

Article

# Differential Regulation of Phytoene Synthase PSY1 During Fruit Carotenogenesis in Cultivated and Wild Tomato Species (*Solanum* section *Lycopersicon*)

Gleb I. Efremov, Maria A. Slugina \*, Anna V. Shchennikova and Elena Z. Kochieva

Institute of Bioengineering, Research Center of Biotechnology, Russian Academy of Sciences, 119071 Moscow, Russia; gleb\_efremov@mail.ru (G.I.E.); shchennikova@yandex.ru (A.V.S.); ekochieva@yandex.ru (E.Z.K.)

\* Correspondence: mashinmail@mail.ru; Tel.: +7-499-1356219

Received: 27 July 2020; Accepted: 7 September 2020; Published: 9 September 2020



**Abstract:** In plants, carotenoids define fruit pigmentation and are involved in the processes of photo-oxidative stress defense and phytohormone production; a key enzyme responsible for carotene synthesis in fruit is phytoene synthase 1 (PSY1). Tomatoes (*Solanum* section *Lycopersicon*) comprise cultivated (*Solanum lycopersicum*) as well as wild species with different fruit color and are a good model to study carotenogenesis in fleshy fruit. In this study, we identified homologous *PSY1* genes in five *Solanum* section *Lycopersicon* species, including domesticated red-fruited *S. lycopersicum* and wild yellow-fruited *S. cheesmaniae* and green-fruited *S. chilense*, *S. habrochaites* and *S. pennellii*. *PSY1* homologs had a highly conserved structure, including key motifs in the active and catalytic sites, suggesting that *PSY1* enzymatic function is similar in green-fruited wild tomato species and preserved in red-fruited *S. lycopersicum*. *PSY1* mRNA expression directly correlated with carotenoid content in ripe fruit of the analyzed tomato species, indicating differential transcriptional regulation. Analysis of the *PSY1* promoter and 5'-UTR sequence revealed over 30 regulatory elements involved in response to light, abiotic stresses, plant hormones, and parasites, suggesting that the regulation of *PSY1* expression may affect the processes of fruit senescence, seed maturation and dormancy, and pathogen resistance. The revealed differences between green-fruited and red-fruited *Solanum* species in the structure of the *PSY1* promoter/5'-UTR, such as the acquisition of ethylene-responsive element by *S. lycopersicum*, could reflect the effects of domestication on the transcriptional mechanisms regulating *PSY1* expression, including induction of carotenogenesis during fruit ripening, which would contribute to red coloration in mature fruit.

**Keywords:** carotenogenesis; fruit color; *Solanum* section *Lycopersicon*; phytoene synthase 1

## 1. Introduction

Tomato (*Solanum lycopersicum* L.) fruit represents a fleshy berry, which during ripening, changes the color from green to red (most common) or yellow, orange, pink, brown, or purple, primarily because of chlorophyll degradation and intensification of carotenoid synthesis and accumulation [1,2]. *S. lycopersicum* belongs to the *Solanum* section *Lycopersicon*, which also includes 12 related wild tomato species with different evolutionary ages. Most wild tomatoes produce green or dark-green fruit, with the exception of three species, which have yellow (*Solanum cheesmaniae* [L. Riley] Fosberg and *Solanum galapagense* S.C. Darwin and Peralta) or red (*Solanum pimpinellifolium* B. Juss.) fruit color depending on the lycopene to  $\beta$ -carotene ratio [3,4]. All tomato species have flower petals of bright yellow-orange color due to the accumulation of carotenoids neoxanthin, violaxanthin, and lutein [5]. Carotenoids not only color flower organs and fruit to attract insects and animals for pollination and seed dispersion but are also involved in photosynthesis, photo-oxidative stress defense, and phytohormone production during plant vegetative growth [6–10].

The biosynthesis of carotenoids, including the underlying genetic and enzymatic networks, is described in many plant species [11]. The carotenoid synthetic pathway is initiated in plastids by phytoene synthase (PSY), which is present in plastid stroma (chloroplasts, chromoplasts and amyloplasts) and catalyzes the condensation of two geranylgeranyl diphosphate (GGPP) molecules to yield 15-*cis*-phytoene [2,12]. The following consecutive biosynthetic reactions catalyzed by phytoene desaturase (PDS),  $\zeta$ -carotene isomerase (ZISO) and  $\zeta$ -carotene desaturase (ZDS) produce *cis*-lycopene, which is then converted by carotenoid isomerase (CrtISO) into *trans*-lycopene (red pigment) further used for the synthesis of  $\alpha$ - and  $\beta$ -carotenes (orange pigments), which in turn may be converted to lutein and xanthophylls, respectively [1,2,13]. In this pathway, PSY is the key enzyme, and its loss blocks carotenoid synthesis during fruit ripening, whereas its overexpression increases carotenogenesis in sink organs in several crop species, including potato and tomato [14,15].

Plants may have from one to four PSY-encoding genes [16–18]. In tomato, two active PSY enzymes were initially described: PSY1 (pTOM5) and PSY2, which provide carotenoid synthesis in fruit and photosynthetic tissues, respectively; both of them are necessary for petal pigmentation [12,19,20]. Paralogous *PSY1* and *PSY2* genes, which are considered to emerge as a result of the *Solanum*-specific whole-genome triplication and subsequent subfunctionalization [12,21], encode proteins with highly homologous structures; however, their biochemical properties are slightly different [12,22]. Thus, the carotenogenic activity of PSY1 is lower than that of PSY2, which may be the reason behind a higher level of *PSY1* gene transcription during carotenoid synthesis in ripe fruit [10]. The fact that all tomato species have the same color of flower petals (yellow) but differ in the color of ripe fruit depending on the evolutionary origin (ancient species have green and recent species—yellow-to-red fruit) may indicate that *PSY1* emerged later than *PSY2* and was initially involved, together with *PSY2*, in flower coloration and only later acquired a distinct function in fruit pigmentation [5]. The recently identified third tomato PSY enzyme (*PSY3*) mainly functions in the synthesis of root apo-carotenoids ( $C_{19}$  strigolactones,  $C_{15}$  abscisic acid, and  $C_{13}/C_{14}$  apo-carotenoids) involved in establishing plant symbiotic and parasitic relationships and adaptation to nutrient deficiency [23]. *PSY3* homologs have been found in both monocots [24–26] and eudicots [23,27,28] up to *Amborella trichopoda* Baill., the most basal angiosperm considered a progenitor of extant angiosperms [23,29], suggesting an earlier origin of the root-specific *PSY3* gene compared to the fruit-specific *PSY1* gene in tomato.

Tomato species (*Solanum* section *Lycopersicon*) are considered a good model to study fleshy fruit ripening, including coloration and evolution of the underlying genetic mechanisms, because of their significant phenotypic diversity [30]. All three *PSY* genes have been detected in cultivated tomato *S. lycopersicum*; however, there are few studies on PSY enzymes in wild tomatoes, which mostly indicate that fruit color depends on PSY1 expression [31–35].

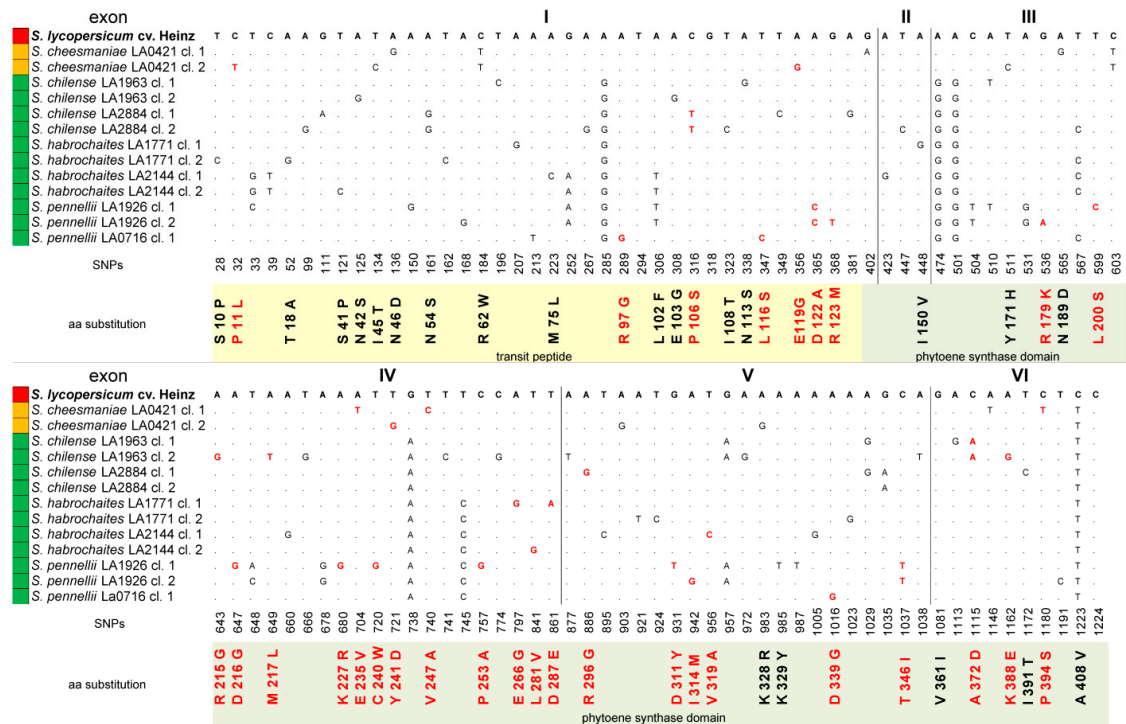
The aim of this study was to evaluate the interspecific variability of *PSY1* homologous genes in cultivated and wild tomato accessions, reveal possible associations of PSY1 structural polymorphisms and expression with carotenoid biosynthesis and fruit color, and determine PSY1 relevance to the evolution of tomato species.

## 2. Results

### 2.1. Identification and Structural Analysis of *PSY1* Homologous Genes in *Solanum* Section *Lycopersicon* Species

A total of 13 complete sequences of *PSY1* homologous genes, including their allelic variants, were amplified from genomic DNA of eight accessions of five tomato species: red-fruited (RF) *S. lycopersicum*, yellow-fruited (YF) *S. cheesmaniae* and green-fruited (GF) *S. chilense*, *S. habrochaites* and *S. pennellii*. The genes were cloned and sequenced, and the data deposited in the NCBI GenBank (accession numbers are shown in Table 1). The size of the identified *PSY1* genes ranged from 4840 bp (*S. chilense* LA1963 clone 2) to 4916 bp (*S. habrochaites* LA1771 clone 2) (Table 1); all genes included 5'-UTR (~1.5 kb), a promoter (~0.9 kb) and a coding (3.3 kb in average from the start to stop codon) sequence. The open reading frame (ORF) was 1239-bp long and contained six exons; the length and number of exons, and,

consequently, the cDNA size did not vary among the analyzed species and were the same as reported for *S. lycopersicum* *PSY1* [12] and other *Solanum* *PSY1* genes available in the NCBI database (Table 1, Figure 1).



**Figure 1.** Polymorphisms in identified tomato *PSY1* homologs. The numbers indicate the positions of single nucleotide polymorphisms (SNPs) in the *PSY1* coding sequences (exons I–VI) relative to *S. lycopersicum* *PSY1*; the resulting amino acid substitutions in the translated *PSY1* proteins are shown below. Non-synonymous SNPs and PROVEAN-predicted radical amino acid substitutions are marked red. Color of the ripe fruit is indicated to the left of the species name.

Compared to the *PSY1* gene of *S. lycopersicum* cv. Heinz 1706, the average variability of the identified *PSY1* genes was 12.8% for complete sequences (641 SNPs in the aligned 4995-bp gene portion): 8.5% for exons (105 SNPs in 1239 bp), 12.2% for introns (298 SNPs in 2117 bp), and 14.6% for 5'-UTR (238 SNPs in 1639 bp); 5'-UTR and exon I were the most variable parts. In total, 51 (48.57%) exonic SNPs were non-synonymous and found mostly in *PSY1* homologs of GF tomatoes (Figure 1).

The same analysis performed for other *Solanum* species indicated that the variability of their *PSY1* genes was mostly lower than that in the *Solanum* section *Lycopersicon* species. Thus, in potato (*Solanum* section *Petota*), it was only 8.00% (414 SNPs in 5150 bp), including 5.2% for exons and 6.98% for 5'-UTR, and in more distant *Solanum* species (*S. sisymbriifolium*, *S. melongena*, *S. prinophyllum*, and *S. torvum*), it was 6.2%, 6.4%, 5.8% and 8.5% for exons, respectively. Other members of the Solanaceae family, such as *Capsicum annuum*, *Lycium* species, *Nicotiana tabacum*, and *Petunia axillaris* had 7.9%, 9.0%, 11.3% and 16.1% variability, respectively, in *PSY1* genes compared to *PSY1* of *S. lycopersicum* cv. Heinz 1706.

