

# Evolutionary Expansion and Expression Analysis of the Oligopeptide Transporter Gene Family Between *Cucurbita Moschata* and *Cucurbita Maxima*

Changwei Shen

Henan Institute of Science and Technology

Jingping Yuan (✉ [jpyuan666@163.com](mailto:jpyuan666@163.com))

Henan Institute of Science and Technology <https://orcid.org/0000-0001-5842-8489>

---

## Research article

**Keywords:** Oligopeptide transporter, *Cucurbita moschata*, *Cucurbita maxima*, Yellow stripe-like transporters, Salt stress

**Posted Date:** September 8th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-68770/v1>

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

---

## Abstract

**Background:** Pumpkin is an important non-saline economic vegetable, salt stress often restricts the growth of pumpkin roots and the transportation and balance of mineral ions in the body. Oligopeptide transporter (OPT) plays an important role in transporting small peptides, secondary amino acids, Glutathione and minerals. However, information about the family of OPT in pumpkin is still limited.

**Results:** In this study, 45 OPT transporters were identified from two cultivated species of *Cucurbita moschata* and *Cucurbita maxima*. Phylogenetic analysis showed that these OPT families were divided into two evolutionary branches: OPT clade and yellow stripe-like (YSL) clade. All of these genes contain the typical structure of OPT superfamily, OPT clade contains 12 conserved motifs, while the YSL clade contains 7 conserved motifs. There are tandem gene replication events on chromosomes 13, 16 and 18 of *Cucurbita moschata* and *Cucurbita maxima*, and 17 and 18 pairs of genes were collinear with *Arabidopsis thaliana*, respectively. Promoter element analysis showed that there were many *cis-acting* elements in the upstream promoters of these genes in *Cucurbita moschata* and *Cucurbita maxima*, which responded to 10 kinds of stress, especially hormones (MeJA and ABA) and hypoxia. The expression patterns based on transcriptome data sources showed that some OPT genes were organ-specific and tissue-specific, which might be involved in plant functional development. Transcriptome and qRT-PCR verification tests showed that *CmoYSL7*, *CmaOPT15* and *CmaYSL7* of *Cucurbita moschata* and *Cucurbita maxima* might play a key role in regulating the balance of metal ions between leaf mesophyll and leaf veins under salt stress.

**Conclusions:** Overall, the data obtained from our study contribute to a better understanding of the complexity of the OPT genes family in pumpkins. These results will provide new insights into the mechanism of salt tolerance and the structure and function of OPT genes in pumpkin.

## Background

Soil salinization is one of the natural stresses often encountered in crop production. Salinized soil is widely distributed in the world. There are about  $6 \times 10$  hm<sup>2</sup> saline-alkali land in China [1]. In a saline environment, plant growth is affected by the interaction between salt ions and many mineral nutrient ions, resulting in an imbalance in the absorption, utilization and distribution of nutrients in plants, but also increases the demand for essential nutrient elements. The harm of salt stress to plants is mainly caused by ion toxicity, osmotic stress and nutritional imbalance [2]. Ion transport is blocked under salt stress, especially microelement in plants. The transport of these elements is closely related to oligopeptide transporters (OPT), which were first described in *Candida albicans* by Lubkowitz et al. [3]. Up to now, great progress has been made in the role of these transporters in their respective organisms. OPTs play a variety of roles in long-distance sulfur distribution, metal homeostasis, nitrogen mobilization and heavy metal chelation by transporting glutathione, peptide and meta-chelates. For example, glutathione transporters are likely to be involved in heavy metal detoxification and long-distance sulfur transport, while other OPTs may be involved in long-distance nitrogen transport. Great changes in nitrogen distribution occur at certain stages of plant life, such as seed formation, senescence and germination, while OPT transporters provide a potentially energy-efficient and rapid way to move large amounts of amino acids [4–9]. Phylogenetic analysis has shown that the OPT family has two different branches, called the yellow stripe (YS) branch and the peptide transport (PT) branch. The OPTs in the YS branch exist in archaea, bacteria, plants and fungi, while the members of the PT branch only exist in plants and fungi. Plants are the only organisms with the characteristics of YS and PT transporters, and all these OPTs seem to be plasma membrane binding proteins [10]. The binding substrates of OPTs can be roughly divided into two categories: small peptides with three to five residues and metal-bound secondary amino acids, especially plant iron carriers and nicotinamide (NA) [4, 6]. OPTs in the branch of YS play an important role in the acquisition and long-distance transportation of metals in plants [11, 12].

There were reported the the characteristics of OPT clade in *Arabidopsis thaliana* [13], rice [6], *Bras-sica juncea* (L.). Czern [14]. The most typical is the analysis of OPT genes family in *Arabidopsis thaliana*. There are 9 OPT genes (*AtOPT1-9*) in *Arabidopsis thaliana*. Except for the weakly expressed *AtOPT8*, other *AtOPTs* are strongly expressed in germinated seedlings, especially in the vascular tissues of cotyledons and hypocotyls. In addition, *AtOPTs* are preferentially expressed in vascular tissues of cotyledons [13], leaves, hypocotyls, roots, flowers, pods and seeds of seedlings and adult plants [13]. *AtOPT3* is associated with metal transport and seems to play a key role in seed development [15]. It has been proved that *AtOPT6* is involved in the long-distance peptide transport or distribution of *Arabidopsis thaliana* [13]. In stems and leaves, the expression of the *AtOPT6-GUS* reporter gene was the strongest in vascular bundles [16]. *AtOPT6* may also play a special role in growth and reproduction; after fertilization, *AtOPT6* is strongly expressed in the vascular system of ovary and style [17]. In addition, the OPTs in the evolutionary branch of PT plays a role in long-distance transport, and its main substrate is most likely to be a peptide, including GSH[17].

YSL1 in YSL clade (PT-YSLs) was first found in maize [11]. There are 8 and 19 homologs of YS1 [18] in *Arabidopsis* and rice genomes, respectively [18]. Some homologs were also found in barley [19] and *Thlaspi caerulescens* [20]. These YSLs are not only involved in the acquisition of iron but also involved in the internal transport of iron [17]. Regardless of the overall classification of YSLs, most OPT proteins are involved in metal transport, detoxification or migration [13, 21]. Members of the YSL subfamily specialize in the long-distance transport of metal chelates and play an important role in metal (Fe, Zn, Cu, Mn, Ni) homeostasis [22].

With the development of High-throughput sequencing technology in recent years, more and more plant species have been sequenced, and the characteristics of OPTs family genes have also been described, such as *Ganoderma lucidum* [23–26] and so on. The growth and development of crops are often affected by external biotic and abiotic environmental stresses, especially abiotic stresses (drought, high temperature, salt stress, freezing injury, etc.). High salt is one of the abiotic stresses that limit the absorption of mineral nutrients by crop roots, leaf photosynthesis and other biological activities in the north. Pumpkin is an important vegetable and cash crop in northern China. People pay more and more attention to its health care function [27, 28], but the soil pH in the northern region is somewhat alkaline, containing a large amount of calcium, iron, magnesium and other microelements as well as sodium ions. With the increasing expansion of the saline-alkali soil area, pumpkin as a non-halophytic vegetable, it is very necessary to study the mechanism of salt tolerance and select salt-tolerant varieties. Although pumpkins are economically important, the latest available genomes (*Cucurbita moschata* and *Cucurbita maxima*) are not available

until recently [27]. Up to now, the effect of salt stress on the expression profile of OPT family members in seedlings of different pumpkin varieties has not been clear.

In this study, 45 OPTs family genes were identified from two pumpkin varieties (*Cucurbita moschata* and *Cucurbita maxima*) grown in different regions. Then these OPTs family members were analyzed by gene structure, phylogenetic tree, chromosome location, gene replication, Ka/Ks, collinearity between pumpkin group and other families, and expression patterns in tissues and organs. To clarify the changes in OPTs family members' expression profile under salt stress, the differential expression patterns of OPTs between *Cucurbita moschata* and *Cucurbita maxima* were analyzed based on published transcriptional ancestral data and RT-qPCR test. Combined with this information, new insight was provided for the analysis of the mechanism of salt tolerance and the structure and function of the *OPT* gene in pumpkin.

## Results

### Genome-wide identification of OPT genes in pumpkin

To identify the members of the *OPT* gene family, we searched the *Cucurbita moschata* and *Cucurbita maxima* genome databases with 17 AtOPTs (AtOPTs and AtYSLs), and further identified 21 and 24 *OPT* genes in two pumpkin genomes by using OPT domain (PF03169) and HMMER software. These genes are named as *CmoOPT1-13*, *CmoYLS1-8*, *CmaOPT1-15* and *CmaYLS1-8* according to the location and distribution of these genes on chromosomes (Fig. 1, Table S1). In *Cucurbita moschata*, *CmoOPT1-13* and *CmoYLS1-8* encoded amino acid lengths from 659 to 1360, pI from 8.38 to 9.33, GRAVY from 0.075 to 0.515, and MW from 72899.67 Da to 127699.53 Da, predicted transmembrane structures from 12 to 26. The prediction of protein subcellular localization showed that except for *CmoOPT2*, *CmoOPT3* and *CmoYSL1* in *Cucurbita moschata*, all proteins were predicted to be located in the cell membrane.

### Evolutionary analysis of OPTs in pumpkin

To study the relationship of *OPT* genes among multi-species, we constructed a phylogenetic tree (Fig. 1) using CmoOPTs and CmaOPTs amino acid sequence sequences and OPT amino acid sequences in Arabidopsis and rice. The results showed that the *OPT* genes of the four plants were divided into OPT clade (type A) and YSL clade (type B), and the OPT clade was subdivided into A1-A4, YSL clade was subdivided into B1-B3. In the OPT clade, A3 contained at least 4 *OPT* genes of pumpkin, A4 contained 12 *OPT* genes of pumpkin, and A1 and A2 both contained 6 *OPT* genes of pumpkin. In the YSL clade, B2 contained at least 2 *YSL* genes of pumpkin, B3 contained at most 8 *YSL* genes of pumpkin, and B1 contained 6 *YSL* genes of pumpkin.

