Cereal Research Communications
https://doi.org/10.1007/s42976-022-00279-w

ORIGINAL PAPER

In silico characterization of *Thinopyrum elongatum*-derived *PsyE1* gene and validation in 7D/7E bread wheat introgression lines open avenues for carotenoid biofortification in wheat

Asish Kumar Padhy[1,2] · Parampreet Kaur[1] · Baljinder Singh[2] · Ravinder Kaur[1] · Sabhyata Bhatia[2] · M. Shamshad[1] · Himanshu Sharma[1] · Satinder Kaur[1] · Puja Srivastava[1] · Achla Sharma[1]

Received: 18 January 2022 / Accepted: 17 May 2022

Abstract

Current global scenario demands agricultural productivity of food grains to be kept at abreast with burgeoning population. Cereals constitute major food stuff for millennia and biofortification of new cereal varieties provides an opportunity to tackle global-scale malnutrition deficiencies without doing major shifts in the diets. Carotenoid biofortification in wheat grains has recently caught the attention of breeders owing to a myriad of health benefits offered by this micronutrient. *Thinopyrum elongatum*-derived *PsyE1* gene encoding for Phytoene Synthase encoding Y gene, is a jackpot to enhance the carotenoid content in wheat. The present study is the first report deciphering detailed in silico characterization of *Thinopyrum elongatum*-derived *PsyE1* gene and its protein. Promoter analysis of chloroplast localized *PsyE1* gene provides clues about its possible role in stress resistance along with enhancing the carotenoid content in both durum and bread wheat. Homology, phylogeny and protein modelling studies of *PsyE1* revealed its closer evolutionary relationship with barley and wheat, as well as provided a preliminary insight into catalytic and secondary structure of the protein. PCR validation of *PsyE1* in 7D/7E bread wheat introgression lines further facilitated development of functional marker that could be used to track its introgression in elite bread wheat varieties. Overall, these detailed *insilico* insights into structure, function and validation of *PsyE1* open doors for its deployment in to produce carotene biofortified hexaploid wheat through facilitating development of functional markers and MAS, as well as to elucidate its mechanism of action and regulation in response to external stimuli.

Keywords *PsyE1* · *Thinopyrum elongatum* · Biofortification · *Insilico* analysis · Wheat

Abbreviations

Psy	Phytoene synthase
Y gene	Yellow pigment gene; later recognized as *PsyE1* gene

Communicated by Molnár-Láng.

✉ Achla Sharma
achla@pau.edu

Asish Kumar Padhy
asishpadhy999@gmail.com

Parampreet Kaur
parampreet162906@pau.edu

Baljinder Singh
baljinder@nipgr.ac.in

Ravinder Kaur
gillravinder1990@gmail.com

Sabhyata Bhatia
sabhyatabhatia@nipgr.ac.in

M. Shamshad
shamshad.rattan@gmail.com

Himanshu Sharma
himanshu.pau@gmail.com

Satinder Kaur
satinder.biotech@pau.edu

Puja Srivastava
pujasrivastava@pau.edu

[1] Punjab Agricultural University, Ferozpur Road, Ludhiana, Punjab 141004, India

[2] National Institute of Plant Genome Research, New Delhi 110067, India

Published online: 09 June 2022

PCR	Polymerase chain reaction
GYPC	Grain yellow pigment content
NCBI	National center for biotechnology information
QTL	Quantitative trait loci
Ils	Introgression lines
CDD	Conserved domain database

Introduction

Wheatgrass genus *Thinopyrum* Löve (Löve 1982; Dewey 1984), a tertiary gene pool wheat relative, houses numerous desirable traits (abiotic and biotic stress, quality, and yield, etc.) for introgression and improvement of cultivated *Triticum* species. This genus got separated from another genus *Elytrigia* in the later twentieth century alongside genera *Lophopyrum* and *Trichopyrum* (Baum and Johnson 2018). *Thinopyrum* species with ploidy ranging from di- to deca-ploidy have been identified (Chen et al 2013; Guo et al 2016; Mao et al 2010); however, the most extensively exploited species targeted for enrichment of cultivated wheat germplasm includes *Th. elongatum* (2n = 2x = 14) (Host) D.R. Dewey [syn. *Agropyron elongatum* (Host) P. Beauv., *Lophopyrum elongatum* (Host) A. Löve] and *Th. bessarabicum* (Savul. & Rayss) Löve [syn. *Agropyron bessarabicum* (Savul. & Rayss)], *Th. intermedium* (2n = 6x = 42) (Host) Barkworth & D.R. Dewey [syn. *Agropyron intermedium* (Host) P. Beauv.), *Elytrigia intermedia* (Host) Nevski] and Th. ponticum (2n = 10x = 70) (Popd.) Barkworth & D.R. Dewey [syn. *Agropyron elongatum* (Host) P. Beauv., *Lophopyrum ponticum* (Popd.) A. Löve, *Elytrigia pontica* (Popd.) Holub]. However, the genomic compositions of the polyploid species of this genera are quite debated, with diploid (Ee) *Th. Elongatum* and *Th. Bessarabicum* (Eb, Ee, Ex or J) (Jauhar 1990; Wang 1992, 2011) along with *Pseudoroegneria strigosa* (St or S, comprising a core genome in *Thinopyrum* perennial species with close relatedness to Ea and Eb genomes) (Cai and Jones 1997; Zhang et al 1996; Wang 2010) to be the donors of polyploid *Th. intermedium* and *Th. elongatum*. *Th. elongatum* and *Th. ponticum* were frequently placed under the same name, *Agropyrum elongatum* (Shepherd and Islam 1988; Li et al. 2017). Different genome compositions for *Th. ponticum*, i.e. EbEbEbEbEbEbEbEbEbEb (Arterburn et al. 2011), JJJJJJJSJSJSJS (Chen et al. 1998) and EeEeEbEbExExStStStSt (Zhang et al. 1996) have been reported.

Most of the genes/QTLs of demonstrated or prospective breeding value of *Thinopyrum* are housed on the long arm of a chromosome 7, named as 7Ag (Sears 1973) or 7el (Sharma and Knott 1966) and are widely used in distant hybridization studies in bread wheat owing to the close homology of its three subgenomes (A, B and D) with those of diploid E genome. For instance, development of 7D/7E substitution lines through Robertsonian translocation in bread wheat (Wang et al. 2020). *Thinopyrum* genus has been used extensively in wheat resistance breeding programs conferring resistance against leaf rust (*Lr19, Lr24, Lr29* and *Lr38*), stem rust (*Sr24, Sr25, Sr26, Sr43, Sr44* and *SrB*), powdery mildew (*Pm40* and *Pm43*) and Fusarium head blight (*Fhb7*) as well as grain pigment content and yield (Ceoloni et al. 2014; Wang et al. 2020). Sequencing of its diploid E genome (Wang et al. 2020; NCBI BioProjectID PRJNA540081) has paved the way forward to plethora of avenues for analysing its economically vital genes and their detailed characterization for transferring to elite hexaploid wheat cultivars.

Since cereals are characterized by low carotenoid content, their biofortification to meet the daily human requirements offers an opportunity to combat wide-scale malnutrition deficiencies without changing the diet structures of population. This requires a comprehensive understanding of various genetic and genomic loci imparting grain yellow pigment content (GYPC) so as to allow molecular breeders to engineer nutritionally important carotenoid pathway in wheat. Several genes of carotenoid biosynthetic pathway have been targeted towards its improved content in wheat grains. In *Triticeae*, major locus affecting GYPC has been mapped in several species onto distal ends of long arms of homeologous group 7 (Elouafi et al 2001; Pozniak et al 2007a, b; Atienza et al 2007). *Psy* (Phytoene synthase) gene has been known to catalyze first committed/rate limiting step of carotenoid biosynthesis to produce 15-cis-phytoene from condensation of 2 molecules of geranyl geranyl pyrophosphate (GGPP); and this gene has been shown to be linked with 7B QTL for GYPC (Gallagher et al 2004; Pozniak et al 2007a, b).

