Potato J (2023) 50 (2): 132-148

ANALYSIS OF CAROTENOIDS METABOLISM GENES LANDSCAPE IN POTATO AND CLONING OF *ZEAXANTHIN EPOXIDASE* GENE

Som Dutt[1*], Kalpana Devi[1], Mamta Chauhan[1], Pinky Raigond[1,2], Dharmendra Kumar[1], Milan Kumar Lal[1], Arvind Kumar Jaiswal[3], Bandana Kaundal[4], Devendra Kumar[4], Satish Kumar Luthra[4], Dinesh Kumar[1] and Brajesh Singh[1]

ABSTRACT: Genome-wide identification of all the genes involved in carotenoids biosynthesis were carried out *in silico* using potato genome sequence. Of the 12 genes analysed [geranylgeranyl pyrophosphate synthase (GGPS), phytoene synthase (PSY), f-carotene desaturase (ZDS), carotenoid isomerise (CRTISO), lycopene ε-cyclase (LCY-ε), lycopene β-cyclase (LCY-β), β-carotene hydroxylase (CHY-β), violaxanthin de-epoxidase (VDE), zeaxanthin epoxidase (ZEP), neoxanthin synthase (NXS), carotenoid cleavage dioxygenase (CCD), 9-cisepoxycarotenoids dioxygenase (NCED)]; three genes namely ZDS, CRTISO and LCY-ε was found as single gene, on chromosome 1, 10 and 12, respectively. GGPS (chromosomes 4 & 11), LCY-β (chromosomes 4 & 10), CHY-β (chromosomes 3& 6) had two genes each. VDE (chromosome 4), NXS (chromosomes 2, 3 & 6) and NCED (chromosome 5, 7 & 8) showed presence of 3, 4 and 5 genes, respectively. CCD (chromosome 1 7 & 8) and PSY (chromosomes 2, 3, 4, 8 & 12) had six genes each in potato genome. These carotenoid biosynthesis genes differed in terms of number of introns, length of open reading frame, molecular weight and isoelectric points of deduced proteins, secondary structure composition etc. This information generated about genes and proteins of carotenoids biosynthesis pathway in potato may be very crucial and of vital utility for genome editing mediated interventions for enhancing total/ targeted carotenoids content in potato. Further, ZEP was cloned full length. The complete ORF of cloned ZEP was 2010 nucleotides long and encoding a protein having 669 amino acids (NCBI acc no. MK852682). It exhibited variations at 9 amino acid positions when compared with the one available in potato genome sequence database. These variations at 9 amino acids might provide some uniqueness to the cloned gene which might be involved in regulation of carotenoid biosynthesis pathway especially beta-carotene, violaxanthin, zeaxanthin carotenoids etc.

KEYWORDS: Genome-wide, genome editing, gene structure, metabolic engineering

INTRODUCTION

Humans and animals do not synthesize carotenoids and thus mainly depends on plant-based diets as source of carotenoids. Carotenoids are necessary to maintain normal health and behavior of animals. Carotenoids acts as precursors of vitamin A and play a range of other important functions in human health. Consumption of carotenoids has been associated with beneficial effects in a number of systemic diseases, eye disorders and cognitive function (Tapiero *et al*., 2004; Lima *et al*., 2016; Buscemi *et al*., 2018; Eggersdorfer and Wyss, 2018). Of more than several hundred carotenoids found in nature only about 50 carotenoids are found in majority of human diet, and around 20 of them are present in human tissue and blood (Fiedor and Burda, 2014). The carotenoids primarily exert antioxidant effects, but individual carotenoids may also act through other mechanisms.

In plants, carotenoids and their oxidative and enzymatic cleavage products are crucial for various biological processes in plants, such as assembly of photosystems and light harvesting antenna complexes

[1]ICAR-Central Potato Research Institute, Shimla - 171001, Himachal Pradesh, India
[2]Present address: ICAR-National Research Centre on Pomegranate, Solapur 413255, Maharashtra, India
[3]ICAR-Central Potato Research Institute, Regional Station, Jalandhar - 144003, Punjab, India
[4]ICAR-Central Potato Research Institute, Regional Station, Modipuram, Meerut - 250110, Uttar Pradesh, India
*Corresponding author; email: Som.Dutt1@icar.gov.in; sd_bio@yahoo.com

for photosynthesis and photoprotection, and regulation of growth and development (Cazzonelli and Pogson, 2010; Ruiz-Sola and Rodrıguez-Concepciona, 2012; Havaux, 2014). Carotenoids are also reported to act as signalling molecules and are thus involved in plant- environment interactions (Walter and Strack, 2011; Cazzonelli, 2011). Due to the pivotal role of carotenoids in nature, the understanding of how plant cells regulate the accumulation and flux of various carotenoids and their metabolites is advancing rapidly. Yet, further insights are essentially required for making plants synthesise desired carotenoids in higher quantities to make them nutritionally more enriched. Carotenoid are C40 tetraterpenoid lipophilic metabolites and are derived from isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP). The IPP and DMAPP used for carotenoid biosynthesis in plants are derived from the MEP pathway (Rodriguez-Concepcion and Boronat, 2002). Various enzymes involved in the carotenoids biosynthesis pathway are depicted in Fig. 1.

In potato tubers total carotenoid content ranges between 0.05 to 0.6 mg per 100 g fresh weight (FW) (Brown, 2005; Xu *et al*., 2003). However, few studies have also reported carotenoid content in potato beyond this range (Bonierbale *et al*., 2009; Tatarowska *et al*., 2019; Bahadori *et al*., 2023). The major carotenoids in potato include; zeaxanthin, antheraxanthin, violaxanthin and lutein, whereas neoxanthin, β-carotene and β-cryptoxanthin are found in trace amounts (Burgos *et al*., 2009, Bonierbale *et al*., 2009). However, their relative proportions have been shown to be varying. Tuber carotenoids content is affected by genotype and growing environment (Hamouz *et al*., 2016). Also, it has been found that total carotenoids content is higher in immature tubers and decreases with tuber maturity (Kotikova *et al*., 2007).

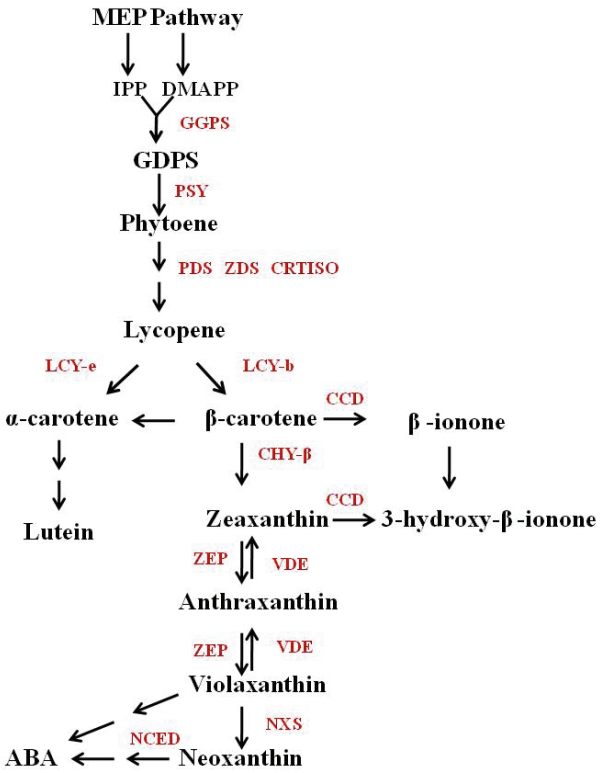

Fig. 1: Enzymes of carotenoids biosynthesis pathway in plants. GGPS: geranylgeranyl pyrophosphate synthase. PSY: phytoene synthase; PDS: phytoene desaturase;ZDS: f-carotene desaturase; CRTISO: carotenoid isomerase; LCY-ε: lycopene ε-cyclase;LCY β, lycopene β-cyclase; CHY-β, β-carotene hydroxylase; VDE, violaxanthin de-epoxidase; ZEP, zeaxanthin epoxidase; NXS, neoxanthin synthase; CCD, carotenoid cleavage dioxygenase; NCED, 9-cisepoxycarotenoids dioxygenase.

Numerous groups have attempted to increase potato carotenoids using transgenic strategies. The strategy commonly used in plants is to increase the biosynthetic capacity by altering the carotenogenic enzyme activities (Romer *et al*., 2002, Ducreux*et al*., 2005; Morris *et al*., 2006; Diretto *et al*., 2007) but due to stringent regulations for genetically modified organisms these transgenic plant lines could not reach the farms and thus were not cultivated. However, now with the advent of advanced genome editing tools and the fact that in most of the countries the plants developed using genome editing approaches (Site-Directed Nuclease 1 (SDN1)

and Site-Directed Nuclease 2 categories) have been kept out of the ambit of bio-safety regulations and thus most probably are speculated reaching fields and the consumers in near future. Having prior knowledge about the molecular machinery (genes and proteins) associated with the trait aimed to be improved using genome editing approach, is pre-requisite. Keeping this in view, the present study was undertaken on genome-wide analysis of carotenoid metabolism in potato (*Solanum tuberosum* L.). Further, an important gene of the carotenoid pathway was cloned from potato and characterized.

MATERIALS AND METHODS

Planting Material

Well sprouted tubers of Kufri Jyoti cultivar of potato were planted in pots and were kept in plant growth chambers having 22 / 18 °C day/ night temperature with 16 h photoperiod. Leaves at 4th position from the top were used for extraction of RNA.

