Introduction

Cardamine hupingshanensis, also known as Cardamine enshiensis, is a unique selenium hyperaccumulator in China that can accumulate more than 1000 mg kg\(^{-1}\) selenium \([1]\). \(C.\) hupingshanensis was first discovered in Yutangba, Enshi, Hubei Province, China, and Huping Mountain, Shimen, Hunan Province, China. Se hyperaccumulators growing in high-selenium environments for a long time have strong selenium tolerance, detoxification and enrichment abilities and have evolved unique molecular mechanisms. Therefore, Se hyperaccumulators have become an important resource for basic theoretical
research on selenium. A chromosome-level genome assembly was performed for *C. enshiensis*, which consists of 443.4 Mb in 16 chromosomes with a scaffold N50 of 24 Mb [2]. Hi-C analysis of chromatin interaction patterns was performed, and genes with compartmental changes after selenium treatment were involved in the metabolism of selenium compounds [2]. Zhou et al. identified the biological pathways and candidate genes of the selenium tolerance mechanism by transcriptomics [3]. Differential expression analysis identified 25 genes in four pathways that are significantly responsive to selenium in *C. hupingshanensis* seedlings [3], including the ATPS genes and research targets in this paper.

Selenium is an essential element in humans and animals that plays a vital role in human health [4, 5]. Long-term severe selenium deficiency is the main cause of Keshan disease and Kashin-Beck disease [6]. Selenium deficiency increases the risk of cancer complications, and appropriate selenium supplementation can help reduce oxidative stress, thereby reducing the incidence of cancer complications [7]. In addition, an appropriate concentration of selenium plays an insulin-like role, but when the concentration of selenium is too high, it will aggravate insulin resistance and lead to type II diabetes [8]. Selenium mediates redox signalling and affects oxidative stress, inflammation and lipid metabolism and plays a certain role in improving the immune level of the human body, alleviating heavy metal toxicity, antiaging, preventing cardiovascular and cerebrovascular diseases, and relieving reproductive disorders [9]. In addition, recent studies have shown that the level of selenium in patients with COVID-19 is lower than that in healthy people [10]. Compared with selenium-deficient areas, selenium-enriched areas have a higher cure rate and lower mortality [11]. Moderate selenium supplementation may help prevent the deterioration of new coronary pneumonia patients [12].

Selenium is also considered to be a beneficial trace element for plants. Low doses of selenium can improve photosynthesis, promote plant growth [13–15], and contribute to the homeostasis of essential nutrients [16], while slightly higher concentrations are toxic. The distinction between selenium deficiency and selenium poisoning is very close, and because of this narrow gap, both selenium deficiency and selenium poisoning are widespread problems worldwide [17]. Although excessive accumulation of selenium can lead to phytotoxicity, low doses of selenium still have stimulatory effects on plants. Food sources of selenium are abundant, such as seafood, meat, cereals, vegetables, and nuts, while selenium from edible plants is a significant source of selenium for humans [18–20]. Therefore, researchers still hope that plants can accumulate selenium to restore the soil environment and alleviate the problem of selenium deficiency in selenium-deficient areas. Additionally, it can be used as a selenium supplement to assist in the treatment of diseases.

The metabolism of selenium in plants is mainly carried out through the metabolic pathway of sulfur in Fig. 1 [17, 21]. SeVI in plant roots is transported to leaf chloroplasts for metabolism, while SeIV can be metabolized in roots [21]. Excess SeIV can also be converted into SeVI by sulfite oxidase for metabolism [3]. Then, ATP sulfurylase (ATPS) catalyses the combination of selenate and ATP to form 5'-adenosine phosphoselenolate (APSe) and release pyrophosphate (PPi) [1, 21]. When APSe is phosphorylated by adenosine phosphosulfate kinase (APK) to generate 3'-phospho-adenosine-5'-phosphoselenolate (PAPSe), which provides a donor molecule for the selenylation of biomolecules, all possible hydroxyl groups of selenide molecules can be catalysed by cytoplasmic sulfotransferases [3]. When APSe is catalyzed by adenosine phosphosulfate reductase (APR) to generate SeI, SeI can be combined with glutathione (GSH) to generate GS-SeO$_3^{2-}$, which is then combined with a molecule of GS-SeO$_3^{2-}$. The combined GSH generates GS-Se-SG, which is further reduced to GS-SeH and cleaved to HSe$^-$ and O-acetylserine [25]. There are five metabolic paths for Sects. [26, 27]. (1) selenocystathione, selenohomocysteine (SeHcys) and selenomethionine (SeMet) are sequentially generated. (2) Methylselenocysteine (MeSec) is generated under the catalysis of selenocysteine methyltransferase. MeSec can be converted to dimethylselenide (DMDSe) by an as yet uncharacterized enzyme. (3) Zerovalent selenium is generated under the catalysis of NifS-like protein or selenocysteinelyase. (4) The SeH group of Sec is oxidized to generate alanine selenate or pyruvate selenate [3] or generate other water-soluble small molecules containing C-Se-C [28]. (5) Participation in the synthesis of selenoproteins or replacement of cysteine into proteins to form damaged or deformed selenoproteins, oxygen proteins and nitroproteins, which can be further removed by the proteasome [29, 30].

