

Characterization of the *GGP* gene family in potato (*Solanum tuberosum* L.) and Capsicum (*Capsicum annuum* L.) and its expression analysis under hormonal and abiotic stresses

Zhiqi Ding

College of Horticulture, Gansu Agricultural University

Kangding Yao

College of Horticulture, Gansu Agricultural University

Yandong Yao

College of Horticulture, Gansu Agricultural University

Xuejuan Pan

College of Horticulture, Gansu Agricultural University

Lizheng Luo

College of Horticulture, Gansu Agricultural University

Long Li

College of Horticulture, Gansu Agricultural University

Chunlei Wang

College of Horticulture, Gansu Agricultural University

Weibiao Liao (✉ liaowb@gsau.edu.cn)

College of Horticulture, Gansu Agricultural University

Article

Keywords:

Posted Date: February 1st, 2024

DOI: <https://doi.org/10.21203/rs.3.rs-3897070/v1>

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

Additional Declarations: No competing interests reported.

Abstract

Potato (*Solanum tuberosum* L.) and capsicum (*Capsicum annuum* L.) are members of the Solanaceae family, and they are widely grown and used worldwide. However, there are significant differences between these two crops in terms of morphology, growing conditions and resistance to stress. But they share some similarities at a genetic level. GDP-L-galactose phosphorylase (GGP) is a key rate-limiting enzyme in plant ascorbic acid synthesis, which plays an important role in plant growth and development as well as stress response. However, the presence of *GGP* and its function in potato and capsicum are not known. In this study, we first identified two GGP genes in the potato genome using a genome-wide search approach. Similarly, two GGP genes were identified in the capsicum genome. We then analyzed their physicochemical properties, conserved domains, protein structures and phylogenetic relationships. Phylogenetic tree analysis revealed that members of the potato and capsicum *GGP* gene families are related to cabbage (*Brassica campestris*), Arabidopsis (*Arabidopsis thaliana* L.), apple (*Malus domestica* cv. Gala) and tomato (*Solanum lycopersicum* L.), with tomato being the most closely related. The promoter sequences mainly contain homeopathic elements such as light-responsive, hormone-responsive and stress-responsive, with light-responsive elements being the most abundant. By analyzing the structure of the genes, it was found that there is no transmembrane structure or signal peptide in the *GGP* gene family of potatoes and capsicums, and that all of its members are hydrophilic proteins. The expression profiles of the different tissues showed that the potato *GGP* was more highly expressed in the leaves and the capsicum *GGP* was more highly expressed in the fruits. qRT-PCR analysis showed that the *GGP* gene was responsive to both ABA and MeJA in both the potato and the capsicum. In addition, the *GGP* genes in both the potato and the capsicum showed different responses to darkness stress, cold stress, salt stress and drought stress. Based on our experimental findings, it is evident that *StGGPs* primarily respond to dark stress, while *CaGGPs* primarily respond to NaCl stress. This study provides a basis for a better understanding of the role of *GGP*-homologous genes from potato and capsicum in the response of plants to abiotic stress, and lays the foundation for further research into the role of *GGPs* in plants.

Introduction

Ascorbic acid (AsA) known as vitamin C is the most abundant water-soluble antioxidant in plant cells. It plays an important role in photosynthesis and stress resistance, and is involved in epigenetic activation¹. Ascorbate mitigates the harmful effects of ROS produced by normal or stressed cellular metabolism, either directly as a ROS scavenger or indirectly as a substrate for the enzyme ascorbate peroxidase². In addition, AsA helps plants to resist adverse environmental conditions, and it has been shown in wheat that ascorbic acid could significantly alleviate salt stress and drought stress³. The exogenous application of AsA significantly increased the activities of POD, CAT and SOD and reduced the Pb-induced oxidative damage in okra under Pb stress. Exogenous AsA played a similar role in alleviating stress damage in rice under aluminum stress⁴. AsA also increases the activity of antioxidant enzymes, photosynthesis capacity and the maintenance of salt stress tolerance in wheat, tomatoes, strawberries and cowpeas, thus counteracting the adverse effects of salt stress on plant growth⁵⁻⁸. There are four pathways used to synthesize ascorbic acid in plants: the L-galactose, L-glutathione, inositol and D-galacturonic acid pathways⁹. The L-galactose pathway is the main pathway for ascorbate biosynthesis in Arabidopsis (*Arabidopsis thaliana* L.) and tomato

(*Solanum lycopersicum* L.)¹⁰. GDP-L-galactose phosphorylase (GGP), also known as GDP-L-galactosyl transferase (GGT), catalyzes the conversion of GDP-L-galactose to L-galactose-1-phosphate and is considered a key control node in the AsA biosynthesis pathway¹¹. In addition, GGPase has been shown to be the rate-limiting step in the L-galactose pathway¹². GGP is the gene for GDP-L-galactose phosphorylase, which catalyzes the conversion of GDP-L-galactose to L-galactose-1-phosphate, and is the last gene cloned in the D-mannose/L-galactose pathway, the main pathway of ascorbate synthesis. The GGP gene plays an important role in the synthesis and regulation of ascorbic acid. It has been studied in crops such as wheat, maize and tomato¹³. It has been found that there are six TaGGP genes in wheat and that TaGGP2 homologues are specifically expressed in the anthers. Meanwhile, transient transformation of the TaGGP coding sequence in leaves resulted in a 5.3-fold increase in ascorbic acid concentration, confirming its activity in plants⁹. The banana genome database identified five GGPs¹⁴. Both rice and maize have a single copy of the GGP gene in their genomes, which is located on the Os11 and Zm6 chromosomes, respectively⁹. There was a severe 80% reduction in leaf ascorbic acid concentration when the OsGGP gene in rice was completely knocked out¹⁵. The expression of the GGP gene showed a good correlation with the corresponding ascorbic acid content in leaves and fruits of kiwifruit (*Actinidia chinensis*) at different development stages¹⁶. GGP proteins not only help to anabolize ascorbate but may also act as growth regulators. In *Arabidopsis vtc2* mutants, VTC2: GUS fusion proteins were found in both cytoplasmic and nuclear locations, indicating that VTC2 might play a catalytic role in cytoplasm and a specific role in nuclear gene expression regulation¹⁷. Potato (*Solanum tuberosum* L.) belongs to the Solanaceae family and is an important tuber crop. It is also the third most important food crop. It is grown and consumed throughout the world¹⁸. Potatoes are grown worldwide due to their adaptable growing habit and ability to grow in varying climatic and soil conditions. Capsicum (*Capsicum annuum* L.) belongs to the nightshade family (Solanaceae) and is a common vegetable crop. It has an important commercial value and is popular with consumers for its variety of flavors and nutritional value¹⁹. More adaptable to warmer climates, capsicum has a greater need for heat and sunlight. Although the potato and capsicum are both members of the nightshade family, they have very different morphology, growing conditions and use. As they grow in different environments and may be exposed to different biotic and abiotic stresses, their tolerance and stress resistance may also differ. At the genomic level, however, both potatoes and capsicum possess the GGP gene family. This is a key gene for the synthesis of ascorbic acid, which plays a crucial role in the growth, development and resistance of plants. Therefore, studying the GGP gene families of potatoes and capsicums may not only help us understand how these two crops adapt to their respective environments by regulating ascorbic acid synthesis, but also provide new strategies for crop variety improvement and increasing the nutritional value of crops. Many studies on potato and capsicum gene families have been reported with the availability of the Solanaceae genome database. However, knowledge of the GGP family in potatoes and capsicum is very limited. Therefore, to better understand the key role of the GGP family in plants, the coding genes of potato and capsicum GGP family members were characterized and analyzed in this study. Members of the potato and capsicum GGPs were analyzed for gene structure, secondary structure, chromosomal localization, conserved motif analysis, cis-acting element analysis, phylogenetic tree and subcellular localization. In addition, the expression patterns of these GGPs in different tissue-specific processes were investigated, as well as gene transcription analysis under different abiotic stresses and hormonal conditions.

This study was designed to lay the groundwork for future research into the role of GGPs in growing, developing and resisting in potato and capsicum.

Results

Identification of members of the *GGP* gene family in potato and capsicum

A total of four (two for potato and two for capsicum) protein sequences containing the *GGP* structural domain were retrieved from the potato and capsicum genome databases (E-value=1e-5). Screening removed incomplete proteins (not containing start or stop codons) and proteins whose annotation information in Blastp (database SSWISS-PROT/Uniprot) in NCBI are not a *GGP* structural domain, resulting in the identification of four (two each for potato and capsicum) *GGP* proteins (Table 1). Based on their chromosomal distribution, potatoes were sequentially named *StGGP1* and *StGGP2*, and capsicum, *CaGGP1* and *CaGGP2*. The distribution of GGP genes on chromosomes was visualized using the genome annotation files of potato and capsicum. The potato *GGP* gene is unevenly distributed on 2 chromosomes and the capsicum *GGP* gene is on the same chromosome, as shown in Figure 1. The number of genes on each chromosome is independent of the size of the chromosome.

