Genome-Wide Identification and Expression Analysis of Anthocyanin Biosynthesis Pathway Genes *PAL* and *CHS* in Colored Potato Tuber

Lina Shang

College of Agronomy and Biotechnology, Southwest University, Chongqing

Mengyuan Wan

College of Agronomy and Biotechnology, Southwest University, Chongqing

Yi Fu

College of Agronomy and Biotechnology, Southwest University, Chongqing

Rongrong Liu

College of Agronomy and Biotechnology, Southwest University, Chongqing

Zhonghui Jin

College of Agronomy and Biotechnology, Southwest University, Chongqing

Jichun Wang

College of Agronomy and Biotechnology, Southwest University, Chongqing

Baigeng Hu

National Engineering Research Center for Potato, Dezhou

Hongju Jian

College of Agronomy and Biotechnology, Southwest University, Chongqing

Dianqiu Lyu (✉ smallpotatoes@126.com)

College of Agronomy and Biotechnology, Southwest University, Chongqing

Research Article

Keywords: colored potato, anthocyanins, PAL, CHS, expression pattern analysis

Posted Date: February 7th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1299381/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background

Colored potato tubers are rich in anthocyanins, which are of great values for human health. Phenylalanine ammonia lyase (PAL) and chalcone synthase (CHS) play key roles in anthocyanin synthesis pathway. The *PAL* and *CHS* genes have been characterized in several plant species, but a comprehensive analysis of these two gene families in potato has not yet been reported.

Results

In this study, Eight *StuPAL* and 12 *StuCHS* genes were identified. Except for *StuPAL4*, *StuPAL8*, *StuCHS3* and *StuCHS7*, which did not contain introns, the other members of the *PAL* and *CHS* family all contained 1-2 intron. The motif characteristics of *StuPAL4* and *StuPAL8* were different from those of other members, but both of them contained L-Aspartase-like feature domains. In addition, the number of *StuPAL* genes in potato genome was twice that of *Arabidopsis* genome, suggesting that genome duplication may occur during the evolution of potato. In the *CHS* family, the motif characteristics of *StuCHS3* and *StuCHS7* were different from other members, and *StuCHS5* only contained three motifs. All *StuCHS* genes members contained Chal_sti_synt_C or Chal_sti_synt_N characteristic structures. Furthermore, 68 and 61 miRNA targets were predicted in *StuPAL* and *StuCHS* genes, respectively, indicating that their functions were accurately regulated. A number of cis-acting elements in response to environmental stresses were identified in the promoter sequences of *StuPAL* and *StuCHS* genes, indicating that their expression level was precisely regulated. The anthocyanin content in peel and pulp among three colored potato varieties was measured. Different expression patterns in different organs were observed for *StuPAL* and *StuCHS* genes. Interestingly, *StuPAL* and *StuCHS* genes exhibited differential expression in peel and pulp. *StuPAL* genes were mainly expressed in the peel of colored potato, but almost not expressed in the pulp. It was speculated that *StuPAL* genes were mainly involved in the synthesis of anthocyanins in the colored potato tuber peel. However, the expression of *StuCHS* family genes were quite different in the peel and pulp of colored potato.

Conclusion

In this work, the function of potato anthocyanin biosynthesis pathway genes was initially analyzed, which laid the foundation for further clarifying the molecular mechanism in colored potato.

Background

Anthocyanins are the most abundant class of plant flavonoid pigments [1]. They are distributed in the stem, leaf, flower, fruit, seed and other tissues of plants, making them present yellow, red, purple, blue and other colors [2]. In plants, flavonoids, including anthocyanins, are involved in many biological processes, such as enabling flower organs to have specific colors to attract pollinators, pollen tube germination, adaptation to ultraviolet radiation, and protection against pests and fungi [3–7]. There are six main chromophores of anthocyanin, namely anthocyanin, cyanidin, methyl anthocyanin, delphinidin, petunia and mallow anthocyanin. Anthocyanin derivatives usually produce red, cyanidin derivatives produce purple, delphinidin appears purple or blue [8]. Anthocyanins have attracted much attention due to their nutritional and pharmacological effects such as scavenging free radicals, anti-platelet coagulation, and prevention of cardiovascular diseases, immunomodulatory activity and anticancer activity [9–11].

Anthocyanin is a branch product of flavonoid metabolic pathway. Its direct precursor is phenylalanine [12]. It can be divided into three stages from phenylalanine to anthocyanin formation. The first stage is the phenylpropane pathway, which is common to many secondary metabolites. In this stage, phenylalanine is deaminated by PAL to form trans-cinnamic acid [13]. Cinnamic acid 4-hydroxylase (C4H) was hydroxylated to form trans-4-coumaric acid [14]. Coumaric acid coenzyme A ligase (4CL) catalyzed the formation of coumaric acid coenzyme A [15]. When coumaric acid is activated, it can enter the flavonoid pathway. The second stage is the key reaction stage of flavonoid metabolism. CHS catalyzes the synthesis of 4-hydroxychalcone (aglycone) from coumaric acid and malonyl coenzyme A, which is converted to naringin by chalcone isomerase (CHI). Naringin forms dihydroflavonols catalyzed by flavanone trihydroxylase [16, 17]. At this stage, there are many branches of flavonoid pathway, such as flavonol, tannin, proanthocyanidins, isoflavones and other important metabolic pathways. The third stage is the formation of anthocyanins catalyzed by dihydroflavonol 4-reductase (DFR) and anthocyanin synthase/colorless anthocyanin dioxygenase (ANS/LDOX) [18, 19]. The formed anthocyanins enter a specific anthocyanin synthesis pathway as substrates. After a series of glycosylation, methylation and acylation modifications, stable anthocyanin glycosides

are formed [20]. Anthocyanin glycosides can be transported to vacuoles via vacuolar transporters such as glutathione S-transferase (GST), and collected and stored in vacuoles [21].

Potato, one of the top four major food crops, is widely cultivated in the world [22]. Potato tubers are rich in nutrients, including starch, sugar and protein, as well as a variety of antioxidants, such as polyphenols, vitamin C, carotenoids, selenium and so on [23]. Anthocyanidin, a water-soluble natural pigment, belongs to polyphenols. Anthocyanin is a natural antioxidant since its antioxidant activity, anti-aging and anti-vascular hardening functions [11]. It is a potential anticancer compound, which are largely abundant in colored potatoes. As an important secondary metabolite in potatoes, anthocyanin has attracted more and more attention, and become one of the hot spots in the development of functional food. However, the research field of potato anthocyanin is relatively backward and still in its infancy.

PAL and *CHS* genes played key roles in anthocyanin synthesis pathway identified in *Arabidopsis thaliana*, so we used the protein sequences of *PAL* and *CHS* genes in *A. thaliana* as the query sequences to search the corresponding homologous gene sequences in the genomes of 18 species including potatoes in this study and the evolutionary relationship were analyzed based on the evolutionary tree. The molecular characteristics of potato *PAL* and *CHS* genes were also predicted. The tissue expression characteristics, stress response, hormone response were analyzed based on the existing expression data. Further, expression differences of colored potato *PAL* and *CHS* genes were analyzed by qRT-PCR. The function of potato anthocyanin biosynthetic pathway genes was preliminarily analyzed, which laid a foundation for further functional analysis.

