Journal of Pharmacognosy and Phytochemistry

Available online at www.phytojournal.com

E-ISSN: 2278-4136
P-ISSN: 2349-8234
JPP 2019; 8(2): 524-530
Received: 26-01-2019
Accepted: 27-02-2019

Suman Dutta
Division of Genetics, ICAR-Indian Agricultural Research Institute, New Delhi, India

Vignesh Muthusamy
Division of Genetics, ICAR-Indian Agricultural Research Institute, New Delhi, India

Rajkumar U Zunjare
Division of Genetics, ICAR-Indian Agricultural Research Institute, New Delhi, India

Firoz Hossain
Division of Genetics, ICAR-Indian Agricultural Research Institute, New Delhi, India

Correspondence
Vignesh Muthusamy
Division of Genetics, ICAR-Indian Agricultural Research Institute, New Delhi, India

Analysis of paralogous genes of *Carotenoid dioxygenase* affecting carotenoid biosynthesis pathway in maize (*Zea mays* L.)

Suman Dutta, Vignesh Muthusamy, Rajkumar U Zunjare and Firoz Hossain

Abstract
Carotenoid dioxygenase contain a group of *carotenoid cleavage dioxygenase* (*CCD*) and *9-cis carotenoid cleavage dioxygenase* (*NCED*) gene(s) playing a vital role in plant kingdom. Family of these genes was involved in apocarotenoid biosynthesis from the precursor carotenoid compounds. In this study, a comprehensive analysis was carried out to find out the paralogous gene of *carotenoids dioxygenase* in maize. Twenty four proteins were found in maize genome with diverse intron exon organization. Out of the 24 genes, 21 genes were found to be of *NCED* type and three genes were found to be of *CCD* type. The proteins were performing in wide range of pH and localized in different organelles indicating their diverse role in maize. All the proteins contain retinal pigment epithelial (RPE) domain which signify their similar kind of function. The proteins were conserved in other plant species as well and may have an evolutionary significance. Analyses of these paralogous genes open the door for deciphering their role in plant development.

Keywords: *Carotenoids dioxygenase*, *CCDs*, *NCEDs*, carotenoids, apocarotenoids, degradation

1. Introduction
Carotenoids are large spectrums of compounds playing a diverse role in plant-morphogenesis and development. Being the second most naturally abundant pigment on earth, they also act as an accessory pigment during photosynthesis, photo-protection and are precursors of apo-carotenoids like strigolactone, abscisic acid (ABA) and some volatile compounds which are important for plant physiological functions (Vallabhaneni *et al.* 2010) [1]. This apocarotenoids are mainly produced by oxidative damage of the carotenoids mediated by *Carotenoid dioxygenases*. *Carotenoids dioxygenase* genes were categorized in two broad types in plants as (a) *carotenoid cleavage dioxygenase* (*CCD*) and (b) *9-cis carotenoids cleavage dioxygenase* (*NCED*) (Hou *et al*. 2016) [2]. In *Arabidopsis* genome, a total of nine *carotenoids dioxygenases* have been reported that include five *NCEDs* and four *CCDs* which showed different substrate specificity (Tan *et al.* 2003) [3]. *Viviparous14* (*vp14*) was the first mutant identified in maize which was later cloned as *NCED1* gene (Tan *et al.* 1997) [4]. NCED enzyme is generally cleaved at 11, 12 double bond position of 9-cis violaxanthin or 9-cis neoxanthin which is the first dedicated step in ABA biosynthesis (Schwartz *et al.* 1997) [5]. CCD1 enzyme is non-heme protein which requires Fe^{2+} cofactor and has broad substrate specificity on the carotenoid compounds with cleavage property at two different positions (Vogel *et al.* 2008) [6]. *CCD1* gene was cloned in maize by RT-PCR and characterized with GC-MS study by expressing in the carotenoids accumulating *E. coli* cell (Sun *et al.* 2008) [7]. White cap formation in maize also reported to be associated with *CCD1* gene (Tan *et al.* 2017) [7] and their copy number varies from genotype to genotype in maize and is directly correlated with lesser retention of carotenoids (Vallabhaneni *et al.* 2010) [1]. So far, *lycopene-ε-cyclase* (*lcyE*) and *β-carotene hydroxylase* (*crtRB1*) has been the two major target genes for provitamin-A (proA) enrichment in maize (Vignesh *et al*. 2012; Muthusamy *et al*. 2014; Zunjare *et al*. 2018) [9, 10, 11]; and these CCDs would be future target in the proA biofortification programmes. Circadian regulation also play important role for expression of the *PhCCD1*gene and produce fragrance of β-ionone, a volatile compound in petunia flowers (Simkin *et al*. 2004) [12]. Novel *carotenoid cleavage dioxygenase* (*CCD2*) catalyzes the first dedicated step in saffron crocin biosynthesis (Frusciantea *et al*. 2014) [13]. Expression level of the potato *CCD4* gene was down-regulated using an RNA interference (RNAi) approach in stable transgenic lines. Down-regulation in tubers resulted in an increased carotenoid content, 2- to 5-fold higher than in control plants.

