



A systematic *in silico* prediction of gibberellic acid stimulated GASA family members: A novel small peptide contributes to floral architecture and transcriptomic changes induced by external stimuli in rice

Izhar Muhammad^{a,1}, Wen-Qiang Li^{a,1}, Xiu-Qing Jing^a, Meng-Ru Zhou^a, Abdullah Shalmani^a, Muhammad Ali^b, Xiao-Yong Wei^a, Rahat Sharif^b, Wen-Ting Liu^a, Kun-Ming Chen^{a,*}

^a State Key Laboratory of Crop Stress Biology in Arid Areas, College of Life Sciences, Northwest A&F University, Yangling 712100, China

^b College of Horticulture, Northwest A&F University, Yangling 712100, China

ARTICLE INFO

Keywords:

GASA gene
Characterization
Rice
GO enrichment analysis
Expression profile
Panicle development

ABSTRACT

The GASA (GA-stimulated *Arabidopsis*) gene family is highly specific to plants, signifying a crucial role in plant growth and development. Herein, we retrieved 119 GASA genes in 10 different plant species in two major lineages (monocots and eudicots). Further, in the phylogenetic tree we classified these genes into four well-conserved subgroups. All the proteins contain a conserved GASA domain with similar characteristics and a highly specific 12-cysteine residue of the C-terminus position. According to the global microarray data and qRT-PCR based analysis, the OsGASA gene family was dominantly expressed in the seedling and transition phase of floral stages. Despite this, OsGASA genes profoundly contribute to rice grain size and length, whereas the highest abundance of transcript level was noticed in stage-2 (Inf 6, 3.0-cm-long spikelet) and stage-3 (Inf 7, 5.0-cm-long spikelet) under GA treatment during panicle formation. Additionally, the maximum expression level of these genes was recorded in response to GA and ABA in young seedlings. Further, in response to abiotic stresses, OsGASA1/8/10 was up-regulated by salt, OsGASA2/5/7 by drought, OsGASA3/6 by cold, and OsGASA4/9 by heat stress. With the exception of OsGASA4, the higher transcription levels of all the other GASA genes were induced by Cd and Cr metal stresses (8–10 fold changes) at various time points. Finally, the GO ontology analysis of GASAs revealed the biological involvement in the GA-mediated signaling pathway and abiotic stresses. Prominently, most of these proteins are localized in cellular components such as the cell wall and extracellular region, where the molecular functions such as ATP binding and protein binding were observed. These results imply that GASAs are significantly involved in rice panicle developmental stages, responses to external stimuli, and hormones.

1. Introduction

Recently a number of studies have been conducted to understand the functional diversity of GASA (GA-stimulated *Arabidopsis*) gene family. So far, GASA peptides have been identified in diverse monocotyledonous and dicotyledonous plant species (Nahirñak et al., 2012b). The first member of *GAST1* gene was identified in tomato and characterized as GA-deficient (*gib1*) mutant gene (Shi et al., 1992). Subsequently, However, the other GASA homologs were identified in *Arabidopsis thaliana* (Aubert et al., 1998; Roxrud et al., 2007), *Triticum aestivum* (Candido et al., 2011), *Oryza sativa* (Furukawa et al., 2006; Zimmermann et al., 2010), *Solanum tuberosum* (Segura et al., 1999; Berrocal-Lobo et al., 2002; Nahirñak, 2015; Nahirñak et al., 2016),

Solanum lycopersicum (Shi et al., 1992; Taylor and Scheuring, 1994), *Petunia hybrida* (Ben-Nissan et al., 2004), *Malus domestica* (Fan et al., 2017), *Gerbera hybrida* (Kotilainen et al., 1999), *Fragaria xananassa* (Fuente et al., 2006), and *Fagus sylvatica* (Alonso-Ramírez et al., 2009). The GASAs encode low-molecular-weight peptides with a precisely conserved 12-cysteine C-terminal, regulated by hormones and participate in signaling pathways that modulate hormonal interaction and endogenous levels of hormone (Nahirñak et al., 2012a; Nahirñak et al., 2012b; Nahirñak et al., 2016; Fan et al., 2017). On the other hand, mutations or missing key peptide residues may cause the non-functional GASA domain and thus loss of biological functions (Sun et al., 2013). In fact, it is evident that the cysteine-rich regions are quite conserved during evolution and, thereby playing significant roles in biological

* Corresponding author.

E-mail addresses: kunmingchen@nwsuaf.edu.cn, kunmingchen@163.com.cn (K.-M. Chen).

¹ These authors have contributed equally to this work.

activities (Rubinovich et al., 2014).

In cereal crops, the spikelet and grain length are the basic units of yield, although there is plenty of evidence suggesting that plant morphology is mainly affected by internal signals, which are regulated by many genes (Liu et al., 2002; Li et al., 2011). To date, only few have been functionally disclosed in the rice GASA gene family, such as *OsGASA1/2/3* influence by phytohormones, triggering internal signal transduction and thus further enhancing panicle growth and development (Furukawa et al., 2006; Lee et al., 2017; Boonpa et al., 2018). In addition, rice *OsGASR7* (Os06g0266800) promotes the development of seed size/length and weight and plays a role in regulating floral meristem identity (Roxrud et al., 2007; Huang et al., 2012; Ling et al., 2013; Dong et al., 2014). Similarly, in wheat (*Triticum urartu*) crop, the two haplotypes (H1 and H2) of *TuGASR7* were determined with pleiotropic effects on grain weight and yield (Dong et al., 2014). Furthermore, based on sequence composition, characterization, and transcriptomic expression patterns, GASA genes are highly active in young proliferating tissues with temporal specificity and contribute to grain size/weight and correspondingly the panicle differentiation process (Furukawa et al., 2006; Zimmermann et al., 2010; Moyano-Canete et al., 2013; Lee et al., 2017; Ke et al., 2018).