Cloning of cDNA of zeaxanthin epoxidase

RNA was isolated from leaf tissue using PureLink®RNA Mini Kit (M/s Lifetechnologies, USA) and digested with DNase I (RNase free) (M/s Thermo Fisher Scientific, USA). Complementary DNA (cDNA) was synthesized from 2 µg of DNase-treated total RNA as a template in 20 µl reaction volume by using cDNA synthesis kit (Invitrogen, USA). Primers (O6629-F-F1 and 06629-F-R1) for ZEP were designed from the conserved regions of corresponding gene reported from different plant sources and the partial gene sequence was amplified by PCR as detailed in Table 1**.** Primers were designed by using Primer3 (http://primer3.ut.ee) and got synthesized from M/s Integrated DNA Technologies (IDT) considering important factors like % GC content, melting temperature for the primer set, and formation of internal hairpin loops and dimerization of oligos. List of primers used are listed in Table 1. The amplicon was cloned in pGEM-T Easy Vector (Promega, USA), plasmids were isolated using GeneJET™ Plasmid Miniprep Kit (M/s Fermentas Inc., USA), and sequencing was performed using Big Dye terminator cycle sequencing mix (Version 3.1; M/s Applied Biosystems, USA) using an automated DNA sequencer (ABI 3130 xl Genetic Analyzer, M/s Applied Biosystems, USA). Protocols

Table 1: Primer used for cloning and expression analysis

S. No.	Name of primer	Sequence of primers (5'-3')	PCR condition
	Primer used for cloning of full length ORF		
1	06629-F-F5	ATGTATTCAACTGTGTTTTACACTTC	Initial denaturation at 94 °C for 3 min, followed by 35 cycles of 94 °C, 30 s; 55–60 °C, 40 s; 72 °C, 50 s. Final extension at 72 °C for 7 min
2	06629-F-R5	GTGAAGCAGTGGGGGCAGCGTAA	
3	O6629-F-F1	ATGTATTCAACTGTGTTTTACACTTC	Initial denaturation at 94 °C for 3 min, followed by 35 cycles of 94 °C, 30 s; 55–60 °C, 40 s; 72 °C, 50 s. Final extension at 72 °C for 7 min
4	06629-F-R1	CATCACTGGTCAAAGGATTA	
5	06629-F-F2	CATCACTGGTCAAAGGATTA	
6	06629-F-R2	CTTTGTTTCTTCAGATGTGG	
7	06629-F-F3	CTTGTTATACTGGAATTGCA	
8	06629-F-R3	GAATGGCTGCAATCATGGCA	
9	06629-F-F4	GAATGGCTGCAATCATGGCA	
10	06629-F-R4	GTGAAGCAGTGGGGGCAGCGTAA	

were followed essentially as described by the respective manufacturer. For sequencing, full length cloned CDS of ZEP, the primers were designed (06629-F-F2, 06629-F-R2, 06629-F-F3, 06629-F-R3, 06629-F-F4, 06629-F-R4) in order to get overlapping fragments amplified which were sub-cloned and sequenced. Sequences of the sub-cloned fragment obtained were overlapped to derive the complete sequence of cloned ORF.

Genome-wide analysis of carotenoids metabolism pathway genes

Carotenoid pathway genes were searched in various published literature and were then looked into potato genome sequence database (http://spuddb.uga.edu/) using PLAZA 5.0 server (https://bioinformatics.psb.ugent.be/plaza/). All the 12 genes were searched individually. All the information obtained about the gene ID, number of genes, etc were enlisted and compiled systematically. The genes found in the potato genome sequence database were then analysed for their chromosome location and size.

Physical mapping of the carotenoids biosynthesis genes and analysis of gene structures

Locations of the genes were drawn using chromosomal coordinates as per the information given in literature (Genome sequence and analysis of the tuber crop potato (The potato genome sequencing consortium. 2011. Nature. 475:189-195). Analysis of genes' structure, introns localization, ORF retrieval etc. were performed using PLAZA 5.0 portal.

Estimation of secondary structure, signal peptide and transmembrane regions

Estimation of secondary structure (for % helix, beta- sheets and coils), signal peptide and transmembrane regions of proteins were performed as per Montgomerie *et al.*, (2008).

Alignment of genes, proteins, computing isoelectric point and molecular weight

Pair-wise/ multiple alignments were performed using Clustal Omega sequence alignment program (https://www.ebi.ac.uk/Tools/msa/clustalo/). For computing isoelectric point and molecular weight of various protein, tools available on Expasy resource portal of the SIB Swiss Institute of Bioinformatics (https://www.expasy.org/) were used.

RESULTS AND DISCUSSION

Genes of carotenoids biosynthesis are present in a range of isoforms and scattered in all the 12 chromosomes except chromosome 9 of potato genome.

All the twelve genes (presented in Fig.1) viz: GGPS, PSY, ZDS, CRTISO, LCY-ε, LCY-β, CHY-β; VDE, ZEP, NXS, CCD and NCED were searched in the potato genome database for their number of gene isoforms and physical location on the potato genome i.e. chromosomes. Six genes each of PSY (IDs: (PGSC0003DMG400021926, PGSC0003DMG400016721, PGSC0003DMG400024063, PGSC0003DMG400029005, PGSC0003DMG400012507, PGSC0003DMG400019372) and CCD (PGSC0003DMG400026032, PGSC0003DMG400018480, PGSC0003DMG400018481, PGSC0003DMG401001968, PGSC0003DMG402001968, PGSC0003DMG400001969, and were found in the data base (Table 1). Five, four and three genes of NCED (PGSC0003DMG400019162, PGSC0003DMG400027633, PGSC0003DMG400015100, PGSC0003DMG400004311, PGSC0003DMG400004312), NXS (PGSC0003DMG400024184, PGSC0003DMG401026407, PGSC0003DMG402019589, PGSC0003DMG400005886) and VDE (PGSC0003DMG400010688, PGSC0003DMG400010690, PGSC0003DMG400020993), respectively, were found in the potato genome sequence database (Table 2). CHY-β (PGSC0003DMG400010169, PGSC0003DMG400028897), LCY-β (PGSC0003DMG400010637, PGSC0003DMG400008159), GGPS (PGSC0003DMG400027856, PGSC0003DMG400015673)

showed two copies each. For rest of the four genes viz: ZDS (PGSC0003DMG400022473), CRTISO (PGSC0003DMG400028224), LCY-ε (PGSC0003DMG400000333), and ZEP (PGSC0003DM400004040) single gene was found in the potato sequence database (Table 2).

The analysis revealed that these 12 genes and their isoforms were scattered throughout

Table 2: List of genes of carotenoids metabolic pathway in potato; alongwith their sequence id, chromosome location and size.

S. No.	Gene	Sequence ID	Chromosome Number	Start Position	End position	Size (basepair)
1	Phytoene synthase (PSY)	PGSC0003DMG400021926	chr02	11294456	11297566	3111
		PGSC0003DMG400016721	chr02	36955673	36959715	4043
		PGSC0003DMG400024063	chr03	2740455	2745217	4763
		PGSC0003DMG400029005	chr04	18612745	18618523	5779
		PGSC0003DMG400012507	chr08	458271	462864	4594
		PGSC0003DMG400019372	chr12	19128104	19129633	1530
2	Beta-carotene hydroxylase (CHY-β)	PGSC0003DMG400010169	chr03	43936158	43938595	2438
		PGSC0003DMG400028897	chr06	37867498	37869654	2157
3	Zeta-carotene desaturase (ZDS)	PGSC0003DMG400022473	chr01	75477194	75485576	8383
4	Carotenoid isomerase (CRTISO)	PGSC0003DMG400028224	chr10	56306412	56312609	6198
5	Lycopene beta cyclase (LCY-β)	PGSC0003DMG400010637	chr04	12331538	12333464	1927
		PGSC0003DMG400008159	chr10	58302687	58304848	2162
6	Lycopene epsilon cyclase (LCY-ε)	PGSC0003DMG400000333	chr12	3547170	3553905	6736
7	Violaxanthin de-epoxidase (VDE)	PGSC0003DMG400010688	chr04	45916315	45921142	4828
		PGSC0003DMG400010690	chr04	46040209	46044982	4774
		PGSC0003DMG400020993	chr04	46104170	46106015	1846
8	Zeaxanthin epoxidase (ZEP)	PGSC0003DM400004040	chr02	43804709	43809973	5265
9	Neoxanthin synthase (NXS)	PGSC0003DMG400024184	chr02	21954500	21958056	3557
		PGSC0003DMG401026407	chr02	41342717	41345874	3158
		PGSC0003DMG402019589	chr03	4736917	4745415	8499
		PGSC0003DMG400005886	chr06	54774962	54776783	1822
10	9-cis-epoxycarotenoid dioxygenase (NCED)	PGSC0003DMG400019162	chr05	48865165	48866694	1530
		PGSC0003DMG400027633	chr07	51778895	51781175	2281
		PGSC0003DMG400015100	chr08	10998923	11001084	2162
		PGSC0003DMG400004311	chr08	39185413	39188702	3290
		PGSC0003DMG400004312	chr08	39213385	39218093	4709
11	Carotenoid cleavage dioxygenase (CCD)	PGSC0003DMG400026032	chr01	69869750	69873902	4153
		PGSC0003DMG400018480	chr08	3615970	3618462	2493
		PGSC0003DMG400018481	chr08	3626959	3631890	4932
		PGSC0003DMG401001968	chr08	47359253	47362257	3005
		PGSC0003DMG402001968	chr08	47360292	47362256	1965
		PGSC0003DMG400001969	chr08	47375775	47378534	2760
12	Geranylgeranyl pyrophosphate synthase (GGPS)	PGSC0003DMG400027856	chr04	69344146	69345467	1322
		PGSC0003DMG400015673	chr11	1682159	1683721	1563

the potato genome, except on chromosome 9 (Fig. 2, Table 2). Maximum numbers of genes/ gene-isoforms were found to be present on chromosome 8, and mostly clustered at the ends of the chromosome. Chromosomes 5, 7 and 11 contained only one gene each, whereas, chromosomes 1, 6, 10 and 12 possessed 2 genes each. These genes were found to be varying size and having a wide range of Open reading frames (ORF) and thus encoding proteins differing in number of amino acids (Table 3). Further, the encoded proteins differed in their calculated molecular weights and isoelectric points (Table 3). Knowledge about the precise physical location of gene(s) on genome is vital to understand the segregation behavior and probability of their crossing & recombination patterns. Genome-wide analysis of a complex pathway like that of carotenoids may assist in devising proper breeding programs for improving total/ targeted carotenoid compound(s) in potato.