Distal end of 7EL chromosome in *Th. elongatum* has also been known to control GYPC (Knott 1968; Zhang et al 2005) as well as possess leaf rust resistance gene Lr19, stem rust resistance gene Sr25 (Knott 1989) and Ug99 (Li and Wang 2009; Liu et al. 2010). In an initial attempt to introgress Lr19 gene from decaploid *Th. Ponticum*, 7Ag (7el1) whole chromosome substitution lines, Argus were developed in durum wheat background (Sharma and Knott 1966), followed by several induced translocations and homeologous exchanges in both common and durum wheat (Gennaro et al 2009 and references therein). However, close association of Lr19 with yellow pigment gene (and linkage drag) and targeted selection for white flour bread wheat varieties in initial breeding programs called for the elimination of negative effects of yellow pigment locus (Y locus, later designated as Psy-E1), consequently developed mutant lines (Agatha-28) (Knott 1980) with non-functional Psy-E1 (Zhang and Dubcovsky 2008) and their utilization to breed rust resistant wheat varieties such as 'Wheatear' (Bariana et al 2007) under CIMMYT breeding programmes. Nevertheless, with an overall increase in health awareness and acceptance of

yellow wheat grains among consumers, breeding for biofortified cereals gained momentum. Detailed characterization of GYPC in tetraploid 7AL/7EL translocation lines, suggested the presence of one additional gene along with PsyE1 in the distal region of homeologous group 7 to mediate yellow pigment content in wheat (Ceoloni et al. 2000; Zhang and Dubcovsky 2008). Subsequent identification of QTLs for GYPC on other homeologous groups (2D, 4D, 4A, 5A, 2A, 4B, 6B), suggested multigenic control of this trait, with major genes located on group 7 chromosomes (Zhang et al. 2005; Pozniak et al. 2007a, b). Among various carotenoids conferred by *Psy-E1*, lutein has been observed to be the most abundant type. Lutein, an antioxidant, reduces the oxidative damage to biological membranes by scavenging peroxi-radicals causing various human diseases (Zakynthinos and Varzakas 2016), aging process and degeneration of food quality. Plants heterozygous for the 7EL segment were observed to develop 56% higher lutein content than the lines without 7EL segment (Zhang et al. 2005). With recent resurgence in nutritional breeding or biofortification of staples as a long term, eco-friendly and sustainable remedial to counter malnutrition, present study takes into account the in silico characterization of *PsyE1* gene, its interaction with other traits along, domain prediction, protein modelling and phylogeny, an in-depth characterization of which is still unavailable in the public domain along with the validation of *PsyE1* in various common wheat introgression lines. The study thus holds significance to open plethora of avenues for *PsyE1* utilization in wheat breeding programs.

Material and methods

Th. ponticum PsyE1 nucleotide sequence was retrieved from NCBI (https://www.ncbi.nlm.nih.gov/) and Blast searched (BLASTn, E value of e^{-10}) against *Th. elongatum* genome. This was followed by gene prediction using FGENESH (http://www.softberry.com/berry.phtml?topic = fgenesh & group = programs & subgroup = gfind) based on monocots specific gene models training sets. Predicted nucleotide sequence was fetched using Bedtools and subsequent removal of five intron sequences to construct mature mRNA sequence and extraction of 3' and 5' UTRs was carried out. In silico protein structure was decoded using ExPASy (https://web.expasy.org/translate/) (default parameters) and compared with the predicted protein and mRNA sequence of the *Th. elongatum* sequences for confirmation. Physical and chemical characterization of protein and domain analysis was carried out using ProtParam (https://web.expasy.org/protparam/); NCBI-CDD (https://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml) and SMART (http://smartembl-heidelberg.de/smart/setmode.cgi?NORMAL = 1)) databases, respectively. Subcellular localization of the protein was deciphered using ProtComp 9.0 tool (http://www.softberry.com/berry.phtml?topic=protcomppl&group=programs&subgroup=proloc). Further, *PSY* protein homologs were searched through BLASTp search in NCBI database and were aligned and visualized using Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/), NCBI Multiple sequence Alignment viewer, version 1.14.0 (https://www.ncbi.nlm.nih.gov/projects/msaviewer/), respectively. Phylogenetic tree of the *PsyE1* homologs was constructed (bootstrap—1000) using MegaX (Sudhir et al. 2018) and edited using iTOL (https://itol.embl.de/). 3D structure of the protein was predicted and modelled using SWISS-MODEL (https://swissmodel.expasy.org/interactive) and Phyre^2 (http://www.sbg.bio.ic.ac.uk/~phyre2/html/page.cgi?id=index) and subsequently validated through analysis of Ramachandran plot (http://mordred.bioc.cam.ac.uk/~rapper/rampage.php)) and edited through Pymol (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.). For promoter characterization, 812 bp sequence upstream of the TSS (Transcription Start Site), the region between the TSS of *PsyE1* and polyA tail of the previous gene was subjected to PlantCARE (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/), PlantPAN3.0 (http://plantpan.itpsncku.edu.tw/promoter.php) and New PLACE (https://www.dna.affrc.go.jp/PLACE/?action=newplace)) databases were utilized for identification of the motifs and cis-acting elements.

Finally, for validation of the predicted gene structure and to develop its specific marker, a set of primer pair (Forward sequence: GTACCAGATAGCCGGTGCAT and Reverse sequence: GCCGACAAACGAAACAATTT) was designed from the 5' UTR of the gene using Primer 3 to produce an amplicon of 184bps. Primer pairs were custom synthesized using the facilities of National Institute of Plant Genome Research, New Delhi and were amplified in genomic DNA isolated from selected 7D/7E bread wheat introgression lines (ILs) of BC_1F_{10} bread wheat, containing 12 ppm GYPC on the higher end range and 3 ppm GYPC on the lower end range. These ILs were developed using HD2967 (ALD/COC//URES/3/HD2160 M/HD 2278), a high yielding wheat variety developed and released by Indian Agricultural Research Institute (IARI), New Delhi as recipient parent. Within three years since its release in 2011 by Indian Agricultural Research Institute (IARI), New Delhi, HD2967 covered much of the area under wheat in India. This variety has an average yield of 5.1 t/ha (potential yield 6.6 t/ha) with bold, round, lustrous and amber coloured grains having a protein content of 12.7%. The donor lines containing the *PsyE1* gene was UC3 having *PsyE1* from *Th. elongatum*. The introgression lines were developed by crossing both the parents (F_1) followed by single back cross with HD2967 (BC_1F_1) and maintained further up to BC_1F_{10} generation by selfing with ear to row progeny method. PCR amplification

was performed in 20-µl aliquots in an ABI (Applied Biosystems) thermal cycler using the following reaction mix: 13.7 µl water, 2 µl of 10X PCR buffer, 1 µM of forward and reverse primer (0.2 µl) each 0.25 µl of commercial Taq Polymerase, 1.6 µl of 25 mM MgCl2, 0.1 µl of 100 mM dNTPs and 2 µl of genomic DNA (~20 ng/µl). PCR conditions were: initial denaturation at 94ºC for 4 min, followed by 35 cycles of 94ºC for 1 min, Ta (58 ºC) for 1 min and 72ºC for 1 min and a final step of extension at 72ºC for 7 min. After amplification, 10 µl of each sample was loaded in 1.5% agarose TBE (0.5X) gels supplemented with ethidium bromide.

