Genome-wide identification of serine acetyltransferase (SAT) gene family in rice (Oryza sativa) and their expressions under salt stress

Fırat Kurt1 · Ertugrul Filiz2 · Adnan Aydın3

Received: 16 April 2021 / Accepted: 3 August 2021 / Published online: 13 August 2021
© The Author(s), under exclusive licence to Springer Nature B.V. 2021

Abstract

Background Assimilation of sulfur to cysteine (Cys) occurs in presence of serine acetyltransferase (SAT). Drought and salt stresses are known to be regulated by abscisic acid, whose biosynthesis is limited by Cys. Cys is formed by cysteine synthase complex depending on SAT and OASTL enzymes. Functions of some SAT genes were identified in Arabidopsis; however, it is not known how SAT genes are regulated in rice (Oryza sativa) under salt stress.

Methods and results Sequence, protein domain, gene structure, nucleotide, phylogenetic, selection, gene duplication, motif, synteny, digital expression and co-expression, secondary and tertiary protein structures, and binding site analyses were conducted. The wet-lab expressions of OsSAT genes were also tested under salt stress. OsSATs have undergone purifying selection. Segmental and tandem duplications may be driving force of structural and functional divergences of OsSATs.

The digital expression analyses of OsSATs showed that jasmonic acid (JA) was the only hormone inducing the expressions of OsSAT1;1, OsSAT2;1, and OsSAT2;2 whereas auxin and ABA only triggered OsSAT1;1 expression. Leaf blade is the only plant organ where all OsSATs but OsSAT1;1 were expressed. Wet-lab expressions of OsSATs indicated that OsSAT1;1, OsSAT1;2 and OsSAT1;3 genes were upregulated at different exposure times of salt stress.

Conclusions OsSAT1;1, expressed highly in rice roots, may be a hub gene regulated by cross-talk of JA, ABA and auxin hormones. The cross-talk of the mentioned hormones and the structural variations of OsSAT proteins may also explain the different responses of OsSATs to salt stress.

Keywords Sulfur · Rice · Hormone · Salt stress · Bioinformatics

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 (SO4^{2−}) and convert it to sulfide (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 plant cells [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 intermediate product O-acetylserine (OAS) using acetyl-CoA and serine. The latter uses the sulfide and OAS for production of Cys as a cofactor in 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 localized in plastids, mitochondria, and cytoplasm, respectively [2]. Mitochondrion is identified as the most important compartment for O-acetylserine (OAS) synthesis in Arabidopsis, 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 γ-zein storage protein during endosperm development [10].

Drought and salt stresses are known to be regulated by abscisic acid, whose biosynthesis is limited by Cys. Cys is formed by cysteine synthase complex depending on SAT and OASTL enzymes. It is reported that *AtSAT2;1* was induced when *Arabidopsis* seedlings were treated by 50 µM ABA [11]. Also, *SAT1* in the chloroplast of *Arabidopsis* were significantly upregulated under salt and light stresses [12]. However, it is not known how *SAT* genes are regulated in rice under salt stress. Therefore, in present study, rice (*Oryza sativa*) SAT (*OsSATs*) genes/proteins were identified at genome-wide level, 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 into their regulations under salt stress.

**Materials and methods**

**Identification of SAT genes and sequence analyses**

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) with an E-value < 1E−20 [13]. Later, the presence of PF06426 (SATase_N) domain structure was validated using Pfam database v32 (https://pfam.xfam.org/) [14]. Genomic and coding sequences corresponding to SATs were collected from Phytozome database for further bioinformatics analyses. The domains of each SAT proteins were searched using HMMER server (https://www.ebi.ac.uk/Tools/hmmer/) [15] and the found domains were identified using multiple sequence alignment tool on Bioedit v7.2.5 [16]. Gene structure analysis was performed by GSDS server [17]. Physicochemical properties of SAT proteins such as length, molecular weight, and isoelectric point (pI) values were obtained from ProtParam tool (web.expasy.org/protparam) [18]. The sequence identity matrix for *Arabidopsis* and rice SATs were also generated using Bioedit v7.2.5 [16]. Annotations of SAT proteins were analyzed using PANNZER (Protein ANNotation with Z-score) server (http://ekhidna2.biocenter.helsinki.fi/sanspanzl/) [19]. 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 [20].