Results

Identification and Evolution Analysis of *PAL* and *CHS* Genes in 18 Plant Species

Genome-wide identification of *PAL* and *CHS* genes in 18 plant species, which not only included major classes like monocots and dicots, but also represent diverse groups such as cereals (*Sorghum bicolor*, *Oryza sativa*, *Setaria italic* and *Zea mays*), fruits (*Vitis vinifera*), trees (*Populus trichocarpa*, *Amygdalus persica*, *Citrus sinensis* and *Malus pumila*), vegetables (*Brassica oleracea*, *Brassica rapa*, *Cucumis sativus* and *Lycopersicon esculentum*), legumes (*Glycine max*), *Gossypium*, *Solanum tuberosum* and model dicot and monocot species (*Arabidopsis thaliana* and *Brachypodium distachyon*). Shown in Table 1, the 13 dicots used in this study have an average genome size of 618.86 Mb compared with that of 780.54 Mb average genome size of the 5 monocots. In addition, higher number of chromosomes were observed in 13 dicots (14.2 in average) compared with that in 5 monocots (9.2 in average). In this study, 135 *PAL* and 479 *CHS* genes were identified in the 18 plant species (Table 1 and Table S2). For *PAL* genes, the maximum number of *PAL* genes was found in *Zea mays* (12) and *Vitis vinifera* (11), respectively. More number of *PAL* genes were found in monocots (9.4 in average) than that in dicots (6.8 in average). Similar results were also obtained in *CHS* genes, namely more gene number in monocots (38.2 in average) than that in dicots (22.2 in average). The maximum number of *CHS* genes was detected in *Glycine max* (43) and *Sorghum bicolor* (55) in dicots and monocots, respectively. Notably, 12 *CHS* and 8 *PAL* genes were identified in *Solanum tuberosum*.

Table 1
Genome-wide identification of *PAL* and *CHS* genes in 18 plant species.

Species	Chromosomes	Genome size (MB)	Gene number	PAL	CHS
<i>Arabidopsis thaliana</i>	5	119.167	39,551	4	4
<i>Brassica oleracea</i>	9	514.431	56,687	6	24
<i>Brassica rapa</i>	10	314.865	47,942	7	29
<i>Cucumis sativus</i>	7	323.986	25,668	9	10
<i>Glycine max</i>	20	984.88	71,219	7	43
<i>Gossypium</i>	26	2189.14	90,927	8	22
<i>Vitis vinifera</i>	19	427.191	76,662	11	22
<i>Malus pumila</i>	17	702.961	51,695	8	29
<i>Citrus sinensis</i>	9	323.528	45,387	4	25
<i>Solanum tuberosum</i>	24	705.934	37,966	8	12
<i>Populus trichocarpa</i>	19	434.29	51,717	5	37
<i>Amygdalus persica</i>	8	212.767	32,595	2	14
<i>Solanum lycopersicum</i>	12	792.038	37,660	9	18
<i>Brachypodium distachyum</i>	5	218.345	37,892	8	21
<i>Zea mays</i>	10	2182.61	52,470	12	33
<i>Setaria italica</i>	9	441.705	35,844	10	39
<i>Sorghum bicolor</i>	10	687.75	39,248	8	54
<i>Oryza sativa</i>	12	372.31	67,393	9	43

Based on protein sequence similarities, the evolutionary trees of *PAL* and *CHS* gene family were constructed using MEGA 7.0 software (Fig. 1). For the *PAL* family, 135 protein sequences were divided into 8 subgroups (A-H). The members of subgroups A and B were monocots specific, while the others belonged to dicots (Fig. 1A). For the *CHS* family, 479 *CHS* protein sequences were divided into six subgroups, of which B and D branches belonged only to monocotyledons, while the other subgroups were common in both monocotyledons and dicotyledon (Fig. 1B).

Molecular Characterization, Chromosomal Location, Gene Duplication, Gene Structure and Motif Analysis of *PAL* and *CHS* Genes in Potato

Information of amino acid number, predicted molecular weight, theoretical isoelectric point (pI) and subcellular localization for StuPALs and StuCHSs were shown in Table 2. These StuPALs varied in length from 378 to 706 aa, with Mw ranging from 42.09 to 76.93 kDa, and pI from 5.50 to 6.28. And these genes encode proteins located mainly in cytoplasm (StuPAL4 and StuPAL8), endoplasmic reticulum membrane (StuPAL3, StuPAL5 and StuPAL7), peroxisome (StuPAL2) and plasma membrane (StuPAL1 and StuPAL6). For StuCHS proteins, they varied in length from 248 to 389 aa, with Mw ranging from 27.66 to 42.96 kDa, and pI from 5.38 to 9.57. Five of them were located in cytoplasm (StuCHS1, StuCHS3, StuCHS7, StuCHS11 and StuCHS12), four in endoplasmic reticulum membrane (StuCHS2, StuCHS4, StuCHS5 and StuCHS6), one in mitochondrial matrix space (StuCHS9), one in nucleus (StuCHS8) and one in plasma membrane (StuCHS10).

Table 2
Molecular characterization of *StuPAL* and *StuCHS* genes.

Gene name	Gene ID	Transcript	PR	MW	pI	Location	EN	SL
<i>StuPAL1</i>	PGSC0003DMG400005492	PGSC0003DMT400014003	650	71.58	5.50	chr05: 36342746- 36347409	2	plasma membrane
<i>StuPAL2</i>	PGSC0003DMG400019386	PGSC0003DMT400049886	689	75.46	6.28	chr10: 6301474- 6305153	2	microbody (peroxisome)
<i>StuPAL3</i>	PGSC0003DMG401021549	PGSC0003DMT400055488	706	76.91	6.08	chr09: 5510117- 5512923	2	endoplasmic reticulum
<i>StuPAL4</i>	PGSC0003DMG402021549	PGSC0003DMT400055489	378	42.09	6.16	chr09: 5502194- 5505885	1	cytoplasm
<i>StuPAL5</i>	PGSC0003DMG402021564	PGSC0003DMT400055531	706	76.91	6.19	chr09: 5530471- 5533348	2	endoplasmic reticulum
<i>StuPAL6</i>	PGSC0003DMG400023458	PGSC0003DMT400060308	703	76.93	6.07	chr05: 51694756- 51698709	2	plasma membrane
<i>StuPAL7</i>	PGSC0003DMG400031365	PGSC0003DMT400080548	694	75.65	5.97	chr10: 51926201- 51930606	2	endoplasmic reticulum
<i>StuPAL8</i>	PGSC0003DMG400031457	PGSC0003DMT400080765	435	48.34	5.84	chr03: 17248040- 17249552	1	cytoplasm
<i>StuCHS1</i>	PGSC0003DMG400001635	PGSC0003DMT400004133	368	40.53	5.59	chr01: 87143395- 8714521	3	cytoplasm
<i>StuCHS2</i>	PGSC0003DMG400008632	PGSC0003DMT400022254	388	42.96	6.18	chr12: 58758677- 58761269	2	endoplasmic reticulum
<i>StuCHS3</i>	PGSC0003DMG400008633	PGSC0003DMT400022255	329	36.03	8.52	chr12: 58741442- 58742633	1	cytoplasm
<i>StuCHS4</i>	PGSC0003DMG400008634	PGSC0003DMT400022258	389	42.66	5.47	chr12: 58746401- 58747957	2	endoplasmic reticulum
<i>StuCHS5</i>	PGSC0003DMG400012670	PGSC0003DMT400032993	248	27.66	9.57	chr02: 41086823- 41089074	2	endoplasmic reticulum
<i>StuCHS6</i>	PGSC0003DMG400016867	PGSC0003DMT400043447	388	42.96	7.49	chr12: 58934717- 58937241	2	endoplasmic reticulum
<i>StuCHS7</i>	PGSC0003DMG400016867	PGSC0003DMT400043449	293	31.85	5.38	chr12: 58934717- 58937241	1	cytoplasm
<i>StuCHS8</i>	PGSC0003DMG400016873	PGSC0003DMT400043464	388	42.57	7.49	chr12: 58954773- 58956693	2	nucleus
<i>StuCHS9</i>	PGSC0003DMG400019110	PGSC0003DMT400049165	385	42.17	5.96	chr05: 48886960- 48888729	2	mitochondrial matrix space
<i>StuCHS10</i>	PGSC0003DMG400027146	PGSC0003DMT400069814	382	42.09	7.71	chr05: 48437160- 48440351	2	plasma membrane

Gene name	Gene ID	Transcript	PR	MW	pl	Location	EN	SL
<i>StuCHS11</i>	PGSC0003DMG400029620	PGSC0003DMT400076178	385	42.15	6.27	chr09: 58356999-58359082	2	cytoplasm
<i>StuCHS12</i>	PGSC0003DMG400029621	PGSC0003DMT400076179	377	41.56	8.47	chr09: 58361561-58364070	2	cytoplasm
Note: PR: Peptide residues; MW: Molecular weight (kDa); EN: Exon number; SL: Subcellular localization.								

