

# Bioinformatics Analyses of Serine Acetyltransferase (SAT) Gene Family in Rice (*Oryza sativa*) and their Expressions under Salt Stress

Firat Kurt

Mus Alparslan Universitesi

Ertugrul Filiz (✉ [ertugrulfiliz@gmail.com](mailto:ertugrulfiliz@gmail.com))

Duzce uni. <https://orcid.org/0000-0001-9636-6389>

Adnan Aydın

Iğdir Üniversitesi: Iğdir Universitesi

---

## Research Article

**Keywords:** Sulfur, Rice, Hormone, Salt, Stress

**Posted Date:** April 26th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-412789/v1>

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

---

**Version of Record:** A version of this preprint was published at Molecular Biology Reports on August 13th, 2021. See the published version at <https://doi.org/10.1007/s11033-021-06620-6>.

## Abstract

Assimilation of sulfur to cysteine occurs in the presence of serine acetyltransferase (SAT). In this study, SAT genes in rice (*Oryza sativa*) were identified and analyzed using bioinformatics approaches. Also, these genes were tested under salt stress. OsSATs have two common motifs, bacterial transferase hexapeptide and acetyltransferase and underwent purifying selection. They have more similar protein sequences compared to *Arabidopsis*. However, there is structural and functional divergence among OsSATs which may be driven by the segmental and tandem duplications. Purifying selection and gene duplications may also have effect leading to variation of specificity and selectivity of OsSATs. In this regard, Asp (D), His (H), Gly (G), Thr (T), Arg (R), Ala (A), and Leu (L) are identified as well-conserved residues in their active sites which have an indicator role on their functions. The OsSATs expressions in different tissues, organs and under hormones showed that jasmonic acid was main hormone inducing the expressions of *OsSAT1;1*, *OsSAT2;1*, and *OsSAT2;2* whereas auxin and abscisic acid only triggered *OsSAT1;1* expression. On the other hand, wet-lab expressions of OsSATs in this study indicated that *OsSAT1;1*, *OsSAT1;2* and *OsSAT1;3* genes were upregulated under different exposure times of salt stress. *OsSAT1;1* is the only OsSAT induced by various stimuli. The findings can be used by plant breeders and genetic engineers to develop new rice varieties having optimal growth and stress tolerance.

## Introduction

Sulfur (S) is a vital macronutrient for plants and involved in many biochemical reactions. S is the fourth most important plant nutrient after N, P, and K. It is found in amino acids (cysteine and methionine), vitamins and cofactors, glutathione and phytochelatins [1]. Plants can use the reduced sulfate ( $\text{SO}_4^{2-}$ ) and convert it to sulfide ( $\text{S}^{2-}$ ). Sulfide is used to integrate S into cysteine (Cys) in the cytosol. Cys biosynthesis is a basic process either as a precursor or donor of key S compounds in plants. While Cys synthesis is observed in the cytosol, plastids, and mitochondria, the sulfate reduction pathway is localized in plastids of plants [2, 3]. The Cys synthesis takes place in the presence of two enzymes: serine acetyltransferase (SAT or SERAT) and O-acetylserine(thiol)lyase (OASTL). The former leads to synthesizing intermediary product O-acetylserine (OAS) from acetyl-CoA and serine. The latter uses the sulfide and OAS for production of Cys as a cofactor in the presence of pyridoxal-5'-phosphate. These two enzymes are called hetero-oligomeric cysteine synthase complex (CSC) [4–6].

The ancestral SAT gene was a host origin gene and not evolved from the cyanobacterial endosymbiont [7]. A total of five serine acetyltransferases (SAT or SERAT; EC 2.3.1.30) were identified in *Arabidopsis thaliana* [8]. AtSAT1, AtSAT3, and AtSAT5 are found to localized in plastid, mitochondria, and cytoplasm, respectively [2]. Additionally, mitochondrion is identified as the most important compartment for O-acetylserine (OAS) synthesis in *Arabidopsis*, which is the precursor of Cys [3]. Similarly, cell-distribution of SAT activities in pea (*Pisum sativum*) leaves is found at the highest level in mitochondria followed by chloroplast and cytoplasm [2]. The other two AtSATs, AtSAT2 and AtSAT4, are localized in cytoplasm but their protein sequences show variations from other SATs and are expressed at low level [9]. The overexpressed of *AtSAT1* in maize does not cause any negative effect on *Arabidopsis*'s growth and it enhances 10-kDa  $\gamma$ -zein storage protein during endosperm development [10].

In this study, rice (*Oryza sativa*) SAT (*OsSATs*) genes/proteins were identified at genome-wide scale, then bioinformatics analyses were performed (sequence, motif, phylogenetic, expression, and protein modeling) to find out structural and functional divergences of SATs in rice. Also, the responses of *OsSAT* genes were tested to offer more insights about their regulations under salt stress.

## Materials And Methods

### Identification of SAT genes and sequence feature analysis

Firstly, Q42588 AtSAT1 (At1g55920), Q8S895 AtSAT2 (At2G17640), Q39218 AtSAT3 (At3g13110), Q8W2B8 AtSAT4 (At4g35640), and Q42538 AtSAT5 (At5g56760) sequences in *Arabidopsis* were retrieved from UniProt database (<https://www.uniprot.org/>). AtSATs were only used for comparative analyses in this study. These protein sequences were searched against rice genome (*Oryza sativa* v7\_JGI) on Phytozome database v12 ([https://phytozome.jgi.doe.gov/pz/portal.html#info?alias=Org\\_Osativa](https://phytozome.jgi.doe.gov/pz/portal.html#info?alias=Org_Osativa)) with an E-value <  $1\text{E}^{20}$  [11]. Later, the presence of PF06426 (SATase\_N) domain structure was validated using Pfam database v32 (<https://pfam.xfam.org/>) [12]. Genomic and coding sequences corresponding to SATs were collected from Phytozome database for further bioinformatics analyses. Physicochemical properties of SAT proteins such as length, molecular weight, and isoelectric point (*pI*) values were obtained from ProtParam tool ([web.expasy.org/protparam](http://web.expasy.org/protparam)) [13]. The sequence identity matrix for *Arabidopsis* and rice SATs were generated using Bioedit v7.2.5 [14]. Annotations of SAT proteins were analyzed using PANNZER (Protein ANnotation with Z-score) server (<http://ekhidna2.biocenter.helsinki.fi/sanspanz/>) [15]. The conserved motif analysis was performed using MEME suit (motif-based sequence analysis tools) (<http://meme-suite.org/>). Max motif number and min–max motif width parameters on the platform were set to, in turn, 10 and 6–50 [16].

### Nucleotide and phylogenetic analyses

Maximum likelihood estimate of transition/transversion bias (R) was estimated under the Kimura [17] using two-parameter model. The estimate of transition/transversion bias and G+C content were calculated using DnaSP v6.12.01 [18]. Molecular phylogenetic analysis was performed according to the maximum likelihood (ML) method using JTT matrix-based model [19]. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances, which are estimated using a JTT model, and then topology with superior log likelihood value was selected. The analysis involved 11 amino acid sequences with 266 positions after all positions containing gaps and missing data were eliminated. Evolutionary analyses were conducted in MEGA7 [20].

## Selection, gene duplication and synteny analyses

Nucleotide sequences of rice *SAT* genes were analyzed using DnaSP v6.12.01 [18], in terms of polymorphic sites, DNA polymorphism, genetic variation level ( $\pi$  and  $\theta$ ), and Tajima's D [21]. *Ka* and *Ks* values were calculated for duplicated gene pairs using DnaSP v6.12.01 [18]. Gene duplications analyses were conducted according to two following criteria: If the alignment of the coding nucleotide sequences covers 70% of the longest genes and the amino acid identity between the sequences are  $\geq 70\%$ , gene of interest is assumed to be duplicated [22].

Synteny analyses were conducted on Circoletto platform [23]. With this aim, processed genes of *Arabidopsis*, *Brachypodium distachyon*, *Solanum lycopersicum*, and *Zea mays* were downloaded from Plant Genome Duplication Database [24]. Then the local blast database was constructed in Bioedit [14] and the blast results were fed to Circoletto.

## Digital expression and co-expression analyses

Digital expression profiles of *OsSAT* genes were obtained from Rice Expression Profile Database (RiceXPro) v3 (<http://ricexpro.dna.affrc.go.jp/>) (Sato et al. 2013a). Co-expression network was constructed using RiceFRIEND co-expression database. RiceFRIEND contains expression profiles obtained from microarray analyses of tissues/organs of rice plant under natural field conditions and various treatments (<http://ricefrend.dna.affrc.go.jp/>) (Sato et al. 2013b)

## Analyses of secondary and tertiary structures, predicted binding sites, and protein interactions

Secondary structure analyses of *OsSAT* proteins were conducted using SOPMA server ([https://npsa-prabi.ibcp.fr/cgi-bin/npsa\\_automat.pl?page=/NPSA/npsa\\_sopma.html](https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html)) [27]. Tertiary models were predicted by using Phyre<sup>2</sup> server (Protein Homology/analogy Recognition Engine V 2.0 (<https://www.sbg.bio.ic.ac.uk/phyre2/>)) [28]. The predicted 3D models were visualized using Swiss-PdbViewer [29]. The independent comparisons of biomolecular 3D structures were performed using CLICK server (<http://www.mspc.bii.a-star.edu.sg/click>) [30]. The binding sites of *OsSAT* protein sequences were identified using InterProScan tool (<http://www.ebi.ac.uk/interpro/search/sequence-search>) [31]. Protein-protein interaction network was constructed using STRING database v11 (<https://string-db.org/>) [32].

## Plant material and growth conditions

Salt tolerant local Karacadag variety was used as plant material. Rice seeds were sown in 10-cm-square pots containing 40% perlite and 60% peat. The plants were grown in the growth cabinet with following cycle: an average of 50% humidity at 25 °C at 8-hour dark stage and 30 °C at 16-hour bright stage. 4-5 plants were grown for each experimental group. Third true leaves of rice plants were exposed to 200 mM NaCl and leaf samples were collected at the 3<sup>rd</sup>-, 12<sup>th</sup>- and 24<sup>th</sup>- hour. Leaf samples were immediately stored in RNA/eater stabilization solution (Invitrogen, Cat No: AM7021).