**Table 1.** Characteristics of *PSY1* homologs from the analyzed *Solanum* section *Lycopersicon* species.

Accession	TGRC Collection Number	Origin	Ripe Fruit Color *	NCBI Gene ID/Solyc No.	Gene, bp	cDNA, bp	Protein, aa	pI	MW, kDa
<i>PSY1</i> genes identified in this study									
<i>S. lycopersicum</i> L. cv. Heinz 1706-BG ( <i>Lycopersicon</i> group)	LA4345		Red	MT664042	4871	1239	412	8.1	46.6
<i>S. cheesmaniae</i> (L. Riley) Fosberg (Esculentum group/ <i>Lycopersicon</i> group)	LA0421	San Cristobal: cliff East of Wreck Bay, Galapagos Islands, Ecuador	Yellow	MN782521 MN782522	4883/4876	1239	412	7.74/7.9	46.6/45.5
<i>S. chilense</i> (Dunal) Reiche (Peruvianum group/ <i>Eriopersicon</i> group)	LA1963	Rio Caplina, Tacna, Peru	Green to whitish green with purple stripes	MN782523 MN812838	4858/4840	1239	412	7.74/7.29	46.6/46.3
	LA2884	Ayaviri, Antofagasta, Chile		MN782524 MN782525	4878/4876	1239	412	7.93/7.90	46.5/46.5
<i>S. habrochaites</i> S. Knapp & D. M. Spooner (Hirsutum group/ <i>Eriopersicon</i> group)	LA1771	Rio Casma, Ancash, Peru	Green with darker green stripes	MN782526 MN782527	4914/4916	1239	412	8.24	46.6
	LA2144	Chanchan, Chimborazo, Ecuador		MN782528 MN782529	4906/4903	1239	412	8.1	46.6
<i>S. pennellii</i> Correll (Hirsutum group/ <i>Neolycopersicon</i> group)	LA1926	Agua Pertida, Ica, Peru	Green	MN782530 MN782531	4898/4901	1239	412	8.5/8.1	46.7/46.6
	LA0716	Atico, Arequipa, Peru		MN782532	4886	1239	412	8.1	46.5
<i>PSY1</i> genes available in NCBI GenBank									
<i>S. lycopersicum</i> cv. Heinz 1706			Red	543988, NC_015440.3 (4350836..4355976); Solyc03g031860.2.1	4872	1239	412	8.1	46.6
<i>S. pimpinellifolium</i>	LA1589	La Libertad, Peru	Red	AGFK01024844.1 (1289..7095)	4872	1239	412	8.1	46.6
<i>S. arcanum</i>	LA2157	Tunel Chotano, Cajamarca, Peru	Green	CBYQ010012533.1 (26020..31886)	4880	1239	412	8.1	46.6
<i>S. habrochaites</i>	LYC4	Peru/Ecuador	Green with dark-green stripes	CBYS010011028.1 (46212..52055)	4910	1239	412	8.1	46.6
<i>S. pennellii</i>	LA0716	Atico, Arequipa, Peru	Green	CCXL01009615.1 (3669..9559)	4949	1239	412	8.1	46.6

\* Fruit color data are indicated according to ref. [36]. pI—isoelectric point; MW—molecular weight; gene sequences were found in the NCBI database.

## 2.2. Structural Analysis of Tomato PSY1 Homologs

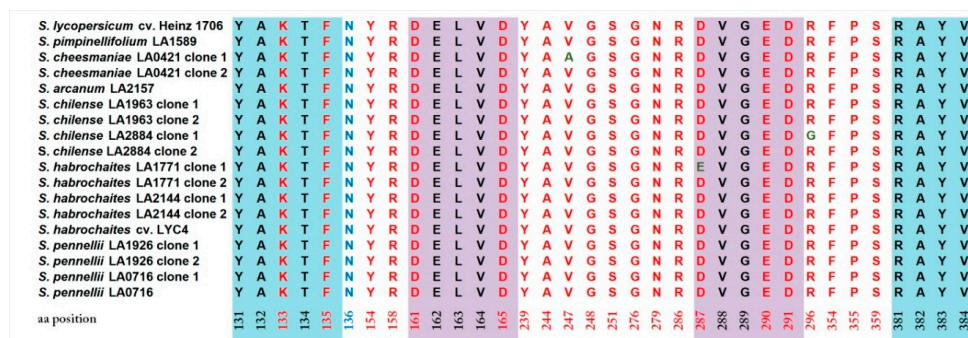
The size of the translated PSY1 proteins was the same for all studied tomato species: 412 amino acids (aa) (Table 1). Similar to *S. lycopersicum* PSY1, the putative PSY1 protein homologs contained a conserved phytoene synthase domain head-to-head (HH)-IPPS (75–405 aa, according to the NCBI-CDD) characteristic for the isoprenoid biosynthesis C1 superfamily. The N-terminal transit peptide (TP) was determined according to a previous study [12]; its predicted cleavage site V61/R62 was the same for the analyzed PSY1 homologs, excluding V61/W62 in PSY1 of *S. cheesmaniae* LA0421. However, according to Resource UniProtKB (Mitofates, Predotar, and TargetP tools) the PSY1 phytoene synthase domain was mapped to 130–412 aa and TP—to 1–129 aa, which is consistent with the plastidic localization of PSY1; in this case, the TP cleavage site was located at A129/E130 and was the same for all analyzed PSY1 proteins.

Compared to the PSY1 of cultivated *S. lycopersicum* (cv. Heinz 1706), the analyzed PSY1 homologs from wild tomatoes had 46 residue substitutions (11.2% of the 412-aa protein), of which 27 (58.7%) were radical (Figure 1). The 62-aa TP encoded by exon I was the most variable PSY1 region, containing seven substitutions (15.2%), of which only one was predicted to be radical (rP11L in *S. cheesmaniae* clone 2). In case of the 129-aa TP, 18 substitutions (39.1%) were found; among them, seven were predicted to be radical, of which four were located close to the cleavage site (Figure 1). Most of the radical replacements were detected in *S. pennellii* LA0716 clone 1 (rR97G, rP106S, rL116S, and rD122A); they were also found in *S. chilense* LA2884 clones 1 and 2 (rP106S), *S. cheesmaniae* LA0421 clone 2 (rE119G), and *S. pennellii* LA1926 clone 2 (rD122A and rR123M) (Figure 1). The remaining 28 substitutions (including 20 radical) were located in the conserved HH-IPPS domain. The HH-IPPS domains of two RF species, *S. pimpinellifolium* and *S. lycopersicum*, differed from each other only by one neutral substitution (nV360I); *S. pimpinellifolium* PSY1 lacked nA408V found in all analyzed accessions (Figure 1). PSY1 of YF *S. cheesmaniae* LA 0421 contained seven aa substitutions (nN189D, rE235V, rV247A, and rP394S in clone 1 and nY171H, rY241D and nK328R in clone 2) (Figure 1).

The majority of residue substitutions (19, including 16 radical) were found in the PSY1 HH-IPPS domain of more ancient GF tomato species. Almost all substitutions were species- and accession-specific; they were found in *S. chilense* LA1963 (rR215G, rM217L, rK388E and rA372D in clone 2 and rA372D in clone 1) and LA2884 (rR296G and nI391T in clone 1), *S. habrochaites* LA1771 (nI150V, rE266G, and nD287E in clone 1) and LA2144 (rL281V in clone 2), and *S. pennellii* LA1926 (rL200S, rD216G, rK227R, rC240W, rD311Y and rT346I in clone 1; rT346I, rR179K and rI314M in clone 2) and LA0716 (rD339G in clone 1) (Figure 1). Thus, among the analyzed tomato species, *S. pennellii* LA1926 PSY1 contained the most variable HH-IPPS domain and *S. pennellii* LA0716 PSY1—the most polymorphic TP, whereas minimal differences were observed between the PSY1 homologs of RF and YF species.

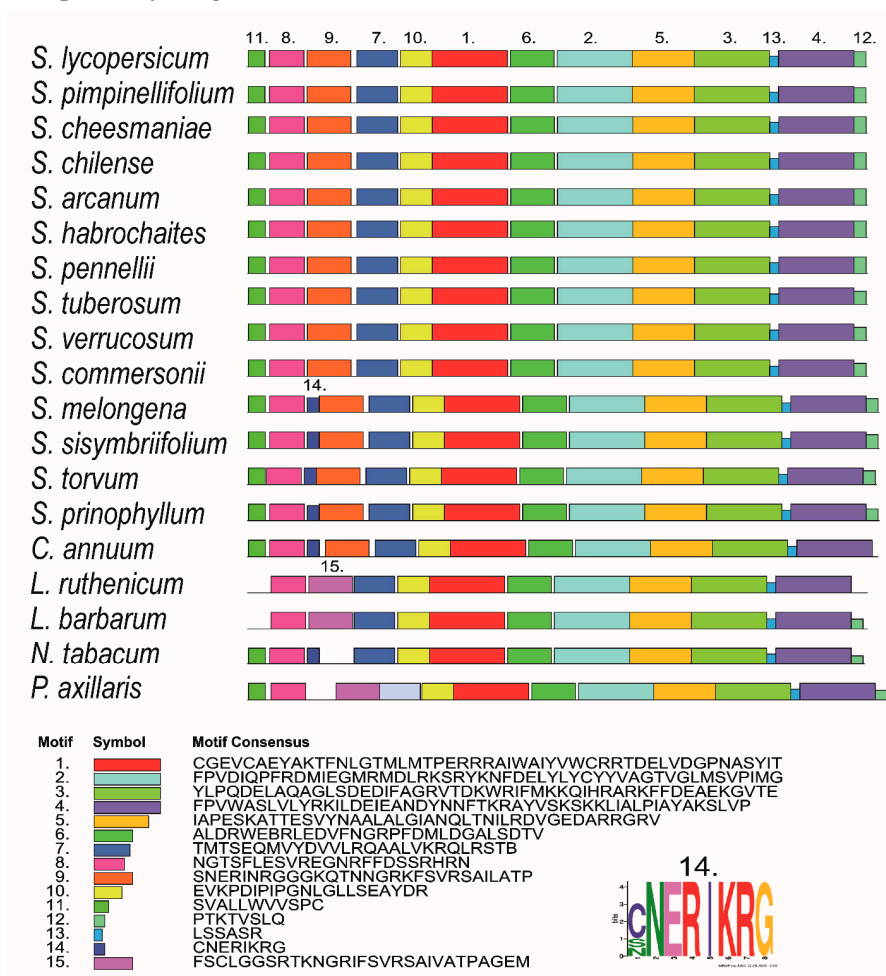
Polymorphisms in PSY1 homologs of species from *Solanum* sections Petota and Melongena, *S. prinophyllum* and *S. torvum*, as well as other Solanaceae members *C. annuum*, *Lycium* species, *N. tabacum*, and *P. axillaris* constituted 4.6%, 5.6%, 6.2%, 7.7%, 8.1%, 11.0%, 6.8% and 14.4%, respectively, compared to *S. lycopersicum* cv. Heinz 1706 PSY1.

To expand the data on the variability of functionally important sites among tomato species, an expanded list of tomato PSY1 homologs (supplemented with PSY1 proteins available from NCBI, in particular, from *S. pimpinellifolium* and *S. arcanum*) was analyzed using the NCBI Conserved Domains Database, and the active site lid residues (131-YAKTF-135 and 381-RAYV-384; define the conformation of the active cleft site), aspartate-rich substrate-Mg<sup>2+</sup> binding sites (161-DELVD-165 and 287-DVGED-291), substrate-binding pockets (K133, F135, Y154, R158, D161, D165, Y239, A244, G248, G276, N279, R286, D287, E290, D291, R296, F354, and P355), and catalytic residues (F135, Y154, D161, D165, Y239, V247, S251, N279, R286, D287, D291, R296, F354, P355, and S359) were revealed (Figure 2). The identified functionally important sites were conserved among the analyzed PSY1 proteins, with the exception of substitutions in the second substrate-Mg<sup>2+</sup> binding site (nD287E in *S. habrochaites* LA1771 clone 1) and the substrate-binding/catalytic site (rR296G in *S. chilense* LA2884 clone 1). Furthermore, all identified tomato PSY1 homologs contained a functionally important mutation Y136N absent in other Solanaceae species (Figure 2).