**Nucleotide and phylogenetic analyses**

Maximum likelihood estimate of transition/transversion bias (R) was estimated by Kimura’s method [21] using two-parameter model. The estimate of transition/transversion bias and G+C content were calculated using DnaSP v6.12.01 [22]. Molecular phylogenetic analysis was performed according to the maximum likelihood (ML) method using JTT matrix-based model [23]. 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 using MEGA7 software [24].

**Selection, gene duplication and synteny analyses**

Nucleotide sequences of rice SAT genes were analyzed using DnaSP v6.12.01 [22] in terms of polymorphic sites, DNA polymorphism, genetic variation level (π and θ), and Tajima’s D [25]. Ka and Ks values were calculated for duplicated gene pairs using DnaSP v6.12.01 [22]. Gene duplication analyses were conducted according to following two criteria: If the alignment of the coding nucleotide sequences covers 70% of the longest genes and the amino acid identity between the sequences are ≥ 70%, the gene of interest is assumed to be duplicated [26].

Synteny analyses were conducted on Circos (2004) platform [27]. With this aim, processed genes of *Arabidopsis thaliana*, *Brachypodium distachyon*, *Solanum lycopersicum*, and *Zea mays* were downloaded from Plant Genome Duplication Database [28]. Then the local blast database was constructed in Bioedit v7.2.5 [16] and the blast results were fed to Circos server.

**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/) [29]. Co-expression
network was constructed using RiceFRIEND co-expression database. RiceFRIEND contains expression profiles obtained from microarray analyses of tissues/organs of rice plants under natural field conditions and various treatments (http://ricefrend.dna.affrc.go.jp/) [30].

**Analyses of secondary and tertiary structures, and predicted binding sites**

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) [31]. Tertiary models were predicted by using Phyre2 server (Protein Homology/analogY Recognition Engine V 2.0 (https://www.sbg.bio.ic.ac.uk/phyre2/) [32]. The independent comparisons of biomolecular 3D structures were performed using CLICK server (http://www.mspc.bii.a-star.edu.sg/click) [34]. The predicted 3D models were visualized using Swiss-PdbViewer [33]. The binding sites of OsSAT protein sequences were identified using InterProScan tool (http://www.ebi.ac.uk/interpro/search/sequence-search) [35].

**Plant material and growth conditions**

The 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-h dark stage and 30 °C at 16-h light period. 4–5 plants were grown for each experimental group. Rice plants at third leaf stage were exposed to 200 mM NaCl and leaf samples were collected at the 3rd-, 12th- and 24th-hour. Leaf samples were immediately stored in RNA later 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 the protocol of the DNaseI kit (Cat: EN0521) into account. 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 [36]. 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 [37]. Relative Quantification \(2^{-\Delta \Delta C_T} [38]\), 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 [39] (http://bioinfo.ut.ee/primer3-0.4.0/) and synthesized by Macrogen Crop. (Amsterdam, The Netherlands).

**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. In terms of gene structures, OsSAT1;1, OsSAT1;3, OsSAT2;2 have similar gene structures whereas OsSAT1;2 and OsSAT2;1 are alike. (Supp. Fig. 1). AtSATS are more diverse and AtSAT1 and AtSAT3 are more similar to each other. All genes contain upstream and downstream regions except for OsSAT1;2 and OsSAT2;1. All SAT proteins in the