## RNA isolation and gene expression analysis

RNA isolation was conducted by TRIzol reagent method (Ambion, Ref: 15596026). The quality and quantity of the isolated RNA were measured using 2% agarose gel and nano-drop spectrophotometer (Maestrogen, MN-913). A total of 12 samples (i.e. 3 biological replicates from each sample) were used in Real-Time qPCR studies. Genomic DNA elimination were carried out with these 12 samples by taking into account the protocol of the DNaseI kit (Cat: EN0521). RNAs were then converted to cDNA according to the kit's protocol. qPCR studies were performed with the QIAGEN Rotor-Gene Q 5-Plex instrument (Bustin et al., 2009). The *OsSATs* were expressed using the AMPIGENE qPCR Green Mix (Cat No: ENZ-NUC104-10000) commercial kit. *OsActinII* gene, identified for salt stress in rice, is used as reference gene [33].

Relative Quantification  $2^{-\Delta\Delta C_T}$  (Livak and Schmittgen 2001) method was used to compare gene expression. Gene expression data were analyzed with Rotor-Gene Q 2.3.5 software. Primers used in the study (Table 1) were designed using Primer3. v. 0.4.0 program (Untergasser et al. 2012) (<http://bioinfo.ut.ee/primer3-0.4.0/>) and synthesized by Macrogen Crop. (Amsterdam, The Netherlands).

Table 1

Primers used for RT-qPCR analysis of SATs genes.

Primer ID	Sequences (5'→3')	Gene Name
OsSAT1;1	F:GACGAGTCGTGGGTCTGG	LOC_Os01g52260.1 CDS
	R:GGTGGCGTAGAGGAACGAC	
OsSAT1;2	F:ACTCCAGCACGAGCATGAC	LOC_Os02g10830.1 CDS
	R:CCTTGATCTGGGACCACAC	
OsSAT3	F:TGAAACTGCCATTGTTGGAA	LOC_Os03g04140.1 CDS
	R:TGACCAATCTTGGGGTGTCT	
OsSAT2;1	F:ATCGTGCTACGGGACGTG	LOC_Os03g08660.1 CDS
	R:GTTACCCCGGAAGCTCT	
OsSAT2;2	F:ACTACGTGTGGGACGTGCT	LOC_Os03g10050.1 CDS
	R:ACAGGACGAGTTCGTGGTAG	
OsSAT1;3	F:ACAACGTCTCCATCCTCCAC	LOC_Os05g45710.1 CDS
	R:CAGAACACCATCCCCAATCT	

## Results

### Genome-wide identification and sequence analyses of SATs

A total of six non-redundant SAT genes were identified in rice genome using five AtSAT protein sequences as references: LOC\_Os01g52260 (OsSAT1;1), LOC\_Os02g10830 (OsSAT1;2), LOC\_Os03g04140 (OsSAT3), LOC\_Os03g08660 (OsSAT2;1), LOC\_Os03g10050 (OsSAT2;2), and LOC\_Os05g45710 (OsSAT1;3) (Table 2). The protein lengths of SATs are ranged from 298 to 391 amino acid residues; molecular weights were found between 30.44 and 42.72 kDa. All SATs are predicted to be acidic character ( $pI \leq 7$ ) except for AtSAT3. Exon numbers range from one to 10. AtSAT2 and 4 (SAT3 family) and OsSAT3 have 10 exons. All SAT proteins in the study contain the serine acetyltransferase N-terminal domain structure (SATase\_N, PF06426). In addition, bacterial transferase hexapeptide (PF00132) domain was identified as one or two repeats.

Table 2

Properties of SAT members in *Arabidopsis* and rice.

Phytozome ID	Species	Exon no	Protein length (aa)	Domain family	Repeat number	Mol. wt. (kDa)	pI
At1g55920 (SAT1, SERAT2;1)	<i>A. thaliana</i>	1	314	PF06426	2	34.25	6.45
At2g17640 (SAT2, SERAT3;1)	<i>A. thaliana</i>	10	323	PF06426	2	34.53	5.81
At3g13110 (SAT3, SERAT2;2)	<i>A. thaliana</i>	1	391	PF06426	1	42.72	7.73
At4g35640 (SAT4, SERAT3;2)	<i>A. thaliana</i>	10	355	PF06426	1	38.42	5.56
At5g56760 (SAT5, SERAT1;1)	<i>A. thaliana</i>	2	312	PF06426	2	32.77	6.70
LOC_Os01g52260 (OsSAT;1)	<i>O. sativa</i>	1	303	PF06426	2	31.85	6.66
LOC_Os02g10830 (OsSAT1;2)	<i>O. sativa</i>	1	298	PF06426	2	30.44	5.33
LOC_Os03g04140 (OsSAT3)	<i>O. sativa</i>	10	354	PF06426	1	37.52	6.37
LOC_Os03g08660 (OsSAT2;1)	<i>O. sativa</i>	1	301	PF06426	1	31.50	5.91
LOC_Os03g10050 (OsSAT2;2)	<i>O. sativa</i>	1	315	PF06426	1	32.92	6.20
LOC_Os05g45710 (OsSAT1;3)	<i>O. sativa</i>	1	314	PF06426	1	33.09	6.38
PF06426 (SATase_N): Serine acetyltransferase, N-terminal, Repeat (PF00132): Bacterial transferase hexapeptide (six repeats)							

To provide more insight about protein sequence diversity of SATs, sequence identity matrix was constructed (Table 3). The lowest identity value was found 0.325 between AtSAT2 and AtSAT3 whilst the highest value was found between OsSAT2;1 and OsSAT2;2 as 0.914. The mean identity value of SAT proteins was 0.476.

Table 3

Sequence identity matrix of SATs in *Arabidopsis* and rice

	AtSAT1	AtSAT2	AtSAT3	AtSAT4	AtSAT5	Os1	Os2	Os3	Os4	Os5	Os6
AtSAT1	ID										
AtSAT2	0.390	ID									
AtSAT3	0.611	0.325	ID								
AtSAT4	0.393	0.700	0.341	ID							
AtSAT5	0.495	0.425	0.409	0.407	ID						
Os1	0.503	0.417	0.426	0.422	0.685	ID					
Os2	0.477	0.398	0.393	0.393	0.609	0.741	ID				
Os3	0.386	0.547	0.358	0.563	0.408	0.405	0.400	ID			
Os4	0.473	0.367	0.383	0.356	0.444	0.506	0.498	0.356	ID		
Os5	0.496	0.374	0.406	0.369	0.476	0.523	0.507	0.372	0.914	ID	
Os6	0.498	0.398	0.407	0.407	0.703	0.818	0.700	0.400	0.481	0.501	ID

Os1: LOC\_Os01g52260 (OsSAT1;1), Os2: LOC\_Os02g10830 (OsSAT1;2), Os3: LOC\_Os03g04140 (OsSAT3), Os4: LOC\_Os03g08660 (OsSAT2;1), Os5: LOC\_Os03g10050 (OsSAT2;2), Os6: LOC\_Os05g45710 (OsSAT1;3).

## Conserved motif analysis

Regarding conserved motifs, 10 conserved motifs were detected (Table 4 and Fig. 1). The motif 3 was found to be related with PF00132 (Bacterial transferase hexapeptide). Motif 4 and 5 were related with PF06426 (serine acetyltransferase) domain structure. The motif 1, 2, 3, 4, and 5 were present in all SATs whereas other motifs were found in different numbers.

Table 4

The details of conserved motif sequences in SAT proteins

No.	Motif sequence	S	W	Domain
1	IGKGILLDHATGVVIGETAVVGBNVSILHGVTLGGTGKESGDRHPKIGDG	11	50	NF
2	YSHCLLNKYGFLALQAHKVAHKLWAQGRKALALQSRVSEVFAVDIHPA	11	50	NF
3	VLIGAGATILGNVIGAGAKIGAGSVVLKDVPPRTTAVGNPARLIGGKDE	11	50	PF00132
4	WDQIKAEAKRDAEKEPILSSFLYASVLSHPSLERALAFHLANKLCNPTLL	11	50	PF06426
5	FAGVLAHPEJRAAVRADLLAAKDRDPAC	11	29	PF06426
6	IPGESMDHTSFISEWSDYTI	6	20	NF
7	QDPSLTMKHDATREFFQHVAVAYKDDKPN	3	29	NF
8	MAACIDKWPTGKPKQ	4	14	NF
9	RLPEKFYCVLPDCTATDRPV	2	20	NF
10	TQLYDL	11	6	NF

W: width, S: sites, NF: not found, PF00132: Bacterial transferase hexapeptide (six repeats), PF06426: Serine acetyltransferase, N-terminal

## Nucleotide and phylogenetic analyses

R value, the ratio of transition to transversion, were estimated to provide insight into DNA sequence evolution and phylogeny reconstruction. Also, G+C contents of OsSATs were calculated to predict probable functional variations of the genes which are important for organisms in gaining adaptation to its environment. The estimated transition/transversion bias (R) was found as 0.71, indicating genetic variations. G+C contents were found 71.82%, 75.25%, 51.92%, 72.63%, 72.47%, and 68.68% for OsSAT1;1, OsSAT1;2, OsSAT3, OsSAT2;1, OsSAT2;2, and OsSAT1;3, respectively.

Phylogenetic analysis indicated that AtSATs and OsSATs split into two major clades (Group A and B). Group A also divided into two subclades (Fig. 2). OsSAT2;1 (LOC\_Os03g08660) and OsSAT2;2 (LOC\_Os03g10050) clustered with AtSAT1 (SAT2;1) and AtSAT3 (SAT2;1) in subclade A1, whereas OsSAT1;3 (LOC\_Os05g45710), OsSAT1;1 (LOC\_Os01g52260), and OsSAT1;2 (LOC\_Os02g10830) grouped with AtSAT5 (SAT1;1) in subgroup A2. OsSAT3

(*LOC\_Os03g04140*) gene clustered with AtSAT2 (SAT3;1) and AtSAT4 (SAT3;2) and separated from rest of SATs in group B (100%), indicating genetic divergence. This separation is related to OsSAT3 having the highest identity scores, in turn, with AtSAT4 and AtSAT2.

