be important in the genetic engineering of crops for abiotic stress tolerance.*

KEYWORDS: *BADH, Conserved regions, evolutionary divergence, Ferritin, P5CS, phylogenetic relationships*

Corresponding Author: Dasanayaka P.N., Email: nilanthiedas@sjp.ac.lk

1 INTRODUCTION

Abiotic stresses including drought and salinity are one of the most decisive limiting factors for stable crop production. This is especially important in developing countries where the largest growth in populations will pose an enormous demand for the stability of food supplies. Environmental stress factors cause depreciation in crop yields by up to 70% when compared to the yields under favorable conditions (Boyer, 1982). Accordingly, the stability of crop yields under changing environmental conditions has become a priority in breeding programs.

Even though a few countries are considered as arid and naturally short of water, there had not been a true awakening to the global threat of water stress caused by the rapidly increasing world population and the accompanying rapid increases in water use for social and economic development (Cosgrove and Rijsberman, 2000). The pressure on water resources will grow significantly in more than 60% of the world, including large areas of Africa, Asia, and Latin America due to the increased water withdrawals (Alcamo *et al*., 1999).

Improving the drought stress tolerance of crops through adaptive strategies is important to ensure food security. To achieve this goal without increasing the area of cultivated land, the emphasis must be on key traits related to plant productivity and adaptation to environmental challenges (Mwadzingeni *et al*., 2016). Different techniques such as marker-assisted breeding, quantitative trait locus mapping, and introgression from the wild gene pool are being employed to improve drought tolerance (Gupta *et al*., 2017; Bhatta *et al*., 2018; Przewieslik-Allen *et al*., 2019). In contrast to conventional breeding, the transgenic technique seems to be a more attractive approach which allows the direct introduction of a single or group of desired genes breaking the species barrier (Gosal *et al*., 2009). To date, only a few genetically-modified crops for drought tolerance have been developed and approved (Khan *et al*., 2019). Drought tolerance is a complex quantitative polygenic trait controlled by a large number of genes, and thus it is difficult to understand the underlying molecular and physiological mechanisms of tolerance mechanisms (Hu and Xiong, 2014; Senapati *et al*., 2018).

Over the years, numerous genes have been isolated from various plants and inserted in transgenic plants to induce stress resistance (Khan *et al*., 2019). Such genes could be either the genes involved in cellular protection including osmoprotectants, membrane stabilization, detoxification, and transport proteins or genes that encode transcription factors and signalling molecules (Vendruscolo *et al*., 2007). The genes**:** Pyrroline-5-carboxylate synthase (*P5CS*), Betaine aldehyde dehydrogenase (*BADH*) and Ferritin have been reported to be involved in delivering tolerance to abiotic stresses in various crop plants.

Proline is a well-known proteogenic amino acid which acts as a compatible osmoprotectant and accumulates under osmotic stress to protect the cellular structure and function (Hmida-Sayari *et* *al*., 2005; Hayat *et al*., 2012). Proline under stress works as an antioxidant and scavenges reactive oxygen species, protects the denaturation of macromolecules, and regulates cytosolic activity (Sawahel and Hassan, 2002). The accumulation of proline in plants is correlated with the tolerance to drought stress. In plants, proline is synthesized from glutamate through the glutamate pathway. The reduction of glutamate to its semialdehyde is catalyzed by P5CS (Delta -1 -pyrroline-5-carboxylate synthase) enzyme, which then reduces to proline. The enzyme, P5CS is a rate-limiting enzyme for the synthesis of proline by feedback inhibition of P5CS (Zhang *et al*, 1995). The gene is functionally well characterized at the molecular level, but there is more to learn about its evolutionary path in the plant kingdom, particularly the drive behind functional (osmoprotective and developmental) divergence of duplication of *P5CS* genes (Rai and Penna, 2013). Glycine betaine is a quaternary ammonium compound which is known to have a protective role against drought stress by maintaining osmotic balance and protecting the quaternary structures of proteins (Giri, 2011). Various genes (*BADH, COD, CDH, and betaA*) involved in the synthesis of glycine betaine, have been inserted into transgenic plants. The *BADH* gene-encoding enzyme synthesizes glycine betaine by catalyzing betaine aldehyde into glycine betaine (Wang *et al*., 2010; Annunziata *et al*., 2011; He *et al*., 2011; Demirkol, 2020).

Transporter genes play an important role in restoring ionic homeostasis under stress. Ferritins are iron-storage proteins which sequester and release iron when needed. Ferritins are highly conserved in plants (Borg *et al*., 2012). Previous studies have revealed the induction of tolerance to various abiotic stresses like cold, heat and drought stresses by the over-expression of the ferritin gene (Borg *et al*., 2012; Zang *et al*., 2017).

Accordingly, in the present study, bioinformatics analysis was used to study the conserved regions, phylogenetic relationships and evolutionary divergence of the genes encoding for Pyrroline-5-carboxylate synthase (*P5CS*), Betaine aldehyde dehydrogenase (*BADH*) and Ferritin in several crop species using biological data retrieved from bioinformatical databases.

2 RESEARCH METHODOLOGY

2.1 Data collection and information resources:

DNA sequences of genes belonging to different plant species were retrieved from the National Center for Biotechnology Information (NCBI) from the databases at URL: (http://www.ncbi.nih.gov). A total of thirty-eight sequences belonging to nine different monocot and dicot crop species were analyzed in the study (Table 1).

Table 1. Details of the plant species subjected to bioinformatic analysis

Scientific name	Common name	Family	Monocot / Dicot
Cajanus cajan	Pigeon pea	Fabaceae	Dicot
Cicer arietinum	Chickpea	Fabaceae	Dicot
Oryza sativa	Asian rice	Poaceae	Monocot
Setaria italica	Foxtail millet	Poaceae	Monocot
Sorghum bicolor	Sorghum	Poaceae	Monocot

Vigna angularis	Adzuki bean	Fabaceae	Dicot
Vigna radiata	Mung bean	Fabaceae	Dicot
Vigna unguiculate	Cowpea	Fabaceae	Dicot
Zea mays	Maize	Poaceae	Monocot

2.2 Bioinformatic analyses

Phylogeny analysis using the retrieved sequences was carried out with ClustalW multiple sequence alignment program using MEGA_5 software (Tamura *et al*., 2011). The maximum likelihood approach implemented in the MEGA_5 software (Tamura *et al*., 2011) was used to construct phylogenetic tree relationships among sequences of the selected three genes. The analysis was performed with 100 replications of Bootstrap analysis by following Kimura-2- parameter model (substitution model) and the classical nearest neighbor interchange (tree inference option). Pairwise distances between each plant species were obtained by following the Maximum Composite Likelihood model (Tamura *et al*., 2004; 2011) using MEGA_5 software (Tamura *et al*., 2011). The DnaSP version 5.0 software (Librado and Rozas, 2009) was used to analyze the conserved DNA regions of the three genes among the studied plant species. Protein domains of the conserved regions were identified using UniProt data available on https://www.uniprot.org/ website.

3 RESULTS

3.1 Retrieval of sequences

The DNA sequences of ***P5CS***, ***BADH*** and Ferritin encoding genes belonging to nine different plant species namely *Cajanus cajan, Cicer arietinum, Oryza sativa, Setaria italica, Sorghum bicolor, Vigna angularis, Vigna radiata, Vigna unguiculate* and *Zea mays* were retrieved from the NCBI website. All the studied coding sequences were complete (Table 2).