ATP sulfurylase is a key catalytic enzyme in selenium metabolism, acting as the first step in the metabolic pathway. The activation reaction of selenate, catalysed by ATPS to form APSe, is a rate-limiting step in the selenium metabolic pathway [31]. ATPS has been found in bacteria, fungi, algae and a variety of higher plants. ATPS plays an important role not only in sulfur metabolism but also in abiotic stress of various heavy metal ions [32, 33]. The catalytic substrate of SpATPS2 in *Stanleya pin nata* can be either sulfate or selenate, which can help plants accumulate selenium [34]. Experiments have shown that transgenic ATPS-overexpressing mustard
plants accumulated more organic selenium and were more tolerant to selenium than wild-type mustard plants [35, 36]. Genome-wide identification of the ATP sulfurylase gene family has been conducted in many species [37–39], while few reports have focused on the functions of this gene family in C. hupingshanensis. Therefore, carrying out bioinformatic analysis of the key catalytic enzyme ATPS gene in the selenium metabolism pathway is preliminarily important to explore the mechanism of selenium accumulation and selenium tolerance in C. hupingshanensis.

Bioinformatics analysis of ATPS in 31 higher plants found that 84% of ATPS were located in the chloroplast, and the rest were located in the cytoplasm [38]. ATPS1-4 are found in Arabidopsis, mainly localized in the chloroplast [40], and selective translation enables Arabidopsis ATPS2 to be expressed in the cytoplasm as well [41, 42]. In plants, ATPS is a homodimer formed by the polymerization of two 48-kDa monomers [43]. Gene structure analysis ATPS contains 4–6 exons, and all ATPS contain the N-terminal domain PF14306 [PUA_2: PUA-like (pseudouridine synthase and archaeosine transglycosylase) domain] and C-terminal catalytic domain PF01747 (ATP-sulfurylase as catalytic domain) [38]. X-crystal diffraction analysis of soybean ATPS revealed that Arg248, Asn249, His255, and Arg349 play important roles in the enzymatic transition state [43]. ATP sulfurylase contains two highly conserved motifs: the HXXH motif and PP-loop [43, 44]. They contain several highly conserved histidine and arginine residues, all of which have functional side chains. ATPS exists in both allosteric and nonallosteric forms [45].

In this study, we aimed to screen and identify the substrate affinity of major members of the C. hupingshanensis ATPS family that respond to selenite stress genome-wide. First, we identified and analysed ATPS genes in C. hupingshanensis using a bioinformatic approach. The protein domain, gene structure, conserved protein motif and evolutionary tree of C. hupingshanensis ATPS gene family members were analysed to clarify the physicochemical properties and basic functions of C. hupingshanensis ATPS members. Second, qRT-PCR was used to screen the main gene from the ATPS family of C. hupingshanensis that reacted to selenite stress. Finally, molecular docking simulations were used to investigate the affinity between ATPS and the substrate.

**Results**

**Identification and analysis of ATPS genes in C. hupingshanensis**

We used the protein sequence of Arabidopsis ATPS in the C. hupingshanensis genome file to search the C. hupingshanensis.
hupingshanensis genome file (the Genome Warehouse BIG Data Center accession number PRJCA005533) with BLASTp to identify potential ChATPS genes in C. hupingshanensis. Nine ChATPS gene family members were identified from the genome of C. hupingshanensis. They were designated ChATPS1-1 to ChATPS4-2 according to the homologous AtATPS. Detailed information about each gene, including gene name, nucleotide length, isoelectric point, predicted protein molecular weight and protein subcellular localization, is given in Table 1 and S1.

As illustrated in Table 1, the length of the nucleotide sequences of the identified ChATPS genes ranged from 618 to 1473 base pairs. All of these genes contain 4 introns. The protein sizes of ChATPS members ranged from 180 aa (ChATPS 2–1) to 490 aa (ChATPS 2–3). Accordingly, the MW of ChATPS members spanned from 20149.25 Da to 54720.22 Da. In addition, the theoretical isoelectric points of the ChATPS members ranged from 5.8 to 8.61. As predicted by the online servers SignalP-5.0 and TMHMM-2.0, all ChATPS proteins have no signal peptide and no transmembrane region. We predicted the subcellular localization of the protein by aligning with the N-terminal mature peptide homologous sequences of chloroplast ATP sulfurylases from Spinach and Arabidopsis, combined with an online server [40, 46, 47]. Preliminary prediction of the subcellular locations of the members showed that ChATPS members are located in the extracellular region, two of which are located in the cytoplasmic region, and the remaining 7 are located in the chloroplast. The gene coding sequences and protein sequences of the ChATPS family members are listed in Table S2.

**Phylogenetic analysis of ATPS genes in C. hupingshanensis**

The protein sequences of 39 ATPSs were used, including 24 from monocotyledons, 4 from dicotyledons and 9 from C. hupingshanensis. We constructed a maximum likelihood (ML) phylogenetic tree by MEGA with default parameters (Fig. 2). We classified ATPS proteins into classes I, II, III, and IV based on bootstrap values and phylogenetic topology (Fig. 2).

Moreover, among the four categories, the second category is farther away from the other three categories, forming a relatively independent branch, suggesting that there may be functional differentiation between ChATPS2 and ChATPS1/3/4. Among them, 3 out of 9 ChATPS genes were distributed in class II. Arabidopsis thaliana, Brassica rapa, Brassica napus and Brassica oleracea have only one ATPS gene in Class I, Class III and Class IV. However, C. hupingshanensis has two ChATPS genes in class I, class III and class IV.
Analysis of the protein motif, conservative domain, Gene structure and sequence alignment of the ChATPSs

Genes have differentiated their regulatory and coding regions due to evolution usually based on gene duplication. As a result, amino acids may be replaced or altered, and the function of genes may be altered to suit different growth conditions. A simpler neighbour-joining phylogenetic tree was constructed from the ATPS protein sequences of *C. hupingshanensis* and *A. thaliana* to adequately recognize the protein motif, conserved domain and gene structure (Fig. 3).

