The structure, function and expression analysis of the nodulin 26-like intrinsic protein subfamily of plant aquaporins in tomato

Yuxiang Zhang, Shihong Fei, Yunmin Xu, Yong He, Zhujun Zhu* & Yuanyuan Liu*

The nodulin 26-like intrinsic protein (NIP) family belonging to a group of aquaporin proteins is unique to plants. NIPs have a wide of transport activities and are involved in developmental processes and stress tolerance. The well reported Lsi1 and Lsi6 belonging to NIP III were characterized as Si transporters. However, except Lsi1 and Lsi6, most NIPs remain unknown. Here, we identified 43 putative aquaporins in tomato. We found there are 12 NIPs, including 8 NIP I proteins, 3 NIP II proteins, and 1 NIP III protein among the 43 aquaporins. Also, there are two Si efflux transporters SlLsi2-1 and SlLsi2-2 identified by using Lsi2 proteins from other species. By analysing the phylogenetic relationships, conserved residues and expression patterns, we propose that three NIP I members (SlNIP-2, SlNIP-3 and SlNIP-11) may transport water, ammonia, urea, and boric acid, and contribute to pollen development. Three NIP II proteins (SlNIP-7, SlNIP-9 and SlNIP-12) may be boric acid facilitators, and affect plant growth and anther development. Overall, the study provides valuable candidates of Si transporters and other NIP proteins to further explore their roles in uptake and transport for silicon, boron, and other substrates in tomato.

Silicon (Si) is a common element, and its abundance in the earth's crust is second only to oxygen, reaching 28.8%1. Si is an essential element for animals and is closely related to animal life activities2. Although Si is not a vital element for plants3, a large number of studies have shown that Si has beneficial effects on plant growth and development, especially under stressed conditions, such as drought, salt, heavy metal, high temperature, radiation damage, freezing, nutrient imbalance4-6. In plants, Si has been proposed as a quasi-essential element, and the International Plant Nutrition Institute (IPNI) defined Si as a “beneficial substance” in 2015. The main forms of Si existing in soil are silica (SiO₂) and silicate, and only a small part exists in the form of silicic acid Si(OH)₄, which is the only form of Si that can be absorbed by plants7.

Although Si is found in most plants7, the content of Si varies greatly among different plants8. According to the differences in Si absorption modes, plants were classified into active, passive and rejective Si absorption plants9. In addition, plants are divided into three types: high Si accumulators, intermediate Si accumulators and low Si accumulators8. Monocots such as rice and wheat belong to high Si accumulation plants and active Si absorption plants9. In contrast, their Si content can reach 1.0–10.0% of dry weight. Most of the eudicots, such as cucumbers, belong to intermediary Si accumulation plants and passive Si absorption plants, and the Si content in their bodies is about 0.2–1.0% of dry weight. Some of the eudicots, such as tomato, belong to low Si accumulation plants and Si rejected plants, and their Si content is only less than 0.2% of dry weight8.

The absorption of Si by plants is achieved mainly by two steps. Firstly, silicic acid is absorbed by roots, and then silicic acid is transported to aboveground by transpiration pull10. Among them, although transpiration pull can promote the transfer of Si in plants from roots to shoots, studies have proved that transpiration pull is not the reason for the difference of Si accumulation in different plants, but the uptake ability of root plays a major role in Si uptake and transport in plants11. The uptake process of Si is mediated by specific transmembrane proteins, Low silicon rice 1 (Lsi1), identified as Si transporters. The first identified Si transporter in higher plants is OsLsi1.

Collaborative Innovation Center for Efficient and Green Production of Agriculture in Mountainous Areas of Zhejiang Province, College of Horticulture Science, Zhejiang A&F University, Hangzhou 311300, Zhejiang, China. *email: zhuzj@zafu.edu.cn; yyliu@zafu.edu.cn
in rice\(^{13}\). OsLsi1 is located in the distal side of the plasma membrane of the exodermis and endodermis cells. Its function is to facilitate Si uptake into cells of exodermis from soil solution, and promote Si uptake into cells of endodermis from aerenchyma\(^{13}\). Shortly after the discovery of OsLsi1, a second Si transporter gene, OsLsi2, was uncovered\(^{14}\). OsLsi2 is located in the proximal side of the plasma membrane of the exodermis and endodermis cells. Its function is to facilitate Si efflux into aerenchyma from exodermis cells and facilitate Si efflux into stele from endodermis cells\(^{14}\). Obviously, OsLsi1 is an influx transporter, while OsLsi2 is an efflux transporter\(^{13,14}\).

A third Si transporter identified as influx transporter is OsLsi6 in rice\(^{15}\). OsLsi6 is primarily located in the parenchyma cells of xylem in the leaf sheaths and leaves, which is responsible for xylem unloading and further distribution of Si in aerial tissues\(^{15,16}\).