The increase in carotenoid content was mainly due to elevated violaxanthin content (Campbell *et al.* 2010) [14]. β-carotene was converted into to carlactone, a strigolactone like plant hormone compound, by the joint action of *D27*,*CCD7* and *CCD8* gene in plants (Alder *et al*. 2012) [15]. Apocarotenoids compounds are also produced from the precursor compounds with action of the *CCDs* genes through different environmental stimuli (Hou *et al.* 2016) [2]. *CCDs* genes have a significant role in plant developmental pathway and also play detrimental role as they degrade carotenoids compounds which have proA activity in animal beings. In the present study, a comprehensive study was carried out to find out the paralogous sequences of the *CCDs* genes in maize. Physicochemical properties of the protein were analyzed for better understanding of this gene family. Maize *CCDs* genes were further analyzed to predict their subcellular localization, biochemical properties, intron-exon organization and their expression. This study will enhance the understanding of the evolution and function of the *CCDs* superfamily in maize. The evolutionary and structural divergence analysis of *CCDs* superfamily in maize presented here provides useful information for further probing the molecular mechanism by which the *CCDs* superfamily contributes to the apocarotenoids biosynthesis.

2. Materials and Methods

2.1 Database search for *Carotenoid dioxygenase* genes in maize genome

Ensembl Plants (https://plants.ensembl.org/index.html; Zerbino *et al*. 2017) [16] databases were searched against maize genome to find out *CCDs* gene or putative *CCD* like genes with oxygenase activity that can cleave carotenoid compounds. Physical Location of all the genes on a particular chromosome was taken from B73 maize reference genome (http://plants.ensembl.org). Unique identification numbers of all genes taken from Gramene database (www.gramene.org; Tello-Ruiz *et al*. 2016) [17]. Proteins encoded by these genes were taken from Universal Protein resource database (www.uniprot.org; UniProt Consortium. 2016) and ENA ID of all the genes were taken from European Nucleotide Archive database (https://www.ebi.ac.uk/ena; Silvester *et al*. 2017) [18].

2.2 Gene structure analysis of *Carotenoids dioxygenase*

The exon/intron organization of five gene families was generated online using Gene structure display server (GSDS; http://gsds.cbi.pku.edu.cn/; Hu *et al.*, 2014) [13] by alignment of the CDS with their corresponding genomic DNA sequences. Sequences of the twenty four genes along with their corresponding coding sequence (CDS) were taken from Ensembl Plants database. Intron phase was included during intron-exon boundary organization.

2.3 Physicochemical property of carotenoids dioxygenase proteins

Conserved domains of the proteins were further confirmed by Pfam (https://pfam.xfam.org/; Finn *et al*. 2015) [20] through HMMER using hidden Markov Models and conserved Pfam motif from start amino acid to last amino acid were listed. The ProtParam tool (https://web.expasy.org/protparam; Gasteiger *et al*. 2005) [21] was used to determine the physicochemical parameters [amino acid length, molecular weight of the protein, iso-electric point of the protein and grand average of hydropathicity (GRAVY)] of *CCDs* gene in maize and the CELLO v2.5 (cello.life.nctu.edu.tw; Yu *et al*. 2006) [22] server was to determine sub-cellular localization of the proteins.

2.4 Analysis of microRNA binding sites in *carotenoids dioxygenase* genes

Carotenoids dioxygenase genes were also analyzed for their regulational level. For analysis of micro-RNA binding site in the CDS (sequences are taken from Ensembl Plants database) regions of the *Carotenoids dioxygenase* gene, updated version of psRNA Target server (plantgrn.noble.org/psRNATarget/; Dai *et al.* 2011) [23] was used for miRNA binding site with the parameter of : expectation level-5, maximum energy to unpair target site (UPE) was 25, flanking length around target site for target accessibility are 17 upstream and 13 downstream nucleotides of target site, translation inhibition range from 10^{th} to 11^{th} nucleotide.

2.5 Phylogenetic analysis of orthologs sequence *CCD1* gene in crop species

Ten orthologous *CCDs* genes from ten different species [*Brachypodium distachyon* (*BRADI4G00335*), *Setaria italica* (*Si021680m*.g), *Oryza sativa* var. *indica* (*BGIOSGA037874*), *Arabidopsis thaliana* (*AT3G63520*), *Triticum aestivum* (TRIAE_*CS42*_U_TGACv1_643498_AA2132620), *Medicago truncatula* (MTR_*8g037275*), *Brassica oleracea* (*Bo8g099460*), *Helianthus annuus* (HannXRQ_*Chr17g0562991*), *Aegilops tauschii* (*F775*_00763), *Zea mays* (*Zm00001d048373*)] were taken for evolutionary study and exon-intron boundary construction using PIECE (Plant Intron Exon Comparison and Evolution) database (https://wheat.pw.usda.gov/piece; Wang *et al.* 2012) [24]. Conserved domains of the proteins were further confirmed by Pfam (https://pfam.xfam.org/; Finn *et al.* 2015) [20] database through HMMER using hidden Markov Models.