Results

Identification and characterization of *PsyE1* gene structure, protein modelling and domain prediction

Genome wide homology-based search for *Th. elongatum PsyE1* gene utilizing already reported *Psy E1* gene in *Th. ponticum PsyE1* gene (ID: EU096095.1, Zhang and Dubcovsky 2008), identified the gene on 7E chromosome (Position GWHABKY00000007; Gene ID GWHABKY043860; Transcript ID Tel7E01G983800.1) (SF1; Fig 1) with 89% identity [Supplementary File 1 (SF1); Fig 2]. Gene sequence was fetched (coordinates: upstream -722414600 bp; downstream-722418823bp) and gene prediction analysis delineated a 4092 bp gene on the negative strand containing 6 coding sequences within it, similar to that of *Th. ponticum PsyE1* gene (Supplementary file 1, Fig 3). The detailed information about various gene and protein structural features of *PsyE1* is given in Table 1. Following the reverse complementation of the extracted nucleotide sequence, removal of the introns and the UTRs, the complete coding sequence of the gene was constructed and 426 amino acid residue containing protein product was translated *in silico*. The translated protein was observed to possess 3 amino acid residues [M (Met), Q(Gln) and H (His)] less at N terminal relative to the protein structure predicted by Wang et al. (2020).

Predicted protein was subjected to *in silico* domain prediction (SF1; Fig 4). Apart from 59 and 52, positive and negatively charged amino acid residues, respectively, several hydrophilic and hydrophobic regions were predicted inside protein structure. Protein was categorized in the unstable category with an instability index (II) of 59.08. Two aspartate rich regions comprising 5 (at position 175–179) and 6 (at position 301–305) amino acids residues were predicted and were observed to be the binding sites for magnesium ions (Fig. 1B). Phytoene synthase domain (PLN02632) was predicted at the interval of 124–420 amino acids residues while Trans-isoprenyl Diphosphate synthase-head to head domain (Trans_IPPS_HH) was predicted at amino acid residue 136 till amino acid residue 403. Two domains for phytoene/squalene synthetase and phytoene/squalene synthase (ERG9 and SQS_*PSY*) which help in lipid metabolism and transport were found between residues 133–414 and 145–402 and all these domains fall under Isoprenoid Biosynthesis C1 super family. Some non-specific hits for squalene

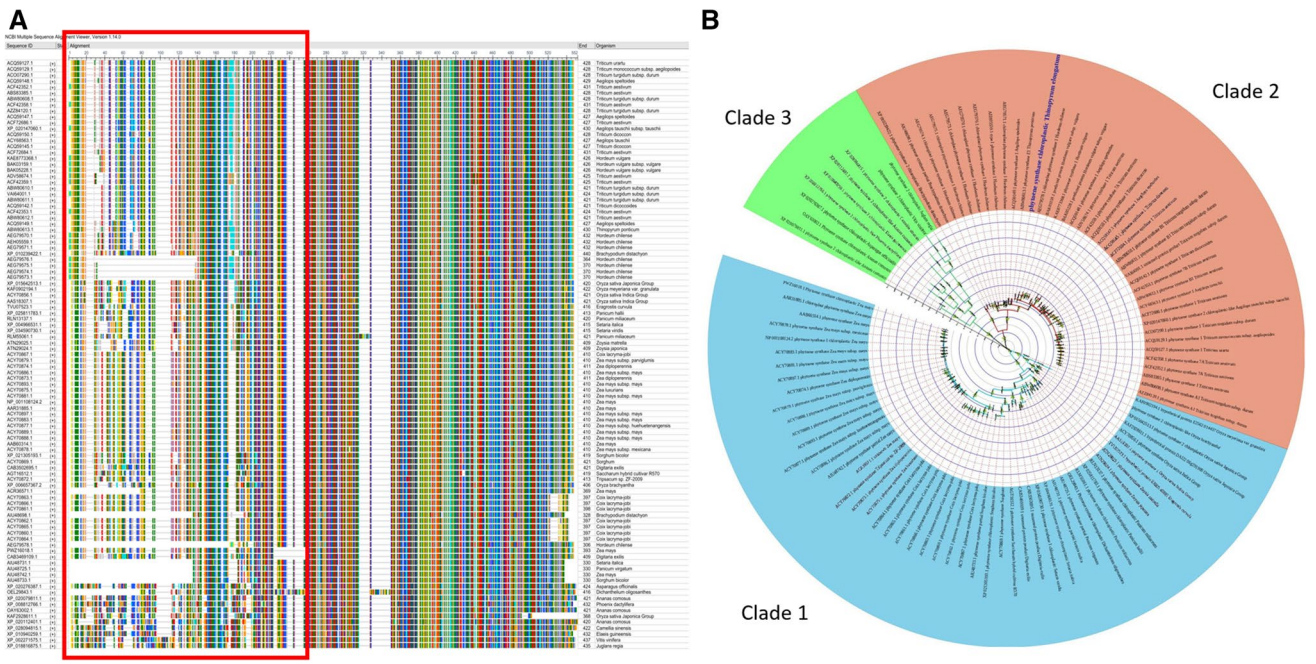

Fig. 1 Alignment of PsyE1 Protein homologs showing diversity in 5' UTR and 1st exonic region shown in red coloured box while it is conserved in later regions (**A**) Phylogeny analysis of PSY homologs (**B** PsyE1 of *Th. elongatum* is marked in blue)

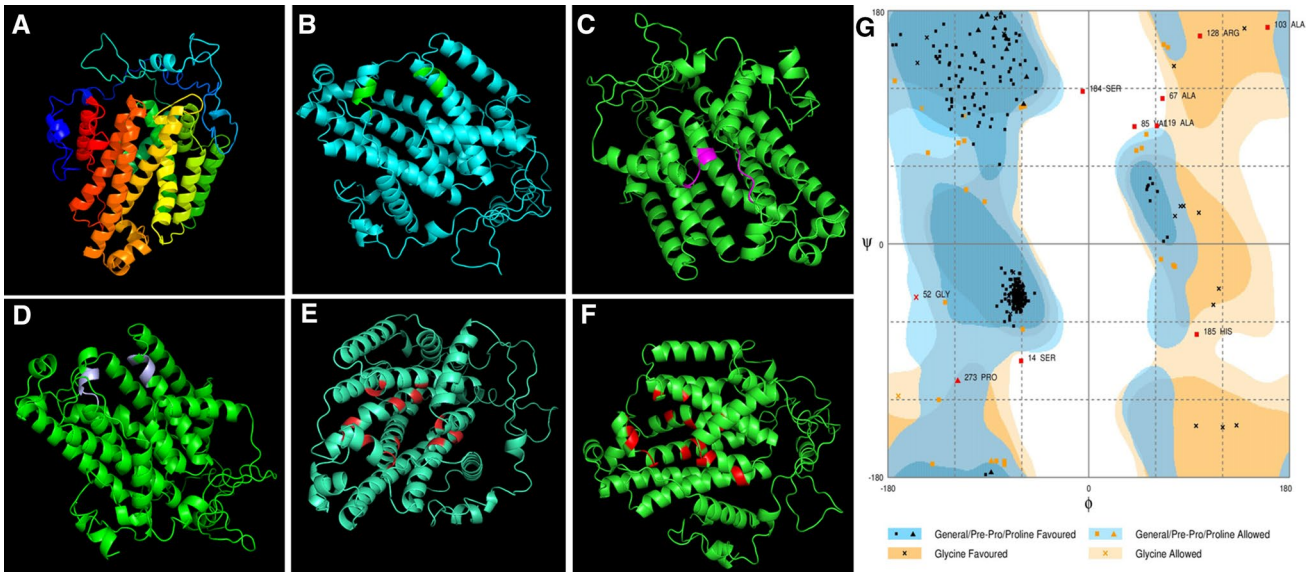

Fig. 2 **A** Predicted protein model of PsyE1 in Phyre2, **B** Magnesium Binding sites, **C** Active-site lid residues, **D** Aspartate rich regions, **E** Catalytic residues **F** Substrate binding pockets, **G** Ramachandran plot for PsyE1; Number of residues in favoured region: 387 (91.3%); Number of residues in allowed region: 27 (6.4%) and Number of residues in outlier region: 10 (2.4%)