It has been reported that Lsi1 and Lsi6 belong to the Nodulin 26-like intrinsic proteins (NIPs) subfamily of the aquaporin (AQP) family\(^{13}\), while Lsi2 falls into the ion transporter superfamily\(^{14}\). AQPs are water channel proteins, also called major intrinsic proteins (MIPs) that facilitate the transport of water and a variety of solutes, such as boric acid, NH\(_4\)+, glycerol or urea\(^{17}\). Based on the sequence similarity and the localization, plant AQPs can be divided into five groups. Except NIPs, the other four groups include the plasma membrane intrinsic proteins (PIPs), the tonoplast intrinsic proteins (TIPs), the small basic intrinsic proteins (SIPs), and the X intrinsic proteins (XIPs)\(^{18-21}\). Among these AQPs, NIPs are unique to plants. Furthermore, NIPs can be classified into three subgroups—NIP I, NIP II and NIP III—with respect to the composition of the aromatic/arginine (ar/R) selectivity filter which is formed by four residues from helices 2 (H2), H5, loop E (LE1) and LE2\(^{22,23}\). The size and chemical properties of the residues in ar/R selectivity filter may determine the size and chemical properties of substrates NIPs can transport. NIP I can transport water, glycerol\(^{24}\) and lactic acid\(^{25}\); NIP II can transport boron\(^{26}\) and urea\(^{27}\), whereas, NIP III have transport activities for silicon, arsenic and selenium\(^{22}\). Indeed, unlike AQP1 belonging to NIP I which contains two bulky amino acid residues in the H2/H5, OsLsi1 belonging to NIP II contains two small residues has a broader pore diameter of the channel\(^{29}\). Therefore, OsLsi1 can transport relatively larger molecules of silicic acid (4.38 Å), while the substitution of larger Phe and His for smaller Gly and Ser at H2 and H5 constrains the AQP1 pore to 2.8 Å permitting the flow of water (diameter 2.4 Å) and the exclusion of bulkier solutes\(^{23,24}\).

Si is extremely low in tomato plant. At present, there is still a lack of research on the Si uptake and transport in tomato, and the reasons for its low Si accumulation are still controversial. One study first reported that SLLsi1 with 109- AA instead of 108- AA spacing between NPA domains showed loss of Si permeability, compared with the functional Lsl1 transporters in high Si accumulators\(^{29}\). The other study believed that SLLsi1 is normally functional, but loss-of-function of SLLsi2 cannot transfer silicic acid absorbed from roots to xylem, thus resulting in low Si accumulation in shoot\(^{30}\). Although tomato is a Si excluder, Si could benefit tomato a lot on many aspects, such as seed germination, shoot growth, the root/shoot ratio, osmotic adjustment, photosynthetic rate, hydraulic conductance, transpiration, antioxidative systems\(^{31}\). Evidences on Si transporters in tomato, even in eudicots, are relatively insufficient. The functions of other NIP proteins other than SLLsi1 in tomato are also still unknown. In the present study, SLLsi1 and SLLsi6 Si influx transporters and SLLsi2 Si efflux transporters from other species were set to blast tomato homologs. 43 aquaporin proteins homologous to Lsi1 and Lsi6, and 2 ion transporter proteins homologous to Lsi2 were identified in tomato. Detailed information, including their phylogenetic relationships, gene structures, conserved domains, subcellular localizations, tissue expression profiles, and cis-acting elements were analyzed. The studies will provide a solid basis for further functional characterization of tomato Si transporters as well as other NIP members, and exploring molecular mechanism of their uptake and transport for silicon, boron, and other substrates in tomato.

**Results**

**Identification of Lsi1/Lsi2/Lsi6 homologues of aquaporin and ion subfamily in tomato.** Tomato Si transporters were identified from the tomato genome database (SGN, https://solgenomics.net) by a BLAST search using the full-length protein sequences of reported Si transporters Lsi1/Lsi2/Lsi6 from rice (Oryza sativa L.) and cucumber (Cucumis sativus L.). The amino acid sequences of Lsi1/Lsi2/Lsi6 from rice and cucumber were downloaded from NCBI database (Supplementary Table S1). A total of 43 tomato aquaporin proteins and 2 ion transporter proteins were obtained after deletion of repetitive sequences.

**Phylogenetic analysis and classification of tomato Lsi1/Lsi2/Lsi6 homologues.** To reveal the phylogenetic relationships of 45 tomato homologues, the amino acid sequences of reported 17 Lsi1/Lsi2/Lsi6 Si transporters from other species were included (Supplementary Table S1) for the construction of phylogenetic tree. These 62 proteins can be classified into five clades (I–V, Fig. 1). Notably, each clade contains tomato homologues, six are assigned to Clade I. Two are classified into Clade II that includes reported Lsi2 Si transporters. Twelve are members of Clade III, together with the well reported Lsi1/Lsi6 Si transporters. Eleven are divided into Clade IV, and fourteen belong to Clade V. Actually, members in Clade I, Clade III, Clade IV and Clade V correspond to members in XIPs, NIPs, TIPs, and PIPs described in current literature\(^{32}\). Twelve tomato homologues were grouped in Clade III with StLsi1 of potato, CsLsi1 and CsLsi6 of cucumber, ZmLsi1 and ZmLsi6 of maize, HvLsi1 and HvLsi6 of barley, OsLsi1 and OsLsi6 of rice, TaLsi1 of wheat, CmLsi1 of pumpkin, SbLsi1 of sorghum. It has been reported that OsLsi1\(^{13}\), OsLsi6\(^{16}\), HvLsi1\(^{30}\), ZmLsi1 and ZmLsi6\(^{16}\), CsLsi1\(^{22}\) belong to NIP III subfamily of aquaporins, suggesting the twelve tomato homologues in Clade III also belong to the NIP subgroup. We name them as SINIP-1 (Solyc03g013340), SINIP-2 (Solyc02g091240), SINIP-3 (Solyc05g008080), SINIP-4 (Solyc03g005980), SINIP-5 (Solyc02g071920), SINIP-6 (Solyc06g073590), SINIP-7 (Solyc08g013730), SINIP-8 (Solyc12g017500), SINIP-9 (Solyc02g117030), SINIP-10 (Solyc02g063310) and SINIP-12 (Solyc01g079890), respectively. Solyc03g025350 and Solyc06g036100 clustered in Clade II with OsLsi2 of rice, HvLsi2 of barley, ZmLsi2 of maize, CsLsi2 of cucumber and CmLsi2 of pumpkin. Based on pro-
tein sequence similarity and their relationship in the phylogenetic tree, the two orthologous Lsi2-like proteins of tomato in Clade II were named as SlLsi2-1 (Solyc06g036100) and SlLsi2-2 (Solyc03g025350) (Supplementary Table S2).