Table 1 Identification of Potatoes and capsicum *GGP* gene family

Gene	Gene ID	Genome matching section/bp	The positive and negative chain
<i>StGGP1</i>	102596542	254672–257608	-
<i>StGGP2</i>	102593176	347521–356656	+
<i>CaGGP1</i>	107860116	142822364–142828691	-
<i>CaGGP2</i>	107860116	142822746–142828631	-

Analysis of conserved structural domains and gene structure of *GGP* family members in potato and capsicum

Figure 2 shows the conserved motifs of the *GGP* gene family and visualization of evolutionary relationships, conserved structures and gene structures of potato and capsicum *GGP* genes. The conserved structural domains showed that both potato and capsicum *GGP* families contained 10 conserved motifs, and all members contained identical conserved motifs (Motif), i.e., Motif 1-Motif 10, and the specific sequence information is shown in Table 2. The structural analysis of the genes showed that the *GGP* genes of the capsicum all contain five exons, in addition to *CaGGP1*, which has four introns, and *CaGGP2*, which has five introns.

However, the gene structures of the potato *GGPs* are very different. *StGGP1* has seven exons and six introns, while *StGGP2* has five exons, four introns and a longer intron. In addition, the values of the selective

evolutionary pressure showed that both the *GGP* genes of the potato and the capsicum were calculated to be less than 1 and that both had a tendency to be purified (Figure 2C).

Table2 Potato and capsicum *GGP* motif information

Species	Motif	Motif Sequence	Length (AA)
Potato	Motif1	WEERQQKGLFRYDVTACETKVIPGECGFIAQLNEGRHLKKRPTEFCIDKV	50
	Motif2	CLPQRIDRDSFAJALHFAREVADPFFRLGYNSLGAFATINHLHFQAYF	48
	Motif3	PQCYAEKQALGEVDQELLDTQVNPAVWEISGHIVLKRKEDYEGASEENAW	50
	Motif4	AGCGRKCLGKCCLPGSKLPLYGFKN	25
	Motif5	NFTKVGQEELLFRFEPSEEDPQLFPGARIDGEKSPSIVAINVSPIEYGH	50
	Motif6	MLKIKRVPTLVSNFQED	17
	Motif7	SCICLQEKNIPFNILADCGKKIF	24
	Motif8	NYPVRGFFVFEGGN	13
	Motif9	FPIEKAPIQKI	11
	Motif10	ASHMPP	6
Capsicum	Motif1	DCLPQRIGRDTFTIALHFAREMADPFFRVGYNSLGAFATINHLHYQAYY	49
	Motif2	NDDNEPMEHNIHTLPEEEYQISFLNNLLLGLWEERMSQGLFRYDVTTCT	50
	Motif3	PTEFRIDKVLQPFDENKFNFNTKVGQDEVLFREFPSTDYKRRYFSGMGVDA	50
	Motif4	GFTFEGENGSTIRDLSEVVVNISCISLQKNIPFNILIAQCGKKIFLFPQC	50
	Motif5	MLTIKRVPTVVSNYQEDVLLESNVVGCGRKCLGKCCLPVSM	41
	Motif6	IPGKCGFIAQLNEGRHLKKR	20
	Motif7	KAPVRKILARKGLGGAGVIVSKLLNYPVR	29
	Motif8	PSIVAINVSPIEYGHVLLIPR	21
	Motif9	PLYSFK	6
	Motif10	SVPFPVE	7

Multiple series comparison and phylogenetic analysis of the *GGP* family of potato and capsicum

The protein sequences of the potato and capsicum *GGP* families have been analyzed. The most conserved sequences are annotated in dark blue, followed by red and blue. As shown in Figure 3A, the homology

between members of the potato and capsicum *GGP* families was as high as 79.62%, indicating that the family is evolutionarily conserved.

To study the evolutionary relationship of potato and capsicum *GGP*s with *GGP*s of the template crop *Arabidopsis thaliana* and other species, a phylogenetic tree of *GGP* protein sequences in *Arabidopsis*, tomato, cabbage, apple, potato, and capsicum was constructed. The results showed that the 11 *GGP* proteins from these plants were categorized into three subgroups (I, II, and III), with *Arabidopsis* AT4G26850.1 and *Brassica napus* BrGGP2 being the most closely related in one subgroup (four members);

The type II subgroup (3 members) contains *GGP* members from apple, potato and tomato. The closest relatives are *StGGP1* from potato and *SlGGP2* from tomato; Three subgroups (four members) are present, with the most closely related being potato *StGGP2* and tomato *SlGGP1*, followed by capsicum *CaGGP1* and capsicum *CaGGP2*. In addition, the closest relative to potato of the three subgroups is tomato, so we speculate that potato and tomato have some evolutionary relationship. Meanwhile, potato and capsicum *GGP* family members are closely related to cabbage, *Arabidopsis* and tomato (Figure 3B).

Analysis of the role of cis-elements of potato and capsicum *GGP* genes

In order to understand the expression regulation of potato and capsicum *GGP* genes, 2000 bp upstream of the potato and capsicum *GGP* genomes were extracted and submitted to PlantCARE for promoter analysis. The potato predictions revealed seven light-responsive elements (Box 4, Box II, GA motif, GATA motif, G-box, GT1 motif and TCT motif); three hormone-responsive elements, including abscisic acid (ABRE), salicylic acid (TCA element) and (MeJA); and one stress element (LTR). The light-responsive elements were unevenly distributed in each of the *GGP* genes in potato, with *StGGP2* having the highest number of elements. The *StGGP1* contains 2 ABREs, the *StGGP2* contains 1 ABRE and the TCA element is distributed in the *StGGP1*. The MeJA response element was distributed in each potato *GGP* and the low temperature response element LTR was distributed in *StGGP1* (Figures 4A, B).

Capsicum predictions revealed the presence of three light responsive elements (Box 4, GT1-motif, TCT-motif), four hormone responsive elements containing methyl jasmonate (CGTCA-motif, TGACG-motif), gibberellin (P-box) and growth hormone (TGA-element) and three stress responsive elements (TC-rich repeats, GC-motif, ARE). The TCT motif, the MeJA, the P box and the TGA element were present in every *GGP* of Capsicum. The stress response element TC-rich repeats were distributed in *CaGGP1* and *CaGGP2* of Capsicum, and the anaerobically induced response element GC-motif, ARE, was distributed in *CaGGP1* (Figures 4C, D).

Structural analysis of proteins from the *GGP* gene family

The analysis of the physicochemical properties of the potato and capsicum *GGP* family revealed that the length of the potato and capsicum *GGP* proteins ranges from 371 aa to 438 aa, and the theoretical isoelectric point pI, except *CaGGP2* which is greater than 7, are distributed between 4.71 and 5.79, which belonged to the acidic proteins. The molecular mass of the *GGP* proteins from the potato and the capsicum ranges from 41.5 kDa to 49.7 kDa. In addition, neither the potato nor the capsicum *GGP* family has a signal peptide or transmembrane structure, according to both signal peptide and transmembrane structure predictions.

Meanwhile, it was predicted that both potato and capsicum *GGP* proteins may be hydrophilic. They contain more hydrophilic peaks (Table 3).

In addition, the potato and capsicum *GGP* genes were found to be distributed mainly in chloroplasts, cytoplasm, nucleus and mitochondria by subcellular localization. However, the distribution of each member was different in different cellular fractions (Table 4).

The secondary structure of the potato and capsicum family *GGP*s revealed that each member of the potato and capsicum family *GGP*s has the highest percentage of irregular curl (pink line) between 44.29 and 49.87 %; The percentage of alpha helices (blue line) is between 22.64 % and 33.33 %, the percentage of elongations (red line) is between 16.67 % and 22.91 %, and the percentage of beta folds (green line) is the smallest, between 4.58 % and 5.71 % (Table 5). Based on the tertiary structure results, the structures of potato and capsicum *GGP* family proteins are consistent with the structural components of the secondary predictions, and the members are modeled very similarly (Figure 5).