Although GASAs are highly stimulated by GA (Furukawa et al., 2006; Li et al., 2011; Nahirñak et al., 2012a), GASAs also show high expression by other growth regulators, such as brassinosteroid (BR), abscisic acid (ABA), salicylic acid (SA), and auxin, depending on the species (Mutasa-Göttgens and Hedden, 2009; Qu et al., 2016). For example, the apple GASA genes show a response to GA and ABA treatments and are highly induced during the seedling and flower stage. Likewise, *Arabidopsis AtGASA2/3/5* and *AtGASA14* are related to ABA signaling and critically participate in flower induction (Roxrud et al., 2007; Zhang et al., 2009, 2016). Additionally, the expression of potato GASA genes is highly induced during seedling and also involved in floral biology as well as in the response to GA and ABA hormones (Nahirñak et al., 2016). Another report also confirmed that rice *OsGASA1/7* genes influence the BR and GA signaling networks (Zhang et al., 2009; Wang et al., 2009; Lee et al., 2015; Boonpa et al., 2018).

Several studies also report that GASA genes are critically involved in abiotic stress response (Porto and Franco, 2013; Sun et al., 2013; Lee et al., 2015). Research has confirmed that overexpression of *Arabidopsis GASA4* confers resistance to heat stress (Ko et al., 2007). Likewise, the overexpression of *FsGASA4* from beechnut (*Fagus sylvatica*) and transgenic lines of *Arabidopsis* improves plant tolerance to salt, oxidative, and heat stresses (Alonso-Ramírez et al., 2009). Another report states that *GASA14* is regulated by abiotic stresses via modulation of reactive oxygen species (ROS) accumulation (Sun et al., 2013). On the other hand, the expression of *GASA5* gene is affected by abiotic stress, as it was inhibited by heat stress (Zhang and Wang, 2011). More recently, wheat *TaGASA1*, which is the homolog of *OsGASA1*, was found to be involved in heat stress (Zhang et al., 2017). In rice, the highest transcript level of rice *OsGASA1* was induced by salt stress and ABA treatments (Lee et al., 2015). Another rice GASA family member, *OsGASA3*, also showed the highest transcript level under salt stress (Lee et al., 2017). However, *OsGASA* family members in responses to biotic stress have not been well elucidated so far.

Additionally, GASA family proteins counteract the inhibitory effects of ions, due to the putative redox-active cysteine region (Rubinovich et al., 2014). *GASA4* overexpression suppresses reactive oxygen species (ROS) and nitric oxide accumulation in *Arabidopsis* plant (Rubinovich and Weiss, 2010). Similarly, petunia (*Petunia hybrida*) plants overexpressing the *GIP2* gene reduce the levels of H_2O_2 and enhance the antioxidant activity in *Arabidopsis* seedlings (Wigoda et al., 2006; Rubinovich and Weiss, 2010). Further confirmation was achieved by the *gasa14* mutant line, where less H_2O_2 accumulation was observed under metal stress (Yamaguchi, 2008). The subcellular localization of GASA proteins differs among various family members as the two proteins GIP2 and GIP5 are located in the cell wall of petunia (Shi et al.,

1992; Ben-Nissan et al., 2004; Zhang and Wang, 2008), but in the cytoplasm, nucleus, and cell wall of rice *OsGASA1/2* and *Glycine soja GsGASA1* (Furukawa et al., 2006; Wang et al., 2009; Zhang and Wang, 2008; Li et al., 2011). More recently, the Snakin-1 and *GASA14* were found to localize to the plasma membrane (Nahirñak et al., 2012a, b; Sun et al., 2013).

In this report, we systematically examine the GASA gene families and their evolutionary relationship. We are also interested in understanding the particular contribution of the rice GASA gene family in multiple stresses and hormonal regulation as well as in the transcriptomic changes induced during panicle development and floral transition stages. This report will be a comprehensive investigation and will hopefully expand the role of GASA genes in rice and also provide a fundamental background for further research on other crop species.

2. Materials and methods

2.1. Data retrieval, identification, and physiochemical properties of GASA genes

To comprehensively annotate the rice GASA protein, we first identified the 15 *AtGASA* protein sequences from the *Arabidopsis* database (TAIR, <http://www.arabidopsis.org>). The obtained sequences were then used as queries to search against the reference genome of the Rice Annotation Project Database (RAP-DB, <http://rapdb.dna.arc.go.jp/index.html>). Further, the protein sequences were integrated into the corresponding protein database for the subsequent identification of GASA genes and their homologs. The candidate proteins were authenticated to confirm their completeness and presence of the core domain via the following online tools: SMART (<http://smart.embl-heidelberg.de/>) (Schultz et al., 1998) and Inter Pro Scan program (<http://www.ebi.ac.uk/interpro/>), and the family-specific GASA domain with accession no (IPR003854) were manually checked. The physiochemical properties of the GASA proteins were analyzed with the online ExPASy proteomics server (<http://web.expasy.org/protparam/>) (Artimo et al., 2012). The WOLF PSORT II program (<http://www.genscript.com/wolf-psort.html>) (Horton et al., 2007) and the TargetP online tool (<http://www.cbs.dtu.dk/services/TargetP/>) (Guo et al., 2016) were used for the prediction of subcellular location. The N-glycosylation sites (Asn) of the GASA protein were searched on NetNGlyc (1.0 server) (<http://www.cbs.dtu.dk/services/NetNGlyc/>) (Additional File Table S1).

2.2. Chromosomal location and sequence alignments

The information about the chromosomal location of *OsGASA* was obtained from Rice Genome (<http://rapdb.dna.arc.go.jp/index.html>), and the physical location of each gene was mapped using MapDraw on a specific chromosome. Nomenclature of the putative GASA genes was assigned based on their chromosomal order. The protein sequences were aligned using DNAMAN program (Version 5.2.2, Lynnon Bio-soft, Canada), and Logos were generated by WebLogo platform (<http://weblogo.berkeley.edu/logo.cgi>). The paralogues were screened on Plant Genome Duplication Database (PGDD) (<http://chibba.agtec.uga.edu/duplication/>) according to (Lee et al., 2012). The presence of transmembrane helices was predicted by TMHMM server v2.0 (<http://www.cbs.dtu.dk/services/TMHMM/>), while the tertiary protein structures were analyzed by PHYRE server v2.0 (<http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index>).