**Table 1 Primers used for RT-qPCR analysis of OsSATS genes**

| Primer ID | Sequences (5′→3′) | Gene ID       |
|-----------|-------------------|---------------|
| OsSAT1;1  | F:GACGAGTCGGTGGGCTGGGCGGT  | LOC_Os01g52260.1 |
|           | R:GGTGGCGTACAGAGGAACGAGC |               |
| OsSAT1;2  | F:ACTCCAGACGAGCATGACGCAG  | LOC_Os02g10830.1 |
|           | R:CCCTGATCTGGGACCACAC       |               |
| OsSAT3   | F:TGAAACTGTCGATTTGTGGAAGCA  | LOC_Os03g04140.1 |
|           | R:TGACCAATCTTGCGGTGCTCTCTCT |               |
| OsSAT2;1  | F:ATCGTGGCTACGGGACGTGACAGG  | LOC_Os03g08660.1 |
|           | R:GTACCCCGGAAAGGTCTCTCTCTCT |               |
| OsSAT2;2  | F:ACAGCAGGACGTGGTGTCTCTCTCT | LOC_Os03g10050.1 |
|           | R:ACAGCAGGACGTGGTGTCTCTCTCT |               |
| OsSAT1;3  | F:ACAGCAGGACGTGGTGTCTCTCTCT | LOC_Os05g45710.1 |
|           | R:ACAGCAGGACGTGGTGTCTCTCTCT |               |
study contain the serine acetyltransferase N-terminal domain structure (SATase_N, PF06426). In addition, bacterial transferase hexapeptide (PF00132) domain was identified as two different sequences (Supp. Fig. 2). According to domain analyses of SAT proteins, WIKM-LEEAKSDVKQPILSNYYASITSHRSALAHILSVKL5NLPSNTLFELFISVLEESPEIIESTKQDLIAV sequence was the conserved serine acetyltransferase domain whereas GVVIGETAVVGDNVSILHGVTLG and IGDGVLIGAGSCILGNITIGEGAKIGGSV sequences were the hexapeptide domains with 100% match score (Supp. Fig. 1). Domain boundaries and the number of hexapeptide domains showed variation among the sequences. The SATs shown with with flicker pink filled ellipses have only the second bacterial transferase hexapeptide domain (Supp. Fig. 2).

To provide more insight about protein sequence diversity of SATs, sequence identity matrix was constructed (Supp. Table 1). The lowest identity value was 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.

Conserved motif analysis

Regarding conserved motifs, 10 conserved motifs were detected (Table 3 and Supp. Fig. 3). 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.

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.

### 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) | A. thaliana | 1        | 314                 | PF06426       | 2             | 34.25          | 6.45 |
| At2g17640 (SAT2, SERAT3;1) | A. thaliana | 10       | 323                 | PF06426       | 2             | 34.53          | 5.81 |
| At3g13110 (SAT3, SERAT2;2) | A. thaliana | 1        | 391                 | PF06426       | 1             | 42.72          | 7.73 |
| At4g35640 (SAT4, SERAT3;2) | A. thaliana | 1        | 355                 | PF06426       | 1             | 38.42          | 5.56 |
| At5g6760 (SAT5, SERAT1;1) | A. thaliana | 2        | 312                 | PF06426       | 2             | 32.77          | 6.70 |
| LOC_Os01g52260 (OsSAT1;1) | O. sativa   | 1        | 303                 | PF06426       | 2             | 31.85          | 6.66 |
| LOC_Os02g10830 (OSSAT1;2) | O. sativa   | 1        | 298                 | PF06426       | 2             | 30.44          | 5.33 |
| LOC_Os03g04140 (OsSAT3) | O. sativa   | 1        | 303                 | PF06426       | 2             | 30.44          | 5.33 |
| LOC_Os03g08560 (OsSAT2;1) | O. sativa   | 1        | 301                 | PF06426       | 1             | 31.50          | 5.91 |
| LOC_Os03g10050 (OsSAT2;2) | O. sativa   | 1        | 315                 | PF06426       | 1             | 32.92          | 6.20 |
| LOC_Os05g45710 (OsSAT1;3) | O. sativa   | 1        | 314                 | PF06426       | 1             | 33.09          | 6.38 |

PF06426 (SATase_N): serine acetyltransferase, N-terminal, Repeat (PF00132): bacterial transferase hexapeptide (six repeats)
Phylogenetic analysis indicated that AtSATs and OsSATs split into two major clades (Group A and B). Group A also divided into two subclades (Fig. 1). OsSAT2;1 (LOC_Os03g08680) 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 (Fig. 1).

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 π and θ parameters, respectively. Tajima’s D was found 1.22, indicating purifying (negative) selection.

In the second selection test, 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 4).

To understand gene duplication dynamics of OsSAT genes four comparative syntenic maps of rice and Arabidopsis were generated (Fig. 2). 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.