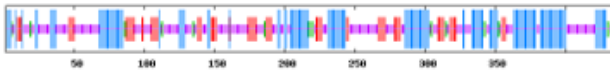
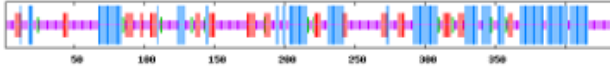
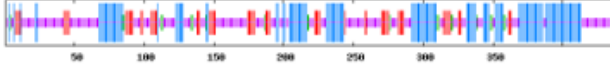
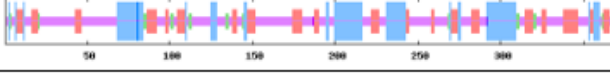
Table 3 Analysis of physicochemical properties of potato and capsicum *GGP*

Gene	Length(aa)	Molecular Weight	pI	Instability index	Grand average of hydropathicity	Signal peptide	Transmembrane domain
StGGP1	438	48703.3	4.71	40.75	-0.211	NO	NO
StGGP2	438	49102.28	5.24	53.53	-0.18	NO	NO
CaGGP1	440	49674.12	5.79	47.78	-0.192	NO	NO
CaGGP2	371	41516.94	8.25	45.74	-0.088	NO	NO

Table 4 Subcellular localization of *GGP* in potato and capsicum

Gene	chloroplast	cytoplasm	nuclear	mitochondria	plasmalemma	extracellular matrix
<i>StGGP1</i>	5	5	2	1.5	–	–
<i>StGGP2</i>	6.5	4	1	–	1	–
<i>CaGGP1</i>	3.8	4	2	–	–	1
<i>CaGGP2</i>	5	5	–	–	1	1

Table 5 Distribution of secondary structural elements in potato and capsicum *GGP*

Gene	Alpha helix (%)	Beta turn (%)	Random coil (%)	Extended strand (%)	Distribution of secondary structure elements
<i>StGGP1</i>	33.33%	5.71%	44.29%	16.67%	
<i>StGGP2</i>	32.42%	4.79%	46.12%	16.67%	
<i>CaGGP1</i>	32.05%	4.77%	46.59%	16.59%	
<i>CaGGP2</i>	22.64%	4.58%	49.87%	22.91%	

Expression pattern analysis of potato and capsicum *GGP* gene family in different tissues

In order to confirm the tissue-specific expression patterns of *StGGPs* and *CaGGPs*, we analyzed the expression levels of the *GGP*s genes at different stages of growth and in different tissues in potato and capsicum. The results showed that the *StGGPs* family has the highest expression in leaves. In addition, the expression of *StGGP1* is higher than that of *StGGP2* in all tissues except stamens. The capsicum *GGP* gene is expressed in roots, stems, leaves, flowers and fruits. Leaves showed the lowest expression, fruits showed the highest expression and red fruits showed higher expression than green ripe fruits. The expression of the capsicum *GGP* gene family in different tissues and during fruit expansion was extracted from previously published data. These results indicate that the *GGP*s genes are mainly active in the leaves of the potato, whereas in the capsicum, the *GGP*s are mainly active in the fruits, especially in the pre-ripening of fruits (Figure 6).

Expression of *GGP* gene in potato and capsicum under MeJA and ABA treatments

In order to understand the expression patterns of the *StGGPs* and *CaGGPs* genes under different hormonal treatments and abiotic stresses, we treated potato and capsicum seedlings with two hormones (ABA and MeJA) as well as with four stress factors (dark, cold, NaCl and PEG). As shown in Figure 7, the expression of *StGGP1*, *StGGP2*, *CaGGP1* and *CaGGP2* was differentially upregulated after 6 h stimulation with ABA and MeJA, and their expression levels all peaked at 24 h under ABA treatment. When treated with MeJA, *StGGP1* and *StGGP2* peaked at 12 h, whereas *CaGGP1* and *CaGGP2* peaked at 24 h, and the expression levels of the genes gradually declined after having peaked.

Expression profiling of the *GGP* gene in potato and capsicum under darkness, low temperature NaCl, and PEG treatments

Under abiotic stress treatments, both *StGGPs* and *CaGGPs* genes were differentially up-regulated after a certain period of stress. Under the duress of darkness, the level of expression of *StGGP1* reached a maximum at 6 h and increased by a factor of 14 compared to the level at 0 h. However, the expression levels of *StGGP2*, *CaGG1* and *CaGG2* were gradually upregulated after dark stress treatment, and all of them reached

the maximum at 24 h. Overall, the relative expression levels of the *StGGPs* were higher than those of the *CaGGPs* under the dark stress treatment (Figure 8A). The expression levels of *StGGPs* and *CaGGPs* were gradually upregulated with time after cold stress treatment, with the difference that the expression levels of *StGGPs* reached the highest after 12 h, while those of *CaGGPs* reached the highest after 24 h (Figure 8B). Under NaCl and PEG stress treatments, the expression levels of both *StGGPs* and *CaGGPs* were differentially up-regulated. Under both stress treatments, the expression levels of *StGGPs* peaked at 24 h, whereas *CaGGPs* peaked at 12 h, and the expression levels at 48 h were similar to those at 0 h (Figure 8C and D).

Discussion

AsA is found in a wide range of plant tissues and is a multifunctional metabolite involved in many physiological processes such as photosynthesis, cell wall biosynthesis, seed germination, flowering, fruit softening and ripening, post-harvest storage, and plant resistance to adverse environments²⁰. The GDP-L-galactose phosphorylase (GGP) gene encodes the rate-limiting enzyme of the L-galactose pathway, the major ascorbate biosynthesis pathway in plants, and is a promising candidate for increasing crop ascorbate⁹. The *GGP* gene has been identified in several species¹³. In this study, a more comprehensive and systematic analysis of the potato and capsicum *GGP* gene families has been undertaken with the aid of bioinformatic techniques. Four GGP genes (two in potato and two in capsicum) have been identified. Meanwhile, the selective evolutionary pressure indicated that there is a trend towards the purification of *GGP* members in both potato and capsicum, and the trend towards the purification of *GGP* in potato is higher than that in capsicum. Using a multi-species phylogenetic tree, we found that the *GGPs* of potato and capsicum are more conserved in evolution than those of *Arabidopsis*, cabbage and tomato. In addition, most plant *GGP* genes have similar gene structures and motif patterns. This suggests that plant *GGP* genes have conserved functions²¹. The analysis of the physicochemical properties of the potato and capsicum *GGP* gene family members showed that the protein sequences of the potato and capsicum *GGP* families differ very little, with amino acid lengths ranging from 371aa to 440aa. With the exception of *CaGGP2*, which has an isoelectric point greater than 7, the isoelectric point of the rest proteins is less than 7. These results suggest that *GGP* proteins have a higher content of acidic amino acids. In addition, the instability indices of the potato and capsicum *GGP* families are greater than 40. This indicates that potato and capsicum *GGPs* are unstable proteins. Meanwhile, transmembrane structure analysis and signal peptide prediction indicated that potato and capsicum *GGP* families do not have transmembrane structures as well as signal peptides. These results show the similarity between the *StGGPs* and the *CaGGPs* in terms of their physico-chemical properties as well as their gene structure. Previous study had also shown that most plant *GGP* genes had similar gene structures and motif patterns²². These similarities suggest that the *GGP* gene may play the same role in growing and developing potato and capsicum.

GGPs have different patterns of expression in different plant tissues. In wheat, the expression of *TaGGP2* in the anthers is about 50 to 580 times higher than in that in the other tissues⁹. In potato we found that the *GGP* family is highly expressed in leaves. In addition, the expression levels of *StGGP1* are higher than those of *StGGP2* in all tissues except stamens. This leads us to speculate that *StGGP1* is the major gene for AsA synthesis in potato. However, in contrast to this, the *GGP* family of capsicums has the lowest expression in

the leaves and the highest expression in the fruits. Meanwhile, the expression of *CaGGPs* genes gradually decreases as the capsicum s ripened. Similarly, previous study showed that most of the 12 banana varieties contained more AsA in the unripe stage than in the ripe stage¹⁴.

Study has shown that AsA is involved in photosynthesis as an electron carrier or electron donor, and that the pathway for the biosynthesis of L-galactose is closely linked to photosynthesis²³. AsA synthesis and *GGP* expression in *Arabidopsis thaliana* were induced by high light stress²⁴. Similarly, analysis of the roles of the homeopathic components in our study showed that the potato *GGP* family responds mainly to light-responsive components (seven) and hormone-responsive components (three), in addition to stress-responsive components. There are also three light-responsive elements in the *GGP* family of capsicum. These results suggest that the *GGP* family plays a major role in light response. In subsequent experiments we found a significant up-regulation of both *StGGPs* as well as *CaGGPs* after dark treatment. This is further evidence that the *GGP* gene family plays an important role in the response to light. In addition, we found that *StGGP1* contains response elements that regulate the circadian rhythm. Meanwhile, previous study has shown that AsA content in leaves is 2-fold higher during the day than at night, which may be related to circadian-regulated response elements²⁵. Therefore, the potato and capsicum *GGP* families may have an effect on AsA levels through participation in photosynthesis.

Previous study has shown that the expression of the *GGP* gene in kiwifruit was induced by light and abiotic stress, and was significantly correlated with the concentration of AsA²⁶. MeJA induces chemical defense in plants by stimulating the expression of plant defense genes. MeJA treatment increased the activity of the promoter, induced the expression of *GGP* in kiwifruit, and enhanced the AsA level²⁶. ABA-responsive elements (ABRE) are also important for the enhancement of *GGP* transcript levels and GUS activities of *GGP* promoters. In both potatoes and capsicum, MeJA-responsive elements were found. However, ABA response elements were found only in potato. Interestingly, in our experiments, after treatment with ABA and MeJA, the expression levels of both *StGGPs* and *CaGGPs* were significantly increased. This may be due to the fact that the ABA treatment regulates the expression of other genes in the capsicum and thus promotes the expression of the *CaGGPs*. *GGP* plays an important role in protecting plants against chilling stress by maintaining the pool and redox state of ascorbate²⁷. In addition, our study showed that *StGGPs* and *CaGGPs* may also have different responses to cold stress, salt stress and drought stress. Previous studies in rice, leek and lettuce have also shown that the *GGP* gene plays an important role under high light and abiotic stress conditions²⁸⁻³⁰. These results suggest that although potato and capsicum are very different in both morphological and environmental characteristics, their *GGP* genes may play the same role under abiotic stress. However, more studies are needed to understand the specific functions of the *GGP* gene family in plant growth and response to abiotic stress.