## Motifs analysis of OPTs protein in pumpkin

Through the conservative domain analysis of 45 *OPT* genes in *Cucurbita moschata* and *Cucurbita maxima*, as shown in Fig. 2A, these genes can be divided into two clades: OPTs and YLSs, the OPTs clade is divided into A1-A4 subfamily, and the YLSs clade is divided into B1-B3 subfamily. All of these genes in Fig. 2B contained the typical structure of OPT superfamily, and the corresponding structure contained different numbers of a motif. Except for *CmaOPT9*, *CmaOPT12* and *CmoOPT7*, other *OPT* genes in OPTs clade have 12 conserved motif patterns (such as Motif 5-9-3-7-8-10-6-4-14-2-1-11), while YSL genes in YLSs clade have 7 conserved motif patterns (such as Motif 9-3-3-13-4-15-12). There were 9 conserved motifs in CmaYSL4 and CmoYSL4 in B2 subfamily, while 8 conserved motifs in B1 and B3 subfamilies. From the intron and exon point of view, the number of exons in the OPTs clade (Fig. 2C) ranged from 4 to 27, while that in the YLSs clade ranged from 6 to 13. B1 and B3 subfamilies contained 8 and 7 introns, respectively, while the B2 subfamily also contained 8 introns, but their distribution was different from that of the B1 subfamily.

### Multiple sequence alignment of OPT genes in pumpkin

By multiple sequence alignment of the amino acid sequences of the *Cucurbita moschata* (21 *OPT* genes) and *Cucurbita maxima* (24 *OPT* genes) by OPT (29) and YLS (16) clades, it was found that 29 OPT amino acid sequences shared 12 transmembrane domains (Fig. S2), and that they also shared the structure of the NPG and KIPPR motifs, containing 26 and 27 amino acid residues, respectively. The YLS amino acid sequences in the YLS clade contained 13 transmembrane domains (Fig. S3).

### Chromosome location of OPT genes and collinearity analysis of in pumpkin genome

According to the distribution of *OPT* genes in *Cucurbita moschata* and *Cucurbita maxima* in chromosomes, it could be seen that these genes were mainly distributed on 9 chromosomes, of which chromosomes 1, 2, 3, 4, 9, 11 and 18 of the two genomes. There was the same number of genes among 9 chromosomes (Fig. 3). *Cucurbita moschata* contained 6, 4 and 4 genes on chromosomes 13, 16 and 18, respectively, while *Cucurbita maxima* contained 3, 5, and 4 genes on chromosomes 13, 16 and 18, respectively. There were gene tandem replication events on these chromosomes. To explore the collinearity of *OPT* genes in the two genomes, this analysis was performed using the MCScanX method (Fig. 3A and 3B, Table S2). We found that there were 4 pairs of fragment duplications in *Cucurbita moschata* chromosomes, however, there were 15 pairs of fragment duplication phenomenon in *Cucurbita maxima* chromosomes. These fragment duplications mainly occur between chromosomes 13, 16 and 18 of the two genomes. By analyzing the Ka, Ks, Ka/Ks and the evolutionary age of the two genomes, it was found that the Ka/Ks of the gene pairs existing in two genomes were all less than 1, which tended to be purified selection, indicating that *OPT* genes in the evolutionary process were relatively conservative, and had not undergone major changes (Table S2). From the perspective of evolution time, the evolutionary time of 4 pairs of gene duplication events in the *Cucurbita moschata* chromosome was 4.56–18.02 MYA, while that of 15 pairs of gene duplication events in the *Cucurbita maxima* chromosome was 0.32–18.17 MYA. It showed that although these genes were conserved in sequence, there were differences in evolutionary time.

### Collinearity of OPT genes between pumpkin and other families

The collinear analysis was performed with the corresponding gene blocks of *Cucurbita moschata* (21 OPT genes), *Cucurbita maxima* (24 OPT genes) and *Arabidopsis thaliana* (16 OPT genes) (Fig. 4, Table S3). It was found that there were 24 pairs of gene collinear between *Cucurbita moschata* and *Cucurbita maxima*, mainly on chromosomes 1, 2, 4, 9, 11, 13, 16 and 18 of two genomes. There were 17 pairs of genes collinear between *Cucurbita moschata* and *Arabidopsis thaliana*, which were mainly concentrated on chromosomes 1, 2, 4, 9, 11, 13, 16, 18 in *Cucurbita moschata* and 1, 3, 4, 5 chromosomes in *Arabidopsis thaliana*. There were 18 pairs of genes collinear between *Cucurbita maxima* and *Arabidopsis thaliana*, which was consistent with the location of chromosomes distributed in *Cucurbita maxima* and *Arabidopsis thaliana*.

### Promoter analysis of OPT genes in pumpkin

To further clarify the function of the OPT family members of *Cucurbita moschata* and *Cucurbita maxima*, we selected the 2000 bp promoter region upstream of the CDS sequence to visualize the *cis-acting* elements by using TBtools software (Fig. 5). There were many *cis-acting* elements in the upstream promoters of these genes. They all responded to 10 kinds of stress (hormone response, anaerobic response, defense and resistance response, drought induction, light response, low-temperature response, etc.). In *Cucurbita moschata*, the *cis-acting* elements responsive to light regulation were distributed in all genes, while *CmoOPT12* responsive to hormone (MeJA) contained as many as 16 *cis-acting* elements, followed by *CmoOPT3* containing up to 14 *cis-acting* elements, and *CmoYSL5* responsive to ABA contained 10 *cis-acting* elements (Fig. 5A). In *Cucurbita maxima*, there were at least 5 *cis-acting* elements in response to light regulation, while *CmaYSL5* in response to hormone-like (ABA) contained as many as 10 *cis-acting* elements; *CmaOPT2*, *CmaOPT11* and *CmaYSL6* in response to IAA all contained 6 *cis-acting* elements; *CmaOPT14* and *CmaOPT10* responsive to hypoxia contained 14 and 8 *cis-acting* elements, respectively (Fig. 5B).

### Expression profiles of OPT genes in different tissues in pumpkin

To analyze the expression of OPTs family genes in different tissues using common transcriptome data (Fig. 6, Table S4), we found that the expression abundance of *CmaOPT1*, *CmoOPT1*, *CmoOPT4*, *CmoOPT11* and *CmaYSL1* was lower in different tissues. In addition, the expression abundance of all genes in fruit was the lowest compared with other tissues. The expression levels of *CmoYSL1*, *CmoYSL2*, *CmoYSL3*, *CmoYSL6*, *CmoOPT2*, *CmoOPT6* and *CmoOPT8* in *Cucurbita moschata* root were higher than those in other tissues. The expression levels of *CmaOPT3*, *CmaOPT5-10*, *CmaOPT12*, *CmaOPT15*, *CmaSYL2*, *CmaSYL4-5* and *CmaSYL8* in *Cucurbita maxima* leaves was higher than those in other tissues.

## Expression profiles of OPTs in pumpkin under salt treatment and qRT-PCR verification

To explore the response of OPTs in pumpkin vein and mesophyll to salt stress, previous RNA-seq data were analyzed (Table S5). There was a significant difference in the expression levels of OPTs family members in mesophyll and vein between *Cucurbita moschata* and *Cucurbita maxima* seedlings under salt stress (Fig. 7). Compared with the control treatment, the expression levels of *CmoOPT2*, *CmoOPT5*, *CmoOPT6*, *CmoOPT7*, *CmoOPT8*, *CmoOPT12*, *CmoYSL1*, *CmoYSL2*, *CmoYSL3*, *CmoYSL4*, *CmoYSL5* and *CmoYSL8* in the mesophyll of *Cucurbita moschata* were up-regulated under the NaCl treatment, while the expression levels of *CmoOPT3*, *CmoOPT9*, *CmoOPT10*, *CmoOPT11*, *CmoOPT13*, *CmoYSL6* and *CmoYSL7* were down-regulated (Fig. 7A). Compared with the control treatment, the expression levels of *CmoOPT5*, *CmoOPT8*, *CmoOPT12*, *CmoYSL4* and *CmoYSL7* in the vein of *Cucurbita moschata* was up-regulated under the NaCl treatment, while the expression levels of *CmoOPT1*, *CmoOPT3*, *CmoOPT10* and *CmoYSL6* were down-regulated (Fig. 7B). Compared with the control treatment, the expression levels of *CmaOPT2*, *CmaOPT3*, *CmaOPT9*, *CmaOPT12*, *CmaOPT13*, *CmaYSL1* and *CmaYSL7* in the mesophyll of *Cucurbita maxima* was up-regulated under the NaCl treatment, while the expression levels of *CmaOPT5*, *CmaOPT6*, *CmaOPT8*, *CmaOPT11*, *CmaOPT14*, *CmaOPT15* and *CmaYSL8* in the mesophyll of *Cucurbita maxima* was down-regulated (Fig. 7C). Compared with the control treatment, the expression levels of *CmaOPT4*, *CmaOPT6*, *CmaOPT12*, *CmaOPT15* and *CmaYSL8* in the vein of *Cucurbita maxima* was up-regulated under the NaCl treatment, while the expression levels of *CmaOPT2*, *CmaOPT9*, *CmaOPT12*, *CmaYSL1*, *CmaYSL6* and *CmaYSL7* was down-regulated (Fig. 7D).

To further verify the response of OPT family members to salt stress, we applied salt stress to the seedling stage of *Cucurbita moschata* "Baimi". It was found that after 12 hours of salt stress, the relative expression of *CmoOPT3* and *CmoYSL7* in the mesophyll of "Baimi" under salt stress was significantly lower than that of the control treatment, while the relative expression of *CmoOPT5-8*, *CmoOPT12-13*, *CmoYSL4* and *CmoYSL8* was significantly higher than that of the control treatment (Fig. 7E). The relative expression of *CmoOPT3*, *CmoYSL5* and *CmoYSL7* in the vein of "Baimi" under salt treatment was significantly higher than that of the control treatment, and the relative expression of *CmoOPT8*, *CmoOPT10* and *CmoOPT12* was significantly higher than that of the control treatment (Fig. 7F).

