Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Gene Identification, Expression Analysis, and Molecular Docking of SAT/OASTL Reveal the Molecular Mechanisms of Selenocysteine Synthesis in Cardamine hupingshanensis

yushan chen
Hubei Key Laboratory of Biological Resources Protection and Utilization, Hubei Minzu University
Cihang Luo
Collage of Biological and Food Engineering, Hubei Minzu University
Yi Liu
Collage of Biological and Food Engineering, Hubei Minzu University
Zhijing Xiao
Hubei Key Laboratory of Biological Resources Protection and Utilization, Hubei Minzu University
Yanke Lu
Collage of Biological and Food Engineering, Hubei Minzu University
Zhi Hou
Collage of Biological and Food Engineering, Hubei Minzu University
Qiang Xiao
Collage of Forestry and Horticulture, Hubei Minzu University
Yifeng Zhou (77416757@qq.com)
Collage of Biological and Food Engineering, Hubei Minzu University
Qiaoyu Tang
Hubei Key Laboratory of Biological Resources Protection and Utilization, Hubei Minzu University

Research Article

Keywords: Serine acetyltransferase, O-acetylserine (thiol) lyase, Selenocysteine synthesis, Se hyperaccumulator plant, Molecular docking, Selenium metabolism, Expression analysis

Posted Date: April 13th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2771418/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Abstract

Background

A complex coupled with serine acetyltransferase (SAT) and O-acetyl serine (thiol) lyase (OASTL) is the key enzyme that catalyses selenocysteine synthesis in plants. The basic bioinformatics and functions of these two gene families have been reported for many plants in addition to *Cardamine hupingshanensis*, and the response of the ChSAT and ChOASTL family members under selenium stress has not been examined to date.

Results

In this study, genome-wide identification and comparative analysis of ChSAT and ChOASTL were performed. The eight genes from the ChSAT family were divided into three branches, and the thirteen genes from the ChOASTL family were divided into four branches by phylogenetic analysis and sequence alignment, indicating the evolutionary conservation of the gene structure and its association with other plant species. The expression of members in the ChSAT and ChOASTL families was studied under selenium stress, and ChSAT1;2 and ChOASTLA1;2 were upregulated by 9.4- and 7.4-fold, respectively, showing that these two proteins are involved in the synthesis of selenocysteine. Likewise, ChCS-like protein was upregulated by 8.8-fold, playing key roles in degradation. In addition, molecular docking simulations showed that ChOASTL binds to the test compound selenophosphate more strongly than selenide and sulfide, and the major motifs that bind the target compound are usually located at residues of amino acids Lys^{46}, Gly^{181}, Thr^{182}, Gly^{183}, Thr^{185} and Ser^{269}.

Conclusions

This study revealed that selenophosphate was the optimal substrate of ChOASTL and participated in selenocysteine synthesis. The results of gene expression and molecular docking indicated that the ChSAT and ChOASTL genes were upregulated under selenium stress, and ChOASTL family genes could both synthesize and degrade cysteine/selenocysteine, which provided a theoretical basis for the regulation of selenocysteine synthesis.

1. Introduction

Selenium is among the essential trace elements for mammals and mainly performs its biological functions as selenocysteine to synthesize selenoproteins [1]. In the past century, selenium deficiency was discovered to be the main reason for endemic diseases, such as Keshan disease and Kashin-Beck disease [2, 3]. Selenocysteine (Sec) is the 21st amino acid, is encoded by TGA, was discovered in 1986 and plays a crucial role in the active centre of some antioxidant enzymes, including the glutathione peroxidase (GSH-Px) family and thioredoxin reductase (TrxR); in addition, Sec can effectively remove hydrogen peroxide (H_2O_2), lipid peroxides, reactive oxygen species (ROS) and other radicals and protect cells from oxidative damage and the integrity of the plasma membrane [4–6]. In recent years, selenium-containing compounds have been applied to treat cancer and prevent diseases of the immune system, and significant results have been achieved [7, 8].

Selenium has also been found to be a beneficial element in plants because it effects plant growth and development [9]. Plants can be divided based on their selenium content into selenium nonaccumulators, selenium accumulators, and selenium hyperaccumulators [10]. In recent years, an increasing number of selenium hyperaccumulator plants have been found, and these plants mainly include *Neptunia amplexicaulis* (legume), *Astragalus disulfide* (legume), *Stanleya pinnata* (brassicaceae), and *Cardamine hupingshanensis* (brassicaceae) [11]; in addition, the accumulation and tolerance mechanisms have become a research focus [12, 13]. Surprisingly, the selenium content of *N. amplexicaulis*, an herbaceous legume from Australia, has been recorded up to 4334 mg kg^{-1} DW in foliar tissue, and in the form of methylselenocysteine (MeSeCys) and selenomethionine (SeMet), these two types of molecules have pharmacological value for dietary supplements due to their high bioavailability [14]. MeSeCys and SeMet were mainly metabolites found in *S. pinnata* and *A.*

disulfide, which are well-known selenium hyperaccumulators from Brassicaceae and legumes, respectively [15–17]. *C. hupingshanensis*, a famous selenium hyperaccumulator plant, was first found in the Yutangba Selenium mineral area, Enshi City, Hubei Province, China [18], with the average Se content of roots, stems, and leaves reaching 2985, 3329, and 2491 mg kg^{-1} DW, respectively [19].

The plant *C. hupingshanensis* is usually used as a research object to explore the mechanism of uptake, transport, and metabolism of selenium because the plant exhibits resistance and hyperaccumulation to selenium [20]. The genome, transcriptome and proteome of *C. hupingshanensis* were analysed, the mechanism of selenium resistance and hyperaccumulation was speculated, and the enzymes of the metabolic pathway (Fig. 1) were used to catalysed biochemical reactions, such as reactions with the ATP sulfurylase family [21–23]. From the schematic, it could be concluded that Sec and SeMet were the junction of the metabolic pathway of selenium, and these two seleno-amino acids were the main carriers for the physiological function of selenium. Cysteine synthase, a complex composed of serine acetyltransferase (SAT) and O-acetylserine (thiol) lyase (OASTL), is an important component of the intracellular selenium metabolic pathway and plays an important role in the synthesis of Sec. The reaction mechanism of Sec synthesis involves the combination of serine and acetyl-CoA catalysed by SAT to produce O-acetylserine (OAS) and then selenide reaction catalysed by OASTL to produce Sec (Fig. 2).

The first important member of the cysteine synthase complex (CSC) is serine acetyltransferase (SAT), which catalyses the synthesis of the intermediate OAS using acetyl-CoA and L-serine [24]. SAT, as an important regulator of Sec biosynthesis in plants, has also been found in animals, microorganisms, and algae. A total of 5 SAT gene members were identified in Arabidopsis thaliana, including AtSAT1, AtSAT3, and AtSAT5 (in the plasmid, mitochondria, and cytoplasm, respectively) and AtSAT2 and AtSAT4 (in the cytoplasm); these genes exhibited lower levels of expression and different protein sequences from the others [25–27]. The SAT monomer consists of an N-terminal α-helix structural domain (α1-α8) and a C-terminal left-parallel β-helix structural domain, which is characteristic of acyltransferases [28]. Additionally, SATs contain the N-terminal domain structure of serine acetyltransferase (SATase_N, PF06426) and the bacterial transferase hexapeptide (PF00132) as their common motif. The SATase_N motif participates in enzyme activity, and the other motif is crucial for the formation of the CSC [29]. SAT not only plays an important role in the synthesis of OAS but also shows resistance to heavy metals such as Ni, Co, Zn, etc., resulting in the accumulation of GSH and increasing antioxidant stress resistance [30]. In addition, SAT is coexpressed with SMT, promoting an increase in OAS, as well as the expression of sulfur transporters and the absorption and accumulation of selenium by plants [31].

The other important member of the cysteine synthase complex (CSC) is O-acetyl serine (thiol) lyase (OASTL), which catalyses the synthesis and degradation of Sec used OAS and selenide with pyridoxal 5'-phosphate (PLP) as a cofactor. OASTL has a molecular weight of 60–70 kDa and usually acts in the form of a homodimer, which is ubiquitous in plant and microbial cellss [32]. The different isoenzymes of OASTL are located in different organelles and play various important functions. Members of the OASTL family are involved in the synthesis and degradation of L-Cys and L-Sec, regulating the concentration of intracellular Cys and Sec homeostasis in plants [33]. Among the nine OASTL family members in *A. thaliana*, the main members of OASTL include cytoplasmic OASTLA1, chloroplast OASTLB, and mitochondrial OASTLC, which are involved in cysteine synthesis [34]. It was shown that in *A. thaliana*, active OASTLA and OASTLB can improve selenium resistance by degrading L-Cys and L-Sec [35]. AtCYSC1, AtCYSD1, and AtCYSD2 have very high homology with β-cyano-alanine synthase and scavenge cyanide toxicity, and AtCS26 and AtCS-like synthesize S-sulfonylcysteine [36, 37]. In addition, OASTL participates in the secondary metabolism of plants and plays an important role in plant growth and development and resistance to abiotic stresses [38].

This study aims to identify the genome of ChSAT/OASTL genes, perform biological information analysis (including sequence, motif, phylogenetic, domain, and protein modelling), and clarify the physical and chemical properties and basic functions of ChSAT/OASTL gene family members. In addition, qRT–PCR was used to identify the main gene from the SAT and OASTL family of *C. hupingshanensis* that reacted to selenite stress. Finally, the molecular docking simulation of

OASTL's affinity to selenium substrates was used to provide a molecular theoretical basis for plant selenium metabolism and Sec biological synthesis.

2. Methods

2.1. Genome-wide ChSAT/OASTL family identification

The genome and annotation file of *C. hupingshanensis* were obtained from the Genome Warehouse BIG Data Center (https://ngdc.cncb.ac.cn/gwh/) under accession number PRJCA005533. To identify the SAT/OASTL gene members of *C. hupingshanensis*, we downloaded the nucleotide and protein sequences of the AtSAT/OASTL genes from the *Arabidopsis* Information Resource (https://www.arabidopsis.org/) as reference sequences and searched for the most similar protein sequences of ChSAT/OASTL using the blast region function of TBtools software [39]. To further confirm whether the extracted protein sequences of ChSAT/OASTL were correct on the NCBI BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi) website, CD-Search (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) was used to analyse conserved domains for ChSAT/OASTL protein sequences.

2.2. Physicochemical properties and phylogenetic analysis of ChSAT/OASTL proteins

Molecular weight (MW), isoelectric point, and other physicochemical property of the ChSAT/OASTL proteins were analysed via ExPASy (https://web.expasy.org/protparam). WoLF PSORT (https://wolfpsort.hgc.jp) was subsequently used to predict subcellular localization.

To further understand the evolutionary relationships of the SAT/OASTL family, the SAT/OASTL family protein sequences of different plants were downloaded from the NCBI website, with 5 SAT and 9 OASTL from *A. thaliana*, 4 SAT and 9 OASTL from *Brassica napus*, 5 SAT and 6 OASTL from *Glycine max*, 5 SAT and 9 OASTL from *Oryza sativa*, and 4 SAT and 12 OASTL from *Zea mays*. The ChSAT and ChOASTL family protein sequences were compared with protein sequences from different plants separately using MEGAX software and the evolutionary tree was constructed by the maximum likelihood method.

2.3. Chromosome location and structure analysis of ChSAT/OASTLs

Information regarding the location of ChSAT/OASTL genes on the chromosome was obtained by analysing the gene annotation file of *C. hupingshanensis*, and the gene location visualization (advance) function of TBtools software was used to draw a location map of the gene on the chromosome. Based on the genome annotation file and obtained NCBI CD-Search conserved domain information, we displayed the ChSAT/OASTL gene structure through the Gene Structure View (advance) of TBtools. For motif patterns, ChSAT/OASTL proteins were submitted and scanned for conserved motifs on the MEME website (http://meme-suite.org/tools/meme) with the MEME motif set to ten. TBtools draws the intron–exon structure of ChSAT/OASTL based on the information extracted from the GFF file of *C. hupingshanensis*. The AtSAT/OASTL and ChSAT/OASTL protein sequences were aligned by ClustalW (https://www.genome.jp/tools-bin/clustalw), and their alignment results were aligned by ESPript 3.0 (https://espript.ibcp.fr/ESPript/cgi-bin/ESPript.cgi) for further processing to output the image.

2.4. Homology modelling and verification of ChSAT/OASTL proteins

The ChSAT/OASTL protein sequences were submitted to SOPMA (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html) for secondary structure prediction. In addition, potential templates for different proteins were searched in the SWISS-MODEL template library. Based on the high similarity score, the best crystal structures were selected as templates, and the ChSAT/OASTL protein sequences were submitted to the SWISS-MODEL web server

(https://swissmodel.expasy.org/) for homology modelling. The protein model quality of the final 3D model of ChSAT/OASTL was authenticated by parameters such as GMQE and QMEANDisCo global [40].

2.5. Ligand Preparation and Molecular Docking

The structures of the ligand compounds (OAS, hydrogen sulfide, hydrogen selenide, selenophosphate) used in this study were obtained from the ChemSpider database, and the 3D structures were plotted with ChemSketch and saved in Protein Data Bank format. Prankweb was used to predict protein active sites [41]. ChOASTL protein and ligand compound docking experiments were performed using AutoDock v4.2 and AutoDock Vina v1.1.2 to calculate the Gasteiger charge of protein–ligand interactions [42].