3. Results and Discussion

3.1 *Carotenoid dioxygenases* superfamily genes and their phylogenetic analysis

A total of 24 genes were found in maize genome from Ensembl Plants database. Out of the 24 genes, 21 genes were found to be of *NCED* type and three genes were found to be of *CCD* type. Five genes (*Zm00001d027592*; *Zm00001d027690*; *Zm00001d031086*, *GRMZM2G407181*; *Zm00001d033222*, *GRMZM2G014392*; *Zm00001d033377*) are found in chromosome 1. Only *NCED4 (Zm00001d007876*, *GRMZM2G408158)* is located on chromosome 2. Two genes (*Zm00001d041319*, *GRMZM5G858784*; *Zm00001d042076*) were found on chromosome 3. *NCED6* (*Zm00001d051556*, *GRMZM2G110192*) is found on chromosome 4. *NCED8* (*Zm00001d017766*, *GRMZM2G150363*) and *NCED9* (*Zm00001d013689*, *GRMZM5G838285*) are present on chromosome 5. Only one gene (*Zm00001d035724*) was found on chromosome 6. Four genes (*Zm00001d018819*, *GRMZM2G417954*; *Zm00001d019813*; *Zm00001d020069*, *GRMZM2G164967*; *Zm00001d022623*, *GRMZM2G330848*) were found on chromosome 7. Two genes were found on chromosome 8 (*Zm00001d008638*; *Zm00001d009286*), chromosome 9 (*Zm00001d045232*; *Zm00001d048373*, *GRMZM2G057243*) and chromosome 10 (*Zm00001d023458*; *Zm00001d023690*). A total of 13 genes with *NCED3* nomenclature were found on six different chromosomes (chromosome 1, 3, 6, 7, 8 and 10). *NCED2*, *NCED4*, *NCED5*, *NCED6*, *NCED7*, *NCED8* and *NCED9* were present in a single copy in the maize genome (Table 1). *viviparous14*

(Vp14) located on chromosome 1 is a mutant identified in maize which is deficient in ABA biosynthesis and *Vp14* cleave 9-*cis*-epoxy-carotenoids to form C25 apo-aldehydes and xanthoxin which act as a precursor of ABA in higher plants (Schwartz *et al.*1997) [5]. UniParc ID, UniPort ID and ENA ID of the protein of the corresponding were listed along with their corresponding chromosomal location (Table 1).

3.2 Gene structure analysis of Carotenoids dioxygenase

From the study of GSDS software, it was found that gene *Zm00001d031086*, *Zm00001d042076*, *Zm00001d009286* and *Zm00001d000095* contain single intron whereas in fourteen genes no intron was found. Gene with the similar structure was clustered together implying their functional similarity. *Zm00001d048373*, *Zm00001d045232* and *Zm00001d020069* contain 12, 11 and 10 introns, respectively and *Zm00001d031086*have largest size of intron (Fig. 1).

3.3 Physicochemical property of CCDs proteins and miRNA binding site in their CDS

Amino acids size of the protein was varied among the *Carotenoids dioxygenase* proteins which range from 77 to 639. Iso-electric point of the protein was varied from 4.5 to 10.77. Therefore Carotenoids dioxygenase proteins are working in wide range of pH. Carotenoids dioxygenase proteins were localized in a wide range of organelles viz., nucleus, chloroplast, mitochondria where few of them are working in the cytoplasm. One protein was found to be a plasma membrane bound whereas another protein showed extra cellular localization (Table 2). Localization of the Carotenoids dioxygenase proteins in different part of the cell implies their diverse role in plant physiology. All the protein contains retinal pigment epithelial (RPE) domain suggesting similar kind of function in plant (Table 2). Among the twenty four *CCDs* genes, twelve genes were found to be complemented with micro-RNA. Among which eleven genes were found to associated with mi-RNA with the help of psRNA Target server and one gene, Zm00001d022623, further included in association with zma-miR393d micro RNA on basis of PMRD mi-RNA data sets (Table 3, Fig. 2). MicroRNA binding site in the CCD sgenes implied their heterogeneity at regulation levels.

3.4 Phyllogenetic analysis of orthologs *carotenoids dioxygenase* gene in crop species

Ten orthologous *CCDs* genes from ten different species were taken for studying the evolutionary study of the *carotenoids dioxygenase* gene and presence of HMM domain of the protein they contain. All the protein contains retinal pigment epithelial (RPE) membrane (Table 4) bound domain which indicated their similar functions across the species. Exon-intron boundary construction using PIECE database revealed their similar kind of structural integrity across the species which signify the evolutionary significance of the gene (Fig. 3). All gene of this family are involved in carotenoids biosynthesis pathway in plant species.