StuPALs and *StuCHSs* genes were unevenly located on chromosomes and they were distributed on four (Chr3, Chr5, Chr9 and Chr10) and five chromosomes (Chr1, Chr2, Chr5, Chr9 and Chr12), respectively (Fig. 2). *StuPALs* and *StuCHSs* genes were mainly distributed on chromosome Chr9 and Chr12, respectively (Fig. 2). Tandem and segmental duplication were the major duplication patterns, which led the expansion of gene families in the process of plant evolution [24, 25]. In this study, tandem duplicated genes of *StuPALs* (*StuPAL3* and *StuPAL4*) and *StuCHSs* (*StuCHS2*, *StuCHS3*, *StuCHS4*, *StuCHS6*, *StuCHS7* and *StuCHS8*) were indicated by yellow blocks in Fig. 2, which suggested that gene duplication played critical roles in the expansion of *StuPALs* and *StuCHSs* gene families.

To better explore the structural features of the *StuPAL* and *StuCHS* genes, the coding domain sequences and corresponding genomic sequences of the same *StuPALs* and *StuCHSs* gene were submitted to GSDS website together to display their exon/intron features. The results revealed the number of introns per gene varied from zero to a maximum of two (Fig. 3). All *StuPAL* and *StuCHS* genes had only one or zero intron except *StuCHS1*, which had two introns. Six *StuPAL* and nine *StuCHS* genes had one intron and two *StuPAL* (*StuPAL4* and *StuPAL8*) and two *StuCHS* (*StuCHS3* and *StuCHS7*) genes had no intron.

To further understand the potential functions of *StuPAL* and *StuCHS* genes, 10 motifs were screened within each protein sequence using the MEME website. The detailed information of predicted motifs was showed in Fig. 4 and Table 3. The length of predicted motifs was 21-50 aa and 8-50 aa for *StuPAL* and *StuCHS* proteins, respectively. Among them, motif 1, 2, 3, 4, 10 were detected in all eight *StuPAL* genes, *StuPAL1* and *StuPAL4* lacked motif 5 as well as *StuPAL4* and *StuPAL8* lacked motif 6, 7, 8, 9. Interestingly, the same motif appears more than once in the same *StuPAL* member, such as motif 1, 4 and 6. The same patterns were also detected in *StuCHS* proteins. To be noteworthy, only three motifs (motif 2, 4, 10) were screened in *StuCHS5* (Fig. 4).

Table 3
Motif detail information of *StuPAL* and *StuCHS* genes predicted using MEME.

	Motif	Sequences	Width	Functional domains
<i>StuPALs</i>	1	DYGFKGAEIAMASYCSELQFLANPVTNHVQSAEQHNQDVNS	41	L-Aspartase-like
	2	KQDRYALRTSPQWLGPQIEVIRAATKMIEREINSVNDNPLIDVSRNKALH	50	L-Aspartase-like
	3	ISARKTAEAVDILKLMSSSTYLVALCQAIDLRHLEENLKNAVKN TVSQVAK	50	L-Aspartase-like
	4	VRSPGEEIDKVFTAMCNGQIIDPLLECLK	29	-
	5	GKPEFTDYLT HKLKHHPGQIEAAIMEHILDGSSYV KAAQKLHEMDPLQK	50	L-Aspartase-like
	6	GNGTETCHTLPHSATRAAMLVRINTLLQGYSGIRFEILEAI	41	Fumarase/histidase N-terminal, L-Aspartase-like
	7	ASSDWVMDSMSKGTDSYGVTTGFGATSHRRTKNGGALQKELIRFLNAGVF	50	Fumarase/histidase N-terminal, L-Aspartase-like
	8	GGFFELQPKEGLALVNGTAVGSGMASMVLFD SNILAVMSEVLSAIFAEVM	50	L-Aspartase-like
	9	MAADSLRGSHLDEVKKMVDEFKPIVKLGG	30	-
	10	ANGELHPARFCEKELLRVVDR	21	Phenylalanine ammonia-lyase
<i>StuCHSs</i>	1	ITHLVFCTTSGVDLPGADYQLTKLLGLEPSVKRFMMYQQGC	41	Chal_sti_synt_N
	2	IGSDPIMNVEKPLFELVFATQTLLPDSEH	29	-
	3	GLKVQIHKDTPMLISKNIERILVEAFQPLDISDWNSIFWVS	41	Chal_sti_synt_C
	4	DQIELKLGLKPEKLKATRNVLSDYGNMASACVLFVLD EMRKTSIKAGLGT	50	Chal_sti_synt_C
	5	PSNCVDQSTYPDYYFRITNSEHKTELKEKF KRMCDKSMIKKRYMHLTEEI	50	Chal_sti_synt_N
	6	NPNICEYMAPSLDARQDIVVVEVPKLGKEAAQKAIKEWGQP	41	Chal_sti_synt_N
	7	VCFRNPNETEVLVAQALFSDGASAVI	28	Chal_sti_synt_N
	8	TGEGLEWGVLF GFGP	15	-
	9	EIRRTQRAMGPATVLAIGTAN	21	-
	10	HPGGRAIL	8	-

For *StuPAL* genes, seven motifs (1-3, 5-8) represented L-Aspartase-like (IPR008948) domain, and motif 10 represented Phenylalanine ammonia-lyase, shielding domain superfamily (IPR023144). While no known domain was detected in motif 4 and 9. Interestingly, Fumarase/histidase N-terminal, active site (IPR022313) domain was also detected in motif 6 and 7. The conserved motifs 1, 5, 6, and 7 represented the Chal-sti-synt-N domain, and motifs 3 and 4 possessed the Chal-sti-synt-C domain and no known domain was detected in the rest motifs among *StuCHSs* (Table 3).

MicroRNA Targeting Prediction of *StuPAL* and *StuCHS* Genes

To explore the potential roles of miRNAs involved in regulation of *StuPALs* and *StuCHSs* genes, the software psRNATarget Server was used to predict possible miRNAs based on the genomic sequences of all *StuPALs* and *StuCHSs* genes. In total, 68 and 61 putative miRNAs targeting all eight *StuPALs* and twelve *StuCHSs* genes, respectively, were detected and constructed the relationship network using Cytoscape software (Fig. 5). We further analysis the regulation network and found that *StuPAL 1*, *StuPAL 6* and *StuPAL 7* were targeted by the top three miRNAs. The stu-miR8015-3p targeted *StuPAL 1*, *StuPAL 2*, *StuPAL 3* and *StuPAL 4* (Fig. 5A). For *StuCHS* genes, *StuCHS10* was targeted by 23 miRNAs and stu-miR8040-3p targeted *StuCHS8*, *StuCHS10* and *StuCHS12*. *StuCHS6* and *StuCHS7* are targeted by some common miRNAs. (Fig. 5B).

StuPAL s and *StuCHSs* Genes Cis-acting Element Analysis

Gene expression was mainly regulated at the transcriptional level and cis-acting elements regulated the precise initiation and transcriptional efficiency of gene transcription by binding to transcription factors. Cis-acting elements in *StuPAL* and *StuCHS* gene

promoters were analyzed using the PlantCARE database. In our study, promoter sequences (1500 bp from the coding start) of *StuPALs* and *StuCHSs* genes were used to screen putative *cis*-acting elements, which involved in auxin (TGA-element and AuxRR-core), salicylic acid (TCA-element), abscisic acid responsiveness (ABRE), gibberellin (TATC-box, GARE-motif and P-box), MeJA (CGTCA-motif and TGACG-motif), defense and stress responsiveness (TC-rich repeats), drought-inducibility (MBS), low-temperature responsiveness (LTR), MYBHv1 binding site (CCAAT-box), endosperm expression (GCN4), cell cycle regulation (MSA-like) and the anaerobic induction (ARE). In total, 70 and 116 *cis*-acting elements were detected in *StuPAL* and *StuCHS* gene promoters (Fig. 6). Among them, 3 (*StuPAL2*)-12 (*StuPAL5* and 7) and 4 (*StuCHS8*)-13 (*StuCHS11*) *cis* acting regulatory elements were varied in two gene families. In *StuPAL* gene promoters, 14 ABREs and 13 AREs counted the top two elements (Fig. 6A). And 21 ABA responsive and 31 MeJA responsive elements were detected in *StuCHS* promoters, indicating that *StuCHS* genes may be regulated by the two hormones (Fig. 6B).