First, the ChOASTL protein-PLP-OAS ternary complex PDB file was formed by docking the binding pocket of each ChOASTL protein molecule to the cofactor PLP and then docking it to the substrate OAS. Then, hydrogen sulfide, hydrogen selenide, and selenophosphate were each docked with the ChOASTL protein-OAS binary complex PDB file vina to form the ChOASTL protein-PLP-OAS-selenophosphate tetrameric complex PDB file. The interactions (hydrogen bonding and hydrophobic interactions) of the ChOASTL protein-OAS-selenophosphate complexes were analysed and visualized using PILP and PyMol, and heatmaps of the docking binding energies were generated using GraphPad Prism version 9.0.0 (GraphPad Software, San Diego, California USA, www.graphpad.comprism).

2.6. RNA isolation and gene expression analysis

C. hupingshanensis seeds were sourced from the Yutangba Colour Mine in Enshi, Hubei Province, China, and were cultivated in our laboratory after seed collection. The seedlings were cultivated at a temperature of approximately 22°C and in a pattern of alternating 12 hours of light and 12 hours of darkness in the plant culture chamber. Thirty-six seedlings of approximately 10 cm in height at the ten-leaf stage were selected, and the roots were flushed with roseite and equalized in Hoagland solution for two days. The samples were treated with different selenium concentrations (0 μg Se L^{-1}, 100 μg Se $L^{-1,}$ and 80,000 μg selenium L^{-1}), with the experimental samples treated with 0 μg Se L^{-1} as the control group. Leaves and roots of three seedlings were isolated from each treatment at 0 h, 3 h, 6 h, 12 h, and 24 h, and these samples were then snap-frozen in liquid nitrogen to extract their RNA.

Total RNA was extracted from root and leaf samples by the *TransZol™* Up Plus RNA Kit and detected by a NanoDrop 2000 for RNA concentration and quality. RNA integrity and gDNA contamination were examined by 1.0% agarose gel electrophoresis. RNA samples were treated with RNase-free DNase to remove residual genomic DNA. Real-time qPCR on ABI StepOne Plus was performed. The expression of target genes in the samples was detected using the Hieff qPCR SYBR Green Mix commercial kit, and gene expression was calculated using the $2^{-\Delta\Delta CT}$ method [43]. The results were analysed, and graphical representation was carried out using GraphPad Prism version 9.0.0, GraphPad Software, San Diego, California USA, www.graphpad.comprism. All analyses were performed in triplicate. The primers used for the qRT—PCR analysis of ChSAT/OASTL are listed in Table S2.

3. Result

3.1. Identification and analysis of SAT/OASTL genes in *C. hupingshanensis*

Eight members of the ChSAT gene family were found in the genome of *C. hupingshanensis* (the Genome Warehouse BIG Data Center accession number PRJCA005533) by comparison with the genome sequences of *A. thaliana*, and the gene names of ChSAT family members were determined by sequence homology of AtSAT. Detailed information about the members of the ChSAT gene family, including the number of amino acid residues in the protein, theoretical isoelectric point, molecular weight, and other related physicochemical property indices, is given in Table 1. The amino acid residues range

from 262 to 398, the molecular weight ranges from 27.95 to 52.23 kDa and the isoelectric point values range from 5.11 to 8.35 of the ChSAT family, and most ChSAT proteins are in the faintly acidic or neutral cellular environment in *C. hupingshanensis* except for ChSAT4. The members of the ChSAT protein family, in addition to ChSAT1;1 and ChSAT1;2, are presumed to be stable proteins by the prediction of instability coefficients. The results of subcellular localization analysis showed that proteins of the ChSAT gene family were distributed in the cytoplasm except for the ChSAT3 protein, which was located in the mitochondrion.

Table 1
Identification of basic physicochemical properties of ChSAT gene family members

Cene ID	Gene name	Amino acid (aa)	Molecular weight (kDa)	Isoelectric point (pI)	Instability index	Subcellular localization
Chu003685	ChSAT1;1	312	34.22	7.18	45.33	Cytoplasm
Chu013066	ChSAT1;2	312	34.11	7.16	41.65	Cytoplasm
Chu020773	ChSAT2;1	262	27.95	6.49	30.35	Cytoplasm
Chu050052	ChSAT2;2	323	34.78	5.78	36.51	Cytoplasm
Chu049653	ChSAT3	398	42.68	5.11	39.01	Mitochondrion
Chu026957	ChSAT4	393	52.23	8.35	38.85	Cytoplasm
Chu011818	ChSAT5;1	309	32.46	6.92	38.44	Cytoplasm
Chu040697	ChSAT5;2	311	32.64	6.89	39.27	Cytoplasm

A total of thirteen members of the OASTL gene family were characterized in the *C. hupingshanensis* genome by homology searches. The ChOASTL proteins ranged from 322 to 430 amino acid residues, molecular weights from 32.01 to 45.96 kDa and isoelectric point values from 5.29 to 8.77 (Table 2). A pI of ChOASTL lower than 7.0 would be similar to the cytoplasmic pI, in contrast to the chloroplast and mitochondrial pI, and similar phenomena were observed in *A. thaliana* and *Sorghum bicolor* [44]. These pI values may indicate the location of OASTL proteins in the cell. The members of the ChOASTL protein family, in addition to ChOASTLB1, ChOASTLB2 and ChOASTLC1, are presumed to be stable proteins by the prediction of instability coefficients. The coding sequences and protein sequences of ChOASTL/SAT genes are shown in Table S1.

Table 2
Identification of basic physicochemical properties of ChOASTL gene family members

Cene ID	Gene name	Amino acid (aa)	Molecular weight (kDa)	Isoelectric point (pI)	Instability index	Subcellular localization
Chu019397	ChOASTLA1;1	322	33.89	5.9	27.5	Cytoplasm
Chu029016	ChOASTLA1;2	322	34.11	5.9	29.41	Cytoplasm
Chu048695	ChOASTLA2	342	36.45	8.25	31.37	Chloroplast
Chu035096	ChOASTLB1	394	41.94	8.38	42.31	Chloroplast
Chu007319	ChOASTLB2	394	41.91	8.68	44.75	Chloroplast
Chu046136	ChOASTLC1	430	45.96	8.78	42.19	Chloroplast
Chu025039	ChOASTLC2	429	45.82	8.73	39.53	Chloroplast
Chu045982	ChCYSC1;1	368	39.76	8.77	29.95	Mitochondrion
Chu026205	ChCYSC1;2	368	39.9	8.11	29.63	Mitochondrion
Chu037773	ChCYSD2;1	323	34.36	5.65	31.13	Cytoplasm
Chu044352	ChCYSD2;2	338	36.59	5.71	36.49	Cytoplasm
Chu023729	ChCS26	405	43.58	7.75	39.09	Chloroplast
Chu037774	ChCS-like	302	32.01	5.29	25.28	Chloroplast

3.2. Chromosomal localization and structure analysis of SAT/OASTLs in *C. hupingshanensis*

The distribution of ChSAT and ChOASTL genes on the sixteen chromosomes of *C. hupingshanensis* is relatively scattered (Fig. 3). Except for chromosome 2, which contains three genes (ChOASTLA2, ChSAT2;2 and ChSAT3), the chromosomes of *C. hupingshanensis* contain one or two ChSAT and ChOASTL genes. A close linkage was found between ChCYSD2 and ChCS-like, which are in the middle of chromosome 10, indicating that members of the ChOASTL gene family have experienced tandem repeats during evolution.

A simple maximum likelihood phylogenetic tree was constructed for the protein sequences of SAT and OASTL of *C. hupingshanensis* to adequately identify the protein motifs, conserved structural domains and gene structures (Fig. 4a, 4e). The online software MEME was used to search the conserved motifs of ChSAT/OASTLs, and ten conserved motifs with diverse architecture for the ChSAT/OASTLs protein sequences were found (Fig. 4b, 4f). The results indicated that motifs 1 to 7 predominantly occurred in most ChSATs, whereas motif 5 was not frequently found in ChSAT2;1. In addition, motif 10 is distinctive to ChSAT5;1 and ChSAT5;2, and motif 9 is also characteristic of ChSAT1;1 and ChSAT1;2. There are eight mainly conserved motifs found in the ChOASTL protein sequences, which are motifs 1, 2, 3, 4, 5, 6, 7, and 8, but motif 6 and motif 4 were not present in ChCS26, and motif 4 was not present in ChCS-like. Moreover, motif 8 was characteristic of the four members of the ChOASTL family, including the ChOASTL B1, ChOASTLB2, ChOASTLC1, and ChOASTLC2 proteins, which might explain the benign variation in the ChSAT/OASTL family during evolution.

The analysis of ChSAT domains using the NCBI CD search server revealed that all proteins contained the conserved serine acetyltransferase N-terminal domain structure (SATase_N, PF06426) and the bacterial transferase hexapeptide (PF00132) domain, and a conserved pyridoxal phosphate-dependent enzyme domain (PF00291, PALP) was shared in all ChOASTL proteins, which could indicate that these enzymes use pyridoxal monohydrate 5'-phosphate (PLP) as a cofactor (Fig. 4c, 4g).

The gene structures of the coding sequences and untranslated regions (UTRs) of the SAT and OASTL genes in *C. hupingshanensis* were visualized using TBtools software, which revealed that the homologous genes from different groups had similar intron or exon numbers and distributions. These results demonstrate that the same subclade of ChSAT/OASTL genes was functionally and structurally conserved, but there were differences between the different subclades (Fig. 4d, 4h).

3.3. Phylogenetic analysis of SAT/OASTLs in *C. hupingshanensis*

To characterize the phylogenetic relationships of SAT and OASTL in different plant species and predict the function of ChSATs and ChOASTLs, 8 ChSAT protein sequences and 13 ChOASTL protein sequences from *C. hupingshanensis* (*Ch*) were used to construct phylogenetic trees with other plants. These sequences included 9 SAT and 18 OASTL protein sequences from *Arabidopsis thaliana* (*At*) and *Brassica napus* (*Bn*) in Brassicaceae, 5 SAT and 6 OASTL protein sequences from *Glycine max* (*Gm*) in Fabaceae, and 9 SAT and 21 OASTL protein sequences from *Oryza sativa* (*Os*) and *Zea mays* (*Zm*) in poaceae (Fig. 5a, 5b). Based on the results of phylogenetic analysis, 8 ChSAT proteins were divided into Groups I ~ III, ChSAT1;1, ChSAT1;2 and ChSAT3 proteins were aggregated in Group I, ChSAT2;1, ChSAT2;2 and ChSAT4 proteins were aggregated in Group II, and ChSAT5;1 and ChSAT5;2 proteins were aggregated in Group III. The same analysis was used for the ChOASTL family, and the results showed that 13 ChOASTL proteins were divided into Groups A ~ D, ChOASTLA1;1, ChOASTLA1;2 and ChOASTLA2 proteins were aggregated in Group A, ChCYSC1;1 and ChCYSC1;2 proteins were aggregated in Group B, ChOASTLB1, ChOASTLB2, ChOASTLC1, ChOASTLC2 and ChCS26 proteins were aggregated in Group C, and ChCYSD2;1, ChCYSD2;2 and ChCS-like proteins were aggregated in Group D. The phylogenetic tree showed that the ChSAT/ChOASTL proteins were more closely related to the AtSAT/AtOASTL proteins, possibly because both *C. hupingshanensis* and *A. thaliana* are dicotyledonous Brassicaceae plants, which exhibit evolutionarily close kinship. Based on the comprehensive phylogenetic analysis, the SAT/OASTL proteins were present in every subgroup of monocotyledonous and dicotyledonous plants; that is, the SAT/OASTL gene family appeared before the differentiation of monocotyledonous and dicotyledonous plants.

3.4. Functional characteristics analysis of ChSAT/OASTL proteins

The alignment of protein sequences in SAT/OASTL processed by ClusterW software showed that the conserved serine acetyltransferase N-terminal domain structure (SATase_N, PF06426) was located in the range of ^{49}W-V^{128}, while the bacterial transferase hexapeptide domain structure (hexapeptide, PF00132) was located in the range of ^{200}G-G^{222} and ^{235}I-V^{265} (Fig. 6, Fig. S1). SATase_N and Hexapeptide are common domains in most SATs that are characteristic of the enzyme. However, a conserved motif devoid of 23 residues was also found in the N-terminus of the Hexapeptide of ChSAT2;1. Previous research has shown that the highly conserved β1c-β2c loop, which we defined as the β7-β8 loop, plays a key role in the binding of SAT proteins to serine [45]. In addition, the 10 residues of the C-terminus of SAT are used to combine the active site of OASTL and form the cysteine synthase complex [46, 47].

According to the conserved motif analysis of OASTL proteins, the highly conserved PLP binding site was present in all ChOASTL sequences and consisted of the amino acid residues Lys^{46}, His^{157}, Gly^{181}, Thr^{182}, Gly^{183}, Thr^{185}, and Ser^{269} (Fig. 7, Fig. S2); furthermore, binding energy originates from the hydrogen bond between cofactor and the enzyme [48]. Because the PLP binding motif is absent, AtOASTLA2 was considered a nonfunctional pseudogene [49]. The substrate-binding site of the members of OASTL also contained another important motif (^{74}TSGNT^{78} loop), which was highly conserved in the three major isoforms of the OASTL family, including OASTLA, OASTLB, and OASTLC. The Ser residue of the loop was substituted by Aln in AtCYSD2, ChCYSD2;1 and ChCYSD2;2, by Thr in AtCS26 and ChCS26, and by Gly in AtCS-like and ChCS-like. The two key amino acids Thr^{74} and Ser^{75} of the ^{74}TSGNT^{78} loop combine with the sulfide or selenide of the substrate, and this site mutation leads to a reduction in the efficiency of cysteine and selenocysteine synthesis [50].