On the basis of the annotated genome structure information, with the exception of ChATPS2-1, the homologous genes from different groups had the same number of introns/exons and a similar distribution. This result indicates that the ChATPS gene is extremely conserved in terms of structure and function.

The conserved motifs and conserved domains of AtATPS and ChATPS were analysed using the online software MEME and the server NCBI CDD to deeply explore the evolutionary relationship between members of different groups of ChATPS. Similar to the gene structure distribution results, the distribution of conserved motifs was conserved in different groups of genes. All ATPS contained all motif types, except for ChATPS2-1, in which motifs 4, 5, 6, 8, 9 and 10 were absent. While motifs 1, 2, 3 and 7 were correlated with the ATP-sulfurylase domain structure (PF01747), motifs 4, 5, 8 and 9 were associated with the PUA_2 (PF14306) domain structure.

Conserved amino acid residues and chromosomal distribution and analysis of ChATPS genes

The distribution of ChATPS genes on the 16 chromosomes of the *C. hupingshanensis* genome is relatively random (Fig. 4). With the exception of chromosome 8, which contains two ChATPS genes (ChATPS2-2 and ChATPS2-3), each of the remaining ChATPS genes is located on a separate chromosome. Notably, most of the ChATPS genes are...
located at the distal end of the chromosome, with five members in a reverse distribution and the other four members in a positive distribution.

From the alignment of the full-length sequences of ChATPS proteins (Fig. S1), the ChATPS protein displayed an N-terminal mature peptide (50GLIEPDGKLVDLVVPEPRR69), which was characterized by transit peptide localization to the chloroplast and had greater than 60% N-terminal homology with the native chloroplast ATPS of *Arabidopsis* and *spinach* [40, 46, 47].

The relatively conserved sequence exists in the C-terminus (Fig. 5). We further analyzed the conservation of amino acid residues in this domain, similar to the analysis in *A. thaliana*. The amino acid residues in the C-terminal domain remained conserved at most loci, which was assumed to be required for ATP sulfurylase. Remarkably, two conserved motifs present in the ChATPSs are the PP-loop (343GANFYIVGRDPAGM360) and HXXH motif (254HNGH257), except for ChATPS2-1.

Furthermore, cysteine residues are key redox targets and play important roles in redox regulatory mechanisms [48]. According to Prioretti et al. [49], algal ATPS proteins contain a large number of cysteine residues and are highly conserved compared to ATPS genes of plants and other organisms. Their research showed that cyanobacteria, marine cyanobacteria, green algae, hyaluronicum and heteroalgae contain five highly conserved cysteine residues. However, the cysteine residues identified in our study are not conserved structures and are few in number. These data are inconsistent with previous findings [49].

Combined with the gene structure, conserved domains, motifs and multiple sequence alignment results, we speculate that ChATPS2-1 may be a mutation or functional redundancy during evolution. Therefore, we will not
perform protein function exploration and gene expression analysis.

Prediction of secondary and tertiary structures of ATPS protein

Secondary structure analysis of the eight ChATPS proteins (Table S3) revealed the presence of α-helices (29.89–32.04%), extended strands (15.70–19.39%), β-turns (5.51–9.44%) and random coils (38.75–46.67%), indicating minor structural differences in ChATPS sequences. BLAST searches were performed on the SWISS-MODEL library to determine a suitable template for the *C. hupingshanensis* ATP sulfurylase. The 3D structures of ChATPS and AtATPS used for docking were also predicted by the SWISS-MODEL server (Fig. 6). The soybean ATP sulfurylase (PDB code: 4MAF) with the highest similarity scores (ranging from 76.62 to 86.07%) was selected as a template (Table 2). The 3D modelled protein structures of ChATPS have high GMQE (0.77–0.89) and QMEAN (0.87–0.89) scores, indicating high confidence in the modelled structures. These ChATPS models were validated with the Structural Analysis and Validation Server (SAVES). Ramachandran plots show that nearly 90% of all models have residues in the favourable region, with ≥95% of residues in the core and allowable regions, which is sufficient to indicate the reliability of the 3D model. Overall quality factor values were greater than 90 in all generated models. The average 3D-1D scores for the nine model residues were higher than 0.2. Furthermore, the plausibility of torsion angles and covalent geometric distributions within the model are indicated by G-factor values, all greater than −0.5. In general, homology models are stable and reliable. ProSA analysis showed that all models had Z scores between −10.99 and −11.65. Finally, an LG score (greater than 3) indicates that the protein model is of good quality. All ChATPS protein structures achieved significant scores, with LG scores greater than 7.3. These results indicate that these models obtained with the homology model are acceptable and can be used for further studies.

Molecular Docking

Molecular docking is a novel technique for identifying binding modes or forces of ligand-protein complexes and is widely used in structural molecular biology and drug discovery [51]. First, we used Prankweb to predict and
Fig. 6 Predicted 3D structures of ChATPSs and AtATPSs by the SWISS-MODEL server
visualize the ligand-binding sites of ATPSs [52]. Several ligand-binding sites were predicted for each ATPS, ranging from 10 to 15 for ChATPSs and 7 to 13 for AtATPSs. The online server numbers each site with Arabic numerals starting from 1 according to the calculated probability score, which is also the basis for our naming of binding sites. Therefore, for different proteins, the same Arabic numerals do not necessarily represent the same spatial position in each protein. For example, part of the ligand binding sites of protein ChATPS1-2 are visualized in Fig. 7. Afterwards, protein-ligand docking and molecular simulations were performed using the AutoDock Vina program [53]. The binding energy of protein-ligand docking, which is an important criterion for interaction, was recorded, with a lower binding energy being considered more stable [53]. The structure of protein-ligand interactions was finally analysed using a protein-ligand interaction analyser (PLIP) and visualized with PyMOL [54, 55].