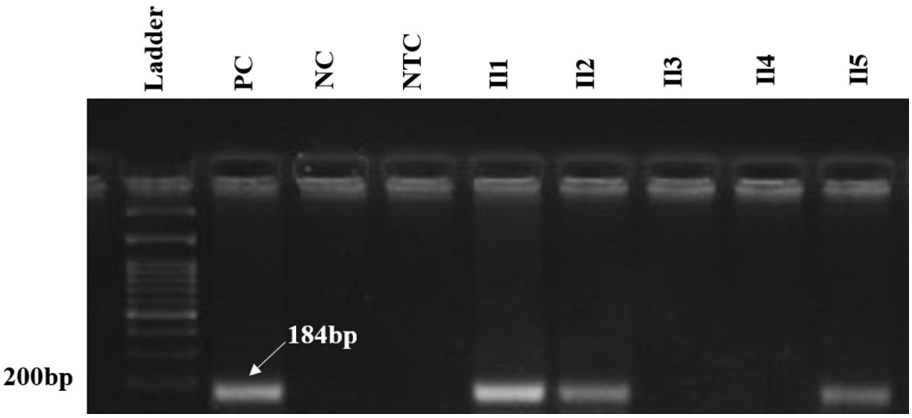

Fig. 3 PCR amplification of PsyE1 gene (Ladder- 100 bp, PC- Positive Control, NC- Negative Control- HD2967, IL—Introgression lines of PsyE1 in HD2967 genetic background at BC1F10; IL1, 2 and 5 are positive introgressions lines and IL3 and 4 are negative introgressions lines

synthase HpnC and HpnD were also found between amino acids 149–410 and 149–403. Domain for Isoprenoid Biosynthesis C1 enzymes (Isoprenoid_Biosyn_C1) and Trans-isoprenyl Diphosphate synthases (Trans_IPPS) was observed between 148 and 378 residues, whereas Farnesyl–diphosphate farnesyltransferase (squal_synth) domain was found between amino acids located at positions 229 till 316. All the observed domains are marked in the predicted protein model.

In silico protein localization

Protein was observed to be localized in the membrane bound organelle, chloroplast with 96.47% sequence similarity, i.e. 411 amino acids out of 426 total residues, with a mean alignment length of 345 amino acids and overall similarity score of 9.1. Further, the protein was predicted to be of extracellular (secreted) origin with score of 0.9. A chloroplast transit peptide of 27 amino acid residues was predicted from position 1–27.

Characterization of *PsyE1* protein homologs-alignment and phylogeny

A total of 99 related *PsyE1* proteins sequences were identified through BLASTp searches from NCBI database and were aligned to analyse the homology between the protein sequences of phytoene synthase family belonging to the super family of Isoprenoid Biosynthesis Class 1 (Fig. 1). Maximum homology (about 90%) was observed between proteins of genus *Oryza, Triticum, Zea* and *Agropyron,* which are closest to *Thinopyrum* genera. The related genes

Table 1 Predicted structural features of gene and protein encoded by *PsyE1*

Structural features of *PsyE1* gene	
Feature	Nucleotide position (basepairs)
Gene length	722,414,669–722,418,761 (4092 bp)
Transcription Start Site (TSS)	4092
Poly A tail site	69
Exons	2686 (454 bp), 2512 (51bps), 1753 (173bps), 805 (236 bp), 488 (193 bp) and 82 (174 bp)
Predicted structural features of the *PsyE1* encoded protein	
Number of amino acid residue	426
Molecular weight	47.46 kDa
Isoelectric point	8.8
Number of positive residues	59 (Arg and Lys)
Number of negative residues	52 (Asp and Glu)
Substrate binding pockets positions	At 147, 149, 168, 175, 179, 253, 258, 262, 290, 293, 300, 301, 304, 305, 310, 368 and 369 amino acid residue positions
Magnesium ion binding sites	At 175–179 and 301–305 amino acid residue position
Active-site lid amino acid residues	At 145–149 and 395–398 amino acid residue position
Catalytic amino acid residues	At 149, 168,175, 179, 253, 261, 265, 293, 300, 301, 305, 310, 368, 369 and 373 amino acid residue position
Superfamily	Isoprenoid Biosynthesis C1 superfamily
Number of specific domains	4
Chloroplast transit peptide	27 amino acid residues

were observed to translate the same kind of protein, except several changes (Indels and Point mutations) were observed among the first 120 amino acid residues representing the first exon sequence from the N-terminal and in the last exonic sequence at the C terminal. Rest of the four exon sequences seemed to be conserved among the related genera. *Hordeum* genus seemed to lack the first exon when compared with the *Thinopyrum* phytoene synthase gene. *Th. ponticum* and *Th. elongatum* were found to have about 98% sequence similarity when aligned. Many synonymous and non-synonymous mutations were observed in the first and last exon though, thus making it exploitable for use in Marker Assisted Selection by developing polymorphic functional markers.

Th. elongatum PsyE1 sequence was used as reference to estimate phylogenetic relations between related sequences. The tree was observed to be divided into three major clades and the divergence of the protein and evolutionary code is displayed in Fig. 2. Clade 1 comprised of 9 species contained the gene evolved in distant genera like *Juglans, Vitis, Camellia, Ananas, Elaeis, Phoenix and Asparagus*. Clade 2 possessed homologs from *Triticum, Aegilops, Hordeum, Thinopyrum, and Brachypodium*. *Thinopyrum elongatum PsyE1* gene belonged to this clade and the divergence analysis marked it closer to *Hordeum vulgare* and *Hordeum chilense* than *Triticum* and *Aegilops*. Eleven distinct genera i.e. *Oryza, Zoysia, Panicum, Dichanthelium, Setaria, Digitaria, Sorghum, Saccharum, Coix, Tripsacum* and *Zea* constituted the third clade.

Protein modelling

Protein structures modelled using SWISSMODEL scored better than those modelled by Phyre[2], but it omitted 125 amino acids residues from the 1st exon, which housed the translation start site of the protein and was thereby rejected. The Phyre[2] protein model (Fig. 2) was compared to the alignment of the predicted protein with already deposited models and sequences. Out of 112 proteins considered for alignment, 20 proteins models were available in Protein Data Base (PDB) and the protein hit the models with more than 50% alignment rate and 97-100% confidence level, particularly belonging to isoprenyl transferase and synthases family. A total of 68, 77 and 89 proteins were depicted to reveal >90%, >80% and > 50% confidence level, respectively, and were used in the modelling of the *PsyE1* protein (Supplementary File 2). Protein secondary structure prediction depicted that 64% of the protein has alpha helix confirmation while 26% residues are disordered. The disordered parts were observed to be localized mostly in first 130 amino acids which had a high confidence rate of disorderness (Supplementary File 2). Phyre[2] modelled protein structure was validated by Ramachandran plot and (Fig. 2) 387 residues (91.3%) were categorized in the favoured region while 6.4% i.e. 27 residues were located in the allowed region (Supplementary File 2). A total of 10 residues (2.4%) fell in the outlier region depicting the model suitability for further progress.

Homology-based promoter and transcription factor-binding Site prediction

Promoter analysis revealed presence of various cis-regulatory elements. WRKY71OS and WBOXNTERF3 (−811), SORLIP1AT (−687), NAPINMOTIFBN (−679), PYRIMIDINEBOXOSRAMY1A (−513, motif CCTTTT), GATABOX (−36) and EBOXBNNAPA and MYC (−194) motifs were observed with their conserved sequences. Light responsive elements such as TCCC-motif (−124), G-box (motif CACGTC at −145) were also observed in the *PsyE1* promoter region. ABRE motif (sequence – ACGTG; involved in the abscisic acid responsiveness) was predicted at −202 amino acid residue position and ARE motif (sequence AAACCA; essential regulatory element for the anaerobic induction) were positioned at −784, −598, −356 amino acid residues. Cis-acting element, TCA at −312 with sequence motif TCAGAAGAGG responsible for n salicylic acid responsiveness was also observed in the present study. CAAT-box, a cis-acting element in promoter and enhancer region of the protein was observed at amino acid residues −800, −327 and −430 of the protein, along with other motifs like MYB1AT (at −784 and −362), IBOXCORE (at −367 and −36) and GT-1 box (also called GT1GMSCAM4) were also predicted.