### Table 3 The details of conserved motif sequences in SAT proteins

| No. | Motif sequence                                                                 | S   | W  | Domain              |
|-----|--------------------------------------------------------------------------------|-----|----|---------------------|
| 1   | IGKGILLDHATGVVIGETAVGVBNVSIHLGVTLLGKESGDRHKIGDG                                  | 11  | 50 | NF                  |
| 2   | YSHCLNYKGFALQAHRVAHKLWQGRKALALQSRVSEQVFAVIIIHA                                  | 11  | 50 | NF                  |
| 3   | VLGAGATILGNVKIGAKIGAKIGAVKVSPKTPAVGNPARLIGGD                                    | 11  | 50 | PF00132             |
| 4   | WDIQKAEKRDAEKPISSLYAFLSHPMLERLAFLHANKLCNPTLL                                    | 11  | 50 | PF06426             |
| 5   | FAVGLAHARAAVADLAAKDRDPAC                                                         | 11  | 29 | PF06426             |
| 6   | IPGEMHDSMFSIEWSYTI                                                               | 6   | 20 | NF                  |
| 7   | QDPLTMKHDATREFFQHVAHVAYKDDKPN                                                    | 3   | 29 | NF                  |
| 8   | MAACIDKWPTGKPQ                                                                   | 4   | 14 | NF                  |
| 9   | RLPEKPYCVLPDCTATDRPV                                                             | 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
The co-expression network of OsSAT genes was constructed using RiceFREND 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. 5). 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 [40]. 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 [41]. 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 [42]. 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 [43].

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 (Supp. Fig. 4). 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 (Supp. Fig. 4A). Serine O-acetyltransferase activity (GO:0009001), terepene synthase activity (GO:0010333), magnesium ion binding (GO:000287), zinc ion binding (GO:0008270), and protein binding (GO:0005515) were predicted molecular functions which are OsSATs are carried out (Supp.

Table 4 Segmental and tandem duplications of SAT paralogous pairs in rice genome

| SAT group   | Chr. location | Duplication type | Ka   | Ks   | Ka/Ks  |
|-------------|---------------|-----------------|------|------|--------|
| OsSAT1;1    | Chr1          | Segmental       | 0.114| 0.184| 0.619  |
| OsSAT1;2    | Chr2          | Segmental       | 0.073| 0.296| 0.247  |
| OsSAT1;3    | Chr1          | Segmental       | 0.146| 0.363| 0.402  |
| OsSAT1;3    | Chr5          | Segmental       | 0.018| 0.023| 0.783  |
| OsSAT2;1    | Chr3          | Tandem          | 0.146| 0.363| 0.402  |
| OsSAT2;2    | Chr3          | Tandem          | 0.073| 0.296| 0.247  |

Non-synonymous (Ka) and synonymous (Ks) indicate the substitution rates. Ka/Ks is non-synonymous/synonymous mutation ratio.

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. 3. For digital expression analysis, five OsSAT genes were obtained from Rice Expression Profile Database (RiceX-Pro) 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 about three fold changes. Also, these two genes were down-regulated 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 (ascbic acid, gibberellin, auxin, brassinosteroid, cytokinin, and jasmonic acid) (Fig. 4). Jasmonic acid (JA) treatment especially caused the upregulation of OsSAT genes in terms of gene expression levels. Particularly OsSAT1;1 gene was upregulated nearly four fold changes by JA. Also, it was found that the optimal time point for the expression to increase was 1 h for OsSAT1;1 gene under JA treatment. Also, OsSAT1;1 responded to the auxin and ABA treatments at the third hour of exposure although its responses were not as high as to those of JA. The upregulation of OsSAT1;1 gene also was observed under ABA and auxin treatments. OsSAT2;1 and OsSAT2;2 genes were upregulated at the first hour of exposure to JA acid. Overall, the levels of gene expressions under hormone treatments supported that OsSAT1;1, OsSAT2;1, and OsSAT2;2 genes were positively regulated by hormone treatments.
Fig. 4B). 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 (Supp. Fig. 4C). All things considered, 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 of OsSATs, 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 2). The predicted 3D structures of OsSAT proteins (Fig. 6) were found reliable due to and their

Fig. 2 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. (Color figure online)
Ramachandran values, ranging from 96 to 99% in core and allowed regions. Overall, structural differences of SATs may indicate their functional differences.

The 3D structural similarities (%) were identified using six rice, five Arabidopsis, and one soybean SAT proteins on CLICK structure comparison server (Table 5). 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. Overall, OsSATs are more structurally similar to AtSATs proving structural divergence among SATs.