Bioinformatics analysis and preliminary study were used to verify the binding ability of selenate, selenite, sulfate, and sulfite with ChATPS according to network analysis and preliminary research. The docking binding energy of each ligand compound to the protein molecule is displayed in the heatmap (Fig. 8). When comparing the docking results between the ligand and ChATPS, the most noticeable difference was the interaction energy, which ranged from $-4.3$ to $2.3$ kcal mol$^{-1}$. We found that all ChATPS had stronger affinity for selenate than other compounds, with ChATPS1-1 ($-4.2$ kcal mol$^{-1}$), ChATPS1-2 ($-4.2$ kcal mol$^{-1}$) and ChATPS3-1 ($-4.3$ kcal mol$^{-1}$) showing stronger affinity for selenate than other genes.

When analysing protein-ligand interactions, it was found that most of the amino acids at site 1 of each protein consisted of two conserved motifs, ATPS, PP-loop and HXXH. It shares a similar spatial structure with the ATPS catalytic role.
site [43, 56]. We call site 1 of each protein the catalytic site, which is abbreviated as CS in Fig. 9. Therefore, we selected the catalytic site and one of the binding sites with the minimum binding energy in the docking simulation to visualize the interaction of the binary ATP-ATPS complex with selenate, including the amino acid residues involved in the interaction and the interaction forces (hydrogen bonds, salt bridges and π-cation interactions). Hydrogen-bond interactions were found to be necessary for the interactions of the binary ATP-ATPS complex with selenate.

The interaction at the catalytic site is similar to GmATPS in that the ligands are surrounded by positively
charged residues Arg250, His254, His257 and Arg350. In previous studies, most of these conserved amino acid residues were shown to interact with the β- or γ-phosphates of ATP \[43\]. In our study, these conserved ATP residues still interact with ATP phosphate and hydrogen bond with ATP adenine and selenate. These may be the reasons for the higher affinity of ChATPS for selenate.

At the maximum affinity binding site, similar to the catalytic site, the ligand is surrounded by positively charged amino acids Arg105 and Arg109. Although not among the characteristic catalytic sites of ATPSs, most of the highest affinity sites in ChATPS are located in similar spatial positions and have similar amino acid compositions (Arg105, Gly106, Arg109, Ser111 and Glu112).

Expression profiles of ATPS genes in different tissues under Se stress

ATP-sulfurylase can participate in plant responses to several abiotic stresses through different sulfides. To better understand the molecular functions of ChATPS genes under abiotic stress conditions, RT-qPCR technology was used to analyse the expression of nine ChATPS genes in *C. hupingshanensis* leaves under different concentrations

---

**Fig. 9** Interactions of the binary ATP-ATPS complex with selenate. The left panel is the overall view, and the right panel is the focused view. The ATPS protein is shown on the surface, the amino acid residues at the binding site are grey-blue, and the ligand (ATP and selenate) is heavy yellow. The blue solid line represents a hydrogen bond, the yellow dashed line represents a salt bridge, and the red dashed line represents a π-cation interaction. CS: putative binding mode of ATP and selenate to model the ATPS protein structure at the catalytic site. MBS: ATP and selenate are in a putative binding mode that mimics the protein structure of ATPS at the site of minimum binding energy, the site of maximum affinity binding.
of Se stress (0 µg Se L⁻¹, 100 µg Se L⁻¹ and 80,000 µg Se L⁻¹).

Among the ChATPS family genes measured by qRT-PCR, under low-concentration selenium stress (100 µg Se L⁻¹), the gene expression in roots was upregulated (Fig. 10 A). Among them, ChATPS1-2 genes were highly upregulated (approximately 29.5-fold) at 6 h. ChATPS1-1 was upregulated approximately 9.4-fold at 6 h, and ChATPS3-1 was upregulated approximately 8.5-fold at 3 h. The upregulation of the remaining six genes in roots was relatively small (1.5- to 4.6-fold) under low-concentration selenium stress. The expression of ChATPS family genes was upregulated in leaves under low-concentration selenium stress (Fig. 10 B). The upregulation of ChATPS1-1 and ChATPS1-2 gene expression was prominent (approximately 29.5-fold and 11.2-fold) at 6 h. ChATPS2-2, ChATPS2-3 and ChATPS3-1 were upregulated at 24 h by approximately 8.8-fold, 8.7-fold and 6.8-fold, respectively. The remaining four genes were upregulated to a smaller extent (2 to 3.7 times) in leaf parts under low-concentration selenium stress.

Under high-selenium stress (80,000 µg Se L⁻¹), only ChATPS4-2 was downregulated in roots. The expression of other members of the ChATPS gene family was upregulated (Fig. 10 C). Among them, ChATPS1-1 and ChATPS1-2 were upregulated approximately 6.7-fold and 10-fold at 6 h, respectively, and ChATPS3-1 was upregulated approximately 10.6-fold at 3 h. The upregulation of the remaining four genes in roots was relatively small (1.4- to 4-fold). ChATPS family members were upregulated to varying degrees in leaves under high-concentration selenium stress (Fig. 10 D). Simultaneously, ChATPS1-1 and ChATPS1-2 were upregulated approximately 6.1-fold and 6-fold at 3 h, respectively. ChATPS2-2 and ChATPS2-3 were upregulated approximately 6.7-fold and 7-fold at 24 h, respectively. The remaining four genes were upregulated to a lesser extent (1.6- to 4.2-fold) in leaf parts under high-concentration selenium stress.