Table 1: Paralogous genes of the *carotenoid dioxygenase* gene in maize

Gene	Gramene ID	Chr.	Position (bp)		UniParc ID	UniPort ID	ENA ID
			From	To			
NCED3	*Zm00001d027592*	1	8589933-8592006		UPI00084457B7	A0A1D6JN83	ONL93499
NCED3	*Zm00001d027690*	1	11168984-11169217		UPI000844AAEB	A0A1D6JNU2	ONL93695
NCED2	*Zm00001d031086 GRMZM2G407181*	1	176495793-176506822		UPI0008430A0D	A0A1D6KGE1 C4J3B4	ONM02149
viviparous14 (Vp14) NCED1	*Zm00001d033222 GRMZM2G014392*	1	255021507-255023321		UPI00022178CD	A0A1D6KX21 O24592	ONM07009
NCED3	*Zm00001d033377*	1	261041495-261042349		UPI000195CFAD	C0PIX7	BT068246 ONM07502
NCED4	*Zm00001d007876 GRMZM2G408158*	2	241535851-241537578		UPI0002218961	A0A1D6F9F2 C0HIM3	ONM27776
NCED3	*Zm00001d041319 GRMZM5G858784*	3	111919013-111920827		UPI00084360AE	A0A1D6MVG7 B6SSJ7	ONM32808
NCED3	*Zm00001d042076*	3	149717341-149723930		UPI000844BB8B	A0A1D6N121	ONM34466
NCED6	*Zm00001d051556 GRMZM2G110192*	4	162737448-162739364		UPI000221A475	K7U7K6	AQK54544
NCED9	*Zm00001d013689 GRMZM5G838285*	5	17465790-17467595		UPI0002207B31	A0A1D6GLI1	AQK64170
NCED8	*Zm00001d017766 GRMZM2G150363*	5	206199093-206201012		UPI0008425F87	A0A1D6HHH7	AQK73987
NCED3	*Zm00001d035724*	6	43483863-43487030		UPI0008433116	A0A1D6LI49	AQK79536
NCED5	*Zm00001d018819 GRMZM2G417954*	7	6342456-6344361		UPI000182B380	B6SV18	EU956583 ONM51390
NCED3	*Zm00001d019813*	7	64605343-64606179		UPI0004DEAA5D	A0A1D6I0E9	ONM53753
Carotenoid cleavage dioxygenase	*Zm00001d020069 GRMZM2G164967*	7	90083031-90086477		UPI000844C1B2	A0A1D6I1X3 B6UEM5	ONM54206
NCED7	*Zm00001d022623 GRMZM2G330848*	7	181400407-181408261		UPI0008435752	A0A1D6IPX2	ONM61280
NCED3	*Zm00001d008638*	8	15485032-15485280		UPI000842EE83	A0A1D6FED1	AQK90328
NCED3	*Zm00001d009286*	8	52122143-52123234		UPI0008444EA0	A0A1D6FIJ1	AQK91613
Carotenoid 910(9'10')-cleavage dioxygenase 1	*Zm00001d045232*	9	16742278-16746086		UPI0008DB33FF	A0A1D6NUR0	AQL01878
Carotenoid 910(9'10')-cleavage dioxygenase 1 (white cap1)	*Zm00001d048373 GRMZM2G057243*	9	155236300-155242285		UPI0000F0779D	A0SMH9 B4FBA4 Q45VT6 Q45VT7 Q5U905	AQL09521 BT034392 EU973966

NCED3	*Zm00001d023458*	10	6279773-6280198	UPI000843EBAB	A0A1D6ITE2	AQK39340
NCED3	*Zm00001d023690*	10	15582214-15582609	UPI0004DE8BC0	A0A1D6IUU2	AQK39788
NCED3	*Zm00001d000095*	-	-	UPI0008450968	A0A1X7YEQ7	-
NCED3	*Zm00001d000157*	-	-	UPI0002216B17	A0A1X7YFJ3	-

Table 2: Physiological properties of the Carotenoids dioxygenase proteins

Gene	Length (AA)	Mol Wt/Dalton	pI	GRAVY	Cello localization	RPE65 (AA)	other
NCED3	426	46542.54	8.55	-0.302	Nuclear (1.654), Chloroplast (1.172)	69-205, 197-326	-
NCED3	77	8757.82	4.5	-0.425	Cytoplasmic (2.226), Nuclear (1.295)	1-77	-
NCED2	569	61066.24	6.84	-0.121	Chloroplast (2.669)	91-561	-
viviparous14 (Vp14) NCED1	604	65462.12	5.68	-0.183	Chloroplast (1.554), Cytoplasmic (1.451)	127-595	-
NCED3	284	29389.43	10.77	-0.263	Nuclear (3.122)	194-246	1-17 (signal peptide)
NCED4	575	62033.42	6.61	-0.236	Chloroplast (1.596), Mitochondria (1.189)	100-567	-
NCED3	604	64758.34	6.30	-0.141	Chloroplast (2.657)	126-596	-
NCED3	420	46596.90	6.31	-0.340	Mitochondria (1.227), Cytoplasmic (1.004)	139-403	-
NCED6	638	68598.81	6.39	-0.212	Chloroplast (2.507)	154-629	1-16 (signal peotide)
NCED9	601	65157.08	6.41	-0.157	Chloroplast(1.730),Mitochondria(1.192),Cytoplasmic(1.113)	125-593	-
NCED8	639	69023.48	6.8	-0.183	Chloroplast (2.772)	156-630	-
NCED3	635	70322.46	8.69	-0.326	Mitochondria (2.484)	205-519	76-182 (Lsy1) 130-150 (signal peptide)
NCED5	573	61674.03	6.58	-0.210	Chloroplast (1.539), Mitochondria (1.059)	95-565	-
NCED3	189	19833.81	10.35	-0.202	Nuclear (2.120)	100-151	1-21 (signal peptide)
Carotenoid cleavage dioxygenase	456	50760.40	5.82	-0.225	Cytoplasmic (2.026)	1-449	-
NCED7	373	41495.13	5.91	-0.247	Cytoplasmic (2.994)	1-253	-
NCED3	82	9271.29	4.53	-0.535	Nuclear (1.959), Cytoplasmic (1.802)	1-81	-
NCED3	252	25955.52	10.80	-0.217	Nuclear (2.833)	162-214	1-15 (signal peptide)
Carotenoid 910(9'10')-cleavage dioxygenase 1	613	68839.37	5.98	-0.456	Cytoplasmic (1.280)	105-346, 344-608	1-18 (signal peptide)
Carotenoid 910(9'10')-cleavage dioxygenase 1 (white cap1)	518	58278.28	5.33	-0.312	Cytoplasmic (2.915)	60-507	-
NCED3	141	15577.61	6.81	-0.167	Extra-cellular (1.774)	1-135	-
NCED3	131	14774.79	4.97	-0.309	Cytoplasmic (3.083)	1-128	-
NCED3	164	17616.31	10.12	0.004	Plasma membrane (1.308), Mitochondria (1.184)	31-161	-
NCED3	105	11658.32	10.33	-0.305	Mitochondria (1.708), nuclear (1.166).	31-105	-