Lsi1 and Lsi6 in other species have been identified as Si influx transporters, and Lsi2 has been characterized as Si efflux transporters. In Lsi1/Lsi6 proteins-containing Clade III, the orthologous SINIP-1 was the most homologous to these Si influx transporters, and reasonably predicted to have Si influx transport activity. Similarly, the orthologous SlNIP-1 was the most homologous to these Si influx transporters, and reasonably predicted to have Si efflux transport activity for Si. However, in a previous study, no Si transport activity of SlLsi1 (designated SlNIP-1 in this study) was detected in Xenopus oocytes29. This report is not consistent with the other study which determined that SlLsi1 was a functional influx transporter30. SlLsi2-L1 (designated SlLsi2-1 in this study) and SlLsi2-L2 (designated SlLsi2-2 in this study) were reported no efflux transport activity for Si30. However, the function of aquaporins in tomato remain to be largely uncharacterized, except SINIP-1, members in Clade III that belong to NIP subgroup are still unknown.

The chromosomal localization of 45 Lsi1/Lsi2/Lsi6 homologous genes in tomato. The chromosomal localizations of 45 tomato Lsi1/Lsi2/Lsi6 homologues were determined to visualize their genomic position information. They were distributed on 10 of 12 chromosomes except for chromosome 4 and chromosome 7 (Fig. 2). Among them, eight genes were mapped to chromosome 3, 6 and 10; two on chromosome 5 and 9, three on chromosome 12, four on chromosome 1 and 8, five on chromosome 2; there was only one gene on chromosome 11 (Fig. 2). It should be noted that gene duplication may occur among these tomato Si transporter homologues, since the loci Solyc10g054790, Solyc10g054800, Solyc10g054810 to Solyc10g054820 are found next to each other on chromosome 10. Also the gene pairs Solyc09g071910/Solyc09g071920, Solyc09g007760/Solyc09g007770 were found on chromosome 2 and chromosome 9, respectively. Interestingly, Solyc10g054790,
Solyc10g054800, Solyc10g054810 and Solyc10g054820 were clustered together in Clade I, Solyc02g071910/ Solyc02g071920 were grouped together in Clade III, and Solyc09g007760/Solyc09g007770 were clustered together in Clade V (Fig. 1), which showed high similarity based on protein sequences, indicating they may share common functions with each other.

**Analysis of exon–intron structure.** The exon and intron information of 45 tomato Lsi1/Lsi2/Lsi6 homologues were searched in SGN database and sorted by GSDS. The color bar on the left side of the gene name represents the clades of these genes in the evolutionary tree. Exon–intron analysis showed that the size and the number of the exons are highly conserved within each clade, but significantly different among the clades. Most members of the 45 tomato Lsi1/Lsi2/Lsi6 homologues contained three to five exons, but a few contained less than three or more than five exons. Members of Clade I and Clade II are characterized by three exons. Most members of Clade III have five exons. The majority of the members of Clade IV features three exons, while most members of Clade V contain four exons (Supplementary Fig. S1A,B).

**Gene structure and motif composition analysis.** Motifs are highly conserved amino acid residues in homologous proteins, which may play important roles in the structure and function of active proteins. The conserved motifs of these 45 tomato Lsi1/Lsi2/Lsi6 homologues were analyzed by MEME. Ten conserved motifs were identified (Supplementary Fig. S1C). Among them, Motif 9 only exists in the members of Clade II and this clade does not contain any other motifs, which might contribute to the functional divergence of Lsi2 proteins in Clade II. It is reasonable Lsi2 was ion transporter protein, differing from NIP proteins. Motifs 1, 2 and 3 were found in Clade III that clustered by Lsi1/Lsi6 proteins, which were also shared in Clade I, IV and V. Interestingly, members of Clade I and V, and some members of Clade IV, but not any members of Clade III, also shared Motif 7. Motif 8 exists in all members of Clade V and some members of Clade I; in addition, members in Clade I solely contain Motifs 5, 6 and 10, and members in Clade V uniquely possess Motif 4 (Supplementary Fig. S1A,C). These observations suggest genes in Clade I, IV and V may have evolved from genes in Clade III, sharing common motifs (Motifs 1, 2 and 3) and similar functions with members in Clade III, and further gained some additional motifs (such as Motif 7) and corresponding functional diversification afterwards.

**Evolutionary analysis of NIP subfamily from tomato, Arabidopsis and rice.** The well-characterized Lsi1 channels belong to the NIP subfamily of aquaporins. In our study, we have found 12 members that in Clade III belong to the NIP subgroup (Fig. 1). However, except SINIP-1, the other members in NIP subgroup are still unknown. Therefore, we have made an in-depth study on these twelve NIP proteins. To well understand their role in substrate selectivity, we constructed phylogenetic analyses of these 12 NIPs with orthologues from Arabidopsis and rice (Fig. 3), the accession numbers of the corresponding proteins were listed in Supplementary Table S3. As phylogenetic tree showed that the Arabidopsis and rice NIPs were distributed in two (I and II) and three (I, II and III) subgroups, respectively, which is well consistent with previous reports. The 12 tomato NIPs fall into three NIP subgroups (Fig. 3). SINIP-1 is grouped with well-defined OsNIP2-1 (OsLsi1) and OsNIP2-2 (OsLsi6) that belong to NIP III subgroup. SINIP-7, SINIP-9 and SINIP-12 fall into NIP II sub-
group, and were tightly clustered together with AtNIP5-1, AtNIP6-1 and AtNIP7-1, respectively. Eight out of the twelve tomato NIPs belong to NIP I, including SlNIP-2, SlNIP-3, SlNIP-4, SlNIP-5, SlNIP-6, SlNIP-8, SlNIP-10 and SlNIP-11. Among them, SlNIP-2, SlNIP-3 and SlNIP-11 showed relatively closed to AtNIP4-1 and AtNIP4-2, whereas SlNIP-4, SlNIP-5, SlNIP-6, SlNBIP-8 and SlNIP-10 showed highly closed to AtNIP1-1, AtNIP1-2, AtNIP2-1, and AtNIP3-1 (Fig. 3).