Expression Analysis of *StuPALs* and *StuCHSs* Genes in Various Tissues of Potato

The expression patterns of genes are usually correlated with their functions. To analysis the expression patterns of *StuCHSs* and *StuPALs* genes in various tissues and organs, RNA-Seq data from PGSC were downloaded and analyzed. A heatmap was generated using the FPKM data of all *StuCHS* and *StuPAL* genes by the software Mev 11.0 (Fig. 7). Thirteen different tissues and organs were included in the analysis. All *StuPALs* genes showed high expression levels in all tissues except *StuPAL1*, which showed very low expression level in all tissues. Notably, *StuPAL4* had the highest expression levels in all tissues except in stamens. *StuPAL5* had relatively low expression levels in stamens (Fig. 7A). For *StuCHS* genes, their expression patterns were very diverse (Fig. 7B). *StuCHS5*, *StuCHS9* and *StuCHS11* (except in callus and stolons) showed constitutive expression in all detected tissues. *StuCHS1* was only expressed in roots and tubers, while *StuCHS6* and *StuCHS7* were only highly expressed in shoots, sepals and petioles. *StuCHS3*, *StuCHS4*, *StuCHS8* and *StuCHS10* showed little expression in detecting tissues. Highly expressed in shoots, sepals, stolons, petioles, petals and fruits was detected in *StuCHS2*, while *StuCHS12* showed high expression level in sepals, stamens, flowers, petals, carpels and fruit (Fig. 7B).

Expression Analysis of *StuPAL* and *StuCHS* Genes in Potato under Environmental Stresses

To further insight into the roles of *StuPAL* and *StuCHS* genes in response to abiotic (salt, drought and heat), biotic (PIL, BTH and BABA) and hormone (ABA, BAP, GA₃ and IAA) stresses, we analysis the fold changes of FPKM using the expression data obtained from the Spud DB database (Fig. 8). For abiotic stresses, seven of eight *StuPAL* genes (except *StuPAL8*) were down-regulated after salt and drought stresses and *StuPAL6* and *StuPAL7* were strongly up-regulated in response to heat stress (Fig. 8A). The expression patterns were varied in *StuCHS* genes in response to abiotic stresses (Fig. 8B). *StuCHS3* and *StuCHS12* were up-regulated ($\log_2$ fold change>1) in both salt and drought stresses, while only *StuCHS8* was significantly down-regulated in salt stress. *StuCHS9*, *StuCHS11* and *StuCHS12* were strongly up-regulated in heat stress. The rest *StuCHS* gene members have little response ($|\log_2$ fold change|<1). For biotic stresses, all *StuPAL* genes were down-regulated in response to PIL (except *StuPAL4* and *StuPAL8*) and BTH (except *StuPAL7*) stresses while *StuPAL4* and *StuPAL6* were significantly up-regulated in response to BABA stress (Fig. 8A). All *StuCHS* genes were down-regulated in response to BABA and PIL (except *StuCHS1*) stress, while five and four *StuCHS* genes were up- and down-regulated in response to BTH stress and the rest three *StuCHS* genes had no response to BTH stress (Fig. 8B). For exogenous hormone treatments, all *StuPAL* genes had no express in response to ABA, GA₃ (except *StuPAL1*) and IAA (except *StuPAL6* and *StuPAL7*), while all except *StuPAL1* were down-regulated in response to BAP treatment (Fig. 8A). No *StuCHS* genes were induced by BAP and IAA, while *StuCHS11* and *StuCHS12* were strongly up-regulated after GA₃ treatment. For ABA treatment, five *StuCHS* genes, namely *StuCHS1*, *StuCHS3*, *StuCHS8*, *StuCHS11* and *StuCHS12*, were up-regulated while *StuCHS2*, *StuCHS4* and *StuCHS10* were down-regulated (Fig. 8B).

Anthocyanin content analysis in three colored potato varieties

Three colored varieties of potato were selected in this study. Among them, JYS are purple peel with purple pulp, while both peel and pulp of 1-11 is red and 16-A2 shows yellow both peel and pulp (Fig. 9A). The content of anthocyanins in the peel and pulp were significant difference among the three materials. The anthocyanin content was the highest in JYS peel while it was the lowest in 16-A2 pulp ($P < 0.05$) (Fig. 9B).

Expression Analysis of *StuPAL* and *StuCHS* Genes in Peel and Pulp of Colored Potato

To further analysis the functions of *StuPAL* and *StuCHS* genes in anthocyanin biosynthesis pathways, the expression levels of *StuPAL* and *StuCHS* genes in three colored potato varieties were detected using qRT-PCR (Fig. 10). For *StuPAL* genes, *StuPAL1*-*StuPAL8* were mainly expressed in the peel of the three potato varieties, but almost not expressed in the pulp (Fig. 10A). Among them, *StuPAL2*, *StuPAL3*, *StuPAL4* and *StuPAL6* had the highest expression in JYS peel, followed by 1-11 peel, and the lowest expression in 16-A2 peel.

The expression patterns of the rest genes are different. *StuPAL 1*, *StuPAL 5*, *StuPAL 7* and *StuPAL 8* had the lowest expression in 1-11 peel. Among them, *StuPAL 1* had the highest expression in JYS peel, followed by 16-A2 peel. The expression levels of *StuPAL 5* and *StuPAL 8* in JYS and 16-A2 peel were basically the same, while *StuPAL 7* had the highest expression in 16-A2 peel, followed by JYS peel. However, the expression of *StuCHS* genes in the peel and pulp of the three varieties changed greatly (Fig. 10B). *StuCHS1*, *StuCHS4*, and *StuCHS6* had almost no changes in the three different colored potatoes peel. *StuCHS3* was highly expressed in JYS peel, followed by 1-11 peel, and almost no expression in 16-A2 peel. *StuCHS5* had the highest expression in JYS peel, and there was little difference between 1-11 and 16-A2 peel. *StuCHS7* had the highest expression in 16-A2 peel and almost no expression between JYS and 1-11 peel. However, *StuCHS9* and *StuCHS11* had the highest expression in 1-11 peel, followed by JYS peel, and almost no expression in 16-A2 peel. Most of the *StuCHS* genes were expressed at low levels in the pulp, except *StuCHS1* was clearly expressed in JYS pulp. *StuCHS3*, *StuCHS9* and *StuCHS11* were hardly expressed in the pulp among three potato varieties. The expression of *StuCHS5* in JYS, 1-11 and 16-A2 pulp did not change much. *StuCHS4*, *StuCHS6* and *StuCHS7* was obviously expressed in JYS and 16-A2 pulp, but almost not expressed in 1-11 pulp. The expression of *StuCHS2*, *StuCHS8*, *StuCHS10* and *StuCHS12* were lower in the three varieties, and the data was not shown.