3.2 Multiple sequence alignment

There were CDS length variations both within and among different species. Multiple sequence alignment showed the presence of conserved regions in the studied three gene sequences. The *P5CS* gene showed 16 conserved regions with $P \leq 0.05$. The conservation value and homozygosity value of the observed regions ranged from 0.550 – 0.559 and 0.781 – 0.812 respectively (Table 3). Furthermore, seven conserved regions were observed for the *BADH* gene with the probability of $P \leq 0.05$. The conservation values of the regions varied between 0.591 and 0.600 whereas the homozygosity value ranged from 0.811 to 0.818 (Table 4) Also, the ferritin gene had six conserved regions with $P \leq 0.05$ and recorded conservation values between 0.593 and 0.629. The least homozygosity value was 0.793 while the highest value was 0.826 out of the observed regions (Table 5). Functional protein domains of conserved regions were identified in *P5CS* (Table 6a) and *BADH* genes (Table 6b). There were not any identifiable protein domains with respect to any of the ferritin gene's conserved regions.

Table 2. Information on CDS utilized in the bioinformatic study

	Delta-1-pyrroline-5-carboxylate synthase		**Betaine aldehyde dehydrogenase (BADH)**		**Ferritin**	
Plant Species	**NCBI reference Sequence**	**Selected region**	**NCBI reference Sequence**	**Selected region**	**NCBI reference Sequence**	**Selected region**
1. Cajanus cajan	XM_020360772.2 XM_020359318.2	84-2246 52-2353	XM_020379819.2	59-1570	XM_020363285.2	128-898
2. Cicer arietinum	XM_027335197.1	51-2330	XM_004508765.3	239-1750	XM_004492435.3	29-805
3. Oryza sativa	XM_015784690.2 XM_015766717.2	212-2362 233-2440	XM_015795403.2 XM_015781605.1	191-1702 107-1624	XM_015762679.2 XM_015762143.2	289-1056 90-848
4. Setaria italica	XM_004961829.4	258-2408	XM_004973405.4 XM_004975822.4	133-1650 179-1696	XM_004977572.4 XM_004978526.4	73-822 94-855
5. Sorghum bicolor	XM_021447400.1 XM_021455806.1	275-2425 355-2544	XM_002444312.2 XM_002447933.2	118-1635 171-1691	XM_021466199.1	92-853
6. Vigna angularis	XM_017574942.1	144-2207	XM_017556443.1	191-1702	NM_001329818.1	1-768
7. Vigna radiata	XM_014640693.2 XM_014661008.2 XM_014649439.2	85-2232 70-2325 144-2294	XM_014654124.2 XM_014641488.2	161-1672 81-1592	XM_014641484.2	131-898
8. Vigna unguiculata			XM_028065277.1	148-1659	XM_028073755.1	133-888
9. Zea mays	NM_001352327.1	145-2295	NM_001164332.1 NM_001112311.2	1-1521 88-1608	NM_001112093.2	107-871

Table 3. Details of the conserved regions of Delta-1-pyrroline-5-carboxylate synthase gene

Region	Start - End	Conservation	Homozigosity	P value	Conserved Motif
1	340-419	0.550	0.781	0.05	ARRYTNRYYMAYAGCAGYTTYKCHGAYMTKCARAANCCNCARVBNRANHTNGAYGGVAARGCHTGYGCNGCYGTBGGDCA
2	342-421	0.550	0.781	0.05	RYTNRYYMAYAGCAGYTTYKCHGAYMTKCARAANCCNCARVBNRANHTNGAYGGVAARGCHTGYGCNGCYGTBGGDCARA
3	343-422	0.550	0.783	0.05	YTNRYYMAYAGCAGYTTYKCHGAYMTKCARAANCCNCARVBNRANHTNGAYGGVAARGCHTGYGCNGCYGTBGGDCARAR
4	344-495	0.553	0.785	0.01	TNRYYMAYAGCAGYTTYKCHGAYMTKCARAANCCNCARVBNRANHTNGAYGGVAARGCHTGYGCNGCYGTBGGDCARARYRKNCTBATGGCKMTYTAYGAYRYNHTRTTYAVYCARCTNGAYGTVWCNTCNKCBCARCTTCTWGTNAMNGAY
5	559-656	0.550	0.805	0.05	YTNARRGTHRTNCCNDTDTTYAAYGARAAYGAYGCHRTYAGYACHMGRARRSMDYCVKAYGAGGGCAAGMACTGYSTGCAGGATTCHTCKGGYRTHTT
6	560-680	0.553	0.812	0.02	TNARRGTHRTNCCNDTDTTYAAYGARAAYGAYGCHRTYAGYACHMGRARRSMDYCVKAYGAGGGCAAGMACTGYSTGCAGGATTCHTCKGGYRTHTTYTGGGAYAAYGAYAGYYTRKCHRS
7	584-681	0.550	0.825	0.05	ARAAYGAYGCHRTYAGYACHMGRARRSMDYCVKAYGAGGGCAAGMACTGYSTGCAGGATTCHTCKGGYRTHTTYTGGGAYAAYGAYAGYYTRKCHRSW
8	586-683	0.550	0.826	0.05	AAYGAYGCHRTYAGYACHMGRARRSMDYCVKAYGAGGGCAAGMACTGYSTGCAGGATTCHTCKGGYRTHTTYTGGGAYAAYGAYAGYYTRKCHRSWYT
9	612-789	0.550	0.801	0.01	DYCVKAYGAGGGCAAGMACTGYSTGCAGGATTCHTCKGGYRTHTTYTGGGAYAAYGAYAGYYTRKCHRSWYTDYTRGCNNHNGARYTNAAWGCHGAYCTYCTWRTYHTRYTNAGYGAYGTDGADGGNYTBTAYAGYGGYCCWCCWAVYGAHCCHHVNTCRAARHTNATHYAYACNTAY
10	1327-1406	0.550	0.794	0.05	RTYYTNGARAARRYWTMWTSHCCWYTRGGWGTNCTNYTNRTYRTWTTYGARTCHMGNCCYGATGCHYTDGTNCAGATWGC
11	1328-1407	0.550	0.793	0.05	TYYTNGARAARRYWTMWTSHCCWYTRGGWGTNCTNYTNRTYRTWTTYGARTCHMGNCCYGATGCHYTDGTNCAGATWGCD
12	1330-1409	0.550	0.792	0.05	YTNGARAARRYWTMWTSHCCWYTRGGWGTNCTNYTNRTYRTWTTYGARTCHMGNCCYGATGCHYTDGTNCAGATWGCDKC
13	1331-1419	0.551	0.797	0.04	TNGARAARRYWTMWTSHCCWYTRGGWGTNCTNYTNRTYRTWTTYGARTCHMGNCCYGATGCHYTDGTNCAGATWGCDKCWYTRGCRATY
14	1341-1501	0.553	0.793	0.01	WTMWTSHCCWYTRGGWGTNCTNYTNRTYRTWTTYGARTCHMGNCCYGATGCHYTDGTNCAGATWGCDKCWYTRGCRATYMGAAGTGGNAAYGGBYTDYTNYTDAAAGGDGGMAARGARGCYMDVMGRTCAAAYRMDRYHYTRCAYAARGTNATWAYTDVDG
15	1531-1803	0.553	0.808	0.00	ATWGGVCWWGTKAMHWVHARRGMHGARATHSCDKABYTRCTHRMGYTKSATGAYGTVATWGATCTDGTVRTHCCHAGAGGHAGYAABMADCTNGTBTCNCARATMAARDVDDCDACHARRATYCCWGTYYTDGGNCATKCHGAYGGWRTHTGYCAYGTHTWYRTHGAYAARWCWGCYRAYDTKRABATRGCAAARMRDATWRTDHKDGAYGCHAARRYWGATTAYCCDGCVGSVTGYAATGCHATGGARACNYTDCTWRTNCAYAWRGAYYTN
16	2047-2223	0.559	0.791	0.00	RADRYKTTYHTRHVBMRRGTHGAYAGTGCYGCKGYDTTYYAYAATGCDAGYACNMGDTTYWSTGAYGGRRCWCGHTTTGGNYTNGGNGCWGAGGTKGGMATWAGYACARGBMGNATHCATGCHMGNGGNCCHGTNGGHGTWGADGGDYTBYTRACHAMNMRMTGSATHHTVMRNGGD