2.3. Phylogenetic relationships, gene structures, and motif analysis

The protein sequences of GASA gene families were initially multiply aligned by ClustalW v2.0 online package (http://www.ebi.ac.uk/Tools/webservices/services/msa/clustalw2_soap), and the phylogenetic tree was constructed using iTOL (<https://itol.embl.de/>) (Letunic and Bork, 2016) via the neighbor-joining (NJ) method with 1000 bootstrap

replicates. The coding sequences were aligned to the corresponding genomic sequences, and the exon/intron structure of each *OsGASA* gene was illustrated by Gene Structure Display Server (GSDS) (<http://gsds.cbi.pku.edu.cn>). The conserved motifs of *GASA* protein sequences were analyzed using online MEME (Multiple Expectation Maximization for Motif Elucidation) platform (version 4.12.0) (<http://meme.sdsc.edu/meme/cgi-bin/meme.cgi>) (Bailey et al., 2006), with default parameter. The maximum number of motifs were set to 10, and the optimum motif width set to > 6 and < 200.

2.4. Analysis of promoter region and GO ontology

The potential *cis*-regulatory elements were scanned on PlantCARE program (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>) searching the 1.5 kb upstream sequence from the transcription start site (start codon ATG) (Lescot et al., 2002) and Neural Network Promoter Prediction (<http://promotor.biosino.org/>) for each candidate *GASA* gene. The GO ontology analysis of *GASA* protein sequences was obtained from Blast2GO program Ver. 2.7.2 (<http://www.blast2go.com>) (Conesa and Gotz, 2008), and three groups of GO classification (molecular functions, biological process, and cellular component) were recovered.

2.5. Developmental expression profile of rice *GASA* genes

The tissue-specific expression profile analysis during different developmental stages of rice *GASA* genes was conducted on the data set as reported by Plant Expression Database (PLEXdb, <http://www.plexdb.org/>) (Dash et al., 2012) and the global microarray data available in the GENEVESTIGATOR V3 database (<https://genevestigator.com/gv/>) (Hruz et al., 2008). The data were based on the Reads Per Kilo-base Per Million mapped reads (RPKM) analysis indicated as the transcriptomic level of *GASA* members, and the results were constructed in heatmap. The rice plants were allowed to grow in experimental field, and the expression profile of different plant organs at various developmental stages including seedling (root, stem, and leaf), booting (node, internode, leaf blade, leaf sheath, and flag leaf), and heading stages (post and pre-inflorescences) were analyzed.

2.5.1. Transcriptomic comparison of *GASA* genes during panicle differentiation

The rice plants were grown in the experimental area of Northwest A & F University Shaanxi, China. When the floral biology was fully developed after the differentiation period of primary and secondary branches, we exogenously sprayed the GA (100 μ M) hormone on rice panicles. The inflorescence samples were collected from three different stages, i.e. In5 (formation of higher-order branches, 2.0-cm-long panicle), In6 (differentiation of glumes, 3.0-cm-long spikelet), and In7 (differentiation of floral organs, 5.0-cm-long spikelet, respectively) immediately frozen in liquid nitrogen and stored at -80°C until the total RNA extraction. The differentiation periods of panicle organs used in this study are relatively closer between three stages with difference one to three days (Ikeda et al., 2004). The biological replicate for every sample of each developmental stage was a mixture of at least 10 individual plants. To minimize the experimental error the growth and developmental status of sampled panicles in a specific development period were highly consistent.

2.6. Biological materials, external treatments, and cDNA library

The rice (*Oryza sativa* L cv. Nipponbare) seeds were obtained from the State Key Laboratory of Crop Stress Biology for Arid Areas (Northwest A&F University, Yangling, China). The detailed methods of growth condition, stress treatments, and concentration for various abiotic, metal stresses, and hormonal application were the same as described in our previous study (Muhammad et al., 2018), but with the

addition of GA (100 μ M) hormone, which was exogenously sprayed on two-week-old rice seedlings, and samples were collected at the same time point as mentioned in our above previous article. All the chemicals used for treatments were purchased from Sigma-Aldrich (Burlington, MA, USA). All the collected samples following the treatments were immediately frozen at -80°C until analysis. Total RNA was extracted from 0.5 g of whole leaf blade samples using RNAiso TM Plus (Takara, Dalian, China). Preceding cDNA synthesis, RNA was treated with RQ1 RNase-free DNase (Promega, Madison, WI, USA) according to the manufacturer's instructions to rule out DNA contamination. The cDNA was synthesized by Super Mix (Trans Gen Biotech, Beijing, China) and for accurate quantification of qPCR cDNA was further diluted upto 50 ng/ μ L for each sample.

The transcript level of rice *GASA* genes was analyzed using SYBR Premix Ex Taq II (TaKaRa, Dalian, China) for qRT-PCR (ABI PRISM 7500 (Applied Biosystems) with SYBR Green master mix (Qiagen) system. All the primers were gene-specific, and BLAST was used to check the specificity of the respective gene. Melting curve analysis of real-time PCR reactions provided further confirmation. The details of primer sequences are listed in Additional File Table S5. To confirm the precision of results, three biological replicates were used and three technical replicates were run for each biological sample. The protocol was referenced to our previous study (Muhammad et al., 2018). The quantitative variation among the replicates was evaluated by relative quantification method $2^{-\Delta\Delta\text{CT}}$ method.

2.7. Statistical analysis

The statistical analysis was performed using SPSS software (version 25.0, SPSS Inc., USA), and 95% confidence intervals ($p \leq 0.05$) were considered statistically significant. The graphs were prepared by GraphPad Prism 6 (GraphPad Software, Inc., La Jolla, CA, USA).