## Selection, gene duplication and synteny analyses

Nucleotide variations of six *OsSAT* genes were analyzed using two selection analyses: Tajima's D and Ka/Ks tests. In Tajima's D test, the number of polymorphic (segregating) sites in *OsSAT* genes were identified as 516, of which 47.1% (243/516) were singleton variable sites and 30.4% (273/516) were parsimony informative sites. 306 sites were invariable (monomorphic) and nucleotide diversity was found 0.33 and 0.27 for  $\pi$  and  $\theta$  parameters, respectively. Tajima's D was found 1.22, indicating purifying (negative) selection.

In the second selection test, the nonsynonymous (*Ka*) and synonymous (*Ks*) substitution rates between the duplicated gene pairs were calculated. *Ka/Ks* values for all *OsSAT* genes were found less than one, validating previous finding that *OsSAT* genes were subjected to purifying selection. Moreover, gene duplications analyses indicated that three segmental duplications and one tandem duplication occurred, suggesting that these duplications are the major force for *SAT* gene expansion (Table 5).

Table 5  
Segmental and tandem duplications of *SAT* paralogous pairs in rice genome

SAT group	Chr. location	Duplication type	Ka	Ks	Ka/Ks
<i>OsSAT1;1 OsSAT1;2</i>	Chr1	Segmental	0.114	0.184	0.619
	Chr2				
<i>OsSAT1;1 OsSAT1;3</i>	Chr1	Segmental	0.073	0.296	0.247
	Chr5				
<i>OsSAT1;2 OsSAT1;3</i>	Chr2	Segmental	0.146	0.363	0.402
	Chr5				
<i>OsSAT2;1 OsSAT2;2</i>	Chr3	Tandem	0.018	0.023	0.783
	Chr3				
Non-synonymous ( <i>Ka</i> ) and synonymous ( <i>Ks</i> ) indicate the substitution rates. <i>Ka/Ks</i> is non-synonymous/synonymous mutation ratio					

To understand gene duplication dynamics of *OsSAT* genes four comparative syntenic maps of rice, associated with four representative species, were generated. (Figure 3). Six *OsSAT* genes had a syntenic relationship with five genes in *Arabidopsis*, followed by four genes in *B. distachyon*, four genes in tomato, and three genes in maize.

## Distinct expression profiles of *OsSAT* genes

To better understand the functions of *OsSAT* genes in the cell metabolism, gene expression levels in different tissues and organs were displayed as a heatmap in Fig 4. For digital expression analysis, five *OsSAT* genes were obtained from Rice Expression Profile Database (RiceXPro) except for *OsSAT1;2* (*LOC\_Os02g10830*) gene. In general, *OsSAT* genes in stem, inflorescence, anther, pistil, lemma, palea, and embryo tissues/organ were expressed at low level. Particularly, *OsSAT2;1* and *OsSAT2;2* genes were down-regulated as about -3-fold change. Also, these two genes were downregulated in inflorescence and anther tissues. In leaf blade and sheath, *OsSAT* genes commonly were upregulated. *OsSAT2;1* and *OsSAT2;2* genes were upregulated about three-fold changes in leaf blade.

In the second step, gene expression profiles of *OsSAT* genes were evaluated at five different time points depending on six types of plant hormone applications (abscisic acid, gibberellin, auxin, brassinosteroid, cytokinin, and jasmonic acid) (Fig. 5). Jasmonic acid treatment especially causes upregulation of *OsSAT* genes in terms of gene expression levels. Particularly *OsSAT1;1* gene was upregulated nearly four-fold by jasmonic acid. Also, it was found that the optimal time point for the expression to increase was one hour for *OsSAT1;1* gene under jasmonic acid treatment. Also, *OsSAT1;1* responded to the auxin and abscisic acid treatments at the third hour of exposure despite its responses were not as high as to that of jasmonic acid. The up-regulation of *OsSAT1;1* gene also was observed under abscisic acid and auxin treatments. *OsSAT2;1* and *OsSAT2;2* genes were upregulated at the first hour of exposure to jasmonic acid. Overall, levels of gene expressions under hormone treatments supported that *OsSAT1;1*, *OsSAT2;1*, and *OsSAT2;2* genes were positively regulated by hormone treatments.

## Co-expression analysis of *OsSAT* genes

The co-expression network of *OsSAT* genes was constructed using RiceFRIEND database. Co-expression network of *OsSAT* genes displayed that Os04g0577500, Os11g0524300, Os06g0167400, Os06g0690700, Os12g0641300, Os04g0488700, and Os07g0589000 were seven first neighbor genes

that co-expressed with OsSATs (Fig. 6). Os04g0488700 (similar to PHY3, AGC kinase) was co-expressed with *OsSAT1;1* gene. The AGC kinase family is one of seven kinase families and they are conserved in all eukaryotic genomes. AGC kinases in plants play roles in modulation of kinase activity by external stimuli [34]. Os12g0641300 (similar to Zn-dependent hydrolases of the beta-lactamase fold) was identified as a co-expressed gene with *OsSAT1;3*. Os07g0589000 (lateral organ boundaries, LOB domain containing protein) was co-expressed with *OsSAT2;1* and *OsSAT2;2*. Lateral organ boundaries domain (LBD) proteins contain lateral organ boundaries (LOB) domain that are key regulators for plant organ development such as photomorphogenesis, plant regeneration, and pollen development [35]. *OsSAT3* was co-expressed with Os04g0577500 (TatD-related deoxyribonuclease family protein), Os11g0524300 (protein of unknown function DUF1001 family protein), Os06g0167400 (di-trans-poly-cis-decaprenylcistransferase family protein), and Os06g0690700 (similar to potential cadmium/zinc-transporting ATPase HMA1). TatD is conserved protein found in all living organisms and participates in DNA fragmentation during apoptosis in eukaryotic cells [36]. Heavy metal pumps (P1B-ATPases) are vital for cellular heavy metal homeostasis. *Arabidopsis thaliana* contains eight *P1B-ATPase* genes (heavy metal ATPases 1–8 (HMA1–HMA8) members [37].

## Annotations of SAT proteins

The gene ontology (GO) analyses of OsSAT proteins were performed using PANNZER server, in terms of biological process, molecular function, and cellular component (Fig. 7). Sulfur amino acid biosynthetic process (GO:0000097), L-serine metabolic process (GO:0006563), biosynthetic process from serine (GO:0006535), cellular amino acid biosynthetic process (GO:0008652), sulfate assimilation (GO:0000103), and response to sulfate starvation (GO:0009970) were identified as biological processes in which OsSATs are involved (Fig. 7A). Serine O-acetyltransferase activity (GO:0009001), terpene synthase activity (GO:0010333), magnesium ion binding (GO:0000287), zinc ion binding (GO:0008270), and protein binding (GO:0005515) were predicted molecular functions which are OsSATs are carried out (Fig. 7B). In addition, cytoplasm (GO:0005737), intrinsic component of membrane (GO:0031224), intracellular part (GO:0044424), and membrane-bounded organelle (GO:0043227) were identified as cellular components in which OsSATs function (Fig. 7C). When the results of GO analyses were evaluated, it was clearly observed that amino acid synthesis is the most prominent biological process for OsSATs.

## Secondary and tertiary structure analyses of OsSATs

According to the secondary structure analyses, there are structural variations among OsSAT proteins. The alpha helix, extended strand, beta turn, and random coil percentages (%) were found between 36.94 -42.62, 17.46 - 21.19, 7.38 - 10.45, and 28.25 -34.65, respectively (Supp. Table 1). The predicted 3D structures of OsSAT proteins (Fig. 8) were found reliable due to and their Ramachandran values ranging from 96% to 99% in core and allowed regions. Homologous proteins from different organisms can be recognized using sequence comparison because amino acid substitutions in particular positions are prevented by strong selective constraints [38]. These structural variations at secondary and tertiary levels may be associated with SAT proteins' functional flexibilities.

The 3D structural similarities (%) were identified using six rice, five *Arabidopsis*, and one soybean SAT proteins on CLICK structure comparison server (Table 6). In general, the 3D structural similarity values were identified between 68.36% and 88.45%. The highest (88.45%) and lowest (68.36%) similarity values were found in turn between OsSAT1;1 & AtSAT4 and OsSAT3 & GmSAT. When *Arabidopsis* and rice similarity values were compared, the similarity was observed over 80%, indicating well-conserved SAT protein structure. The protein having the lowest similarity values with other SATs was OsSAT3 protein, proving structural divergence among SATs [9].

Table 6

The 3D structure overlap (%) of rice, *Arabidopsis*, and soybean SATs using CLICK structure comparison server

	Os1	Os2	Os3	Os4	Os5	Os6
AtSAT1	82.84	87.58	79.30	85.71	82.80	79.94
AtSAT2	83.50	87.25	84.21	85.38	84.76	81.85
AtSAT3	87.79	88.26	73.45	86.05	86.03	87.90
AtSAT4	88.45	88.26	72.88	86.71	84.44	84.71
AtSAT5	79.87	82.21	84.62	80.73	79.17	81.73
GmSAT	79.87	81.21	68.36	81.73	77.14	77.71
Os1: LOC_Os01g52260 (OsSAT1;1), Os2: LOC_Os02g10830 (OsSAT1;2), Os3: LOC_Os03g04140 (OsSAT3), Os4: LOC_Os03g08660 (OsSAT2;1), Os5: LOC_Os03g10050 (OsSAT2;2), Os6: LOC_Os05g45710 (OsSAT1;3), Gm: <i>Glycine max</i> (PDB code: 4N69)						

## Predicted active sites of OsSATs

The identification of catalytic residues of enzymes is an indispensable step for understanding the functions of enzymes [39]. In this study, active site predictions of OsSATs were performed using InterPro 74.0 server (Table 7). Particularly, Asp (D), His (H), Gly (G), Thr (T), Arg (R), Ala (A), and Leu (L) residues were conserved at different positions in all OsSATs; in contrast, some residues such as 248M (Met), 249Q (Gln), and 292A (Ala) residues were only

identified in LOC\_Os03g04140 (OsSAT3) protein, suggesting functional divergence of SAT3 protein in rice. In general perspective, it was found that similar amino acid residues were present in the predicted active binding sites.