### Prediction of physicochemical properties and identification of conserved domain.

It is well known that NIPs of aquaporin subfamily were characterized for six transmembrane domains (TMDs) connected by five loops (loop A—loop E)\(^3\), two highly conserved NPA (asparagine-proline-alanine) motifs on loop B and loop E, an ar/R selectivity filter and Froger's positions\(^3\). Sequence alignments of twelve tomato NIPs were performed with ClustalX2 (Fig. 4). The TMDs of each protein were predicted by TMHMM2.0 online tool, the positions of TMDs are marked with gray areas, and the numbers of TMDs of each protein are recorded in Table 1. All the identified Si transporters of Lsi1/Lsi6 in other species contain six TMDs, and most of the twelve tomato NIPs also contain six TMDs. However, there is an exception for SlNIP-10 and SlNIP-11, which possess four and three TMDs, respectively. The second TMD in SlNIP-10 is incomplete, while SlNIP-11 missed the fourth, fifth and sixth TMDs (Table 1).

For the members of aquaporin family, NPA domains will generate electrostatic repulsion to protons, and then form a channel, while ar/R selectivity filter and Froger’s positions will affect the specific substrate of the channel. All Si transporters identified in other species have two NPA domains, only with an exception for CsLsi1, which have NPV on Loop E with the alanine (A) in the third position replaced by threonine (T). However, NPA motif changed into NPV does not disturb Si transport activity of CsLsi1\(^3\). As for tomato, 11 of the 12 tomato NIPs all carried two NPA/S/V/T motifs, except for SlNIP-10 only having one NPA motif on Loop B. A specific length of 108 AA between the two NPA domains is a necessary and selectivity feature for Si among all Si-transporting plants\(^2\). These already validated Si-transporting Lsi1 proteins all presented 108 AA residues, such as CsLsi1, HvLsi1, TaLsi1, SbLsi1, CmLsi1, ZmLsi1 and OsLsi1. Whereas most of the tomato NIPs have 109 residues in the spacing, and only SlNIP-7 and SlNIP-9 have 108 AA. Moreover, SlNIP-8 has 112 AA and SlNIP-10 has 115 AA between the two NPA domains. However, the spacing itself is unlikely to be required for Si permeability. Supporting evidences showed that SlLsi1/SlNIP-1 possesses 109 AA in the spacing and shows transport activity for Si\(^1\).

For most members of NIP III subgroup from various species, including SlNIP-1 from tomato, four amino acids from helix 2 (H2), helix 5 (H5), and loop E (LE1 and LE2) that constitute ar/R selectivity filter are GSGR, except for CsLsi6, where a small G in the H2 is substituted with the bulkier C residue (Table 1). However, the residue at the H2 position is not critical for Si transport. For example, when glycine (G) was substituted by alanine

![Figure 3. Evolutionary analysis of twelve, nine, twelve NIP proteins from tomato, Arabidopsis and rice. The polygenetic tree was constructed by MEGAX using maximum likelihood method, and was modified by iTOL. Members in NIP I, II and III were shown by red, blue and yellow colors, respectively.](https://www.nature.com/sdata/12022-9180-f3/fig3.png)
At H2 of OsLsi1, the transport activities for Si were unaffected. Compared to residue at H2, residue at H5 is required for Si transport activity. When serine (S) at the H5 position was substituted by isoleucine (I), the transport activity of OsLsi1 for Si was totally lost. In addition, when both residues at the H2 and H5 were substituted, Si transport activities of OsLsi1 as well as transport activities for other solutes are completely lost. Most of tomato NIPs belong to NIP I subgroup possessing an ar/R filter consisting of W, V, A, and R, except SlNIP-6 with an ar/R selectivity filter of WIA, and SlNIP-10 and SlNIP-11 whose ar/R selectivity filters are incomplete. For SlNIP-7, SlNIP-9, and SlNIP-12, which belong to the NIP II subgroup, their ar/R selectivity filters are SIAR, TIAR, and AVGR, respectively. It is proposed that SlNIP-6, SlNIP-7, and SlNIP-9 with I at H5, and SlNIP-10 and SlNIP-11 with incomplete ar/R selectivity filters might have no transport activities for Si. It is worth testing the transport activities of SlNIP-1, SlNIP-2, SlNIP-3, SlNIP-4, SlNIP-5, SlNIP-8 with ar/R selectivity filter consisting of WVAR and SlNIP-12 with ar/R selectivity filter consisting of AVGR.