## Discussion

### Evolutionary relationship of OPTs in *Cucurbita moschata* and *Cucurbita maxima*

Pumpkin (*Cucurbita moschata* Duch.) is cultivated by *C. moschata* (*Cucurbita moschata* Duch.), *C. Pepo* (*Cucurbita Pepo* L.), *C. maxima* (*Cucurbita maxima* Duch.) and several wild species [29]. *C. moschata* and *C. maxima* are both economic vegetables and widely cultivated, and the genomes of *C. moschata* and *C. maxima* are only recently reported [27]. In this study, a total of 45 OPT family members in two cultivars were found by searching the published pumpkin genome data. The comprehensive analysis of the phylogenetic tree, exon/intron gene structure and conserved motif of these genes in two cultivated species enables us to make some predictions and generalizations on the possible origin and interrelationship of OPTs. Phylogenetic analysis of two cultivated species, rice and *Arabidopsis* showed that the members of the OPT family were divided into two branches (OPT clade and YSL clade), which is basically consistent with the previous classification of *Arabidopsis thaliana* [13], rice [6] and several other species [10], indicating that the members of the OPT family were relatively conservative in the process of evolution. The structural differences of exons/introns can be used to determine the evolutionary history of gene families to some extent. From the point of view of introns and exons (Fig. 2C), the range of exon number of OPT clades was wider than that of the YSL clade, indicating that the variation of OPT clade members was larger in the process of evolution, while YSLs clade members were relatively conservative. There were 27 exons in *CmaOPT4* and 4 exons in *CmoOPT12*. The numbers of exons in *Arabidopsis thaliana* and rice were usually 4–6 and 1–7, respectively [6]. The

results showed that the introns and exons of OPTs members in the OPT clade of pumpkin experienced strong gene replication in the process of evolution. This is also confirmed by the collinearity (Fig. 3 and Fig. 4) of OPT genes among *C. moschata*, *C. maxima* and *Arabidopsis thaliana*. In addition, the gene replication of *C. moschata* and *C. maxima* mainly occurred on chromosomes 13, 16 and 18. According to the evolutionary time, the chromosome gene repetition event of *C. moschata* occurred at 4.56–18.02 Mya, while that of *C. maxima* occurred at 0.32–18.17MYA. Some studies have shown that the earliest replication event of Cruciferae is expected to occur in 34MYA [30]. This showed that the gene replication of the two cultivated species occurred at almost the same time. At the same time, the Ka/Ks values of gene pairs in OPT and YSL clade members were less than 1 (Table S2), which further indicates that these genes were purified in the process of evolution.

- **Tissue expression of OPT genes in *Cucurbita moschata* and *Cucurbita maxima***
- The expression profiles of family members in crops can provide useful clues for gene function [25]. In this study, the expression levels of 45 OPT genes were analyzed using publicly available RNA sequence data from the genomes of *Cucurbita moschata* and *Cucurbita maxima* [27]. It was found that the expression abundance of all genes in the fruits of the two cultivars was the lowest compared with other tissues. The expression abundance of 7 genes in *Cucurbita moschata* was higher than that in other tissues, and the expression abundance of 13 genes in *Cucurbita maxima* was higher than that in other tissues (Fig. 7). The expression levels of *CmoYSL3*, 5, 7 and *CmaYSL3*, 5, 7 were the highest in the roots, and these genes belonged to the B3 branch, in which the *OsYSL12* and *OsYSL13* gene was closely related to its evolution. It has been reported that the highest expression of *OsYSL12* and *OsYSL13* is in the roots of 14-day-old seedlings in rice [31]. These showed that these OPT genes were tissue-specific, and had multiple functions in different tissues. On the other hand, some genes might play a role in specific tissues. Some studies have shown that *AtOPT3* [32] and *AtOPT6* [16] in *Arabidopsis thaliana* expressed in seed tissues during development, which might be involved in the process of embryo development. *OsOPT1*, *OsOPT3*, *OsOPT4* and *OsOPT7* were expressed in rice embryos [6]. In this study, except for A3 clade, other genes were closely related to these genes in rice, suggesting that these OPTs might be related to embryo development. In addition, little is known about the role of OPT genes on stress response in pumpkin. Through promoter element analysis and tissue expression analysis, we obtained some clues about the role of pumpkin OPT genes on stress response [33]. Some studies indicated that GT1-motif and TGACG-motif were identified as salt stress response elements [34], ABRE, G-box, MBS and TGA-element had a regulatory role under salt stress. Therefore, we speculated that OPTs in *Cucurbita moschata* and *Cucurbita maxima* might play a key role in salt stress.

#### Expression profile of OPT genes in *Cucurbita moschata* and *Cucurbita maxima* under salt stress

In plants, Na<sup>+</sup> is either distributed in specialized cells [35], or isolated in vacuoles in the mesophyll, so the cytoplasmic Na<sup>+</sup> level of cytoplasm in functional cells is kept at a low level [36, 37]. The ion changes in plants under salt stress are closely related to K<sup>+</sup>, Mg<sup>2+</sup>, Ca<sup>2+</sup> and microelements. Functional diversification is the result of the evolutionary expansion of gene families through gene duplication, usually accompanied by changes in the expression profiles of gene family members. The OPT family is divided into two subfamilies (OPT subfamily and YSL subfamily) [11], in which the YSL subfamily is considered to be able to transport iron compounds, and the OPT subfamily is considered to be able to transport oligopeptides with 45 amino acid residues [38]. However, some members of the OPT subfamily have also been identified as iron transporters, such as *AtOPT3* [13], *OsOPT1*, *OsOPT4* and *OsOPT7* [6]. According to the published RNA-seq data, the transcriptional data of *CmoOPTs* and *CmaOPTs* in leaf mesophyll and leaf vein under salt stress was analyzed. It was found that the expression levels of *CmoYSL7*, *CmaOPT15* and *CmaYSL7* in the mesophyll and veins of the two cultivated species were opposite under salt treatment. It was speculated that these genes played a key role in regulating the balance of metal ions between mesophyll and veins under salt stress (Fig. 7). The qRT-PCR technology was used to further verify the expression of these genes in the veins and mesophyll of *Cucurbita moschata* (Fig. 8). This result is almost similar to the transcriptional result. In this study, the response of pumpkin to salt stress was discussed from the expression profile of OPTs family genes, but the specific function and mechanism of these OPT genes in pumpkin need to be further studied.

## Conclusions

In summary, the genomic analysis of OPT genes family in *Cucurbita moschata* and *Cucurbita maxima* was carried out, and the phylogeny, gene structure and homologous repetition history of OPTs family were discussed. We identified a total of 45 OPTs in the pumpkin genome and provided genetic information such as chromosome location, exon-intron structure, conserved domain and repetitive genes. In particular, we examined the expression profiles of these OPT genes in different tissues. At the same time, we detected the response of 45 OPT genes to salt stress and verified by qRT-PCR, and found that several key genes played a role in regulating the balance of metal ions between mesophyll and veins. These data may provide valuable information for the functional study of the gene family in the future.

## Methods

### Genome-wide identification of OPT genes in pumpkin

We downloaded the whole genome data from the *Cucurbita moschata* and *Cucurbita maxima* genome database (<http://cucurbitgenomics.org/>), and identified the possible members of the OPTs gene family in these two genomes by BLASTP tool. Seventeen OPT proteins [10] from *Arabidopsis thaliana* were used as query sequences to search the two genome databases. We choose the protein sequences (E value is lower than 1E<sup>-10</sup> value) as the candidate proteins. The obtained sequences were further confirmed through the CDD (conservative domain database) in NCBI and the PFAM website (<http://pfam.xfam.org/>) download OPT domain (PF03169). The integrity of the OPT domain in candidate genes was verified by ExPASy (<http://prosite.expasy.org/>). Each OPTs gene was given a unique name according to its exact location on the chromosome (each chromosome from top to bottom). Prediction of the transmembrane helix was performed with TMHMM Server v. 2.0 (<http://www.cbs.dtu.dk/services/TMHMM/>) [39]. The molecular weight, isoelectric point and grand average hydropathy (GRAVY) values of each OPT gene were estimated using the ExPASy Proteomics Server (<https://web.expasy.org/tools/protparam.html/>) [40]. The subcellular location of OPTs gene was predicted using the Plant-mPLoc (<http://www.csbio.sjtu.edu.cn/bioinf/plant-multi>) [41].

# Multiple sequence alignment and phylogenetic tree analysis of OPT genes in pumpkin

The OPT protein sequences were downloaded from the genomes of *Arabidopsis thaliana*, *Oryza sativa*, *Cucurbita moschata* and *Cucurbita maxima*, respectively. The protein sequences of these species were compared by ClustalW2 (<http://www.genome.jp/tools-bin/CLUSTORW>) software, and the phylogenetic analysis of the OPT proteins was constructed using the neighbor-joining method in MEGA 7.0 (1000 Bootstrap repeats) [42].

## Structural analysis and chromosome mapping of OPT genes in pumpkin

The exon/intron organizations of the OPT genes were analyzed using the Gene structure display server 2.0 (GSDS2.0; <http://gsds.cbi.pku.edu.cn/>). The CDS of each gene was compared with its corresponding genomic DNA sequence, [43]. The structural motif of OPTs protein sequence was annotated using online software Multiple EM for Motif Elicitation (MEME) (<http://meme-suite.org/tools/meme>) [44], and the maximum value of motif site was set to 15. The results of the exon/intron and motifs were visualized by the TBtools software (<https://github.com/CJ-Chen/TBtools>) [45].