**Figure 2.** Functionally important sites in PSY1-encoded proteins. Active site residues (131-YAKTF-135 and 381-RAYV384) are highlighted in blue, aspartate-enriched substrate-Mg<sup>2+</sup> binding sites (161-DELVD-165 and 287-DVGED-291)—violet, and additional residues that form the active site are in red font. Amino acid positions are indicated according to *S. lycopersicum* cv. Heinz PSY1. Position N136 (in blue font) indicates the Y136N replacement compared to *C. annuum* PSY1.

MEME-based analysis revealed 13 conserved motifs in PSY1 proteins of the analyzed tomato accessions and potato species. Additional group-specific motifs were found in *S. melongena*, *S. sisymbriifolium*, *S. torvum*, *S. prinophyllum*, and *C. annuum* (motif 14; 41–48 aa), and in *Lycium* and *Petunia* species (motif 15; 42–70 and 60–88 aa, respectively) (Figure 3).



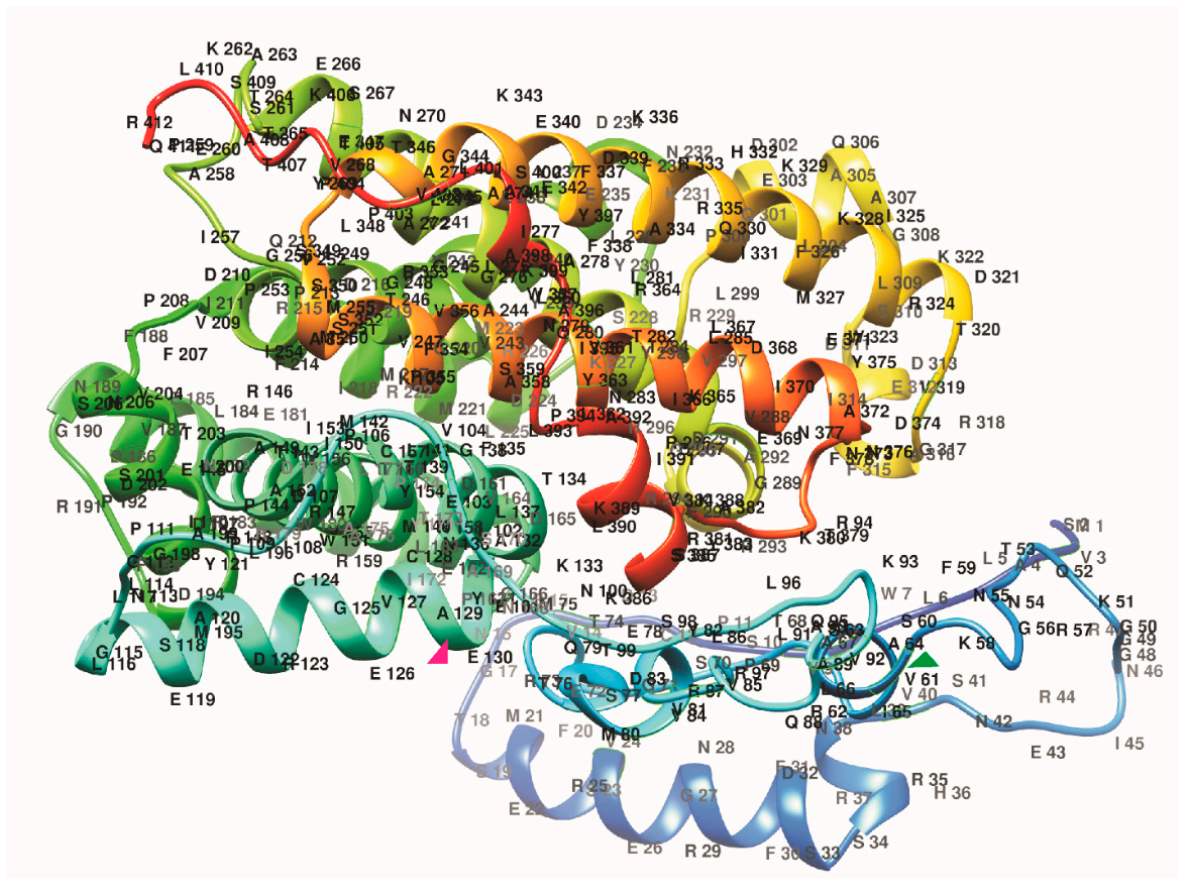
**Figure 3.** Distribution of conserved motifs in PSY1 homologs of Solanaceae species identified using the MEME search tool. The length of each box corresponds to that of the motif; the order of the motifs corresponds to their position in each protein.

### 2.3. Spatial Structure of Tomato PSY1 Homologs

The three-dimensional (3D) structure of tomato PSY1 homologs was predicted based on crystal structures of four related enzymes: the C(30) carotenoid dehydrosqualene synthase 2 from *Staphylococcus aureus* (25–26% identity; PDB: 3W7F), dehydrosqualene synthase (29–31%; 5IYS), squalene synthase HpnC from *Alicyclobacillus acidocaldarius* (27–29%; 4HD1), and putative phytoene/squalene synthase YisP from *Bacillus subtilis* (26–27%; 3WE9). The results indicated that similar to the structures of reference enzymes, that of tomato PSY1 was rich in anti-parallel  $\alpha$ -helices, which formed a large central cavity making the catalytic site.

Overall, we could model 288–300 aa-long fragments representing 69–71% of the PSY1 protein (including the 109–408-aa region with the HH-IPPS domain) with more than 90% confidence. Similar results (284–300-aa; 68–73%) were obtained for PSY1 homologs from *S. tuberosum*, *S. commersonii*, *S. verrucosum* (sect. Petota), *L. ruthenicum*, *L. chinense*, *L. barbarum*, *S. torvum*, *S. sisymbriifolium*, *S. melongena*, *S. prinophyllum*, *C. annuum*, *N. tabacum*, *P. axillaris*, and *A. thaliana*. The remaining 27–32% residues were modeled *ab initio*.

No species or fruit color-specific structural features were observed in TP or the HH-IPPS domain. Considering that reliable modeling covered the 114–401-aa region of PSY1, which included a domain conserved in dehydrosqualene synthases (also belonging to the isoprenoid biosynthesis C1 superfamily), the obtained data may support the predicted 1–129-aa location of TP in PSY1, which could extend as far as the end of the fourth  $\alpha$ -helix (L117–E129), i.e., up to the active site (130-YAKTF-134) (Figure 4). Although the TP sequence varied among plant PSY1 enzymes, its C-terminus ( $\alpha$ -helix with motif LSEAYDRCGEVCAE) was modeled with more than 90% confidence, suggesting its conservation as the region preceding the cleavage site.

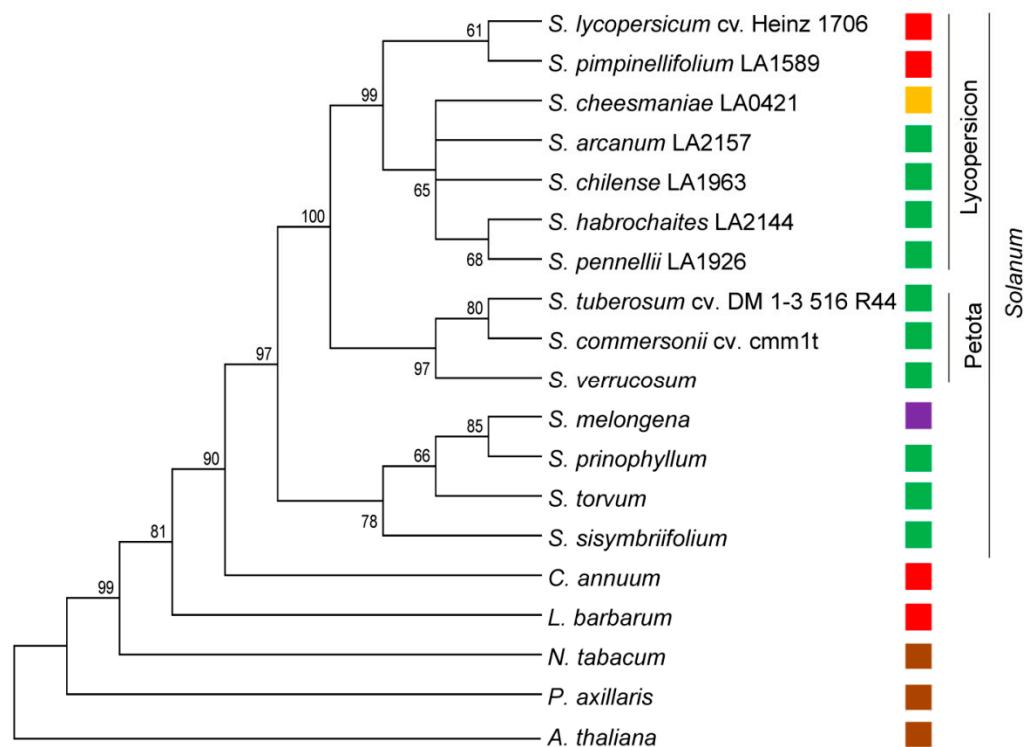


**Figure 4.** Spatial structure of *S. lycopersicum* cv. Heinz 1706 PSY1 built using Phyre2. Chains are colored according to the rainbow spectrum from the N- to C-terminus. Two residues predicted to flank TP (V61 and A129) are indicated by green and pink triangles, respectively.

#### 2.4. PSY1-Based Phylogeny of Tomato Species

The coding sequences of the identified and already known *PSY1* genes were used to analyze the phylogeny of the Solanaceae family. The foundation of the Solanaceae phylogenetic tree was occupied by *P. axillaris*, followed by *N. tabacum*, *L. barbarum*, and *C. annuum* (Figure 5). The *Solanum* species were the most recent; eggplant and the related species were grouped in the basic sub-cluster, whereas sister clades of the Petota and Lycopersicon species were positioned at the top of the tree (Figure 5). Tomatoes were divided into two main clades containing RF and GF species, respectively. The GF clade was rooted by the most ancient species *S. habrochaites* and *S. pennellii* grouped together, and included *YF S. cheesmaniae* (Figure 5).



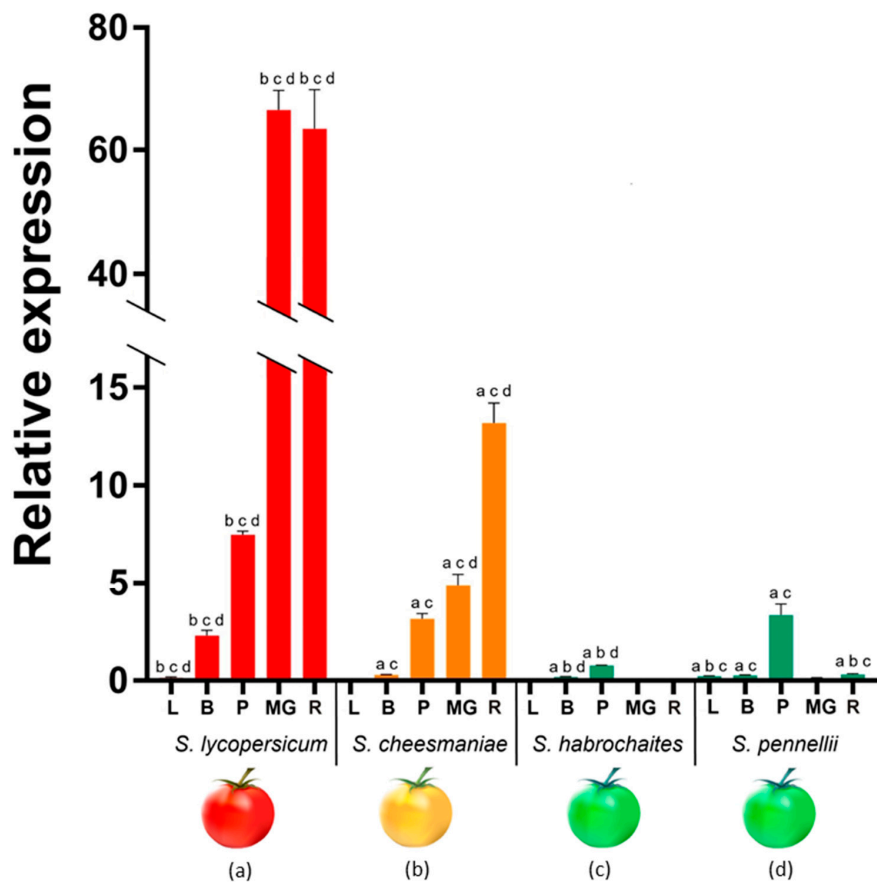


**Figure 5.** Evolutionary relationship among Solanaceae species (*Solanum lycopersicum*, *Solanum pimpinellifolium*, *Solanum cheesmaniae*, *Solanum arcanum*, *Solanum chilense*, *Solanum habrochaites*, *Solanum pennellii*, *Solanum tuberosum*, *Solanum commersonii*, *Solanum verrucosum*, *Solanum melongena*, *Solanum prinophyllum*, *Solanum torvum*, *Solanum sisymbriifolium*, *Capsicum annuum*, *Lycium barbarum*, *Nicotiana tabacum*, and *Petunia axillaris*) based on the PSY1 coding sequences. *Arabidopsis thaliana* (Brassicaceae) PSY was used as an outgroup. Analysis was performed using the neighbor-joining method. The optimal tree with the sum of branch length = 2.35980228 is shown. Percentages of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Jukes–Cantor method and are in the units of the number of base substitutions per site. For each species, the color of the ripe fruit is indicated to the right of the name.