Validation of *PsyE1* in 7D/7E introgression lines

Designed primer pair successfully amplified the *PsyE1* gene in the ILs possessing higher GYPC with a desired band of approx. 180 bps (Fig. 3). Similar band was absent in the ILs with a lower range of GYPC, demonstrating lack of introgression of gene in those lines. Thus, the primer pair so developed could be successfully used to characterize the 7E/7D bread wheat introgression lines to reveal the stable introgression of *PsyE1* in them or vice versa.

Discussion

Carotenoids are pivotal category of isoprenoids produced in plastids with multifarious roles, such as cell protection from photooxidative damage, accessory pigments of photosynthetic machinery (Nisar et al. 2015; Kromdijk et al. 2016), and production of apocarotenoids (phytohormones, strigolactones) through their oxidative cleavage (Beltran and Stange 2016) and others. Carotenoid biosynthesis is subjected to three level regulation, i.e. transcription of genes involved in carotenoid metabolism (biosynthesis and degradation); activity and stability of enzymes catalyzing its metabolism, and presence of storage structures of carotenoids in plastids (Li et al. 2016). In other words, overall product of carotenoid metabolic flux regulation is a function of key enzymes involved in its biosynthesis (phytoene synthase/*PSY*), degradation (carotenoid cleavage dioxygenase/CCDs) and protein stability (orange protein/OR) (Cerda et al. 2020), that eventually governs abundance of these pigments. Biosynthesis of carotenoids is well elucidated (Fraser and Bramley 2004; Ruiz-Sola and Rodriguez-Concepcion 2012) and several reports have established the *Psy* gene catalyzed first step of the pathway, is the key point of carotenoid flux regulation (Rodriguez-Villalon et al. 2009). The present study thus holds significance in providing a first detailed insight into *in silico* characterization of *PsyE1* gene of *Th. elongatum* in terms of its gene structure, protein structure, domain and motif identification as well as development of functional marker from the gene to aid in validation of its introgression lines.

In the current study, BLAST homology searches identified a single copy of *PsyE1* homologue in *Th. elongatum*. In some crops, single gene copy encodes for *Psy* gene, such as in *A. thaliana*. However, in many other crops such as rice, wheat, melon, tomato, carrot, maize and others, *Psy* gene families demonstrating organ or environmental stimuli are present (Fraser et al. 1999; Just et al. 2007; Welsch et al. 2008; Zhang and Dubcovsky 2008; Qin et al. 2011). Since *PSY* gene is the rate controlling step in carotenoid biosynthesis and thus has been targeted in numerous studies to improve the carotenoid content of the produce. For instance, overexpression of *PSY* gene in tomatoes was observed to be responsible for production of tomato plants with reduced height but an enhanced carotenoid content (Fray et al. 1995). Similarly, rice has also been genetically modified with the *PSY* gene from daffodil to increase carotenoid content in the endosperm to enhance its nutrient value (Burkhardt et al. 1997). Likewise, an increase in wheat carotenoid content up to 50–100% has been reported with the introgression of *PsyE1* gene from *Thinopyrum* sources (Zhang et al. 2005; Padhy 2019). Biofortification of provitamin A in tetraploid durum wheat cultivars utilizing Tilling approach based on *lcy* gene has been successfully demonstrated by Sestilli et al (2019), wherein the complete mutants were observed to possess approximately 75% more β-carotene relative to control. But the same studies utilizing *PsyE1* aiming at biofortified cultivar development are still scarce. Gene structure of *PsyE1* of *Th. elongatum* observed in the current study was quite similar to that *Th. ponticum*, providing a level of confidence to perform subsequent studies with the retrieved sequence. Comparative analysis of nucleotide sequences from related organisms provides opportunities to deduce homology, evolutionary relations, and infer functions of newly sequenced genes. But there is a high proportion of homology among all the homologues of *Psy* genes. *Triticum and Hordeum* being the cultivated ones, are more of significance in the context of its nutritive role and biofortification, out of which substitution lines have already been developed

in *Triticum* and are exploited in durum varieties (Gennaro et al. 2003; Kuzmanović et al. 2018). Considerable variation was observed in the starting and terminal region of the gene, signifying the importance of an in-depth characterization of *PsyE1* gene structure and its encoded protein. This also demonstrates the potential of the same to develop functional markers in distant hybridization studies for validating successful introgression from *Th. elongatum*, eventually leading towards development of biofortified wheat varieties. Four conserved exon sequences among the related genera seem critical for functionality of the protein. *Hordeum* genus was seen to lack the first exon when compared with the *Thinopyrum* phytoene synthase gene. In many cases, deletion has been observed in cross-genera comparisons of proteins, implicating the significance of developing size-based polymorphic markers, while in cross-species comparisons, SNP markers hold the potential to be more polymorphic as the protein lengths were almost similar. Homology analysis among related *Psy* genes revealed relation between the observed variations to the evolutionary history of that crop. The analysis clearly revealed an early divergence of second and third clade in the timeline, thereby the encoded protein in these clades have not evolved much, while the members of first clade shows the divergence time in near past deciphering the evolution of the gene in this clade.

Domain and motif prediction play a pivotal role in determining biological sequence features and its function owing to their evolutionary significance. Motifs predicted in the current study such as ABE, ABRE, GT-1 box, Salicylic acid responsive elements and others has been earlier reported in *PsyE1* gene in wheat (Flowerika et al. 2016), rice (Welsch et al. 2008) and sorghum (Li et al. 2009). The presence of GT-1 box (also called GT1GMSCAM4) at the promoter region contributes to its induction in response to pathogen and sodium chloride salt (Park et al. 2004). Presence of these motifs links the probable expression of *PsyE1* to stress conditions. Similar results have also been reported by Welsch et al (2008). Álvarez et al (2016) described an instant increase in carotenoid flux requirement under salt stress or upon sudden changes in light intensity while, Fray et al. (1995) has mentioned the possible role of *Psy* genes in the Abscisic acid pathway. Presence of ARE motif and TCA motif in the promoter region further throws light to deduce the possible role of the gene in the stress conditions. Role of carotenoids in various processes such as light harvesting, photoprotection, and phytohormone biosynthesis have also been pointed by Cerda et al. (2020).

Comparative information obtained from well characterized species tends to enhance the genetics and genomics of those less elucidated ones. Similarly, identification of well conserved regions and those that harbours mutations through synteny studies provides an insight to the course of evolution of *PSY* gene. Advanced protein structure analysis inferred amino acid residues present at position 100–400 to function as putative substrate binding and catalytic domains as well as a conserved C-terminal. Observed sequence diversity, primarily in N terminal region of *Psy* gene encoded protein sequence of different species, reflects regions referred to as non-essential for enzymatic activity (Tran et al. 2009), but crucial for dimerization and targeting of protein, and eventually dire for maintaining enzyme stability and activity (Sun et al. 1996; Cunningham and Gantt 1998). Detailed studies targeting N-terminal diversity and other variable regions of *Psy* between species could help gain insight into its distinct regulatory mechanisms (Sharma et al. 2010; Shumskaya et al. 2012). Such understanding would assist production of foreign proteins through bioengineering of *Psy* promoter. Lastly, variation harbouring the 5'UTR of the gene sequence was successfully deployed in differentiating successful *PsyE1* introgression from those lacking it. Selection of the lines on the basis of their GYPC biochemical estimations and presence/absence amplification of *PsyE1* demonstrates the utility of designed marker to characterize positive introgression among different common wheat genetic backgrounds.