Conclusion

In conclusion, we identified four *GGP* genes in the whole genomes of potato and capsicum, of which potato and capsicum each contains two, and all of their members show a trend of purified selection. Meanwhile, there are strong homology and conservation of the *GGP* family in potato, capsicum, tomato, Arabidopsis, apple and cabbage. The potato and capsicum *GGP* genes have a genetic basis for responding to a wide

range of hormones and stresses. *StGGPs* and *CaGGPs* are expressed in a variety of tissues, but at different levels, with *StGGPs* being more highly expressed in leaves and *CaGGPs* being more highly expressed in fruits. The expression patterns of *GGP* genes indicate tissue-specific roles across different plant species. In addition, *StGGPs* and *CaGGPs* may be involved in the mitigation of abiotic stress and hormonal responses. Meanwhile, *StGGPs* responded mainly to dark stress, whereas *CaGGPs* responded mainly to NaCl stress. Therefore, this study may provide further research on the involvement of the *GGPs* gene family in growth regulation and stress response in potato and capsicum. It may also provide a theoretical basis for further exploration of the functions of plant *GGP* members.

Materials and methods

Identification of members of the *GGP* gene family in potato and capsicum

The whole genome data and annotation files for potato and capsicum were downloaded from NCBI (<https://www.NCBI.nlm.nih.gov/>). Membership information and protein sequences of the Arabidopsis *GGP* gene family were downloaded from TAIR (arabidopsis.org), and the Arabidopsis database was used to search for identified members of the *GGP* gene family as template genes. The Blast Compare Two Seqs function of the TBtools software was used for a preliminary comparison to obtain the transcriptome IDs of the candidate members of the potato and capsicum *GGP* gene families. The structural domains of the potato and capsicum *GGP* nominees were viewed using the NCBI Web CD search tool, and the TBtools visualization of NCBI CDD patterns. Candidate proteins with *GGP* structural domains were members of the potato and chili capsicum *GGP* gene families, and the genes corresponding to these proteins were named *StGGP* (potato) and *CaGGP* (capsicum) genes.

Chromosomal localization of the *GGP* gene family in potato and capsicum

The chromosomal location distribution of the *GGP* gene was analyzed using the "Gene Location Visualize from GTF/GFF" function of the TBtools software. The *GGP* genes for potatoes and capsicum were renamed based on their location on the chromosome.

Conserved Motif and Protein Conserved Domain Analysis

The MEME website was used to conduct an online analysis of the shared conserved structural domains of potato and capsicum *GGP* gene families. The structure of the potato and capsicum *GGP* gene families was visualized using the TBtools "Gene Structure View (Advanced)" function.

GGP gene family multiple sequence comparison and phylogenetic tree construction

Potato and capsicum *GGPs* were subjected to Multiple Sequence Comparison Analysis using DNAMAN Software. The *GGP* protein sequences of Arabidopsis, cabbage (*Brassica campestris*), apple (*Malus domestica* cv. *Gala*), and tomato were downloaded from the Arabidopsis, Brassica, NCBI, and Solanaceae databases, respectively. The evolutionary tree was constructed using the MEGA11 software, using the neighbor joining method, with the number of replicates set to 1000 and the rest of the options set to default values.

The evolutionary tree was then further modified using an online website (<https://www.evolgenius.info//evolview/#mytrees/clcle/123>).

Prediction of the role of homeopathic elements of the *GGP* gene family

TBtools was used to extract data 2000 bp upstream of the potato and capsicum *GGP* genes from the potato and capsicum gene databases. The data was submitted to the PlantCARE database (<http://bioinformatics.psb.ugent.be/webtools/plantcare/>) for gene homeotic element (promoter) analysis and was visualized using TBtools.

Analyzing protein structure and expression patterns

The ExPASy online website (https://web.expasy.org/compute_pi/) was used to analyze the molecular weights, instability coefficients, isoelectric points and hydrophilicity of individual members of the potato and capsicum *GGP* genes. The WoLF PSORT online website (<https://wolfpsort.hgc.jp/>) was used to analyze the subcellular localization of each member of the potato and capsicum *GGPs*. The signal peptides of the *GGP* proteins of potatoes and capsicums were predicted using SignalP (<https://services.healthtech.dtu.dk/service.php?SignalP-5.0>). The NPS online website (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_sopma.html) was used to predict the secondary structure. The TMHMM online website (<https://services.healthtech.dtu.dk/service.php?TMHMM-2.0>) was used for transmembrane structural analysis. The UniPort online website (<https://www.uniprot.org/>) was used to extract expression data of the potato *GGP* family in different tissues. The capsicum *GGP* expression profile was analyzed using the transcriptome data set obtained from the sequencing of the capsicum genome and visualized using TBtools³¹.

Transcription analysis of the *GGP* gene under different abiotic stresses and hormonal conditions

The potato "Atlantic" seedlings were selected, and the medium (4.43 g.L⁻¹ MS+1.0 mg.L⁻¹ NAA (naphthalene acetic acid) +0.2 mg.L⁻¹ 6-BA (6-benzylaminopurine) +30 g.L⁻¹ sucrose+6.5 g.L⁻¹ agar) was used as the basic culture medium. The cultures were incubated in an artificial climate chamber at 23 °C, 16 h of light/16 °C, 8 h of darkness. Potato seedlings with good growth similarity and free of contamination were selected for subsequent treatment after 35 d of incubation. For NaCl, ABA, Methyl jasmonate (MeJA), and PEG 6000 treatments, seedlings were transplanted into 1/2 Hoagland nutrient solution containing NaCl (200 mM), ABA (100 µM), MeJA (50 µM), and PEG 6000 (10%) and incubated for 0, 6, 12, 24, and 48 h. The treatment concentrations used were obtained based on pre-experiments. For the cold treatments, the seedlings were placed in 1/2 Hoagland nutrient solution and in a refrigerator (Qingdao Haier Specialty Appliances Co., Ltd., Qingdao, China) at 4 °C for 0, 6, 12, 24, and 48 h. For the dark treatment, potato seedlings were treated by placing them in black, airtight, breathable boxes. Leaf samples were collected for qRT-PCR experiments after treatment at 0, 6, 12, 24, and 48 h.

The capsicum variety "Qiang feng 101", provided by the Vegetable Breeding Laboratory, College of Horticulture, Gansu Agricultural University, was used as the test material for this experiment. The capsicum seeds were soaked and allowed to germinate, then sown in hollow trays. The seedlings were grown under

day/night temperatures of 25°C/20°C, relative humidity of about 60%, 14 h of light and 10 hours of darkness. When the capsicum seedlings had six true leaves, they were treated in the same way as potatoes.

The collected samples were immediately frozen with liquid nitrogen and stored in a vertical ultra-low-temperature refrigerator at -80 °C (Qingdao Haier Special Electric Appliance Co., Ltd., Qingdao, China). Each treatment contained three biological replicates.

RNA Extraction and qRT-PCR Fluorescence Quantification

Total RNA was extracted from the samples using TRIzol reagent (Invitrogen, Carlsbad, CA, USA), taking advantage of the FastQuant First Strand cDNA Synthesis Kit (Tianen, Beijing, China) to synthesize cDNA (The reaction system was 2 µL RNA, 2 µL 5× Evo M-MLV Reagent Premix and 6µL ddH₂O). Dilute the cDNA concentration to 500 ng.µL⁻¹. These reactions were carried out under the following conditions: 37 °C for 15 min, 85 °C for 5 s, and finally ending at 4 °C. LightCycler 480 Real-Time PCR System (Roche Applied Science, Penzberg, Germany) and SYBR Green Premix Pro Taq HS Premix kit were used for qRT-PCR. The reaction system was 2×SYBR Green Pro Taq HS Premix 10 µL, primer (The concentration was 10 µ mol.L⁻¹) F 0.4 µL, primer R 0.4 µL, cDNA 2 µL, and ddH₂O 7.2 µL. The primers used in the qRT-PCR were designed using Primer 5.0 and their specificity was confirmed by melting curve analysis. PCR cycling was performed as follows: 2 min at 95 °C followed by 39 rounds of 5 s at 95 °C, 30 s at the optimal annealing temperature and finally, 1 cycle of 5 s at 65 °C. A melting curve (65–95 °C; at increments of 0.5 °C) was generated to verify the specificity of primer amplification. Three replicates of each tissue sample were used to monitor possible sampling and experimental errors. The efficiency of each primer pair was determined by generating a standard curve using sequence dilutions of the cDNA. ct values were within the linear amplification range to ensure the reliability of the data. The potato and capsicum ACTIN genes were used to normalize relative expression levels, respectively. The specific primer sequences are shown in Table S1. The qRT-PCR data were analyzed using the $2^{-\Delta\Delta Ct}$ calculation method³².