3.5. Secondary and tertiary structure prediction of SAT/OASTL proteins in *C. hupingshanensis*

Analysis of the secondary structures of ChSAT/OASTL using SOPMA online outlets showed that they were enriched in the alpha helix (33.76%-47.13%), extended strand (16.50%-24.43%) and random coil (27.56%-43.22%) and contained small amounts of beta turn (6.28%-11.99%), indicating some variability in protein structures on ChSAT/OASTLs (Table 3). To better visualize the tertiary structure, ChSAT/OASTL protein sequences were submitted to the SWISS-MODEL online server and blast-searched for the protein template with the best match. The serine acetyltransferase (SMTL ID: 4n69.1. A) of soybean was used as the template for crystal structures in the homology modelling of AtSAT and ChOASTL, and the predicted tertiary structures of AtSAT and ChSAT are shown in Fig. 8. Similarly, the crystal structure of the *A. thaliana* O-acetyl serine (thiol) enzyme (SMTL ID: 4aec.1. A) was chosen as the template for homology modelling of AtOASTL and ChOASTL, and the predicted tertiary structures of AtOASTL and ChOASTL are shown in Fig. 9. The quality of the 3D structural models of ChSAT/OASTLs predicted by SWISS-MODEL used metrics of sequence identity, coverage, and global model evaluation, and the results are presented in Table 3. The sequence identity of the ChSAT/OASTL protein models was above 55.00%, and the coverage reached above 0.66. The global model quality evaluation (GMQE) and QMEANDisCo global give an overall quality measure between 0 and 1, with larger numbers indicating a higher expected quality. The valuation results of the quality of the ChSAT/OASTL 3D structures showed that the quality of the models provided by SWISS-MODEL met our expectations and could be used in the next study.

Table 3
Quality parameters of ChSAT/OASTL protein models and secondary structure analysis.

Gene	Template	Sequence identity	Coverage	GMQE	QMEAN	Alpha helix	Beta turn	Extended strand	Random coli
ChSAT1;1	4n69.1.A	60.07%	0.89	0.74	0.84	39.10%	7.05%	18.91%	34.94%
ChSAT1;2	4n69.1.A	61.09%	0.88	0.74	0.84	42.95%	8.01%	21.47%	27.56%
ChSAT2;1	4n69.1.A	55.94%	0.77	0.65	0.79	47.13%	6.90%	17.62%	28.35%
ChSAT2;2	4n69.1.A	55.94%	0.73	0.63	0.83	35.29%	8.67%	19.81%	36.22%
ChSAT3	4n69.1.A	60.36%	0.70	0.56	0.82	33.92%	6.28%	16.58%	43.22%
ChSAT4	4n69.1.A	55.00%	0.66	0.51	0.73	37.66%	7.89%	24.43%	30.03%
ChSAT5;1	4n69.1.A	86.23%	0.89	0.79	0.89	44.34%	7.44%	17.80%	30.42%
ChSAT5;2	4n69.1.A	86.59%	0.89	0.78	0.89	45.43%	7.72%	18.33%	28.06%
ChOASTLA1;1	4aec.1. A	72.50%	0.98	0.91	0.88	40.37%	10.87%	18.94%	29.81%
ChOASTLA1;2	4aec.1. A	69.30%	0.98	0.90	0.88	40.37%	10.56%	17.70%	31.37%
ChOASTLA2	4aec.1. A	68.34%	0.93	0.83	0.82	37.72%	11.99%	19.30%	30.99%
ChOASTLB1	4aec.1. A	82.34%	0.98	0.83	0.86	35.79%	9.14%	17.51%	37.56%
ChOASTLB2	4aec.1. A	82.99%	0.87	0.79	0.86	33.76%	8.88%	16.50%	40.86%
ChOASTLC1	4aec.1. A	92.25%	0.99	0.79	0.87	38.37%	9.77%	18.37%	33.49%
ChOASTLC2	4aec.1. A	92.72%	0.99	0.79	0.86	41.03%	7.93%	17.25%	33.80%
ChCYSC1;1	4aec.1. A	60.69%	0.86	0.78	0.86	38.59%	9.78%	16.58%	35.05%
ChCYSC1;2	4aec.1. A	60.38%	0.86	0.78	0.85	37.50%	9.78%	18.21%	34.51%
ChCYSD2;1	4aec.1. A	67.29%	0.99	0.89	0.85	39.94%	10.84%	19.81%	29.41%
ChCYSD2;2	4aec.1. A	61.37%	0.95	0.80	0.79	37.87%	11.24%	20.71%	30.18%
ChCS26	4aec.1. A	63.06%	0.78	0.67	0.82	35.06%	10.37%	20.00%	34.57%
ChCS-like	4aec.1. A	64.55%	0.99	0.88	0.84	35.43%	10.26%	20.20%	34.11%

3.6. Molecular docking

Molecular docking is a bioinformatics-based theoretical simulation method that is commonly used to study the interaction forces between molecules (protein receptors and ligands) and to predict their binding modes and affinity at the minimum energy state using a computer program [51]. In recent years, molecular docking techniques have been widely used in various fields, such as molecular biology, medicinal chemistry, and food science, and exhibit several advantages, including being relatively fast and economical [52, 53]. The members of the ChOASTL protein were predicted with eight potential binding pockets and numbered from highest to lowest probability score of the pocket using the Prankweb online server. Molecular docking was performed for the potential binding pocket of each ChOASTL using the Auto Dock Vina program, and the

protein–ligand docking binding energy was recorded. PyMOL was used to export the protein–ligand complexes by PDB files, which were imported into PILP (Protein–Ligand Interaction Profiler) for visualization and to analyse the protein–ligand complexes [54, 55]. The sulfide or selenide of substrates were used as ligands for molecular docking with ChOASTL, and the results showed that the interaction energy between sulfide or selenide and ChOASTLs ranged from − 0.8 to − 0.5 kcal·mol^{-1}, which may indicate that sulfide or selenide compounds experience difficulty when binding to ChOASTL molecules. When selenophosphate was used as a ligand for molecular docking, satisfactory results were obtained, showing that the binding energy was between − 4.4 and − 2.3 kcal.mol^{-1}; thus, selenophosphate may be a potential substrate for interaction with ChOASTL protein molecules.

The web of PLIP was used to process visualized protein–ligand complexes and gain insight into the protein–ligand binding mode constructed by hydrogen bonds, salt bridges and hydrophobic interactions from amino acid residues, substrates (including selenophosphate and OAS) and cofactor PLP (Fig. 11). The hydrogen bond is necessary for interactions between the ternary complex ChOASTL-PLP-OAS with selenophosphate. In this study, the highest affinity for substrates was found at two conserved sequences of ChOASTL, which included the PLP binding site and TSTNT loop and the catalytic site. At the catalytic site, the ligand was surrounded by Lys^{46}, Gly^{181}, Thr^{182}, Gly^{183}, Thr^{185}, and Ser^{269}, conserved amino acid residues that are binding sites for the cofactor PLP involved in OAS and cysteine synthesis [26]. The ligands were surrounded by the amino acid residues of the TSTNT loop at the catalytic sites of most ChOASTL and formed hydrogen bonds. However, the ligands were surrounded by one or two amino acid residues of the TSTNT loop in the ChCYSC1;1, ChCYSC1;2, and ChCYSD1;1 proteins. This phenomenon might be explained by the various abilities of ChOASTL family members to bind to selenophosphate at the catalytic site.

3.7. Expression analysis of ChSAT/OASTL genes in different tissues under selenium stress

qRT–PCR technology was used to analyse the expression and investigate the response of ChSAT/OASTL family members when the seedlings of *C. hupingshanensis* were treated with selenite. In these tests, the selenium concentration was 100 μg Se L^{-1} and 80,000 μg Se L^{-1}, the control group was not treated with selenite, and plant tissues included leaves and roots. As the seedlings of *C. hupingshanensis* treated with 100 μg Se L^{-1} selenite arrived at 24 hours, the expression of all ChSAT family members was significantly upregulated in leaves, ChSAT3 was highly upregulated by approximately 9.1-fold and ChSAT5;1 was upregulated by approximately 4.8-fold. The expression levels of other genes were not obviously upregulated or slightly downregulated at other times (Fig. 12a). On the other hand, the expression of ChSAT1;2 in roots was shown to be highly upregulated at 3 h, 6 h, 12 h, and 24 h, with a significant upregulation of approximately 6.3-fold at 12 h (Fig. 12c),and the expression levels of the remaining ChSAT were not changed significantly.

When the seedlings of *C. hupingshanensis* were treated with 80,000 μg Se L^{-1} selenite at 3 h and 24 h, the expression of ChSAT family members was upregulated in leaves, the ChSAT1;1 and ChSAT3 genes were upregulated approximately 5.3- and 3.8-fold at 24 h, respectively, and the ChSAT1;2 gene was upregulated by approximately 3.4-fold at 3 h (Fig. 12b). For the ChSAT family members in roots, ChSAT1;2 was significantly upregulated by approximately 9.4-fold, and the expression levels of ChSAT2;1, ChSAT2;2, ChSAT4 and ChSAT5;1 appeared to be downregulated at 3 h, 6 h, 12 h, and 24 h (Fig. 12d).

The expression levels of the ChOASTL family were examined under selenium stress using the same method. As the seedlings of *C. hupingshanensis* treated with 100 μg Se L-1 selenite arrived at 24 h, the ChOASTLA1;2 gene was highly upregulated by approximately 7.4-fold in leaves, and the expression levels of ChCS-like were highly upregulated at 3 h, 6 h, and 12 h, with a significant upregulation of approximately 8.8-fold at 12 h (Fig. 13a). For ChOASTL family members in the roots, ChOASTLA2 was significantly upregulated by approximately 6.3-fold at 3 h (Fig. 13c).

When the seedlings of *C. hupingshanensis* were treated with 80,000 μg Se L^{-1} selenite, the expression of most ChOASTL members in leaves was upregulated (Fig. 13b). Compared to the other ChOASTL, ChCS-like was more significantly

upregulated, with an upregulation of approximately 6.9-fold at 3 h, 9.8-fold at 6 h and 6.0-fold at 12 h, which indicated that it exhibits a high selenium tolerance. The upregulation of ChOASTL genes in leaves was more obvious than that in roots under selenium stress (Fig. 13). Based on these data, members in the ChSAT and ChOASTL families may play an important role in the response to selenium.

4. Discussion

In our study, the alignment of members in the SAT and OASTL families presents conserved functions, and the greater number of members exhibit redundancy. The analysis of conserved motifs and gene structures of the ChSAT/OASTL gene family showed that the conserved amino acid sequences and gene structures of members in the same subclade were similar, which was consistent with the results of *Solanum lycopersicum*, *Sorghum bicolor* and *Oryza sativa* [44, 56, 57]. However, the number of members in the SAT gene family varies significantly between plants as follows: 8 ChSAT in *C. hupingshanensis*, 5 AtSAT in *Arabidopsis thaliana*, 6 OsSAT in *Oryza sativa*, and 6 VvSAT in *Vitis vinifera* [58]. The OASTLs in plants also have multiple isozymes, and the number of members of the OSATL gene family also varies between plants (13 ChOASTL in *C. hupingshanensis*, 9 AtOASTL in *Arabidopsis thaliana*, 9 OsSAT in *Oryza sativa*, 9 SlOASTL in *Solanum lycopersicum*, 10 PtOASTL in *Populus trichocarpa*, 7 SbOASTL in *Sorghum bicolor* and 6 GmOASTL in *Glycine max*) [50]. Therefore, we presume that the gene and function redundancy of the ChSAT and ChOASTL family of *C. hupingshanensis* originate from the harsh primary environmental region for plants, survival conditions including high altitude, lower growth temperature, multiple heavy metal ions (such as cadmium and mercury ions) and anions (such as fluorine, sulfate and sulfite, selenate and selenite) and other hazardous substances.

The synthesis of Sec is completed by the balance between dissociation and formation of the cysteine synthase complex. The form of the cysteine synthase complex occurred under selenium/sulfur sufficiency conditions, and when the concentration could not satisfy the needs of the plant cell, the complex uncoupled to release free SAT and OASTL (Fig. 14) [45]. The 10 peptides on the C-terminus of SAT occupied the OAS binding site located on the N-terminus of OASTL with the conserved sequence TNSTNT, which caused Sec and Cys to unform and trigger the accumulation of OAS. When the cellular OAS concentration reached some point, OAS competed with the binding site, which dissociated the complex, and then the free OASTL catalysed Sec and Cys formation. Once the cellular selenium and sulfur concentrations decline to some point, the reaction for the synthesis of Sec and Cys stops, and the cysteine synthase complex remodels. In *A. thaliana*, the formation of the CSC complex was dominated by the interaction of key active site residues Gln^{147}, and Thr^{74} and Ser^{75} in the TSTNT loop, a conserved motif of the OASTL protein, with the C10 peptide of the SAT protein [59]. The mutation in the TSTNT loop of ChOASTL members resulted in the inability of ChOASTL to synthesize Sec and Cys and to form a complex with SAT.