Based on these data, ChATPS may play an important role in selenium detoxification by promoting selenide production.

**Discussion**

In our study, more *C. hupingshanensis* ATPS genes were identified from the genome database using bioinformatics methods, which is a potential reason for Se tolerance and Se accumulation capacity in Se hyperaccumulators. In total, 9 ATPS genes were recognized in *C. hupingshanensis*, which is relatively more abundant than that of *Arabidopsis thaliana* (4 AtATPS), *Brachypodium distachyon*
The ATPS family genes of *C. hupingshanensis* shared conserved structures and motifs but had a stronger affinity for selenate. This allows inorganic selenium to enter metabolic pathways faster, helping plants accumulate selenium. First, ChATPSs exhibited high similarity to AtATPSs in motifs, ATPS-type domains, CDS regions, and exons (Fig. 3B), which makes them persist in catalytic function. Based on their interfamily similarities, relation with their homologues from other species, such as *Arabidopsis thaliana*, and motif distribution, the ChATPS gene family was classified into four subfamilies. The obtained results are in line with those of previously reported studies in *Arabidopsis*. Moreover, the conserved MEME motifs of ATPS proteins also exhibited corresponding permutation and combination with their phylogenetic relationship (Fig. 3 C). These results implied the possibility that the ATPS gene family may function in a conserved manner in *C. hupingshanensis* and *Arabidopsis*. Analysis of the characteristics of the identified ATPS proteins showed that all ATPSs have no signal peptides and no transmembrane region. The absence of a signal peptide, transmembrane domain and cysteine residues and disulfide bonds indicate that these proteins are likely intracellular in nature. In addition, according to the conserved amino acid residue analysis, two conserved domains present in ChATPSs are the PP-loop and HXXH motif (Fig. 5) [44, 56, 58], which constitutes Site1 of these proteins. In addition, the results of molecular docking calculated by computer algorithms cannot completely simulate the actual conformational changes of proteins. Therefore, when we docked the protein as a semiflexible molecule to the substrate, the optimal binding region did not appear in the catalytic domain. We are still conducting further biological experiments to explore the true binding form. On the other hand, after molecular docking simulations, it was found that the ligand selenate has a stronger affinity for the ChATPS protein than AtATPS. In terms of amino acid composition at the affinity site, hydrogen bonds, salt bridges and n-cation interactions together form affinity interactions. Additionally, we found that the sites with higher affinity in the ChATPS protein were enriched in the following residues: Arg105, Gly106, Arg109, Ser111 and Glu112. These residues constitute a relatively conserved sequence $^{105}$RGXXRXSE$^{112}$ in the ChATPS genes (Fig. 5).

In conclusion, the results of this study provide important insights into the function of ChATPS genes in Se hyperaccumulators and their responses to selenium stress conditions. In this context, nucleotide and protein sequence analysis and phylogeny, determination of gene expression profiles of *C. hupingshanensis* under selenium stress, and 3D structure prediction of ChATPS was performed. Notably, the ATPS gene is highly conserved. In addition, the *C. hupingshanensis* ATPS gene showed...
different expression patterns according to time and stress intensity, indicating dynamic regulation. The results of this study may support the understanding of the selenium assimilation pathway in higher plants under abiotic stress conditions.

**Materials and methods**

**Genome-wide identification of ATPS Family genes**

*C. hupingshanensis* genome and its annotation file were obtained from the Genome Warehouse BIG Data Center under accession number PRJCA005533. To identify the ATPS gene family in *C. hupingshanensis*, the AtATPS protein sequences were downloaded from the Arabidopsis Information Resources (TAIR) database (https://www.arabidopsis.org/). Using AtATPS as the query sequence, the most representative ChATPS protein sequence was extracted by the Blast Zone of TBtools software [59]. In addition, the extracted ChATPS proteins were further checked by NCBI BLAST (https://blast.ncbi.nlm.nih.gov/blast/Blast.cgi). The conserved domains of ChATPS proteins were analysed by CD-search (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi).

**Bioinformatic analysis of the ATPS genes**

ChATPS chromosomal location information was extracted from the *C. hupingshanensis* genome GFF file and plotted by “Gene Location Visualize from GTF/GFF” of TBtools software. In addition, the molecular weight (MW), isoelectric point (pI) and other physical and chemical properties of the ChATPS family can be predicted and analysed using the online tool ExPASy (https://web.expasy.org/protparam/). The online software SignalP 5.0 (http://www.cbs.dtu.dk/services/SignalP/) was used to predict signal peptides. WoLF PSORT (https://wolfsort.hgc.jp/) was used for ChATPS gene subcellular localization predictions. The transmembrane regions of proteins were analysed by TMHMM 2.0 (http://www.cbs.dtu.dk/services/TMHMM/) [61]. The AtATPS and ChATPS protein sequences were aligned by ClustalW (https://www.genome.jp/tools-bin/clustalw), and the alignment result was further processed by ESPript 3.0 (https://esprit.ibcp.fr/ESPrift/cgi-bin/ESPrift.cgi) to output the image [50].