Table 4. Details of the conserved regions of Betaine aldehyde dehydrogenase gene

Region	Start - End	Conservation	Homozigosity	P value	Conserved Motif
1	439-531	0.591	0.811	0.04	WSNYAYVTHCKNARRGARCCHMTYGGDGTWGTWGSRYTDATHACWCCHTGGAAYTATCCKMTSYTRATGGCDACDTGGAARGTHGCWCCTKCY
2	443-601	0.597	0.825	0.00	AYVTHCKNARRGARCCHMTYGGDGTWGTWGSRYTDATHACWCCHTGGAAYTATCCKMTSYTRATGGCDACDTGGAARGTHGCWCCTKCYYTKGCWGCKGGBTGTRCHRCWRTRYTRAARCCNTCWGARYTKKCWTCYBTRASHTGYTTRGAGCTDGSTG
3	751-871	0.593	0.821	0.02	DCHGCWKCHCMDMTRRYYAAGCCTGTTWCDYTRGARCTTGGDGGVAAAAGYCCWHTHRTDGTNTTYGAKGAYRTTSGTGAYVTYGAHAARRCTGYTGARTGGRCHMTSTTTGGBWKYTTYK
4	780-875	0.591	0.816	0.04	DYTRGARCTTGGDGGVAAAAGYCCWHTHRTDGTNTTYGAKGAYRTTSGTGAYVTYGAHAARRCTGYTGARTGGRCHMTSTTTGGBWKYTTYKBDAM
5	784-876	0.600	0.818	0.03	GARCTTGGDGGVAAAAGYCCWHTHRTDGTNTTYGAKGAYRTTSGTGAYVTYGAHAARRCTGYTGARTGGRCHMTSTTTGGBWKYTTYKBDAMH
6	1343-1434	0.598	0.808	0.03	TNTGGRTNAAYTGYKCNCARCCMWSCTTHDBYCADGCYCCHTGGGGHGGBRWHAARCGNAGYGGHTTYGGHCGNGARYTNGGASARKGGGGH
7	1374-1467	0.596	0.812	0.03	YCADGCYCCHTGGGGHGGBRWHAARCGNAGYGGHTTYGGHCGNGARYTNGGASARKGGGGHMTNGANAAYTAYHTRAVHRTBAARCARGTSACB

Table 5. Details of the conserved regions of ferritin gene

Region	Start - End	Conservation	Homozigosity	P value	Conserved Motif
1	158-246	0.593	0.805	0.04	SHASDKYBVCBGGSAARGGSAAGGAGGTGYTHASYGGSGTBRTNTTYSARCCMTTYGARGAGVTYAAGRRGGRWRAKCTYKCSCTCGTV
2	160-251	0.596	0.807	0.03	ASDKYBVCBGGSAARGGSAAGGAGGTGYTHASYGGSGTBRTNTTYSARCCMTTYGARGAGVTYAAGRRGGRWRAKCTYKCSCTCGTVYYCCM
3	165-255	0.591	0.793	0.04	BVCBGGSAARGGSAAGGAGGTGYTHASYGGSGTBRTNTTYSARCCMTTYGARGAGVTYAAGRRGGRWRAKCTYKCSCTCGTVYYCCMRNYV
4	169-256	0.600	0.803	0.03	GGSAARGGSAAGGAGGTGYTHASYGGSGTBRTNTTYSARCCMTTYGARGAGVTYAAGRRGGRWRAKCTYKCSCTCGTVYYCCMRNYVH
5	288-535	0.629	0.826	0.00	KYYGRYGANTGYGARBCYGYVMTYARYGARCAGATHAAYGTGGARTWCAAYRYHTCSTAYGYVTACCAYTCCHTNTTYGCMTACTTYGAYMGBGAYAACRTNGCTCTSAARGGAYTYGCCAARTTCTTYAARGAATCHAGYRAHGARGARAGRGADCAYGCWGARAARCTYATSRARTAYCARAACAHDCGTGGWGGVAGRGTDVKDCTYCAVYCBATYRWSAATGYDCCYTTRACHGARTTYGASCA
6	655-806	0.610	0.819	0.00	TGRCHSACTTCRTHGARAGYGARTTYYTBDVKGADCAGGTBRAAKMMATHAABAAKATMKCHRAGTAYGTBKCYCARYTGAGRAGRGTBGGMAARKGKCAYGGKGTKTGGCACTTYGAYCARADDCTDCTTSABGADGRRMATGCTSCYTRA

Table 6 (a). Protein domains associated with conserved regions of the Delta-1-pyrroline-5-carboxylate synthase gene (*P5CS*) gene

Conserved region	Protein domain
4	Laminin G domain protein
	Laminin subunit alpha-5
	Transmembrane emp24 domain-containing protein 5
	Laminin-like protein epi-1
	Paramyosin-like protein 1 (Fragment)
	BMA EPL 1, isoform v
	JmjC domain-containing protein
	transcription intermediary factor 1-alpha isoform X5
	transcription intermediary factor 1-alpha isoform X3
	transcription intermediary factor 1-alpha isoform X1
	DgyrCDS12560
	LAMAS protein (Fragment)
	DEAD/DEAH box helicase
	RING finger protein 207
	Zine finger Isd1 subclass family protein
5	Receptor protein serine/threonine kinase
9	V-set and immunoglobulin domain
10	Lipase_3 domain
11	Lipase_3 domain
12	Lipase_3 domain
	Dipeptidyl peptidase
14	Lipase_3 domain
15	2-hydroxy-acid oxidase Bromodomain

Table 6 (b). Protein domains associated with conserved regions of the Betaine aldehyde dehydrogenase (*BADH*) gene