The concentration of Sec of in *C. hupingshanensis* in vivo was regulated by members of the OASTL gene family, which could maintain the dynamic balance of the concentration of Sec and may function a protective mechanism to reduce the toxicity of selenium in plants. In previous studies, the three OASTL isoforms OASTLA1, OASTLB, and OASTLC, which are most highly expressed in *A. thaliana* cells, catalyse the synthesis of cysteine; in contrast, CS-like primarily catalyses the degradation of cysteine rather than its synthesis and catalyses the formation of sulfide, ammonia and pyruvate in a 1:1:1 stoichiometric ratio to maintain the dynamic balance of intracellular cysteine [24, 35, 60].Due to the chemical similarity of S and Se, the expression of different OASTL family genes may regulate the dynamic homeostasis mechanism of the product Sec concentration. In addition, the results of molecular docking suggest that OASTL family proteins exhibit a higher affinity for the ligand selenophosphate; therefore, selenophosphate may be the optimal substrate for the synthesis of Sec (Figs. 9, 10). In our study, the gene members of ChOASTL showed a high expression level under selenium stress, especially the expression of ChOASTLA1;1, ChOASTLA1;2 and ChCS-like genes, which were upregulated at high times (Fig. 13). This result possibly suggests that ChOASTLA-type and CS-like were regulated together by the concentration of Sec in vivo. In the major OASTL isoforms, the highly conserved ^{74}TSTNT^{78} loop acts as the active centre for cysteine or Sec synthesis by doping with sulfide or selenide [48]. Sequence comparisons show that the ^{75}Ser of the ^{74}TSTNT^{78} loop was replaced by Gly in CS-

like (Fig. 6), which eliminated an important group and caused a change in spatial conformation and function; furthermore, the function of CS-like was changed to degrade Sec or Cys [61]. In contrast, the residue responsible for anchoring PLP is highly conserved in all encoded OASTL-like proteins. Analogous to pathways in animal cells, the results should indicate that CS-like participated in regulating the concentration of Sec and Cys in vivo and selenoprotein biosynthesis [62].

5. Conclusions

This study provides new insights into the function and tolerance of the SAT and OASTL genes under selenium stress in *C. hupingshanensis*. Furthermore, we used bioinformatics to screen ChSAT and OASTL genes in *C. hupingshanensis* to determine the expression profile of ChSAT/OASTL genes under selenium stress and the 3D structure prediction of ChSAT/OASTL to obtain deeper knowledge. The results of gene expression and molecular docking indicated that the ChSAT and ChOASTL genes were upregulated under selenium stress, and OASTL family genes could synthesize and degrade cysteine/selenocysteine, which provides a theoretical basis for selenocysteine synthesis.

Declarations

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

Finding: 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 (T2020020), and the Open Foundation of Hubei Engineering Research Center of selenium food nutrition and health intelligent technology (PT082102) to Yifeng Zhou. The research was mainly finished at the Hubei Key Laboratory of Biological Resources Protection and Utilization.

Author Contributions: Y. Z and Q. T designed the research. Y.C., Y. L. and Z. H prepare plant materials. Y.C, Q X, and Y. Z completed the bioinformatics analysis of *C. hupingshanensis* SAT/OASTL genes. Y.L, Y.C, and Z. X isolated RNA and analyzed the differential expression of the gene. Y. C, Y. Z, and C. L wrote the main manuscript text and draw the figures. All authors reviewed the manuscript.

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*, were sourced from the Yutangba Colour Mine in Enshi, Hubei Province, China, and were cultivated in our laboratory after seed collection, in compliance with ethical standards and regulations. 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.

Conflicts of Interest: The authors declare no conflict of interest.

Data Availability Statement: The datasets generated and/or analysed during the current study are available in the Genome Warehouse BIG Data Center (https://ngdc.cncb.ac.cn/gwh/) under accession number PRJCA005533.

Supplementary Materials: Fig. S1: Multiplexed alignment of full sequences of SAT protein in *C. hupingshanensis*; Fig. S2: Multiplexed alignment of full sequences of OASTL protein in *C. hupingshanensis*; Table S1: The coding sequences and protein sequences of ChSAT/OASTL genes; TableS2: The primers of ChSAT/OASTL for qRT-PCR.

References

1. Schomburg L, Schweizer U, Köhrle J: Selenium and selenoproteins in mammals: extraordinary, essential, enigmatic. Cellular Molecular Life Sciences CMLS 2004, 61:1988-1995.
2. Jia Y, Wang R, Su S, Qi L, Wang Y, Wang Y, Zou Y, Liu X, Zhang Y, Hou J: A County-Level Spatial Study of Serum Selenoprotein P and Keshan Disease. Frontiers in Nutrition 2022, 9:47.
3. Wang K, Yu J, Liu H, Liu Y, Liu N, Cao Y, Zhang X, Sun D: Endemic Kashin-Beck disease: A food-sourced osteoarthropathy. Seminars in Arthritis and Rheumatism 2020, 50(2):366-372.
4. Tindell R, Wall SB, Li Q, Li R, Dunigan K, Wood R, Tipple TE: Selenium supplementation of lung epithelial cells enhances nuclear factor E2-related factor 2 (Nrf2) activation following thioredoxin reductase inhibition. Redox biology 2018, 19:331-338.
5. Brigelius-Flohé R, Flohé L: Regulatory phenomena in the glutathione peroxidase superfamily. Antioxidants redox signaling 2020, 33(7):498-516.
6. Chambers I, Frampton J, Goldfarb P, Affara N, McBain W, Harrison PR: The structure of the mouse glutathione peroxidase gene: the selenocysteine in the active site is encoded by the 'termination'codon, TGA. The EMBO journal 1986, 5(6):1221-1227.
7. Kuršvietienė L, Mongirdienė A, Bernatonienė J, Šulinskienė J, Stanevičienė I: Selenium anticancer properties and impact on cellular redox status. Antioxidants 2020, 9(1):80.
8. Wallenberg M, Misra S, Björnstedt M: Selenium cytotoxicity in cancer. Basic clinical pharmacology toxicology 2014, 114(5):377-386.
9. Yang H, Yang X, Ning Z, Kwon SY, Li M-L, Tack FM, Kwon EE, Rinklebe J, Yin R: The beneficial and hazardous effects of selenium on the health of the soil-plant-human system: An overview. Journal of Hazardous Materials 2022, 422:126876.
10. Szőllősi R, Molnár Á, Oláh D, Kondak S, Kolbert Z: Uptake and Metabolism of Selenium in Plants: Recent Progress and Future Perspectives. Selenium Nano-Selenium in Environmental Stress Management Crop Quality Improvement 2022:79-90.
11. Wangeline AL, Valdez JR, Lindblom SD, Bowling KL, Reeves FB, Pilon-Smits EA: Characterization of rhizosphere fungi from selenium hyperaccumulator and nonhyperaccumulator plants along the eastern Rocky Mountain Front Range. American Journal of Botany 2011, 98(7):1139-1147.
12. Das S, Biswas AK: Comparative study of silicon and selenium to modulate chloroplast pigments levels, Hill activity, photosynthetic parameters and carbohydrate metabolism under arsenic stress in rice seedlings. Environmental Science and Pollution Research 2022, 29(13):19508-19529.
13. Kushwaha A, Goswami L, Lee J, Sonne C, Brown RJ, Kim K-H: Selenium in soil-microbe-plant systems: Sources, distribution, toxicity, tolerance, and detoxification. Critical Reviews in Environmental Science Technology 2022, 52(13):2383-2420.
14. Harvey M-A, Erskine PD, Harris HH, Brown GK, Pilon-Smits EA, Casey LW, Echevarria G, Van Der Ent A: Distribution and chemical form of selenium in *Neptunia amplexicaulis* from Central Queensland, Australia. Metallomics 2020, 12(4):514-527.
15. Lima LW, Castleberry M, Wangeline AL, Aguirre B, Dall'Acqua S, Pilon-Smits EAH, Schiavon M: Hyperaccumulator *Stanleya pinnata*: In Situ Fitness in Relation to Tissue Selenium Concentration. Plants 2022, 11(5):690.
16. Sors TG, Martin CP, Salt DE: Characterization of selenocysteine methyltransferases from Astragalus species with contrasting selenium accumulation capacity. The Plant Journal 2009, 59(1):110-122.
17. Chao W, Rao S, Chen Q, Zhang W, Liao Y, Ye J, Cheng S, Yang X, Xu F: Advances in Research on the Involvement of Selenium in Regulating Plant Ecosystems. Plants 2022, 11(20):2712.
18. Chi S: Differences in Selenium Uptake and Transport, and Related Gene Expression in Three Brassica Vegetables. Applied Ecology and Environmental Research 2018, 16(3):2781-2793.

19. Guo Z, Zhu B, Guo J, Wang G, Li M, Yang Q, Wang L, Fei Y, Wang S, Yu T: Impact of selenium on rhizosphere microbiome of a hyperaccumulation plant *Cardamine violifolia*. Environmental Science Pollution Research 2022, 29(26):40241-40251.
20. Cui L, Zhao J, Chen J, Zhang W, Gao Y, Li B, Li YF: Translocation and transformation of selenium in hyperaccumulator plant *Cardamine enshiensis* from Enshi, Hubei, China. Plant Soil 2018, 425(1):577-588.
21. Huang C, Ying H, Yang X, Gao Y, Li T, Wu B, Ren M, Zhang Z, Ding J, Gao J: The *Cardamine enshiensis* genome reveals whole genome duplication and insight into selenium hyperaccumulation and tolerance. Cell discovery 2021, 7(1):1-14.
22. 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*. Scientific reports 2018, 8(1):1-17.
23. Xiao Z, Lu Y, Zou Y, Zhang C, Ding L, Luo K, Tang Q, Zhou Y: Gene Identification, expression analysis and molecular docking of ATP sulfurylase in the selenization pathway of *Cardamine hupingshanensis*. BMC Plant Biology 2022, 22(1):1-17.
24. Wirtz M, Hell R: Functional analysis of the cysteine synthase protein complex from plants: structural, biochemical and regulatory properties. Journal of plant physiology 2006, 163(3):273-286.
25. Droux M: Plant serine acetyltransferase: new insights for regulation of sulphur metabolism in plant cells. Plant Physiology Biochemistry 2003, 41(6-7):619-627.
26. Hell R, Jost R, Berkowitz O, Wirtz M: Molecular and biochemical analysis of the enzymes of cysteine biosynthesis in the plant *Arabidopsis thaliana*. Amino acids 2002, 22(3):245-258.
27. Kawashima CG, Berkowitz O, Hell R, Noji M, Saito K: Characterization and expression analysis of a serine acetyltransferase gene family involved in a key step of the sulfur assimilation pathway in Arabidopsis. Plant Physiology 2005, 137(1):220-230.
28. Johnson CM, Roderick SL, Cook PF: The serine acetyltransferase reaction: acetyl transfer from an acylpantothenyl donor to an alcohol. Archives of biochemistry biophysics 2005, 433(1):85-95.
29. Yeon JY, Yoo SJ, Takagi H, Kang HA: A novel mitochondrial serine O-acetyltransferase, OpSAT1, plays a critical role in sulfur metabolism in the thermotolerant methylotrophic yeast *Ogataea parapolymorpha*. Scientific Reports 2018, 8(1):1-12.
30. Hasanuzzaman M, Bhuyan M, Anee TI, Parvin K, Nahar K, Mahmud JA, Fujita M: Regulation of ascorbate-glutathione pathway in mitigating oxidative damage in plants under abiotic stress. Antioxidants 2019, 8(9):384.
31. Çakır Ö, Turgut-Kara N, Arı Ş: Selenium metabolism in plants: molecular approaches. Advances in selected plant physiology aspects 2012:209-232.
32. Leustek T, Martin MN, Bick J-A, Davies JP: Pathways and regulation of sulfur metabolism revealed through molecular and genetic studies. Annual review of plant biology 2000, 51(1):141-165.
33. Assylay K, Aizat B, Aigerim S, Dinara O, Zhadyrassyn N, Sudhakar S, Poonam T, Kumar DA, Moshe S: Active O-acetylserine-(thiol) lyase A and B confer improved selenium resistance and degrade l-Cys and l-SeCys in Arabidopsis. Journal of Experimental Botany 2022, 73(8):2525-2539.
34. Romero LC, Aroca MÁ, Laureano-Marín AM, Moreno I, García I, Gotor C: Cysteine and cysteine-related signaling pathways in *Arabidopsis thaliana*. Molecular plant 2014, 7(2):264-276.
35. Heeg C, Kruse C, Jost R, Gutensohn M, Ruppert T, Wirtz M, Hell Rd: Analysis of the Arabidopsis O-acetylserine (thiol) lyase gene family demonstrates compartment-specific differences in the regulation of cysteine synthesis. The Plant Cell 2008, 20(1):168-185.
36. Birke H, Haas FH, De Kok LJ, Balk J, Wirtz M, Hell R: Cysteine biosynthesis, in concert with a novel mechanism, contributes to sulfide detoxification in mitochondria of *Arabidopsis thaliana.* Biochemical Journal 2012, 445(2):275-283.