AA: Amino acid; pI: Iso-electric focusing; GRAVY: grand average of hydropathicity; RPE: Retinal pigment epithelial

Table 3: Micro RNA binding sites in *Carotenoid dioxygenase* genes in maize

miRNA_Acc.	Target_Acc.	Expectation	UPE	Target_start	Target_end	miRNA_aligned_fragment
zma-miR167d-3p	Zm00001d009286	5	14.286	739	761	GGUCAUGCUGCUGCAGCCUCACU
zma-miR166k-5p	Zm00001d018819	4.5	23.584	133	153	GGAUUGUUGUCUGGCUCGGGG
zma-miR166n-5p	Zm00001d018819	4.5	23.584	133	153	GGAUUGUUGUCUGGCUCGGUG
zma-miR167d-3p	Zm00001d019813	5	9.302	553	575	GGUCAUGCUGCUGCAGCCUCACU
zma-miR528a-5p	Zm00001d019813	4	23.497	74	94	UGGAAGGGGCAUGCAGAGGAG
zma-miR528b-5p	Zm00001d019813	4	23.497	74	94	UGGAAGGGGCAUGCAGAGGAG
zma-miR167f-3p	Zm00001d020069	4.5	24.776	282	303	GAUCGUGCUGCGCAGUUUCACC
zma-miR2275d-5p	Zm00001d020069	5	11.63	1575	1595	AGAGUUGGAGGAAAGAAAACU
zma-miR159e-5p	Zm00001d020069	5	14.045	991	1011	CAGCUCCUGCAGCAUCUGUUC
zma-miR164g-3p	Zm00001d020069	4.5	24.294	555	575	CACGUGCUCCCCUUCUCCACC
zma-miR399b-5p	Zm00001d023690	4.5	21.752	124	144	GUGCAGCUCUCCUCUGGCAUG
zma-miR399g-5p	Zm00001d023690	5	21.752	124	144	GGGCAACCCCCCGUUGGCAGG
zma-miR399h-5p	Zm00001d023690	5	21.752	124	144	GUGCAGUUCUCCUCUGGCACG
zma-miR399i-5p	Zm00001d023690	5	21.752	124	144	GUGCGGCUCUCCUCUGGCAUG
zma-miR164e-3p	Zm00001d035724	5	23.51	1353	1373	CAUGUGUCCGCCCUCUCCACC
zma-miR528a-5p	Zm00001d035724	5	11.663	1552	1572	UGGAAGGGGCAUGCAGAGGAG
zma-miR528b-5p	Zm00001d035724	5	11.663	1552	1572	UGGAAGGGGCAUGCAGAGGAG
zma-miR164a-5p	Zm00001d041319	4	13.683	16	36	UGGAGAAGCAGGGCACGUGCA
zma-miR164d-5p	Zm00001d041319	4	13.683	16	36	UGGAGAAGCAGGGCACGUGCA
zma-miR164b-5p	Zm00001d041319	4	13.683	16	36	UGGAGAAGCAGGGCACGUGCA
zma-miR164c-5p	Zm00001d041319	4	13.683	16	36	UGGAGAAGCAGGGCACGUGCA
zma-miR167c-3p	Zm00001d041319	5	13.683	1	23	GAUCAUGCUGUGGCAGCCUCACU

zma-miR164f-5p	Zm00001d041319	5	13.683	15	36	UGGAGAAGCAGGGCACGU-GCU
zma-miR164g-5p	Zm00001d041319	4	13.683	16	36	UGGAGAAGCAGGGCACGUGCA
zma-miR164h-5p	Zm00001d041319	4	13.683	16	36	UGGAGAAGCAGGGCACGUGUG
zma-miR444a	Zm00001d042076	5	16.937	512	532	UGCAGUUGUUGUCUCAAGCUU
zma-miR444b	Zm00001d042076	5	16.937	512	532	UGCAGUUGUUGUCUCAAGCUU
zma-miR166g-5p	Zm00001d045232	4.5	16.683	1288	1308	GGAAUGUUGUCUGGUUGGAGA
zma-miR166c-5p	Zm00001d045232	5	16.683	1288	1308	GGAAUGUUGUCUGGCUCGAGG
zma-miR166k-5p	Zm00001d045232	4	16.683	1288	1308	GGAUUGUUGUCUGGCUCGGGG
zma-miR166n-5p	Zm00001d045232	4	16.683	1288	1308	GGAUUGUUGUCUGGCUCGGUG
zma-miR171a-5p	Zm00001d048373	5	21.397	1624	1643	UAUUGGCGAGGUUCAAUCAGA
zma-miR408a	Zm00001d048373	4.5	24.99	1695	1716	CUGCACUGCCUCUUC-CCUGGC
zma-miR395d-5p	Zm00001d048373	5	22.608	393	414	GUUCUAUGCAAGCACUUCACGA
zma-miR395e-5p	Zm00001d048373	4.5	22.608	393	414	GUUCCCUUCAAGCACUUCACAU
zma-miR395g-5p	Zm00001d048373	5	22.608	393	414	GUUCUAUGCAAGCACUUCACGA
zma-miR395h-5p	Zm00001d048373	4.5	22.608	393	414	GUUCCCUUCAAGCACUUCACAU
zma-miR395j-5p	Zm00001d048373	4.5	22.608	393	414	GUUCCCUUCAAGCACUUCACAU
zma-miR395k-5p	Zm00001d048373	5	22.608	393	414	GUUUCCUUCAAGCACUUCACAU
zma-miR395o-5p	Zm00001d048373	5	22.608	393	414	GUUCUCUUCAAGCACUUCACGA
zma-miR395p-5p	Zm00001d048373	4.5	22.608	393	414	GUUCCCUUCAAGCACUUCACAU
zma-miR408b-3p	Zm00001d048373	4.5	24.99	1695	1716	CUGCACUGCCUCUUC-CCUGGC
zma-miR167c-3p	Zm00001d051556	4.5	22.465	1859	1881	GAUCAUGCUGUGGCAGCCUCACU
zma-miR393d	Zm00001d022623	4.5	19.283	930	950	UCCAAAGGGAUUGCACUGAUC
zma-miR393e	Zm00001d022623	4.5	19.283	930	950	UCCAAAGGGAUUGCACUGAUC