Table 7

Predicted active sites of OsSAT proteins

Protein name	Active sites
OsSAT1;1 (LOC_Os01g52260)	189D, 190H, 209L, 210H, 216G, 217T, 224R, 225H, 236A, 251K, 253G, 254A, 257L, 259L, 267T, 272P
OsSAT1;2 (LOC_Os02g10830)	182D, 183H, 202L, 203H, 209G, 210T, 217R, 218H, 229A, 244K, 246G, 247A, 250L, 252L, 260T, 265P
LOC_Os03g04140 (OsSAT3)	228D, 229H, 248M, 249Q, 255G, 256T, 263R, 264H, 275A, 290M, 292A, 293A, 296L, 298L, 306M, 311P
OsSAT2;1 (LOC_Os03g08660)	197D, 198H, 217L, 218H, 224G, 225T, 232R, 233H, 244A, 259E, 261G, 262A, 265I, 267L, 275T
OsSAT2;2 (LOC_Os03g10050)	203D, 204H, 223L, 224H, 230G, 231T, 238R, 239H, 250A, 265K, 267G, 268A, 271V, 273L, 281T, 286P
OsSAT1;3 LOC_Os05g45710	195D, 196H, 215L, 216H, 222G, 223T, 230R, 231H, 242A, 257K, 259G, 260A, 263V, 265L, 273T, 278P

## The expression of OsSATs under salt stress

In this study, *OsSATs* responses under 3, 12, 24-h salt treatments were investigated (Fig. 9). *OsSAT2;1*, *OsSAT2;2* and *OsSAT3* were downregulated under all exposure times. The magnitude of *OsSAT2;1* expression to all salt treatments was the lowest compared to other *OsSATs*. On the other hand, *OsSAT2;2* and *OsSAT3* responded to exposure times in a similar way. The expressions of *OsSAT1;1* and *OsSAT1;2* and *OsSAT1;3* increased depending on the 3, 12, and 24-hour NaCl treatments. *OsSAT1;2* was generally expressed at the highest level under all exposure times. Lastly, the responses of *OsSAT1;2* and *OsSAT1;3* to salt exposure times were the highest at 24-h NaCl treatment. Overall, *OsSAT1;1* and *OsSAT1;2* and *OsSAT1;3* are responsive genes to different salt exposure times; and *OsSAT1;2* and *OsSAT1;3* were particularly upregulated by 24-hour salt treatment.

## Discussion

Sequence, nucleotide and phylogenetic analyses of *Arabidopsis* and rice proteins showed that there are sequential and phylogenetic divergences among SATs (Kawahishima et al. 2005). AtSATs and OsSATs were not separated from each other completely in phylogenetic analysis due to having the same number of exons and relatively higher identity scores. For example, *AtSAT2/4* and *OsSAT3* genes have 10 exons with identity values above 50 % unlike the rest of SATs. In terms of identity values of plants of interest, OsSATs have more similar (0.541) protein sequences compared to AtSATs (0.450). Nonetheless, the selection analyses of *OsSAT* genes showed that SATs are subjected to the purifying selection. Purifying selection, also known as background selection, reduces genetic diversity and shapes it in natural population [40]. Consequently, it may be suggested that the genetic diversity of *OsSAT* genes decreased as a result of purifying selection.

R value, the ratio of transition to transversion, were estimated for DNA sequence evolution and phylogeny reconstruction [41]. In any genome, transitions (T↔C and A↔G) are observed at higher frequencies than transversions (T↔A, T↔G, C↔A, and C↔G). In this study, the estimated transition/transversion bias (R) was found as 0.71, indicating genetic variations. Proving this result, GC contents of OsSATs also showed a considerable variation. Genomic DNA base composition (GC content) affects genes' functions and adaptation of species to its environment and it may play roles in complex gene regulation [42]. Consequently, it can be concluded that the action of purifying selection may increase the specificity and the selectivity of SATs in rice metabolism, leading to variations of G+C contents. In addition to that, we found that segmental and tandem duplications are driving force of OsSATs evolution. It is known that gene duplication is one of types of genomic change that can lead to evolutionary changes. The duplicated genes can contribute to the evolution of novel functions including adaptation to stress, induction of disease resistance, production of floral structures [43] and expansion of gene families [44]. As is found in this study, it is reported that *SAT2;1/SAT2;2* and *SAT3;1/SAT3;2* isoforms in *Arabidopsis* were also duplicated gene pairs [45].

Although all SATs contain serine acetyltransferase N-terminal domain structure (SATase\_N, PF06426) and bacterial transferase hexapeptide (PF00132) as their common motifs, there is still variations in motif structures. The presence of protein motifs may play important roles in protein function. The motifs in active sites of proteins are well-conserved [46], suggesting that these variations in motif structures may be connected with functional diversities of SATs in plants.

Expression patterns of OsSATs showed difference according to tissues and organs, and treatments. This result is validated by the co-expression maps of OsSAT genes showing their involvement in various metabolic pathways and their association with gene families with different functions. When the expression levels of OsSATs by organs are taken into consideration, it was observed that the expression levels of *OsSAT2;1* and *OsSAT2;2* genes showed more dynamic profiles. Jasmonic acid is the only hormone activating expression of *OsSAT1;1*, *OsSAT2;1*, and *OsSAT2;2*. Besides, *OsSAT1;1* was the only OsSAT induced by auxin and abscisic acid hormones. The expressions of OsSATs under different salt treatments conducted within study showed that *OsSAT1;1*, *OsSAT1;2* and *OsSAT1;3* were generally upregulated depending on different salt exposure times. However, *OsSAT1;2* showed the highest expression levels when exposed to 3, 12, and 24-hour salt treatments. Conversely, *OsSAT2;2*, *OsSAT3* and particularly *OsSAT2;1* were downregulated depending on salt exposure times. Watanabe et al. [45] stated that three cytosolic isoforms, *SAT1;1*, *SAT3;1*, and *SAT3;2* genes in *Arabidopsis* contribute to



seed development and *SAT* gene family plays essential roles for plant survive. Also, *SAT3;1* and *SAT3;2* isoforms of *Arabidopsis* play roles in plant development. In *Arabidopsis*, *AtSAT2* and *AtSAT4* are 10–100 times less transcribed compared to the major expressed SAT isoenzymes such as *AtSAT1*, 3 and 5 [9, 47]. Everything considered, the varied expression patterns of *OsSATs* under various conditions show that they are dynamically regulated and *OsSAT1;1* is the only *OsSAT* induced by various stimuli.

The expressional similarities among *OsSATs* may be originated from preserved residues in their active sites taking effect on their functions. As stated earlier, Asp (D), His (H), Gly (G), Thr (T), Arg (R), Ala (A), and Leu (L) are the conserved residues in active sites of *OsSATs*. His169, Asp154, His18, Arg203 and His204, Lys230, Arg253 in soybean were identified as active residues involved in reactions such as catalysis, oxyanion reaction intermediate, serine binding, CoA binding according to crystal structures and analysis of site-directed mutation data [48]. In this study, similar residues were identified in predicted active sites of *OsSAT* proteins.

## Conclusion

A total of six *OsSAT* genes were identified in rice genome and variations at gene and protein structures were identified using bioinformatics approaches. It is found that jasmonic acid induced the expressions of *OsSAT1;1*, *OsSAT2;1*, and *OsSAT2;2* whereas auxin and abscisic acid induced only *OsSAT1;1*. On the other hand, *OsSAT1;1*, *OsSAT1;2* and *OsSAT1;3* genes were upregulated under different exposure times of salt stress. *OsSAT1;1* is the only *OsSAT* gene induced by various stimuli.

*SAT* gene family play specific roles in plant metabolism, particularly sulfur assimilation pathway depending on purifying selection which decreased genetic diversity of them. In addition to purifying selection, segmental and tandem duplications may lead to *OsSATs* to have more specific and selective roles in metabolic pathways which may have an effect on the plant's responses to abiotic and biotic stress conditions. Therefore, the findings can be used by plant breeders and genetic engineers to develop new rice varieties having optimal growth and stress tolerance.

## Declarations

## Author contribution

E.F and F.K. designed the study. E.F., F.K. and A.A. conducted the experiments and analyzed the data. All authors contributed to writing and editing the manuscript.