37. Bermúdez MA, Páez-Ochoa MA, Gotor C, Romero LC: Arabidopsis S-sulfocysteine synthase activity is essential for chloroplast function and long-day light-dependent redox control. The Plant Cell 2010, 22(2):403-416.
38. Rajabi-Farshmi M, Shirzadian-khorramabad R, Abedi A, Rezadoost M-h: Phylogenetic, structural and expression study of OASTL-Like gene family in the genome wide of Arabidopsis and rice using in silico approaches. Journal of Plant Research 2022.
39. Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, Xia R: TBtools: an integrative toolkit developed for interactive analyses of big biological data. Molecular plant 2020, 13(8):1194-1202.
40. Studer G, Rempfer C, Waterhouse AM, Gumienny R, Haas J, Schwede T: QMEANDisCo-distance constraints applied on model quality estimation. Bioinformatics 2020, 36(6):1765-1771.
41. Jendele L, Krivak R, Skoda P, Novotny M, Hoksza D: PrankWeb: a web server for ligand binding site prediction and visualization. Nucleic acids research 2019, 47(W1):W345-W349.
42. Eberhardt J, Santos-Martins D, Tillack AF, Forli S: AutoDock Vina 1.2. 0: New docking methods, expanded force field, and python bindings. Journal of Chemical Information Modeling 2021, 61(8):3891-3898.
43. Livak KJ, Schmittgen TD: Analysis of relative gene expression data using real-time quantitative PCR and the 2-ΔΔCT method. methods 2001, 25(4):402-408.
44. Akbudak MA, Filiz E, Uylas S: Identification of O-acetylserine (thiol) lyase (OASTL) genes in sorghum (*Sorghum bicolor*) and gene expression analysis under cadmium stress. Molecular biology reports 2019, 46(1):343-354.
45. Yi H, Dey S, Kumaran S, Lee SG, Krishnan HB, Jez JM: Structure of soybean serine acetyltransferase and formation of the cysteine regulatory complex as a molecular chaperone. Journal of Biological Chemistry 2013, 288(51):36463-36472.
46. Jez JM, Dey S: The cysteine regulatory complex from plants and microbes: what was old is new again. Current Opinion in Structural Biology 2013, 23(2):302-310.
47. Bogdanova N, Hell R: Cysteine synthesis in plants: protein-protein interactions of serine acetyltransferase from *Arabidopsis thaliana*. The Plant Journal 1997, 11(2):251-262.
48. Bonner ER, Cahoon RE, Knapke SM, Jez JMJJoBC: Molecular basis of cysteine biosynthesis in plants: structural and functional analysis of O-acetylserine sulfhydrylase from *Arabidopsis thaliana*. 2005, 280(46):38803-38813.
49. Jost R, Berkowitz O, Wirtz M, Hopkins L, Hawkesford M, Hell R: Genomic and functional characterization of the oas gene family encoding O-acetylserine (thiol) lyases, enzymes catalyzing the final step in cysteine biosynthesis in *Arabidopsis thaliana*. Gene 2000, 253(2):237-247.
50. Kopriva S, Mugford SG, Matthewman C, Koprivova A: Plant sulfate assimilation genes: redundancy versus specialization. Plant cell reports 2009, 28(12):1769-1780.
51. Guedes IA, de Magalhães CS, Dardenne LE: Receptor-ligand molecular docking. Biophysical reviews 2014, 6:75-87.
52. Tao X, Huang Y, Wang C, Chen F, Yang L, Ling L, Che Z, Chen X: Recent developments in molecular docking technology applied in food science: a review. International Journal of Food Science Technology 2020, 55(1):33-45.
53. Sethi A, Joshi K, Sasikala K, Alvala M: Molecular docking in modern drug discovery: Principles and recent applications. Drug discovery development-new advances 2019, 2:1-21.
54. Adasme MF, Linnemann KL, Bolz SN, Kaiser F, Salentin S, Haupt V J, Schroeder M: PLIP 2021: expanding the scope of the protein-ligand interaction profiler to DNA and RNA. Nucleic Acids Research 2021, 49(W1):W530-W534.
55. Seeliger D, de Groot BL: Ligand docking and binding site analysis with PyMOL and Autodock/Vina. Journal of computer-aided molecular design 2010, 24(5):417-422.
56. Liu D, Lu J, Li H, Wang J, Pei Y: Characterization of the O-acetylserine (thiol) lyase gene family in *Solanum lycopersicum* L. Plant molecular biology 2019, 99(1):123-134.
57. Kurt F, Filiz E, Aydin A: Genome-wide identification of serine acetyltransferase (SAT) gene family in rice (*Oryza sativa*) and their expressions under salt stress. Mol Biol Rep 2021, 48(9):6277-6290.

58. Tavares S, Wirtz M, Beier MP, Bogs J, Hell R, Amâncio S: Characterization of the serine acetyltransferase gene family of Vitis vinifera uncovers differences in regulation of OAS synthesis in woody plants. Frontiers in plant science 2015, 6:74.
59. Francois JA, Kumaran S, Jez JM: Structural Basis for Interaction of O-Acetylserine Sulfhydrylase and Serine Acetyltransferase in the Arabidopsis Cysteine Synthase Complex. The Plant Cell 2006, 18(12):3647-3655.
60. A lvarez C, Calo L, Romero LC, Garci a I, Gotor C: An O-Acetylserine(thiol)lyase Homolog with l-Cysteine Desulfhydrase Activity Regulates Cysteine Homeostasis in Arabidopsis. Plant Physiology 2010, 152(2):656-669.
61. Hartl FU, Bracher A, Hayer-Hartl M: Molecular chaperones in protein folding and proteostasis. Nature 2011, 475(7356):324-332.
62. Saito Y: Selenium transport mechanism via selenoprotein P-Its physiological role and related diseases. Frontiers in Nutrition 2021, 8:685517.

Figures

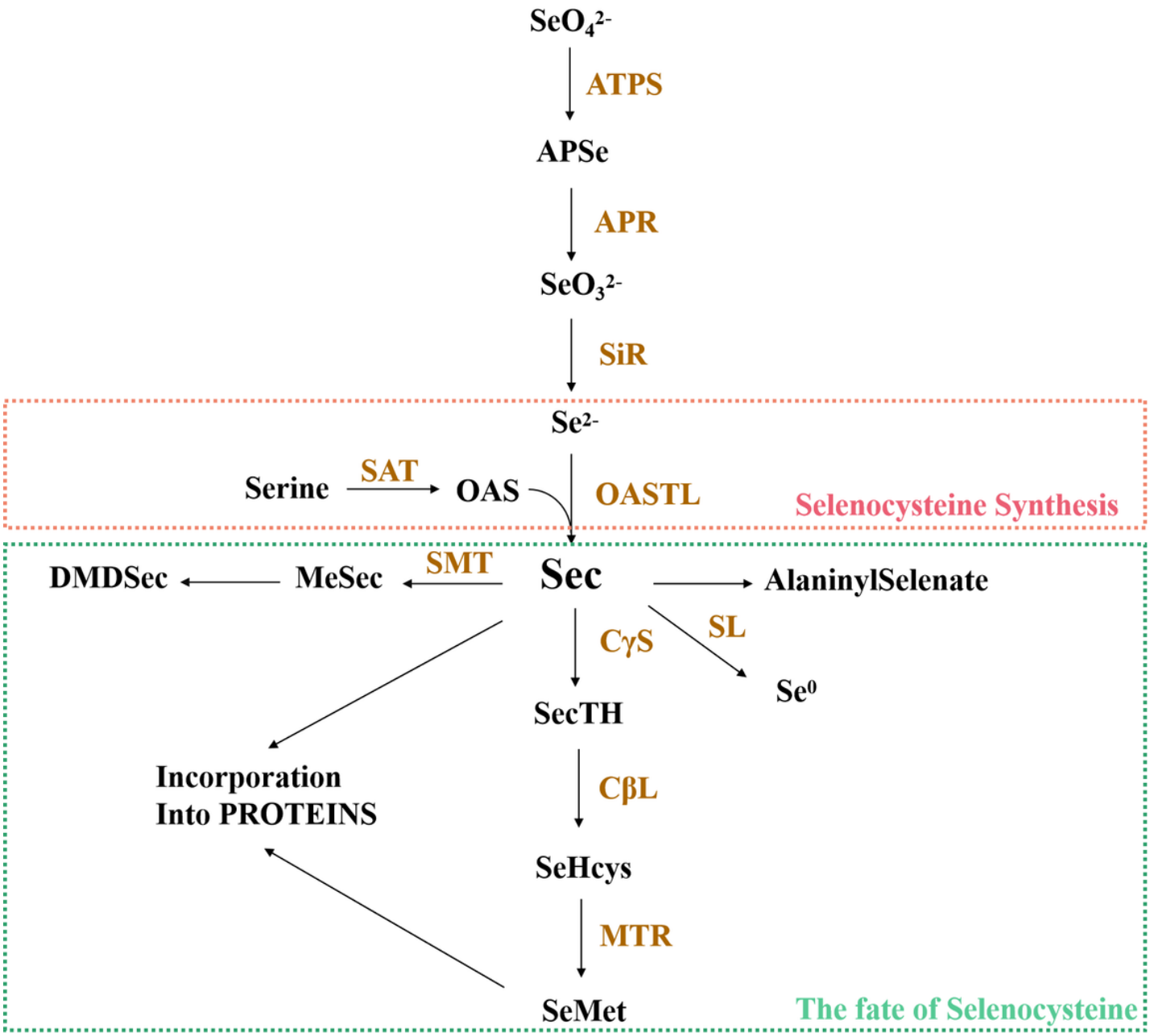


Figure 1

The metabolic pathway of selenium in plants. Selenate: SeO_4^{2-}; ATPS: ATP sulfurylase; APSe: adenosine 5'-phosphoselenate; APR: adenosine 5'-phosphosulfate reductase; selenite: SeO_3^{2-}; SiR: sulfite reductase. Selenocysteine Synthesis: SAT: serine acetyltransferase; OAS: O-acetylserine; OASTL: O-acetylserine (thiol) lyase; Sec: selenocysteine. The fate of selenocysteine: CγS: cystathionine gamma synthase; SecTH: selenocystathionine; CβL: cystathionine beta lyase; SeHcys: selenohomocysteine; MTR: methionine synthase; SeMet: selenomethionine; MeSec: methyl-selenocysteine; DMDSe: dimethyl diselenide; SL: selenocysteine lyase.