Having genome-wide information of a gene of gene families is vital for systematic

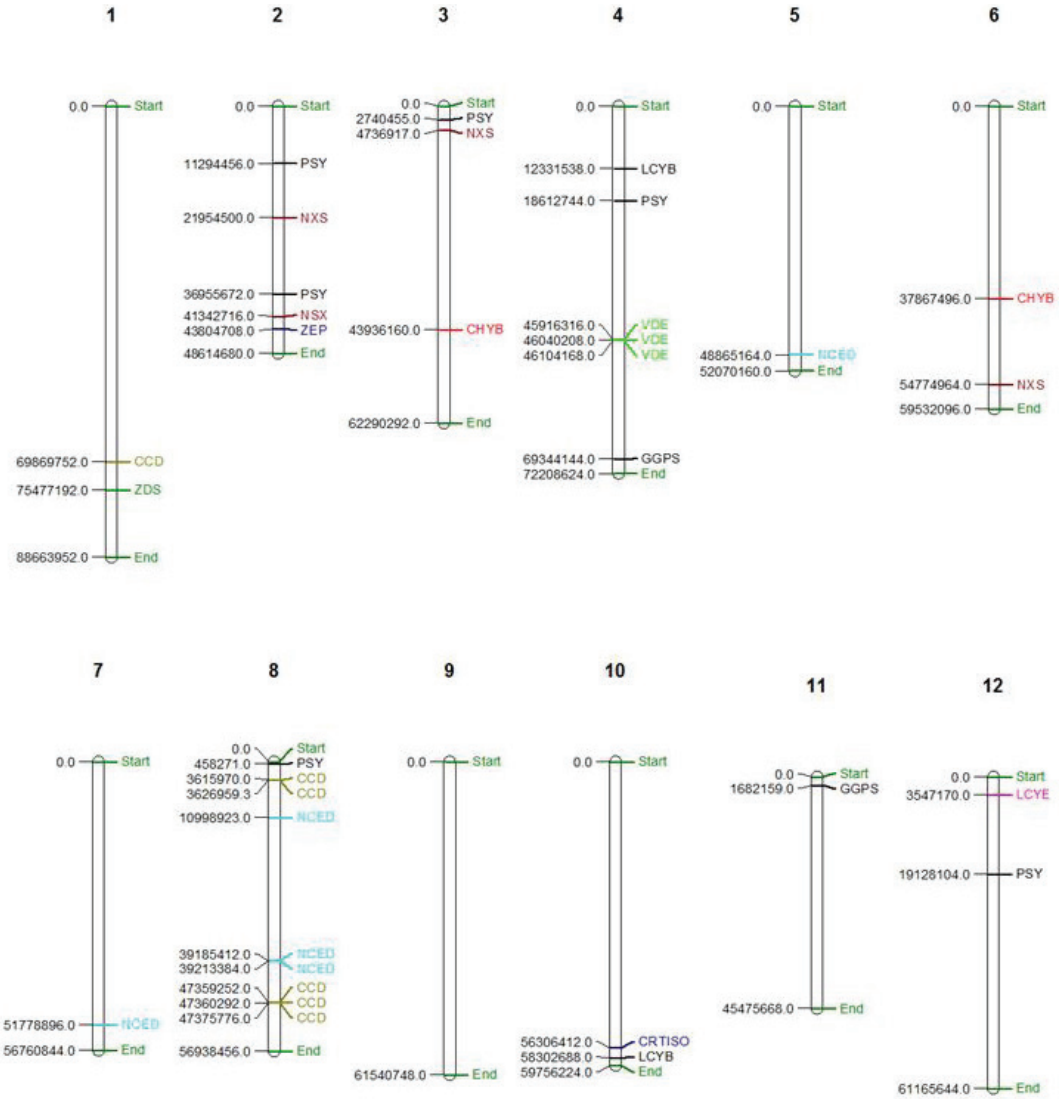

Fig. 2: Genome-wide landscape of genes of carotenoids metabolic pathway in potato

understanding the role of the gene (s) and their further utilization for improving target trait in plants. For example, Valcarcel *et al*., (2016) found no relationship between transcript levels of phytoene synthase (PSY) gene and total carotenoid accumulation in potato. The

Table 3: Structural features (number of introns, open reading frame size) of carotenoid pathway genes and characteristics of the deduced proteins (number of amino acids, molecular weight, theoretical isoelectric point) of carotenoids metabolic pathway in potato.

Gene	Sequence ID	Number of introns	Open reading frame (nucleotides)	Number of amino acids in deduced protein	Theoretical Isoelectric point/ molecular weight (Dalton)
Phytoene synthase (PSY)	PGSC0003DMG400021926	3	480	159	9.00 / 18404.59
	PGSC0003DMG400016721	5	1317	438	8.76 / 49415.49
	PGSC0003DMG400024063	5	1239	412	8.44 / 46484.11
	PGSC0003DMG400029005	4	681	226	4.84 / 25548.35
	PGSC0003DMG400012507	0	912	303	9.10 / 33727.67
	PGSC0003DMG400019372	2	462	153	4.22 / 17438.83
Beta-carotene hydroxylase (CHY-β)	PGSC0003DMG400010169	6	945	314	9.12 / 35067.69
	PGSC0003DMG400028897	6	930	309	9.30 / 34305.64
Zeta-carotene desaturase (ZDS)	PGSC0003DMG400022473	13	1767	588	8.60 / 64699.45
Carotenoid isomerase (CRTISO)	PGSC0003DMG400028224	12	1848	615	6.92 / 67516.46
Lycopene beta cyclase (LCY-β)	PGSC0003DMG400010637	0	1503	500	6.84 / 56097.91
	PGSC0003DMG400008159	0	1503	500	7.15 / 56336.39
Lycopene epsilon cyclase (LCY-ε)	PGSC0003DMG400000333	2	531	176	5.03 / 19960.65
Violaxanthin de-epoxidase (VDE)	PGSC0003DMG400010688	4	1437	478	6.22 / 54427.16
	PGSC0003DMG400010690	2	927	308	8.57 / 35309.31
	PGSC0003DMG400020993	3	423	140	4.82 / 15702.77
Zeaxanthin epoxidase (ZEP)	PGSC0003DM400004040	15	2010	669	6.27 / 73025.23
Neoxanthin synthase (NXS)	PGSC0003DMG400024184	4	705	234	9.77 / 25860.44
	PGSC0003DMG401026407	5	750	249	9.60 / 27788.67
	PGSC0003DMG402019589	5	681	226	9.93 / 25513.05
	PGSC0003DMG400005886	0	1497	498	9.13 / 56490.70
9-cis-epoxycarotenoid dioxygenase (NCED)	PGSC0003DMG400019162	0	1530	509	5.96 / 56996.91
	PGSC0003DMG400027633	0	1521	506	5.48 / 56493.34
	PGSC0003DMG400015100	0	1743	580	6.12 / 64708.54
	PGSC0003DMG400004311	12	1824	607	5.80 / 69005.68
	PGSC0003DMG400004312	12	1842	613	6.17 / 69220.04
Carotenoid cleavage dioxygenase (CCD)	PGSC0003DMG400026032	6	1938	645	6.61 / 72869.07
	PGSC0003DMG400018480	2	1143	380	6.11 / 42748.44
	PGSC0003DMG400018481	2	1728	575	6.53 / 64702.59
	PGSC0003DMG401001968	1	723	240	5.82 / 26862.87
	PGSC0003DMG402001968	0	1029	342	8.71 / 38025.08
	PGSC0003DMG400001969	1	1776	591	6.04 / 65237.43
Geranylgeranyl pyrophosphate synthase (GGPS)	PGSC0003DMG400027856	0	1128	375	5.20 / 40657.93
	PGSC0003DMG400015673	0	1098	365	6.38 / 39977.10

probable reason for this lack of correlation might be the analysis of only one isoform of PSY and it might be possible that expression behavior of other isoform(s) of potato might be associated with carotenoid content.

Carotenoid biosynthesis genes exhibit variations in their genomic structures.