## Conflict of interest

The authors declare that they have no conflict of interest

## Ethical approval

This article does not contain any studies with human participants or animals performed by any of the authors

## References

1. Leustek T (2002) Sulfate Metabolism. *Arab B* 1:e0017. <https://doi.org/10.1199/tab.0017>
2. Droux M (2003) Plant serine acetyltransferase: new insights for regulation of sulphur metabolism in plant cells. *Plant Physiol Biochem* 41:619–627. [https://doi.org/10.1016/s0981-9428\(03\)00083-4](https://doi.org/10.1016/s0981-9428(03)00083-4)
3. Haas FH, Heeg C, Queiroz R, et al (2008) Mitochondrial Serine Acetyltransferase Functions as a Pacemaker of Cysteine Synthesis in Plant Cells. *Plant Physiol* 148:1055–1067. <https://doi.org/10.1104/pp.108.125237>
4. Droux M, Ruffet M-L, Douce R, Job D (1998) Interactions between serine acetyltransferase and O-acetylserine (thiol) lyase in higher plants. Structural and kinetic properties of the free and bound enzymes. *Eur J Biochem* 255:235–245. <https://doi.org/10.1046/j.1432-1327.1998.2550235.x>
5. Wirtz M, Hell R (2006) Functional analysis of the cysteine synthase protein complex from plants: Structural, biochemical and regulatory properties. *J Plant Physiol* 163:273–286. <https://doi.org/10.1016/j.jplph.2005.11.013>
6. Romero LC, Aroca MÁ, Laureano-Marín AM, et al (2014) Cysteine and Cysteine-Related Signaling Pathways in *Arabidopsis thaliana*. *Mol Plant* 7:264–276. <https://doi.org/10.1093/mp/sst168>
7. Kopriva S, Patron NJ, Keeling P, Leustek T (2008) Phylogenetic Analysis of Sulfate Assimilation and Cysteine Biosynthesis in Phototrophic Organisms. *Sulfur Metab. Phototrophic Org.* 31–58
8. Hell R, Jost R, Berkowitz O, Wirtz M (2002) Molecular and biochemical analysis of the enzymes of cysteine biosynthesis in the plant *Arabidopsis thaliana*. *Amino Acids* 22:245–257. <https://doi.org/10.1007/s007260200012>
9. Kawashima CG, Berkowitz O, Hell R, et al (2005) Characterization and Expression Analysis of a Serine Acetyltransferase Gene Family Involved in a Key Step of the Sulfur Assimilation Pathway in *Arabidopsis*. *Plant Physiol* 137:220–230. <https://doi.org/10.1104/pp.104.045377>

10. Xiang X, Wu Y, Planta J, et al (2018) Overexpression of serine acetyltransferase in maize leaves increases seed-specific methionine-rich zeins. *Plant Biotechnol J* 16:1057–1067. <https://doi.org/10.1111/pbi.12851>
11. Goodstein DM, Shu S, Howson R, et al (2012) Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res* 40:D1178–D1186. <https://doi.org/10.1093/nar/gkr944>
12. Finn RD, Coghill P, Eberhardt RY, et al (2015) The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res* 44:D279–D285. <https://doi.org/10.1093/nar/gkv1344>
13. Gasteiger E, Hoogland C, Gattiker A, et al (2005) Protein Identification and Analysis Tools on the ExPASy Server. *Proteomics Protoc. Handb.* 571–607
14. Hall TA (1999) Hall(1999)*Nucleic\_Acids\_Symp\_Ser\_41,95*.[\[BioEdit\].pdf](#). *Nucleic Acids Symp Ser* 41:95–98. <https://doi.org/citeulike-article-id:691774>
15. Törönen P, Medlar A, Holm L (2018) PANNZER2: a rapid functional annotation web server. *Nucleic Acids Res* 46:W84–W88. <https://doi.org/10.1093/nar/gky350>
16. Bailey TL, Johnson J, Grant CE, Noble WS (2015) The MEME Suite. *Nucleic Acids Res* 43:W39–W49. <https://doi.org/10.1093/nar/gkv416>
17. Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol* 16:111–120. <https://doi.org/10.1007/bf01731581>
18. Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, et al (2017) DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets. *Mol Biol Evol* 34:3299–3302. <https://doi.org/10.1093/molbev/msx248>
19. Jones DT, Taylor WR, Thornton JM (1992) The rapid generation of mutation data matrices from protein sequences. *Bioinformatics* 8:275–282. <https://doi.org/10.1093/bioinformatics/8.3.275>
20. Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol* 33:1870–1874. <https://doi.org/10.1093/molbev/msw054>
21. Tajima F, Misawa K, Innan H (1998) The amount and pattern of DNA polymorphism under the neutral mutation hypothesis. *Mutat. Evol.* 103–107
22. Yang S, Zhang X, Yue J-X, et al (2008) Recent duplications dominate NBS-encoding gene expansion in two woody species. *Mol Genet Genomics* 280:187–198. <https://doi.org/10.1007/s00438-008-0355-0>
23. Darzentas N (2010) Circoletto: visualizing sequence similarity with Circos. *Bioinformatics* 26:2620–2621. <https://doi.org/10.1093/bioinformatics/btq484>
24. Lee T-H, Tang H, Wang X, Paterson AH (2012) PGDD: a database of gene and genome duplication in plants. *Nucleic Acids Res* 41:D1152–D1158. <https://doi.org/10.1093/nar/gks1104>
25. Sato Y, Takehisa H, Kamatsuki K, et al (2013) RiceXPro Version 3.0: expanding the informatics resource for rice transcriptome. *Nucleic Acids Res* 41:D1206–D1213. <https://doi.org/10.1093/nar/gks1125>
26. Sato Y, Namiki N, Takehisa H, et al (2013) RiceFRIEND: a platform for retrieving coexpressed gene networks in rice. *Nucleic Acids Res* 41:D1214–D1221. <https://doi.org/10.1093/nar/gks1122>
27. Geourjon C, Deléage G (1995) SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. *Bioinformatics* 11:681–684. <https://doi.org/10.1093/bioinformatics/11.6.681>
28. Kelley LA, Sternberg MJE (2009) Protein structure prediction on the Web: a case study using the Phyre server. *Nat Protoc* 4:363–371. <https://doi.org/10.1038/nprot.2009.2>
29. Guex N, Peitsch MC, Schwede T (2009) Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. *Electrophoresis* 30:S162–S173. <https://doi.org/10.1002/elps.200900140>
30. Nguyen MN, Tan KP, Madhusudhan MS (2011) CLICK - Topology-independent comparison of biomolecular 3D structures. *Nucleic Acids Res* 39:24–28. <https://doi.org/10.1093/nar/gkr393>
31. Jones P, Binns D, Chang H-Y, et al (2014) InterProScan 5: genome-scale protein function classification. *Bioinformatics* 30:1236–1240. <https://doi.org/10.1093/bioinformatics/btu031>
32. Szklarczyk D, Franceschini A, Wyder S, et al (2014) STRING v10: protein–protein interaction networks, integrated over the tree of life. *Nucleic Acids Res* 43:D447–D452. <https://doi.org/10.1093/nar/gku1003>
33. Kim JH, Lim SD, Jang CS (2020) *Oryza sativa* drought-, heat-, and salt-induced RING finger protein 1 (OsDHSRP1) negatively regulates abiotic stress-responsive gene expression. *Plant Mol Biol* 103:235–252. <https://doi.org/10.1007/s11103-020-00989-x>
34. Rademacher EH, Offringa R (2012) Evolutionary Adaptations of Plant AGC Kinases: From Light Signaling to Cell Polarity Regulation. *Front Plant Sci* 3. <https://doi.org/10.3389/fpls.2012.00250>
35. Xu C, Luo F, Hochholdinger F (2016) LOB Domain Proteins: Beyond Lateral Organ Boundaries. *Trends Plant Sci* 21:159–167. <https://doi.org/10.1016/j.tplants.2015.10.010>
36. Chen Y-C, Li C-L, Hsiao Y-Y, et al (2014) Structure and function of TatD exonuclease in DNA repair. *Nucleic Acids Res* 42:10776–10785. <https://doi.org/10.1093/nar/gku732>
37. Bækgaard L, Mikkelsen MD, Sørensen DM, et al (2010) A Combined Zinc/Cadmium Sensor and Zinc/Cadmium Export Regulator in a Heavy Metal Pump. *J Biol Chem* 285:31243–31252. <https://doi.org/10.1074/jbc.m110.111260>
38. Marks DS, Colwell LJ, Sheridan R, et al (2011) Protein 3D Structure Computed from Evolutionary Sequence Variation. *PLoS One* 6:e28766. <https://doi.org/10.1371/journal.pone.0028766>

39. Sankararaman S, Sha F, Kirsch JF, et al (2010) Active site prediction using evolutionary and structural information. *Bioinformatics* 26:617–624. <https://doi.org/10.1093/bioinformatics/btq008>
40. Cvijović I, Good BH, Desai MM (2018) The Effect of Strong Purifying Selection on Genetic Diversity. *Genetics* 209:1235–1278. <https://doi.org/10.1534/genetics.118.301058>
41. Yang Z, Yoder AD (1999) Estimation of the Transition/Transversion Rate Bias and Species Sampling. *J Mol Evol* 48:274–283. <https://doi.org/10.1007/pl00006470>
42. Šmarda P, Bureš P, Horová L, et al (2014) Ecological and evolutionary significance of genomic GC content diversity in monocots. *Proc Natl Acad Sci* 111:E4096–E4102. <https://doi.org/10.1073/pnas.1321152111>
43. Panchy N, Lehti-Shiu M, Shiu SH (2016) Evolution of gene duplication in plants. *Plant Physiol* 171:2294–2316. <https://doi.org/10.1104/pp.16.00523>
44. Kuo Y-T, Chao Y-T, Chen W-C, et al (2019) Segmental and tandem chromosome duplications led to divergent evolution of the chalcone synthase gene family in *Phalaenopsis* orchids. *Ann Bot* 123:69–77. <https://doi.org/10.1093/aob/mcy136>
45. Watanabe M, Mochida K, Kato T, et al (2008) Comparative Genomics and Reverse Genetics Analysis Reveal Indispensable Functions of the Serine Acetyltransferase Gene Family in *Arabidopsis*. *Plant Cell* 20:2484–2496. <https://doi.org/10.1105/tpc.108.060335>
46. Ben-Hur A, Brutlag D (2006) Sequence Motifs: Highly Predictive Features of Protein Function. *Featur. Extr.* 625–645
47. Wirtz M, Droux M (2005) Synthesis of the sulfur amino acids: cysteine and methionine. *Photosynth Res* 86:345–362. <https://doi.org/10.1007/s11220-005-8810-9>
48. Yi H, Dey S, Kumaran S, et al (2013) Structure of Soybean Serine Acetyltransferase and Formation of the Cysteine Regulatory Complex as a Molecular Chaperone. *J Biol Chem* 288:36463–36472. <https://doi.org/10.1074/jbc.m113.527143>