### 2.5. PSY1 Expression Pattern

PSY1 expression was tested in young leaves, flower buds with green petals, yellow petals, and mature green (MG) and ripe fruit of four tomato species differing in fruit color—GF *S. habrochaites* LA2144 and *S. pennellii* LA0716, YF *S. cheesmaniae* LA0421, and RF *S. lycopersicum* (cv. Heinz 1706-BG). The aim was to analyze a possible correlation between the level of PSY1 expression and the content of carotenoids in fruits, and the choice of species, in addition to the ripe fruit color, was determined by the species phylogeny: RF *S. lycopersicum* as the most recent species, YF *S. cheesmaniae* as an evolutionary older species, and also the two most ancient GF species *S. pennellii* and *S. habrochaites*. *S. chilense* was not included in the comparative analysis; this species is another GF tomato, evolutionary located between the recent RF and YF and the most ancient GF species, and its fruits contained the same amount of carotenoids as other GF species.

In YF *S. cheesmaniae* and RF *S. lycopersicum*, PSY1 mRNA was expressed in all tested tissues, except for the leaves of *S. cheesmaniae*; PSY1 expression was consistently upregulated from a very low level in leaves to a high level in ripe fruit in both species; however, it was about five times higher in *S. lycopersicum* than in the *S. cheesmaniae* fruit (Figure 6).



**Figure 6.** *PSY1* mRNA expression in leaves (L), young flower buds (B), yellow petals (P), and mature green (MG), and ripe (R) fruit of *S. lycopersicum* cv. Heinz 1706, *S. cheesmaniae* LA0421, *S. habrochaites* LA2144, and *S. pennellii* LA0716. Low-case letters above the bars indicate statistically significant differences ( $p$ -value < 0.005) between gene expression levels in the same tissue of different species: *S. lycopersicum* (a), *S. cheesmaniae* (b), *S. habrochaites* (c), and *S. pennellii* (d). For example, *PSY1* expression in *S. lycopersicum* leaves differed significantly from those in *S. cheesmaniae*, *S. habrochaites* and *S. pennellii* leaves, which is, respectively, denoted by three letters “bcd” above the L-bar.

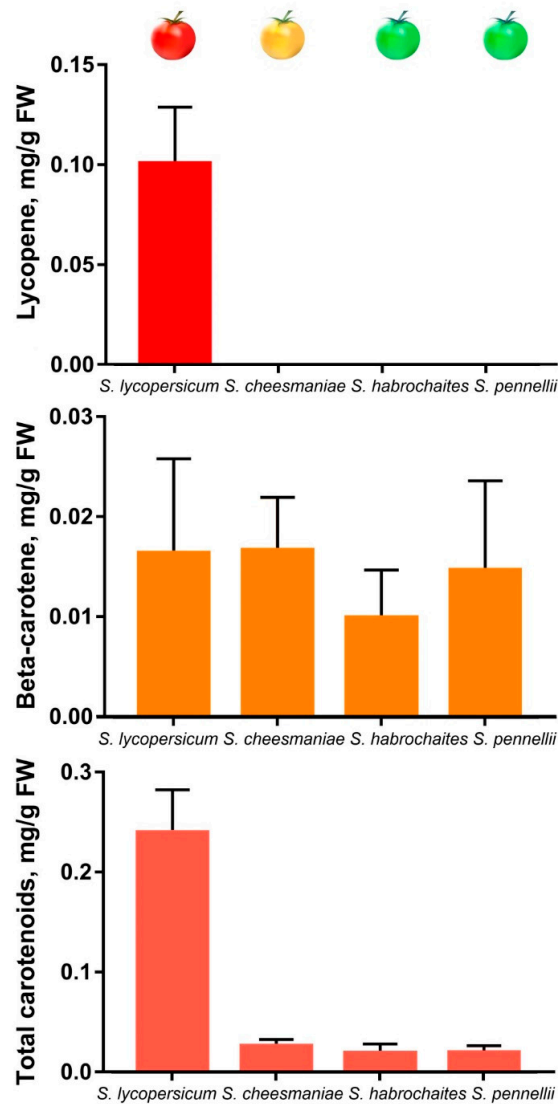
In GF *S. pennellii*, *PSY1* transcription was also detected in all tested tissues; however, in contrast to YF and RF species, it was maximal in yellow petals and minimal in MG and ripe fruit, where its level was approximately 5–13 times lower than in *S. cheesmaniae*, and 65 times lower than in *S. lycopersicum*. In another GF species *S. habrochaites*, *PSY1* mRNA was detected, at a very low level, only in buds and petals (Figure 6).

*PSY1* expression was minimal or absent in the leaves of all four tomato species and very low in buds of all wild accessions. The highest *PSY1* mRNA level in yellow petals was observed in *S. lycopersicum*, where it exceeded that in *S. cheesmaniae*/*S. pennellii* and *S. habrochaites* by two and 10 times, respectively (Figure 6).

## 2.6. Carotenoid and Chlorophyll Content

Accumulation of total (x + c) and specific carotenoids (lycopene and  $\beta$ -carotene), as well as colorless carotenoid precursors (phytoene and phytofluene) and chlorophylls *a* and *b*, was assessed in ripe fruit and leaf of GF species *S. habrochaites* (LA2144), *S. pennellii* (LA0716) and *S. chilense* (LA1963), YF *S. cheesmaniae* (LA0421) and RF *S. lycopersicum* (cv. Heinz 1706-BG). The results indicated that the fruit of wild tomatoes lacked lycopene and contained 10–20 times less of total carotenoids than those of *S. lycopersicum*; however, fruit  $\beta$ -carotene content was the same for all analyzed species (Table 2,

Figure 7). At the same time, the content of total carotenoids in fruit was significantly lower than in leaf (Table 2). However, leaf-specific carotenogenesis is mainly mediated by another phytoene synthase isoform PSY2; thus, a comparison of the carotenoid content in fruit and leaf could be correct if the ripe green fruit consisted only of photosynthetic tissues, which is obviously not the case. Chlorophyll content in the fruit of GF tomatoes was considerably reduced compared to the leaves, whereas in the fruit of YF *S. cheesmaniae* and RF *S. lycopersicum* it was absent at all (Table 2).



**Figure 7.** Carotenoid content in ripe fruit of tomato species *S. lycopersicum* cv. Heinz 1706, *S. cheesmaniae* LA0421, *S. habrochaites* LA2144, and *S. pennellii* LA0716.

**Table 2.** Chlorophyll and carotenoid content in ripe fruits and leaves of analyzed tomato accessions.

Accession	Ripe Fruit Pigment Content, $\mu\text{g/g}$ FW					Leaf Pigment Content, $\mu\text{g/g}$ FW	
	Chlorophyll (a + b)	Lycopene	Total Carotenoids (x + c)	$\beta$ -carotene	Other x + c ( $-\beta$ -carotene)	Total Carotenoids (x + c)	Chlorophyll (a + b)
<i>Solanum lycopersicum</i> cv. Heinz 1706	N/D	$0.09 \pm 0.02$	$0.24 \pm 0.030$	$0.01 \pm 0.008$	$\sim 0.23$	$0.94 \pm 0.098$	$3.19 \pm 0.390$
<i>Solanum cheesmaniae</i> LA 0421	N/D	N/D	$0.03 \pm 0.004$	$0.02 \pm 0.003$	$\sim 0.01$	$0.96 \pm 0.210$	$3.33 \pm 0.820$
<i>Solanum chilense</i> LA 1963	$0.03 \pm 0.004$	N/D	$0.01 \pm 0.004$	$0.01 \pm 0.001$	$\sim 0.00$	$0.89 \pm 0.070$	$2.78 \pm 0.260$
<i>Solanum habrochaites</i> LA 2144	$0.06 \pm 0.010$	N/D	$0.02 \pm 0.006$	$0.01 \pm 0.003$	$\sim 0.00$	$0.93 \pm 0.004$	$3.15 \pm 0.060$
<i>Solanum pennellii</i> LA 0716	$0.09 \pm 0.050$	N/D	$0.02 \pm 0.005$	$0.02 \pm 0.004$	$\sim 0.00$	$0.96 \pm 0.110$	$3.51 \pm 0.400$

N/D—not detected; FW—fresh weight.

In the fruit of GF species, the content of total carotenoids was almost equal to the  $\beta$ -carotene content (Figure 7, Table 2), which means the lowest possible amount of other types of colorful carotenoids. In the fruit of YF *S. cheesmaniae*,  $\beta$ -carotene accounted for about 2/3 of the total carotenoids (Figure 7, Table 2), which may indicate the accumulation of  $\alpha$ -carotene and lutein, as was previously shown for *yellow-fruited tomato 1* mutant fruit [37]. We assume that the fruit of *S. cheesmaniae* produce so little lycopene that it could be fully converted into  $\alpha$ - and  $\beta$ -carotene by lycopene cyclases.

The potential accumulation of phytoene and phytofluene was assessed by the absorption spectra of fruit extracts. The characteristic spectral features of colorless carotenoid precursors were revealed only in the red ripe fruit of *S. lycopersicum* cv. Heinz, while in the spectra of ripe fruit of GF and YF accessions, these spectral features were not recorded, which indicates a minor content of carotenoid precursors (Supplementary File).

### 2.7. Promoter and 5'-UTR Analysis

We next analyzed the promoter and 5'-UTR of the *PSY1* gene in two tomato species, RF *S. lycopersicum* cv. Heinz 1706 and its most distant wild relative GF *S. pennellii* LA0716. A comparison of the identified *S. lycopersicum* sequences with those from the NCBI database (*S. lycopersicum* cv. Heinz 1706; *PSY1* Gene ID: 543988, NC\_015440.3 [4350836 ... 4355976]) revealed no differences, whereas in the identified *S. pennellii* LA0716 *PSY1*, 10 nucleotides inserted and 31 SNPs were detected compared to *S. pennellii* LA0716 *PSY1* (Gene ID: CCXL01009615.1 [3669 ... 9559]).

The sizes of the promoter/5'-UTR were 906/1535 bp in *S. lycopersicum* and 919/1550 bp in *S. pennellii*. Compared to the *S. lycopersicum* *PSY1* promoter and 5'-UTR, the corresponding *S. pennellii* *PSY1* regions contained 56 and 63 SNPs, indicating 6% and 3.98% variability, respectively; these variability levels were significantly higher than those in the whole gene (from start to stop codon), cDNA and protein: 2.54%, 0.8% and 0.95% (86, 10 and 3 substitutions), respectively.

Search for sites important for *PSY1* transcription revealed that the *PSY1* promoter and 5'-UTR contained 37 types of regulatory elements related to tissue-specific expression and response to phytohormones and stress factors. Among them, 32 were common for *PSY1* of both species, including core promoter elements and 11 were species-specific; among the latter, there were light- and ethylene-responsive elements, sites associated with reactions to salicylic acid, cold and wounding and binding sites of the MYB transcription factor implicated in drought response (Table 3). Compared to the ancient *S. pennellii*, the *S. lycopersicum* *PSY1* promoter/5'-UTR region lost light-(chs-Unit 1 m1 and TCT motif), salicylic acid-(two TCA sites) and drought-(MBS) responsive elements and an AAGAA motif with unknown function, while acquiring other light-responsive elements (3-AF3 binding site and GATA motif) as well as ethylene-(ERE) and wound-(WUN-motif) responsive elements and an AACCTAACCT motif with unknown function (Table 3).