Conserved region	Protein domain
1	RING-type domain
2	RING-type domain
3	Integrin beta
	VWFD domain-containing protem (Fragment)
	PAP2 superfamily protein
	X8 domain-containing protein (Fragment)
	Heterogeneous nuclear ribonucleo A2 B1-like protein
	CFEM domain-containing protein
	Phosphatase PAP2 family protein
	Dockerin
	Chaplin
	(S) ureidoglycine glyoxylate aminotransferase
	Type I secretion C-terminal target domain-containing protein
	EGF-like domain-containing protein
	Hint_2 domain-containing protein
	Uncharacterized protein
4	Phosphatase PAP2 family protein
	VWFD domain-containing protein (Fragment)
	Heterogeneous nuclear ribonucleo A2 B1-like protein
	X8 domain-containing protein (Fragment)
	EGF-like domain containing protein
	CFEM domain-containing protein
	Dockerin
	Chaplin
	(S)-ureidoglycine-glyoxylate aminotransferase
	PAP2 superfamily protein
	Type I secretion C-terminal target domain-containing protein
	Hint_2 domain-containing protein
	Uncharacterized protein
5	PAP2 superfamily protein
	VWFD domain-containing protein (Fragment)
	PAP2 superfamily protein
	X8 domain-containing protein (Fragment)
	Heterogeneous nuclear ribonucleo A2 B1-like protein
	CFEM domain-containing protein
	Phosphatase PAP2 family protein
	Dockerin
	Chaplin
	(S)-ureidoglycine-glyoxylate aminotransferase
	Type I secretion C-terminal target domain-containing protein
	EGF-like domain containing protein
	Hint_2 domain-containing protein
	MYB DNA binding protein (Tbf1), putative
	Uncharacterized protein
7	KDM2B demethylase

3.3 Phylogeny analysis

Phylogenetic trees of the three studied genes were reconstructed by the maximum likelihood method and the results recorded distinct clusters. Phylogenetic tree relationships among the sequences of the *P5CS* gene revealed three major clusters (Figure 1). The first cluster consisted of four entries (*Vigna radiata* (LOC106774140), *Cajanus cajan* (LOC109798886), *Cicer arietinum* (LOC101512568) and *Cajanus cajan* (LOC109800085)) and the second cluster consisted of three entries (*Vigna radiata* (LOC106757860), *Vigna radiata* (LOC106764975) and *Vigna angularis*

(LOC108338195)) while the rest of the studied CDS entries (*Sorghum bicolor* (LOC110433516), *Oryza sativa* (LOC4324853), *Oryza sativa* (LOC4338979), *Setaria italica* (LOC101765114), *Sorghum bicolor* (LOC110430201) and *Zea mays* (LOC103630430)) grouped into the third cluster.

Phylogenetic cluster analysis of the *BADH* gene revealed three distinct clusters (Figure 2). The CDS sequences of *Sorghum bicolor* (LOC8075627), *Zea mays* (LOC606443), *Setaria italica* (LOC101780729) and *Oryza sativa* (LOC4345606)) were grouped into the first cluster while that of *Oryza sativa* (LOC4336081), *Setaria italica* (LOC101753329), *Sorghum bicolor* (LOC8057487) and *Zea mays* (LOC541949) grouped into the second cluster. The third cluster comprised *Vigna radiata* (LOC106758570), *Cajanus cajan* (LOC109815177), *Cicer arietinum* (LOC101506136), *Vigna unguiculata* (LOC114179073), *Vigna radiata* (LOC106768797) and *Vigna angularis* (LOC108318717).

Only two major clusters were obtained for the cluster analysis of the ferritin gene (Figure 3). Those clusters consisted of six (*Setaria italica* (LOC101771422), *Setaria italica* (LOC101778084), *Sorghum bicolor* (LOC110437661), *Zea mays* (LOC542553), *Oryza sativa* (LOC4351264) and *Oryza sativa* (LOC9269178)) while five (*Cicer arietinum* (LOC101503152), *Cajanus cajan* (LOC109802075), *Vigna_unguiculata* (LOC114185823), *Vigna angularis* (LOC108329281) and *Vigna radiata* (LOC106758565)) formed the second cluster.

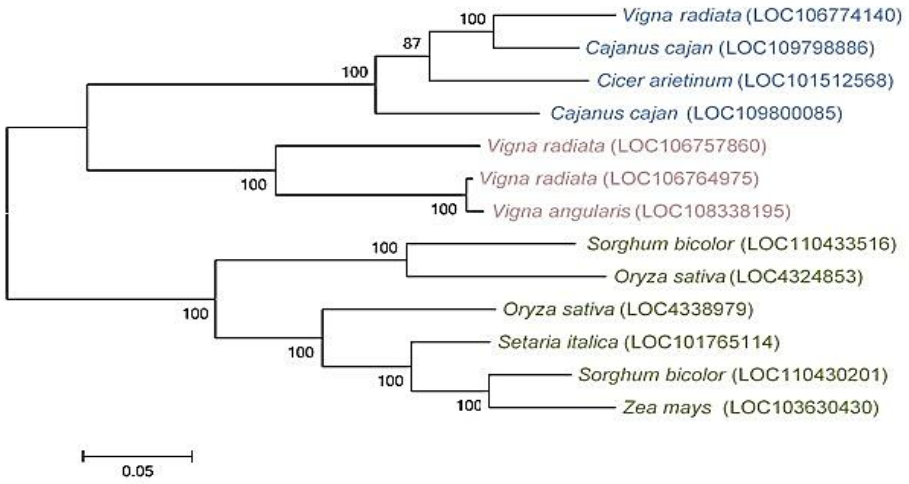

100
87
100
Vigna radiata (LOC106774140)
Cajanus cajan (LOC109798886)
Cicer arietinum (LOC101512568)
Cajanus cajan (LOC109800085)
Vigna radiata (LOC106757860)
100
Vigna radiata (LOC106764975)
100
Vigna angularis (LOC108338195)
100
Sorghum bicolor (LOC110433516)
Oryza sativa (LOC4324853)
100
Oryza sativa (LOC4338979)
100
Setaria italica (LOC101765114)
100
Sorghum bicolor (LOC110430201)
100
Zea mays (LOC103630430)
0.05

Figure 1. Maximum likelihood Tree of Delta-1-pyrroline-5-carboxylate synthase gene, with the percentage bootstrap support from 100 replications above the branches

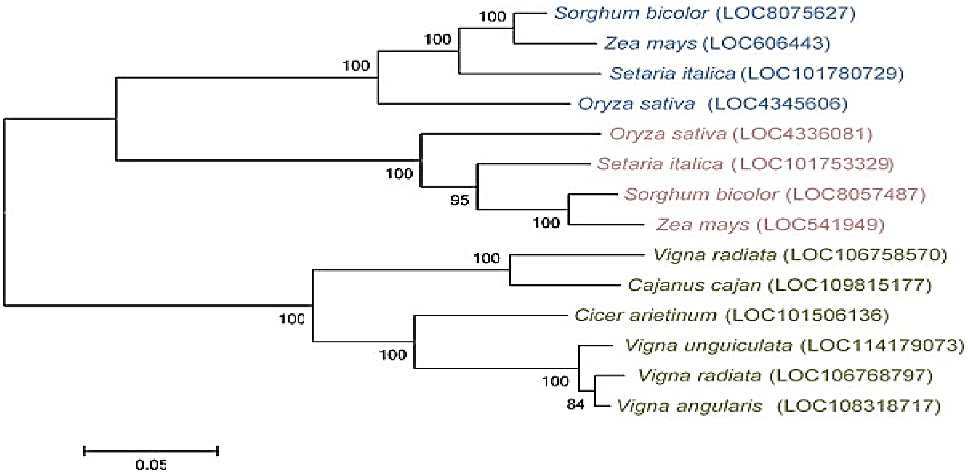

Sorghum bicolor (LOC8075627)
Zea mays (LOC606443)
Setaria italica (LOC101780729)
Oryza sativa (LOC4345606)
Oryza sativa (LOC4336081)
Setaria italica (LOC101753329)
Sorghum bicolor (LOC8057487)
Zea mays (LOC541949)
Vigna radiata (LOC106758570)
Cajanus cajan (LOC109815177)
Cicer arietinum (LOC101506136)
Vigna unguiculata (LOC114179073)
Vigna radiata (LOC106768797)
Vigna angularis (LOC108318717)
100
100
100
100
95
100
100
100
100
100
84
0.05

Figure 2. Maximum likelihood tree of Betaine aldehyde dehydrogenase gene, with the percentage bootstrap support from 100 replications above the branches.