## Figures

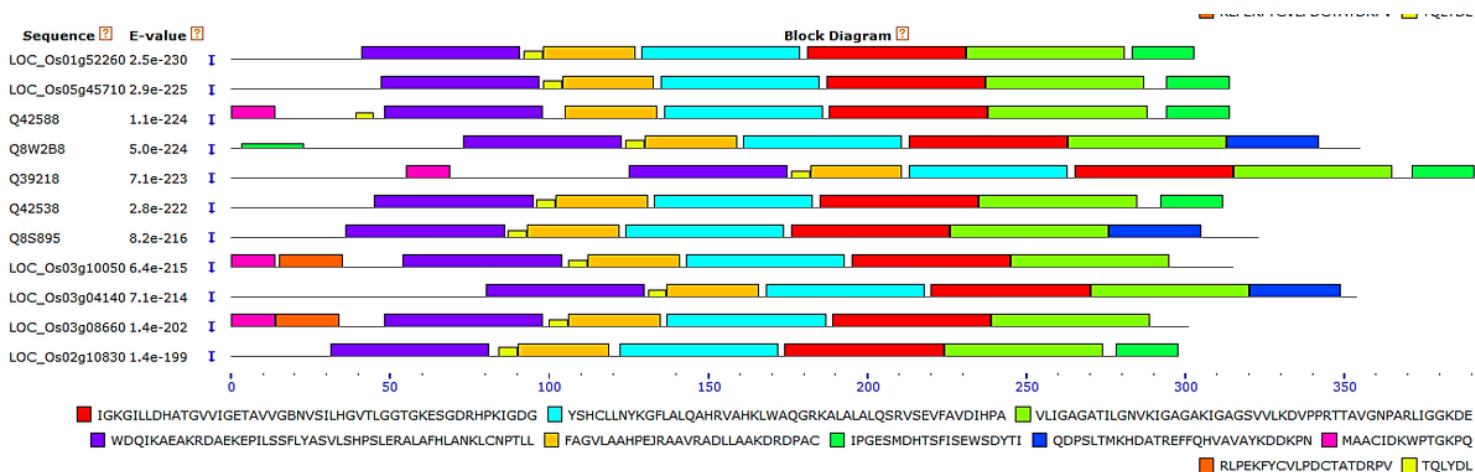
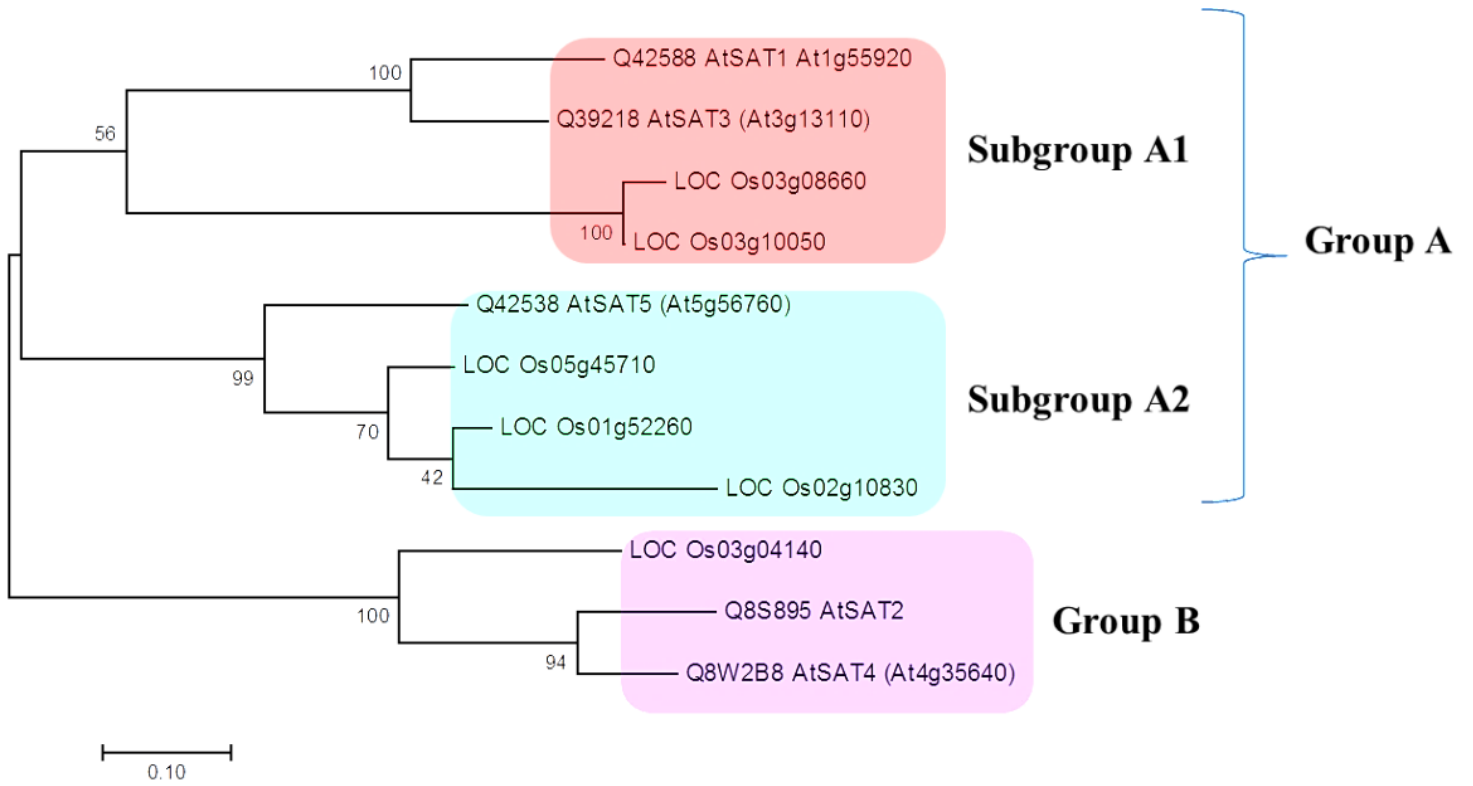


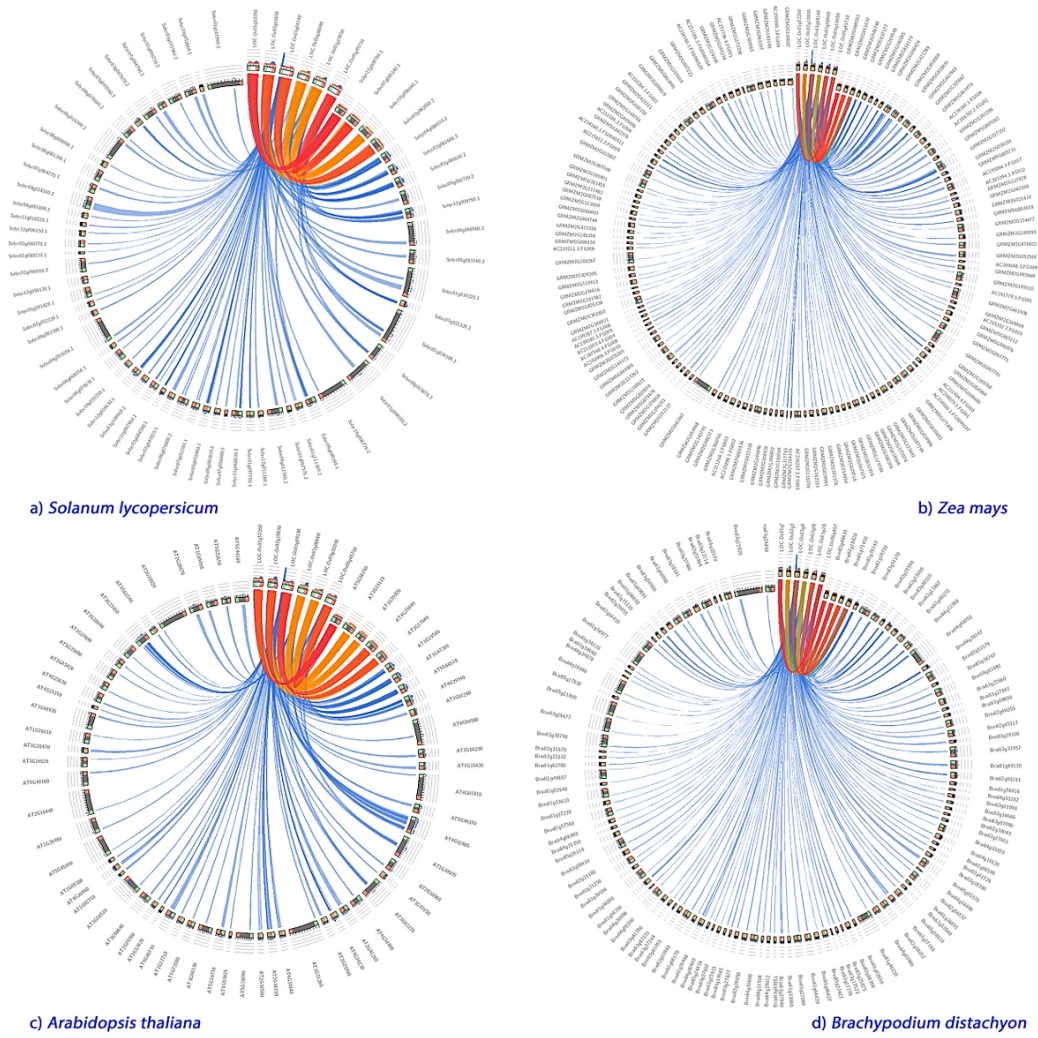
Figure 1

Block diagram of conserved motifs in 11 SAT protein sequences from rice and Arabidopsis using MEME server. Each color represents the distinct motif structure.



**Figure 2**

Phylogenetic tree of SAT protein family based on SAT amino acid sequences of rice and Arabidopsis.



**Figure 3**  
 Synteny analyses of *OsSAT* genes in tomato (a), maize (b), Arabidopsis (c), and *B. distachyon* (d). Blue lines in the background indicate the collinear blocks within rice and other plant genomes whilst the red and yellow lines show the syntenic SAT gene pairs.