Discussion

Bioinformatics analysis of PAL and CHS genes in higher plants

As the first and rate-limiting enzyme in the phenylpropanoid pathway, PAL plays a key role in plant growth, development and environmental adaptation [26]. In this study, the BLAST alignment program was performed and eight *StuPAL* genes were obtained (Table 1). Our results showed that significant differences of the *PAL* gene family members among various species were observed (Fig. 1). The genome duplication may occur in the evolution of potato because the number of this gene in potato genome is twice that in *A. thaliana* genome. Duplication events including genome duplications, segmental duplications and tandem duplications may contribute to the expansion and evolution of gene families [27]. *StuPAL3* and *StuPAL4* were identified tandem duplication gene pair based on their chromosomal distribution and phylogenetic relationships (Fig. 2). In previous studies, the *CHS* gene family acts key roles in plant growth and development and contains smaller gene family in most plant genomes [28]. For example, eight *CHS* genes in *Petunia hybrid* [29], six in *Ipomoea* [30], four in *Arabidopsis thaliana* [31] and fourteen in *Zea mays* [28] were detected. In our study, 12 *CHS* genes were identified in potato and it was considered that tandem duplication was the major force for the gene family expansion during evolutionary process. Six tandemly duplicated genes were found in 12 *StuCHS* genes (Fig. 2). Gene structure and conserved motif analysis provide more evidences for the evolutionary relationship of multigene family and gene members with close evolutionary relationships have similar gene structure and motif composition [32, 33]. In this study, gene structure and motif characteristics of *PAL* and *CHS* gene members in potato also accorded with this pattern (Fig. 3 and Fig. 4). In the *PAL* family, except that *StuPAL4* and *StuPAL8* do not contain introns, the other members all contain one intron (Fig. 3). Similarly, the motif characteristics of *StuPAL4* and *StuPAL8* are different from those of other members, but both contain L-Aspartase-like feature domains. In the *CHS* gene family, *StuCHS3* and *StuCHS7* do not contain introns, and the other members all contain 1-2 introns. The motif characteristics of *StuCHS3* and *StuCHS7* are different from other members, and *StuCHS5* only contains three motifs (Fig. 4). All gene members contain Chal_sti_synt_C or Chal_sti_synt_N characteristic structures (Table 3). In addition, the software psRNATarget Server was used to predict the potential miRNA of *StuPAL* and *StuCHS* genes. The results showed that *StuPAL* and *StuCHS* genes had 68 and 61 miRNA respectively, indicating that their functions were accurately regulated. *StuCHS2*, *StuCHS5*, *StuCHS10* and *StuPAL* genes were regulated by multiple miRNAs (Fig. 5). As binding sites of transcription factors, cis-acting elements determine the expression pattern of genes to some extent [34]. A large number of cis-acting elements in response to environmental stresses were identified in the promoter sequences of *StuPAL* and *StuCHS* genes, indicating that their expression level was precisely regulated by environmental stresses (Fig. 6).

Expression profiles of *StuPAL* and *StuCHS* genes in response to environment stresses and different tissues

Gene expression pattern is an important aspect of gene function. High throughput sequencing provides a convenient way to study the gene expression patterns. As the first key genes in the phenylpropanoid pathway, *PAL* genes were precisely regulated in all respects, such as the pre- and post-transcriptional levels [35]. Different expression patterns in different organs were observed in many species. However, very few studies on the *PAL* genes in potatoes were reported. In this study, expression data were downloaded from PGSC. As shown in Fig. 7A, *StuPAL 1* was specific expressed in carpels. *StuPAL2*, *StuPAL3*, *StuPAL5* and *StuPAL8* highly expressed in all detected organs except in stamens. The rest members, *StuPAL4*, *StuPAL6* and *StuPAL7* were highly expressed in all detected organs. In previous study, *PAL* genes play key roles in response to drought stress after being treated for 24 and 36 h in *Salvia miltiorrhiza* [26], while there is no response to drought stress for all *StuPAL* gene members (Fig. 8A). This may be due to the different simulators (PEG-6000 and mannitol for *S. miltiorrhiza* and potatoes, respectively) of the treatments, the different treatment time and the different phase of plant growth and

development. Heat treatment prior to storage alleviated chilling injury and enhanced PAL activity in banana fruit [36]. *StuPAL6* and *StuPAL7* were strongly up-regulated after heat stress in this study. In *S. miltiorrhiza*, all *SmPALs* were down-regulated after 12 h MeJA treatment, while *SmPAL1* was up-regulated against after 48 h treatment [26]. However, all eight *StuPAL* genes except *StuPAL1* were down-regulated after BAP treatment and no response to the rest three hormones (IAA, ABA and GA₃) treatments in our study (Fig. 8A). For *StuCHS* genes, the expression patterns among different tissues were very diverse (Fig. 7B). *StuCHS5*, *StuCHS9* and *StuCHS11* (except in callus and stolons) showed constitutive expression in all detected tissues. While the rest *StuCHS* genes were expressed in specific tissues, suggesting functional diversification of *StuCHS* genes (Fig. 7B). Plant growth and development are strongly influenced by environmental factors, such as salt, drought, heat, diverse pathogens and chemical inducers during their life cycles. Many stress-related genes were induced or restrained to adapt the adverse environments [37, 38]. The *CHS* family has been reported to be regulated by hormone treatment [39] and light [40]. However, few studies about *CHS* genes responding to environmental stresses have been reported in potatoes. Thus, the expression changes of *StuCHS* genes in response to environmental stresses based on RNA-Seq data have been explored in this study. *StuCHS3* and *StuCHS12*, induced by salt, drought, ABA, GA₃ and BTH, were identified as key *CHS* genes in response to environmental stresses (Fig. 8B).

The contents of anthocyanin are varied in colored potato peel and pulp

As one of the most important food crops, potato is regarded as an important source of antioxidants for the human diet. Potato tubers contain a lot of polyphenols. Anthocyanins, the major count of the visible polyphenols, are very rich in different colored potatoes [41]. Anthocyanin not only plays important physiological roles in plant growth, development and stress response, but also have functions in the treatment of cardiovascular diseases as an antioxidant in human body [11]. Therefore, cultivating potato varieties with high anthocyanin content is an important measure to enhance people's health. In purple and red potatoes, the levels of anthocyanins are significantly higher than those in white and yellow tubers [42]. As shown in Fig. 9B, the levels of anthocyanin in peel of colored potato is significantly higher than that in pulp. Among them, purple potato peel had the highest anthocyanin level, followed by red potato peel, and yellow peel had the lowest anthocyanin level.

Expression analysis of *StuPAL* and *StuCHS* genes in colored potato tubers

Detecting the expression levels of *StuPAL* and *StuCHS* genes in potato tubers will help to further understand their functions in anthocyanin biosynthesis pathways. For *StuPAL* genes, the expression levels of all *StuPAL* genes were mainly expressed in the peel of the three potato varieties, but almost not expressed in the pulp. It was speculated that they were mainly involved in the accumulation of anthocyanin in peel, but not in pulp. The expression changes of *StuPAL2*, *StuPAL3*, *StuPAL4* and *StuPAL6* were consistent with the changes in anthocyanin content (Fig. 10A). It is speculated that these four genes were more likely to be involved in the anthocyanin synthesis. However, the *StuCHS* genes were quite different. *StuCHS* family members were expressed in both peel and pulp. The expression levels of *StuCHS3*, *StuCHS4*, *StuCHS5*, *StuCHS6*, *StuCHS9* and *StuCHS11* were higher in peel than pulp among three potato varieties, while *StuCHS1* is only highly expressed in purple pulp (Fig. 10B). These results suggested that *StuCHS* genes could be involved in both peel and pulp anthocyanin accumulation. This also showed that *StuPAL* and *StuCHS* genes had different functions in the regulation of anthocyanin synthesis in different colored potato varieties. In conclusion, these findings provided in this study lay an important basis to further explore the biological functions of the *StuPAL* and *StuCHS* gene family in colored potato.

Conclusion

Anthocyanins are the most abundant class of plant flavonoid pigments. In our study, 135 *PAL* and 479 *CHS* genes were comprehensively identified in the 18 plant species. Among them, Eight *StuPAL* and 12 *StuCHS* genes were analysed in potato for the first time. These genes were mainly distributed on chromosome Chr9 and Chr12, respectively. *StuPAL* genes mainly contained L-Aspartase-like feature domains and *StuCHS* genes members contained Chal_sti_synt_C or Chal_sti_synt_N characteristic structures. They are fairly conserved and their gene structures were related to their functions. *StuPAL* genes were mainly expressed in the peel of colored potato, *StuPAL2*, *StuPAL3*, *StuPAL4* and *StuPAL6* had the highest expression, consistent with the content of anthocyanins. *StuCHS* family genes were expressed in the peel and pulp. These results suggested that *StuPAL* genes could be involved in the synthesis of anthocyanins in the colored potato tuber peel, while *StuCHS* genes might performed functions both in peel and pulp. These researches provided insights into the characteristics of *PAL* and *CHS* genes and could facilitate to further explore the biological functions of the *StuPAL* and *StuCHS* gene family in colored potato.