Acc.: Accession; UPE: unpair target site

Table 4: Orthologous of *Carotenoids dioxygenase* genes in ten species

Crop	Gene	Description of protein domain
Brachypodium distachyon	BRADI4G00335	Retinal pigment epithelial membrane protein (57-536)
Setaria italica	Si021680m.g	Retinal pigment epithelial membrane protein (54-533)
Oryza sativa var. indica	BGIOSGA037874	Retinal pigment epithelial membrane protein (51-531)
Arabidopsis thaliana	AT3G63520	Retinal pigment epithelial membrane protein(52-528)
Triticum aestivum	TRIAE_CS42_U_TGACv1_643498_AA2132620	Retinal pigment epithelial membrane protein (68-553)
Medicago truncatula	MTR_8g037275	Retinal pigment epithelial membrane protein (88-565)
Brassica oleracea	Bo8g099460	Retinal pigment epithelial membrane protein (133-554)
Helianthus annuus	HannXRQ_Chr17g0562991	Retinal pigment epithelial membrane protein(53-533)
Aegilops tauschii	F775_00763	Retinal pigment epithelial membrane protein (68-550)
Zea mays	Zm00001d048373	Retinal pigment epithelial membrane protein (60-539)

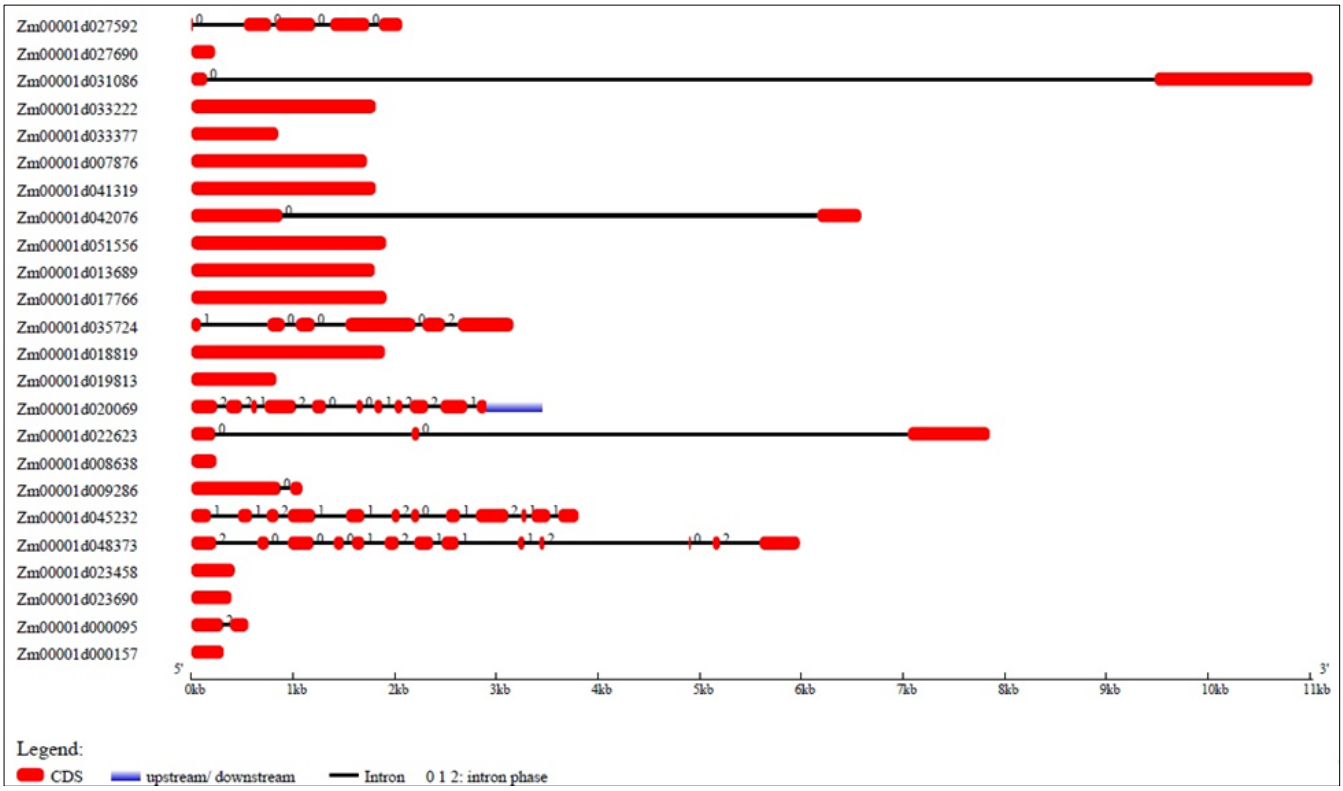

Fig 1: Intron exon organization of *Carotenoid dioxygenase* genes in maize using GSDS server