**Table 3.** Regulatory elements found in the *PSY1* promoter and 5'-UTR of RF *S. lycopersicum* and GF *S. pennellii*.

No	Type	Sequence	<i>S. lycopersicum</i> cv. Heinz 1706		<i>S. pennellii</i> LA0716		Comments
			Strand	Position	Strand	Position	
1	chs-Unit 1 m1			n/d	-	-1110	Part of a light responsive element
2	Box II	ACACGTAGA	-	-1713	-	-1735	
3	GATA-motif	AAGGATAAGG	+	-2207	n/d		
4	GTGGC-motif	GATTCTGTGGC	+	-564	+	-566	
5	TCT-motif			n/d	+	-2419	
6	AE-box	AGAAACAA	-	-528	+	-530	Part of a conserved DNA module involved in light response
7	Box 4	ATTAAT	+	-2373	+	-2402	
8	GA-motif	ATAGATAA		n/d	-	-2143	
9	I-box	TAGATAACC		n/d	+	-28	
10	3-AF3 binding site	CACTATCTAAC	+	-2321	n/d		Light response
11	GT1-motif	GGTAA	+	-1101	+	-1102	
				n/d	+	-1079	
12	ABRE	TACGTGTC	+	-1711	+	-1733	The binding sites for AREB/ABF factors involved in the abscisic acid response
		ACGTG	+	-675	+	-677	
			+	-1710	+	-1732	
13	ABRE3a	TACGTG	+	-1711	+	-1733	
			+	-676	+	-678	
14	ABRE4	CACGTA	-	-1711	-	-1733	
			-	-676	-	-678	
15	AT~ABRE	TACGTGTC	+	-1711	+	-1733	
16	ERE	ATTTTAAA	+	-1272	n/d		Ethylene-response
17	TGA-element	AACGAC	+	-2241	+	-2270	Auxin-response
			+	-2219	+	-2248	
18	CGTCA-motif	CGTCA	-	-1846	-	-1876	MeJA response
19	STRE	AGGGG	+	-1384	n/d		Defense and stress response
			-	-1459			
			+	-712	+	-714	
20	TC-rich repeats	GTTTTCTTAC	+	-16	+	-16	
			-	-1988	n/d		Essential for the anaerobic induction
			+	-1964			
21	ARE	AAACCA	-	-265	-	-265	
			+	-1527	+	-1560	
			-	-204	-	-204	
			-	-682	-	-684	

Table 3. Cont.

No	Type	Sequence	<i>S. lycopersicum</i> cv. Heinz 1706		<i>S. pennellii</i> LA0716		Comments
			Strand	Position	Strand	Position	
22	WUN-motif	AAATTACT	-	-102	-	-102	Wounding response
23	CAT-box	GCCACT	-	-1496	-	-1495	Related to meristem-specific expression
			-	<b>-1221</b>	-	<b>n/d</b>	
24	AT-TATA-box	TATATA	<b>n/d</b>	<b>-1528</b>	-	<b>n/d</b>	Enriched near transcription start. TATA-box-like, putative TBP-binding
			+	<b>-1391</b>	-	<b>n/d</b>	
			-	<b>-973</b>	-	<b>n/d</b>	
			-	-1007	-	-1007	
25	CAAT-box	CAAT/CAAAAT	-	-1159	-	-1164	Common cis-acting element in promoter and enhancer
			43 repeats		42 repeats		
26	TATA-box	TATAAAAT; TATAAATA; TATAAAAT; TATAAA; TATAA; TATA	multiple repeats		multiple repeats		Core promoter element
27	W box	TTGACC	-	-1894	-	-1924	WRKY TF binding site
28	MBS	CAACTG	<b>n/d</b>	<b>-1825</b>	-	<b>-1825</b>	MYB TF binding site; drought response
29	MRE	AACCTAA	-	-2275	-	-2286	MYB TF binding site; light response
			<b>n/d</b>	<b>-2304</b>	-	<b>-580</b>	
30	MYB	TAACCTG	+	-551	+	-553	MYB TF binding site
			<b>n/d</b>	<b>-1825</b>	-	<b>-1825</b>	
			+	<b>-289</b>	-	<b>n/d</b>	
			-	-517	-	-519	
			<b>n/d</b>	<b>-1538</b>	-	<b>-1538</b>	
			+	<b>-24</b>	-	<b>n/d</b>	
31	MYC	CATGTG	-	-787	-	-788	Originally found in some viral and bacterial T-DNA promoters. Pathogen-inducible
			-	-278	-	-278	
32	AS-1 (activation sequence-1)	TGACG	+	-1846	+	-1876	Binding site of AT-rich DNA binding protein (ATBP-1)
33	AT1-motif	AATTATTTTATT	-	<b>-2111</b>	-	<b>n/d</b>	
34	G-box	TACGTG	+	-1711	+	-1733	Multifunctional
			+	-676	+	-678	
35	Unnamed_2	AACCTAACCT	-	<b>-1107</b>	-	<b>n/d</b>	Unknown function
36	Unnamed_4	CTCC	13 repeats		13 repeats		
37	AAGAA-motif	GAAAGAA	<b>n/d</b>	<b>-1515</b>	+	<b>-1515</b>	

The elements that are different (present or not) between the two species are highlighted in bold and pink.

### 3. Discussion

The existing wide variety of tomato cultivars, which differ in fruit color and other economically important traits, are derived from the most evolutionary recent species *S. lycopersicum*. It is proposed that tomato domestication began in South America, when wild RF species *S. pimpinellifolium* gave rise to *S. lycopersicum* var. *cerasiforme*, further evolution of which in Mesoamerica resulted in *S. lycopersicum* var. *lycopersicum* [38,39] that, together with its 12 wild relatives, now form *Solanum* section *Lycopersicon* [3].

Tomato diversification due to domestication is distinct from the natural species divergence as it passed through various genetic bottlenecks caused by the selection of a limited set of traits valuable for humans, which significantly limited genomic diversity among modern *S. lycopersicum* cultivars [36,40,41]. As a result, *S. lycopersicum* differs from its wild ancestors in a wide range of acquired morphophysiological characteristics (so-called domestication syndrome) and in genetic diversity, which constitutes no more than 5% of that existing among wild tomatoes [39]. Fruit color and carotenoid content are included in the domestication syndrome; these features are controlled by qualitative trait genes, including *PSY1* encoding phytoene synthase, the key enzyme of carotenogenesis in fruit, which has a conserved structure but also contains polymorphisms associated with the evolution of plant species [42,43].

In the present study, *PSY1* homologous genes were identified and characterized in one cultivated (*S. lycopersicum* cv. Heinz) and four wild (*S. pennellii*, *S. habrochaites*, *S. chilense*, and *S. cheesmaniae*) tomato species that differ in ripe fruit color (Table 1). All identified homologs were highly conserved and had the same functionally important residues, indicating that the enzymes should have the same catalytic mechanism underlain by the formation of the C-terminal catalytic site—a large central cavity composed of two anti-parallel alpha-helices and motifs DELVD and DVGED at the opposite sides of the cavity (Figures 2 and 4) [44]. These data indicate that the catalytic function of the *PSY1* homologs in carotenogenesis is conserved in wild and domesticated tomato species.

The N-terminal TP is responsible for *PSY* translocation into plastids [10,12], and its polymorphism may be one of the main species-specific signatures in plants [45]. Indeed, in the analyzed tomato species, the N-terminal TP was found to be the most polymorphic *PSY1* region, which distinguished *Lycopersicon* species from the other *Solanaceae* members (Figures 3 and 4). Furthermore, it appears that in the course of evolution, TPs of *Capsicum* and *Solanum* *PSY1* enzymes have lost the FSCLGGSRTKNGRIFSVRSIAIVATPAGEM motif and acquired the CNERIKRG motif instead, which was later deleted in *Petota* and *Lycopersicon* species (Figure 3).

It has been suggested that carotenogenic enzymes can be used to shed light on the molecular evolution of different plant species [43]. Phylogenetic analysis based on whole-genome sequence data or trait-encoding genes produces similar results, to which exons contribute more evolution-related information than introns as they are under higher selective pressure [42]. Therefore, in this study, we examined the possibility of using *PSY1* coding sequences to evaluate *Solanaceae* phylogeny (Figure 5). The results indicated that the analyzed tomato species were grouped according to the generally accepted evolution history of nightshades, which suggests major splits between tomato and potato (eight million years ago [Mya]), eggplant and tomato/potato (14 Mya), *Solanum* and *Capsicum* (19 Mya), and *Solanum* and *Nicotiana* (24 Mya) [46]. Fleshy fruit developed during the latter split had the ability to acquire color because of carotenogenesis. In this study, the analyzed tomatoes could be further divided into two main clades of RF and GF species based on the *PSY1* sequence (Figure 5), indicating that *PSY1* may be a good phylogenetic marker to investigate the recent evolution of tomatoes.

In the GF clade, the most distant *S. habrochaites* and *S. pennellii* were grouped together, which is consistent with the incompletely resolved dichotomy between the two species belonging to different informal groups, *Eriopersicon* and *Neolycopersicon*, respectively [47]; sometimes, *S. pennellii* is considered to be an intermediate species between *S. habrochaites* and *S. arcanum* [42]. The other two informal groups are *Lycopersicon*, which includes all the YF-RF tomato species, and *Arcanum* [47]. Carotenoid accumulation is the main feature of *Lycopersicon* species: the red color of *S. lycopersicum* and *S. pimpinellifolium* fruit is caused by the accumulation of lycopene (up to 500-fold during ripening [40]), which in



*S. cheesmaniae* and *S. galapagense* is further processed into  $\beta$ -carotene accounting for yellow-to-orange fruit color [48,49]. In *Neolycopersicon*, *Eriopersicon*, and *Arcanum* species, the level of carotenoid biosynthesis in fruit is similar to that in photosynthetic tissues [49].

Considering the conserved structure of *PSY1* as the key carotenogenic enzyme in all tomato species, the question is why the fruit of *S. lycopersicum* turns red, whereas those of wild tomatoes remain green or only slightly yellow. The answer may lie in differential regulation of *PSY1* expression, the level of which positively correlates with carotenoid content. Thus, in maize, *PSY1* is transcribed in yellow endosperm but not in the carotenoid-deficient white endosperm [17,50] and in rice, *PSY1* overexpression makes white endosperm turn yellow because of  $\beta$ -carotene accumulation [51]. In RF *S. lycopersicum*, *PSY1* expression is low in photosynthetic tissues and high in petals and ripe fruit [11,12], whereas in GF species (*S. peruvianum*, *S. chilense*, *S. pennellii* and *S. chmielewskii*) *PSY1* is not expressed in ripe fruit, which do not accumulate lycopene and contain 100 times less  $\beta$ -carotene than those of RF tomatoes [31,32].

However, the question of why wild tomato ripe fruits remain green cannot be answered simply by the *PSY1* expression level. For instance, overexpression of *PSY* in *A. thaliana* does not turn photosynthetic leaves into orange or red, and the content of carotenoids there remains unchanged, while  $\beta$ -carotene accumulation in non-photosynthetic tissues sharply increases [52]. The green color of the ripe fruit in GF tomatoes may be due to the lack of special globular structures for carotenoid accumulation, which typically present in the chromoplasts of non-photosynthetic tissues, or to the impaired chromoplast formation [33,35,53,54]. In this regard, it is interesting that *PSY1* expression and *PSY1* enzymatic activity are quite significant in the fruit of GF *S. habrochaites* (accession LA1777); however, there is no carotenoid accumulation because of impaired chromoplast formation [33] due to a mutation in the ORANGE protein [35,54].

Consistent with these observations, in the current study, we detected *PSY1* transcription in yellow-colored petals and red- and yellow-colored fruits but not in leaves (Figure 6), which confirms chromoplast specificity of *PSY1* expression in tomato species [12].

The absence of yellow-to-red pigments in the ripe green fruit may result from a metabolic bottleneck and blockage of reactions following *PSY1*-catalyzed phytoene biosynthesis. Then the ripe green fruit could accumulate an excess of colorless precursors, which in RF tomatoes are rapidly processed to form colorful carotenoids. In this study, significant amounts of phytoene and phytofluene were observed in the ripe fruit of only RF species (and these amounts directly correlated with the content of colorful carotenoids), but not in the GF or YF tomatoes (Supplementary file). Apparently, there is no blocking of the reactions following the synthesis of phytoene and phytofluene, and the low content of colorful carotenoids in the ripe fruit of GF and YF tomatoes may be explained by low amounts of carotenoid precursors. Thus, the final content of carotenoids in the ripe fruit is largely controlled by the level of *PSY1* expression, determining the amount of colorless precursors.

Overall, in this study, the content of colorful carotenoids in ripe fruit directly correlated with the level of *PSY1* expression—highest in *S. lycopersicum*, moderate in *S. cheesmaniae*, and lowest in *S. habrochaites* and *S. pennellii* (Figures 6 and 7; Table 2).

*PSY1* of *S. habrochaites* LA2144 behaved differently from that of *S. habrochaites* LA1777 [33] and similar to the other analyzed GF species in that it was weakly expressed in ripe fruit and produced small amounts of carotenoids (Figures 6 and 7). The latter may be the result of  $\beta$ -carotene biosynthesis in remaining chloroplasts of green fruit tissues due to the expression of both *PSY1* and *PSY2* [12]. Ripe fruit of YF *S. cheesmaniae* had five times lower *PSY1* expression than those of RF *S. lycopersicum* cv. Heinz and accumulated  $\beta$ -carotene instead of lycopene (Figures 6 and 7), which was probably due to increased activity of chromoplast-specific lycopene beta-cyclase caused by polymorphisms [42,48].