These genes of carotenoid biosynthesis of potato when analyzed for their genomic structures revealed variation in their isoforms. ZEP possessed highest i.e. 15 number of introns. ZDS and CRTISO contained 13 and 12 introns, respectively (Fig. 3, Table 2). Interestingly, of the five genes of NCED, PGSC0003DMG400004311 and PGSC0003DMG400004312 genes contained 12 introns each, whereas remaining three genes did not possess any introns. Both the genes LCY-β and GGPS each, lacked the introns (Fig. 3, Table 2). Number of introns in five phytoene synthase gene isoforms varied from 0 to 5. The presence and absence of introns and their numbers in a gene are reported to play important roles in regulating their expression in response to various intrinsic factors (related to growth and development) and external factors. Information about presence/absence of *cis- elements* such as introns is vital to understand the transcriptional behaviour of genes. This is because of the reason that in many eukaryotes, including mammals, plants, yeast, and insects, introns are reported to increase gene expression through a phenomenon termed 'intron-mediated enhancement' (Akua and Shaul, 2013; Shaul, 2017). It has been suggested that introns can increase transcript levels by affecting the rate of transcription, nuclear export, transcript stability and also can increase the efficiency of mRNA translation (Shaul, 2017). Further, the protein encoded by these genes were analysed for their secondary structure composition, and for the presence/absence of

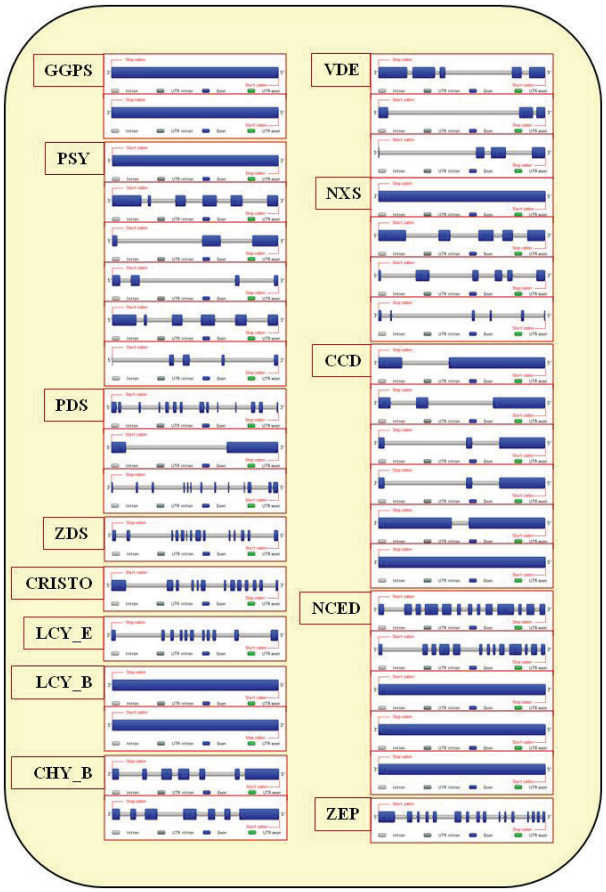

Fig.3: Structure of carotenoids biosynthesis pathway genes of potato (- exon; -intron)

signal peptide and transmembrane domain. Different protein isoforms of these 12 genes differed greatly in terms of their secondary structure composition, signal peptide and transmembrane region. PSY genes encoded six proteins contained 53 -71%, 0-8 %, and 29-42 % helix, beta sheet and random coil secondary structures, respectively (Table 3). Of the six PSY proteins, two were found to have the transmembrane regions. Similarly, CCD genes encoded six proteins contained 2 -10%, 30-44 %, and 53-62 % helix, beta sheet and random coil secondary structures, respectively (Table 4). NCED encoded five proteins did not exhibit much variations in their secondary structure composition. NXS genes encoded four proteins contained 34-

64%, 5-25 %, and 24-42 % helix, beta sheet and random coil secondary structures, respectively (Table 4). Three proteins encoded by VDE genes contained 6-39 %, 19-33 %, and 42-61 % helix, beta sheet and random coil secondary structures, respectively (Table 4). Two of the six, PSY proteins (PGSC0003DMG400021926, PGSC0003DMG400019372), both CHY-β proteins (PGSC0003DMG400010169 and PGSC0003DMG400028897) and three of the

Table 4: Predicted secondary structure, signal peptide and transmembrane content of carotenoid pathway genes encoded proteins of potato (http://www.proteus2.ca/proteus2/index.jsp)

Gene	Sequence ID	Secondary structure (%)			Signal Peptide (%)	Transmembrane (%)
		Helix	Beta Sheet	Random Coil		
Phytoene synthase (PSY)	PGSC0003DMG400021926	53	6	42	0	14
	PGSC0003DMG400016721	58	2	39	0	0
	PGSC0003DMG400024063	59	2	39	0	0
	PGSC0003DMG400029005	71	0	29	0	0
	PGSC0003DMG400012507	71	0	29	0	0
	PGSC0003DMG400019372	71	0	29	0	15
Beta-carotene hydroxylase (CHY-β)	PGSC0003DMG400010169	55	6	39	3	20
	PGSC0003DMG400028897	53	8	39	0	15
Zeta-carotene desaturase (ZDS)	PGSC0003DMG400022473	31	19	50	0	0
Carotenoid isomerase (CRTISO)	PGSC0003DMG400028224	38	18	44	0	0
Lycopene beta cyclase (LCY-β)	PGSC0003DMG400010637	30	23	46	0	0
	PGSC0003DMG400008159	31	24	45	0	0
Lycopene epsilon cyclase (LCY-ε)	PGSC0003DMG400000333	51	8	41	0	0
Violaxanthin de-epoxidase (VDE)	PGSC0003DMG400010688	39	19	42	0	0
	PGSC0003DMG400010690	20	19	61	0	0
	PGSC0003DMG400020993	6	33	61	0	0
Zeaxanthin epoxidase (ZEP)	PGSC0003DM400004040	25	21	54	0	0
Neoxanthin synthase (NXS)	PGSC0003DMG400024184	58	10	32	0	29
	PGSC0003DMG401026407	64	12	24	0	37
	PGSC0003DMG402019589	56	5	39	0	29
	PGSC0003DMG400005886	34	25	42	0	0
9-cis-epoxycarotenoid dioxygenase (NCED)	PGSC0003DMG400019162	6	39	55	0	0
	PGSC0003DMG400027633	7	37	57	0	0
	PGSC0003DMG400015100	6	38	56	0	0
	PGSC0003DMG400004311	9	35	57	0	0
	PGSC0003DMG400004312	7	34	59	0	0
Carotenoid cleavage dioxygenase (CCD)	PGSC0003DMG400026032	7	31	62	0	0
	PGSC0003DMG400018480	3	44	53	0	0
	PGSC0003DMG400018481	8	37	55	0	0
	PGSC0003DMG401001968	2	42	56	0	0
	PGSC0003DMG402001968	10	30	60	0	0
	PGSC0003DMG400001969	5	35	60	0	0
Geranylgeranyl pyrophosphate synthase (GGPS)	PGSC0003DMG400027856	61	0	39	0	0
	PGSC0003DMG400015673	61	2	36	0	0

four NXS proteins (PGSC0003DMG400024184, PGSC0003DMG402019589) were found to possess transmembrane regions (Table 4). Except for one of the proteins encoded by CHY-β gene (PGSC0003DMG400010169) all of the proteins encoded by 12 genes of carotenoid biosynthesis pathway were found to be devoid of signal peptides (Table 3). Sub-cellular localisation of the gene(s) in the cell also has been shown to affect accumulation of corresponding metabolites (Pasare *et al*., 2013), hence, information about the signal peptide in a gene indicates the probable location of the protein/ enzyme that would be encoded by that gene.

Phytoene desaturase is present as phytoene dehydrogenase

When the potato genome sequence database was searched for phytoene desaturate gene, strangely, it showed no gene/ entry. However, searching potato phytoene desaturase gene from other databases such as NCBI GenBank revealed presence of partially cloned Phytoene desaturase gene (GenBank accession number AY484445) of potato (Supplementary Fig. S1). Keeping in view this interesting and exceptional finding (lack of phytoene desaturate gene in potato genome sequence database), the partially cloned Phytoene desaturase gene available in NCBI Genbank was Blast searched for sequence similarity. Surprisingly, it was observed that the searched potato Phytoene desaturase gene showed highest per cent sequence similarity with that of phytoene dehydrogenase gene of potato (Fig. 3). This was further confirmed when proteins encoded by these two genes were searched for presence of the conserved domains. Both these proteins possessed phytoene desaturase conserved domain (Fig.4). Based on this finding, potato genome sequence database was then searched for phytoene

>AY484445.1 *Solanum tuberosum* phytoene desaturase (pds) mRNA, partial cds; nuclear gene for chromoplast product