3. Results

3.1. Identification and annotation of rice *GASA* genes in the plant kingdom

A Hidden Markov Model (HMM) search was carried out to comprehensively investigate rice *GASA* genes and their homologs in other species, where the BLASTp search corresponding to *Arabidopsis GASA* genes was used as a query and a total of 119 *GASA* genes were identified in 10 different species (Fig. 1A and B, Table 1). Previously, nine rice *GASA* genes were reported (Zimmermann et al., 2010). Here, we identified one more *GASA* gene (*OsGASA10*; Os10t0115550) in rice. The newly identified gene was confirmed in the rice database, and the commonly shared conserved 12-cysteine domain was also verified in Conserved Domain Database (CDD). In addition, as a simplified nomenclature for the identified gene, we adopted the acronyms *OsGASA1* to *OsGASA10* based on their order of appearance from chromosome 1–10. The *GASA* genes were extensively distributed on *Oryza sativa*, *Zea mays*, *Brachypodium distachyon*, *Sorghum bicolor*, *Solanum tuberosum*, and *Setaria italic* in monocots, and *Arabidopsis thaliana*, *Cucumis sativus*, *Solanum lycopersicon*, and *Capsicum annuum* in eudicots (Fig. 1A and B). The genes with alternative transcript variants were manually checked, and the longest transcript was chosen for further analysis. The SMART and Pfam results showed that the *GASA* core domain, the gibberellic-acid regulated domain, was found in all *OsGASA* genes (Fig. 2E). The physicochemical characteristics of *GASAs* were evaluated on EXPASY PROTOPARAM online tool (Additional File Table S1). The deduced rice *GASA* proteins have low-molecular-weight peptides ranging from 92 (*OsGASA5*) to 152 (*OsGASA4*) amino acid residues, with CDS region ranging from 279 to 456 (Additional File Table S2). The isoelectric points are close to 9, indicating the alkaline nature of the protein, and the instability index (index value ≤ 40) showed that most of the *OsGASA* proteins are unstable. The predicted subcellular localization of *OsGASA* proteins revealed that they are mostly localized in the cell

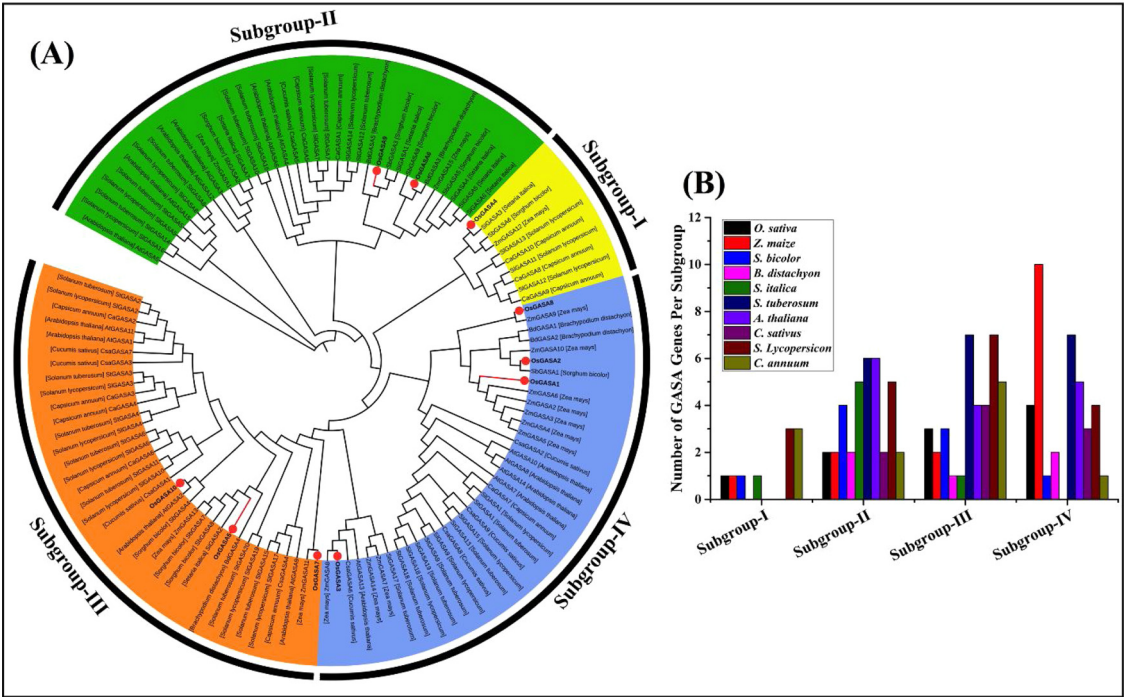


Fig. 1. Systematic evolutionary relationships of GASA gene families in 10 plant species. (A) The phylogenetic relationship among the rice GASA family members. The four well-conserved subgroups are marked by different colors and represented as Subgroup-I, Subgroup-II, Subgroup-III, and Subgroup-IV, respectively. The OsGASA genes are marked with a red line and red bold dot in the phylogenetic tree. (B) The distribution of GASAs in the subgroup I-IV from ten plant species. The separate color bars indicate number of genes per species. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

wall, extracellular membrane, chloroplast, vacuole, and/or mitochondria. The grand average of hydropathicity (GRAVY) values indicates that all the GASA proteins are of hydrophilic nature. The aliphatic indexes of the GASA proteins ranged from 49.30 (OsGASA10) to 68.82 (OsGASA8), and are regarded as a positive factor for the increase in thermostability of globular proteins. In terms of amino acid residuals, Cys and Lys were the highly enriched residues, while Leu, Ser, Asp, Val, Pro, and Thr were detected in low number. The SMART results depicted that transmembrane regions existed in OsGASA1/6/7/10 (Additional File Figs. S1, A), and the tertiary structure of OsGASA genes suggest the presence of a hypothetical GASA protein, with a 98.3% probability of homology (Additional File Figs. S1, B).