The expression level of *PSY1* in RF tomatoes was much higher than in RF *C. annuum* [55], which was thought to compensate for low *PSY1* activity in tomato caused by Y136N mutation [10] that was detected in all *PSY1* homologs identified in the current study. Two other known substitutions that can affect the level of carotenoid biosynthesis (the C-terminal P to L substitution in grass *PSY-E1*, found

as a result of white endosperm selection, and capable of reducing the biosynthesis of carotenoids in grains [56], and A191D, leading to an increase in PSY enzymatic activity [57]) were not detected in the identified tomato PSY1 homologs.

In wheat, the phenotypic variation of grain yellow pigments is significantly associated with three *Psy1-A1* alleles, differing in the effect on carotenoid content [58]. In the fourth intron of *Psy1-A1a*, related to a large reduction in grain pigment, the 676-bp INDEL was detected, which could be linked to mutations in a regulatory region of the gene that alters its expression [58].

It may be proposed that species-specific differences in *PSY1* expression may depend on variations in the promoter/5'-UTR sequence, in particular, the presence and location of regulatory elements and transcription factor-binding sites. Thus, *S. lycopersicum* *PSY1* has been shown to contain 5'-UTR with MBS, ABRE, and TC-rich repeats [59]. In the 2.5-kb *PSY1* region comprising the promoter and 5'-UTR, we found 37 types of regulatory elements (Table 3), indicating that *PSY1* transcription may be regulated in response to light, abiotic stresses, and hormones such as methyl jasmonate (MeJA), auxins, abscisic acid (ABA), and ethylene. There were variations in the presence of regulatory elements between RF *S. lycopersicum* and GF *S. pennellii*, such as the loss of some and acquisition of the other light-responsive elements, and the lack of drought-responsive elements in *S. lycopersicum*, which may indicate lower resistance of tomato cultivars to abiotic stresses compared to their wild relatives. ABA is an apo-carotenoid involved in plant developmental processes, including seed maturation and dormancy and stress tolerance [60], whereas JA and auxins together promote resistance to necrotrophic pathogens [61]. The presence of hormone-responsive elements in *PSY1* 5'-UTR/promoter of both RF and GF species indicate the involvement of *PSY1* in the regulation of hormone-mediated signaling in tomato. Gibberellin (GAs)-responsive elements have been previously reported only in 5'-UTR of the *S. lycopersicum* *PSY2* gene [60] but not in the *PSY1* gene, despite the fact that GAs participate in fruit senescence and have the same precursor (GGPP) as carotenoids [61]. Although we also did not identify GAs-responsive elements in the *PSY1* 5'-UTR/promoter, we found them in *PSY1* intron and exon sequences—P-box (CCTTTTG) in introns IV and V of all GF species and GARE (TCTGTTG) in exon I and intron II of both RF and GF species and in intron III of *S. habrochaites* LA2144.

The fruit ripening process, including ethylene production, was shown to be independent of carotenogenesis [62]; nevertheless, we observed that *S. lycopersicum* acquired ethylene-responsive elements in the *PSY1* promoter/5'-UTR (Table 3), which could account for stronger induction of carotenogenesis during fruit ripening compared to *S. pennellii*, and, consequently, contribute to red pigmentation in fruit.

However, despite these differences, the overall arrangement of regulatory elements in the *PSY1* 5'-UTR/promoter is similar in RF *S. lycopersicum* and GF *S. pennellii*, suggesting a significant degree of conservation of the *PSY1* transcriptional regulatory mechanism between RF and GF species.

## 4. Materials and Methods

### 4.1. Plant Material

Accessions of tomato species (*Solanum* section *Lycopersicon*) (Table 1) were kindly provided by the Tomato Genetics Resource Center (USA; <https://tgrc.ucdavis.edu/>). Plants were grown in a greenhouse with temperature kept at 28 °C/23 °C during a 16-h/8-h day/night light cycle (light intensity, 300–400  $\mu\text{mol m}^{-2} \text{s}^{-1}$ ). Young leaves, whole flower buds (closed flowers in the pre-anthesis stage with green petals), colored (yellow) petals, and MG and ripe fruit were collected in two biological replicates and homogenized in liquid nitrogen. The MG stage was defined as firm green fruit of a final (maximal) size, and ripe fruit in GF species was defined by softness and in YF or RF species—by a change of color from green to yellow or red, respectively, due to chlorophyll degradation and carotenoid accumulation [36,63,64].

#### 4.2. Gene Identification

To amplify the full-length *PSY1* genes from tomato species, gene-specific primers (Table 4) were designed based on *PSY1* genomic sequences of *S. lycopersicum* (cv. Heinz 1706) available in NCBI GenBank (Gene ID: 543988, NC\_015440.3 [4350836 . . . 4355976], <http://www.ncbi.nlm.nih.gov/Genbank>, genome annotation releases), and Sol Genomics Network (Solyc03g031860.2.1, <https://solgenomics.net/>); manual revision of sequence polymorphisms and additional evaluation were performed using Primer3 (<http://frodo.wi.mit.edu/primer3/>). Genomic DNA was isolated from young leaves of a single plant of each species accession, as previously described [65], and used as a template (100 ng) for PCR amplification at the following conditions: initial denaturation at 94 °C for 10 min, 35 cycles of denaturation at 94 °C for 30 s, primer annealing at 55 °C for 30 s, and extension at 65 °C for 4.5 min, and a final extension at 65 °C for 10 min. The amplified PCR products of the expected size were purified by using the QIAEX® II Gel Extraction kit (QIAGEN, Hilden, Germany), cloned in the pGEM®-T Easy (Promega, Madison, WI, USA), and sequenced (3–5 clones for each accession) on ABI Prism 377 DNA Sequencer (Applied Biosystems, Waltham, MA, USA) using the designed primers (Table 4).

**Table 4.** List of primers for *PSY1* gene amplification, sequencing, and expression analysis.

Primer	Sequence (5'→3')	Application
PSY1geneF	AGTGGGAATCTACTAGGAGT	Gene amplification and sequencing
PSY1geneR	TTATCTTTGAAGAGAGACAGTTT	
tPSY_F6	CTATCTGGGCAATATATGGTG	Gene sequencing
tPSY_F7	TCTCGTCCTAGATACTACAC	
tPSY_F8	CAGTGACGAGCCATGATC	
tPSY_F9	TTGAGCTTGTCGTTCTCAGT	qRT-PCR, <i>PSY1</i>
PSY1rtF	CTGAGATCTACCAATGAGTTAG	
PSY1rtR	TCTCGGGAGTCATTAGCATAG	Promoter/5'-UTR amplification and sequencing
prPSY1-F	GTTGGATTTGCATGTAGACC	
tPSY_F2	GGCTAAATCGAAAATYGAATC	
tPSY_F3	TAACCTTCTATTGCTTTGCTAGTG	
tPSY_F4	TGGTAGGTAATATTGCTGATTTTG	qRT-PCR, reference genes
Actin 2/7-F	CATTGTGCTCAGTGGTGGTTC	
Actin 2/7F-R	TCTGCTGGAAGGTGCTAAGTG	
Expressed-F	GCTAAGAACGCTGGACCTAATG	
Expressed-R	TGGGTGTGCCTTTCTGAATG	

#### 4.3. Structural and Phylogenetic Analysis

Multiple sequence alignments and structural and phylogenetic analyses of *PSY1* genes and encoded proteins were conducted with MEGA 7.0 [66]. Phylogenetic dendrograms were constructed based on complete gene, cDNA, and protein sequences using the neighbor-joining [67] and Jukes-Cantor [68] methods; confidence for tree topologies was estimated by bootstrap values of 1000 replicates.

Predicted proteins were characterized in terms of molecular weight, isoelectric point (pI) (<http://isoelectric.ovh.org>; accessed on the 20 March 2020; [69]), secondary, tertiary, and quaternary structures (PHYRE2 [70]; visualization by Chimera-1.11.2, <http://www.cgl.ucsf.edu/chimera/download.html>; downloaded on February 6, 2017), the functional importance of amino acid residue substitutions (PROVEAN; [71]), and conserved domains, sites, and motifs (NCBI-CDD, <https://www.ncbi.nlm.nih.gov/cdd>; UniProt, <https://www.uniprot.org/>; and Multiple Em for Motif Elicitation (MEME 5.1.1) [72], <http://meme-suite.org/tools/meme>).

For comparative structural analysis, the complete sequences and/or transcripts of *PSY1* gene homologs were extracted from the NCBI database. We used *PSY1* from the model plant *Arabidopsis thaliana* (Gene ID: 831587) and from Solanaceae species: *Solanum* section Petota—*S. tuberosum* (NW\_006238968.1: 1057573-1062303), *S. verrucosum* (FYAE01567738.1: 183863-187835) and *S. commersonii* cv. cmm1t (JXZD01122484.1: 1000-7000); other *Solanum* sections—*S. sisymbriifolium* (GGFC02031958.1: 553-1815) *S. melongena* (GBGZ01079994.1: 146-1408), *S. prinophyllum* (374-1636), and *S. torvum* (51-1307); *Capsicum annuum* (NM\_001324967.1: 205328571-205334820); *Lycium* species—*L. ruthenicum* (KF957704.1), *L. chinense* (KJ624406.1), and *L. barbarum* (KF957680.1); *Nicotiana tabacum* (GDGU01109046.1: 486-1900); *Petunia axillaris* (GBRU01036418.1: 1-1560).

#### 4.4. Gene Expression

Total RNA was isolated from individual samples of young leaves, flower buds with green petals, yellow petals, and fruit using the RNeasy Plant Mini Kit (QIAGEN, Hilden, Germany), qualified by gel electrophoresis, and used for first-strand cDNA synthesis (Reverse Transcription System; Promega, Madison, WI, USA) with an oligo-dT primer. The RNA and cDNA concentrations were quantified by fluorimetry (Qubit® Fluorometer, Thermo Fisher Scientific, Waltham, MA, USA). Quantitative real-time (qRT)-PCR was performed with 2.5 ng of cDNA, SYBR Green RT-PCR mixture (Syntol, Moscow, Russia), and specific primers (Table 4) at the following cycling conditions: initial denaturation at 95 °C for 5 min and 40 cycles of denaturation at 95 °C for 15 s and annealing/extension at 60 °C for 40 s. To normalize the levels of *PSY1* expression, two reference tomato genes *Expressed* (SGN-U346908) and *actin 2/7* (NM\_001330119.1) [73,74] were used.

The qRT-PCR results were statistically analyzed with Graph Pad Prism version 7.02 (GraphPad Software Inc., San Diego, CA, USA; <https://www.graphpad.com/scientific-software/prism/>). The data were expressed as the mean ± standard deviation (SD) based on three technical replicates of two biological replicates. The unequal variance (Welch's) *t*-test was applied to assess statistical significance (*p*-value < 0.05) of differences in gene expression between tissues within the same species and between the same tissues of different tomato species.

#### 4.5. Carotenoid and Chlorophyll Content

Total carotenoid content was measured by spectrophotometry in two biological and three technical replicates using a modified Folch method [75,76]. Briefly, 0.2 g of plant tissue was homogenized in Folch solution (2:1 chloroform-methanol [v/v]) in the presence of trace Mg<sub>2</sub>CO<sub>3</sub> amounts [76], incubated at 4 °C for 1 h, and centrifuged at 4000 rpm for 10 min at 4 °C. The lower chloroform phase was collected and used for spectrophotometric assay of chlorophyll, lycopene, β-carotene, and total carotenoid contents were measured in acetone-hexane solutions, as previously described [77] using a spectrophotometer (Basic, Eppendorf, Hamburg, Germany) and the following equations:

$$\text{Chlorophyll } a \text{ (}\mu\text{g/mL)} = 11.47 (A_{666} - A_{750}) - 2 (A_{648} - A_{750})$$

$$\text{Chlorophyll } b \text{ (}\mu\text{g/mL)} = 21.85 (A_{648} - A_{750}) - 4.53 (A_{666} - A_{750})$$

$$\text{Total carotenoids (x + c) (}\mu\text{g/mL)} = [1000 (A_{480} - A_{750}) - 1.33 \text{ Chl } a - 23.93 \text{ Chl } b]/202$$

$$\text{Lycopene (mg/100 mL)} = 0.204 A_{645} - 0.0458 A_{663} + 0.372 A_{505} - 0.0806 A_{453}$$

$$\beta\text{-carotene (mg/100 mL)} = 0.216 A_{663} - 1.22 A_{645} - 0.304 A_{505} + 0.452 A_{453}$$

where  $A_{666}$ ,  $A_{648}$ ,  $A_{750}$ ,  $A_{663}$ ,  $A_{645}$ ,  $A_{505}$ ,  $A_{453}$  and  $A_{480}$  are absorbance values at the indicated wavelengths; x + c—the sum of xanthophylls and carotenes. The extracts were then evaporated in a stream of nitrogen and the pellets were re-dissolved in hexane. The absorption spectrum was measured in the range of 250–800 nm and used for screening for the presence of phytoene and phytofluene based on their distinct spectral signature (maximum absorption of phytoene—at 285 nm, and phytofluene—at 331, 347 and 365 nm) in the UV range (Supplementary file) [78].