Figurine 4. Identification of conserved domains of Si transporter Lsi1/6 homologues (SlNIP-1 to SlNIP-12) in tomato and Si transporters Lsi1/6 reported in other plants. Protein sequence alignment of functionally known Si transporters Lsi1/6 in other species and the newly identified homologues in tomato. The gray region represents the transmembrane (TM) domains, the red region represents the conserved NPA domains, the yellow region represents the ar/R selectivity filter sites, and the blue region represents the Froger’s residue sites.
and was preferentially expressed in anthesis flowers, while SlLsi2-2
SlNIP-3
SlNIP-6
Script abundance.
SlNIP-8
all tested tissues, and four genes (SlNIP-5, SlNIP-7
highly in anthesis flowers.
examined tissues, especially in root, indicating it may function in root.
SlNIP-7
SlNIP-2
and was highly expressed in all
10 days post anthesis fruit (10 DPA), 20 days post anthesis fruit (20 DPA) and ripening fruit (33 DPA) (Fig. 5).
ocotyledons, young leaves, mature leaves, vegetative meristems, young flower buds, flowers at anthesis (0 DPA),
FUNCTIOPNAL GENOMICS DATABASE in 11 various tissues of tomato including whole root, hypocotyl,
SlNIP-3, we firstly investigated their expression atlas based on the RNA-seq data derived from the TOMATO
Tissue-specific expression analysis. To explore the functions of tomato NIP genes as well as the two
SILSI2 genes, we firstly investigated their expression atlas based on the RNA-seq data derived from the TOMATO
FUNCTIONNAL GENOMICS DATABASE in 11 various tissues of tomato including whole root, hypocotyl,
cotyledons, young leaves, mature leaves, vegetative meristems, young flower buds, flowers at anthesis (0 DPA),
10 days post anthesis fruit (10 DPA), 20 days post anthesis fruit (20 DPA) and ripening fruit (33 DPA) (Fig. 5).
SlNIP-7 and SlNIP-2 exhibited high and organ-specific expression pattern. SlNIP-7 was highly expressed in all
examined tissues, especially in root, indicating it may function in root. SlNIP-2 was preferentially expressed
highly in anthesis flowers. SlNIP-5, SlNIP-9 and SILSII-1 were constitutively expressed at a high level in nearly
all tested tissues, and four genes (SlNIP-8, SlNIP-10, SlNIP-11 and SlNIP-12) were expressed with very low transcript abundance. SlNIP-3 was preferentially expressed in anthesis flowers, while SlNIP-6 and SILSII-2 was spe-
SlNIP-3, SlNIP-7 all have ABA responsive elements, while eight tomato Si transporter homologues except SlNIP-4, SlNIP-2, these genes are possibly sensitive to plant hormones. Eleven tomato Si transporter homologues except SlNIP-4, SlNIP-2, cis-acting elements of this group in the promoters of 14 tomato Si transporter homologues, indicating that these genes may be less induced by signaling from the specific growth element (CAT-box), seed-specific regulation element (RY-element) and mesophyll cell differentiation element (HD-Zip1). There are only 0–3 cis-acting elements of this group in promoters of all 14 tomato Si transporter homologues, which indicates that these genes may function in many aspects of tomato growth and development.

**Promoter analysis.** Specific cis-acting elements can combine with particular transcription factors to regulate downstream gene transcription and expression. In order to determine the promoter cis-acting elements of 14 tomato Si transporter homologous proteins, the 2000 bp sequences upstream of the start site (ATG) of these genes were downloaded as putative promoters from SGN database. The promoter sequences were analyzed by using PlantCARE online tool, and the identified cis-acting elements were sorted into three groups related to plant growth and development, hormone response and biotic/abiotic stresses (Fig. 7). Cis-acting elements related to plant growth and development included endosperm expression element (GCN4_motif), circadian rhythm regulation element (circadian), zein metabolism regulation element (O$_2$-site), meristem expression element (CAT-box), seed-specific regulation element (RY-element) and mesophyll cell differentiation element (HD-Zip1). There are only 0–3 cis-acting elements of this group in promoters of all 14 tomato Si transporter homologous genes, which indicates that these genes may be less induced by signaling from the specific growth and development stage of plants. Cis-acting elements belonging to hormone response included salicylic acid responsive element (TCA-element), abscisic acid (ABA) responsive element, methyl jasmonate responsive element (CGTCA-motif, TGACG-motif), ethylene responsive element (ERE), auxin responsive element (TGA-element, AuxRR-core) and gibberellin responsive element (TATC-box, GARE-motif, P-Box), which has 3–14 cis-acting elements of this group in the promoters of 14 tomato Si transporter homologues, indicating that these genes are possibly sensitive to plant hormones. Eleven tomato Si transporter homologues except SlNIP-2, SlNIP-4 and SlNIP-7 all have ABA responsive elements, while eight tomato Si transporter homologues except SlNIP-3, SlNIP-6, SlNIP-9, SlNIP-11, SlNIP-12 and SlLsi2-1 all have ethylene responsive element (ERE). ABA and
Figure 6. Expression patterns of homologous genes of Si transporters in tomato in different tissues were examined using qRT-PCR. Tomato ACTIN2 gene was used as an internal standard. Values are means ± SD (n = 3). Gray represents the roots of tomato; yellow represents the stems of tomato; blue represents the leaves of tomato; red represents the flowers of tomato; green represents the green mature fruits of tomato; pink represents the red mature fruits of tomato.

Figure 7. Prediction of cis-acting elements in promoter regions of Si transporter homologues in tomato. (A) The different colors and numbers of the grid indicated the numbers of rech predicted elements in these Si transporter homologues promoter region. (B) The different colored histogram represented the sum of the cis-acting elements in each category.
ethylene play vital roles in plant growth and development as well as in mediating plant response to a wide range of stresses. It should be noted that tomato Si transporter homologues may be regulated by ABA or ethylene and participate ABA or ethylene-mediated plant development and stress responsiveness. Cis-acting elements related to biotic/abiotic stress included drought response element (MBS), anaerobic induction response element (ARE), defense and stress response element (TC-rich repeats), low temperature response element (LTR) and mechanical injury response element (WUN-motif). There are 1–7 cis-acting elements of this group in the promoters of 14 tomato Si transporter homologues. Among them, three WUN-motifs, MBS, TC-rich repeats were found in \( \text{SlNIP-4} \), \( \text{SlLsi2-2} \), \( \text{SlNIP-7} \) and \( \text{SlNIP-9} \) promoters, respectively. The cis-regulatory elements identified in the promoters of tomato Si transporter homologues suggested that they may play an important role in hormone signaling and various stress response.