TTTCAGTAAAATGCCTCAAATTGGACTTGTTTCTGCTGTTAATTTGAGAGTCCAAGGTAGTTCAGCTTAT
CTTTGGAGCTCGAGGTCGTCTTCTTTGGGAACTGAAAGTCGAGATGGTTGCTTGCAAAGGAATTCGTTAT
GTTTTGCTGGTAGCGAATCAATGGGTCATAAGTTAAAGATTCGTGCCACGACCAGAAGATTGGTTAAGGA
CTTGGGGCCTTTAAAGGTAGTATGCATTGATTATCCAAGACCAGAGCTAGACAATACAGTTAACTATTTG
GAGGCTGCATTCTTATCATCAACATTCCGTGCTTCTCCGCGCCCAACTAAACCATTGGAGATTGTTATTG
CTGGTGCAGGTTTGGGTGGTTTGTCTACAGCAAAATATTTGGCAGATGCTGGTCACAAACCGATACTACT
GGAGGCAAGGGATGTTCTAGGTGGAAAGGTAGCTGCATGGAAAGATGATGATGGAGATTGGTACGAGACT
GGTTTGCATATATTCTTTGGGGCTTACCCAAATATTCAGAACCTGTTTGGAGAATTAGGGATTAATGATC
GATTGCAGTGGAAGGAACATTCAATGATATTTGCAATGCCAAGTAAGCCAGGAGAATTTAGCCGCTTTGA
TTTCCCCGAAGCTTTACCCGCTCCTTTAAATGGAGTTTTGGCCATCCTAAAGAACAATGAAATGCTTACA
TGGCCAGAGAAAGTCAAATTTGCAATTGGACTCTTGCCAGCAATGCTTGGAGGGCAATCTTATGTTGAAG
CTCAAGACGGGATAAGTGTTAAGGACTGGATGAGAAAGCAAGGTGTGCCGGATAGGGTGACAGATGAAGT
GTTCATCGCCATGTCAAAGGCACTTAACTTTATAAACCCTGACGAACTGTCAATGCAGTGCATCTTGATC
GCATTGAACAGGTTTCTCCAGGAGAAACATGGTTCAAAAATGGCCTTTTAGATGGTAATCCTCCTGAGA
GACTTTGCATGCCGATTGTTGAACACATCGAGTCAAAAGGTGGCCAAGTCAGATTGAACTCACGAATAAA
AAAGATTGAGTTGAATGAGGATGGGAGTGTCAAGTGTTTTATACTGAATGACGGTAGTACAGTTGAGGGC
GATGCTTTTGTGTTTGCCACTCCAGTGGATATTTTCAAGCTGCTTTTGCCTGAAGACTGGAAAGAGATTC
CATATTTCCAAAAGTTGGAGAAGTTAGTCGGAGTACCTGTTATAAATGTACATATATGGTTCGACAGAAA
ACTGAAGAACACATATGATCATTTGCTCTTCAGCAGAAGCTCACTGCTCAGTGTGTATGCTGACATGTCT
GTCACATGTAAGGAATATTACAACCCCAATCAGTCTATGTTGGAATTGGTTTTTGCACCTGCAGAAGAGT
GGATATCTCGCAGCGACTCAGAAATAATTGATGCTACGATGAAGGAACTAGCAACACTTTTTCCTGATGA
AATTTCAGCAGATCAAAGCAAAGCAAAAATATTAAAGTATCATGTTGTCAAAACTCCAAGGTCTGTTTAT
AAAACAGTGCCAGGTTGTGAACCTTGTCGGCCATTGCAAAGATCCCCTATAGAGGGGTTTTATTTAGCCG
GTGA

Supplementary Fig. S1: Partially cloned phytoene desaturase gene from Solanum tuberosum available in NCBI GenBank (accessed on March 2023)

dehydrogenase gene, which revealed presence of three phytoene dehydrogenase genes viz: PGSC0003DMG400009156, PGSC0003DMG400026550, and PGSC0003DMG402003505. Of the three genes, PGSC0003DMG400009156 was found to be located on chromosome 3 and other two on chromosome 2 (Table 5). The sizes of these three genes were found to be ranging from 868 to 8921 basepair. Thirteen, one and twelve introns were found to be present in PGSC0003DMG400009156, PGSC0003DMG400026550, and PGSC0003DMG402003505 genes respectively (Fig. 5), and possessing ORF of 1752, 345, and 1398 nucleotides, respectively (Table 5). Phytoene dehydrogenase genes encoded three proteins contained 15-48 %, 0-92 %, and 19-43 % helix, beta sheet and random coil secondary structures, respectively. All these three phytoene dehydrogenase proteins lacked signal peptide sequence and transmembrane regions (Table 5). The result of this study suggests that thorough analysis of genes is important before concluding absence of a gene in a plant.

Supplementary Fig.S2: Pairwise alignment of nucleotide sequence cloned KJ-ZEP gene cloned from potato with that of available in Potato Genome Sequence Database (PGSC0003DMG400004020).

Fig.3: Pair-wise alignment of Phytoene Desaturase gene sequence available on NCBI GenBank (NCBI Accession No. AY484445) with that of Phytoene Dehydrogenase (NCBI Accession No. XM_006342818) of potato. The Blast analysis of Phytoene Desaturase from potato showed maximum identity (90 %) with that of the Phytoene Dehydrogenase.

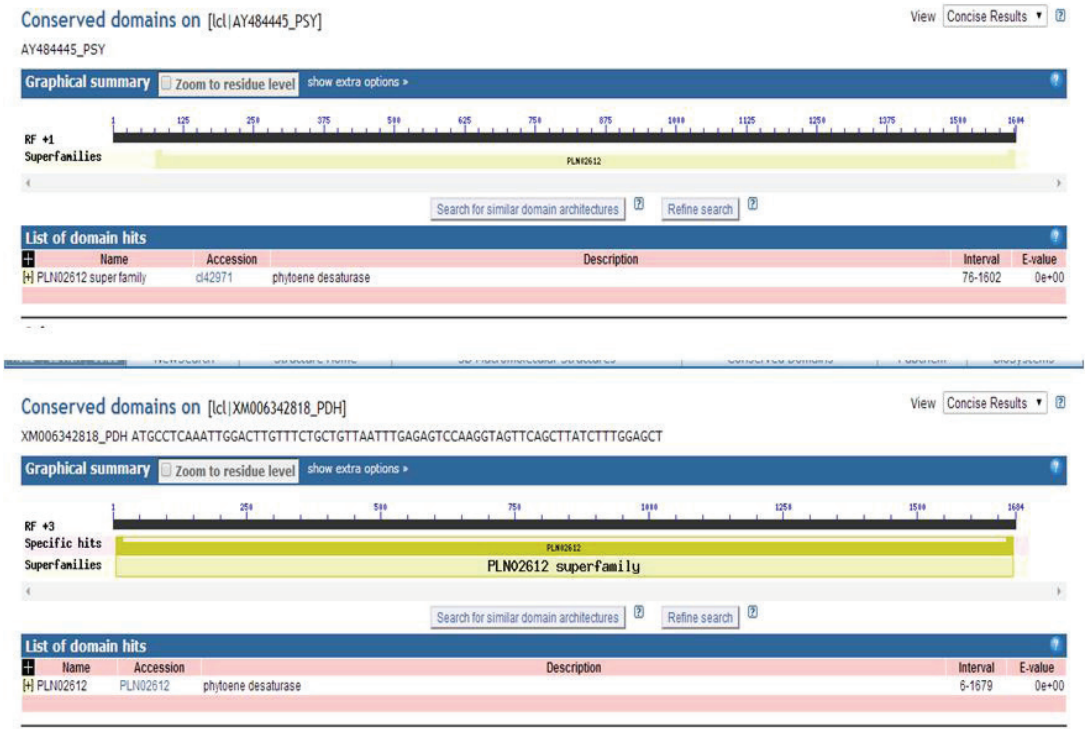

Fig.4: Conserved domain (CD) search analysis of Phytoene Desaturase gene sequence available on NCBI GenBank (NCBI Accession No. AY484445) with that of Phytoene Dehydrogenase (NCBI Accession No. XM_006342818) of potato. The CD analysis of both these genes showed presence of Phytoene Desaturase domain in both the proteins encoded by these genes and thus conforms that both the genes are similar in sequence, structure and function.

Table 5: Features of Phytoene Dehydrogenase genes of potato

Feature	Phytoene Dehydrogenase genes of potato		
	PGSC0003 DMG400009156	PGSC0003 DMG400026550	PGSC0003 DMG402003505
Chromosome Number	chr03	chr02	chr02
Start Position	61872343	24917817	40493434
End position	61879640	24918684	40502354
Size (basepair)	7298	868	8921
Number of introns	13	1	12
Open reading frame (nucleotides)	1752	345	1398
Number of amino acids in deduced protein	583	114	465
Theoretical Isoelectric point/ molecular weight (Dalton)	6.41 / 64967.98	4.25 / 12878.51	7.01 / 50960.10
Secondary structure			
(i) Helix (%)	37	92	38
(ii) Beta Sheet (%)	15	0	19
(iii) Random Coil (%)	48	8	43
Signal Peptide (%)	0	0	0
Transmembrane (%)	0	0	0

Zeaxanthin epoxidase gene cloned from Kufri Jyoti cultivar of potato exhibited sequence variations at several positions.

Furthering the work on carotenoid metabolism in potato a *Zeaxanthin epoxidase* (*ZEP*)gene was cloned from Kufri Jyoti cultivar of potato using different set of primers listed in the Table 1. As the size of the cloned genes was more than 2 kilobase pair, and thus, sequencing of the complete gene as such was not possible (the used gene sequencer gives accurate sequencing upto 600 base pair). For this reason, the overlapping fragments were sub-cloned (Fig. 6) and then sequenced in order to get the overlapping sequences required for alignment and for deriving the final complete sequence of the gene. The cloned and sequenced ZEP gene from Kufri Jyoti cultivar of potato is hereinafter referred to as KJ-ZEP. The KJ-ZEP gene sequence was submitted to NCBI GenBank and was assigned NCBI accession number MK852682. The complete ORF of cloned KJ-ZEP was 2010 nucleotides long (Fig. 7) and encoding a protein having 669 amino acids (Fig 8). It exhibited variations at 9 amino acid positions when compared with the one available in potato genome sequence database (Fig. 9, Table 6) and at 14 nucleotides (Supplementary Fig S2). Based on these variations the clone KJ-ZEP may be considered as new isoform of ZEP in potato. These variations at 9 amino acids might provide some uniqueness to the cloned gene which might be involved in regulation of carotenoid biosynthesis pathway especially beta-carotene, violxanthin, zeaxanthin carotenoids etc. Potential of zeaxanthin epoxidase as a target gene for increasing zeaxanthin and total carotenoid content has already been reported in potato (Romer *et al*., 2002).