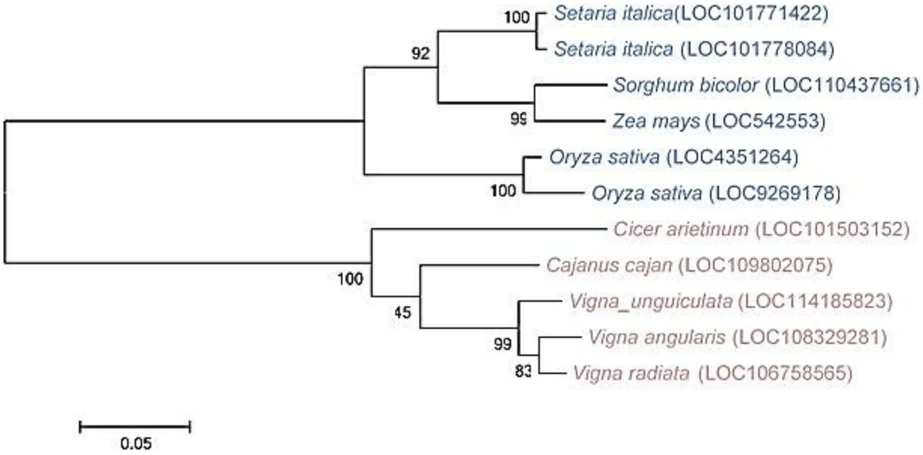

Setaria italica(LOC101771422)
Setaria italica (LOC101778084)
Sorghum bicolor (LOC110437661)
Zea mays (LOC542553)
Oryza sativa (LOC4351264)
Oryza sativa (LOC9269178)
Cicer arietinum (LOC101503152)
Cajanus cajan (LOC109802075)
Vigna_unguiculata (LOC114185823)
Vigna angularis (LOC108329281)
Vigna radiata (LOC106758565)
100
92
99
100
100
45
99
83
0.05

Figure 3. Maximum likelihood tree of ferritin gene, with the percentage bootstrap support from 100 replications above the branches

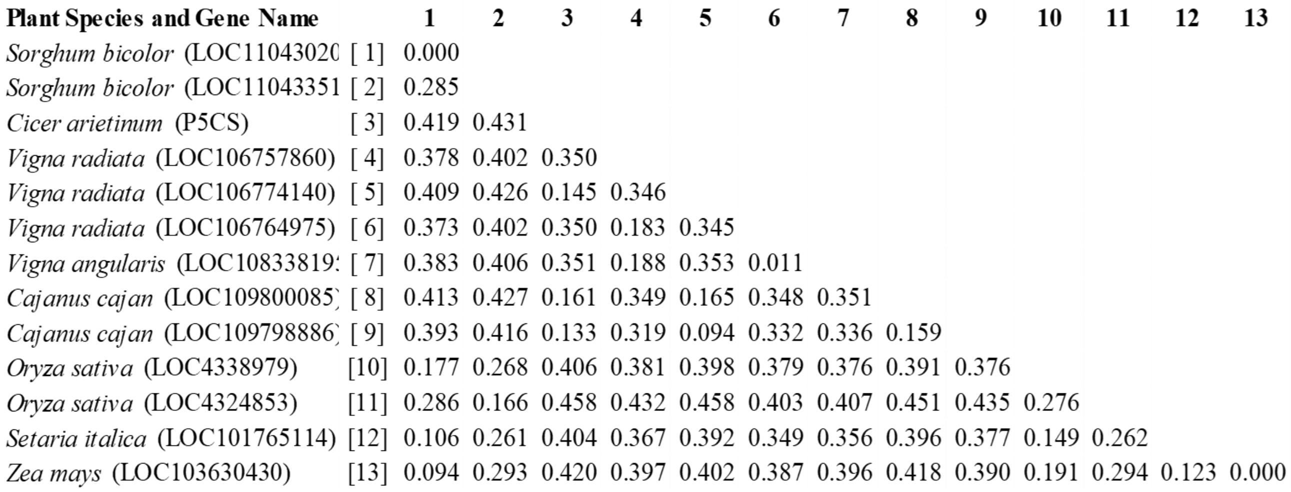
3.4 Evolutionary Divergence

A significant evolutionary divergence existed among *P5CS* gene coding sequences (Table 7). The lowest divergence (0.011) was observed between the coding sequences of *Vigna angularis* and *Vigna radiata* while *Oryza sativa* and *Cicer arietinum* had the highest divergence (0.458).

Concerning coding sequences of the *BADH* gene (Table 8), the lowest divergence was 0.017 (between *Vigna angularis* and *Vigna radiata*) while 0.406 was the highest divergence (between *Oryza sativa* and *Vigna radiata*).

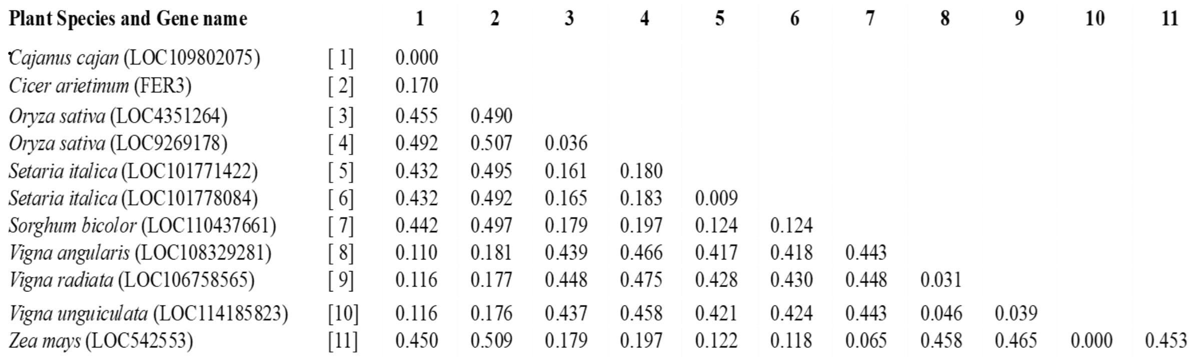
The two ferritin gene coding sequences (Table 9) of *Setaria italica* (gene LOC101771422 and gene LOC101778084) showed the lowest divergence of 0.009 whereas the highest divergence value of 0.509 was observed between ferritin gene coding sequences of *Zea mays* and *Cicer arietinum.*