Materials And Methods

Sequence Retrieval and Phylogenetic Trees Constructed of PAL and CHS Genes in 18 Plant Species

Protein sequences of *PAL* and *CHS* genes of *A. thaliana* were downloaded from the website TAIR (<https://www.arabidopsis.org/>). These protein sequences were used as query sequences and Blastp program was performed in Phytozome 12 (<https://phytozome.jgi.doe.gov/pz/portal.html#>), with E value $\leq 10^{-20}$, similarity $\geq 75\%$, and protein length covered greater than 60%. *PAL* and *CHS* genes in *Amygdalus persica*, *Brachypodium distachyum*, *Brassica oleracea*, *B. rapa*, *Citrus sinensis*, *Cucumis sativus*, *Glycine max*, *Gossypium*, *Lycopersicon esculentum*, *Malus pumila*, *Oryza sativa*, *Populus trichocarpa*, *Setaria italica*, *Solanum tuberosum*, *Sorghum bicolor*, *Vitis vinifera* and *Zea mays* species were downloaded. The sequences were then checked for the corresponding protein characteristic domains using SMART (<http://smart.embl-heidelberg.de/>) [43] and Pfam (<http://pfam.xfam.org/>) [44]. ClustalX (ver.1.83) was used to conduct multiple sequence alignment and phylogenetic trees was built by MEGA 7.0 software (www.megasoftware.net) [45] using the maximum likelihood method.

Molecular Characterization, Gene Structure, Chromosome Distribution and Gene Duplication of Anthocyanin Synthesis Pathway Genes in Potato

The molecular weight (Mw) and isoelectric points (pI) of *StuPALs* and *StuCHSs* were predicted the online ExPASy program (<https://www.expasy.org/>). The proteins location were predicted using PSORT Prediction (<http://psort1.hgc.jp/form.html>). The gene structures were visualized using the online GSDS v2.0 (<http://gsds.cbi.pku.edu.cn/>) [46] and motifs were explored using MEME (<http://meme-suite.org/>) [47], setting the maximum number of motifs at 10 and the optimum width of each motif, between 6 and 50 residues. Mapchart 2.2 software was used to map *PAL* and *CHS* genes on the potato chromosomes. The duplicated events of anthocyanin synthesis genes were confirmed by searching the Plant Genome Duplication Database (<http://chibba.agtec.uga.edu/duplication/index/locus>).

Promoter Sequence and miRNA/target Analyses

To analysis the cis-acting regulatory elements present in the promoter regions of the *StuPAL* and *StuCHS* genes, 1500 bp upstream regions from the transcription start site of these genes were retrieved from Spud DB (<http://solanaceae.plantbiology.msu.edu/index.shtml>) [48] and supplied to PlantCARE database (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>) [49]. To predict these genes-targeted miRNAs, genomic DNA sequences of these genes were searched against the published *Solanum tuberosum* miRNAs sequences using psRNATarget Server with default parameters (<http://plantgrn.noble.org/psRNATarget/>) [50]. The interaction networks of the predicted miRNAs and the corresponding target genes were visualized using Cytoscape software (<https://cytoscape.org/>).

Expression Profiling of StuPAL and StuCHS Genes

Expression analysis of anthocyanin synthesis pathway genes in various tissues, stress response and exogenous hormones response were performed using RNA-Seq data obtained in Potato Genome Sequencing Consortium (NCBI accession: SRA030516). The gene expression levels were transformed by $\log_2$ (FPKM+0.01) and visualized using HemI software (<http://hemi.biocuckoo.org/down.php>). In the dataset, tissues (including flower, leaves, stems and roots) in normal condition, stress response (including biotic stress, such as *Phytophthora infestans* (PIL) and elicitors acibenzolar-S-methyl (BTH), DL-b-amino-n-butyric acid (BABA) and abiotic stress, such as salinity (NaCl, 150 mmol L⁻¹), drought (mannitol, 260 μ mol L⁻¹) and heat (35°C)), and four hormone treatments such as abscisic acid (ABA, 50 μ mol L⁻¹), 6-benzylaminopurine (BAP, 10 μ mol L⁻¹), gibberellic acid (GA₃, 50 μ mol L⁻¹), and indole-3-acetic acid (IAA, 10 μ mol L⁻¹). For abiotic stresses, whole plants were detached at 24 h in vitro after exposed to abiotic stresses (heat, salt and drought). For rest stresses, samples were pooled at 24, 48 and 72 h after stresses.

Plant samples and Determination of anthocyanin content

Three potato cultivars were selected in this study based on the color of tuber peel and pulp: JYS, 1-11 and 16-A2 (Fig. 9A). These varieties were grown in fields located in Wulong, Chongqing, during the 2020/2021 season and tubers were harvested at their physiological maturity.

Anthocyanin content of tuber peel and pulp of three potato cultivars was measured using sodium phosphate, dibasic-citric acid buffer method [51]. All samples were ground to powder and about 2 g powder was placed into 40 mL of extraction buffer (Sodium phosphate, dibasic-Citric Acid buffer) and incubated at 60°C for two hours. After cooling, keep volume to 100 mL and let stand for 2 hours.

Absorbance values (A525) of the supernatant were measured using spectrophotometer after centrifugation for 5 min at 5,000 rpm. The anthocyanin content was calculated as:

$$C \text{ (mg/100g potato)} = 1/958 * V/100 * 1/m * ABS * 100000$$

The characters noted as follows: C: anthocyanin content, 958: the empirical coefficient of mass to volume ratio, V (mL): volume, m (g): Sample quality, ABS: absorbance value, 100000: unit conversion coefficient.

qRT-PCR analysis

Total RNA of skin and tuber flesh from three potato cultivars was extracted using the Pure Plant RNA Kit (Sangon, China) and convert it to cDNA by reverse transcription. In addition, primers for qRT-PCR of *StuPAL* and *StuCHS* genes were designed using Premier 5.0 (**Table S1**) and synthesized by Sangon Biotech (Shanghai, China). Three biological repetitions and three technical repetitions were performed for each sample. Real-time PCR reactions were performed in a total reaction volume of 20 µL using the following conditions: 95°C /30 s; 39 cycles (95°C /5 s, 60°C /30 s); 95°C /5 s; 60°C /1 min; dissolution curve 60-95°C, increment 0.5°C. Relative gene expression was analyzed using the $2^{-\Delta\Delta Ct}$ method.

Abbreviations

PAL
phenylalanine ammonia lyase
CHS
chalcone synthase
DFR
dihydroflavonol 4-reductase
ANS
anthocyanin synthase
LDOX
colorless anthocyanin dioxygenase
C4H
Cinnamic acid 4-hydroxylase
CHI
chalcone isomerase
GST
glutathione S-transferase
pI
isoelectric point
Mw
molecular weight.

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

The gene expression data were downloaded from the PGSC website with accession number SRA030516 in NCBI.

Competing interest

The authors declared no conflict of interest in the authorship and publication of this document.

Funding

This study was supported by the National Key Research and Development Program of China (2018YFE0127900), Science and Technology Partnership Program, Ministry of Science and Technology of China (KY201904016), National Natural Science Foundation of China (32101659), the Talent Introduction Program of Southwest University Project (SWU019008) and Special Project of Germplasm Creation of Southwest University.

Authors' contributions

D.L. and H.J. conceived and designed the experiments. L.S. performed the experiments. L.S., M.W., Y.F., R.L., Z.J., J.W. and B.H. analyzed the data. L.S. and H.J. wrote the paper. All authors reviewed the manuscript.