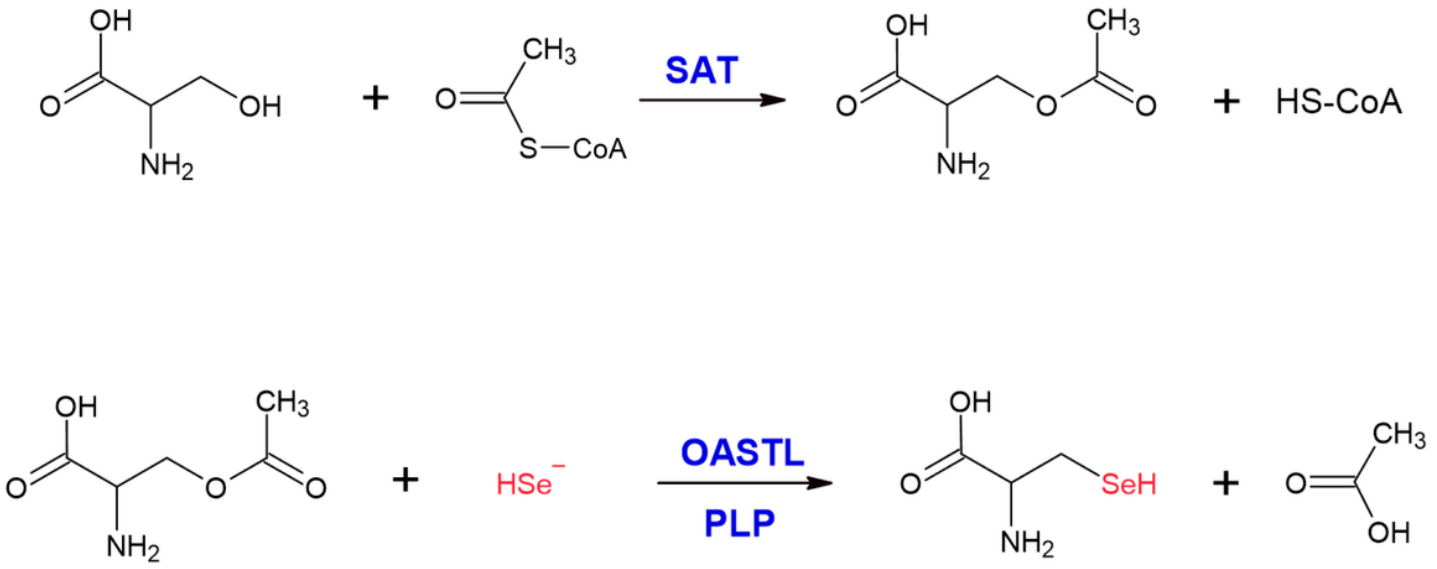


Figure 2

Reaction process of selenocysteine synthesis

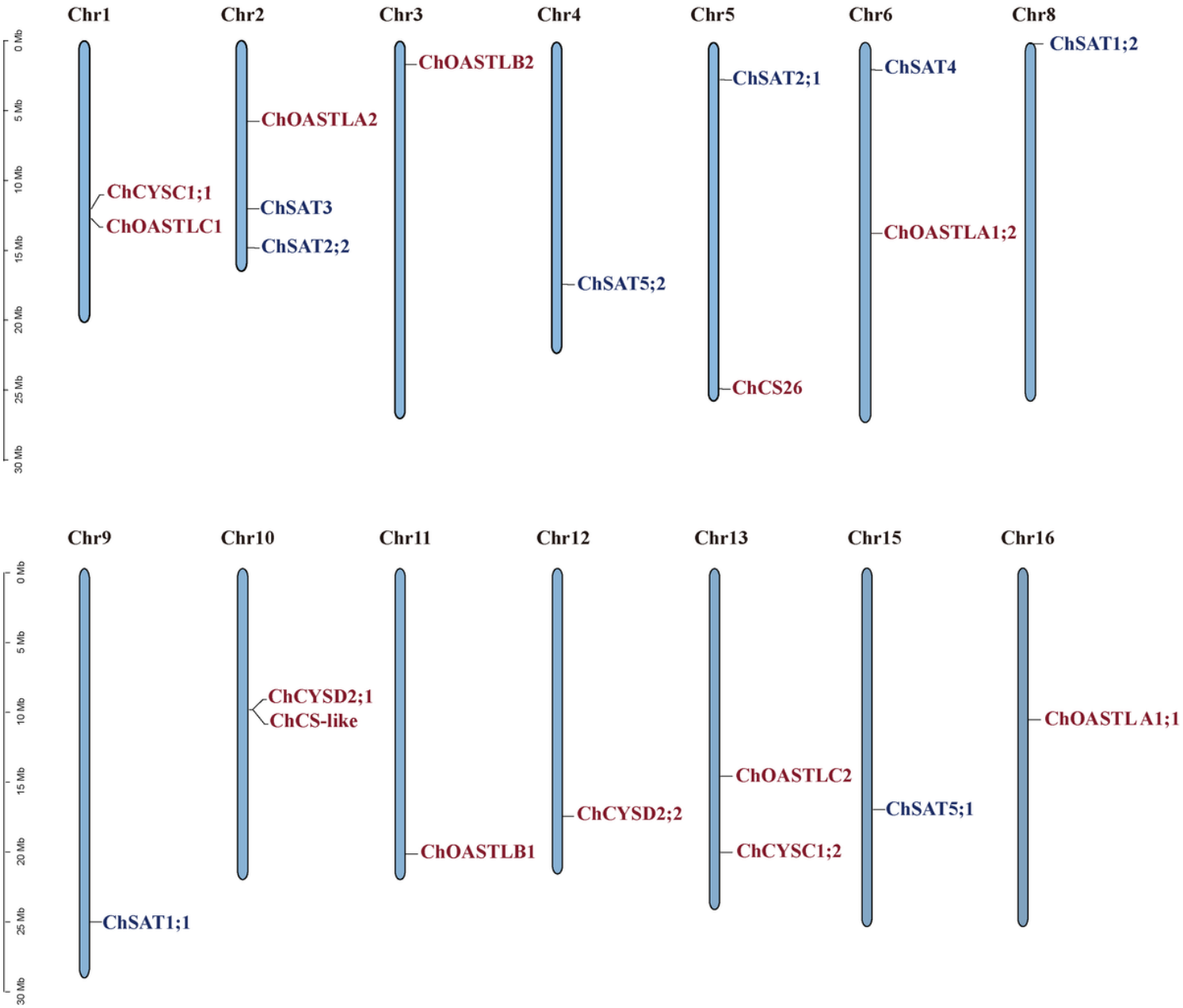


Figure 3

Chromosomal localization of the SAT/OASTL genes in *C. hupingshanensis*.

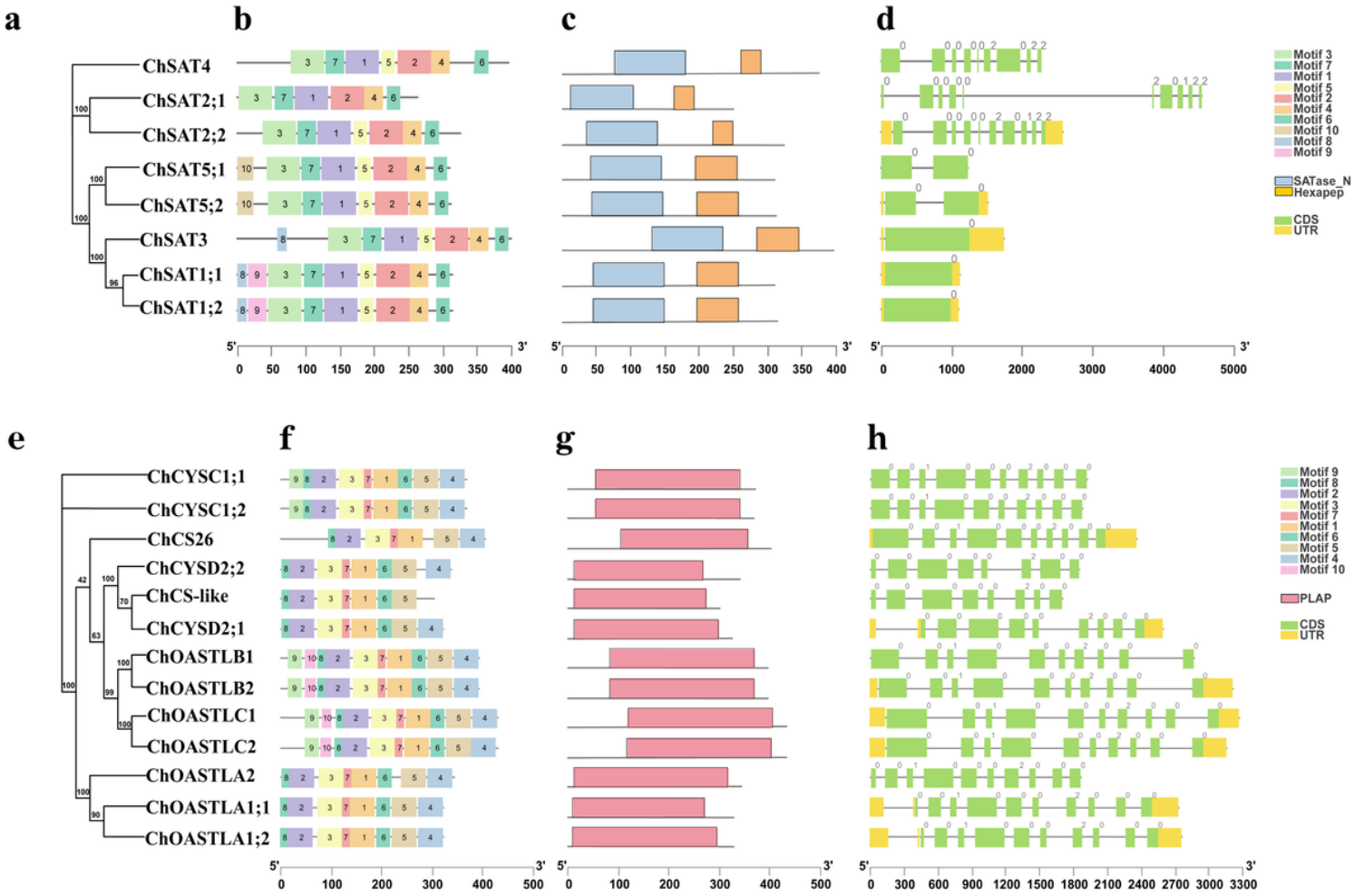


Figure 4

Phylogenetic tree, conserved motifs, conserved structural domains, and gene structure of SAT/OASTL genes in *C. hupingshanensis*. (a) The phylogenetic tree of ChSAT proteins. (b, c) The conserved motifs and structural domains of ChSAT proteins; different conserved motifs and structural domains are marked by different colours. (d) Intron and exon structures of ChSAT genes; exons and introns are represented by green squares and black lines. (e) The phylogenetic tree of ChOASTL proteins. (f, g) The conserved motifs and structural domains of ChOASTL proteins; different conserved motifs and structural domains are marked by different colours. (h) Intron and exon structures of ChOASTL genes; exons and introns are represented by green squares and black lines.

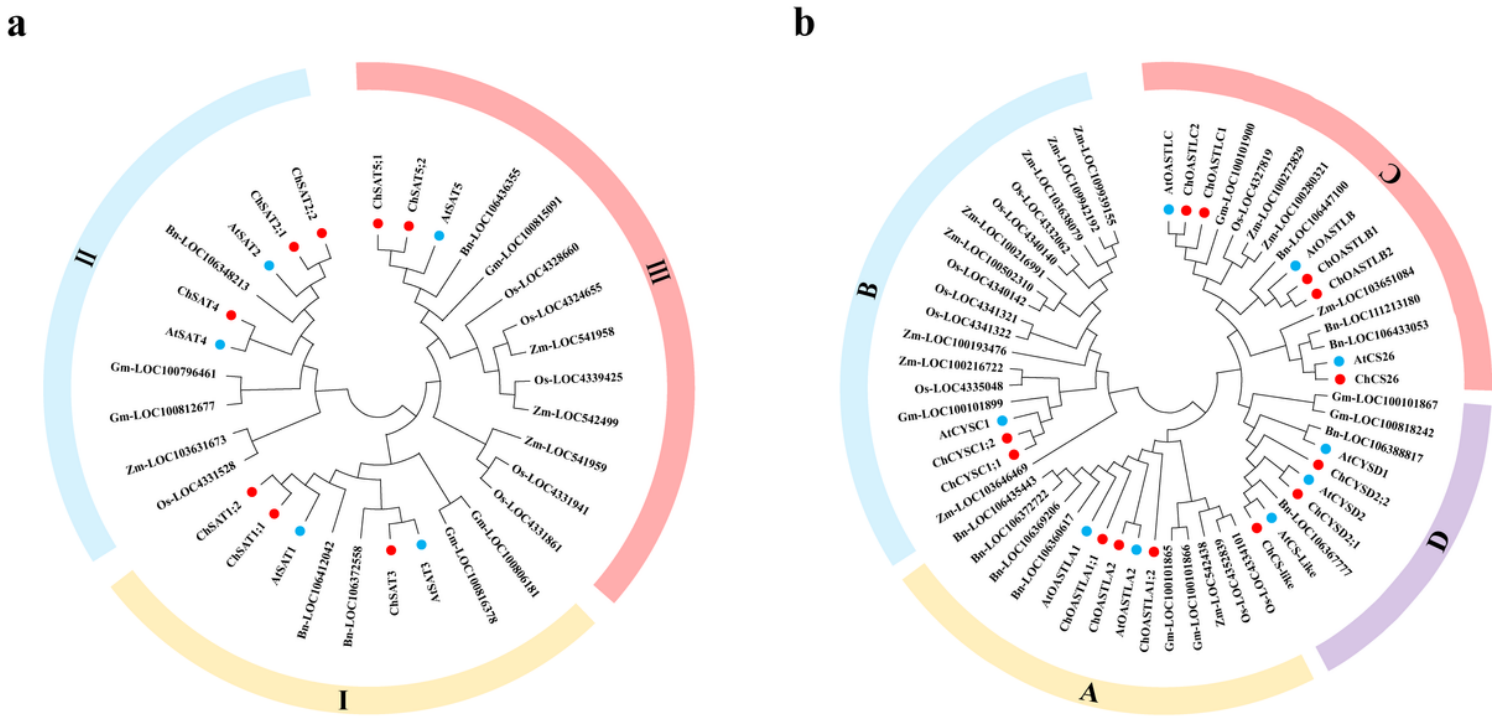


Figure 5

a. Phylogenetic analysis of the ChSAT family of proteins in *C. hupingshanensis*; b. Phylogenetic analysis of the ChOASTL family of proteins in *C. hupingshanensis*. (At: *Arabidopsis thaliana*; Bn: *Brassica napus*; Ch: *Cardamine hupingshanensis*; Gm: *Glycine max*; Os: *Oryza sativa*; Zm: *Zea mays*)