With regard to enhancement of coloured pigments in wheat grain for nutritional security, *PsyE1* gene derived from *Th. elongatum or Th. ponticum* holds considerable significance for further exploration and utilization in nutritional wheat breeding as it confers a high level of yellow pigments. The *Y* gene carrying alien chromosome segment has been used in wheat breeding mainly for linked resistance genes and even yield enhancing factors. However, yellow pigments contributes towards an increased nutritional value of wheat and its derived products (Bast et al. 1996; Adom et al. 2003; Mezzomo and Ferreira 2016; Zakynthinos and Varzakas 2016) and has demonstrated their efficacy in treating several cases of inflammations, immune-modulations, inter-cellular communications, degenerative-, cardiovascular-diseases, obesity and hypolipidemias (Nishino et al. 2002; Mezzomo and Ferreira 2016; Zakynthinos and Varzakas 2016).

Gennaro et al (2003) quantified 5–8.9 ppm of carotenoids in ILs *Th. ponticum* in bread wheat. Similarly, Gazza et al (2016) estimated 6–11 ppm of carotenoids in different perennial bread wheat lines derived from *Th. ponticum, Th. elongatum* and *Th. Intermedium*. Ceoloni et al. (2017) reported up to 9 ppm of carotenoids in the 7D/7E lines derived from *Th. Elongatum* and up to 11 ppm in the 7D/7el lines derived from *Th. ponticum*. In the present study, 12.3 ppm of carotenoids have been estimated in *Th. elongatum*-derived ILs. The major step ahead of the present study is the introgression of *PsyE1* in high yielding commercial variety HD2967 which can directly be used as a new variety, while the former studies have characterized the gene in some genetic stocks and in breeding lines (Ceoloni et al. 2017; Gazza et al. 2016; Gennaro 2003). Although it is stated that *Th. ponticum*

originated *PsyE1* homologue performs better than that of *Th. elongatum* in increasing the carotenoid content (Ceoloni et al. 2017), but being diploid isolation, cloning, transfer and using of *Th. elongatum* originated *PsyE1* in molecular breeding programmes will be far easy than that of *Th. ponticum* which is a polyploid species. Thus, biofortifying carotenoids in the staple crops constitute a cost-effective and sustainable approach to reduce the instances of nutritional deficiency. Few such success stories of biofortification has already been achieved, such as quality protein maize biofortified with lysine and tryptophan (QPM) and Vitamin A rich orange sweet potato.

The present study thus constitutes the first report to provide a detailed *in silico* characterization of *PSY* gene, which is a potential player for carotenoid biofortification in wheat and opens avenues to achieve the same. The study also constitutes a valuable resource for synteny-based studies, comparative structural genomics and to facilitate the map-based gene cloning, identification and deployment of functional markers for transferring the trait to the elite hexaploid wheat cultivars.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s42976-022-00279-w.

Acknowledgements The authors are thankful for financial assistance received under the ad hoc project, "Addressing food security through nutritionally enriched improved cultivars and technologies for *swasth bharat* under PURSE program" funded by Department of Science & Technology, Govt. of India, grant number: SR/PURSE Phase 2/25(G), 28.09.2017.

Author contributions AKP, AS and PK conceived the theme of study. AKP, AS and RK did the wet lab marker analysis. AKP, BS, PK, SK and MS did the bioinformatics part of the study. AKP, PK, PS, HS, SK and AS drafted the manuscript. AS and SB provided overall guidance and edited the manuscript. All authors read and approved the final version of the manuscript.

Funding Funding Agency: Department of Science & Technology, Govt. of India (partial funding) Award Number: DST file number SR/PURSE Phase 2/25(G), 28.09.2017. Recipient: Dr Achla Sharma.

Declarations

Conflict of interest Authors declare no conflict of interest.

Ethical Approval Authors worked in compliance with research ethics.

References

López-Emparán Ada, Daniela QM, Matías ZB, Víctor C, Federico I, Laura MF (2014) Functional analysis of the brassica napus L. phytoene synthase (PSY) gene family. PLoS One 9(12):e114878. https://doi.org/10.1371/journal.pone.0114878

Adom KK, Sorrells ME, Hai LR (2003) Phytochemical profiles and antioxidant activity of wheat varieties. J Agric Food Chem 51:7825–7834

Alok A, Kumar J, Thakur N, Pandey A, Pandey AK, Upadhyay SK, Tiwari S (2016) Characterization and expression analysis of phytoene synthase from bread wheat (Triticum aestivum L.). PloS one 11(10):e0162443

Álvarez D, Voß B, Maass D, Wüst F, Schaub P, Beyer P, Welsch R (2016) Carotenogenesis is regulated by 5′UTR-mediated translation of phytoene synthase splice variants. Plant Physiol 172(4):2314–26

Arango J, Wust F, Beyer P, Welsch R (2010) Characterization of phytoene synthases from cassava and their involvement in abiotic stress-mediated responses. Planta 232(5):1251–1262

Arterburn M, Kleinhofs A, Murray T, Jones S (2011) Polymorphic nuclear gene sequences indicate a novel genome donor in the polyploidy genus Thinopyrum. Hereditas 148:8–27

Atienza SG, Avila CM, Martin A (2007) The development of a PCR-based marker for PSY1 from Hordeum chilense, a candidate gene for carotenoid content accumulation in tritordeum seeds. Aust J Agric Res 58:767–773

Bariana HS, Brown GN, Bansal UK et al (2007) Breeding triple rust resistant wheat cultivars for Australia using conventional and marker-assisted selection technologies. Austr J Agric Res 58:576–587

Bast A, Vanderplas RM, Vandenberg H, Haenen G (1996) Betacarotene as antioxidant. Eur J Clin Nutr 50:54–56

Baum BR, Johnson DA (2018) Lophopyrum Á Löve (1980), Thinopyrum Á Löve (1980) and Trichopyrum Á Löve (1986): one, two or three genera? a study based on the nuclear 5S DNA. Genet Resour Crop Ev 65:1–26

Beltran JCM, Stange C (2016) Apocarotenoids: a new carotenoid-derived pathway. Subcell Biochem 79:239–272

Bouis HE, Hotz C, McClafferty B, Meenakshi JV, Pfeiffer WH (2011) Biofortification: a new tool to reduce micronutrient malnutrition. Food Nutr Bull 32(1_suppl1):S31–S40. https://doi.org/10.1177/15648265110321S105

Burkhardt PK, Beyer P, Wünn J, Klöti A, Armstrong GA, Schledz M, von Lintig J, Potrykus I (1997) Transgenic rice (Oryza sativa) endosperm expressing daffodil (Narcissus pseudonarcissus) phytoene synthase accumulates phytoene, a key intermediate of provitamin a biosynthesis. Plant J 11:1071–1078

Cai X, Jones S (1997) Direct evidence for high level of autosyndetic pairing in hybrids of Thinopyrum intermedium and Th. ponticum with Triticum aestivum. Theor Appl Genet 95:568–572

Ceoloni C, Forte P, Ciaffi M, Nenno M, Bitti A, De Vita P, D'Egidio MG (2000) Chromosomally engineered durum wheat: the potential of alien gene introgressions affecting disease resistance and quality. In: Seminar on durum wheat improvement in the mediterranean region: new challenges, Spain, Zaragoza, 12–14

Ceoloni C, Forte P, Gennaro A, Micali S, Carozza R, Bitti A (2005) Recent developments in durum wheat chromosome engineering. Cytogenet Genome Res 109:328–334

Ceoloni C, Kuzmanović L, Gennaro A, Forte P, Giorgi D, Rosaria Grossi M, Bitti A (2014) Genomes, chromosomes and genes of the wheatgrass genus Thinopyrum: the value of their transfer into wheat for gains in cytogenomic knowledge and sustainable breeding. In: Genomics of plant genetic resources, Springer, Dordrecht, 333–358