Subcellular localization of \( \text{SlNIP-1} \) and its mutated proteins. Plasma membrane location is required for the uptake of Si by \( \text{Lsi1} \) transporters. Functional \( \text{Lsi1} \) Si influx transporters are unexceptionally localized in the plasma membrane, such as \( \text{OsLsi133} \), \( \text{HvLsi133} \), \( \text{ZmLsi134} \), \( \text{TaLsi122} \), \( \text{SlLsi129} \), \( \text{CsLsi135} \), and \( \text{NsLsi141} \). Subcellular localization can largely affect the Si transport activities. \( \text{CmLsi1} \) from the bloomless rootstock that was localized at the plasma membrane have transport activity, whereas the one from the bloomless rootstock was localized at the endoplasmic reticulum and had no transport activity for Si. In order to verify the subcellular localization of tomato \( \text{SlNIP-1} \), we cloned the 849 bp CDS sequence of \( \text{SlNIP-1} \), and introduced it into the vector \( \text{pGWB5} \) to obtain the GFP fusion vector \( \text{35S::SlNIP-1-GFP} \). The vector was transformed into \( \text{Agrobacterium tumefaciens} \) strain \( \text{GV3101} \) to generate transgenic \( \text{Arabidopsis} \) plants by floral dip method. \( \text{T1} \) transgenic \( \text{Arabidopsis} \) plants were used for observing GFP fluorescence signal. Consistent with WoLF’s prediction in Table 1 and previous report\(^{29}\), \( \text{SlNIP-1} \) protein was localized in plasma membrane (Fig. 8). A 108-AA spacing between NPA domains is essential to Si influx transport activity for \( \text{SlLsi129} \). Predicted by PROVEAN, we found five site deletions (140 V, 141 T, 142 K, 143 N and 144 V) result in neutral effects. Then, the localization of various \( \text{SlNIP-1} \) mutants of single amino acid deletion was confirmed by transgenic plants. As the results showed that, like \( \text{SlNIP-1} \), these five mutated \( \text{SlNIP-1} \) proteins (\( \text{SlNIP-1Δ140V} \), \( \text{SlNIP-1Δ141T} \), \( \text{SlNIP-1Δ142K} \), \( \text{SlNIP-1Δ143N} \), and \( \text{SlNIP-1Δ144V} \)) with a spacing of 108 AA were also located on the plasma membrane (Fig. 8). In 2015, Deshmukh et al. had reported that deletion 140 V in \( \text{SlNIP-1} \) exhibited Si permeability relative to the native \( \text{SlNIP-1} \) protein\(^{29}\). Thus, the function of the other four \( \text{SlNIP-1} \) mutants need to be tested.

**Discussion**

Tomato Si transporter homologues \( \text{SlNIP-1} \) to \( \text{SlNIP-12} \) belong to NIP subfamily (Figs. 1, 3). Six of the twelve proteins with ar/R filter consisting of Trp (H2), Val (H5), Ala (LE1), and Arg (LE2), WVAR, as well as \( \text{SlNIP-6} \) with WIAAR, belong to the NIP I subgroup, whereas \( \text{SlNIP-7} \) with SIAR, \( \text{SlNIP-9} \) with TIAR, and \( \text{SlNIP-12} \) with AVGR were classified into NIP II group\(^{42,43}\). Additionally, \( \text{SlNIP-1} \) with GSGR belong to NIP III group\(^{28}\). Compared to NIP III subgroups, which have small amino acid residues Gly in H2 and Ser in H5; members of NIP I subgroups have two bulky residues Trp and Val in H2 and H5 positions, respectively\(^{29}\). So, considering the composition of ar/R filter, we propose that most members of tomato NIPs possessing WVAR (WIAAR in \( \text{SlNIP-6} \)) do not have transport activities for silicic acid, but may transport small molecules of water, glycerol...
and lactic acid as members in NIP I. To support the propose, one study showed that replacement of AIGR by WIGR endows AtNIP6;1 with water transport activity, suggesting the possible water-transport activity as well as glycerol transport of SINP-6 which was presented by WIAR of the ar/R region.

The Arabidopsis NIP II subgroup has three member genes: AtNIP5;1, AtNIP6;1, and AtNIP7;123, and each of them has an ortholog in tomato, represented by SINP-7, SINP-9 and SINP-12, respectively.45 Current study further confirmed the remarkable orthologous relationships between AtNIP5;1, AtNIP6;1, AtNIP7;1 and SINP-7, SINP-9, SINP-12. Firstly, structural similarity is within pore determinant regions, such as SINP-7 and AtNIP5;1 share the a/R filter of SIAR; SINP-9 and AtNIP6;1 share the a/R filter of TIAR; SINP-12 together with AtNIP7;1 share the a/R filter of AVGR (Table 1). Secondly, similar tissue expression profiles, RT-qPCR combined with RNA-seq data showed that SINP-7 was mainly expressed in the roots, SINP-9 was constitutively expressed in all examined tissues with high expression level in stem, and SINP-12 was specifically expressed in flower (Fig. 5), which were similar to previous studies.43-45 Thirdly, similar to previous observations with AtNIP5;146, AtNIP6;146 and AtNIP7;145, SINP-7 and SINP-9 were predicted to be localized on plasma membrane, and SINP-9 was predicted to be localized on vacuole membrane (Table 1). As AtNIP5;1, AtNIP6;1, and AtNIP7;1 were reported to mediate boron uptake and contribute to the development of root, rosette leaves, inflorescences and pollen grain.27,43-45 In summary, the results suggested SINP-7, SINP-9 and SINP-12 might act as major channel proteins mediating boron acid transport. It will be important to investigate the substrate specificity of the SINP-7, SINP-9 and SINP-12 and the possible involvement of SINP-7, SINP-9 and SINP-12 in plant growth and development.