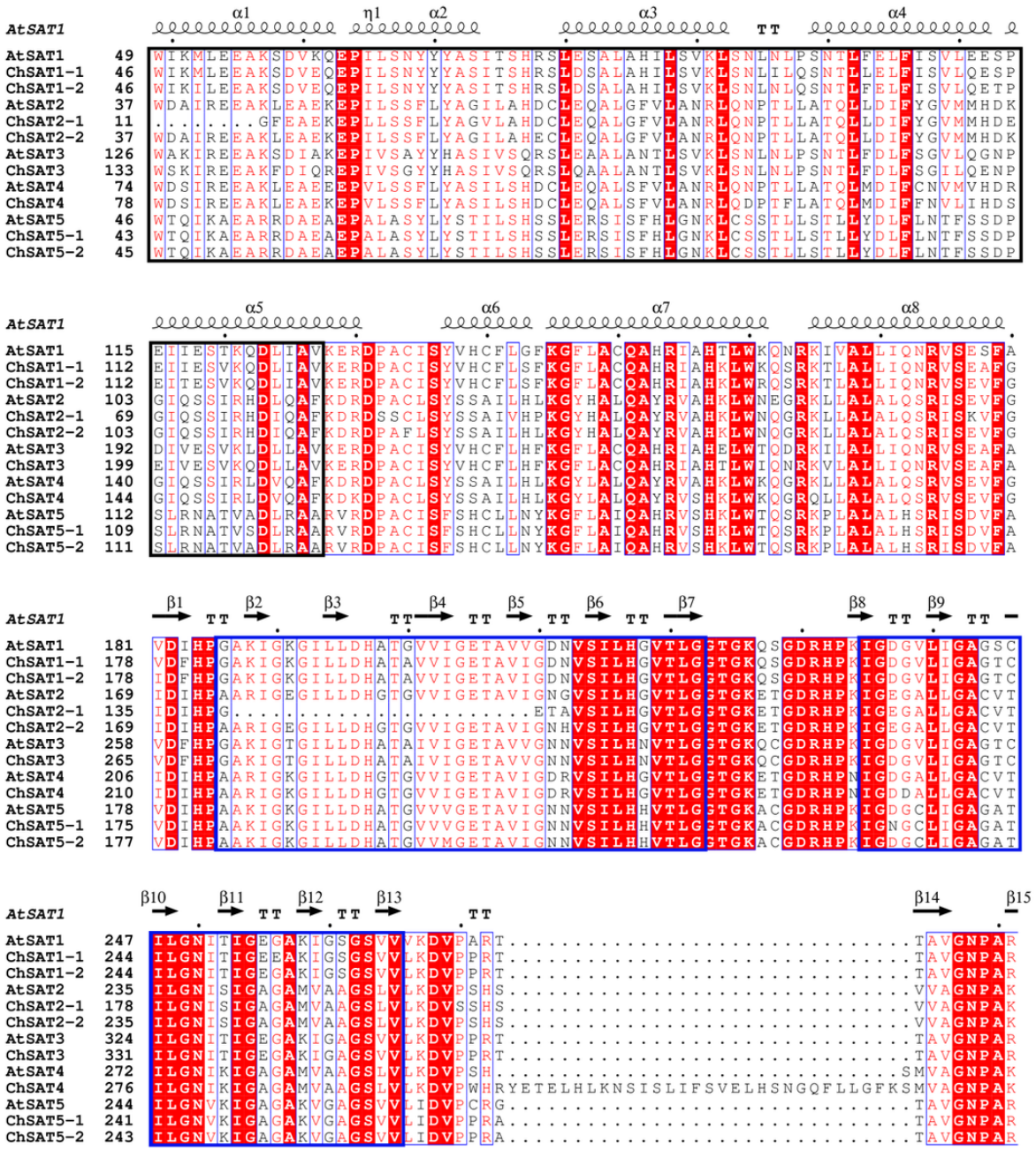


Figure 6

Multiplexed alignment of partial sequences of SAT protein in *C. hupingshanensis*. Amino acid residues forming the structural domain of the serine acetyltransferase N-terminal domain structure (SATase_N, PF06426) are marked by black boxed underlines, and those of the bacterial transferase hexapeptide domain structure (PF00132) are marked by blue box underlines.

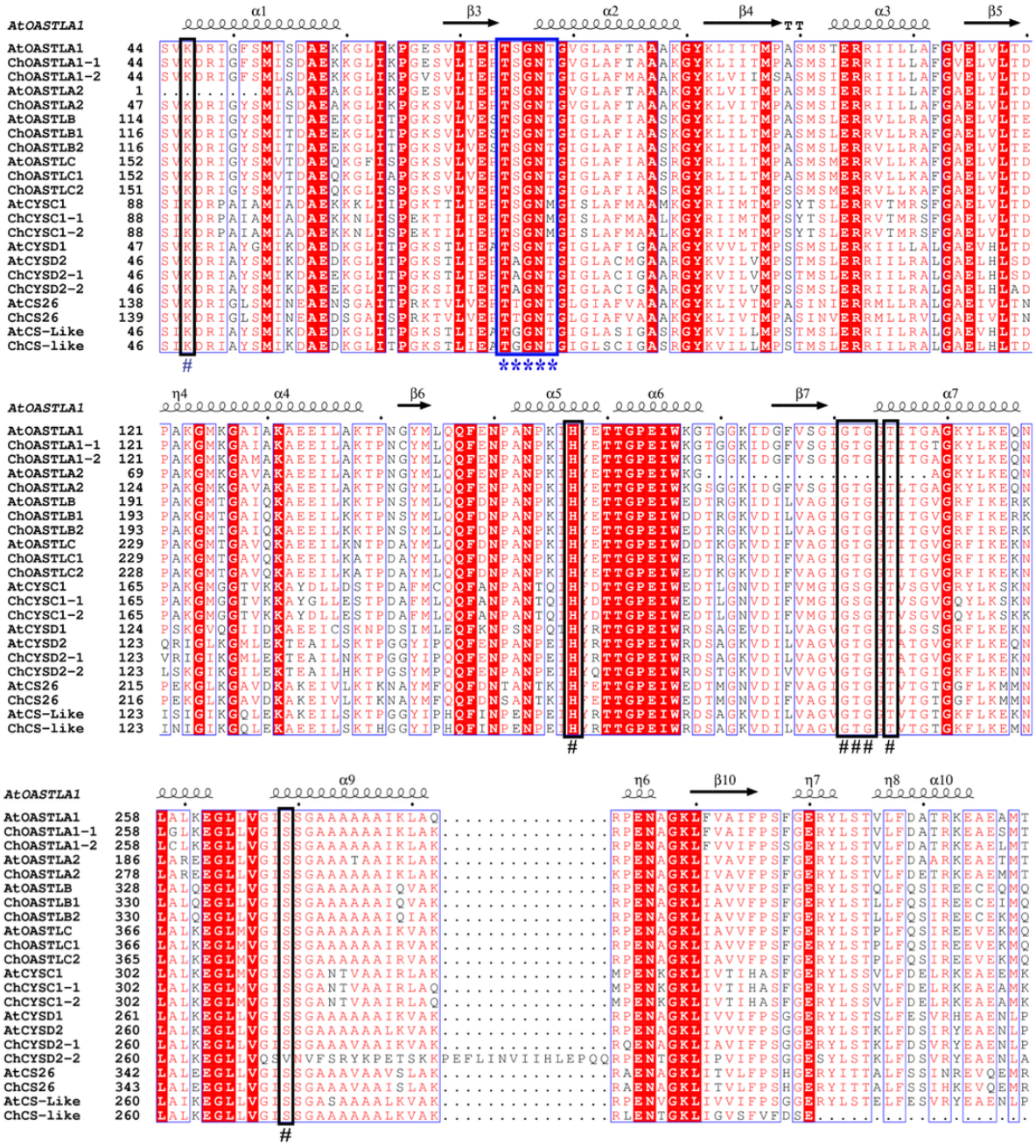


Figure 7

Multiplexed alignment of the partial sequences of OASTL protein in *C. hupingshanensis*. Amino acid residues bound to PLP are marked by black box lines and "#", and those bound to substrate constant are indicated by blue box lines and " ".

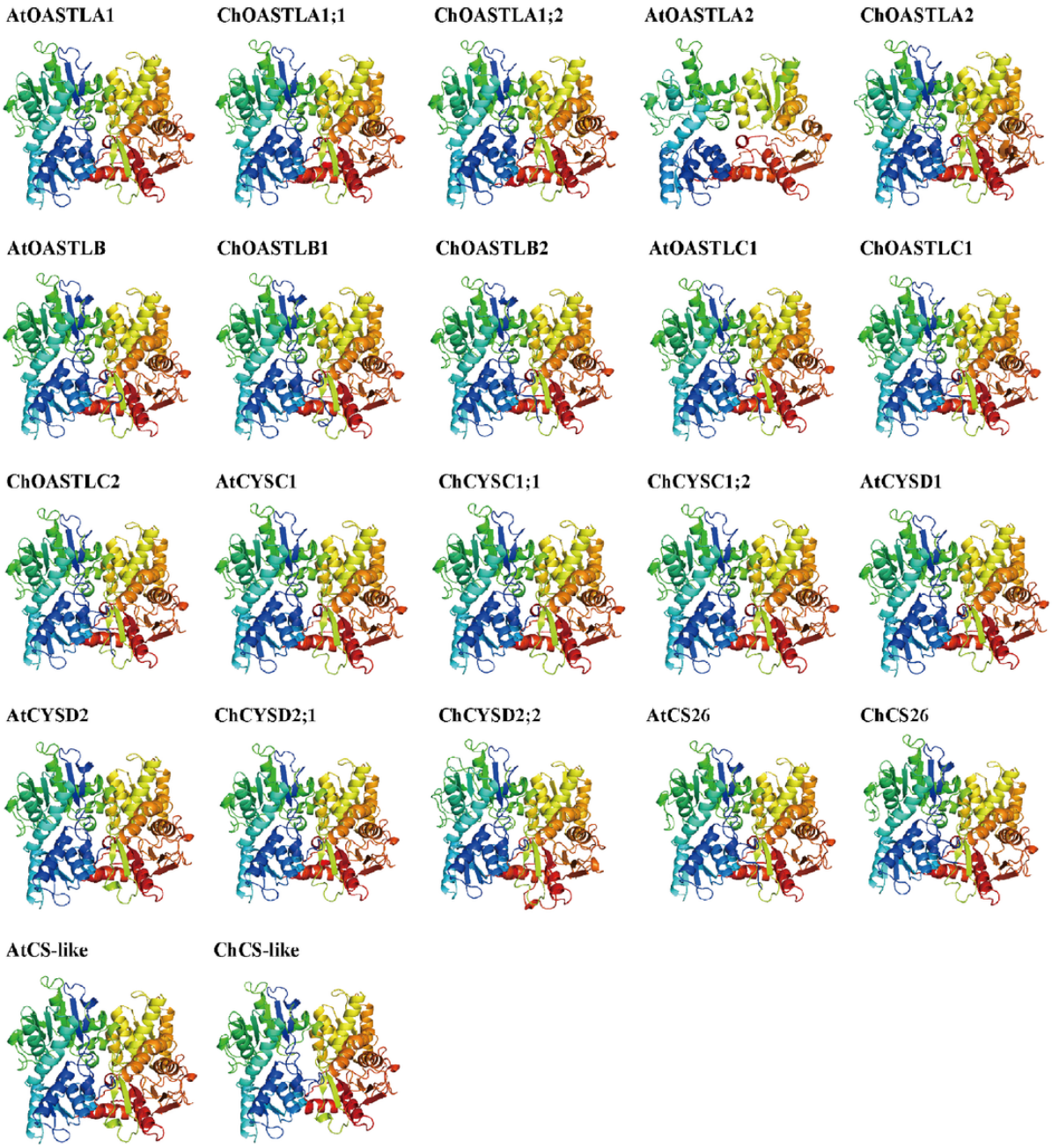


Figure 9

The 3D structures of ChOASTLs and AtOASTLs predicted by the SWISS-MODEL online server.

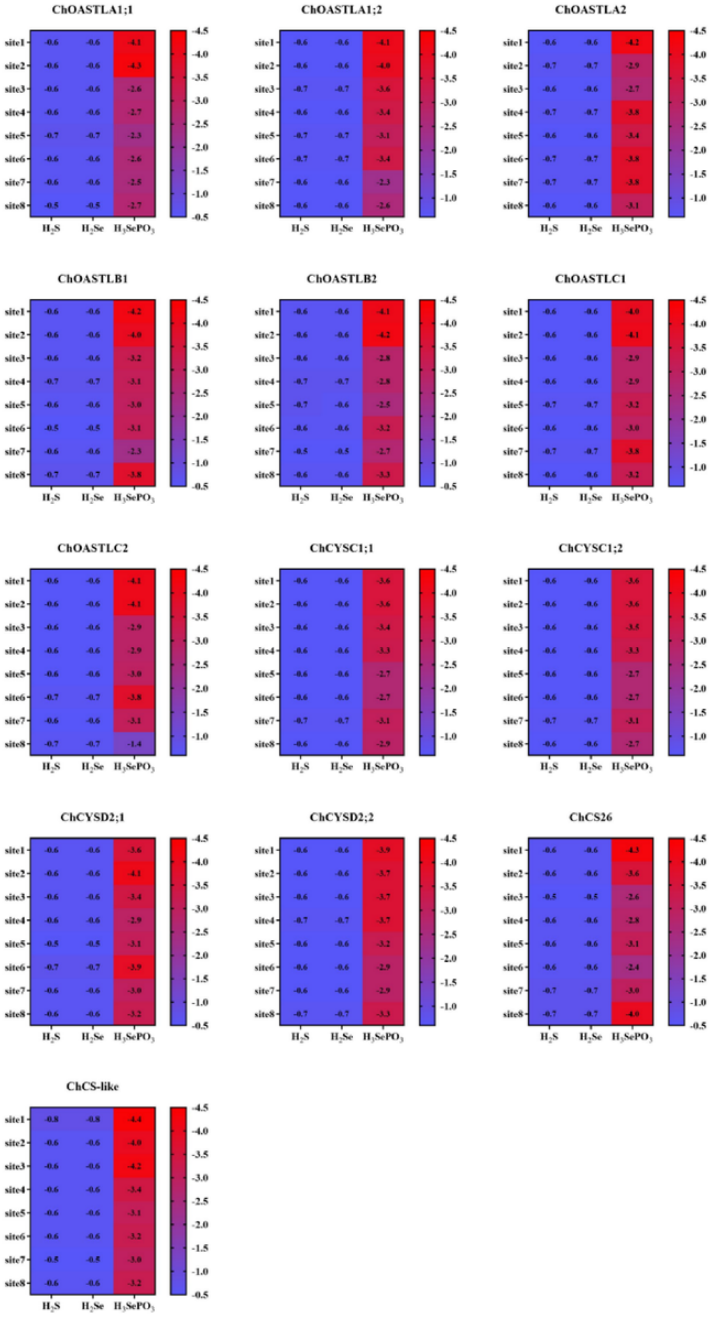


Figure 10

Binding energies of complexes to different ligand binding sites of ChOASTLs (unit: kcal mol^{-1}).