3.2. Phylogeny, gene structure, and motif analysis of GASA gene family in rice

To clarify the evolutionary relationship among GASA genes, the 119 amino acid sequences obtained from different species were multiply aligned. Subsequently, a neighbor-joining (NJ) phylogenetic tree was constructed using iTOL. All the GASA genes were further divided into

four well-conserved subgroups (Subgroup-I, II, III, and IV) according to the topological association/consistency or similarities in functional behavior among the subgroups (Fig. 1A and B). Based on our analysis, the OsGASA4 gene was clustered in Subgroup-I, OsGASA6/9 in Subgroup-II, OsGASA5/710 in Subgroup-III, and OsGASA1/2/3/8 in Subgroup-IV (Fig. 2A). Interestingly, all the GASA members within subgroups are highly consistent in the evolutionary history, and the OsGASA gene clustering formation within certain subgroups exhibits an identical percentage of protein sequences, the presence of an integral core domain, the same number of exons/introns, and the similarities/differences in functional divergence. The structural diversity of the GASA genes was examined based on exon/intron distribution. The gene structural organization was attained from the alignment of coding region sequences with respective genomic full-length sequences using online GSDS (<http://gsds.cbi.pku.edu.cn/>) tool. The OsGASA gene structural classification (exon/intron) can be divided into four subgroups, I–IV (Fig. 2C). The rice GASA genes share conserved structures within subgroups, while the length, arrangement, and position of introns are relatively less conserved. For example, the subgroup-I and II have two introns and three exons. Only OsGASA7, cluster member of

Table 1
Details of plant genomes analyzed in this study.

Lineages	Plant Species	Genome (Mbp)	Available Database
Monocots	<i>Oryza sativa</i>	430	http://rice.plantbiology.msu.edu/index.shtml
	<i>Zea mays</i>	2,400	http://www.maizesequence.org/index.html
	<i>Sorghum bicolor</i>	699	http://www.plantgdb.org/SbGDB
	<i>Solanum tuberosum</i>	705	http://www.plantgdb.org/StGDB
	<i>Brachypodium distachyon</i>	355	http://www.phytozome.net/brachy.php
	<i>Setaria italica</i>	490	http://plants.ensembl.org/Setaria_italica/Info/Index/
Eudicots	<i>Arabidopsis thaliana</i>	157	http://www.arabidopsis.org/
	<i>Cucumis sativus</i>	367	http://cucurbitgenomics.org/blast
	<i>Solanum Lycopersicon</i>	900	https://solgenomics.net/organism/solanum_lycopersicon/genome
	<i>Capsicum annuum</i>	2700	http://peppersequence.genomics.cn/page/species/blast.jsp

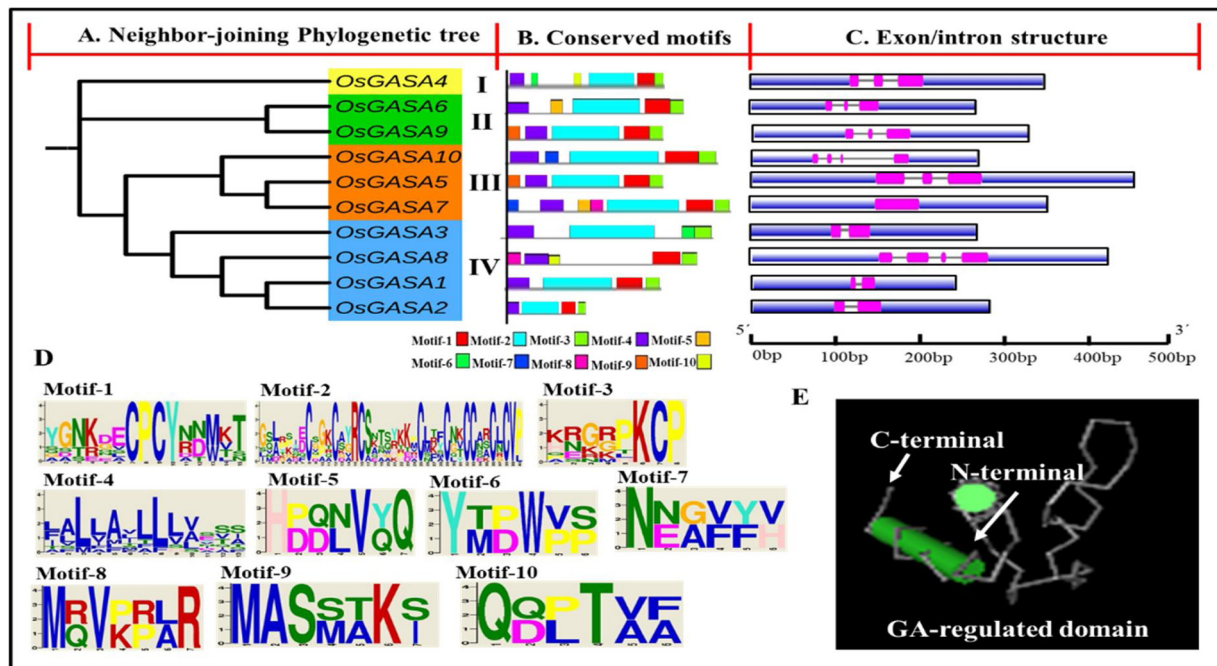


Fig. 2. Phylogenetic relationship, conserved motif analysis, exon/intron distribution, and domain organization of GASA family genes in rice. (A) Phylogenetic analysis and clustering of rice genes into different subgroups. (B) Conserved motif analysis of rice GASA family genes. (C) Exon/intron distribution of GASA genes. The blue boxes represent the UTR region, pink boxes represent exons, and introns are represented by black lines between two exons. (D) The logos represent the 10 conserved motifs. (E) 3D picture of conserved GASA domain. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

subgroup-III, has no intron and one exon. Notably, the *OsGASA* genes have fewer exons/introns, or the coding sequences are disrupted by one to three introns that may allow the distribution of *OsGASA* genes among the subgroups (Fig. 2C). Furthermore, we detected that the members of *OsGASA* proteins commonly share motif 2/3/4, except *OsGASA8*, a member of Subgroup-IV that does not carry motif-2. However, motif-5 is present only in *OsGASA6/7*, motif-6 in *OsGASA2/4*, motif-7 in *OsGASA7/10*, motif-9 in *OsGASA5/9* and motif-10 only in *OsGASA4/8*, respectively. Thus, motif-2/3/4 is frequently found in GASA proteins, while the other *OsGASA* proteins carry a single motif. Therefore, motif-1/2/3/4 shows the highest homology in GASA proteins (Fig. 2B, D).