Table 7. Estimates of the evolutionary divergence between the coding sequences of Delta-1-pyrroline-5-carboxylate synthase gene



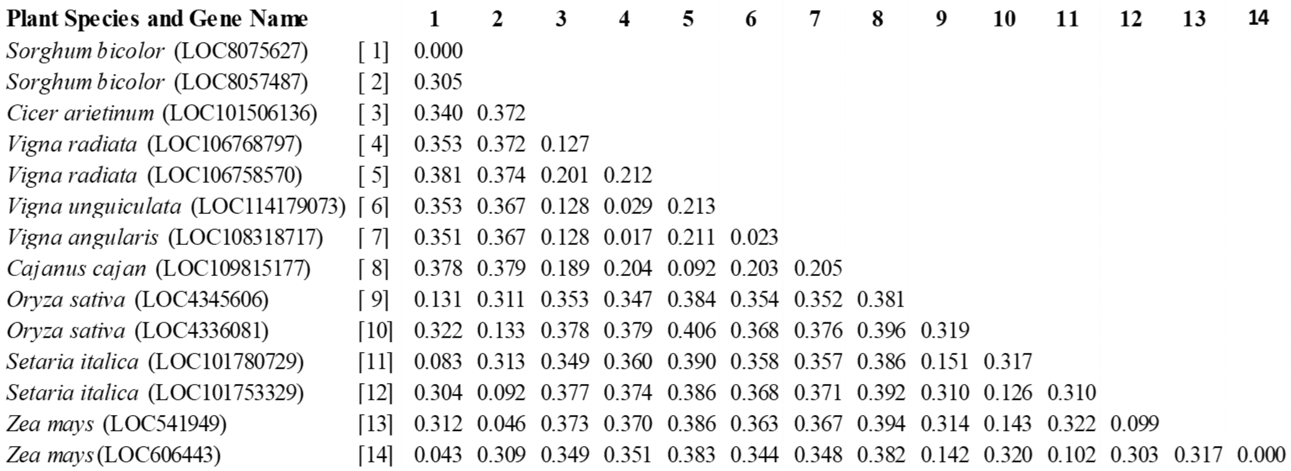
Plant Species and Gene Name		1	2	3	4	5	6	7	8	9	10	11	12	13
Sorghum bicolor (LOC11043020	[1]	0.000												
Sorghum bicolor (LOC11043351	[2]	0.285												
Cicer arietinum (P5CS)	[3]	0.419	0.431											
Vigna radiata (LOC106757860)	[4]	0.378	0.402	0.350										
Vigna radiata (LOC106774140)	[5]	0.409	0.426	0.145	0.346									
Vigna radiata (LOC106764975)	[6]	0.373	0.402	0.350	0.183	0.345								
Vigna angularis (LOC10833819:	[7]	0.383	0.406	0.351	0.188	0.353	0.011							
Cajanus cajan (LOC109800085`	[8]	0.413	0.427	0.161	0.349	0.165	0.348	0.351						
Cajanus cajan (LOC109798886`	[9]	0.393	0.416	0.133	0.319	0.094	0.332	0.336	0.159					
Oryza sativa (LOC4338979)	[10]	0.177	0.268	0.406	0.381	0.398	0.379	0.376	0.391	0.376				
Oryza sativa (LOC4324853)	[11]	0.286	0.166	0.458	0.432	0.458	0.403	0.407	0.451	0.435	0.276			
Setaria italica (LOC101765114)	[12]	0.106	0.261	0.404	0.367	0.392	0.349	0.356	0.396	0.377	0.149	0.262		
Zea mays (LOC103630430)	[13]	0.094	0.293	0.420	0.397	0.402	0.387	0.396	0.418	0.390	0.191	0.294	0.123	0.000

Table 8. Estimates of the evolutionary divergence between the coding sequences of Betaine aldehyde dehydrogenase gene



Plant Species and Gene name		1	2	3	4	5	6	7	8	9	10	11
Cajanus cajan (LOC109802075)	[1]	0.000										
Cicer arietinum (FER3)	[2]	0.170										
Oryza sativa (LOC4351264)	[3]	0.455	0.490									
Oryza sativa (LOC9269178)	[4]	0.492	0.507	0.036								
Setaria italica (LOC101771422)	[5]	0.432	0.495	0.161	0.180							
Setaria italica (LOC101778084)	[6]	0.432	0.492	0.165	0.183	0.009						
Sorghum bicolor (LOC110437661)	[7]	0.442	0.497	0.179	0.197	0.124	0.124					
Vigna angularis (LOC108329281)	[8]	0.110	0.181	0.439	0.466	0.417	0.418	0.443				
Vigna radiata (LOC106758565)	[9]	0.116	0.177	0.448	0.475	0.428	0.430	0.448	0.031			
Vigna unguiculata (LOC114185823)	[10]	0.116	0.176	0.437	0.458	0.421	0.424	0.443	0.046	0.039		
Zea mays (LOC542553)	[11]	0.450	0.509	0.179	0.197	0.122	0.118	0.065	0.458	0.465	0.000	0.453

Table 9. Estimates of the evolutionary divergence between the coding sequences of the ferritin gene



Plant Species and Gene Name		1	2	3	4	5	6	7	8	9	10	11	12	13	14
Sorghum bicolor (LOC8075627)	[1]	0.000													
Sorghum bicolor (LOC8057487)	[2]	0.305													
Cicer arietinum (LOC101506136)	[3]	0.340	0.372												
Vigna radiata (LOC106768797)	[4]	0.353	0.372	0.127											
Vigna radiata (LOC106758570)	[5]	0.381	0.374	0.201	0.212										
Vigna unguiculata (LOC114179073)	[6]	0.353	0.367	0.128	0.029	0.213									
Vigna angularis (LOC108318717)	[7]	0.351	0.367	0.128	0.017	0.211	0.023								
Cajanus cajan (LOC109815177)	[8]	0.378	0.379	0.189	0.204	0.092	0.203	0.205							
Oryza sativa (LOC4345606)	[9]	0.131	0.311	0.353	0.347	0.384	0.354	0.352	0.381						
Oryza sativa (LOC4336081)	[10]	0.322	0.133	0.378	0.379	0.406	0.368	0.376	0.396	0.319					
Setaria italica (LOC101780729)	[11]	0.083	0.313	0.349	0.360	0.390	0.358	0.357	0.386	0.151	0.317				
Setaria italica (LOC101753329)	[12]	0.304	0.092	0.377	0.374	0.386	0.368	0.371	0.392	0.310	0.126	0.310			
Zea mays (LOC541949)	[13]	0.312	0.046	0.373	0.370	0.386	0.363	0.367	0.394	0.314	0.143	0.322	0.099		
Zea mays (LOC606443)	[14]	0.043	0.309	0.349	0.351	0.383	0.344	0.348	0.382	0.142	0.320	0.102	0.303	0.317	0.000

4 DISCUSSION

Molecular characterization of the genes or gene families helps in understanding their functions and species relationships. The availability of completely sequenced plant genes and their orthologs has enabled researchers to study the molecular evolution of genes.

The *P5CS* gene encodes a bifunctional enzyme that catalyzes the rate-limiting reaction in proline biosynthesis in living organisms. Proline has a wide range of multifunctional roles in stress defence. Out of all proline biosynthetic genes, especially, P5CS is commonly used in metabolic engineering for inducing proline overproduction in plants in order to confer stress tolerance (Rai and Penna, 2013).