Declarations

Data Statistics and Analysis

All experiments were conducted with at least three independent biological replicates, and all reported data were presented as mean ± standard deviation (SD). Statistical analysis was performed using SPSS statistical software 22.0 (SPSS Inc. Chicago, IL, USA). Multiple comparisons were performed with the Tukey's test when one-way ANOVA showed a significant effect. (* $p \leq 0.05$, ** $p \leq 0.01$). Means ± SD was based on three replicates. Graphs were constructed using GraphPad prism 9.0.0 (GraphPad Software, San Diego, CA).

Author contributions statement

W.L. conceived the project and original research plans; Z.D., K.Y., C.W. and Y.Y. performed the experiments; L.L., Z.D. and L.L. analyzed the data; Z.D. and X.P. wrote the manuscript with input from the other authors. All authors have read and agreed to the published version of the manuscript.

Competing interests

The authors declare no competing interests.

Author Contribution

W.L. conceived the project and original research plans; Z.D., K.Y., C.W. and Y.Y. performed the experiments; L.L., Z.D. and L.L. analyzed the data; Z.D. and X.P. wrote the manuscript with input from the other authors. All authors have read and agreed to the published version of the manuscript.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Nos. 32360743, 32072559, 31860568, 31560563, and 31160398); the Key Research and Development Program of Gansu Province, China (No. 21YF5WA096); the National Key Research and Development Program (2018YFD1000800); the Research Fund of Higher Education of Gansu, China (No. 2018C-14 and 2019B-082); and the Natural Science Foundation of Gansu Province, China (Nos. 1606RJZA073). The funders had no role in the study design, data collection and analyses, decision to publish, or preparation of the manuscript.

Data availability

All data generated or analyzed during this study are included in this article.

References

1. De Tullio, M. C. Is ascorbic acid a key signaling molecule integrating the activities of 2-oxoglutarate-dependent dioxygenases? Shifting the paradigm. *Environmental and Experimental Botany* **178**, 104173, doi:https://doi.org/10.1016/j.envexpbot.2020.104173 (2020).
2. Foyer, C. H. & Noctor, G. Ascorbate and glutathione: the heart of the redox hub. *Plant Physiol* **155**, 2-18, doi:10.1104/pp.110.167569 (2011).
3. Hafez, E. & Gharib, H. Effect of exogenous application of ascorbic acid on physiological and biochemical characteristics of wheat under water stress. *International Journal of Plant Production* **10**, 579-596 (2016).
4. Zhou, S. *et al.* Arabidopsis CaM1 and CaM4 Promote Nitric Oxide Production and Salt Resistance by Inhibiting S-Nitrosoglutathione Reductase via Direct Binding. *PLoS Genet.* **12** (2016).
5. Athar, H. U. R., Khan, A. & Ashraf, M. Exogenously applied ascorbic acid alleviates salt-induced oxidative stress in wheat. *Environmental & Experimental Botany* **63**, 224-231 (2008).
6. Mittova, V., Guy, M., Tal, M. & Volokita, M. Response of the cultivated tomato and its wild salt-tolerant relative *Lycopersicon pennellii* to salt-dependent oxidative stress: increased activities of antioxidant enzymes in root plastids. *Free radical research* **36**, 195-202, doi:10.1080/10715760290006402 (2002).

7. Crizel, R. L. *et al.* Absciscic acid and stress induced by salt: Effect on the phenylpropanoid, L-ascorbic acid and absciscic acid metabolism of strawberry fruits. *Plant Physiol Biochem* **152**, 211-220, doi:10.1016/j.plaphy.2020.05.003 (2020).
8. Nagda, J., Bhanu, A., Katiyar, D., Hemantaranjan, A. & Yadav, D. Mitigating effect of foliar applied ascorbic acid on morpho-physiological, biochemical changes and yield attributes induced by salt stress in *Vigna radiata*. *Agricultural Science Digest - A Research Journal* **37**, doi:10.18805/asd.v37i2.7984 (2017).
9. Broad, R. C. *et al.* Genome-wide identification and characterization of the GDP-L-galactose phosphorylase gene family in bread wheat. *BMC Plant Biol.* **19**, 515, doi:10.1186/s12870-019-2123-1 (2019).
10. Baldet, P. *et al.* Investigating the role of vitamin C in tomato through TILLING identification of ascorbate-deficient tomato mutants. *Plant Biotechnology* **30**, 309-314, doi:10.5511/plantbiotechnology.13.0622b (2013).
11. Laing, W. A. *et al.* An upstream open reading frame is essential for feedback regulation of ascorbate biosynthesis in *Arabidopsis*. *Plant Cell* **27**, 772-786, doi:10.1105/tpc.114.133777 (2015).
12. Zhang, G. Y. *et al.* Manipulation of the rice L-galactose pathway: evaluation of the effects of transgene overexpression on ascorbate accumulation and abiotic stress tolerance. *PLoS One* **10**, e0125870, doi:10.1371/journal.pone.0125870 (2015).
13. Liu, F. *et al.* Higher transcription levels in ascorbic acid biosynthetic and recycling genes were associated with higher ascorbic acid accumulation in blueberry. *Food chemistry* **188**, 399-405, doi:https://doi.org/10.1016/j.foodchem.2015.05.036 (2015).
14. Chaturvedi, S. *et al.* Overexpression of banana GDP-L-galactose phosphorylase (GGP) modulates the biosynthesis of ascorbic acid in *Arabidopsis thaliana*. *International Journal of Biological Macromolecules* **237**, 124124, doi:https://doi.org/10.1016/j.ijbiomac.2023.124124 (2023).
15. Höller, S. *et al.* Ascorbate biosynthesis and its involvement in stress tolerance and plant development in rice (*Oryza sativa* L.). *Plant Mol Biol* **88**, 545-560, doi:10.1007/s11103-015-0341-y (2015).
16. Bulley, S. M. *et al.* Gene expression studies in kiwifruit and gene over-expression in *Arabidopsis* indicates that GDP-L-galactose guanyltransferase is a major control point of vitamin C biosynthesis. *J. Exp. Bot.* **60**, 765-778, doi:10.1093/jxb/ern327 (2009).
17. Müller-Moulé, P. An expression analysis of the ascorbate biosynthesis enzyme VTC2. *Plant Mol Biol* **68**, 31-41, doi:10.1007/s11103-008-9350-4 (2008).
18. Spooner, D. M., Ghislain, M., Simon, R., Jansky, S. H. & Gavrilenko, T. Systematics, Diversity, Genetics, and Evolution of Wild and Cultivated Potatoes. *The Botanical Review* **80**, 283-383, doi:10.1007/s12229-014-9146-y (2014).
19. Tripodi, P. & Kumar, S. 1-8 (2019).
20. Gallie, D. R. L-ascorbic Acid: a multifunctional molecule supporting plant growth and development. *Scientifica* **2013**, 795964, doi:10.1155/2013/795964 (2013).
21. Wang, Y. *et al.* Identification of the GDP-L-Galactose Phosphorylase Gene as a Candidate for the Regulation of Ascorbic Acid Content in Fruits of *Capsicum annuum* L. *Int. J. Mol. Sci.* **24**, 7529 (2023).

22. Tao, J., Hao, Z. & Huang, C. Molecular evolution of GDP-L-galactose phosphorylase, a key regulatory gene in plant ascorbate biosynthesis. *AoB PLANTS* **12**, doi:10.1093/aobpla/plaa055 (2020).
23. Tóth, S. Z., Schansker, G. & Garab, G. The physiological roles and metabolism of ascorbate in chloroplasts. *Physiol Plant* **148**, 161-175, doi:10.1111/ppl.12006 (2013).
24. Dowdle, J., Ishikawa, T., Gatzek, S., Rolinski, S. & Smirnoff, N. Two genes in *Arabidopsis thaliana* encoding GDP-L-galactose phosphorylase are required for ascorbate biosynthesis and seedling viability. *Plant J* **52**, 673-689, doi:10.1111/j.1365-313X.2007.03266.x (2007).
25. Tedone, L., Hancock, R. D., Alberino, S., Haupt, S. & Viola, R. Long-distance transport of L-ascorbic acid in potato. *BMC Plant Biol* **4**, 16, doi:10.1186/1471-2229-4-16 (2004).
26. Li, J., Liang, D., Li, M. & Ma, F. Light and abiotic stresses regulate the expression of GDP-L-galactose phosphorylase and levels of ascorbic acid in two kiwifruit genotypes via light-responsive and stress-inducible cis-elements in their promoters. *Planta* **238**, 535-547, doi:10.1007/s00425-013-1915-z (2013).
27. Wang, L. Y., Li, D., Deng, Y. S., Lv, W. & Meng, Q. W. Antisense-mediated depletion of tomato GDP-L-galactose phosphorylase increases susceptibility to chilling stress. *J. Plant Physiol.* **170**, 303-314, doi:10.1016/j.jplph.2012.10.015 (2013).
28. Broad, R. C. *et al.* Effect of Rice GDP-L-Galactose Phosphorylase Constitutive Overexpression on Ascorbate Concentration, Stress Tolerance, and Iron Bioavailability in Rice. *Front. Plant Sci.* **11**, doi:10.3389/fpls.2020.595439 (2020).
29. Anisimova, O. K., Seredin, T. M., Shchennikova, A. V., Kochieva, E. Z. & Filyushin, M. A. Estimation of the Vitamin C Content and GDP-L-Galactose Phosphorylase Gene (VTC2) Expression Level in Leek (*Allium porrum* L.) Cultivars. *Russian Journal of Plant Physiology* **68**, 85-93, doi:10.1134/S1021443720060023 (2021).
30. Zha, L. *et al.* Regulation of Ascorbate Accumulation and Metabolism in Lettuce by the Red:Blue Ratio of Continuous Light Using LEDs. *Front. Plant Sci.* **11**, doi:10.3389/fpls.2020.00704 (2020).
31. Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol Biol Evol* **35**, 1547-1549, doi:10.1093/molbev/msy096 (2018).
32. Ding, Z. *et al.* SISRK3B Promotes Tomato Seedling Growth and Development by Regulating Photosynthetic Capacity. *Int. J. Mol. Sci.* **25** (2024).