3.3. Sequence analysis and chromosomal localization of rice GASA genes

The conserved domains of the GASA proteins were confirmed by Pfam, SMART, Inter Pro Scan, and Conserved Domain Database (CDD) databases. The putative *OsGASA* protein possesses three distinct domain/regions; a highly conserved C-terminal GASA domain containing 12 conserved cysteines, a highly divergent intermediate downstream region which is diverse in different GASA families, and an N-terminal signal peptide, an individual stretch of 18 to 34 amino acids according to the SignalP.v4.0 software, which is a cleavable signal sequence. The amino acid sequence comparison between these proteins reveals an overall homology, and the C-termini are approximately 74% identical, while the N-terminal region of these proteins is not well conserved. However, the multiple sequence alignment illustrates that the perfect conservation of this domain is family specific (Fig. 3B, Additional File Fig. S3). Further, this finding confirms that the rice GASA gene family consists of at least 10 members (Fig. 3D). These genes were numbered according to the nomenclature established in the rice genome database (<http://rice.plantbiology.msu.edu/index.shtml>). The chromosome localization of rice GASA genes indicated that all of them correspond to different loci in the genome, where they are randomly mapped on 10 chromosomes (Fig. 3D). Chromosome 4, 7, 9, and 10 carry only one copy of gene *OsGASA3*, *OsGASA8*, *OsGASA9* and *OsGASA10*, respectively,

while chromosome 3, 5, and 6 contain one pair of rice genes such as *OsGASA1/2*, *OsGASA4/5*, and *OsGASA6/7*, respectively. According to gene duplication analysis, no tandem and segmental duplication were observed in our study.

3.4. Promoter region analysis of rice GASA genes

In silico analyses revealed that the rice GASA genes upstream region carries multiple potential *cis*-regulatory elements. A total of 22 *cis*-regulatory elements were predicted, of which 11 corresponded to environmental stresses and 11 were hormone related. Moreover, all the stresses or hormone-responsive elements are unevenly distributed on the 10 *OsGASA* promoter region (Additional File Fig. S2, Table S3). The highest number of Gibberellin responsive elements P-Box and TATC-box were identified in all *OsGASA* genes, followed by ABA-responsive element (ABRE), which is only missing *OsGASA7/8*. Meanwhile, TCA (responsible for salicylic acid) and MeJA-responsive elements (CGTCA-motif) were found in the promoter region of five *OsGASA* genes each. Additionally, the MYB binding site involved in drought-inducibility (MBS), resistance, light responsiveness (G-BOX and Sp1), and endosperm-expression responsiveness (Skn-1-motif and GCN4-motif) was also identified in the promoter region of various GASA genes (Additional File Fig. S2). The *cis*-regulatory elements are clustered in certain promoter regions, implying the vital contribution of these elements in the regulation of the genes in various biological processes.

3.5. Spatio-temporal regulation of rice GASA genes

We systematically approached to elucidate the expression pattern of GASA genes. First, a set of microarray data for the specific GASA gene expression in various organs/tissues were attained from GENEVESTIGATOR.v3 (Hruz et al., 2008) (Fig. 4A), and then we experimentally ran quantitative real-time polymerase chain reaction (qRT-PCR) (Fig. 4B). Transcriptomic profiles were generated from 14 different tissues in three (seedlings, booting, and flowering) developmental