Ceoloni C, Forte P, Kuzmanović L et al (2017) Cytogenetic mapping of a major locus for resistance to fusarium head blight and crown rot of wheat on Thinopyrum elongatum 7EL and its pyramiding with valuable genes from a Th. ponticum homoeologous arm onto bread wheat 7DL. Theor Appl Genet 130:2005–2024. https://doi.org/10.1007/s00122-017-2939-8

Cerda A, Moreno JC, Acosta D, Godoy F, Cáceres JC, Cabrera R, Stange C (2020) Functional characterisation and in silico modelling of MdPSY2 variants and MdPSY5 phytoene synthases from malus domestica. J Plant Physiol 249:153166

Chen Q, Conner RL, Laroche A, Thomas JB (1998) Genome analysis of Thinopyrum intermedium and Thinopyrum ponticum using genomic in situ hybridization. Genome 41:580–586

Chen S, Huang Z, Dai Y, Qin Y, Zhang L, Gao Y, Chen J (2013) The development of 7E chromosome-specific molecular markers for Thinopyrum elongatum based on SLAF-seq technology. PLoS ONE 8:e65122

Clydesdale FM (1993) Color as a factor in food choice. Crit Rev Food Sci Nutr 33:83–101

Cunningham FX Jr, Gantt E (1998) Genes and enzymes of carotenoid biosynthesis in plants. Annu Rev Plant Biol 49(1):557–583

Dewey DR (1984) The genomic system of classification as a guide to intergeneric hybridization with the perennial Triticeae. In: Gustafson JP (ed) Gene manipulation in plant improvement. Plenum Publ Corp, NewYork, pp 209–279

Elouafi I, Nachit MM, Martin LM (2001) Identification of a microsatellite on chromosome 7B showing a strong linkage with yellow pigment in durum wheat (Triticum turgidum L. var, durum). Hereditas 135(2–3):255–61

Flowerika Alok A, Kumar J, Thakur N, Pandey A, Pandey AK et al (2016) Characterization and expression analysis of phytoene synthase from bread wheat (Triticum aestivum L). PLoS ONE 11(10):e0162443. https://doi.org/10.1371/journal.pone.0162443

Fraser PD, Bramley PM (2004) The biosynthesis and nutritional uses of carotenoids. Prog Lipid Res 43(3):228–265

Fraser PD, Kiano JW, Truesdale MR, Schuch W, Bramley PM (1999) Phytoene synthase-2 enzyme activity in tomato does not contribute to carotenoid synthesis in ripening fruit. Plant Mol Biol 40(4):687–698

Fray RG, Wallace A, Fraser PD, Valero D, Hedden P, Bramley PM, Grierson D (1995) Constitutive expression of a fruit phytoene synthase gene in transgenic tomatoes causes dwarfism by redirecting metabolites from the gibberellin pathway. Plant J 8:693–701

Gallagher CE, Matthews PD, Li FQ, Wurtzel ET (2004) Gene duplication in the carotenoid biosynthetic pathway preceded evolution of the grasses. Plant Physiol 135:1776–1783

Gazza L et al. (2016) Qualitative traits of perennial wheat lines derived from different Thinopyrum species. Genet Res Crop Evol 63(2):209–219

Gennaro A, Borrelli GM, D'Egidio MG, De Vita P, Ravaglia S, and Ceoloni C (2003) A chromosomally engineered durum wheat-Thinopyrum pontificum recombinant line with novel and promising attributes for varietal development. In Proceedings of the 10th International Wheat Genetics Symposium 2: 881–3

Gennaro A, Koebner R, Ceoloni C (2009) A candidate for Lr19, an exotic gene conditioning leaf rust resistance in wheat. Funct Integr Genomics 9(3):325–334

Guo J, Yu X, Yin H, Liu G, Li A, Wang H, Kong L (2016) Phylogenetic relationships of Thinopyrum and Triticum species revealed by SCoT and CDDP markers. Plant Syst Evol 302:1301–2130

Han Y, Zheng QS, Wei YP, Chen J, Liu R, Wan HJ (2015) In silico identification and analysis of phytoene synthase genes in plants. Genet Mol Res 14(3):9412–22. https://doi.org/10.4238/2015.August.14.5

He XY, Zhang YL, He ZH, Wu YP, Xiao YG, Ma CX, Xia XC (2008) Characterization of phytoene synthase 1 gene (Psy1) located on common wheat chromosome 7A and development of a functional marker. Theor Appl Genet 116:213–221

Jauhar PP (1990) Multidisciplinary approach to genome analysis in the diploid species, Thinopyrum bessarabicum and Th. elongatum (Lophopyrum elongatum), of the Triticeae. Theor Appl Genet 80:523–536

Just BJ, Santos CA, Fonseca ME, Boiteux LS, Oloizia BB, Simon PW (2007) Carotenoid biosynthesis structural genes in carrot (Daucus carota): isolation, sequence-characterization, single nucleotide polymorphism (SNP) markers and genome mapping. Theor Appl Genet 114(4):693–704

Kim IJ, Ko KC, Nam TS, Kim YW, Chung WI, Kim CS (2003) Expression and activity of citrus phytoene synthase and $\beta $-carotene hydroxylase in escherichia coli. J Microbiol 41(3):212–218

Knott DR (1968) Translocations involving Triticum chromosomes and Agropyron chromosomes carrying rust resistance. Can J Genet Cytol 10:695–696

Knott DR (1980) Mutation of a gene for yellow pigment linked to Lr19 in wheat. Can J Genet Cytol 22:651–654

Knott DR (1989) The effect of transfers of alien genes for leaf rust resistance on the agronomic and quality characteristics of wheat. Euphytica 44:65–72

Kromdijk J, Glowacka K, Leonelli L, Gabilly ST, Iwai M, Niyogi KK, Long SP (2016) Improving photosynthesis and crop productivity by accelerating recovery from photoprotection. Science 354(6314):857–861

Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol 35(6):1547–1549. https://doi.org/10.1093/molbev/msy096

Kuzmanović L, Ruggeri R, Able JA, Bassi FM, Maccaferri M, Tuberosa R, Ceoloni C (2018) Yield of chromosomally engineered durum wheat-Thinopyrum ponticum recombinant lines in a range of contrasting rain-fed environments. Field Crop Res 228:147–157

Lao YM, Xiao L, Ye ZW, Jiang JG, Zhou SS (2011) In silico analysis of phytoene synthase and its promoter reveals hints for regulation mechanisms of carotenogenesis in Duanliella bardawil. Bioinformatics 27(16):2201–2208. https://doi.org/10.1093/bioinformatics/btr371

Li H, Wang X (2009) Thinopyrum ponticum and Thinopyrum intermedium: the promising source of resistance to fungal and viral diseases of wheat. J Genet Genomics 36:557–565

Li L, Yuan H, Zeng Y, Xu Q (2016) Plastids and carotenoid accumulation. Subcell Biochem 79:273–293

Li X, Jiang X, Chen X, Song J, Ren C, Xiao Y, Gao X, Ru Z (2017) Molecular cytogenetic identification of a novel wheat-Agropyron elongatum chromosomes translocation line with powdery mildew resistance. PLoS ONE 12(9):e0184462

Liu M, Hambleton S (2010) Taxonomic study of stripe rust, Puccinia Striiformis sensu lato, based on molecular and morphological evidence. Fungal Biol 114:881–899

Löve A (1982) Generic evolution of the wheatgrasses. Biol Zentralbl 101:199–212

Maass D, Arango J, Wüst F, Beyer P, Welsch R (2009) Carotenoid crystal formation in arabidopsis and carrot roots caused by increased phytoene synthase protein levels. PLoS ONE 4(7):e6373. https://doi.org/10.1371/journal.pone.0006373

Mao P, Huang Y, Wang X, Meng L, Mao P, Zhang G (2010) Cytological evaluation and karyotype analysis in plant germplasms of Elytrigia Desv. Agr Sci China 9:553–1560

Mezzomo N and Ferreira Sandra RS (2016) Carotenoids functionality, sources, and processing by supercritical technology: a review. J Chem