Acknowledgements

Not applicable

References

1. Kim HW, Kim SR, Lee YM, Jang HH, Kim JB. **Analysis of Variation in Anthocyanin Composition in Korean Coloured Potato Cultivars by LC-DAD-ESI-MS and PLS-DA**. *Potato Res.* 2018;61(1):1-17.
2. Bulgakov VP, Avramenko TV, Tsitsiashvili GS. **Critical analysis of protein signaling networks involved in the regulation of plant secondary metabolism: focus on anthocyanins**. *Crit Rev Biotechnol.* 2017;37(6):685-700.
3. Muhlemann JK, Younts TLB, Muday GK. **Flavonols control pollen tube growth and integrity by regulating ROS homeostasis during high-temperature stress**. *P Natl Acad Sci USA.* 2018;115(47):E11188-E97.
4. Bhat R, Stamminger R. **Impact of ultraviolet radiation treatments on the physicochemical properties, antioxidants, enzyme activity and microbial load in freshly prepared hand pressed strawberry juice**. *Food Sci Technol Int.* 2015;21(5):354-63.
5. Treptow TC, Comarella CG, Franco FW, Rodrigues E, Domingues F, Bochi VC, Sautter CK. **Thermal Pest Control in 'Tannat' grapes: Effect on anthocyanins, sensory and color of one-year-old wines**. *Food Res Int.* 2017;100:113-21.
6. Yang F, Li MQ, Xin YY, Gao DM, Wu FZ. **Effects of flavonoids from potato-onion on Fusarium wilt fungus of tomato**. *Allelopathy J.* 2019;47(1):119-26.
7. Katsumata S, Toshima H, Hasegawa M. **Xylosylated Detoxification of the Rice Flavonoid Phytoalexin Sakuranetin by the Rice Sheath Blight Fungus *Rhizoctonia solani***. *Molecules.* 2018;23(2).
8. Hwang JW, Natarajan SB, Kim YS, Kim EK, Lee JW, Moon SH, Jeon BT, Park PJ. **Biosynthesis of Oligomeric Anthocyanins from Grape Skin Extracts**. *Molecules.* 2017;22(3).
9. Igwe EO, Charlton KE, Roodenrys S, Kent K, Fanning K, Netzel ME. **Anthocyanin-rich plum juice reduces ambulatory blood pressure but not acute cognitive function in younger and older adults: a pilot crossover dose-timing study**. *Nutr Res.* 2017;47:28-43.
10. Bontempo P, De Masi L, Carafa V, Rigano D, Scisciola L, Iside C, Grassi R, Molinari AM, Aversano R, Nebbioso A *et al.* **Anticancer activities of anthocyanin extract from genotyped *Solanum tuberosum* L. "Vitelotte"**. *J Funct Foods.* 2015;19:584-93.
11. Zhou F, Wang T, Zhang BL, Zhao HF. **Addition of sucrose during the blueberry heating process is good or bad? Evaluating the changes of anthocyanins/anthocyanidins and the anticancer ability in HepG-2 cells**. *Food Res Int.* 2018;107:509-17.
12. Salvatierra A, Pimentel P, Moya-Leon MA, Caligari PD, Herrera R. **Comparison of transcriptional profiles of flavonoid genes and anthocyanin contents during fruit development of two botanical forms of *Fragaria chiloensis* ssp. *chiloensis***. *Phytochemistry.* 2010;71(16):1839-47.
13. Vogt T. **Phenylpropanoid biosynthesis**. *Mol Plant.* 2010;3(1):2-20.
14. Yamamura Y, Ogihara Y, Mizukami H. **Cinnamic acid 4-hydroxylase from *Lithospermum erythrorhizon*: cDNA cloning and gene expression**. *Plant Cell Rep.* 2001;20(7):655-62.
15. Chen HC, Song JN, Williams CM, Shuford CM, Liu J, Wang JP, Li QZ, Shi R, Gokce E, Ducoste J *et al.* **Monolignol Pathway 4-Coumaric Acid: Coenzyme A Ligases in *Populus trichocarpa*: Novel Specificity, Metabolic Regulation, and Simulation of Coenzyme A Ligation Fluxes**. *Plant Physiol.* 2013;161(3):1501-16.
16. Sun LL, Li Y, Li SS, Wu XJ, Hu BZ, Chang Y. **Identification and characterisation of DfCHS, a chalcone synthase gene regulated by temperature and ultraviolet in *Dryopteris fragrans***. *Cell Mol Biol.* 2014;60(6):1-7.

17. Wang HL, Wang W, Zhan JC, Yan AL, Sun L, Zhang GJ, Wang XY, Ren JC, Huang WD, Xu HY. **The accumulation and localization of chalcone synthase in grapevine (*Vitis vinifera* L.).** Plant Physiol Bioch. 2016;106:165-76.
18. Lim SH, You MK, Kim DH, Kim JK, Lee JY, Ha SH. **RNAi-mediated suppression of dihydroflavonol 4-reductase in tobacco allows fine-tuning of flower color and flux through the flavonoid biosynthetic pathway.** Plant Physiol Bioch. 2016;109:482-90.
19. Ben-Simhon Z, Judeinstein S, Trainin T, Harel-Beja R, Bar-Ya'akov I, Borochoy-Neori H, Holland D. **A "White" Anthocyanin-less Pomegranate (*Punica granatum* L.) Caused by an Insertion in the Coding Region of the Leucoanthocyanidin Dioxygenase (LDOX; ANS) Gene.** Plos One. 2015;10(11).
20. Tanaka Y, Brugliera F, Kalc G, Senior M, Dyson B, Nakamura N, Katsumoto Y, Chandler S. **Flower Color Modification by Engineering of the Flavonoid Biosynthetic Pathway: Practical Perspectives.** Biosci Biotech Bioch. 2010;74(9):1760-69.
21. Pietrini F, Iannelli MA, Massacci A. **Anthocyanin accumulation in the illuminated surface of maize leaves enhances protection from photo-inhibitory risks at low temperature, without further limitation to photosynthesis.** Plant Cell Environ. 2002;25(10):1251-59.
22. Plantenga FDM, Bergonzi S, Abelenda JA, Bachem CWB, Visser RGF, Heuvelink E, Marcelis LFM. **The tuberization signal StSP6A represses flower bud development in potato.** J Exp Bot. 2019;70(3):937-48.
23. Cho K, Cho KS, Sohn HB, Ha IJ, Hong SY, Lee H, Kim YM, Nam MH. **Network analysis of the metabolome and transcriptome reveals novel regulation of potato pigmentation.** J Exp Bot. 2016;67(5):1519-33.
24. Vision TJ, Brown DG, Tanksley SD. **The origins of genomic duplications in Arabidopsis.** Science. 2000;290(5499):2114-7.
25. Kong H, Landherr LL, Frohlich MW, Leebens-Mack J, Ma H, dePamphilis CW. **Patterns of gene duplication in the plant SKP1 gene family in angiosperms: evidence for multiple mechanisms of rapid gene birth.** Plant J. 2007;50(5):873-85.
26. Hou X, Shao F, Ma Y, Lu S. **The phenylalanine ammonia-lyase gene family in *Salvia miltiorrhiza*: genome-wide characterization, molecular cloning and expression analysis.** Mol Biol Rep. 2013;40(7):4301-10.
27. Chalhoub B, Denoeud F, Liu S, Parkin IA, Tang H, Wang X, Chiquet J, Belcram H, Tong C, Samans B *et al.* **Plant genetics. Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome.** Science. 2014;345(6199):950-3.
28. Han Y, Ding T, Su B, Jiang H. **Genome-Wide Identification, Characterization and Expression Analysis of the Chalcone Synthase Family in Maize.** Int J Mol Sci. 2016;17(2).
29. Koes RE, Spelt CE, van den Elzen PJ, Mol JN. **Cloning and molecular characterization of the chalcone synthase multigene family of *Petunia hybrida*.** Gene. 1989;81(2):245-57.
30. Durbin ML, Learn GH, Jr., Huttley GA, Clegg MT. **Evolution of the chalcone synthase gene family in the genus *Ipomoea*.** Proc Natl Acad Sci U S A. 1995;92(8):3338-42.
31. Wang WK, Schaal BA, Chiou YM, Murakami N, Ge XJ, Huang CC, Chiang TY. **Diverse selective modes among orthologs/paralogs of the chalcone synthase (Chs) gene family of *Arabidopsis thaliana* and its relative *A. halleri* ssp. *gemma*.** Mol Phylogenet Evol. 2007;44(2):503-20.
32. Hou Z, Cao J. **Comparative study of the P2X gene family in animals and plants.** Purinergic Signal. 2016;12(2):269-81.
33. Cao J, Li X, Lv Y. **Dynein light chain family genes in 15 plant species: Identification, evolution and expression profiles.** Plant Sci. 2017;254:70-81.
34. Liu Y, Wang L, Xing X, Sun LP, Pan JW, Kong XP, Zhang MY, Li DQ. **ZmLEA3, a Multifunctional Group 3 LEA Protein from Maize (*Zea mays* L.), is Involved in Biotic and Abiotic Stresses.** Plant Cell Physiol. 2013;54(6):944-59.
35. Yan F, Li H, Zhao P. **Genome-Wide Identification and Transcriptional Expression of the PAL Gene Family in Common Walnut (*Juglans Regia* L.).** Genes (Basel). 2019;10(1).
36. Chen JY, He LH, Jiang YM, Wang Y, Joyce DC, Ji ZL, Lu WJ. **Role of phenylalanine ammonia-lyase in heat pretreatment-induced chilling tolerance in banana fruit.** Physiol Plant. 2008;132(3):318-28.
37. Albrecht V, Weinl S, Blazevic D, D'Angelo C, Batistic O, Kolukisaoglu U, Bock R, Schulz B, Harter K, Kudla J. **The calcium sensor CBL1 integrates plant responses to abiotic stresses.** Plant J. 2003;36(4):457-70.
38. Kasuga M, Liu Q, Miura S, Yamaguchi-Shinozaki K, Shinozaki K. **Improving plant drought, salt, and freezing tolerance by gene transfer of a single stress-inducible transcription factor.** Nat Biotechnol. 1999;17(3):287-91.
39. Yu HN, Wang L, Sun B, Gao S, Cheng AX, Lou HX. **Functional characterization of a chalcone synthase from the liverwort *Plagiochasma appendiculatum*.** Plant Cell Rep. 2015;34(2):233-45.
40. Koduri PK, Gordon GS, Barker EI, Colpitts CC, Ashton NW, Suh DY. **Genome-wide analysis of the chalcone synthase superfamily genes of *Physcomitrella patens*.** Plant Mol Biol. 2010;72(3):247-63.