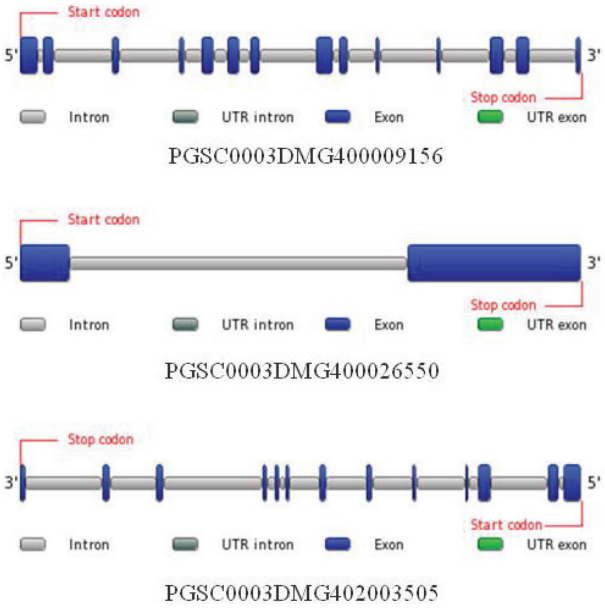

Fig.5: Structure of Phytoene Dehydrogenase genes of potato (- exon; -intron).

Table 6: Variations in amino acid sequences of cloned KJ-ZEP encoded protein when compared with the one encoded by ZEP gene available in Potato Genome Sequence Database (PGSC0003DMG400004020).

S. No.	Position of amino acid	Amino acid present in PGSC0003DMG400004020 encoded protein	Amino acid present in cloned KJ-ZEP gene encoded protein
1	8	Threonine (T)	Serine (S)
2	16	Isoleucine (I)	Valine (V)
3	17	Leucine (L)	Phenylalanine (F)
4	32	Threonine (T)	Alanine (A)
5	38	Leucine (L)	Isoleucine (I)
6	60	Isoleucine (I)	Leucine (L)
7	61	Threonine (T)	Alanine (A)
8	71	Serine (S)	Asparagine
9	643	Alanine (A)	Valine (V)

CONCLUSION

Although carotenoid biosynthesis and their regulation have been subjected to extensive studies, yet, much remains to be elucidated. Various enzymes have been shown to contribute to the regulation of carotenoids biosynthesis. Despite extensive studies, our understanding of carotenoid metabolism, regulation, and roles of carotenoid derivatives is still developing. A comprehensive understanding of the regulation of carotenoid accumulation