Predicted active sites of OsSATs

The identification of catalytic residues of enzymes is an indispensable step for understanding the functions of enzymes [44]. In this study, active site predictions of OsSATs were performed using InterPro 74.0 server (Table 6). 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 in rice. In a nutshell, it was found that similar amino acid residues were present in the predicted active binding sites.
The expressions of OsSATs under salt stress

In this study, OsSATs’ responses were investigated under 3, 12, 24-h salt treatments under 3, 12, 24-h salt treatments (Fig. 7). OsSAT2;1, OsSAT2;2 and OsSAT3 were downregulated at 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 salt stress exposure times in a similar way. The expressions of OsSAT1;1, OsSAT1;2, and OsSAT1;3 increased depending on the 3, 12, and 24-h NaCl treatments. OsSAT1;2 was generally expressed at the highest level at 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-h salt treatment.

Discussion

Sequence, nucleotide and phylogenic analyses of Arabidopsis and rice proteins showed that there are sequential and phylogenetic divergences among SATs [9]. 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 genes 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 [45]. Consequently, it may be suggested that the genetic diversity of OsSAT genes decreased as a result of purifying selection.
**Fig. 5** Co-expression networks of *OsSAT* genes (blue-circled). The each network was generated using the data on RiceFREND gene co-expression database. (Color figure online)

**Fig. 6** The predicted 3D model of *Arabidopsis* and rice SAT proteins

© Springer
R value, the ratio of transition to transversion, were estimated for DNA sequence evolution and phylogeny reconstruction [46]. 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 [47]. 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

Table 5 The 3D structure overlap (%) of rice, Arabidopsis, and soybean SATs using CLICK structure comparison server

| SAT   | 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: Glycine max (PDB code: 4N69)

Table 6 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, 306M, 311P |
| OsSAT2;1 (LOC_Os03g08660) | 197D, 198H, 217L, 218H, 224G, 225T, 232R, 233H, 244A, 259E, 261G, 262A, 265L, 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 |

Fig. 7 The expression profiles of OsSATs under salt stress. Each gene was shown in different color. The error bars on the graphs show the standard deviations of expressions. (Color figure online)
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 [48] and expansion of gene families [49]. 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 [50].

Although all SATs contain serine acetyltransferase N-terminal domain structure (SATase_N, PF06426) and bacterial transferase hexapeptide (PF00132) as their common motifs, there are 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 [51], suggesting that these variations in motif structures may be connected with functional diversities of SATs in plants. Homologous proteins from different organisms can be recognized using sequence comparison because amino acid substitutions in particular positions are prevented by strong selective constraints [52]. The structure of proteins affect their functions [9] SAT proteins showed structural variations indicating SAT proteins’ functional flexibilities. Therefore, SATs may be involved in more specific processes on the same pathway.

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 [53]. In this study, similar residues were identified in predicted active sites of OsSAT proteins.

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. All SATs but OsSAT1;1 were expressed in leaf blade. OsSAT1;1 and OsSAT2;1 were particularly induced in roots. JA is the only hormone activating expression of OsSAT1;1, OsSAT2;1, and OsSAT2;2. Besides, OsSAT1;1 was the only OsSAT induced by auxin, JA and ABA hormones. The wet-lab 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 exposure times. OsSAT1;2 showed the highest expression levels when exposed to 3, 12, and 24-h salt treatments. Conversely, OsSAT2;2, OsSAT3 and particularly OsSAT2;1 were downregulated depending on exposure times. Watanabe et al. [50] 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 survival. 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, 54]. The varied expression patterns of OsSATs under various conditions show that they are dynamically regulated and OsSAT1;1 is the only gene induced by ABA, JA and auxin hormones. Interestingly, Wang et al. [55] reported that ABA stimulates JA synthesis in rice seedlings through SAPK10 (osmotic stress/ABA-activated protein kinase) bZIP72-AOC pathway. Conversely, JA is also reported to activate ABA, ethylene and salicylic acid. JA is known to increase salt tolerance of plants. Exogenous JA application also increase ABA and proline contents in cells along with antioxidant enzyme activity rendering them more tolerant to salt stress [56]. Similarly, auxin, ethylene, salicylic acid, brassinosteroids, and gibberellin are also triggered by JA [57]. Consequently, the findings showed that SAT enzymes (EC 2.3.1.30) in rice may be regulated by cross-talk of ABA, JA and auxin hormones.