Figures

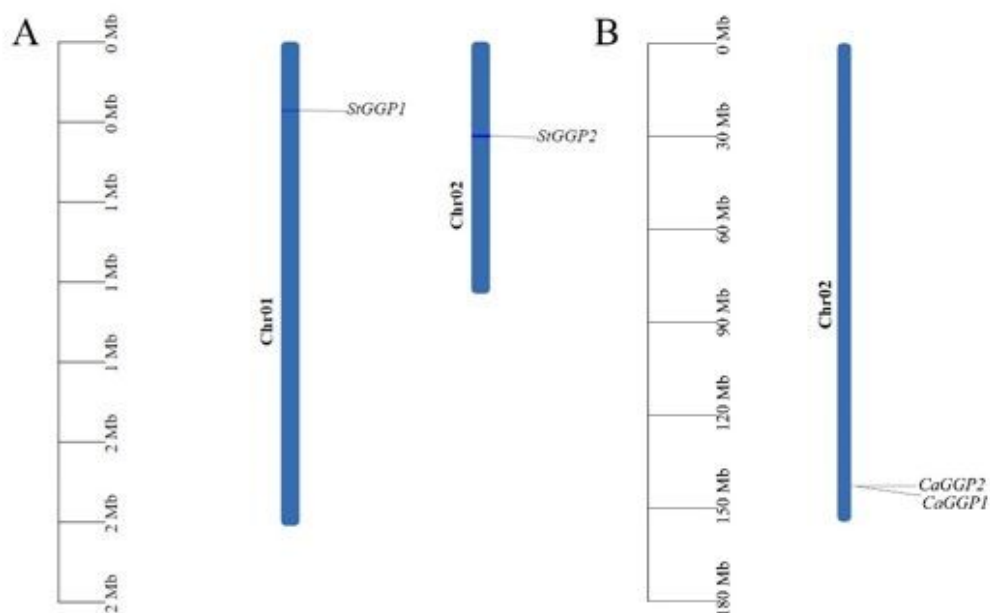


Figure 1

Chromosomal location of the *GGP* gene. (A) Chromosomal location of the potato *GGP* gene, (B) Chromosomal location of the capsicum *GGP* gene, chromosomal numbers shown at top of each bar. Gene names are shown in black. Scale bars are shown on the left.

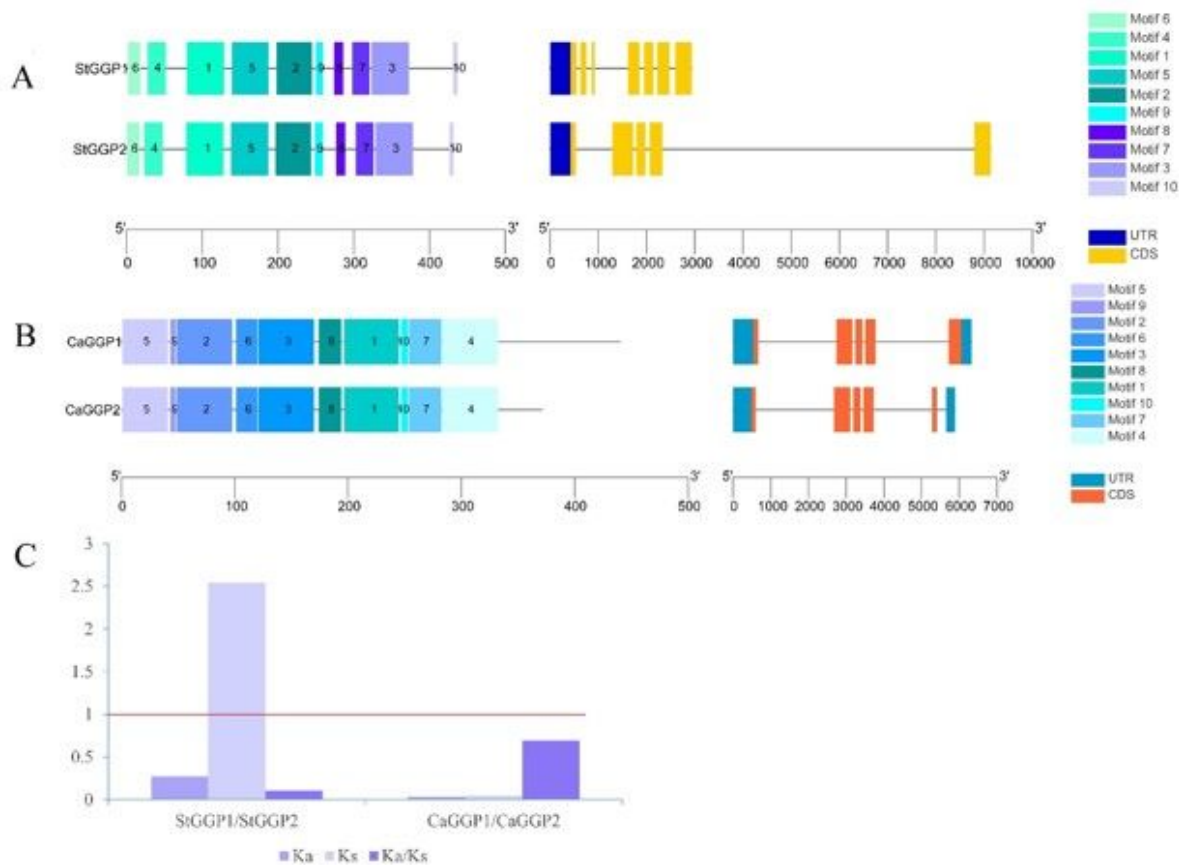


Figure 2

Motif, gene structure and selective evolutionary pressures in the potato and capsicum *GGP* gene families. (A) motif and gene structure of the potato *GGP* gene family; (B) motif and gene structure of the capsicum *GGP* gene family; on the left is the motif, with different colors representing different motifs; on the right is the gene structure; (C) selective evolutionary pressures (Ka/Ks).

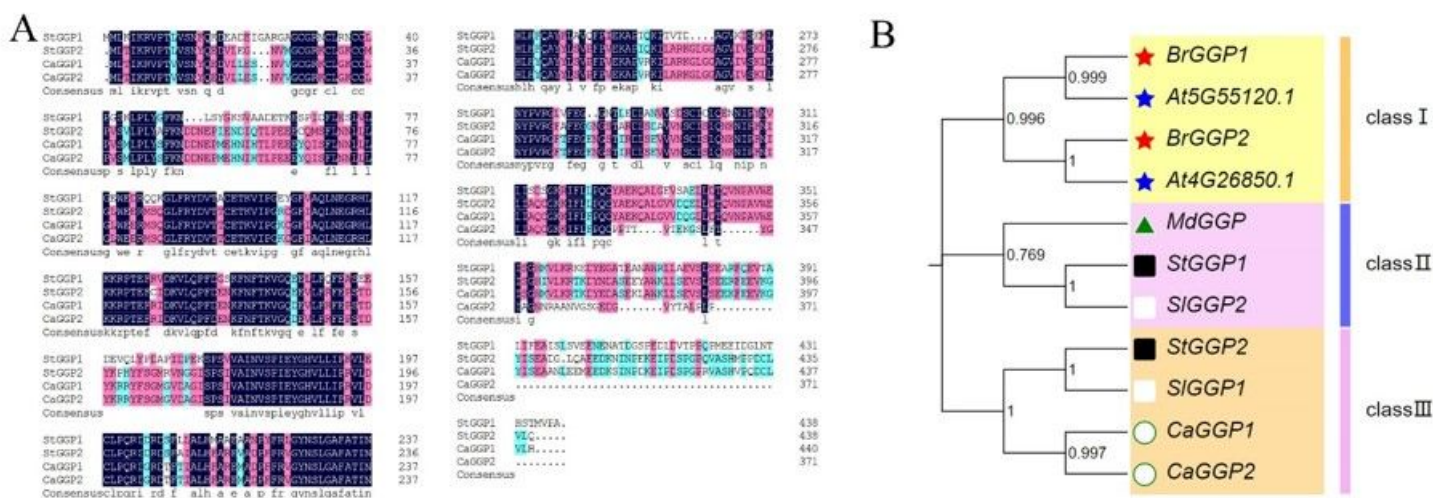


Figure 3

Multiple sequence comparison of potato and capsicum *GGP* gene families and phylogenetic tree of potato, capsicum, tomato, cabbage, Arabidopsis and apple *GGP* gene families. (A) Multiple sequence comparison of potato and capsicum *GGP*. (B) Phylogenetic trees of the *GGP* families of potatoes, capsicum, tomatoes, cabbages, Arabidopsis and apples. The 6 different species are represented by the 6 different colored shapes. The red pentagram represents cabbage. The blue pentagram represents Arabidopsis thaliana. The green triangles represent apples. The black squares represent potatoes. The white squares represent tomatoes. The white circles represent capsicum.