Nisar N, Li L, Lu S, Khin NC, Pogson BJ (2015) Carotenoid metabolism in plants. Mol Plant 8(1):68–82

Nishino H, Murakoshi M, Li T, Takemura M, Kuchide M, Kanazawa M, Yang Mou X, Wada S, Masuda M, Ohsaka Y, Yogosawa S, Satomi Y, Jinno K (2002) Carotenoids in cancer chemoprevention. Cancer Metastasis Rev 21(3):257–264

Padhy AK (2019) Marker assisted selection in wheat for high carotenoid and protein content in grains (M.Sc. thesis, Punjab Agricultural University, Ludhiana). http://krishikosh.egranth.ac.in/handle/1/5810127613

Park HC, Kim ML, Kang YH, Jeon JM, Yoo JH, Kim MC, Park CY, Jeong JC, Moon BC, Lee JH, Yoon HW, Lee S, Woo WS, Lim CO, Lee SY, Hong JC, Cho MJ (2004) Pathogen- and NaCl-induced expression of the SCaM-4 promoter is mediated in part by a GT-1 box that interacts with a GT-1-like transcription factor. Plant Physiol 135(4):2150–2161. https://doi.org/10.1104/pp.104.041442

Pozniak CJ, Knox RE, Clarke FR, Clarke JM (2007) Identification of QTL and association of a phytoene synthase gene with endosperm colour in durum wheat. Theor Appl Genet 114(3):525–37

Pozniak CJ, Knox RE, Clarke FR, Clarke JM (2007a) Identification of QTL and association of a phytoene synthase gene with endosperm colour in durum wheat. Theor Appl Genet 114(3):525–537

Pozniak CJ, Knox RE, Clarke FR, Clarke JM (2007b) Identification of QTL and association of a phytoene synthase gene with endosperm colour in durum wheat. Theor Appl Genet 114:525–537

Qin X, Coku A, Inoue K, Tian L (2011) Expression, subcellular localization, and cis-regulatory structure of duplicated phytoene synthase genes in melon (Cucumis melo L.). Planta 234(4):737–748

Rodriguez-Villalon A, Gas E, Rodriguez-Concepcion M (2009) Phytoene synthase activity controls the biosynthesis of carotenoids and the supply of their metabolic precursors in dark-grown arabidopsis seedlings. Plant J 60(3):424–435

Rosas-Saavedra C, Stange C (2016) Biosynthesis of carotenoids in plants: enzymes and color. Subcell Biochem 79:35–69

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

Sears ER (1972a) Agropyron-wheat transfers through induced by homoeologous pairing. Can J Genet Cytol 14:736

Sears ER (1972) Chromosome engineering in wheat, vol 4. Stadler Symposia, Columbia, pp 23–38

Sears ER (1973) *Agropyron*-wheat transfers induced by homoeologous pairing. In: Sears ER, Sears LMS (eds) Proceeding of the 4th International Wheat Genetics Symposium. Univ. of Missouri, Columbia, MO, USA, pp 191–199

Sestili F, Garcia-Molina MD, Gambacorta G, Beleggia R, Botticella E, De Vita P et al (2019) Provitamin a biofortification of durum wheat through a tilling approach. Int J Mol Sci 20(22):5703

Sharma D, Knott DR (1966) The transfer of leaf rust resistance from Agropyron to Triticum by irradiation. Can J Genet Cytol 8:137–143

Sharma A, Punetha A, Grover A, Sundar D (2010) Insights into the key enzymes of secondary metabolites biosynthesis in camellia sinensis. J Bioinform Seq Anal 2(5):53–68

Shepherd KW, Islam AKMR (1988) Fourth compendium of wheat alien chromosome lines. In: Miller TE, Koebner RMD (eds) Proceedings of the 7th international genetic symposium. Bath Press, Bath, pp 1373–1398

Shewmaker CK, Sheehy JA, Daley M, Colburn S, Ke DY (1999) Seed-specific overexpression of phytoene synthase: increase in carotenoids and other metabolic effects. Plant J 20:401–412

Shumskaya M, Bradbury LM, Monaco RR, Wurtzel ET (2012) Plastid localization of the key carotenoid enzyme phytoene synthase is altered by isozyme, allelic variation, and activity. Plant Cell 24(9):3725–3741

Singh G, Saini JS, Bains NS, Singh RP (2014) Positive influence of Lophopyrum ponticum derived Y gene on yellow pigment content- a major durum wheat quality trait, India. Indian J Genet Plant Breed 74:651–655. https://doi.org/10.5958/0975-6906.2014.00905.5

Sun Z, Gantt E, Cunningham FX (1996) Cloning and functional analysis of the β-carotene hydroxylase of arabidopsis thaliana. J Biol Chem 271(40):24349–24352

Tran D, Haven J, Qiu WG, Polle JE (2009) An update on carotenoid biosynthesis in algae: phylogenetic evidence for the existence of two classes of phytoene synthase. Planta 229(3):723–729

Wang Q, Xiang J, Gao A et al (2010) Analysis of chromosomal structural polymorphisms in the St, P, andY genomes of Triticeae (Poaceae). Genome 53:241–249

Wang RR-C (1992) Genome relationships in the perennial Triticeae based on diploid hybrids and beyond. Hereditas 116:133–136

Wang RR-C (2011) Agropyron and Psathyrostachys. In: Kole C (ed) Wild crop relatives: genomic and breeding resources, cereals. Springer, Berlin, pp 77–108

Wang H, Sun S, Ge W, Zhao L, Hou B, Wang K, Lyu Z, Chen L, Xu S, Guo J, Li M (2020) Horizontal gene transfer of Fhb7 from fungus underlies fusarium head blight resistance in wheat. Science. https://doi.org/10.1126/science.aba5435

Welsch R, Wust F, Bar C, Al-Babili S, Beyer P (2008) A third phytoene synthase is devoted to abiotic stress-induced abscisic acid formation in rice and defines functional diversification of phytoene synthase genes. Plant Physiol 147(1):367–380

Zakynthinos G, Varzakas T (2016) Carotenoids from plants to food industry. Curr Res Nutr Food Sci 4(Special Issue):38

Zhang W, Dubcovsky J (2008) Association between allelic variation at the phytoene synthase 1 gene and yellow pigment content in the wheat grain. Theor Appl Genet 116:635–645

Zhang W, Lukaszewski A, Kolmer J, Soria M, Goyal S, Dubcovsky J (2005) Molecular characterization of durum and common wheat recombinant lines carrying leaf rust resistance (Lr19) and yellow pigment (Y) genes from Lophopyrum ponticum. Theor Appl Genet 111:573–582

Zhang XY, Koul A, Petroski R et al (1996) Molecular verification and characterization of BYDV resistant germ plasms derived from hybrids of wheat with Thinopyrum ponticum and Th. intermedium. Theor Appl Genet 93:1033–1039

Online references

http://bioinformatics.psb.ugent.be/webtools/plantcare/html/
http://mordred.bioc.cam.ac.uk/~rapper/rampage.php
http://plantpan.itps.ncku.edu.tw/promoter.php
http://smartembl-heidelberg.de/smart/setmode.cgi?NORMAL=1.
http://www.sbg.bio.ic.ac.uk/~phyre2/html/page.cgi?id=index
http://www.softberry.com/berry.phtml?topic=fgenesh&group=programs&subgroup=gfind
http://www.softberry.com/berry.phtml?topic=protcomppl&group=programs&subgroup=proloc
https://itol.embl.de/
https://swissmodel.expasy.org/interactive
https://web.expasy.org/protparam/
https://web.expasy.org/translate/
https://www.dna.affrc.go.jp/PLACE/?action=newplace
https://www.ebi.ac.uk/Tools/msa/clustalo/
https://www.ncbi.nlm.nih.gov/
https://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml
https://www.ncbi.nlm.nih.gov/projects/msaviewer/
The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.