**Conclusion**

Sulfur is one of macronutrient which is necessary for plants’ growth and development. On the other hand, Cys, one of sulfur-containing amino acids, is used in biological systems to synthesize sulfur-containing compounds. SAT gene family play specific roles in plant metabolism, particularly in sulfur assimilation pathway. In this study, 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 purifying selection along with segmental and tandem duplications led OsSATs to have more specific and selective roles in metabolic pathways which may have an effect on the rice plant’s responses to abiotic and biotic stress conditions.

SATs appear to be induced by cross-talk of several hormones. More explicitly, JA induced the expressions of OsSAT1;1, OsSAT2;1, and OsSAT2;2 whereas auxin and ABA induced only OsSAT1;1. OsSAT1;1 was the only gene induced by ABA, JA and auxin in digital expression analysis. As is known, salt stress also activates ABA synthesis and ABA-related gene expressions. The wet-lab data in present study showed that OsSAT1;1, OsSAT1;2 and OsSAT1;3 genes were upregulated when exposed to different
time periods of salt stress. Consequently, \textit{OsSAT1;1} may be a hub gene regulated by cross-talk of the mentioned hormones. Interestingly, \textit{OsSAT1;1} is the only \textit{SAT} gene lowly expressed in leaf blade but highly expressed in root cells. The cross-talk of hormones and the structural variations of OsSAT proteins may also explain the different responses of OsSATS to salt stress. In this regard, the effect of cross-talk of hormones and how to fine-tuning OsSATS expressions under abiotic stress conditions may be considered as research topics for future studies.

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1007/s11033-021-06620-6.

**Author contributions** EF and FK designed the study. EF, FK and AA conducted the experiments and analyzed the data. All authors contributed to writing and editing the manuscript.

**Declarations**

**Conflict of interest** The authors declare that they have no conflict of interest.

**Ethical approval** This article does not contain any studies with human interest.

**References**

1. Leustek T (2002) Sulfate metabolism. Arabidopsis Book 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

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, Raffet 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.jpph.2005.11.013

6. Romero LC, Aroca MA, Laureano-Marin AM et al (2014) Cysteine and cysteine-related signaling pathways in \textit{Arabidopsis thaliana}. Mol Plant 7:264–276. https://doi.org/10.1093/mp/ssu168

7. Kopriva S, Patron N, Keeling P, Leustek T (2008) Phylogenetic analysis of sulfate assimilation and cysteine biosynthesis in phototrophic organisms. In: Hell R, Dahl C, Leustek T (eds) Sulfur metabolism in phototrophic organisms. Springer, Dordrecht, The Netherlands, pp 33–60

8. Hell R, Jost R, Berkowitz O, Wirtz M (2002) Molecular and biochemical analysis of the enzymes of cysteine biosynthesis in the plant \textit{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 \textit{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. Cao MJ, Wang Z, Zhao Q et al (2014) Sulfate availability affects ABA levels and germination response to ABA and salt stress in \textit{Arabidopsis thaliana}. Plant J 77:604–615. https://doi.org/10.1111/tpj.12407

12. Dominguez-Solis JR, He Z, Lima A et al (2008) A cyclophilin links redox and light signals to cysteine biosynthesis and stress responses in chloroplasts. Proc Natl Acad Sci USA 105:16386–16391. https://doi.org/10.1073/pnas.0808240105

13. 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

14. Finn RD, Coggill 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

15. Finn RD, Clements J, Arndt W et al (2015) HMMER web server: 2015 update. Nucleic Acids Res 43:W30–W38. https://doi.org/10.1093/nar/gkv397

16. Hall TATA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser 41:95–98. https://doi.org/10.1098/rpst.1999.0453

17. Hu B, Jin J, Guo A-Y et al (2015) GSDS 2.0: an upgraded gene feature visualization server. Bioinformatics 31:1296–1297. https://doi.org/10.1093/bioinformatics/btu817

18. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A (2005) Protein identification and analysis tools on the ExPASy server. In: Walker JM (ed) The proteomics protocols handbook. Humana Press, pp 571–607