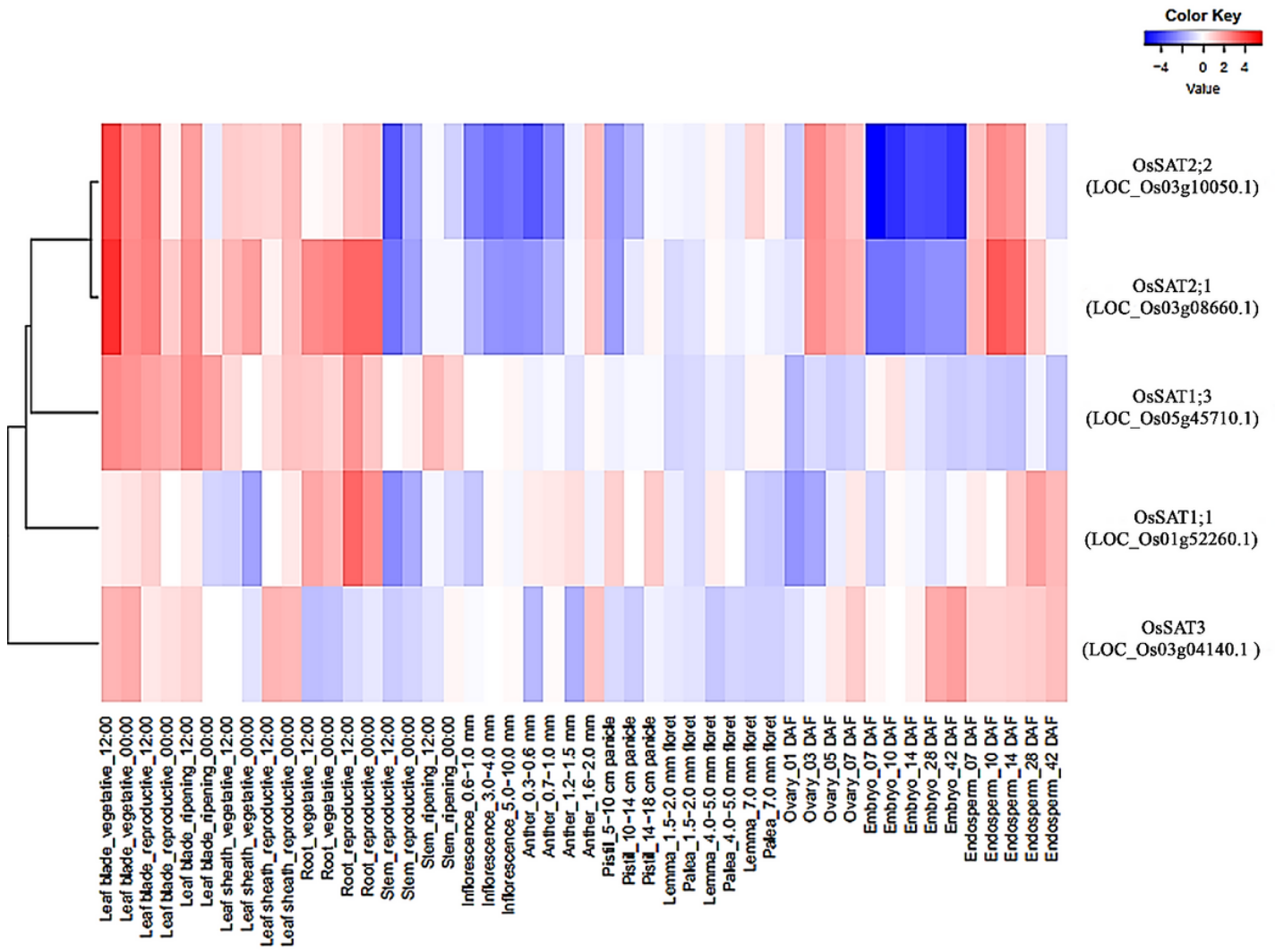


Figure 4

Spatio-temporal OsSAT gene expression of various tissues/organs throughout entire growth in the field. The expression data is retrieved from Rice Expression Profile Database (RiceXPro)

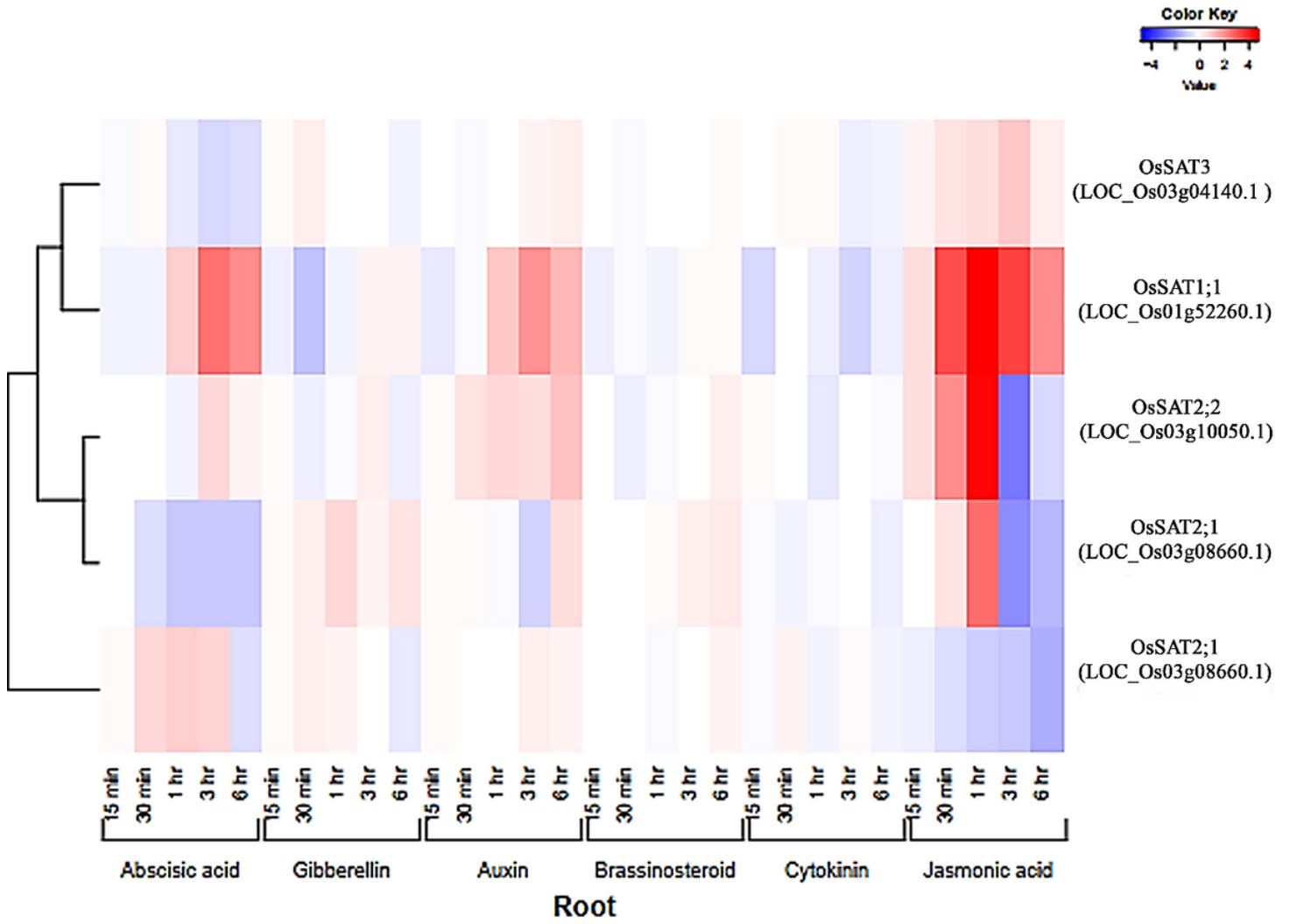


Figure 5  
Global OsSAT gene expression profile in response to plant hormones. The expression data is retrieved from Rice Expression Profile Database (RiceXPro)