Submit the ChATPS and AtATPS protein sequences to perform a conserved motif scan on the MEME website (http://meme-suite.org/tools/meme) with the MEME-motif set to 10. ChATPS and AtATPS protein sequences were submitted to CDD: NCBI’s conserved domain database (https://www.ncbi.nlm.nih.gov/Structure/bwpsb/bwpsb.cgi) to obtain conserved domain information. The intron-exon gene structure information of the ChATPS and AtATPS genes was extracted from the GFF files of the respective genomes. Submit the Newick Tree file output by MEGA, the xml file obtained from the MEME website, the HitDate file and genome GFF file obtained by NCBI-CDD, and visualize by “Gene Structure View (advanced)” of TBtools.

**Phylogenetic analysis of ChATPS**

To explore the phylogenetic relationship of ChATPS family genes, 4 from *Arabidopsis thaliana* (At), 8 from *Camelina sativa* (Cs), 4 from *Brassica rapa* (Br), 4 from *Brassica napus* (Bn), 4 from *Brassica oleracea* (Bo), 2 from *Oryza sativa* (Os) and 2 from *Zea mays* (Zm) were downloaded from NCBI for multiple sequence alignment. The amino acid sequences were aligned using Clustal W, and then a maximum likelihood (ML) tree was constructed with all of the ChATPS protein sequences using MEGA 11, bootstrap = 1000 repetitions.

**Homology modelling and validation of ChATPS**

SOPMA was used to predict the protein secondary structure [62]. Then, we searched and selected the best crystal structure as a template in the SWISS-MODEL (https://swissmodel.expasy.org/) template library and used the SWISS-MODEL web server to model the ChATPS protein homology. The final 3D models of ChATPS were validated using the online server SAVES 5.0 (https://services.mbi.ucla.edu/SAVES/) with various functions.

**Ligand Preparation and Molecular Docking**

The compounds (ATP, selenate, selenite, sulfate, sulfite) used in this study were selected from the Chemspider database. Their structures were sketched with ChemSketch. The compounds were downloaded from the Chemspider database. Their structures were sketched with ChemSketch and saved in protein data bank format. PrankWeb [52] was used to predict protein active sites.

Experiments with the docking of proteins and ATP sulfurylase were performed using AutoDock v4.2 and AutoDock Vina v1.1.2 [53]. We used AutoDock v4.2 to modify proteins and ligand compounds, adding all hydrogens, incorporating nonpolar hydrogens and calculating Gasteiger charges. We used AutoDock Vina to perform the molecular docking of compounds with ATP sulfurylase proteins with the exhaustiveness setting at 10. The best aptamer conformations are selected based on their minimal binding energies.

First, the ATP-Protein complex PDB file was formed by docking the substrate ATP with the binding pockets of each protein molecule by AutoDock Vina and PyMOL. Only pockets with negative binding energy can be docked in the next step. Then, the small molecule ligand and ATP-protein complex PDB files were docked by AutoDock Vina to form a small molecule-ATP-protein complex PDB file. The interaction of the small molecule-ATP-protein complex (hydrogen bond and hydrophobic interaction) was analysed and visualized by PLIP and PyMol, and the docking binding energy was visually analysed by GraphPad Prism version 9.0.0 for Windows,
GraphPad Software, San Diego, California USA, www.graphpad.com.

Gene expression analysis
Seeds of *C. hupingshanensis* were collected from the Yutangba Colour Mine in Enshi, Hubei Province, China. Plants were placed in a room with a constant temperature of 22 ± 1 °C, a photoperiod of 16 h, and an irradiance of 1500 mol m⁻² s⁻¹. Thirty-nine seedlings approximately 10 cm tall and four months old were selected, and the roots were rinsed with vermiculite and equilibrated in Hoagland's solution for two days. The samples were treated with selenium at different concentrations (0 µg Se L⁻¹, 100 µg Se L⁻¹, 80,000 µg Se L⁻¹), the actual concentration of elemental selenium coming from the selenite (analytical reagent, Sinopharm Chemical Reagent Co., Ltd, Shanghai, China), and the samples treated with 0 µg Se L⁻¹ were used as experimental controls. Leaves and roots of 3 seedlings were isolated from each treatment at 0 h, 3 h, 6 h, 12 and 24 h, and these samples were snap frozen in liquid nitrogen for RNA extraction.

Total RNA was extracted from root and leaf samples by the TRIzol method, and the RNA concentration and quality were detected by a NanoDrop 2000. RNA integrity and genomic DNA contamination were detected by gel electrophoresis. RNA samples were treated with RNase-Free DNase to remove residual genomic DNA. Real-time PCR was performed on ABI StepOne Plus. The primers used in qRT-PCR analysis for ChATPS are shown in Table S4. The EviScript RNA SYBR Green I Master Kit (Roche) was used to perform qRT-PCR analysis with a *C. hupingshanensis* gene expression analysis. Colony graphs were generated using GraphPad Prism version 9.0.0 for Windows, GraphPad Software, San Diego, California USA, www.graphpad.com. All assays were carried out in triplicate.

**Supplementary Information**
The online version contains supplementary material available at https://doi.org/10.1186/s12870-022-03872-7.

**Acknowledgements**
We would like to acknowledge Chuying Huang for support and assistance with the study subjects.

**Author contributions**
Q.T. and Y.Z. designed the research, Q.T. and K.L. prepared the plant materials, Z.X., Y.L. and Y.Z. completed the bioinformatics analysis of *C. hupingshanensis* ATPS genes and prepared tables. Y.L., Z.X. and Q.T. isolated RNA and analysed the differential expression of genes. Z.X. wrote the main manuscript and prepared the figures. All authors reviewed and approved the final manuscript.