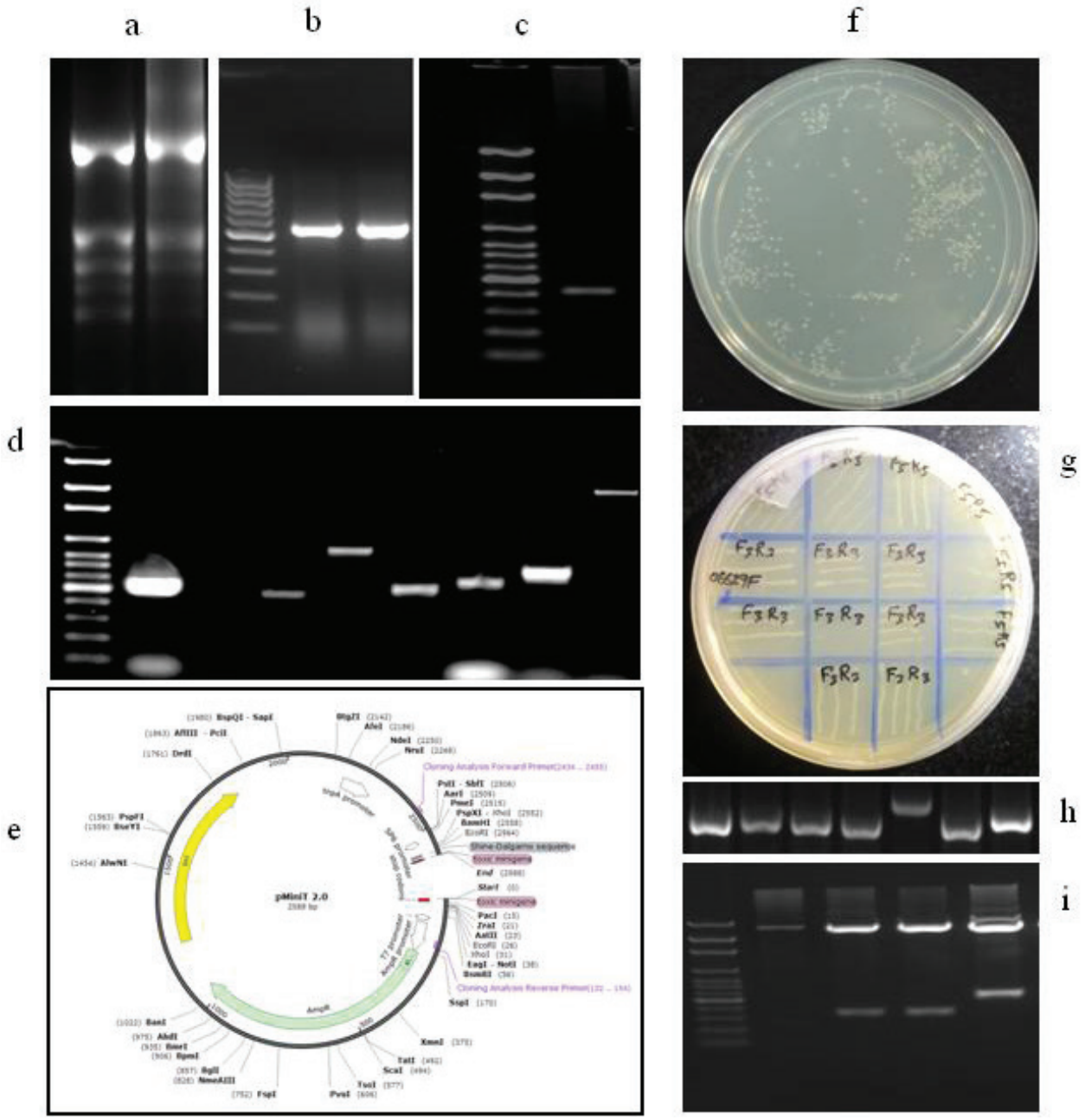

Fig.6: cloning of Zeaxanthin epoxidase (ZEP) from Kufri Jyoti cultivar of Potato. (a) RNA confirmation gel, (b) 26s rRNA confirmation (L1-100 BP ladder, L2 and L3 amplified 26s rRNA). (c): partially amplified gene (L1- 100bp ladder, L2 partially amplified amplicon primer used 06629-F-F2R2). (d): amplicons obtained upon PCR (L1- 100bp ladder, L2- 26s, L3 -06629-F-F1R1, L4-06629-F-F1R2, L5- 06629-F-F2R2, L6-06629-F3R3,L7 06629-F-F4R4,L8-06629-F-F5R5). (e): Map of vector pMiniT2.0 used in cloning. (f): Transformed colonies of the E.coli on ampicillin plate of ZEP gene,(g): colonies selected from the transformed plate on the selection media/plate. (h): agarose gel electrophoretic analysis of colony PCR. (i) Restriction digestion confirmation of inserted gene (L1 100BP ladder, L2-L5 restriction digestion).

```
ATGTATTCAACTGTGTTTTACTCTTCAGTTCATCCTTCCACTTCAGTTTTTTCAAGAAAGCA
GCTGCCTTTATTGATTTCCAAGGACTTTCCTGCAGAGTTATATCATTCTATACCTTGTAGGA
GCTTGGAAAATGGGCATATCAAGAAGGTTAAAGGAGTAAAAGTAAAAGCCACACTAGCTGAA
GCTCCGGTTACTCCTACAGAGAAGAATGACTCTGGGGTTAACGGTGATTTGAAGGTTCCACA
GAAGAAGTTGAAAGTACTTGTTGCCGGTGGTGGGATTGGAGGGTTAGTTTTTGCTTTAGCAG
CAAAGAAAAAGGGGTTTGATGTGTTGGTGTTTGAGAGGGATTTAAGTGCTATCAGAGGAGAG
GGACAATATAGAGGTCCAATTCAGATACAGAGCAATGCATTGGCTGCTTTGGAAGCAATTGA
TATGGATGTTGCTGAAGACATCATGAATGCTGGCTGCATCACTGGTCAAAGGATTAATGGCT
TGGTTGATGGTATTTCTGGCAACTGGTATTGCAAGTTTGATACGTTCACTCCAGCAGTGGAA
CGTGGACTTCCCGTGACAAGAGTCATCAGCCGCATGACTTTGCAGCAGATCCTTGCACGTGC
TGTTGGGGAGGATACAATTATGAATGAAAGTAATGTTGTAGACTTTGAGGATGATGGGGAGA
AGGTTAGTGTGGTTCTTGAGAATGGACAACGATTTACAGGTGATCTTCTGGTTGGTGCTGAT
GGCATAAGGTCTAAGGTACGGACTAATTTGTTTGGACCCAGTGAAGTTACTTACTCTGGCTA
CACTTGTTATACTGGAATTGCAGATTTTGTTCCTGCTGATATTGATACAGTTGGGTACCGAG
TCTTTTTGGGCCACAAACAGTACTTTGTTTCTTCAGATGTGGGGGGAGGCAAGATGCAGTGG
TATGCATTTTACAATGAACCAGCTGGTGGTGTGGATGCTCCAAATGGTAAAAAGGAAAGATT
GCTTAAAATATTTGGGGGATGGTGTGACAATGTCATAGACCTATTAGTTGCCACAGATGAAG
ATGCAATTCTTCGTCGTGACATCTATGATAGACCCACCAACTTTTAGTTGGGGAAGAGGTCAT
GTTACATTGCTTGGGGACTCTGTCCATGCTATGCAGCCTAATTTGGGTCAAGGGGGATGCAT
GGCCATAGAGGATAGCTATCAACTAGCACTGGAACTTGACAAAGCATGTAGTCGAAGTGCCG
AGTCAGGAAGCCCTGTGGATATCATCTCATCTTTAAGGAGCTATGAAAGTGCTAGAAAACTT
CGAGTTGGAGTCATCCATGGACTGGCTAGAATGGCTGCAATCATGGCATCTACTTACAAAGC
TTATCTTGGCGTCGGACTAGGTCCACTATCATTTTTGACGCAGTATAGAATACCACATCCTG
GAAGAGTTGGTGGAAGAGTATTTATTGACTTGGGAATGCCTCTGATGTTAAGTTGGGTTCTA
GGAGGCAATGGGGACAAGCTTGAAGGCAGAATAAAACATTGCAGGCTATCTGAGAAAGCAAA
TGACCAATTGAGAAAATGGTTTGAAGATGATGATGCATTAGAGCGTGCTACTGATGCAGAGT
GGTTACTTTTACCTGCGGGGAATGGCACTTCTGGTTTAGAAGCTATTGTTTTAAGCAGAGAT
GAGGATGTCCCTTGCACTATCGGGTCTGTCTCACATACAAACATTCCTGGAAAATCAGTAGT
TTTACCTTTGCCACAGGTGTCTGAAATGCACGCCCGAATATCCTGCAAAGACGGAGCATTTT
TTGTAACTGATTTACAAAGCGAACATGGCACCTGGGTTACAGATAATGAAGGCAGAAGATAC
CGGACGTCTCCAAACTTCCCTACACGTTTTCATCCATCAGATGTTATCGAATTTGGTTCTGA
TAAGGTAGCATTTCGTGTTAAGGCAATGAAATTTCCTCCAAAAACTACTGAAAGGAAGGAAG
AGCGTGAAGCAGTGGGGGCAGCGTAA
```

Fig. 7: Complete Open Reading Frame (ORF) of KJ-ZEP gene cloned from potato cultivar Kufri Jyoti.

```
MYSTVFYSSVHPSTSVFSRKQLPLLISKDFPAELYHSIPCRSLENGHIKKVKGVKVKATLAEAPVTP
TEKNDSGVNGDLKVPQKKLKVLVAGGGIGGLVFALAAKKKGFDVLVFERDLSAIRGEGQYRGPIQIQ
SNALAALEAIDMDVAEDIMNAGCITGQRINGLVDGISGNWYCKFDTFTPAVERGLPVTRVISRMTLQ
QILARAVGEDTIMNESNVVDFEDDGEKVSVVLENGQRFTGDLLVGADGIRSKVRTNLFGPSEVTYSG
YTCYTGIADFVPADIDTVGYRVFLGHKQYFVSSDVGGGKMQWYAFYNEPAGGVDAPNGKKERLLKIF
GGWCDNVIDLLVATDEDAILRRDIYDRPPTFSWGRGHVTLLGDSVHAMQPNLGQGGCMAIEDSYQLA
LELDKACSRSAESGSPVDIISSLRSYESARKLRVGVIHGLARMAAIMASTYKAYLGVGLGPLSFLTQ
YRIPHPGRVGGRVFIDLGMPLMLSWVLGGNGDKLEGRIKHCRLSEKANDQLRKWFEDDDALERATDA
EWLLLPAGNGTSGLEAIVLSRDEDVPCTIGSVSHTNIPGKSVVLPLPQVSEMHARISCKDGAFFVTD
LQSEHGTWVTDNEGRRYRTSPNFPTRFHPSDVIEFGSDKVAFRVKAMKFPPKTTERKEEREAVGAA
```

Fig. 8: deduced protein from cloned Zeaxanthin epoxidase (ZEP) from Kufri Jyoti cultivar of Potato (KJ-ZEP)

is critical in providing novel strategies for such endeavour. The present findings about presence of different isoforms of the carotenoids biosynthesis genes in potato and their genome-wide distribution pattern will provide deeper insights into the mechanism of carotenoid synthesis, spatial distribution and assist plant breeders and biotechnologists to devise and apply precision breeding tools for enhancing total carotenoid content or

```
PGSC0003DMG400004020    MYSTVFYTSVHPSTSILSRKQLPLLISKDFPTELYHSLPCRSLENGHIKKVKGVKVKATI    60
KJ-ZEP                  MYSTVFYSSVHPSTSVFSRKQLPLLISKDFPAELYHSIPCRSLENGHIKKVKGVKVKATL    60
                        *******:********::**************:****:********************:

PGSC0003DMG400004020    TEAPVTPTEKSISGVNGDLKVPQKKLKVLVAGGGIGGLVFALAAKKKGFDVLVFERDLSA    120
KJ-ZEP                  AEAPVTPTEKNISGVNGDLKVPQKKLKVLVAGGGIGGLVFALAAKKKGFDVLVFERDLSA    120
                        :.********.**************************************************

PGSC0003DMG400004020    IRGEGQYRGPIQIQSNALAALEAIDMDVAEDIMNAGCITGQRINGLVDGISGNWYCKFDT    180
KJ-ZEP                  IRGEGQYRGPIQIQSNALAALEAIDMDVAEDIMNAGCITGQRINGLVDGISGNWYCKFDT    180
                        ************************************************************

PGSC0003DMG400004020    FTPAVERGLPVTRVISRMTLQQILARAVGEDTIMNESNVVDFEDDGEKVSVVLENGQRFT    240
KJ-ZEP                  FTPAVERGLPVTRVISRMTLQQILARAVGEDTIMNESNVVDFEDDGEKVSVVLENGQRFT    240
                        ************************************************************

PGSC0003DMG400004020    GDLLVGADGIRSKVRTNLFGPSEVTYSGYTCYTGIADFVPADIDTVGYRVFLGHKQYFVS    300
KJ-ZEP                  GDLLVGADGIRSKVRTNLFGPSEVTYSGYTCYTGIADFVPADIDTVGYRVFLGHKQYFVS    300
                        ************************************************************

PGSC0003DMG400004020    SDVGGGKMQWYAFYNEPAGGVDAPNGKKERLLKIFGGWCDNVIDLLVATDEDAILRRDIY    360
KJ-ZEP                  SDVGGGKMQWYAFYNEPAGGVDAPNGKKERLLKIFGGWCDNVIDLLVATDEDAILRRDIY    360
                        ************************************************************

PGSC0003DMG400004020    DRPPTFSWGRGHVTLLGDSVHAMQPNLGQGGCMAIEDSYQLALELDKACSRSAESGSPVD    420
KJ-ZEP                  DRPPTFSWGRGHVTLLGDSVHAMQPNLGQGGCMAIEDSYQLALELDKACSRSAESGSPVD    420
                        ************************************************************

PGSC0003DMG400004020    IISSLRSYESARKLRVGVIHGLARMAAIMASTYKAYLGVGLGPLSFLTQYRIPHPGRVGG    480
KJ-ZEP                  IISSLRSYESARKLRVGVIHGLARMAAIMASTYKAYLGVGLGPLSFLTQYRIPHPGRVGG    480
                        ************************************************************

PGSC0003DMG400004020    RVFIDLGMPLMLSWVLGGNGDKLEGRIKHCRLSEKANDQLRKWFEDDDALERATDAEWLL    540
KJ-ZEP                  RVFIDLGMPLMLSWVLGGNGDKLEGRIKHCRLSEKANDQLRKWFEDDDALERATDAEWLL    540
                        ************************************************************

PGSC0003DMG400004020    LPAGNGTSGLEAIVLSRDEDVPCTIGSVSHTNIPGKSVVLPLPQVSEMHARISCKDGAFF    600
KJ-ZEP                  LPAGNGTSGLEAIVLSRDEDVPCTIGSVSHTNIPGKSVVLPLPQVSEMHARISCKDGAFF    600
                        ************************************************************

PGSC0003DMG400004020    VTDLQSEHGTWVTDNEGRRYRTSPNFPTRFHPSDVIEFGSDKAAFRVKAMKFPPKTTERK    660
KJ-ZEP                  VTDLQSEHGTWVTDNEGRRYRTSPNFPTRFHPSDVIEFGSDKVAFRVKAMKFPPKTTERK    660
                        *****************************************.******************

PGSC0003DMG400004020    EEREAVGAA          669
KJ-ZEP                  EEREAVGAA          669
                        *********
```

Fig. 9: Pairwise alignment of the amino acid sequence of protein deduced from KJ-ZEP gene cloned from potato with that of available in Potato Genome Sequence Database (PGSC0003DMG400004020).

increasing the targeted carotenoid such as beta-carotene, violaxanthin, zeaxanthin, neoxanthin, lutein *etc*. in potato. Potato being third most important food crop after rice and wheat will have higher impacts of nutritional improved cultivar consumptions on consumers' health.