In some plant species, two copies of the P5CS gene (P5CS1 and P5CS2) with distinct functions have been recognized. These two forms (P5CS1 and P5CS2) show varying temporal and spatial expression patterns, while P5CS2 records a more predominantly cytoplasmic localization than P5CS1 (Verdoy et al., 2006; Gruszka et al., 2007; Sze´kely et al., 2008). In such cases, only P5CS1 which has more chloroplastic localization was retrieved for the study.

Observation of statistically significant 16 conserved regions with significantly high conservation and homozygosity values indicated the possible functional value of the sequences. This study revealed protein domains to nine conserved regions of the P5CS gene.

Rai and Penna (2013) reported that most of the key enzymes of metabolic pathways are generally encoded by redundant genes, which may be generated by gene duplication events during evolution. There is evidence suggesting independent duplication events in the P5CS gene (Zhang, 2003; Turchetto-Zolet et al., 2009). Interspecific phylogenetic trees of full-length cDNA sequence of P5CS shows the existence of two groups, separating the P5CS gene of monocot from that of dicots (Rai and Penna, 2013).

In this study, all the monocot entries (*Sorghum bicolor*, *Oryza sativa*, *Setaria italica* and *Zea mays*) were grouped into one distinct cluster when comparing the coding sequences of P5CS genes found in different species. Though dicot entries were separated into two groups, coding sequences of *Vigna radiata* were observed in both groups (LOC106774140 in the first group; LOC106757860 and LOC106764975 in the second group). However, the cluster that comprised monocots only was further bifurcated into two internal branches. In that case, coding sequences of *Sorghum bicolor* and *Oryza sativa* (LOC110433516 and LOC4324853) were observed to be phylogenetically more similar to each other compared to their other studied coding sequence (LOC110430201 of *Sorghum bicolor* and LOC4338979 of *Oryza sativa*).

The significant evolutionary divergence that existed among *P5CS* gene coding sequences of the studied plants also indicated the uniqueness of each entry. Moreover, significant divergence was observed between the different *P5CS* gene coding sequences of the same species as 0.285 for *Sorghum bicolor* (LOC110430201 - LOC110433516), 0.346, 0.183 and 0.345 for *Vigna radiata* (LOC106757860 - LOC106774140, LOC106757860 - LOC106764975 and LOC106774140 - LOC106764975, respectively) and 0.276 for *Oryza sativa* (LOC4338979 - LOC4324853).

BADH catalyzes the last step in the synthesis of the osmoprotectant glycine betaine from choline in higher plants. However, transgenic plants with glycine betaine synthesizing genes could accumulate lower levels of glycine betaine than natural accumulators, still enhancing the tolerance to various abiotic stresses (Khan et al., 2019).

Seven statistically significant conserved regions of BADH gene coding sequences with significantly high conservation and homozygosity values indicated that these sequences may have a functional value. Protein domains to six conserved regions of the BADH gene were revealed in the present study. A complete BADH sequence homology comparison between Ophiopogon japonicus and other reported species was conducted by Liu et al. (2010) and the result reported a high homology in Ophiopogon japonicus' BADH gene and that of the Chenopodiaceae plant.

BADH coding sequences of dicots (*Vigna radiata*, *Cajanus cajan*, *Cicer arietinum*, *Vigna unguiculate* and *Vigna angularis*) grouped into one distinct cluster. Monocots were clustered into two groups as coding sequences of the same species clustered separately (The first cluster - *Sorghum bicolor* (LOC8075627), *Zea mays* (LOC606443), *Setaria italica* (LOC101780729) and *Oryza sativa (LOC4345606)); the second cluster - Oryza sativa* (LOC4336081), *Setaria italica* (LOC101753329), *Sorghum bicolor* (LOC8057487) and *Zea mays* (LOC541949)).

Not only a significant evolutionary divergence observed among BADH gene coding sequences of the studied plants but also the same species coding sequences of *Sorghum bicolor* (LOC8075627 - LOC8057487), *Vigna radiata* (LOC106768797 - LOC106758570), *Oryza sativa* (LOC4345606 - LOC4336081), *Setaria italica* (LOC101780729 - LOC101753329) and *Zea mays* (LOC541949 - LOC606443) showed a significant divergence (0.305, 0.212, 0.319, 0.310 and 0.317, respectively) from each other.

Ferritin is an iron-storage protein which comprises 24 homologous or heterologous subunits. Goto et al. (1999) explored the possibility of introducing the soybean ferritin gene into rice plants by Agrobacterium-mediated transformation. The results indicated that the soybean ferritin could be highly expressed and accumulated in the endosperm tissue of heterogeneous rice plants with the transgenic seeds storing up to three times more iron than the normal seeds.

In the present study, the observation of six statistically significant conserved regions of ferritin coding sequences with significantly high conservation and homozygosity values indicates that these sequences may have a functional value.

When the phylogeny reconstruction was performed with the studied ferritin coding sequences, dicots and monocots were separated into two clusters.

However, two ferritin gene coding sequences of *Oryza sativa* (LOC4351264 - LOC9269178) were observed to be significantly divergent from each other (0.036).

5 CONCLUSION AND RECOMMENDATIONS

Highly conserved DNA sequences are considered to have functional values. The results of this study revealed several regions that have been conserved throughout the evolutionary process in the three genes studied. On the other hand, most of the key enzymes of the metabolic pathways are generally encoded by redundant genes, which may be generated by gene duplication events during evolution. Functional redundancy due to gene duplications is a typical feature of many biological systems and phylogeny reconstruction of the studied three genes showed the existence of two groups, separating monocot from dicot plants. As even transgenic plants with newly introduced genes could accumulate lower levels of desired products than natural accumulators, having better knowledge about the conserved region of a gene throughout the evolutionary process and the evolutionary divergence of genes among different plant species and families to the coding sequences is vitally important for the crop genetic engineering processes and such information are uncovered in this study.

ACKNOWLEDGMENT

The authors would like to acknowledge the University research council, University of Sri Jayewardenepura for the financial assistance (Research grant no: ASP/1/RE/SCI/2017/07, ASP/1/RE/SCI/2021/08) and the Department of Botany, Faculty of Applied Sciences, University of Sri Jayewardenepura.

REFERENCES

Alcamo, J, Leemans, R & Kreileman, E 1999, Global Change Scenarios of the 21st Century, ‘*Results from the Image 2.1 Model’*, Elsevier Science, The Netherlands.

Annunziata, MG, Ciarmiello, LF, Woodrow, P, Dell’Aversana, E & Carillo, P 2019, ‘Spatial and Temporal Profile of Glycine Betaine Accumulation in Plants Under Abiotic Stresses’, *Frontiers in Plant Sciences*. vol.10, pp.230-233.

Bhatta, M, Morgounov, A, Belamkar, V & Baenziger, PS 2018, ‘Genome-Wide Association Study Reveals Novel Genomic Regions for Grain Yield and Yield-Related Traits in Drought-Stressed Synthetic Hexaploid Wheat’, *International Journal of Molecular Sciences*, vol.19, pp. 3011-3014.

Borg, S, Brinch-Pedersen, H, Tauris, B, Madsen, LH, Darbani, B & Noeparvar, S 2012, ‘Wheat ferritins: Improving the iron content of the wheat grain’, *Journal of Cereal Science*, vol.56, pp.204–213.