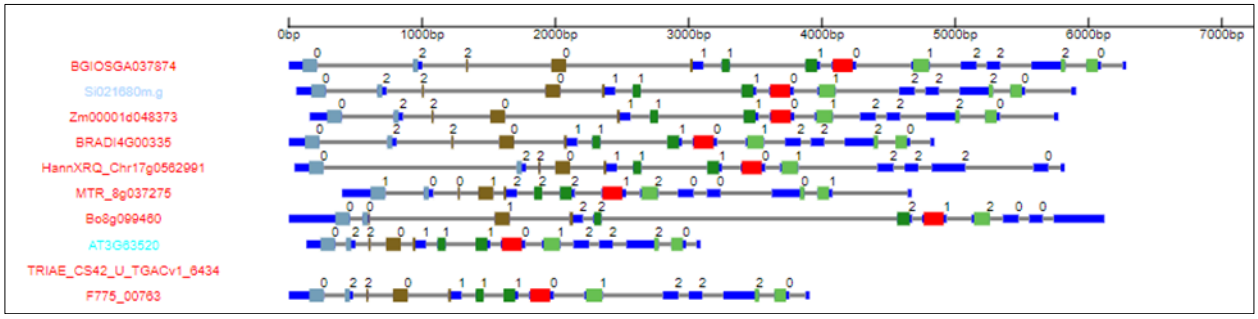

Fig 2: Intron exon organization of *Carotenoid dioxygenase* genes in maize using PIECE (Plant Intron Exon Comparison and Evolution database)

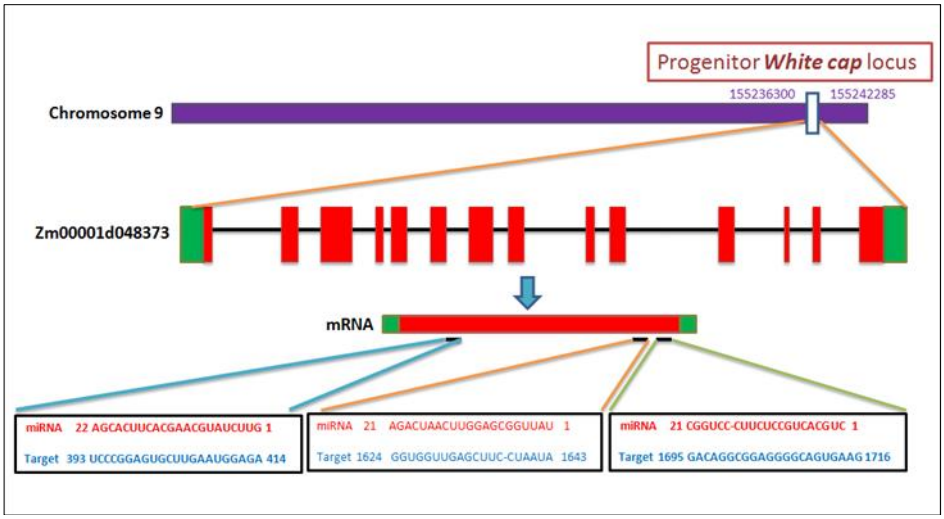

Fig 3: Micro RNA binding site in *CCD1* gene of maize at the progenitor *White cap* locus

4. Conclusion

In this study maize genome has been analyzed to identify and characterize *carotenoids dioxygenase* gene by using different bioinformatics tool. HMM profile search was employed to varify candidate *carotenoids dioxygenase* genes. Iso-electric point was obtained from ProtParam tool. Chromosomal location, locus ID was retrieved from Ensemble plant database. Exon-intron distribution in *carotenoids dioxygenase* genes were analyzed using GSDS server. This comprehensive analysis of *carotenoids dioxygenase* genes paves the way for deciphering the functional role of all *carotenoids dioxygenase* genes in maize. *CCD1* belonging to the *carotenoids dioxygenase* gene family has a detrimental role in carotenoids biosynthesis pathway. This analysis presented here may be useful in selection of candidates genes for further functional characterization related to maize carotenoids biosynthesis pathway.

Acknowledgements

The first author is thankful to Indian Council of Agricultural Research for the Junior Research Fellowship to pursue the master's degree programme. Financial support from the ICAR funded CRP on Biofortification is thankfully acknowledged.