## Gene duplication and evaluation of Ka, Ks and Ka/Ks in pumpkin

To identify gene duplication events, all CDS sequences of wheat OPT genes were subjected to BLAST searches against each other (identity > 80%, e value <  $1e^{-10}$ ) by using the local Blast program. Gene alignment coverage was then acquired by pair-wise alignment using the previously calculated method: Gene alignment coverage = (alignment length - mismatch length) / the length of larger genes. When the gene alignment coverage is more than 0.75, it is considered to be a duplicate gene pair. In addition, in the 100 kb region, two genes separated by five or fewer genes are considered as tandemly duplicated genes [46]. The values of Ka/ Ks between Ka and Ks and between paired genes were calculated by DnaSP software (<http://www.ub.edu/dnasp/>) [47]. For the timing of duplication events, the formula:  $T = Ks/2\lambda \times 10^{-6}$  Mya was used to calculate divergence time (T) in millions of years (Mya), where  $\lambda = 6.5 \times 10^{-9}$  represented the rate of replacement of each locus per herb plant year [48]. All the OPTs gene location information was obtained from the genome database of *Cucurbita moschata* and *Cucurbita maxima*. The chromosome distribution of these genes and tandem repeats in *Cucurbita moschata* and *Cucurbita maxima* were visualized using the Advanced Circos tool (<https://github.com/CJ-Chen/TBtools>) [45].

### Collinearity of OPT genes between pumpkin and other families

The collinearity among *Arabidopsis thaliana* (17 OPTs), *Cucurbita moschata* (21 OPTs) and *Cucurbita maxima* (24 OPTs) was analyzed and visualized with Dual Synteny Plotter software (<https://github.com/CJ-Chen/TBtools>) [45].

### Analysis of cis-acting elements of OPT genes promoter in Pumpkin

To analyze the promoter elements of OPT genes in *Cucurbita moschata* and *Cucurbita maxima*, we downloaded the sequence of the transcriptional initiation site (ATG) pre-2000 bp of these genes from the pumpkin database. The cis-acting elements of these genes were predicted by online tool PLANTCARE (<https://bioinformatics.psb.ugent.be/webtools/plantcare/html/>) [49] and visualized by TBtools (<https://github.com/CJ-Chen/TBtools>) [45].

### Tissue expression of OPT genes in pumpkin and its response to salt stress

To analyze the expression patterns of CmoOPTs and CmaOPTs in different organs (root, stem, leaf, fruit), the transcriptome data (BioProject: PRJNA385310) of different organs of *Cucurbita moschata* and *Cucurbita maxima* were downloaded [27]. To determine the response of *Cucurbita moschata* and *Cucurbita maxima* OPTs gene family to salt stress, we analyzed two different Cucurbita cultivars, "Rifu" (*Cucurbita moschata*) and "Rimu" (*Cucurbita maxima*). The transcriptome data in 2018 (BioProject: PRJNA464060) was excavated and analyzed the transcription profiles of OPTs in the leaf veins and leaf mesophylls under salt stress.

All the published transcriptome data were represented by RPKM (Reads per kilobase of exon model per million mapped reads), which has been converted to  $\log_2$  (RPKM) when plotting heat map. The heat map was visualized using the TBtools Heatmap tools (<https://github.com/CJ-Chen/TBtools>) [45].

### Experimental materials and stress treatment

To further analyze the response of OPTs in *Cucurbita moschata* to salt stress, the experiments were performed with "Baimi9" as the research material. They were provided by the pumpkin team of School of Horticulture and Landscape Architecture, Henan Institute of Science and Technology. The seeds were sown in a tray containing a matrix-meteorite (3:1) mixture and grown in a plant growth chamber. The artificial growth conditions were set as light intensity of 350  $\mu\text{mol}/\text{m}^2/\text{sec}$ , 65% relative humidity and 25 °C, 16 h light / 16 °C, 8 h dark. The two-month-old seedlings were cultured in Hoagland solution, pH 6.5. After 5 days of adaptation, some of the seedlings were cultured with 75 mM NaCl. Leaf veins and leaf mesophylls were collected at 12 h after the NaCl treatment. The control plants (CK) refers to the seedlings cultured normally without NaCl treatment. Each treatment has three independent biological replications. The Control and salt-treated samples were frozen in liquid nitrogen and stored at -70 °C for further analysis.

### Quantitative real-time PCR (qRT-PCR) analysis

Total RNA was removed with RNase-free DNase I (Takara, Tokyo, Japan) to avoid genomic DNA contamination. First-strand cDNAs were synthesized using the PrimeScript RT Reagent Kit with gDNA Eraser (Takara, Dalian, China) according to the manufacturer's protocol. The qRT-PCR assays were performed with the Primer Script RT Reagent Kit (Takara, Dalian, China). The specific primer of OPTs gene in pumpkin was designed on Primer-BLAST program (<https://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi>) (Table S6). The *Cucurbita moschata*  $\beta$ -Actin was used as the internal reference gene. Data were analyzed with Opticon monitor software (Bio-Rad). All primers for qRT-PCR were designed using Primer 6.0 software and primer sequences are listed in Table S3. The PCR conditions were as follows: 95 °C for 10 s and 40 cycles of 95 °C for 5 s and 60 °C for 30 s. The qRT-PCRs were performed using an ABI Step One

Plus. All the experiments were performed with three biological replicates. One sample constitutes a mixture of three seeds. The relative expression was calculated using the  $2^{-\Delta\Delta Ct}$  method [50].

Table 1  
Properties and locations of the predicted OPT proteins in *Cucurbita moschata* and *Cucurbita maxima*.

Species	Number	Gene name	Gene ID	Chr	Chr_Start	Chr_End	CDS Length	AA Length	Number of TM domains	PI	Molecular Weight Mw/Da
<i>Cucurbita moschata</i>	1	<i>CmoOPT1</i>	<i>CmoCh01G015170.1</i>	Cmo_Ch01	11712931	11718482	2232	743	14	8.38	83528.95
	2	<i>CmoOPT2</i>	<i>CmoCh02G009460.1</i>	Cmo_Ch02	5781968	5790763	4083	1360	14	5.36	152528.2
	3	<i>CmoOPT3</i>	<i>CmoCh03G007040.1</i>	Cmo_Ch03	6112323	6121326	3456	1151	14	6.49	127699.5
	4	<i>CmoOPT4</i>	<i>CmoCh09G005910.1</i>	Cmo_Ch09	2907414	2911576	2172	723	14	8.67	81089.38
	5	<i>CmoOPT5</i>	<i>CmoCh09G010240.1</i>	Cmo_Ch09	5472652	5477279	2250	749	14	9.06	84219.54
	6	<i>CmoOPT6</i>	<i>CmoCh13G003970.1</i>	Cmo_Ch13	4991452	4995450	2241	746	16	9.29	84109.87
	7	<i>CmoOPT7</i>	<i>CmoCh13G003990.1</i>	Cmo_Ch13	5034029	5039719	1980	659	12	8.88	73977.45
	8	<i>CmoOPT8</i>	<i>CmoCh16G003020.1</i>	Cmo_Ch16	1387655	1394389	2202	733	16	8.93	82620.95
	9	<i>CmoOPT9</i>	<i>CmoCh16G008560.1</i>	Cmo_Ch16	4545098	4547862	2193	730	16	9.33	82279.91
	10	<i>CmoOPT10</i>	<i>CmoCh16G008790.1</i>	Cmo_Ch16	4725582	4731270	3738	1245	26	9.16	140372.0
	11	<i>CmoOPT11</i>	<i>CmoCh16G009920.1</i>	Cmo_Ch16	6844868	6848814	2139	712	14	9.18	79796.11
	12	<i>CmoOPT12</i>	<i>CmoCh18G005810.1</i>	Cmo_Ch18	5481596	5485783	2238	745	14	9.20	83858.34
	13	<i>CmoOPT13</i>	<i>CmoCh18G013470.1</i>	Cmo_Ch18	12742452	12745655	2202	733	14	9.15	82221.74
	14	<i>CmoYSL1</i>	<i>CmoCh01G002660.1</i>	Cmo_Ch01	1210950	1218355	3183	1060	13	8.74	117738.5
	15	<i>CmoYSL2</i>	<i>CmoCh02G009500.1</i>	Cmo_Ch02	5802855	5810249	3153	1050	13	9.07	115361.4
	16	<i>CmoYSL3</i>	<i>CmoCh04G031180.1</i>	Cmo_Ch04	21643174	21647543	2109	702	14	9.15	77484.77
	17	<i>CmoYSL4</i>	<i>CmoCh11G016030.1</i>	Cmo_Ch11	11477985	11483729	2028	675	12	5.66	73801.34
	18	<i>CmoYSL5</i>	<i>CmoCh13G009930.1</i>	Cmo_Ch13	8495819	8500285	2082	693	12	9.30	77128.39
	19	<i>CmoYSL6</i>	<i>CmoCh16G000330.1</i>	Cmo_Ch16	155618	158339	2010	669	14	8.52	72899.67
	20	<i>CmoYSL7</i>	<i>CmoCh18G000780.1</i>	Cmo_Ch18	522094	527184	2043	680	12	8.96	75432.47
	21	<i>CmoYSL8</i>	<i>CmoCh18G013650.1</i>	Cmo_Ch18	12815949	12820091	2013	670	14	8.64	73480.46
<i>Cucurbita maxima</i>	22	<i>CmaOPT1</i>	<i>CmaCh01G014630.1</i>	Cma_Ch01	10293617	10299333	2232	743	14	8.08	83648.92