41. Tengkun N, Dongdong W, Xiaohui M, Yue C, Qin C. **Analysis of Key Genes Involved in Potato Anthocyanin Biosynthesis Based on Genomics and Transcriptomics Data.** *Front Plant Sci.* 2019;10:603.
42. Strygina KV, Kochetov AV, Khlestkina EK. **Genetic control of anthocyanin pigmentation of potato tissues.** *BMC Genet.* 2019;20(Suppl 1):27.
43. Letunic I, Bork P. **20 years of the SMART protein domain annotation resource.** *Nucleic Acids Res.* 2018;46(D1):D493-D96.
44. El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, Qureshi M, Richardson LJ, Salazar GA, Smart A *et al.* **The Pfam protein families database in 2019.** *Nucleic Acids Res.* 2019;47(D1):D427-D32.
45. Kumar S, Stecher G, Tamura K. **MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets.** *Mol Biol Evol.* 2016;33(7):1870-4.
46. Hu B, Jin J, Guo AY, Zhang H, Luo J, Gao G. **GSDS 2.0: an upgraded gene feature visualization server.** *Bioinformatics.* 2015;31(8):1296-7.
47. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS. **MEME SUITE: tools for motif discovery and searching.** *Nucleic Acids Res.* 2009;37(Web Server issue):W202-8.
48. Potato Genome Sequencing C, Xu X, Pan S, Cheng S, Zhang B, Mu D, Ni P, Zhang G, Yang S, Li R *et al.* **Genome sequence and analysis of the tuber crop potato.** *Nature.* 2011;475(7355):189-95.
49. Lescot M, Dehais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, Rouze P, Rombauts S. **PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences.** *Nucleic Acids Res.* 2002;30(1):325-7.
50. Dai X, Zhuang Z, Zhao PX. **psRNATarget: a plant small RNA target analysis server (2017 release).** *Nucleic Acids Res.* 2018;46(W1):W49-W54.
51. Yang GC CX, Wang JC, Du C, Zhang K, Cai YM, Gao X, Tang DB, Zhao Y. **Anthocyanin extraction from purple sweet potato by citric acid-disodium hydrogen phosphate buffer.** *Journal of southern agriculture.* 2013;44(4):653-56.

Figures

Figure 1

Phylogenetic relationships among the identified PAL and CHS proteins in 18 plant species. The PAL and CHS protein sequences were aligned by ClustalX, and the phylogenetic tree was constructed using MEGA7 by the Maximum Likelihood method analysis (1000 replicates). Genes of 18 plant species are indicated at the end of the branches. Subgroups were named according to *Arabidopsis thaliana*. A: PAL family; B: CHS family.

Figure 2

The distribution of *StuPALs* and *StuCHSs* in *Solanum tuberosum* L. chromosomes. The chromosome name is at the top of each bar. The scale of the chromosome is in millions of bases (Mb).

Figure 3

Exon-intron structures of *StuPAL* and *StuCHS* genes. The phylogenetic tree was constructed using MEGA 7 software. Yellow rectangles represent CDS, blue rectangles represent the upstream and downstream, black thin lines represent intron. A: *StuPAL* genes; B: *StuCHS* genes.

Figure 4

Phylogenetic relationships and architecture of the conserved protein motifs in *StuPAL* and *StuCHS* genes. The motifs, numbered 1–10, are displayed in different colored boxes. The length of the protein can be estimated using the scale at the bottom. A: *StuPAL* genes; B: *StuCHS* genes.

Figure 5

miRNA–gene interactions network. Green nodes represent miRNAs and yellow nodes represent genes. A: *StuPAL* genes; B: *StuCHS* genes.

Figure 6

The promoter cis-elements analysis of *StuDALs* and *StUCHSs*. The 1500 bp DNA fragments upstream of the ATG starting code of genes were analyzed using online analysis software PlantCARE (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>). A: *StuDAL* genes; B: *StUCHS* genes.

Figure 7

Expression levels of *StuPALs* and *StuCHSs* in different tissues. A heat map created with clusters based on the FPKM value. The boxes display different colors represent different expression levels. Red color indicates a high expression of the relevant gene. A: *StuPAL* genes; B: *StuCHS* genes.

(A)

(B)

Figure 8

The expression patterns of *StuPALs* and *StuCHSs* under stresses. A heat map created with clusters based on the FPKM value. The boxes display different colors represent different expression levels. Red color indicates a high expression of the relevant gene. A: *StuPAL* genes; B: *StuCHS* genes.

Figure 9

Morphological characteristics and anthocyanin content determination of three potato varieties. A: For each variety, the complete tuber shape and peel color are shown on the left, and the tuber profile and pulp color are shown on the right. B: Tuber determination of anthocyanin content in peel and pulp. The results are presented as the mean $\pm$ SD of three replicates. Different letters represent significant differences at $P < 0.05$ between different samples. Error bars were obtained from three measurements. JYS-pe: peel of JYS variety; JYS-pu: pulp of JYS variety; 1-11-pe: peel of 1-11 variety; 1-11-pu: pulp of 1-11 variety; 16-A2-pe: peel of 16-A2 variety; 16-A2-pu: pulp of 16-A2 variety.

Figure 10

The relative expression levels of *StuPALs* and *StuCHSs* in the peel and pulp among three colored potato varieties. A: *StuPAL* genes; B: *StuCHS* genes. JYS-pe: peel of JYS variety; JYS-pu: pulp of JYS variety; 1-11-pe: peel of 1-11 variety; 1-11-pu: pulp of 1-11 variety; 16-A2-pe: peel of 16-A2 variety; 16-A2-pu: pulp of 16-A2 variety.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [TableS1RealTimeqPCRprimersofsequences.xlsx](#)
- [TableS2IdentificationofPALandCHSgenesamong18plantspecies.xlsx](#)