19. 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

20. Bailey TL, Johnson J, Grant CE, Noble WS (2009) MEME Suite. Nucleic Acids Res 37:W108–W113. https://doi.org/10.1093/nar/gkn416

21. 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

22. 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

23. Jones DT, Taylor WR, Thornton JM (1992) The rapid generation of mutation data matrices from protein sequences. Bioinformatics 8:275–282. https://doi.org/10.1093/1093/bioinformatics/83.2.275

24. 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

25. Tajima F, Misawa K, Innan H (1993) The amount and pattern of DNA polymorphism under the neutral mutation hypothesis. Genetica 90:103–107

26. 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

27. Darzentas N (2010) Circoletto: visualizing sequence similarity with Circos. Bioinformatics 26:2620–2621. https://doi.org/10.1093/bioinformatics/btp484
28. 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
29. Sato Y, Takehisa H, Kamatsuki K et al (2013) RiceXPro version 3.0: expanding the informatics resource for rice transcriptome. Nucleic Acids Res 41:D2106–D2113. https://doi.org/10.1093/nar/gks1125
30. Sato Y, Namiki N, Takehisa H et al (2013) RiceFREND: a platform for retrieving coexpressed gene networks in rice. Nucleic Acids Res 41:D1214–D1221. https://doi.org/10.1093/nar/gks1122
31. 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
32. 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
33. 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
34. Nguyen MN, Tan KP, Madhusudhan MS (2011) CLICK—topology-independent comparison of biomolecular 3D structures. Nucleic acids Res 39(Web Server issue):W24–W28. https://doi.org/10.1093/nar/gjr393
35. 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
36. Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL, Vandesompele J, Wittwer CT (2009) The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. Clin Chem 55:611–622. https://doi.org/10.1373/clinchem.2008.112797
37. 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
38. Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2-ΔΔCT method. Methods 25:402–408. https://doi.org/10.1006/meth.2001.1262
39. Wirtz M, Droux M (2005) Synthesis of the sulfur amino acids: cysteine and methionine. Photosynth Res 86:345–362. https://doi.org/10.1007/s11120-005-0041-5
40. Rademacher EH, Offringa R (2012) Evolutionary adaptations of plant AGC kinases: from light signaling to cell polarity regulation. Front Plant Sci. https://doi.org/10.3389/fpls.2012.00250
41. Xu C, Luo F, Hochhaldinger 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
42. 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/gkt732
43. Berggaard L, Mikkelsen MD, Sorensen 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
44. 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/btp008
45. Cvićovj I, Good BH, Desai MM (2018) The effect of strong purifying selection on genetic diversity. Genetics 209:1235–1278. https://doi.org/10.10334/genetics.118.30105
46. Yang Z, Yoder AD (1999) Estimation of the transition/transversion rate bias and species sampling. J Mol Evol 48:274–283. https://doi.org/10.1007/3-540-40064-7
47. 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
48. 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
49. 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
50. Watanabe M, Mochida K, Kato T et al (2008) Comparative genomics and reverse genetics analysis reveal indispensable functions of the serine acetyltransferase family gene in Arabidopsis. Plant Cell 20:2484–2496. https://doi.org/10.1105/tpc.108.06335
51. Ben-Hur A, Brutlag D (2006) Sequence motifs: highly predictive features of protein function. In: Guyon I, Nikravesh M, Gunn S, Zadeh LA (eds) Feature extraction. Studies in fuzziness and soft computing, vol 207. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-35488-8_32
52. 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
53. 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
54. Wirtz M, Droux M (2005) Synthesis of the sulfur amino acids: cysteine and methionine. Photosynth Res 86:345–362. https://doi.org/10.1007/s11120-005-0880-9
55. Wang Y, Hou Y, Qiu J et al (2020) Abscisic acid promotes jasmonic acid biosynthesis via a ‘SAPK10-bZIP72-AOC’ pathway to synergistically inhibit seed germination in rice (Oryza sativa). New Phytol 228:1336–1353. https://doi.org/10.1111/nph.16774
56. Wang J, Song L, Gong X et al (2020) Functions of jasmonic acid and other plant hormone signaling highlight the involvement of jasmonic acid as a core component in plant response to biotic and abiotic stresses. Front Plant Sci 10:1–12. https://doi.org/10.3389/fpls.2019.01349

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