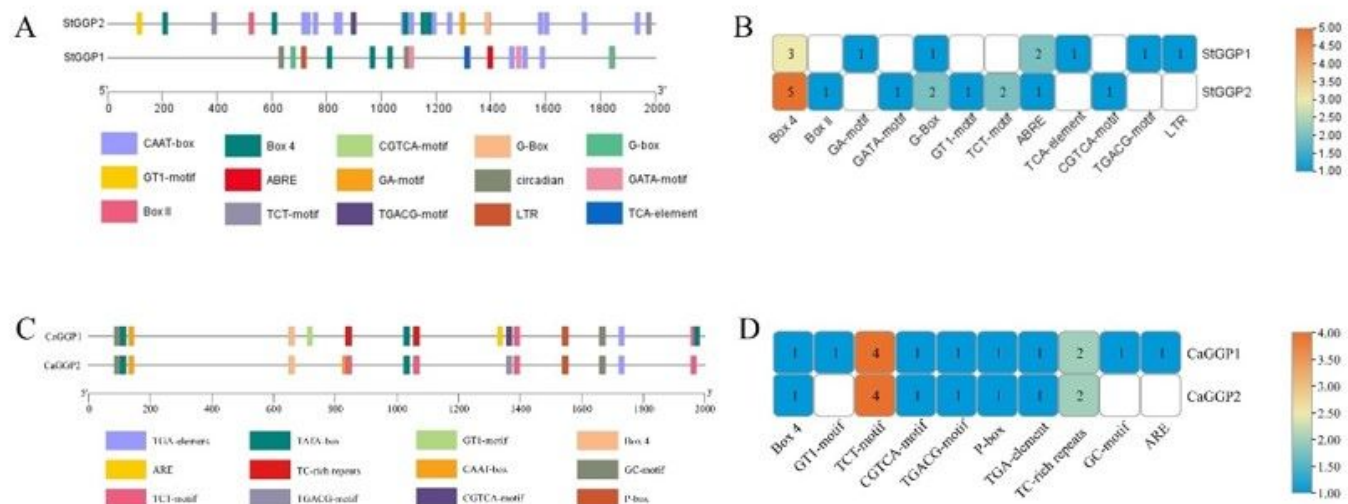


Figure 4

Analysis of cis-acting elements in potato and capsicum. (A) Potato homeopathic component visualization. (B) Potato homeopathic component type statistics. (C) Capsicum Homeopathic Component Visualization. (D) Capsicum Homeopathic Component Type Statistics. Different colored wedges represent different cis elements. The length and position of each *GGP* gene were mapped to scale. The scale bar represents the length of the DNA sequence. The length and position of each *GGP* gene were mapped to scale. The scale bar represents the length of the DNA sequence.

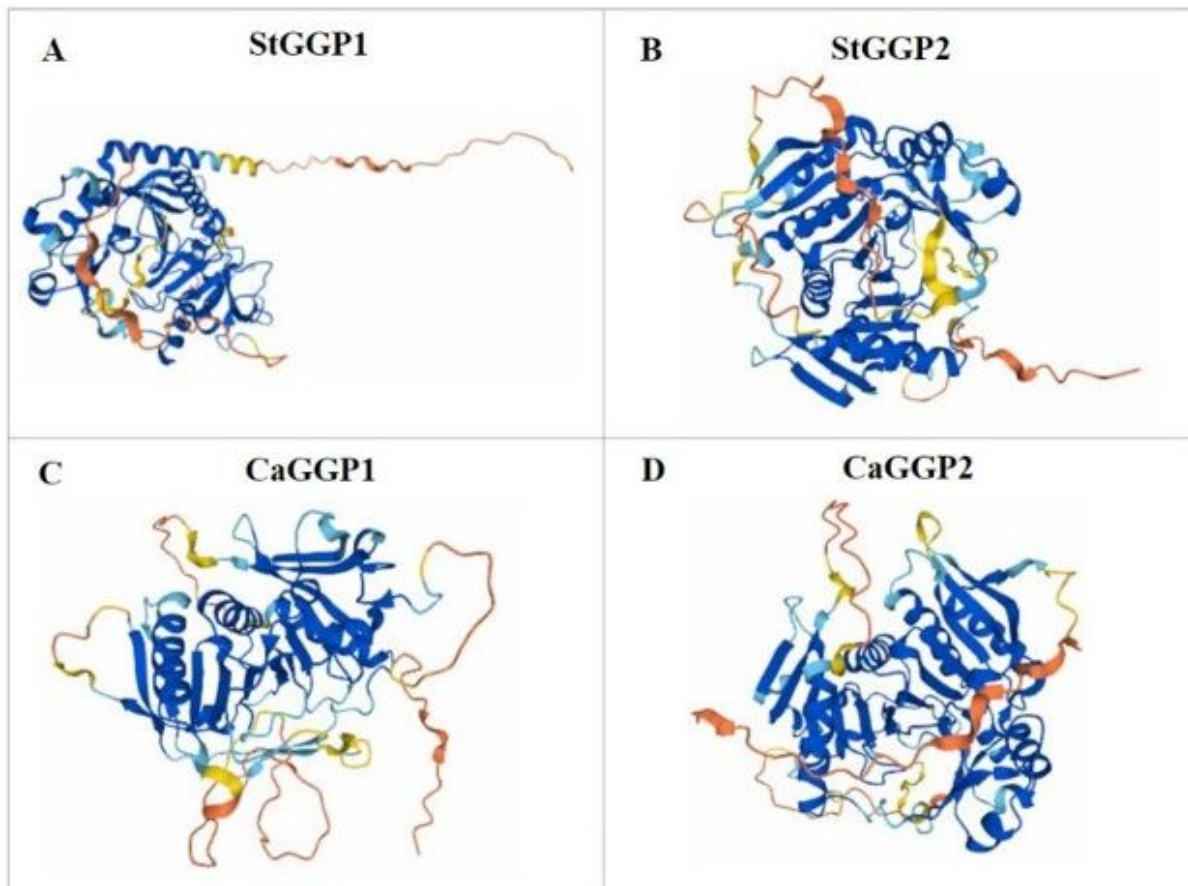


Figure 5

Three-stage structure model of potato and capsicum *GGP*. (A-B) Tertiary structural modeling of the potato *GGP* gene family. (C-D) Tertiary structural modeling of the *GGP* gene family in capsicum.