ACKNOWLEDGEMENT

Authors thank Indian Council of Agricultural Research for funding the projects entitled "Post harvest management and nutritional enhancement of potatoes (IXX15884)" and "Genetic enhancement of potato through molecular and genomic tools (IXX15910)".

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest

ETHICAL STATEMENT

This article does not contain any studies with human participants or animals performed by any of the authors

LITERATURE CITED

Akua T, Shaul O (2013). The *Arabidopsis thaliana* MHX gene includes an intronic element that boosts translation when localized in a 5′ UTR intron. J Exp Bot. 64:4255-4270.

Bahadori S, Giglou MT, Esmaielpour B, Dehdar B, Estaji A, Hano C, Gohari G, Vergine M, Vita F (2023). Antioxidant compounds of potato breeding genotypes and commercial cultivars with yellow, light Yellow, and white flesh in Iran. *Plants*. 12:1707.

Bonierbale, M., W. Grüneberg, W. Amoros, G. Burgos, E. Salas, E. Porras, and T. ZumFelde (2009). Total and individual carotenoid profiles in *Solanum phureja* cultivated potatoes: II. Development and application of near-infrared reflectance spectroscopy (NIRS) calibrations for germplasm characterization. *Journal of Food Composition and Analysis* 22: 509-516.

Brown CR, Culley D, Yang C, Durst R, Wrolstad R (2005). Variation of anthocyanins and carotenoid contents and associated antioxidant values in potato breeding lines. *Journal of the American Society for Horticultural Science* 130:174-80.

Burgos G, Salas E, Amoros W, Auqui M, Munoa L., Kimura M, Bonierbale M (2009). Total and individual carotenoid profiles in *Solanum phureja* of cultivated potatoes: I. Concentrations and relationships as determined by spectrophotometry and HPLC. *Journal of Food Composition and Analysis* 22: 503-508.

Buscemi S, Corleo D, Di Pace F, Petroni ML, Satriano A, Marchesini G (2018). The effect of lutein on eye and extra-eye health. Nutrients. 10:1321.

Cazzonelli CI, Pogson BJ (2010). Source to sink: regulation of carotenoid biosynthesis in plants. *Trends Plant Sci*.15:266-274.

Cazzonelli CI (2011). Carotenoids in nature: insights from plants and beyond. *Funct. Plant Biol*. 38:833-847.

Diretto G, Al-Babili S, Tavazza R, Papacchioli V, Beyer P, Giuliano G (2007). Metabolic engineering of potato carotenoid content through tuber-specific overexpression of a bacterial mini-pathway. *PLoS ONE* 2: e350.

Diretto G, Welsch R, Tavazza R, Mourgues F, Pizzichini D, Beyer P, Giuliano G (2007). Silencing of beta-carotene hydroxylase increases total carotenoid and beta-carotene levels in potato tubers. *BMC Plant Biol* 2:7-11.

Ducreux LJ, Morris WL, Hedley PE, Shepherd T, Davies HV, Millam S, Taylor MA (2005). Metabolic engineering of high carotenoid potato tubers containing enhanced levels of beta-carotene and lutein. *J Exp Bot* 56:81-89.

Eggersdorfer M, Wyss A (2018). Carotenoids in human nutrition and health. Arch BiochemBiophys. 652:18-26.

Fiedor J, Burda K (2014). Potential role of carotenoids as antioxidants in human health and disease. *Nutrients* 6:466-488.

Hamouz KK, Pazderu J, Lachman J, Cepl, and Z. Kotikova (2016). Effect of cultivar, flesh colour, locality and year on carotenoid content in potato tubers. *Plant, Soil and Environment* 62 (2): 86–91.

Hamouz K, Pazderu K, Lachman J, Cepl J, Kotikova Z (2016). Effect of cultivar, flesh colour, locality and year on carotenoid content in potato tubers. *Plant Soil Environ*. 62:86-91.

Kotikova Z, Hejtmankova A, Lachman J, Hamouz K, Trnkova E, Dvorak P (2007). Effect of selected

factors on total carotenoid content in potato tubers (*Solanum tuberosum* L.). *Plant Soil and Environment 53*: 355.

Leman, JK, Mueller R, Karakas M, Woetzel N, Meiler J (2013). Simultaneous prediction of protein secondary structure and transmembrane spans. *Proteins* 81:1127-1140.

Lima VC, Rosen RB, Farah MI (2016). Macular pigment in retinal health and disease. *Int. J. Retina Vitreous*.2:19.

Van Bel M, Silvestri F, Weitz EM, Kreft L, Botzki A, Coppens F, Vandepoele K (2022). PLAZA 5.0: extending the scope and power of comparative and functional genomics in plants. *Nucleic Acids Res*.50:D1468-D1474.

Montgomerie S, Cruz JA, Shrivastava S, Arndt D, Berjanskii M, Wishart DS (2008). PROTEUS2: a web server for comprehensive protein structure prediction and structure-based annotation. *Nucleic Acids Res*. 36(Web Server issue):W202-9.

Morris WL, Ducreux LJ, Fraser PD, Millam S, Taylor MA (2006). Engineering ketocarotenoid biosynthesis in potato tubers. *Metab Eng* 8:253-263.

Pasare S, Wright K, Campbell R, Morris W, Ducreux L, Chapman S, Bramley P, Fraser P, Roberts A, Taylor M (2013). The sub-cellular localisation of the potato (*Solanum tuberosum* L.) carotenoid biosynthetic enzymes, CrtRb2 and PSY2. *Protoplasma* 250:1381-1392.

Potato Genome Sequencing Consortium; Xu X, Pan S, Cheng S, Zhang B, Mu D, Ni P, Zhang G, Yang S, Li R, Wang J, Orjeda G, Guzman F, Torres M, Lozano R, Ponce O, Martinez D, De la Cruz G, Chakrabarti SK, Patil VU, Skryabin KG, Kuznetsov BB, Ravin NV, Kolganova TV, Beletsky AV, Mardanov AV, Di Genova A, Bolser DM, Martin DM, Li G, Yang Y, Kuang H, Hu Q, Xiong X, Bishop GJ, Sagredo B, Mejía N, Zagorski W, Gromadka R, Gawor J, Szczesny P, Huang S, Zhang Z, Liang C, He J, Li Y, He Y, Xu J, Zhang Y, Xie B, Du Y, Qu D, Bonierbale M, Ghislain M, Herrera Mdel R, Giuliano G, Pietrella M, Perrotta G, Facella P, O'Brien K, Feingold SE, Barreiro LE, Massa GA, Diambra L, Whitty BR, Vaillancourt B, Lin H, Massa AN, Geoffroy M, Lundback S, DellaPenna D, Buell CR, Sharma SK, Marshall DF, Waugh R, Bryan GJ, Destefanis M, Nagy I, Milbourne D, Thomson SJ, Fiers M, Jacobs JM, Nielsen KL, Sønderkær M, Iovene M, Torres GA, Jiang J, Veilleux RE, Bachem CW, de Boer J, Borm T, Kloosterman B, van Eck H, Datema E, HekkertBt, Goverse A, van Ham RC, Visser RG (2011). Genome sequence and analysis of the tuber crop potato. *Nature* 475:189-195.

Rodríguez-Concepción M, Boronat A (2002). Elucidation of the methylerythritol phosphate pathway for isoprenoid biosynthesis in bacteria and plastids. A metabolic milestone achieved through genomics. *Plant Physiol*. 130:1079-1089.

Romer S, Lubeck J, Kauder F, Steiger S, Adomat C, Sandmann G (2002). Genetic engineering of a zeaxanthin-rich potato by antisense inactivation and co-suppression of carotenoid epoxidation. *Metab Eng* 4:263-272.

Ruiz-Sola MA, Rodríguez-Concepcion M (2012). Carotenoid biosynthesis in Arabidopsis: a colorful pathway. *Arabidopsis Book*. 10:e0158.

Shaul O (2017). How introns enhance gene expression. *Int J Biochem Cell Biol*. 91:145-155.

Tapiero H., Townsend D.M., Tew K.D (2004). The role of carotenoids in the prevention of human pathologies. Biomed. Pharmacother. 58:100-110

Tatarowska B, Milczarek D, Wszelaczyńska E, Pobereżny J, Keutgen N, Keutgen AJ, Flis B (2019). Carotenoids variability of potato tubers in relation to genotype, growing location and year. *American Journal of Potato Research 96*:493-504.

Valcarcel J, Reilly K, Gaffney M, O'Brien NM (2016). Levels of potential bioactive compounds including carotenoids, vitamin C and phenolic compounds, and expression of their cognate biosynthetic genes vary significantly in different varieties of potato (*Solanum tuberosum* L.) grown under uniform cultural conditions. *J Sci Food Agric*. 96:1018-1026.

Walter MH, Strack D (2011). Carotenoids and their cleavage products: biosynthesis and functions. *Nat Prod Rep*. 28:663-692.

Xu J, Li Y, Kaur L, Singh J, Zeng F (2023). Functional Food Based on Potato. *Foods* 12: 2145.

MS Received : June 22, 2023; Accepted : September 13, 2023