Species	Number	Gene name	Gene ID	Chr	Chr_Start	Chr_End	CDS Length	AA Length	Number of TM domains	PI	Molecular Weight Mw/Da
	23	<i>CmaOPT2</i>	<i>CmaCh02G009360.1</i>	Cma_Chr02	5566860	5570919	2196	731	14	6.70	82014.62
	24	<i>CmaOPT3</i>	<i>CmaCh03G006790.1</i>	Cma_Chr03	5623861	5627489	2265	754	15	8.25	84248.23
	25	<i>CmaOPT4</i>	<i>CmaCh09G006040.1</i>	Cma_Chr09	2824917	2841073	4200	1399	10	7.81	155646.51
	26	<i>CmaOPT5</i>	<i>CmaCh09G010210.1</i>	Cma_Chr09	5245061	5249312	2583	860	10	9.27	96985.30
	27	<i>CmaOPT6</i>	<i>CmaCh13G003660.1</i>	Cma_Chr13	4190289	4193336	2241	746	16	9.36	83849.63
	28	<i>CmaOPT7</i>	<i>CmaCh13G003670.1</i>	Cma_Chr13	4194142	4197493	1788	595	12	9.53	67549.72
	29	<i>CmaOPT8</i>	<i>CmaCh13G003680.1</i>	Cma_Chr13	4209423	4212306	2193	730	16	9.38	82266.00
	30	<i>CmaOPT9</i>	<i>CmaCh13G003690.1</i>	Cma_Chr13	4242559	4250972	2124	707	13	9.30	79619.74
	31	<i>CmaOPT10</i>	<i>CmaCh13G003700.1</i>	Cma_Chr13	4274274	4277157	2193	730	16	9.37	82265.95
	32	<i>CmaOPT11</i>	<i>CmaCh16G002820.1</i>	Cma_Chr16	1329545	1335746	2202	733	16	8.86	82544.80
	33	<i>CmaOPT12</i>	<i>CmaCh16G007990.1</i>	Cma_Chr16	4422653	4424298	1230	409	7	9.03	46182.06
	35	<i>CmaOPT13</i>	<i>CmaCh16G009530.1</i>	Cma_Chr16	7388795	7392172	2235	744	16	9.08	83087.38
	36	<i>CmaOPT14</i>	<i>CmaCh18G006110.1</i>	Cma_Chr18	5410583	5414858	2238	745	14	9.09	83751.19
	37	<i>CmaOPT15</i>	<i>CmaCh18G013170.1</i>	Cma_Chr18	10108928	10112175	2202	733	14	9.22	82361.89
	38	<i>CmaYSL1</i>	<i>CmaCh01G002550.1</i>	Cma_Chr01	1182517	1185564	2058	685	14	8.99	75294.37
	39	<i>CmaYSL2</i>	<i>CmaCh02G009410.1</i>	Cma_Chr02	5580942	5584935	1971	656	13	9.21	71851.87
	40	<i>CmaYSL3</i>	<i>CmaCh04G029970.1</i>	Cma_Chr04	19609442	19613723	2109	702	14	8.98	77455.51
	41	<i>CmaYSL4</i>	<i>CmaCh11G015190.1</i>	Cma_Chr11	9929680	9935313	2028	675	12	6.13	73665.28
	42	<i>CmaYSL5</i>	<i>CmaCh13G009590.1</i>	Cma_Chr13	7510416	7514832	2082	693	12	9.26	77207.42
	43	<i>CmaYSL6</i>	<i>CmaCh16G000280.1</i>	Cma_Chr16	134110	136820	2010	669	14	8.61	72868.54
	44	<i>CmaYSL7</i>	<i>CmaCh18G001010.1</i>	Cma_Chr18	476546	481516	1950	649	12	8.87	71856.36
	45	<i>CmaYSL8</i>	<i>CmaCh18G013340.1</i>	Cma_Chr18	10176428	10179894	2013	670	14	8.64	73564.52

## Declarations

## Ethics approval and consent to participate

The *Cucurbita moschata* “Baimi9” seeds used in this study were cultivated by Prof. Xinzheneg Li from the School of Horticulture and Landscape Architecture, Henan Institute of Science and Technology, Xinxiang, China.

## Consent for publication

Not applicable.

## Availability of data and materials

All raw transcriptome data have been deposited in NCBI's BioProjects (PRJNA385310 and PRJNA464060).

## Competing interests

The authors declare that they have no competing interests.

## Funding

This work was funded by the Scientific Research Foundation for High-level Talent (2017034/103010620001/015). These funding sources provided support for the data collection and manuscript writing.

## Authors' contributions

CWS and JPY conceived, designed and supervised the experiments; CWS and JPY wrote the manuscript; CWS performed the experiments; CWS and JPY provided support to the lab experiment and data analyses; CWS and JPY analysed the data. All the authors read and approved the manuscript.

## Acknowledgements

We are indebted to Prof. Xinzheneg Li for the support provided with the experimental seed materials.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Author details

<sup>1</sup>School of Resources and Environmental Sciences, Henan Institute of Science and Technology, Xinxiang 453003, China; <sup>2</sup>School of Horticulture and Landscape Architecture, Henan Institute of Science and Technology, Xinxiang, China.

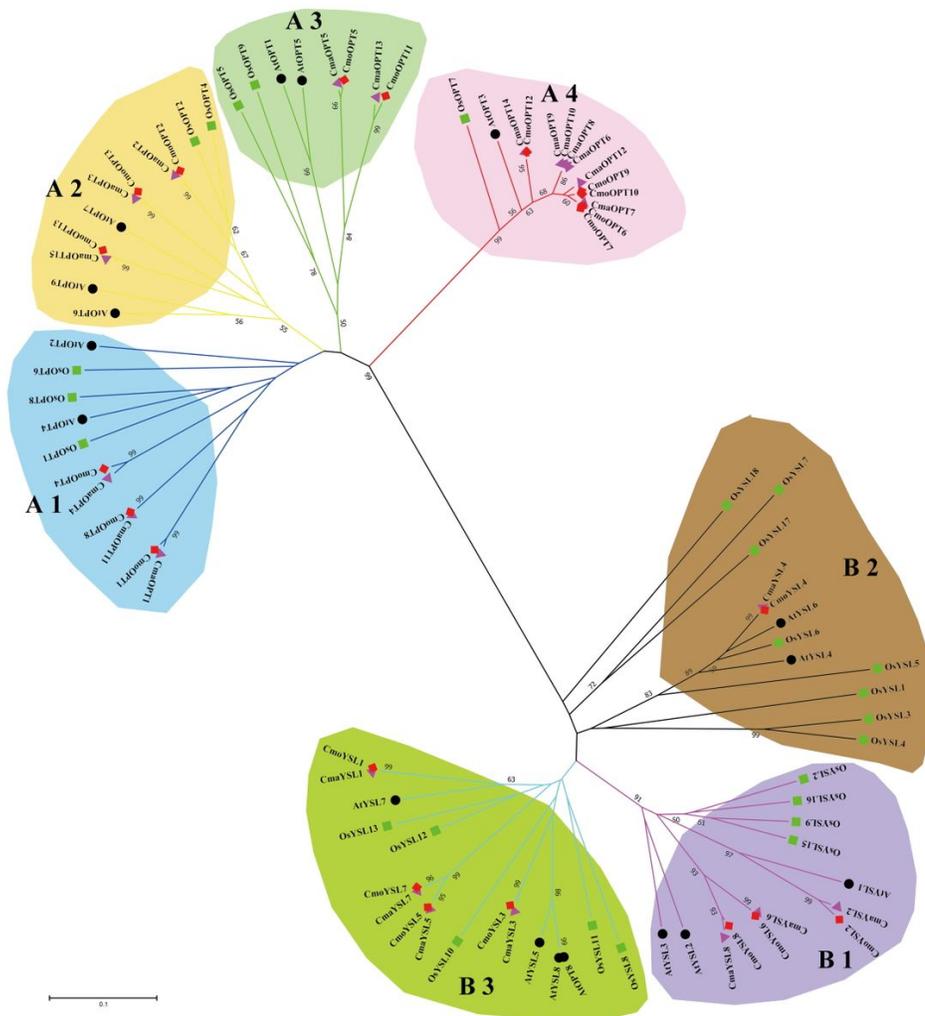
## References

1. Xu H, Huang X, Zhong T, Chen Z, Yu J. **Chinese land policies and farmers' adoption of organic fertilizer for saline soils.** *Land Use Policy.* 2014;**38(38):541–549.**
2. Mengliang N, Junjun X, Chen C, Haishun C, Jingyu S, Qiusheng K, Sergey S, Lana S, Yuan H, Zhilong B. **Salt tolerance mechanisms in pumpkin species (*Cucurbita*).** *J. Exp. Bot.* 2018;**69(20).**
3. Lubkowitz MA, Hauser L, Breslav M, Naider F, Becker JM. **An oligopeptide transport gene from *Candida albicans*.** *Microbiology.* 1997;**143(2):387–96.**
4. Koh S, Wiles AM, Sharp JS, Naider F, Becker JM, Stacey G. An oligopeptide transporter gene family in Arabidopsis. *Plant Physiol.* 2002;**128(1):21–9.**
5. Bogs J, Bourbonloux A, Cagnac O, Wachter A, Rausch T, Delrot S. **Functional characterization and expression analysis of a glutathione transporter, BJGT1, from *Brassica juncea*: evidence for regulation by heavy metal exposure.** *Plant Cell Environ.* 2003;**26(10):1703–11.**
6. Vasconcelos MW, Li GW, Lubkowitz M, Grusak MA. Characterization of the PT clade of oligopeptide transporters in rice. *Plant Genome.* 2008;**1(2):77–88.**
7. Mendozacoatl DG, Xie Q, Akmakjian GZ, Jobe TO, Patel A, Stacey MG, Song L, Demoin DW, Jurisson SS, Stacey G. OPT3 is a component of the iron-signaling network between leaves and roots and misregulation of OPT3 leads to an over-accumulation of cadmium in seeds. *Mol Plant.* 2014;**7(9):1455–69.**
8. Zhang Z, Xie Q, Jobe TO, Kau AR, Wang C, Li Y, Qiu B, Wang Q, Mendozacoatl DG, Schroeder JI. Identification of AtOPT4 as a plant glutathione transporter. *Mol Plant.* 2016;**9(3):481–4.**
9. Wongkaew A, Asayama K, Kitaiwa T, Nakamura S, Kojima K, Stacey G, Sekimoto H, Yokoyama T, Ohkamaohtsu N. **AtOPT6 protein functions in long-distance transport of glutathione in *Arabidopsis thaliana*.** *Plant Cell Physiol.* 2018;**59(7):1443–51.**
10. Lubkowitz M. The Oligopeptide Transporters: A small gene family with a diverse group of substrates and functions? *Mol. Plant.* 2011;**4(3):407–15.**
11. Curie C, Panaviene Z, Loulergue C, Dellaporta SL, Briat J, Walker EL. **Maize yellow stripe1 encodes a membrane protein directly involved in Fe(III) uptake.** *Nature.* 2001;**409(6818):346–349.**
12. Roberts LA, Pierson AJ, Panaviene Z, Walker EL. Yellow Stripe1. Yellow stripe1. expanded roles for the maize iron-phytosiderophore transporter. *Plant Physiol.* 2004;**135(1):112–20.**
13. Stacey MG, Osawa H, Patel A, Gassmann W, Stacey G. Expression analyses of Arabidopsis oligopeptide transporters during seed germination, vegetative growth and reproduction. *Planta.* 2006;**223(2):291–305.**