In addition to ar/R selectivity filter, NPA motif and inter-NPA distance can also affect the Si permeability of Lsi1 channels. It has been reported that two NPA motifs localized at loop B and loop E form part surface of the narrow aqueous pore.46 Residues composed of NPA motifs influence the size of constriction filter, and then the specificity of substrates. For example, the pore diameter of the NPA regions of AtNIP6;1 is narrower than that of Nodulin 26, which may result from a larger Val substitution for the small Ala in the second NPA motif (loop E) of AtNIP6;1 compared with Nodulin 26. Most members of the tomato NIPs have two conserved NPA motifs, while the first NPA motif in SINP-4 and SINP-7 was changed to NPV. In addition, the second NPA motif in SINP-7 was also replaced by NPT. Moreover, the second motif in SINP-9 and SINP-10 was changed to NPT and NPT, respectively. However, it seems that the second NPA motif is not a crucial factor for the substrate specificity of NIP proteins. In the case of AtNIP6;1, a substitution of Ala for Val in the second NPA motif has little effect on the transport selectivity of AtNIP6;1.26.

An inter-NPA distance of 108- AA is a common feature that is critical for the function of Si transporters. SINP-1/SLSi1 owned an ar/R filter composing of GSGR belonging to the NIP III group and might be a Si transporter, but 109 AA in the inter-NPA region of SINP-1 may cause the loss function of Si transport activities. However, Si permeability of SLSi1 is still controversial.26,28 In addition to the precise space between the NPA domains, other molecular determinants can affect the function of Lsi1 transporters. NsLsi1 from tobacco possessing typical molecular signatures, a GSGR selectivity filter and a 108 AA inter-NPA region, were found to be Si-impermeable.41 Further investigation showed that Si transport activities of NsLsi1 can be compromised by P125F substitution. P125F substitution increased plasma membrane localization of NsLsi1P125F compared to NsLsi1WT, thus enhanced the transport capacity.41 It was also reported that analysis of boric acid permeability of SLSi1/SINP-1 did not match previous results.30 This discrepancy may reflect that SLSi2 was not expressed constantly but show high expression levels in flower, suggesting these genes may play specific roles in flower. Similarly, SINP-3 and SINP-10 were also mainly expressed in flower. Polygenetic tree indicated that SINP-2, SINP-3 and SINP-11 were highly homologous with AtNIP4;1 and AtNIP4;2. It has been reported that AtNIP4;1 and AtNIP4;2 were expressed in flower specific to pollen grains and pollen tubes, and characterized as important regulators for the proper pollen development, pollen germination, and pollen tube growth.45 Similar to AtNIP4;1 and AtNIP4;2, SINP-2, SINP-3 and SINP-11 may participate in the male gametophyte development.
Conclusions

There are 43 aquaporins including 8 NIP I proteins, 3 NIP II proteins, and 1 NIP III protein in tomato. SINIP-1 was the reported Si influx transporters SILsi1. Mutated SINIP-1 proteins of single amino acid deletion—SINIP-1Δ141T, SINIP-1Δ142K, SINIP-1Δ143N, and SINIP-1Δ144V—resembled the native SINIP-1 showing plasma membrane localization. Though polygenetic tree analysis, conserved structure analysis, and gene expression patterns analysis, we predict SINIP-7, SINIP-9, SINIP-12 may be boric acid facilitators, and SINIP-2, SINIP-3 and SINIP-11 may be involved in the pollen development. Considering the important biological functions of NIP subfamily of plant aquaporins, further analysis of the functions of these genes in specific tissue is deserved. In addition, we also propose that functional identification of SINIP-1, SILsi2-1 and SILsi2-2 should be obtained through transgenic plants and furthermore experimental studies.

Materials and methods

Screening and identification of Lsi1 and Lsi6 homologues and Lsi2 homologues in tomato. The amino acid sequences of Si transporters Lsi1, Lsi2 and Lsi6 of rice and cucumber were obtained from NCBI database (https://www.ncbi.nlm.nih.gov/). Sequences of tomato were searched by the BLAST program provided by tomato genome database SGN (https://solgenomics.net/), and then the sequences of retrieved homologous proteins with E value less than e−6 was selected. The amino acid sequences of barley (Hordeum vulgare L.), wheat (Triticum aestivum L.), corn (Zea mays L.), sorghum (Sorghum bicolor L.) and potato (Solanum tuberosum L.) were downloaded from NCBI database.

Phylogenetic tree construction and chromosome location. Sequence alignments of Si transporters of tomato and other species were carried out by ClustalX2 software (http://www.clustal.org/)49, and sequence alignment and evolutionary tree construction were further carried out by MEGA 7.0 software (https://www.megasoftware.net/)39. Neighbor-joining method was selected as the evolutionary tree construction method with 1000 bootstrap replications. Finally, the phylogenetic tree was illustrated using Interactive Tree of Life (IToL, http://itol.embl.de) and Notepad++ v7.8.8 software (https://notepad-plus-plus.org/downloads/v7.8.8/). The chromosomal locations of tomato Lsi1, Lsi2 and Lsi6 homologous genes were searched in tomato genome database SGN, and the distribution of gene location on chromosome was drawn by MapInspect 1.0 software (https://mapinspect.software.informer.com/).

Analysis of exon–intron structure and conserved motifs. The gene sequences of tomato Si transporter homologous proteins were searched from tomato genome database SGN, and the exon and intron of these genes were found on GSWS website (http://gsds.cbi.pku.edu.cn/); the motifs of these proteins was searched by MEME website (https://meme-suite.org/meme/). The method was set to Any Number of Repetitions, and the number of motifs searched was set to 10. The transmembrane domains were analyzed by TMHMM online tool (http://www.cbs.dtu.dk/services/TMHMM/) prediction; subcellular localizations were predicted by WoLF PSORT Tool (http://wolfsort.hgc.jp/). The basic physical and chemical properties of protein were analyzed by ProtParam online tool (http://web.expasy.org/protparam/); the positions of ar/R selectivity filter and Froger’s residues were marked after sequence alignments.