Cosgrove, WJ & FR Rijsberman 2000, *World Water Vision: Making Water Everybody's Business*, London: Earthscan Publications.

Demirkol, G 2020, 'The role of *BADH* gene in oxidative, salt, and drought stress tolerances of white clover', *Turkish Journal of Botany*, vol. 44, pp.214-221.

Giri, J 2011, 'Glycinebetaine and abiotic stress tolerance in plants', *Plant Signaling and Behavior*, vol.6, pp.1746–1751.

Gosal, SS, Wani, SS & Kang, MS 2009, 'Biotechnology and drought tolerance', *Journal of Crop Improvement*, vol.23, pp.19–54.

Goto, F, Yoshihara, T, Shigemoto, N, Toki, S & Takaiwa, F 1999, 'Iron fortification of rice seed by the soybean ferritin gene', *Nature biotechnology*, vol.17, pp.282- 286.

Gruszka, VEC, Schuster, I, Pileggi, M, Scapim, CA, Marur, CJ & Vieira, LG 2007, 'Stress induced synthesis of proline confers tolerance to water deficit in transgenic wheat', *Journal of Plant Physiology*, vol.164, pp.1367–1376.

Gupta, PP, Balyan, SS & Gahlaut, V 2017, 'QTL analysis for drought tolerance in wheat: Present status and future possibilities', *Agronomy*, vol.7, pp.5.

Hayat, S, Hayat, Q, Alyemeni, MN, Wani, AS, Pichtel, J & Ahmad, A 2012, 'Role of proline under changing environments: A review', *Plant Signaling and Behavior*, vol.7, pp.1456–1466.

He, C, Zhang, W, Gao, Q, Yang, A, Hu, X & Zhang, J 2011, 'Enhancement of drought resistance and biomass by increasing the amount of glycinebetaine in wheat seedlings' *Euphytica*, vol.177, pp.151–167.

Hmida-Sayari, A, Gargouri-Bouzid, R, Bidani, A, Jaoua, L, Savoure, A & Jaoua, S 2005, 'Overexpression of Δ1-pyrroline-5-carboxylate synthetase increases proline production and confers salt tolerance in transgenic potato plants', Plant Science, vol.169, pp.746–752.

Hu, H & Xiong, L 2014, 'Genetic engineering and breeding of drought-resistant crops', *Annual Review of Plant Biology*, vol.65, pp.715–741.

Khan, S, Anwar, S, Yu, S, Sun, M, Yang, Z & Gao, Z 2019, 'Review - Development of Drought-Tolerant Transgenic Wheat: Achievements and Limitations', *International Journal of Molecular Sciences*, vol.20, pp.3350.

Librado, P & Rozas, J 2009, 'DnaSP v5: A software for comprehensive analysis of DNA polymorphism data', *Bioinformatics,* vol.25, pp.1451-1452.

Liu, J, Zenga, H, Lia, Z, Xua, L, Wanga, Y, Tang, W & Han, L 2010, 'Isolation and Characterization of the Betaine Aldehyde Dehydrogenase Gene in *Ophiopogon japonicus', The Open Biotechnology Journal*, vol.4, pp.18-25.

Mwadzingeni, L, Shimelis, H, Dube, E, Laing, MD & Tsilo, TJ 2016, 'Breeding wheat for drought tolerance: Progress and technologies', *Journal of Integrative Agriculture*, vol.15, pp.935–943.

Przewieslik-Allen, AM, Burridge, AA, Wilkinson, PP, Winfield, MM, Shaw, DD, McAusland, L, King, J, King, II, Edwards, KK & Barker, GLA 2019, 'Developing a High-Throughput SNP-Based Marker System to Facilitate the Introgression of Traits from *Aegilops* Species into Bread Wheat (*Triticum aestivum*), *Frontiers in Plant Science*, vol.9, pp.1993.

Rai, A & Penna, S 2013, 'Molecular evolution of plant P5CS gene involved in proline biosynthesis', *Molecular Biology Reports*, vol.40, pp.6429–6435.

Sawahel, WW & Hassan, AH 2002, 'Generation of transgenic wheat plants producing high levels of the osmoprotectant proline', *Biotechnology Letters*, vol.24, pp.721–725.

Senapati, N, Stratonovitch, P, Paul, MJ & Semenov, MA 2018, 'Drought tolerance during reproductive development is important for increasing wheat yield potential under climate change in Europe', *Journal of Experimental Botany*, vol.70, pp.2549–2560.

Szekely, G, Abraha´m, E, Cseplo, A, Rigo, G, Zsigmond, L, Csiszar, J, Ayaydin, F, Strizhov, N, Jasik, J, Schmelzer, E, Koncz, C & Szabados, L 2008, 'Duplicated *P5CS* genes of Arabidopsis play distinct roles in stress regulation and developmental control of proline biosynthesis', *Plant Journal*, vol.53, pp.11–28.

Tamura K, Nei M, & Kumar S 2004, 'Prospects for inferring very large phylogenies by using the neighbor-joining method', *Proceedings of the National Academy of Sciences (USA)*, vol.101, pp.11030-11035.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, & Kumar S 2011, 'MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods', *Molecular Biology and Evolution,* vol.28, pp.2731-2739.

Turchetto-Zolet, AN, Margis-Pinheiro, M & Margis, R 2009, 'The evolution of pyrroline-5-carboxylate synthase in plants: a key enzyme in proline synthesis', *Molecular Genetics and Genomics, vol.*281, pp.87–97.

Vendruscolo, ECG, Schuster, I, Pileggi, M, Scapim, CC, Molinari, HBC, Marur, CC & Vieira, LGE 2007, 'Stress-induced synthesis of proline confers tolerance to water deficit in transgenic wheat', *Journal of Plant Physiology*, vol.164, pp.1367–1376.

Verdoy, D, De la Pena, TC, Redondo, FJ, Lucas, MM & Pueyo, JJ 2006, Transgenic *Medicago truncatula* plants that accumulate proline display nitrogen-fixing activity with enhanced tolerance to osmotic stress Plant', *Cell Environment*, vol.29, pp.1913–1923.

Wang, GG, Hui, Z, Li, F, Zhao, MM, Zhang, J & Wang, W 2010, 'Improvement of heat and drought photosynthetic tolerance in wheat by overaccumulation of glycinebetaine', *Plant Biotechnology Reports*, vol.4, pp.213–222.

Zang, X, Geng, X, Wang, F, Liu, Z, Zhang, L, Zhao, Y, Tian, X, Ni, Z, Yao, Y & Xin, M 2017, ‘Overexpression of wheat ferritin gene TaFER-5B enhances tolerance to heat stress and other abiotic stresses associated with the ROS scavenging’, *BMC Plant Biology*, vol.17, pp.14.

Zhang, J 2003, ‘Evolution by gene duplication: an update’, *Trends in Ecology and Evolution*, vol.18, pp.292–298.

Zhang, CC, Lu, Q & Verma, DPS 1995, ‘Removal of feedback inhibition of Δ1-pyrroline-5-carboxylate synthetase, a bifunctional enzyme catalyzing the first two steps of proline biosynthesis in plants’, *Journal of Biological Chemistry*, vol.270, pp.20491–20496.