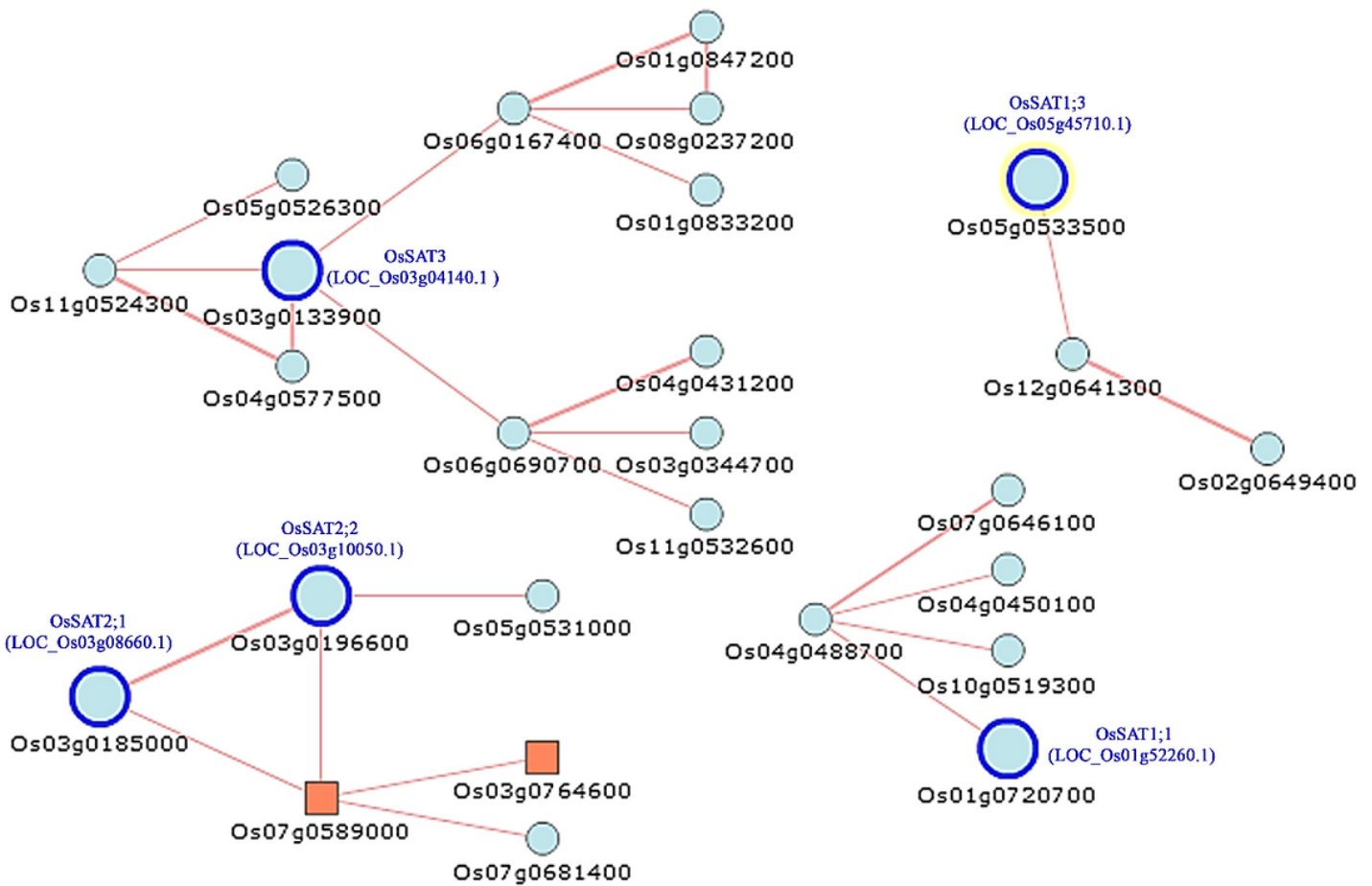


Figure 6

Co-expression networks of OsSAT genes (blue-circled). The each network was generated using the data on RiceFRIEND gene co-expression database.

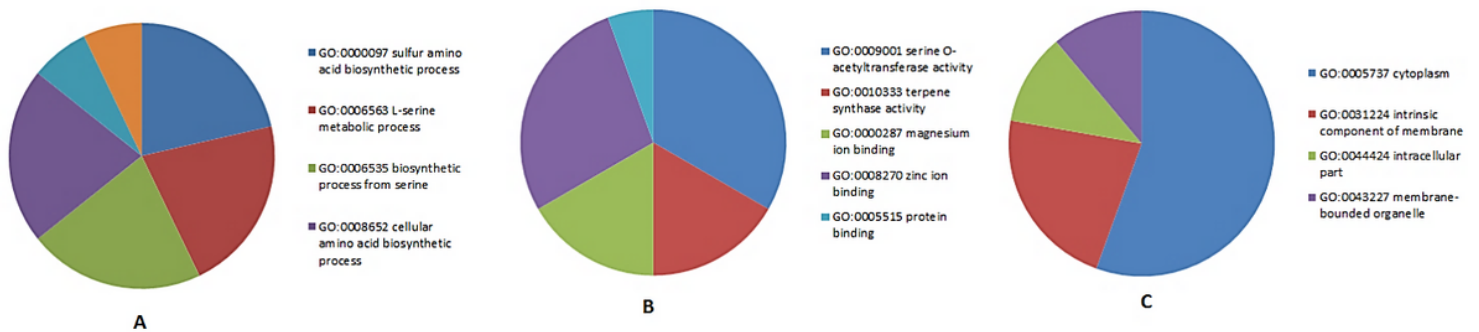


Figure 7

GO annotation of OsSAT proteins using PANNZER server. The annotation was done on three categories: (a) biological process, (b) molecular function, and (c) cellular component.



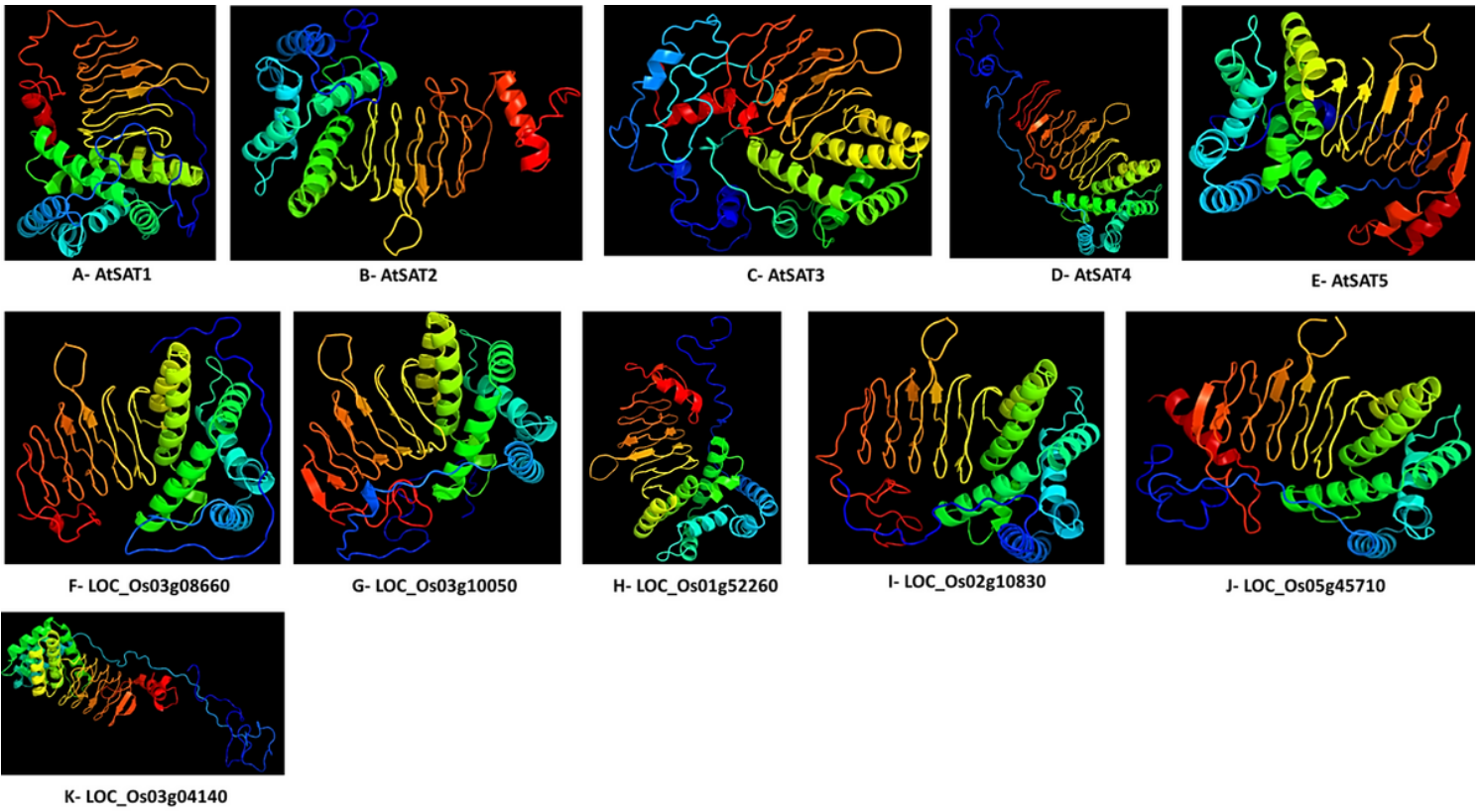


Figure 8

The predicted 3D model of Arabidopsis and rice SAT proteins

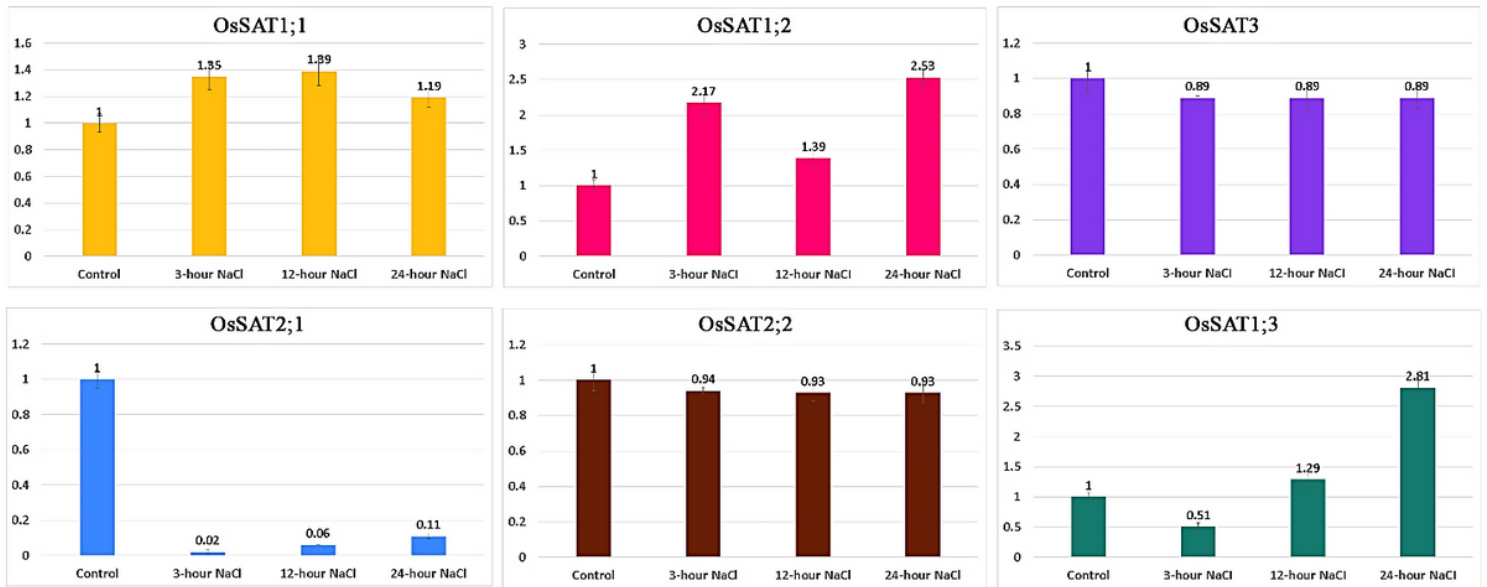


Figure 9

The expression of OsSATs under salt stress. Each gene was shown in different color. The error bars on the graphs show the standard deviations of expressions.