Evolutionary analysis by maximum likelihood method. The evolutionary history was inferred by using the Maximum Likelihood method and JTT matrix-based model. The tree with the highest log likelihood (~ 14,866.98) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the JTT model, and then selecting the topology with superior log likelihood value. This analysis involved 33 amino acid sequences. There were a total of 474 positions in the final dataset. Evolutionary analyses were conducted in MEGA 7.0.

Tissue-specific expression analysis. The expression patterns analysis of tomato Si transporter homologue genes in different tissues of different development stages of tomato was carried out according to RNA-seq data from Tomato Functional Genomics Database (TFGD, http://ted.bti.cornell.edu/cgi-bin/TFGD/digital/home.cgi). Transcriptome analysis of 11 tissues in wild species S. pimpinellifolium, LA1589 using Illumina RNA-seq. The tissues include: newly developed leaves around 5 mm long, mature green leaflets, flower buds 10 days before anthesis or younger, flowers at anthesis, fruit of 10 days post anthesis (DPA), 20 DPA fruit, and breaker stage ripening fruit. RPKM values are an average of 4 replicates. Gene expression data of RPKM values were log2 transformed before analysis, and heatmaps were generated by TBTools v1.0986988 software41.

Plant materials and growth conditions. Tomato (Solanum lycopersicum) cv. Ailsa Craig (AC) was used as the wild type in this study. Seeds of tomato cultivar AC (Accession LA2838A) were obtained from the Tomato Genetics Resources Center at the University of California, Davis (https://tgrc.ucdavis.edu/). Tomato seeds were placed on moistened filter paper for 48 h for germination. Tomato seedlings were transferred to growth chambers and maintained under a long-day photoperiod (16 h of light/8 h of dark) with a white light intensity of 200 mmol photons m−2 s−1 at 25 °C during the subjective day and at 18 °C during the subjective night.

Arabidopsis (Arabidopsis thaliana) ecotype Columbia was used as the wild type in this study. Arabidopsis plants were grown in Murashige and Skoog (MS) media at 22 °C under a long-day (16 h light/8 h dark) photoperiod (light intensity of 120 mM photons m−2 s−1).
The plant materials were well used for research project and comply with relevant institutional, national, and international guidelines and legislation.

**RNA extraction and quantitative RT-PCR (qRT-PCR) analysis.** TRLzol reagent (Invitrogen) was used for RNA extraction. For quantitative real-time polymerase chain reaction (qRT-PCR) analysis, one microgram of DNA-free RNA was transcribed into first strand cDNA by Prime-Script RT Master Mix (TaKaRa). The qRT-PCR was carried out with the UVP ChemStudio (analyticjena) using TB Green Premix Ex Taq (TaKaRa). The reaction conditions were 95 °C for 30 s, and 40 cycles at 95 °C for 5 s, 60 °C for 30 s. Expression levels of target genes were normalized relative to ACTIN2 gene. Primers used to quantify gene expression levels are listed in Supplementary Table S4. Each reaction was performed with three biological replicates.

**Cis-elements analysis.** Regions 2000 bp upstream of the start codon (ATG) of each gene were downloaded in the tomato genome database SGN as the predicted promoter sequences. Cis-acting elements of the promoters were predicted by PlantCare (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/).

**Plasmid construction and plant transformation.** To generate the 35S::SINIP-1-GFP construct, the coding sequence without termination codon of SINIP-1 was cloned into pENTR and recombined with the binary vector PGWB5 (35S promoter, C-GFP). A similar approach was used to generate the 35S::SINIP-1A140V-GFP (35S promoter, C-GFP), 35S::SINIP-1A141T-GFP, 35S::SINIP-1A142K-GFP, 35S::SINIP-1A143N-GFP, 35S::SINIP-1A144V-GFP constructs. All primers used for DNA construct generation are listed in Supplementary Table S4. The above constructs were then transformed into Agrobacterium tumefaciens strain GV3101, which was used for transformation of Arabidopsis plants via the floral dip method. T1 transgenic plants were selected based on their resistance to hygromycin.

**Microscopy observations.** For subcellular localization analysis, roots of 10-day-old T1 transgenic plants with corresponding constructs were observed and imaged under a laser-scanning confocal microscope (Olympus fluoview FV3000). For imaging GFP, the 488-nm lines of the laser were used for excitation, and emission was detected at 510 nm.

**Data availability**

The amino acid sequences analyzed in the current study are available in the Sol Genomics Network repository (https://solgenomics.net/), the NCBI database (https://www.ncbi.nlm.nih.gov/), the Arabidopsis Information Resource (TAIR) (http://www.arabidopsis.org/) and the rice genome annotation database (http://rice.uga.edu/), respectively. The accession numbers are included in the Supplementary Tables.
Author contributions
Z.J.Z. and Y.Y.L. designed the experiments. Y.X.Z. and S.H.F. performed the experiments. Y.Y.L. and S.H.F. analyzed the data. Y.Y.L. wrote the original draft of the manuscript. Y.M.X., Y.H., Z.J.Z. and Y.Y.L. revised article. Z.J.Z. and Y.Y.L. designed the experiments. Y.X.Z. and Y.Y.L. performed the experiments. Y.X.Z. and Y.Y.L.

Acknowledgements
The study was supported by the National Key Research and Development Program of China (2018YFD1000800), the Natural Science Foundation of Zhejiang Province (LZ20C150001), the National Natural Science Foundation of China (32102402) and the Scientific Research and Development Foundation of Zhejiang A & F University (2019FR045).

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-022-13195-0.
Open Access  This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2022