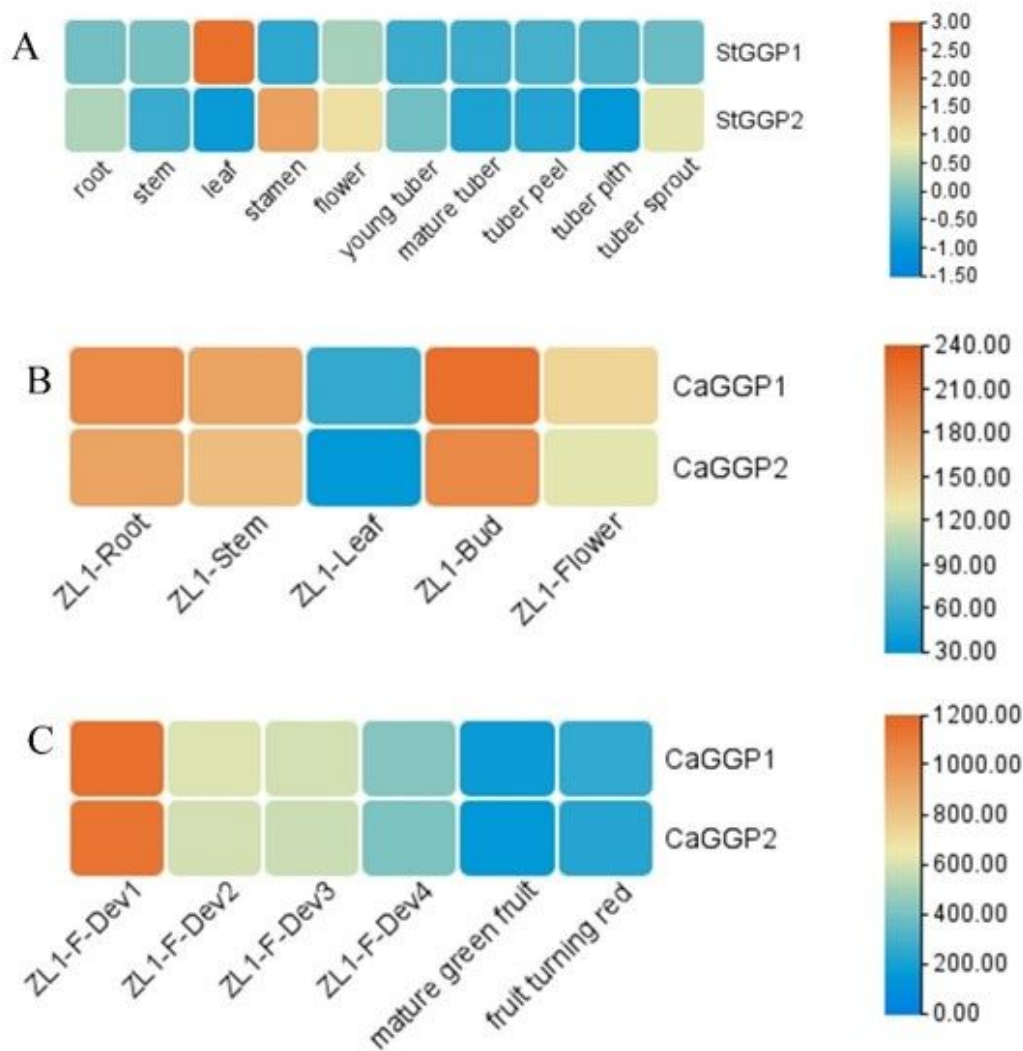


Figure 6

The heatmap is shown in blue/yellow/orange colors indicating low/medium/high expression, respectively. In the figure, ZL1-F-Dev1–ZL1-F-Dev4 represents fruits with lengths of 0-1cm, 1-3cm, 3-4cm and 4-5cm.

Expression analysis of different tissues of potato and capsicum *GGPs* gene families. (A) Expression pattern of potato *GGP* in different tissues. (B) Expression pattern of capsicum *GGP* in different tissues. (C) *GGP* expression patterns at different times during the ripening of the fruit of the capsicum. Color scale represents fold change normalized by $\log_2$ -transformed data.

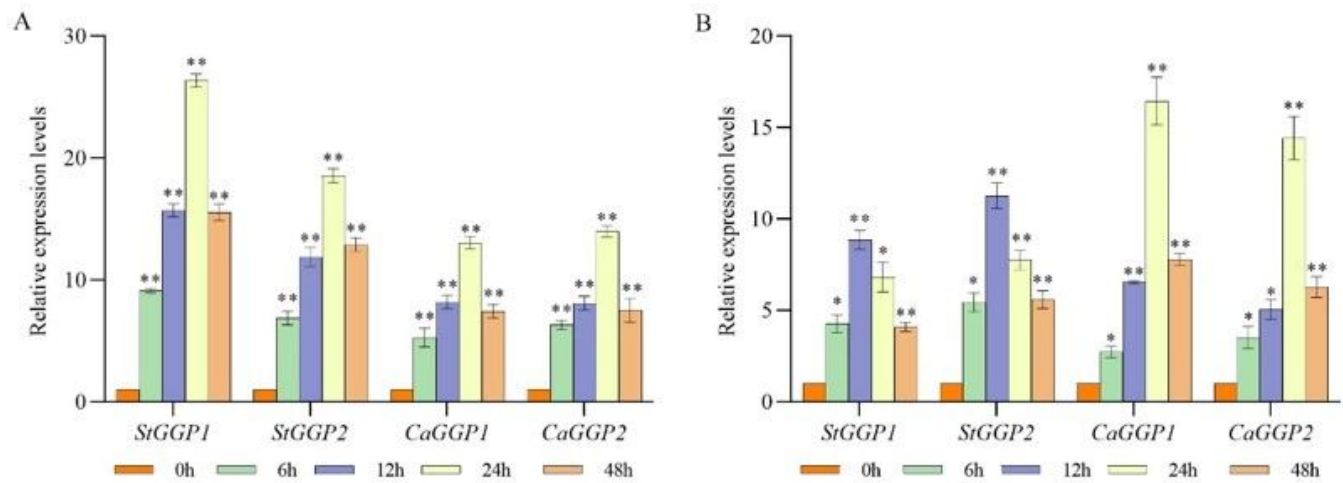


Figure 7

Relative expression levels of *StGGPs* and *CaGGPs* under different hormone treatments. (A) ABA treatment. (B) MeJA treatment. Expression values were normalized to the control (0h) for each gene. An asterisk indicates a significant difference between the stress-treated and control groups. ($p \leq 0.05$, $** p \leq 0.01$, one-way ANOVA, Tukey's test)

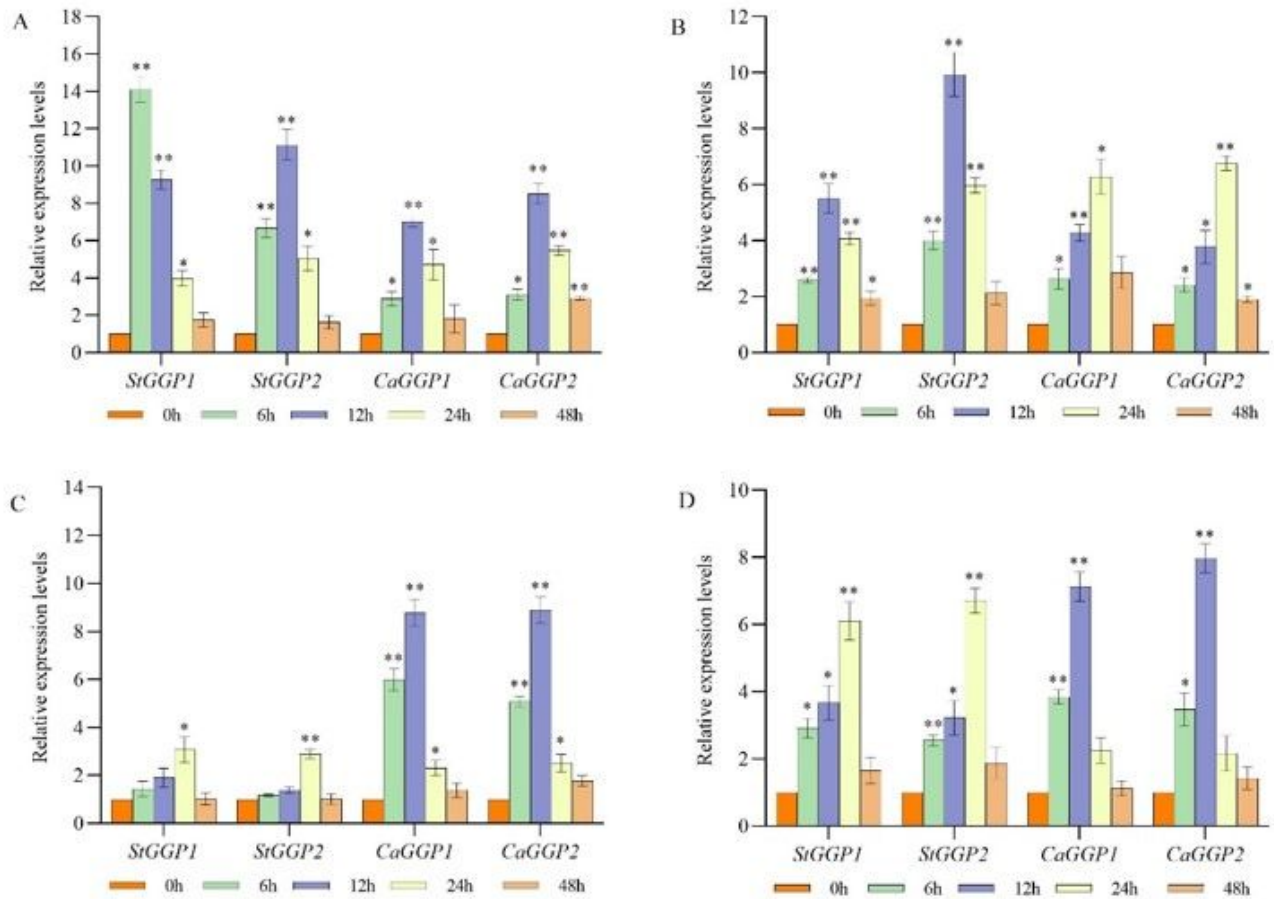


Figure 8

Relative expression levels of *StGGPs* and *CaGGPs* under different hormone treatments. (A) Dark stress. (B) Cold stress. (C) NaCl stress. (D) PEG stress. Expression values were normalized to the control (0h) for each gene. An asterisk indicates a significant difference between the stress-treated and control groups. ($p \leq 0.05$, ** $p \leq 0.01$, one-way ANOV A, Tukey's test)

Supplementary Files

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

- [FileS1.xlsx](#)