**Funding**
This work was supported by the National Natural Science Foundation of China (32260070), the Excellent Young and Middle-aged Scientific and Technological Innovation Team Projects of Colleges and Universities in Hubei Province (T20200020), the Enshi Autonomous Prefecture Science and Technology Program (D2018009) and the Open Foundation of Hubei Key Laboratory of Biological Resources Protection and Utilization (PT2012102) to Yifeng Zhou. The research was mainly finished at the Hubei Key Laboratory of Biological Resources Protection and Utilization.

**Data availability**
The data that support the findings of this study are available from the corresponding author upon reasonable request.

**Declarations**

**Ethics approval and consent to participate**
We declare that all experimental procedures, including the use of plants and the collection of seeds of *C. hupingshanensis*, comply with ethical standards and legislation. This article does not contain any studies with human participants or animals and did not involve any endangered or protected species.

**Consent for publication**
Not applicable.

**Conflict of interest**
The authors declare no conflict of interest.

Received: 13 June 2022 / Accepted: 4 October 2022
Published online: 18 October 2022

**References**

1. White PJ. Selenium accumulation by plants. Ann Botany. 2015;117(2):217–35.
2. Chuying H, Hongqin Y, Xibiao Y, Yuan G, Tuo L, Bo W, Meng R, Zixiong Z, Jun D, Jianhua G, et al: The Cardamine enshiensis genome reveals whole genome duplication and insight into selenium hyperaccumulation and tolerance. Cell Discovery 2021, 7(1).
3. Zhou Y, Tang Q, Wu M, Mou D, Liu H, Wang S, Zhang C, Ding L, Luo J. Comparative transcriptomics provides novel insights into the mechanisms of selenium tolerance in the hyperaccumulator plant Cardamine hupingshanensis. Sci Rep. 2018,8(1):2789.
4. Kieliszek M, Bano I, Zare H. A Comprehensive Review on Selenium and Its Effects on Human Health and Distribution in Middle Eastern Countries. Biol Trace Elem Res. 2022,200(3):971–87.
5. Minich WB. Selenium Metabolism and Biosynthesis of Selenoproteins in the Human Body. Biochem (Moscow). 2022,87(1):168–77.
6. Liu H, Wang X, Zhang B, Han Z, Wang W, Chi Q, Zhou J, Nie L, Xu S, Liu D. Concentration and distribution of selenium in soils of mainland China, and implications for human health. J Geochem Explor. 2021,220:106654.
7. Carlisle AE, Lee N, Matthew-Onabanjo AN, Spears ME, Park SJ, Youkana D, Doshi MB, Peppers A, Li R, Joseph AB. Selenium detoxification is required for cancer-cell survival. Nat Metabol. 2020,2(7):603–11.
8. Steinbrenner H, Speckmann B, Pinto A, Sies H. High selenium intake and increased diabetes risk: experimental evidence for interplay between selenium and carbohydrate metabolism. J Clin Biochem Nutr. 2010,48(1):40–5.
9. Alcántara DBDA, Artur AG, Silveira BKS, Lopes AF, Guedres JAC, Luz LR, Nascimento RF, Lopes GS, Hermsdorff HHM, Zoccolo GJ. Selenium in Brazil nuts: An overview of agronomical aspects, recent trends in analytical chemistry, and health outcomes. Food Chemistry 2021.
10. Younesian O, Khodabakhshi A, Abdollahi N, Norouzi A,Behnampour N, Hossein Zadeh S, Aliar S, Joshiaghdam H. Decreased serum selenium levels of COVID-19 patients in comparison with healthy volunteers. Biol Trace Elem Res. 2022;2020(4):1562–7.
11. Hou J, Zhu L, Chen C, Feng H, Li D, Sun S, Xing Z, Wan X, Wang X, Li F. Association of selenium levels with the prevention and control of Keshan disease: A cross-sectional study. J Trace Elem Med Biol. 2021;68:126832.
12. Fakhrolmobasheri M, Mazarieh-Teherani S, Keliszek M, Zeinalian M, Abbasii M, Karimi F, Mozaffari AM. COVID-19 and selenium deficiency: a systematic review. Biological Trace Element Research 2021:1–12.
13. Six J. Plant nutrition for sustainable development and global health. Plant Soil. 2011;339(1–2):1073–90.
14. Mirza H, Bhuyan MRMB, Ali R, Barbara H-N, Renata M-G, Jubayer Al M, Kamran M, Masayuki F. Selenium in plants. Boin or bane? Environmental and Experimental Botany. 2020. 178.
15. Fatima Salwa N, Mohammad Y, Tanveer AK, Qazi F, Agil A. Low level of selenium increases the efficacy of 24-epibrassinolide through altered physiological and biochemical traits of Brassica juncea plants. Food Chem. 2015;165:441–8.
16. Chauhan R, Avasthi S, Tripathi P, Mishra S, Dwivedi S, Niranjain A, Mallick S, Tripathi P, RD. Selenium regulates the level of phenolics and nutrient element to alleviate the toxicity of arsenic in rice (Oryza sativa L.). Ecotocool Environ Saf. 2017;138:47–55.
17. Schiavo M, Plon-Smits EH. Selenium Biofortification and Phytoremediation Phytotechnologies: A Review. Environ J. 2017;46(1):10–9.
18. Moreda-Piñeiro J, Moreda-Piñeiro A, Romaris-Hortas V, Dominguez-González R, Alonso-Rodríguez E, López-Mahía P, Munuategui-Lorenzo S, Prada-Rodríguez D, Bermejo-Barrera P. In vitro bioavailability of total selenium and selenium species from seafood. Food Chem. 2013;139:872–7.
19. Moreda-Piñeiro J, Moreda-Piñeiro A, Bermejo-Barrera P. In vivo and in vitro testing for selenium and selenium compounds bioavailability assessment in foodstuff. Crit Rev Food Sci Nutr. 2017;57(4):805–33.
20. Rayman MP. Selenium and human health. The Lancet. 2010;373(9661):1189–99.
21. Chauhan R, Avasthi S, Srivastava S, Dwivedi S, Pilon-Smits EH, Dhanakhor OP. Understanding selenium metabolism in plants and its role as a beneficial element. Crit Rev Environ Sci Technol. 2019;49(21):1937–58.
22. Ganther HE. Selenosulfurides. Formation by the reaction of thiols with selenious acid. Biochemistry. 1968;7(8):2986–905.
23. Ganther HE. Reduction of the selenosulfuride derivative of glutathione to a persulfide analog by glutathione reductase. Biochemistry. 1971;10(22):4089–98.
24. Tarze A, Dauplais M, Grigoras I, Lazard M, Ha-Duong N-T, Barbier F, Blanquet R, AS. Plastid-cytosol partitioning and integration of two cDNA clones encoding a persulfide analog by gluthathione reductase. Biochemistry. 1968;7(8):2898–905.
25. Van Hoewyk D. A tale of two toxicities: malformed selenoproteins and oxidative stress both contribute to selenium stress in plants. Ann Botany. 2007;282(12):8759–67.
26. Ouerdane L, Both EB, Xiang J, Yin H, Kang Y, Shao S, Kiszelák K, Jókai Z, Robert X, Gouet P. Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res. 2014;42(W1):W320–4.
27. Prioretti L, Gontero B, Hell R, Giordano M. Diversity and regulation of ATP sulfurylase in photosynthetic organisms. Frontiers in Plant Science 2014, 5.
28. Abdollahi K, Ince C, Condict L, Hung A, Kasapis S. Combined spectroscopic and molecular docking study on the pH dependence of molecular interactions between β-lactoglobulin and ferulic acid. Food Hydrocolloids 2019, 101.
29. Le Duc DL, Abdel-Samie M, Monetés-Bayon M, Wu CP, Reisiger SJJ, Terry N. Overexpressing both ATP sulfurylase and selenocysteine methyltransferase enhances selenium phytoaccumulation traits in Indian mustard. Environ Pollut. 2006;144(1):70–6.
30. Pilon-Smits EA, Hwang R, Van Hoewyk D, Mol优越 C, Zhan Y, Terry N, Kawaka H, Chen Y, Leustek T, Terry N. Overexpression of ATP sulfurylase in Indian mustard leads to increased selenate uptake, reduction, and tolerance. Plant Physiol. 1999;119(1):123–32.
31. Hatzfeld Y, Lee M, Lee S, Leustek T, Saito K, Takahashi H. Alternative translational initiation of ATP sulfurylase underlying dual localization of sulfation pathways in plants and cytosol and Arabidopsis thaliana. Frontiers in Plant Science 2014, 5.
32. Moreda-Piñeiro J, Moreda-Piñeiro A, Bermejo-Barrera P. In vivo and in vitro testing for selenium and selenium compounds bioavailability assessment in foodstuff. Crit Rev Food Sci Nutr. 2017;57(4):805–33.
33. Robert X, Gouet P. Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res. 2014;42(W1):W320–4.
34. Deyrup AT, Singh B, Krishnan S, Lyle S, Schwartz NB. Chemical modification and site-directed mutagenesis of conserved HXXH and PP-loop motif arginines and histidines in the murine bifunctional ATP sulfurylase/adenosine 5’-phosphosulfate kinase. J Biol Chem. 1999;274(41):28929–36.
35. Prioretti L, Gontero B, Hell R, Giordano M. Diversity and regulation of ATP sulfurylase in photosynthetic organisms. Frontiers in Plant Science 2014, 5.
36. Schrodinger LLC: The PyMOL Molecular Graphics System, Version 1.8. In.; 2022:22:491.
56. Kristian P, Günter Fritz, Ulrich, Ermler, Peter, Kroneck: Conserving energy with sulfate around 100°C—structure and mechanism of key metal enzymes in hyperthermophilic Archaeoglobus fulgidus. Metallomics Integr Biometal Sci. 2013;5:267–406.