#### 4.6. Promoter and 5'-UTR Analyses

1. The search of specific cis-elements in promoters and 5'-UTRs was performed using the PlantCARE database, which provides an evaluation of cis-regulatory elements, enhancers, and repressors [79]; (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>; accessed May 31, 2020).

### 5. Conclusions

In this study, we identified *PSY1* homologous genes in RF, YF, and GF tomato species of *Solanum* section *Lycopersicon*. *PSY1* homologs shared high similarity of structure, conserved motifs, and functionally important sites. *PSY1* transcription levels were species- and plant organ-specific, and directly correlated with carotenoid content in ripe fruit. Analysis of the *PSY1* promoter and 5'-UTR sequences revealed differences between GF and RF tomatoes, which could be attributed to domestication. Our results provide valuable data for further functional and evolutionary characterization of the carotenogenesis pathway in fleshy fruit in the Solanaceae family.

**Supplementary Materials:** The supplementary materials are available online at <http://www.mdpi.com/2223-7747/9/9/1169/s1>.

**Author Contributions:** Performed the experiments: E.Z.K. Provided plant material and carried out the experiment: M.A.S. and G.I.E. Analyzed the data: E.Z.K., A.V.B., A.V.S., O.A.S. Wrote the paper: A.V.S., G.I.E., M.A.S. and G.I.E. contributed to the manuscript equally. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Russian Science Foundation (grant no. 19-16-00016) and partially by the Ministry of Science and Higher Education of the Russian Federation.

**Acknowledgments:** We would like to thank Alexey Solovchenko for valuable advice on carotenoid spectrometry and Marina Chuenkova for English language editing.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

### References

1. Yoo, H.J.; Park, W.J.; Lee, G.M.; Oh, C.S.; Yeam, I.; Won, D.C.; Kim, C.K.; Lee, J.M. Inferring the genetic determinants of fruit colors in tomato by carotenoid profiling. *Molecules* **2017**, *22*, 764. [[CrossRef](#)] [[PubMed](#)]
2. Liu, L.; Shao, Z.; Zhang, M.; Wang, Q. Regulation of carotenoid metabolism in tomato. *Mol. Plant* **2015**, *8*, 28–39. [[CrossRef](#)] [[PubMed](#)]
3. Peralta, I.E.; Knapp, S.; Spooner, D.M. Nomenclature for wild and cultivated tomatoes. *Rep. Tomato Genet. Coop.* **2006**, *56*, 6–10.
4. Peralta, I.E.; Spooner, D.M. History, origin and early cultivation of tomato (solanaceae). *Genet. Improv. Solanaceous Crop.* **2006**, *2*, 1–24.
5. Galpaz, N.; Ronen, G.; Khalfia, Z.; Zamir, D.; Hirschberg, J. A chromoplast-specific carotenoid biosynthesis pathway is revealed by cloning of the tomato white-flower locus. *Plant Cell* **2006**, *18*, 1947–1960. [[CrossRef](#)]
6. Nambara, E.; Marion-Poll, A. Abscisic acid biosynthesis and catabolism. *Annu. Rev. Plant Biol.* **2005**, *56*, 165–185. [[CrossRef](#)]
7. Niyogi, K.K.; Truong, T.B. Evolution of flexible non-photochemical quenching mechanisms that regulate light harvesting in oxygenic photosynthesis. *Curr. Opin. Plant Biol.* **2013**, *16*, 307–314. [[CrossRef](#)]
8. Al-babili, S.; Bouwmeester, H.J. Strigolactones, a novel plant hormone. *Annu. Rev. Plant Biol.* **2015**, *66*, 161–186. [[CrossRef](#)]
9. Hashimoto, H.; Uragami, C.; Cogdell, R.J. Carotenoids and photosynthesis. *Sub-Cell. Biochem.* **2016**, *79*, 111–139. [[CrossRef](#)]
10. Cao, H.; Luo, H.; Yuan, H.; Eissa, M.A.; Thannhauser, T.W.; Welsch, R.; Hao, Y.J.; Cheng, L.; Li, L. A neighboring aromatic-aromatic amino acid combination governs activity divergence between tomato phytoene synthases. *Plant Physiol.* **2019**, *180*, 1988–2003. [[CrossRef](#)]
11. Giuliano, G. Plant carotenoids: Genomics meets multi-gene engineering. *Curr. Opin. Plant Biol.* **2014**, *19*, 111–117. [[CrossRef](#)]

12. Giorio, G.; Stigliani, A.L.; D'Ambrosio, C. Phytoene synthase genes in tomato (*Solanum lycopersicum* L.)—New data on the structures, the deduced amino acid sequences and the expression patterns. *FEBS J.* **2008**, *275*, 527–535. [[CrossRef](#)] [[PubMed](#)]
13. Moise, A.R.; Al-Babili, S.; Wurtzel, E.T. Mechanistic aspects of carotenoid biosynthesis. *Chem. Rev.* **2014**, *114*, 164–193. [[CrossRef](#)] [[PubMed](#)]
14. Ducreux, L.J.M.; Morris, W.L.; Hedley, P.E.; Shepherd, T.; Davies, H.V.; Millam, S.; Taylor, M.A. Metabolic engineering of high carotenoid potato tubers containing enhanced levels of  $\beta$ -carotene and lutein. *J. Exp. Bot.* **2005**, *56*, 81–89. [[CrossRef](#)] [[PubMed](#)]
15. Chen, L.; Li, W.; Li, Y.; Feng, X.; Du, K.; Wang, G.; Zhao, L. Identified trans-splicing of YELLOW-FRUITED TOMATO 2 encoding the PHYTOENE SYNTHASE 1 protein alters fruit color by map-based cloning, functional complementation and RACE. *Plant Mol. Biol.* **2019**, *100*, 647–658. [[CrossRef](#)] [[PubMed](#)]
16. Zhou, X.; Welsch, R.; Yang, Y.; Álvarez, D.; Riediger, M.; Yuan, H.; Fish, T.; Liu, J.; Thannhauser, T.W.; Li, L. Arabidopsis OR proteins are the major posttranscriptional regulators of phytoene synthase in controlling carotenoid biosynthesis. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 3558–3563. [[CrossRef](#)] [[PubMed](#)]
17. Gallagher, C.E.; Matthews, P.D.; Li, F.; Wurtzel, E.T. Gene duplication in the carotenoid biosynthetic pathway preceded evolution of the grasses. *Plant Physiol.* **2004**, *135*, 1776–1783. [[CrossRef](#)]
18. Fu, X.; Feng, C.; Wang, C.; Yin, X.; Lu, P.; Grierson, D.; Xu, C.; Chen, K. Involvement of multiple phytoene synthase genes in tissue- and cultivar-specific accumulation of carotenoids in loquat. *J. Exp. Bot.* **2014**, *65*, 4679–4689. [[CrossRef](#)]
19. Bartley, G.E.; Viitanen, P.V.; Bacot, K.O.; Scolnik, P.A. A tomato gene expressed during fruit ripening encodes an enzyme of the carotenoid biosynthesis pathway. *J. Biol. Chem.* **1992**, *267*, 5036–5039.
20. Bartley, G.E.; Scolnik, P.A. cDNA cloning, expression during development and genome mapping of PSY2, a second tomato gene encoding phytoene synthase. *J. Biol. Chem.* **1993**, *268*, 25718–25721.
21. Sato, S.; Tabata, S.; Hirakawa, H.; Asamizu, E.; Shirasawa, K.; Isobe, S.; Kaneko, T.; Nakamura, Y.; Shibata, D.; Aoki, K.; et al. The tomato genome sequence provides insights into fleshy fruit evolution. *Nature* **2012**, *485*, 635–641. [[CrossRef](#)]
22. Fraser, P.D.; Schuch, W.; Bramley, P.M. Phytoene synthase from tomato (*Lycopersicon esculentum*) chloroplasts—Partial purification and biochemical properties. *Planta* **2000**, *211*, 361–369. [[CrossRef](#)]
23. Stauder, R.; Welsch, R.; Camagna, M.; Kohlen, W.; Balcke, G.U.; Tissier, A.; Walter, M.H. Strigolactone levels in dicot roots are determined by an ancestral symbiosis-regulated clade of the PHYTOENE SYNTHASE gene family. *Front. Plant Sci.* **2018**, *9*. [[CrossRef](#)] [[PubMed](#)]
24. Li, F.; Vallabhaneni, R.; Wurtzel, E.T. PSY3, a new member of the phytoene synthase gene family conserved in the poaceae and regulator of abiotic stress-induced root carotenogenesis. *Plant Physiol.* **2008**, *146*, 1333–1345. [[CrossRef](#)] [[PubMed](#)]
25. Dibari, B.; Murat, F.; Chosson, A.; Gautier, V.; Poncet, C.; Lecomte, P.; Mercier, I.; Bergès, H.; Pont, C.; Blanco, A.; et al. Deciphering the genomic structure, function and evolution of carotenogenesis related phytoene synthases in grasses. *BMC Genom.* **2012**, *13*. [[CrossRef](#)] [[PubMed](#)]
26. Ahrazem, O.; Diretto, G.; Argandoña Picazo, J.; Fiore, A.; Rubio-Moraga, Á.; Rial, C.; Varela, R.M.; Macías, F.A.; Castillo, R.; Romano, E.; et al. The specialized roles in carotenogenesis and apocarotenogenesis of the phytoene synthase gene family in saffron. *Front. Plant Sci.* **2019**, *10*. [[CrossRef](#)] [[PubMed](#)]
27. Arango, J.; Wüst, F.; Beyer, P.; Welsch, R. Characterization of phytoene synthases from cassava and their involvement in abiotic stress-mediated responses. *Planta* **2010**, *232*, 1251–1262. [[CrossRef](#)]
28. Costa, M.G.C.; Moreira, C.D.; Melton, J.R.; Otoni, W.C.; Moore, G.A. Characterization and developmental expression of genes encoding the early carotenoid biosynthetic enzymes in *Citrus paradisi* Macf. *Mol. Biol. Rep.* **2012**, *39*, 895–902. [[CrossRef](#)]
29. Soltis, D.E.; Albert, V.A.; Leebens-Mack, J.; Palmer, J.D.; Wing, R.A.; DePamphilis, C.W.; Ma, H.; Carlson, J.E.; Altman, N.; Kim, S.; et al. The Amborella genome: An evolutionary reference for plant biology. *Genome Biol.* **2008**, *9*. [[CrossRef](#)]
30. Giovannoni, J.; Nguyen, C.; Ampofo, B.; Zhong, S.; Fei, Z. The epigenome and transcriptional dynamics of fruit ripening. *Annu. Rev. Plant Biol.* **2017**, *68*, 61–84. [[CrossRef](#)]
31. Meléndez-Martínez, A.J.; Fraser, P.D.; Bramley, P.M. Accumulation of health promoting phytochemicals in wild relatives of tomato and their contribution to in vitro antioxidant activity. *Phytochemistry* **2010**, *71*, 1104–1114. [[CrossRef](#)] [[PubMed](#)]