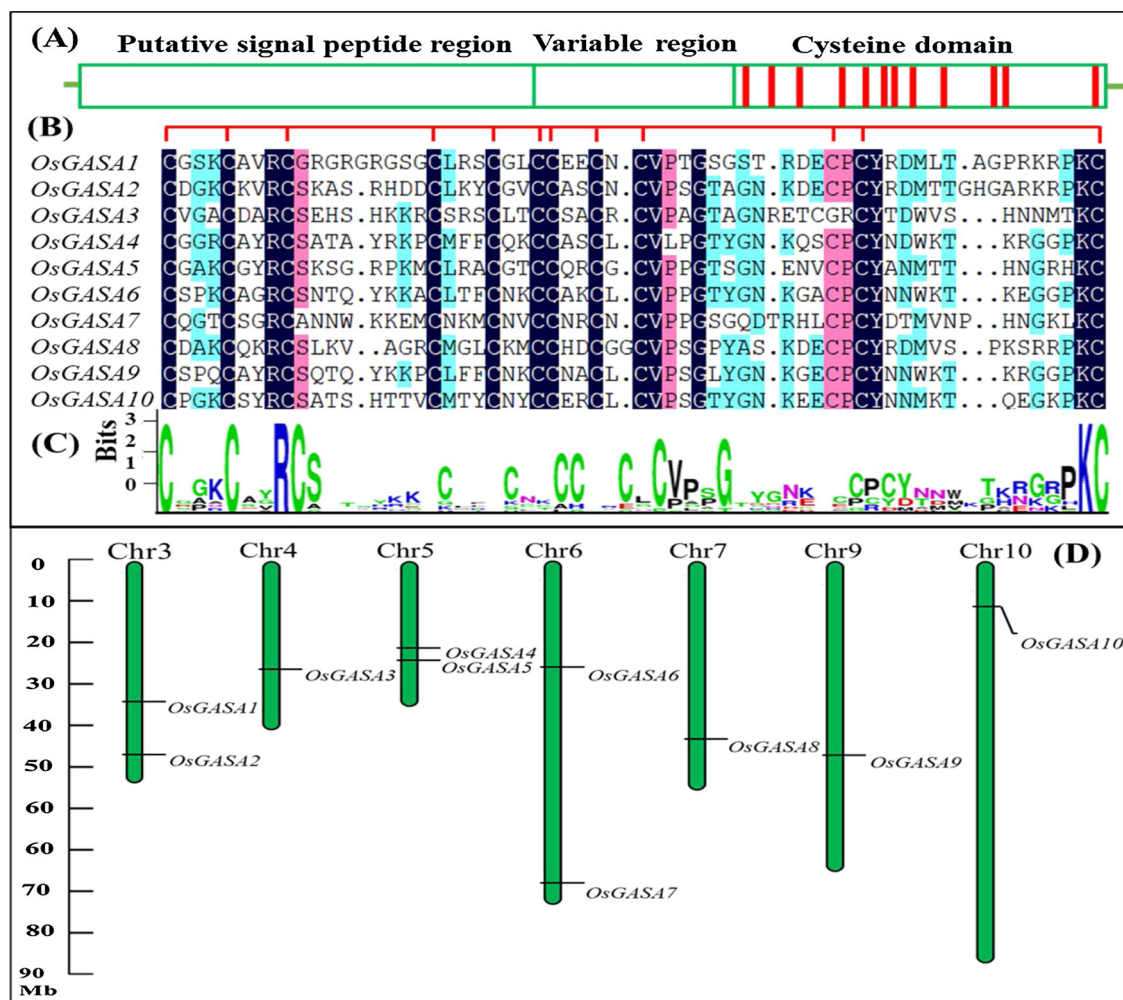


Fig. 3. Amino acid sequence alignment of the rice *OsGASA* proteins. (A) represents analysis of *GASA* protein structures; the red columns denote their conserved 12-cysteine-rich regions. (B) indicates the multiple alignments of the *OsGASA* protein sequences and their conserved *GASA* domains. (C) denotes the sequence logo analysis of the conserved *GASA* domains. (D) indicates the chromosomal localization of *GASA* genes. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

stages. The expression obtained from qRT-PCR and microarray data have shown a correlation in multiple tissues, while the dissimilar expression level in some tissues may be due to the difference in tissue age/type. The qRT-PCR analysis revealed the highest expression level during inflorescence emergence stages, followed by seedling stage. More expansively, *OsGASA3/5* is significantly expressed during anthesis, pistil, and booting stage, and *OsGASA1/2/4/7* are highly expressed in young seedlings, apical meristem, and inflorescence emergence (milking stage). The expression of *OsGASA6/9* is high in root, flag leaf, and post and pre-inflorescence emergence, respectively, while *OsGASA8/10* are dominantly expressed in flowering stage, particularly in young panicle before milking stage (Fig. 4B). The expression of *OsGASA* genes in other organs was much lower, at an almost undetectable level, especially in node, internode, leaf sheath, and leaf blade. Therefore, the overall results signify the unique strong spatio-temporal and tissue specificity in rice plant. The strong spatio-temporal and tissue-specific expressions suggest an obvious biological role, as well as long-term evolutionary relationships, among the *OsGASA* family members in rice.

3.6. *OsGASA* expression pattern in response to gibberellic acid during rice panicle differentiation period and contribution to rice grain size

As the *GASA* genes possess the gibberellic acid (GA) *cis-regulatory* element in the promoter region, we extended our study to elucidate the

new insights of *OsGASA* genes and their transcriptional variation induced by GA during panicle developmental stages (Fig. 5, Additional File Table S3). Initially, we obtained the expression of *OsGASA* genes from GENEVESTIGATOR.v3 (Hruz et al., 2008) presented in the form of a heatmap, from green-black-red, reflecting the percentage expression (Fig. 5A). Subsequently, the transcriptomic analysis (qRT-PCR) of rice panicle at three different stages (S1: stage-1, inf-5 (formation of higher-order branches); S2: stage-2, inf-6 (differentiation of glumes); and S3: stage-3, inf-7 (differentiation of floral organs)) under GA regulation were compared. As can be seen in Fig. 5B, *OsGASA* gene regulation by GA is significant between stages with $\log_2 FC > 2.0$ when compared to control plants. Moreover, *OsGASA3/5/4* are up-regulated (13-, 7-, and 4.5-fold, respectively) by GA during stage S1 and S2 (stage-1 and 2), but down-regulated in S3 (stage-3). In contrast, *OsGASA10* is up-regulated (10-fold) only in S3, whereas *OsGASA6/8* is highly expressed during S1 and *OsGASA1/2* is induced by GA (12- and 18-fold, respectively) during S2 and S3. Moreover, *OsGASA7/9* is up-regulated (7.5-fold) in all stages by exogenous application of GA (Fig. 5B). Overall the trend of their expression varied at three different stages, but most of the genes are expressed during stage-2 and stage-3. Therefore, according to transcriptomic analysis, it would be reasonable that these genes are assuming a dynamic role in the regulation of rice panicle formation. Further, we also investigated the rice panicle morphology, which is the key factor for successful crop productivity. When rice panicle was