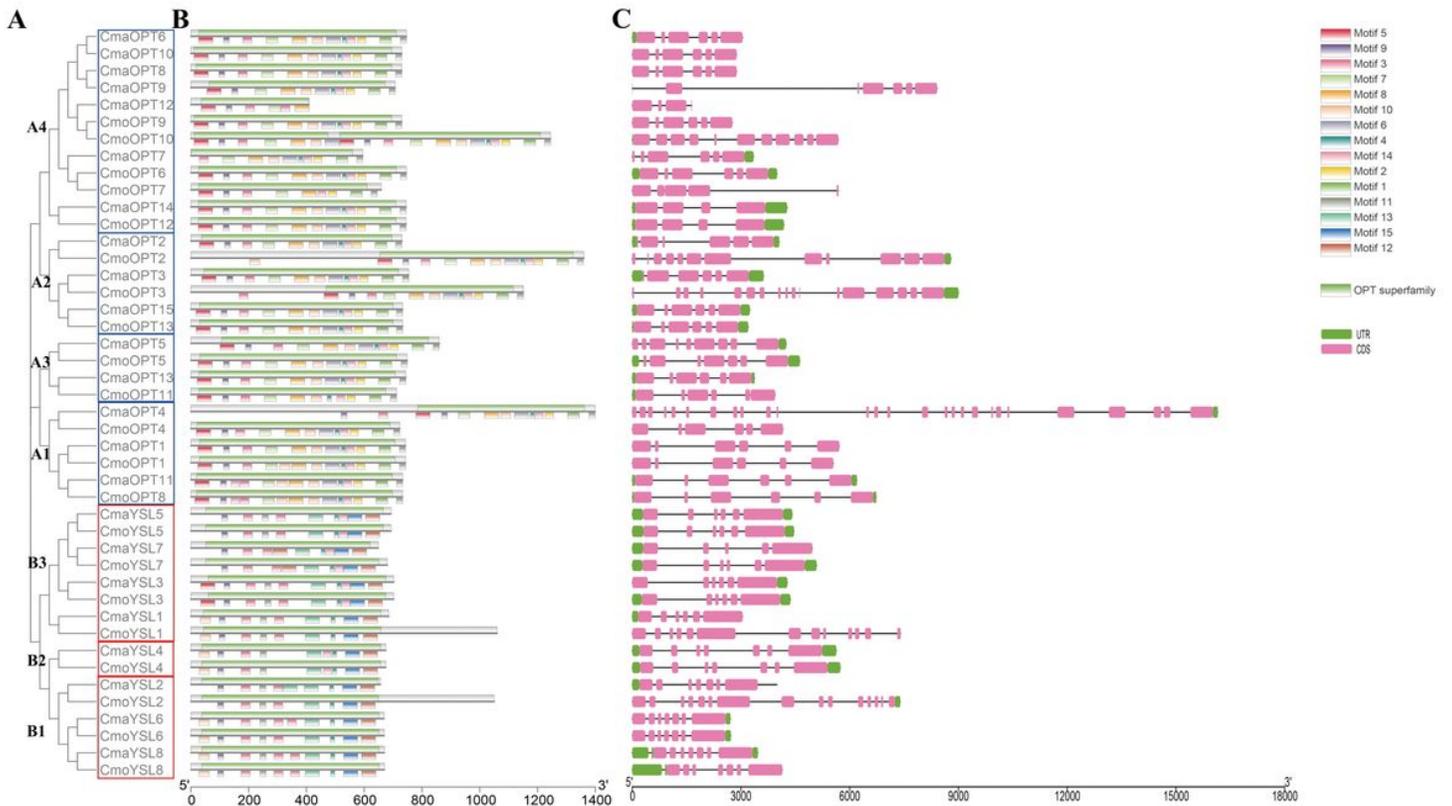
14. Osawa H, Stacey G, Gassmann W. ScOPT1 and AtOPT4 function as proton-coupled oligopeptide transporters with broad but distinct substrate specificities. *Biochem J.* 2006;393(1):267–75.
15. Wintz H, Fox TC, Wu Y, Feng V, Chen W, Chang H, Zhu T, Vulpe CD. **Expression profiles of *Arabidopsis thaliana* in mineral deficiencies reveal novel transporters involved in metal homeostasis.** *J Biol Chem.* 2003;278(48):47644–53.
16. Cagnac O, Bourbonloux A, Chakrabarty D, Zhang M, Delrot S. AtOPT6 transports glutathione derivatives and is induced by primisulfuron. *Plant Physiol.* 2004;135(3):1378–87.
17. Pike S, Patel A, Stacey G, Gassmann W. Arabidopsis OPT6 is an oligopeptide transporter with exceptionally broad substrate specificity. *Plant Cell Physiol.* 2009;50(11):1923–32.
18. Jean ML, Schikora A, Mari S, Briat J, Curie C. A loss-of-function mutation in AtYSL1 reveals its role in iron and nicotianamine seed loading. *Plant J.* 2005;44(5):769–82.
19. Murata Y, Ma JF, Yamaji N, Ueno D, Nomoto K, Iwashita T. **A specific transporter for iron(III)–phytosiderophore in barley roots.** *Plant J.* 2006;46(4):563–572.
20. Gendre D, Czernic P, Conéjéro G, Pianelli K, Mari S. **TcYSL3, a member of the YSL gene family from the hyper-accumulator *Thlaspi caerulescens*, encodes a nicotianamine-Ni/Fe transporter.** *Plant J.* 2010;49(1):1–15.
21. Aoyama T, Kobayashi T, Takahashi M, Nagasaka S, Usuda K, Kakei Y, Ishimaru Y, Nakanishi H, Mori S, Nishizawa NK. OsYSL18 is a rice iron(III)–deoxymugineic acid transporter specifically expressed in reproductive organs and phloem of lamina joints. *Plant Mol Biol.* 2009;70(6):681–92.
22. Hu YT, Ming F, Chen WW, Yan JY, Xu ZY, Li GX, Xu CY, Yang JL, Zheng SJ. **TcOPT3, a member of oligopeptide transporters from the hyperaccumulator *Thlaspi caerulescens*, is a novel Fe/Zn/Cd/Cu transporter.** *PLoS ONE.* 2012;7(6).
23. Xiang Q, Shen K, Yu X, Zhao K, Gu Y, Zhang X, Chen X, Chen Q. **Analysis of the oligopeptide transporter gene family in *Ganoderma lucidum*: structure, phylogeny, and expression patterns.** *Genome.* 2017;60(4):293–302.
24. Su H, Chu Y, Bai J, Gong L, Huang J, Xu W, Zhang J, Qiu X, Xu J, Huang Z. **Genome-wide identification and comparative analysis for OPT family genes in *Panax ginseng* and eleven flowering plants.** *Molecules.* 2018;24(1):15.
25. Pu Y, Yang D, Yin X, Wang Q, Chen Q, Yang Y, Yang Y. **Genome-wide analysis indicates diverse physiological roles of the turnip (*Brassica rapa* var. *rapa*) oligopeptide transporters gene family.** *Plant Diversity.* 2018;40(2):57–67.
26. Kumar A, Kaur G, Goel P, Bhati KK, Kaur M, Shukla V, Pandey AK. Genome-wide analysis of oligopeptide transporters and detailed characterization of yellow stripe transporter genes in hexaploid wheat. *Funct Integr Genomic.* 2019;19(1):75–90.
27. Sun H, Wu S, Zhang G, Jiao C, Guo S, Ren Y, Zhang J, Zhang H, Gong G, Jia Z. Karyotype stability and unbiased fractionation in the paleo-allotetraploid cucurbita genomes. *Mol Plant.* 2017;10(10):1293–306.
28. Huang H, Yu T, Li J, Qu S, Wang M, Wu T, Zhong Y. Characterization of *Cucurbita maxima* fruit metabolomic profiling and transcriptome to reveal fruit quality and ripening gene expression patterns. *J Plant Biol.* 2019;62(3):203–16.
29. Guo WL, Chen BH, Chen XJ, Guo YY, Yang HL, Li X, Wang GY. Transcriptome profiling of pumpkin (*Cucurbita moschata* Duch.) leaves infected with powdery mildew. *PLoS ONE.* 2018;13(1).
30. Schranz ME, Lysak MA, Mitchell-Olds T. The ABC's of comparative genomics in the Brassicaceae: building blocks of crucifer genomes. *Trends Plant Sci.* 2006;11(11):535–42.
31. Liu T, Zeng J, Xia K, Fan T, Li Y, Wang Y, Xu X, Zhang M. **Evolutionary expansion and functional diversification of oligopeptide transporter gene family in rice.** *Rice.* 2012;5(1):12–12.
32. Stacey MG, Koh S, Stacey BG. AtOPT3, a member of the oligopeptide transporter family, is essential for embryo development in Arabidopsis. *Plant Cell.* 2002;14(11):2799–811.
33. *Saeediazar S, Najafi-Zarrini H, Ranjbar GA, Heidari P. Identification and study of cis regulatory elements and phylogenetic relationship of TaSRG and other salt response genes. J. Biodiversity Environ. Sci. (JBES) 2014;5:1–5.*
34. Buyuk I, Inal B, Ilhan E, Tanriseven M, Aras S, Erayman M. Genome-wide identification of salinity responsive HSP70s in common bean. *Mol Biol Rep.* 2016;43(11):1251–66.
35. Shabala S, Pottosin I. Regulation of potassium transport in plants under hostile conditions: implications for abiotic and biotic stress tolerance. *Physiol Plant.* 2014;151(3):257–79.
36. Shapira O, Khadka S, Israeli Y, Shani U, Schwartz A. **Functional anatomy controls ion distribution in banana leaves: significance of Na<sup>+</sup> seclusion at the leaf margins.** *Plant Cell Environ.* 2009;32(5):476–485.
37. Munns R, Gilliham M. Salinity tolerance of crops - what is the cost? *New Phytol.* 2015;208(3):668–73.
38. Lubkowitz M. The OPT family functions in long-distance peptide and metal transport in plants. *Genet Eng.* 2006;27:35.
39. Krogh A, Larsson B. **Von Heijne G, Sonnhammer ELL. Predicting transmembrane protein topology with a hidden markov model: application to complete genomes.** *J Mol Biol.* 2001;305(3):567–80.
40. Wilkins MR, Gasteiger E, Bairoch AM, Sanchez JE, Williams KL, Appel RD, Hochstrasser DF. Protein identification and analysis tools in the Expasy server. *Methods Mol Biol.* 1999;112:531–52.
41. Fang Z, Jiang W, He Y, Ma D, Liu Y, Wang S, Zhang Y, Yin J. Genome-wide identification, structure characterization, and expression profiling of dof transcription factor gene family in wheat (*Triticum aestivum* L.). *Agronomy.* 2020;10(2):294.