32. Bolger, A.; Scossa, F.; Bolger, M.E.; Lanz, C.; Maumus, F.; Tohge, T.; Quesneville, H.; Alseekh, S.; Sørensen, I.; Lichtenstein, G.; et al. The genome of the stress-tolerant wild tomato species *Solanum pennellii*. *Nat. Genet.* **2014**, *46*, 1034–1038. [[CrossRef](#)] [[PubMed](#)]
33. Kilambi, H.V.; Manda, K.; Rai, A.; Charakana, C.; Bagri, J.; Sharma, R.; Sreelakshmi, Y. Green-fruited *Solanum habrochaites* lacks fruit-specific carotenogenesis due to metabolic and structural blocks. *J. Exp. Bot.* **2017**, *68*, 4803–4819. [[CrossRef](#)] [[PubMed](#)]
34. Luo, Z.; Zhang, J.; Li, J.; Yang, C.; Wang, T.; Ouyang, B.; Li, H.; Giovannoni, J.; Ye, Z. A STAY-GREEN protein SISR1 regulates lycopene and  $\beta$ -carotene accumulation by interacting directly with SPSY1 during ripening processes in tomato. *New Phytol.* **2013**, *198*, 442–452. [[CrossRef](#)]
35. Osorio, C.E. The role of orange gene in carotenoid accumulation: Manipulating chromoplasts toward a colored future. *Front. Plant Sci.* **2019**, *10*. [[CrossRef](#)]
36. Tanksley, S.D.; McCouch, S.R. Seed banks and molecular maps: Unlocking genetic potential from the wild. *Science* **1997**, *277*, 1063–1066. [[CrossRef](#)]
37. Gao, L.; Zhao, W.; Qu, H.; Wang, Q.; Zhao, L. The yellow-fruited tomato 1 (yft1) mutant has altered fruit carotenoid accumulation and reduced ethylene production as a result of a genetic lesion in ETHYLENE INSENSITIVE2. *Theor. Appl. Genet.* **2016**, *129*, 717–728. [[CrossRef](#)]
38. Blanca, J.; Cañizares, J.; Cordero, L.; Pascual, L.; Diez, M.J.; Nuez, F. Variation revealed by SNP genotyping and morphology provides insight into the origin of the Tomato. *PLoS ONE* **2012**, *7*, e48198. [[CrossRef](#)]
39. Mata-Nicolás, E.; Montero-Pau, J.; Gimeno-Paez, E.; Garcia-Carpintero, V.; Ziarsolo, P.; Menda, N.; Mueller, L.A.; Blanca, J.; Cañizares, J.; van der Knaap, E.; et al. Exploiting the diversity of tomato: The development of a phenotypically and genetically detailed germplasm collection. *Hortic. Res.* **2020**, *7*, 66. [[CrossRef](#)]
40. Bai, Y.; Lindhout, P. Domestication and breeding of tomatoes: What have we gained and what can we gain in the future? *Ann. Bot.* **2007**, *100*, 1085–1094. [[CrossRef](#)]
41. Sim, S.C.; Robbins, M.D.; Van Deynze, A.; Michel, A.P.; Francis, D.M. Population structure and genetic differentiation associated with breeding history and selection in tomato (*Solanum lycopersicum* L.). *Heredity* **2011**, *106*, 927–935. [[CrossRef](#)] [[PubMed](#)]
42. The 100 Tomato Genome Sequencing Consortium. Exploring genetic variation in the tomato (*Solanum* section *Lycopersicon*) clade by whole-genome sequencing. *Plant J.* **2014**, *80*, 136–148. [[CrossRef](#)] [[PubMed](#)]
43. Liang, M.H.; Zhu, J.; Jiang, J.G. Carotenoids biosynthesis and cleavage related genes from bacteria to plants. *Crit. Rev. Food Sci. Nutr.* **2018**, *58*, 2314–2333. [[CrossRef](#)]
44. Shumskaya, M.; Bradbury, L.M.T.; Monaco, R.R.; Wurtzela, E.T. Plastid localization of the key carotenoid enzyme phytoene synthase is altered by isozyme, allelic variation and activity. *Plant Cell* **2011**, *24*, 3725–3741. [[CrossRef](#)] [[PubMed](#)]
45. Da Silva Mendes, A.F.; Fontes Soares, V.n.L.c.; Cardoso Costa, M.G. Carotenoid biosynthesis genomics. *Pigment. Fruits Veg.* **2015**, 9–29.
46. Särkinen, T.; Bohs, L.; Olmstead, R.G.; Knapp, S. A phylogenetic framework for evolutionary study of the nightshades (Solanaceae): A dated 1000-tip tree. *BMC Evol. Biol.* **2013**, *13*. [[CrossRef](#)] [[PubMed](#)]
47. Peralta, I.E.; Spooner, D.M.; Knapp, S. Taxonomy of wild tomatoes and their relatives (*Solanum* sect. *Lycopersicoides*, sect. *Juglandifolia*, sect. *Lycopersicon*; Solanaceae). *Syst. Bot. Monogr.* **2008**, *84*, 1–186.
48. Ronen, G.; Carmel-Goren, L.; Zamir, D.; Hirschberg, J. An alternative pathway to  $\beta$ -carotene formation in plant chromoplasts discovered by map-based cloning of Beta and old-gold color mutations in tomato. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 11102–11107. [[CrossRef](#)]
49. Paran, I.; Van Der Knaap, E. Genetic and molecular regulation of fruit and plant domestication traits in tomato and pepper. *J. Exp. Bot.* **2007**, *58*, 3841–3852. [[CrossRef](#)] [[PubMed](#)]
50. Palaisa, K.A.; Morgante, M.; Williams, M.; Rafalski, A. Contrasting effects of selection on sequence diversity and linkage disequilibrium at two phytoene synthase loci. *Plant Cell* **2003**, *15*, 1795–1806. [[CrossRef](#)]
51. Ye, X.; Beyer, P. Engineering the provitamin A ( $\beta$ -carotene) biosynthetic pathway into (carotenoid-free) rice endosperm. *Science* **2000**, *287*, 303–305. [[CrossRef](#)] [[PubMed](#)]
52. Lätari, K.; Wüst, F.; Hübner, M.; Schaub, P.; Beisel, K.G.; Matsubara, S.; Beyer, P.; Welsch, R. Tissue-specific apocarotenoid glycosylation contributes to carotenoid homeostasis in *Arabidopsis* leaves. *Plant Physiol.* **2015**, *168*, 1550–1562. [[CrossRef](#)] [[PubMed](#)]

53. Fraser, P.D.; Enfissi, E.M.A.; Halket, J.M.; Truesdale, M.R.; Yu, D.; Gerrish, C.; Bramley, P.M. Manipulation of phytoene levels in tomato fruit: Effects on isoprenoids, plastids and intermediary metabolism. *Plant Cell* **2007**, *19*, 4131–4132. [[CrossRef](#)] [[PubMed](#)]
54. D'Andrea, L.; Rodriguez-Concepcion, M. Manipulation of plastidial protein quality control components as a new strategy to improve carotenoid contents in Tomato fruit. *Front. Plant Sci.* **2019**, *10*. [[CrossRef](#)] [[PubMed](#)]
55. Romer, S.; Huguency, P.; Bouvier, F.; Camara, B.; Kuntz, M. Expression of the genes encoding the early carotenoid biosynthetic enzymes in *Capsicum annuum*. *Biochem. Biophys. Res. Commun.* **1993**, *196*, 1414–1421. [[CrossRef](#)]
56. Zhang, W.; Dubcovsky, J. Association between allelic variation at the Phytoene synthase 1 gene and yellow pigment content in the wheat grain. *Theor. Appl. Genet.* **2008**, *116*, 635–645. [[CrossRef](#)]
57. Welsch, R.; Arango, J.; Bär, C.; Salazar, B.; Al-Babili, S.; Beltrán, J.; Chavarriaga, P.; Ceballos, H.; Tohme, J.; Beyer, J. Provitamin A accumulation in cassava (*Manihot esculenta*) roots driven by a single nucleotide polymorphism in a phytoene synthase gene. *Plant Cell* **2010**, *22*, 3348–3356. [[CrossRef](#)]
58. Singh, A.; Reimer, S.; Pozniak, C.J.; Clarke, F.R.; Clarke, J.M.; Knox, R.E.; Singh, A.K. Allelic variation at Psy1-A1 and association with yellow pigment in durum wheat grain. *Theor. Appl. Genet.* **2009**, *118*, 1539–1548. [[CrossRef](#)]
59. Qin, X.; Coku, A.; Inoue, K.; Tian, L. Expression, subcellular localization and cis-regulatory structure of duplicated phytoene synthase genes in melon (*Cucumis melo* L.). *Planta* **2011**, *234*, 737–748. [[CrossRef](#)]
60. Dong, T.; Park, Y.; Hwang, I. Abscisic acid: Biosynthesis, inactivation, homeostasis and signalling. *Essays Biochem.* **2015**, *58*, 29–48. [[CrossRef](#)]
61. Bürger, M.; Chory, J. Stressed Out about hormones: How plants orchestrate immunity. *Cell Host Microbe* **2019**, *26*, 163–172. [[CrossRef](#)] [[PubMed](#)]
62. Fraser, P.D.; Truesdale, M.R.; Bird, C.R.; Schuch, W.; Bramley, P.M. Carotenoid biosynthesis during tomato fruit development. Evidence for tissue-specific gene expression. *Plant Physiol.* **1994**, *105*, 405–413. [[CrossRef](#)]
63. Grumet, R.; Fobes, J.F.; Herner, R.C. Ripening behavior of wild tomato species. *Plant Physiol.* **1981**, *68*, 1428–1432. [[CrossRef](#)] [[PubMed](#)]
64. Chevalier, C.; Nafati, M.; Mathieu-Rivet, E.; Bourdon, M.; Frangne, N.; Cheniclet, C.; Renaudin, J.P.; Gvaudant, F.; Hernould, M. Elucidating the functional role of endoreduplication in tomato fruit development. *Ann. Bot.* **2011**, *107*, 1159–1169. [[CrossRef](#)] [[PubMed](#)]
65. Puchooa, D. A simple, rapid and efficient method for the extraction of genomic DNA from lychee (*Litchi chinensis* Sonn.). *Afr. J. Biotechnol.* **2004**, *3*, 253–255.
66. Kumar, S.; Stecher, G.; Tamura, K. MEGA7: Molecular evolutionary genetics analysis version 7.0. molecular biology and evolution. *Mol. Biol. Evol.* **2016**, *33*, 1870–1874. [[CrossRef](#)]
67. Saitou, N.; Nei, M. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **1987**. [[CrossRef](#)]
68. Jones, D.T.; Taylor, W.R.; Thornton, J.M. The rapid generation of mutation data matrices. *Comput. Appl. Biosci.* **1992**, *8*, 275–282. [[CrossRef](#)]
69. Kozlowski, L.P. IPC—Isoelectric Point Calculator. *Biol. Direct* **2016**, *11*, 55. [[CrossRef](#)]
70. Kelley, L.A.; Mezulis, S.; Yates, C.M.; Wass, M.N.; Sternberg, M.J.E. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **2015**, *10*, 845–858. [[CrossRef](#)]
71. Choi, Y.; Sims, G.E.; Murphy, S.; Miller, J.R.; Chan, A.P. Predicting the functional effect of amino acid substitutions and indels. *PLoS ONE* **2012**, *7*, e46688. [[CrossRef](#)] [[PubMed](#)]
72. Bailey, T.L.; Elkan, C. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* **1994**, *2*, 28–36. [[PubMed](#)]
73. Expósito-Rodríguez, M.; Borges, A.A.; Borges-Pérez, A.; Pérez, J.A. Selection of internal control genes for quantitative real-time RT-PCR studies during tomato development process. *BMC Plant Biol.* **2008**, *8*, 131. [[CrossRef](#)]
74. Bemer, M.; Karlova, R.; Ballester, A.R.; Tikunov, Y.M.; Bovy, A.G.; Wolters-Arts, M.; de Barros Rossetto, P.; Angenent, G.C.; de Maagd, R.A. The tomato fruitfull homologs tdr4/ful1 and mbp7/ful2 regulate ethylene-independent aspects of fruit ripening. *Plant Cell* **2012**, *24*, 4437–4451. [[CrossRef](#)] [[PubMed](#)]
75. Folch, J.; Lees, M.; Sloane, G.H. A simple method for the isolation and purification of total lipides from animal tissues. *J. Biol. Chem.* **1953**, *226*, 497–509.
76. Solovchenko, A.E.; Chivkunova, O.B.; Merzlyak, M.N.; Reshetnikova, I.V. A spectrophotometric analysis of pigments in apples. *Russ. J. Plant Physiol.* **2001**, *48*, 693–700. [[CrossRef](#)]



77. Nagata, M.; Yamashita, I. A simple method for simultaneous determination of chlorophyll and carotenoids in tomato fruit. *J. Jpn. Soc. Food Sci. Technol.* **1992**, *39*, 925–928. [[CrossRef](#)]
78. Meléndez-Martínez, A.J.; Mapelli-Brahm, P.; Benítez-González, A.; Stinco, C.M. A comprehensive review on the colorless carotenoids phytoene and phytofluene. *Arch. Biochem. Biophys.* **2015**, *572*, 188–200. [[CrossRef](#)]
79. Rombauts, S.; Déhais, P.; Van Montagu, M.; Rouzé, P. PlantCARE, a plant cis-acting regulatory element database. *Nucleic Acids Res.* **1999**, *27*, 295–296. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).