5. References

1. Vallabhaneni R, Bradbury LM, Wurtzel ET. The carotenoid dioxygenase gene family in maize, sorghum, and rice. Archives of Biochemistry and Biophysics. 2010; 504(1):104-111.
2. Hou X, Rivers J, Leon P, McQuinn RP, Pogson BJ. Synthesis and function of apocarotenoid signals in plants. Trends in Plant Science. 2016; 21(9):792-803.
3. Tan BC, Joseph LM, Deng WT, Liu L, Li QB, Cline K *et al*. Molecular characterization of the Arabidopsis *9-cis epoxy carotenoids dioxygenase*gene family. The Plant Journal. 2003; 35(1):44-56.
4. Tan BC, Schwartz SH, Zeevaart JA, McCarty DR. Genetic control of abscisic acid biosynthesis in maize. Proceedings of the National Academy of Sciences. 1997; 94(22):12235-12240.
5. Schwartz SH, Tan BC, Gage DA, Zeevaart JA, McCarty DR. Specific oxidative cleavage of carotenoids by *VP14* of maize. Science. 1997; 276(5320):1872-1874.
6. Vogel JT, Tan BC, McCarty DR, Klee HJ. The *carotenoid cleavage dioxygenase 1* enzyme has broad substrate specificity, cleaving multiple carotenoids at two different bond positions. Journal of Biological Chemistry. 2008; 283(17):11364-11373.
7. Sun Z, Hans J, Walter M, Matusova R, Beekwilder J, Verstappen F *et al*. Cloning and characterisation of a maize *carotenoid cleavage dioxygenase* (*ZmCCD1*) and its involvement in the biosynthesis of apocarotenoids with various roles in mutualistic and parasitic interactions. Planta. 2008; 228(5):789-801.
8. Tan BC, Guan JC, Ding S, Wu S, Saunders JW, Koch KE *et al*. Structure and origin of the White Cap locus and its role in evolution of grain color in maize. Genetics. 2017; 206:135-150.
9. Vignesh M, Hossain F, Nepolean T, Saha S, Agarwal PK, Guleria SK *et al*. Genetic variability for kernel β-carotene and utilization of *crtRB1* 3'TE gene for biofortification in maize *(Zea mays* L.). Indian J Genet. 2012; 72(2):189-194.
10. Muthusamy V, Hossain F, Thirunavukkarasu N, Choudhary M, Saha S, Bhat JS *et al*. Development of β-

carotene rich maize hybrids through marker-assisted introgression of β-carotene hydroxylase allele. PLoS ONE. 2014; 9(12):e113583.
11. Zunjare RU, Hossain F, Muthusamy V, Baveja A, Chauhan HS, Bhat JS *et al*. Development of Biofortified Maize Hybrids through Marker-Assisted Stacking of *β-carotene hydroxylase*, *lycopene-ε-cyclase* and *opaque2* Genes. Frontiers in Plant Science. 2018; 9:178.
12. Simkin AJ, Schwartz SH, Auldridge M, Taylor MG, Klee HJ. The tomato *carotenoid cleavage dioxygenase* 1 genes contribute to the formation of the flavor volatiles β-ionone, pseudoionone, and geranylacetone. The Plant Journal. 2004; 40(6):882-892.
13. Frusciante S, Diretto G, Bruno M, Ferrante P, Pietrella M, Prado-Cabrero A *et al*. Novel carotenoid cleavage dioxygenase catalyzes the first dedicated step in saffron crocin biosynthesis. Proceedings of the National Academy of Sciences. 2014; 111(33):12246-12251.
14. Campbell R, Ducreux LJ, Morris WL, Morris JA, Suttle JC, Ramsay G *et al*. The metabolic and developmental roles of carotenoid cleavage dioxygenase 4 from potato (Solanumtuberosum L). Plant Physiology, 2010, 110.
15. Alder A, Jamil M, Marzorati M, Bruno M, Vermathen M, Bigler P *et al*. The path from β-carotene to carlactone, a strigolactone-like plant hormone. Science. 2012; 335(6074):1348-1351.
16. Zerbino DR, Achuthan P, Akanni W, Amode MR, Barrell D, Bhai J *et al*. Nucleic Acids Research. 2018-2017; 46(D1):D754-D761.
17. Tello-Ruiz MK, Stein J, Wei S, Youens-Clark K, Jaiswal P, Ware Gramene D. a resource for comparative analysis of plants genomes and pathways. In Plant Bioinformatics. Humana Press, New York, NY, 2016, 141-163.
18. Silvester N, Alako B, Amid C, Cerdeno-Tarraga A, Clarke L, Cleland I *et al*. The european nucleotide archive in 2017. Nucleic Acids Research. 2017; 46(D1):D36-D40.
19. Hu B, Jin J, Guo AY, Zhang H, Luo J, Gao G. GSDS 2.0: an upgraded gene feature visualization server. Bioinformatics. 2014; 31(8):1296-1297.
20. Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL *et al*. The Pfam protein families database: towards a more sustainable future. Nucleic Acids Research. 2015; 44(D1):D279-D285.
21. Gasteiger E, Hoogland C, Gattiker A, Wilkins MR, Appel RD, Bairoch A. Protein identification and analysis tools on the ExPASy server. In The Proteomics Protocols Handbook, 2005, 571-607.
22. Yu CS, Chen YC, Lu CH, Hwang JK. Prediction of protein subcellular localization. Proteins: Structure, Function and Bioinformatics. 2006; 64(3):643-651.
23. Dai X, Zhao PX. psRNA Target: a plant small RNA target analysis server. Nucleic Acids Research. 2011; 39(2):W155-W159.
24. Wang Y, You FM, Lazo GR, Luo MC, Thilmony R, Gordon S *et al*. PIECE: a database for plant gene structure comparison and evolution. Nucleic Acids Research. 2012; 41(D1):D1159-D1166.