42. Zhu Y, Yang L, Liu N, Yang J, Zhou X, Xia Y, He Y, He Y, Gong H, Ma D. Genome-wide identification, structure characterization, and expression pattern profiling of aquaporin gene family in cucumber. *BMC Plant Biol.* 2019;19(1):345–5.
43. Hu B, Jin J, Guo A, Zhang H, Luo J, Gao G. **GSDS 2.0: an upgraded gene feature visualization server.** *Bioinformatics.* 2015;31(8):1296–1297.
44. Bailey TL, Williams N, Misleh C, Li WW. MEME: discovering and analyzing DNA and protein sequence motifs. *Nucleic Acids Res.* 2006;34:369–73.
45. Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, Xia R. TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Mol Plant.* 2020;13(8):1194–202.
46. Wang L, Guo K, Li Y, Tu Y, Hu H, Wang B, Cui X, Peng L. Expression profiling and integrative analysis of the CESA/CSL superfamily in rice. *BMC Plant Biol.* 2010;10(1):282–2.
47. Yang Z, Nielsen R. Estimating synonymous and nonsynonymous substitution rates under realistic evolutionary models. *Mol Biol Evol.* 2000;17(1):32–43.
48. Yang Z, Gu S, Wang X, Li W, Tang Z, Xu C. Molecular evolution of the CPP-like gene family in plants: insights from comparative genomics of Arabidopsis and rice. *J Mol Evol.* 2008;67(3):266–77.
49. Lescot M, Dehais P, Thijs G, Marchal K, Moreau Y, De Peer YV, 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. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the  $2^{-\Delta\Delta CT}$  method. *Methods.* 2001;25(4):402–408.

## Figures

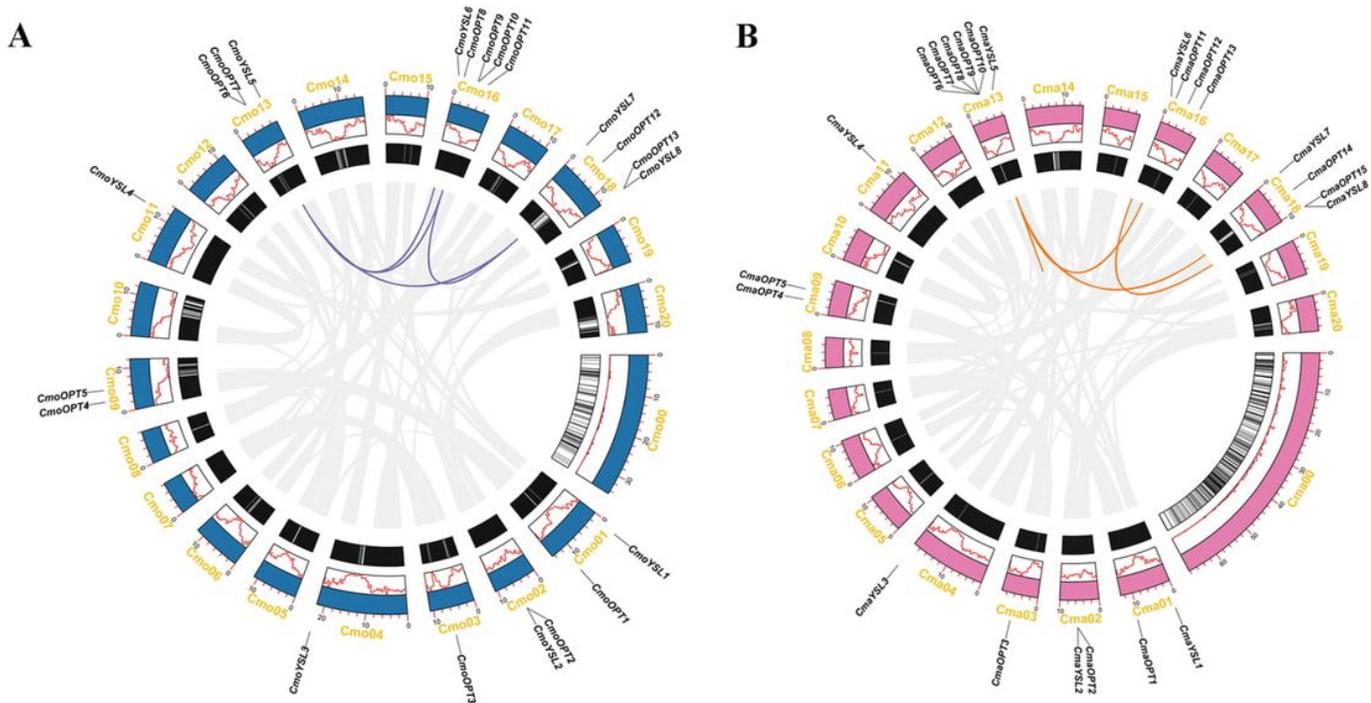


**Figure 1**  
 Phylogenetic tree of OPT family proteins among *Arabidopsis thaliana*, *Oryza sativa*, *Cucurbita moschata* and *Cucurbita maxima*. The OPT family proteins were classified into two OPT clade (type A) and YSL clade (type B), and the OPT clade was subdivided into A1-A4, YSL clade was subdivided into B1-B3. At, *Arabidopsis thaliana*; Os, *Oryza sativa*; Cmo, *Cucurbita moschata*; Cma, *Cucurbita maxima*.



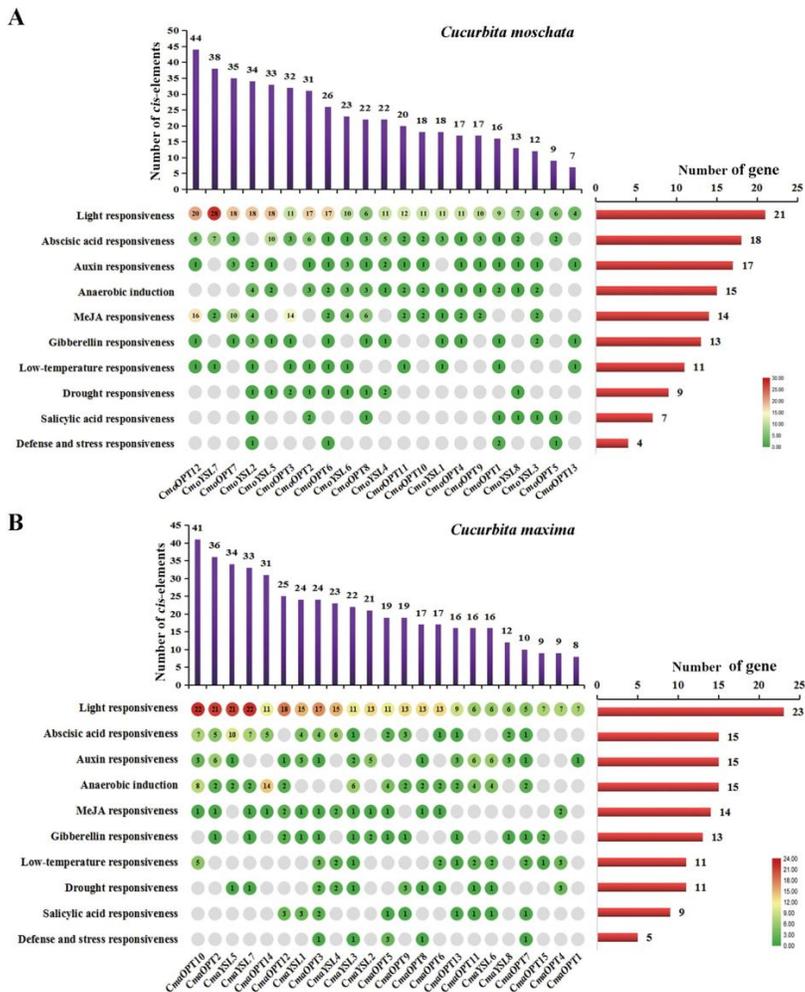
**Figure 2**

Exon-intron structures and motif compositions of OPT proteins in *Cucurbita moschata* and *Cucurbita maxima*. A, the phylogenetic tree was constructed using full length OPT protein sequences from the *Cucurbita moschata* and *Cucurbita maxima*. The 11 major groups designated from 1 to 11 are marked with different color backgrounds. B, A schematic representation of conserved motifs (obtained using MEME) in OPT proteins is displayed in the middle panel. C, Exon/intron structures of the OPT genes are shown in the right panel. Pink boxes represent exons and black lines represent introns. Different motifs are represented by different colored boxes. The sequence information for each motif is provided in Fig. S1.