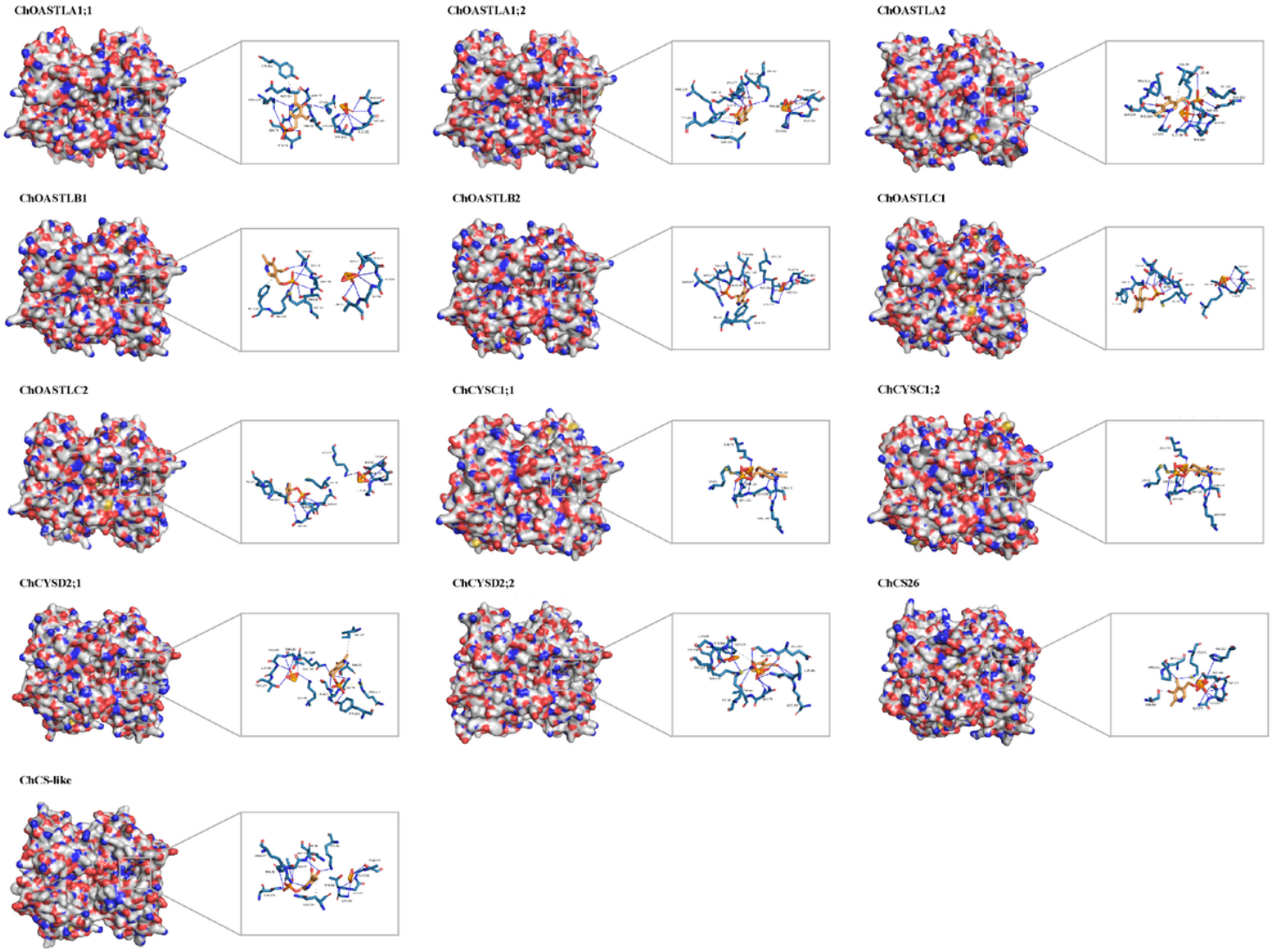


Figure 11

Interactions of the ternary ChOASTL-PLP-OAS complex with selenophosphate, simulating the binding form of ChOASTL proteins to selenophosphate at the catalytic site. The left panel shows the overall picture, and the right panel shows the details. Protein is shown on the surface, the amino acid residues at the binding site are shown in blue—green and the ligands (PLP, OAS, and selenophosphate) are shown in dark yellow. The grey dotted line represents hydrophobic interactions, the solid blue line represents hydrogen bonds, and the dashed yellow line represents salt bridges.

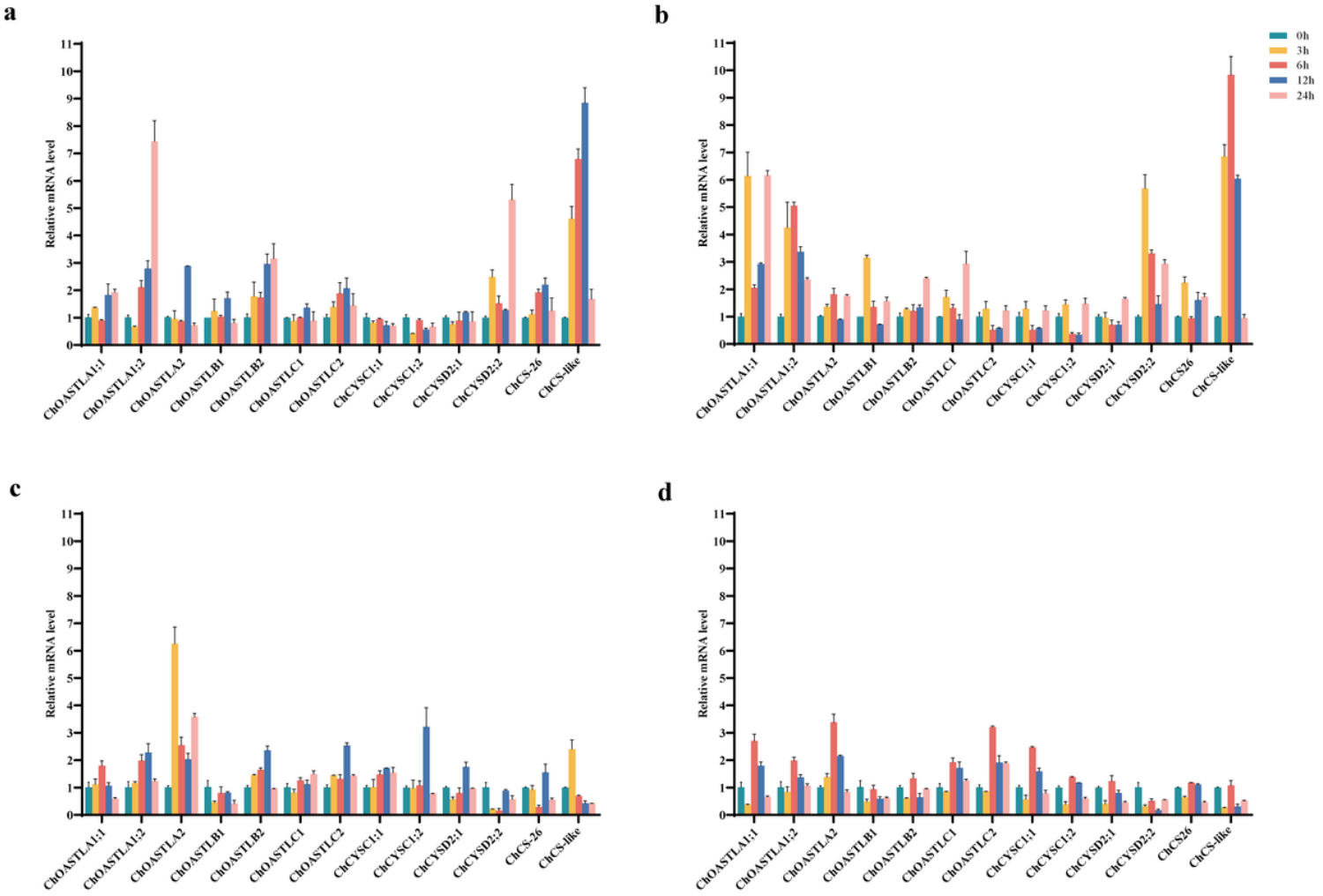


Figure 13

Expression of ChOASTL genes in leaf and root tissues in response to selenium stress. (a) Expression of ChOASTL genes in leaves in response to low selenium stress (100 μg Se L^{-1}). (b) Expression of ChOASTL genes in leaves in response to high selenium stress (80,000 μg Se L^{-1}). (c) Expression of ChOASTL genes in roots in response to low selenium stress (100 μg Se L^{-1}). (d) Expression of ChOASTL genes in roots in response to high selenium stress (80,000 μg Se L^{-1}). The horizontals represent the ChOASTL genes and the vertices represent the relative expression levels at each time point (0, 3, 6, 12, 24 h) with different treatments (0 μg Se L^{-1}, 100 μg Se L^{-1}, 80,000 μg Se L^{-1}). Samples at 0, 3, 6, 12, and 24 h are indicated in green, yellow, red, blue, and pink, respectively. Each data point represents the mean ± standard deviation (SD) (n = 3). Error bars represent standard deviations.

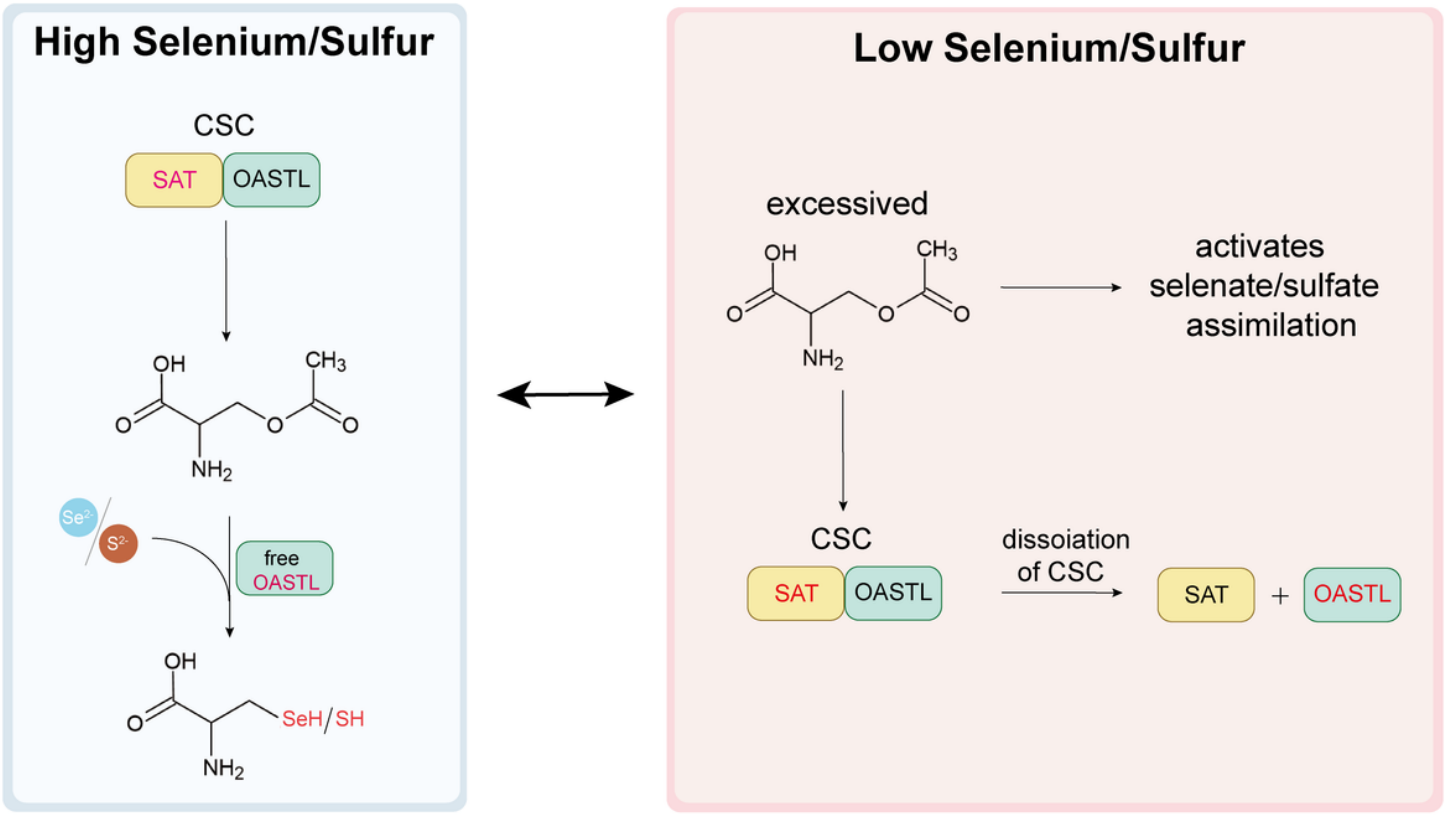


Figure 14

Formation of the cysteine synthase complex (CSC) to regulate Sec/Cys synthesis. Red font represents the active form of SAT and OASTL; black font represents the inactive form of SAT and OASTL.

Supplementary Files

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

- SupplementaryFigures.doc
- SupplementaryTableS1andS2.xls