57. Schiavon M, Pilon M, Malagoli M, Pilon-Smits EAH. Exploring the importance of sulfate transporters and ATP sulphurylases for selenium hyperaccumulation—a comparison of Stanleya pinnata and Brassica juncea (Brassicaceae). Frontiers in Plant Science. 2015; 6.

58. Venkatachalam KV, Fuda H, Koonin EV, Strott CA. Site-selected mutagenesis of a conserved nucleotide binding HXGH motif located in the ATP sulfurylase domain of human bifunctional 3’-phosphoadenosine 5’-phosphosulfate synthase. J Biol Chem. 1999;274(5):2601–4.

59. Chengjie C, Hao C, Yi Z, Hannah RT, Margaret HF, Yehua H, Rui X. TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. Mol Plant. 2020;13(8):1194–202.

60. Gasteiger EHC, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A. Protein identification and analysis tools on the ExPASy server. The proteomics protocols handbook 2005:571–607.

61. Möller S, Croning MDR, Apweiler R. Evaluation of methods for the prediction of membrane spanning regions. Bioinformatics. 2001;17(7):646–53.

62. Combet C, Blanchet C, Geourjon C, Deléage G. NPS@: Network Protein Sequence Analysis. Trends Biochem Sci. 2000;25(3):147–50.

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