**Figure 3**  
 Circular collinearity plots of OPT genes in *Cucurbita moschata* and *Cucurbita maxima*. A, Cmo00-20 represented chromosomes in *Cucurbita moschata*. Red lines linked the syntenic orthologous in *Cucurbita moschata*. B, Cma00-20 represented chromosomes in *Cucurbita maxima*. Blue lines linked the syntenic orthologous in *Cucurbita maxima*. The grey lines were as enomic collinearity background in *Cucurbita moschata* and *Cucurbita maxima* genomes.

**Figure 4**  
 Synteny analysis of OPT genes in *Cucurbita moschata*, *Cucurbita maxima* and *Arabidopsis thaliana*. Cmo00-20 represented chromosomes in *Cucurbita moschata*. Cma00-20 represented chromosomes in *Cucurbita maxima*. AtChr1-5 represented chromosomes in *Arabidopsis thaliana*. Red lines linked the duplicated gene pairs between *Cucurbita moschata* and *Cucurbita maxima*. Blue lines linked the duplicated gene pairs between *Cucurbita moschata* and *Arabidopsis thaliana*. Green lines linked the duplicated gene pairs between *Cucurbita maxima* and *Arabidopsis thaliana*. At, *Arabidopsis thaliana*; Cmo, *Cucurbita moschata*; Cma, *Cucurbita maxima*.



**Figure 5**

Cis-acting elements distribution and statistical analysis of OPT genes promoters. The numbers in circles depicts the quantity of the predicted cis-acting elements in the promoter region.

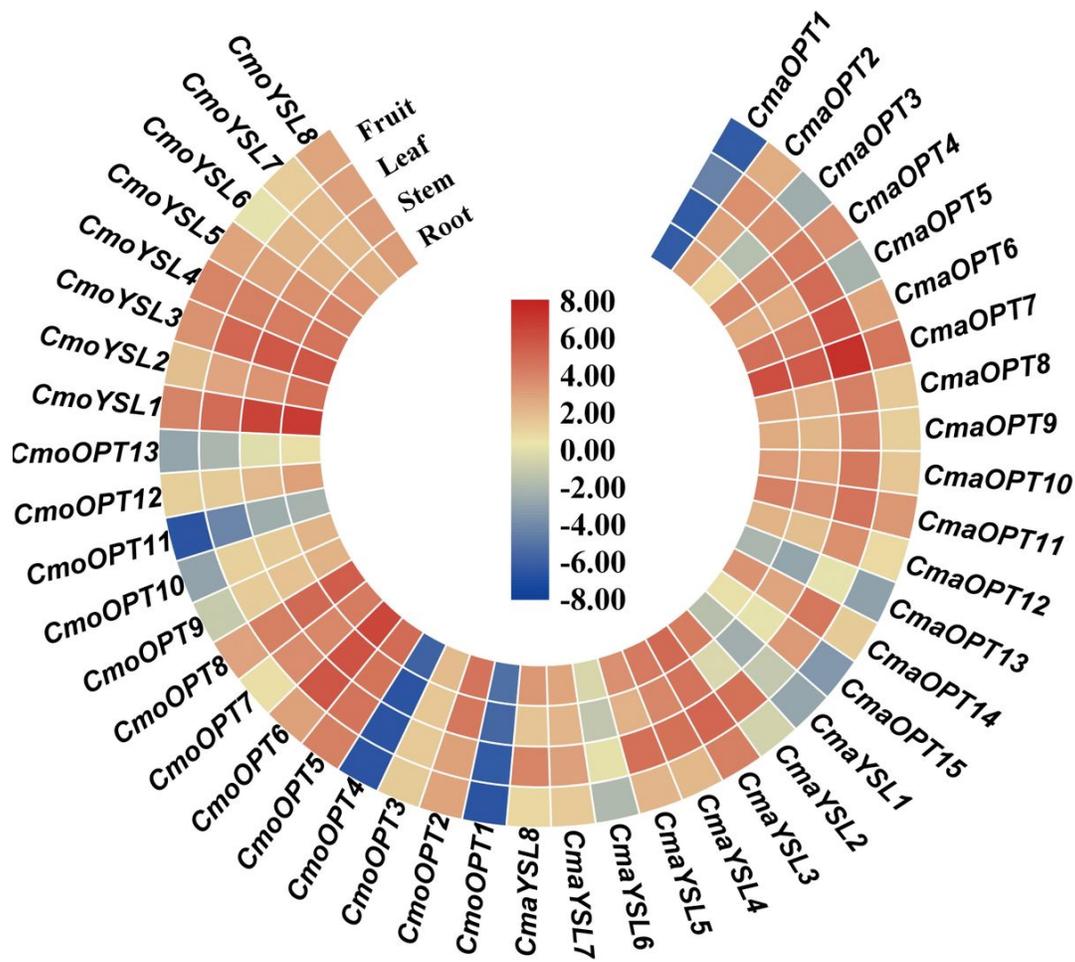
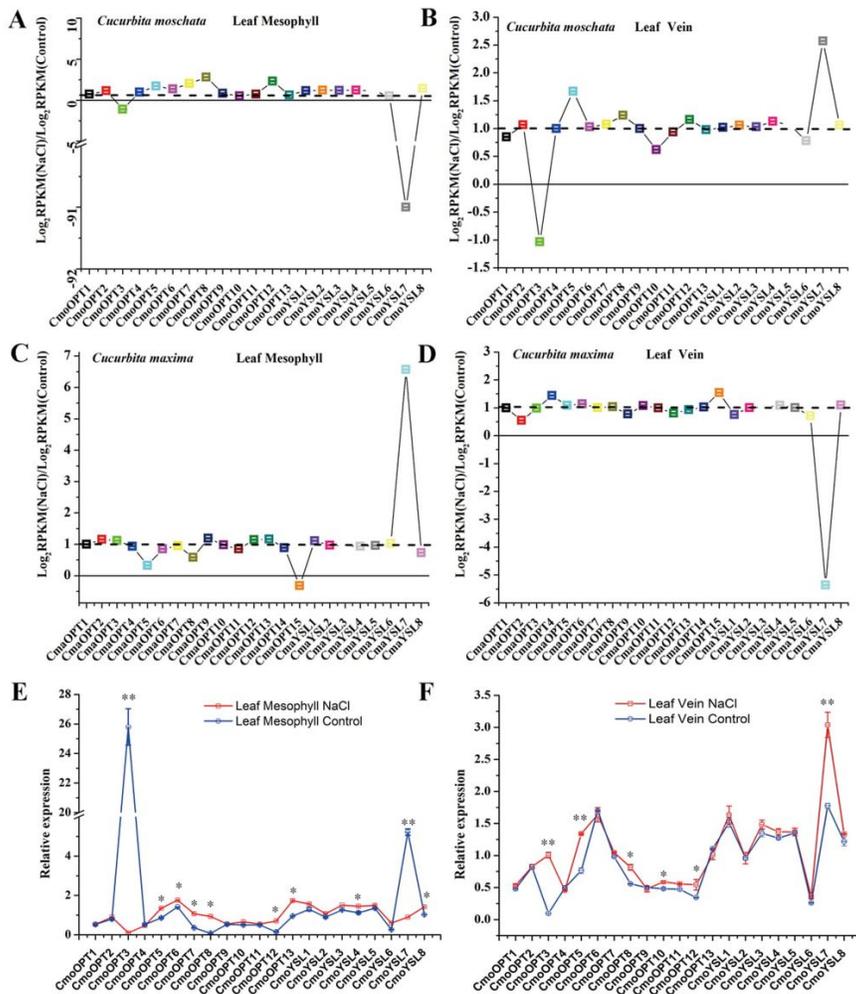


Figure 6

Expression profile of OPT genes at different tissues of *Cucurbita moschata* and *Cucurbita maxima*. Gene expression of OPT genes was calculated with transcriptome data (BioProject: PRJNA385310). All data of the heat map were standardized by Log2 (RPKM)



**Figure 7**

Expression profiles and qRT-PCR verification of OPTs in pumpkin vein and mesophyll under salt treatment. A-D, Expression levels of OPT genes were recalculated with transcriptome data (BioProject: PRJNA464060). All data were standardized by  $\text{Log}_2\text{RPKM}(\text{NaCl})/\text{RPKM}(\text{Control})$ . NaCl: NaCl treatment, Control: normal conditions. E-F, The data represented the relative expression of OPT genes at 12 h after NaCl treatment. Control referred to untreated plants (control plants) under normal conditions. The data was calculated via the  $2^{-\Delta\Delta\text{Ct}}$  method, and the reference gene ( $\beta$ -Actin) was used to correct the expression level of target genes. Errors bars represent the standard errors of three biological replicates, and asterisks indicated a significant difference in expression level between NaCl and Control treatments (\* $P < 0.05$ , \*\* $P < 0.01$ )

## Supplementary Files

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

- [Fig.S3.pdf](#)
- [Fig.S2.pdf](#)
- [Fig.S1.pdf](#)
- [TableS6.xlsx](#)
- [TableS5.xlsx](#)
- [TableS4.xlsx](#)
- [TableS3.xlsx](#)
- [TableS2.xlsx](#)
- [TableS1.xlsx](#)