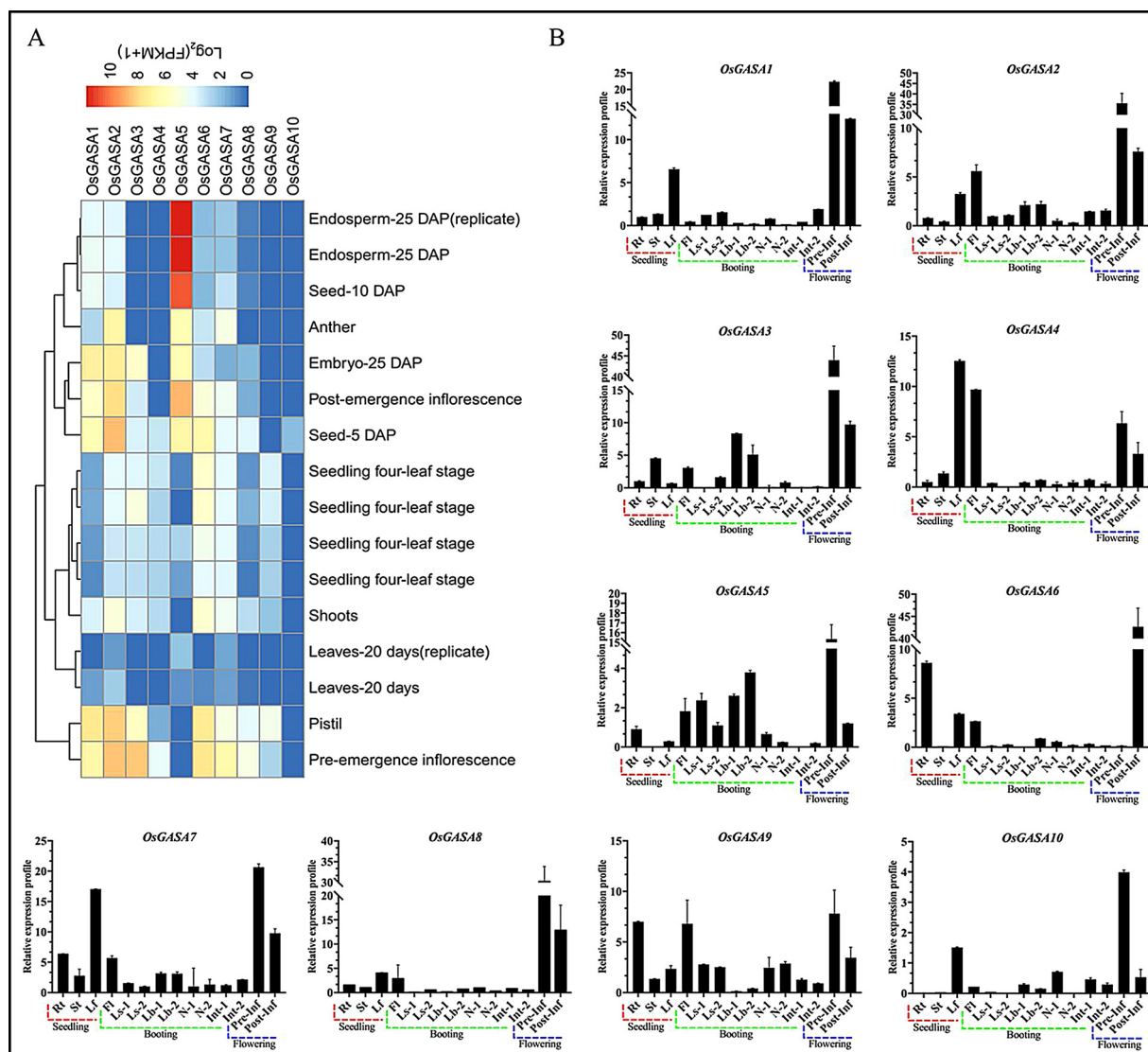


Fig. 4. The expression patterns of GASA family genes in rice at various developmental stages in different tissues. (A) The developmental expression patterns obtained from GENEVESTIGATOR v3, in which 16 different developmental stages (seedling, booting, seed, post-emergence inflorescences, pre-emergence inflorescences, anthesis, pistil) during rice growth were investigated. Results are shown as heat maps in blue accent/red coding that reflect relative signal values; deep blue accent boxes represent either lower expression or no expression, and dark to darker red boxes represent the highest expression level. (B) Expression profiles of ten rice GASA genes in 14 different tissues analyzed by qRT-PCR. The analyzed tissues on the Y-axis are as follows: seedling stage; (Rt-root, St-stem, and Lf-leaf), Booting stage; (Fl-flag leaf, Ls-1,-leaf sheath-1, Ls-II,-leaf sheath-II, Lb-I,-leaf blade-I, Lb-II,- Leaf blade-II, N-I,-Node-I, N-II,-Node-II, Int-I,-Internode-I, Int-II,- Internode-II, Flowering stage; Pre-Inf,- Pre-inflorescences; and Post-Inf, Post-inflorescence). The data were obtained from three independent experiments. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

treated with GA, significant changes were noticed in grain length (Fig. 6). Moreover, the grains showed high quality and appeared healthier than non-treated panicles; in particular, the grain length was significantly elevated as indicated in Fig. 6B and D. No significant differences were observed in grain number per spikelet. In summary, these results clearly indicate that *OsGASA* genes are responsive to GA and play a vital role in signal transduction and interplay/crosstalk mechanisms by the complex hormonal regulation network during panicle formation and developmental stages in rice.

3.7. Inducible expression profiles of *OsGASAs* under abiotic stresses and hormonal treatments

To elucidate the differential expression pattern of rice *GASA* genes and their possible involvement under abiotic stresses and hormonal regulation, we first searched the *cis-acting* elements in the upstream region of each *GASA* gene using the PlantCARE program (Additional

File, Table. S3). The findings revealed several responsive elements for biotic/abiotic stresses, like drought (MBS, MYB binding sites), anaerobic conditions (ARE and GC-motif), endosperm expression (Skn-1-motif), and some hormonal regulation responsive elements such as abscisic acid (ABA), methyl jasmonic acid (MeJA), salicylic acid (SA), and gibberellic acid (GA) (Additional File Table. S3). Subsequently, to explore the expression analysis of *GASA* genes subjected to abiotic stresses, we again examined the rice microarray data GENEVESTIGATOR database (Additional File Fig. S4). We also cross-checked the ePF expression browser for further confirmation. As shown in Table 2, the expression profiles of *GASA* genes show high fluctuations by an external factor with stress specificity. According to the qRT-PCR analysis, significant transcriptomic changes were observed under salt stress (200 mM, NaCl), As the *OsGASA1/8/10* have shown up-regulation with 16-, 19-, 14-fold changes, respectively at 12 h and 24 h time points. While *OsGASA2/5/7* showed highest transcriptional regulation at 6 h under drought stress (20% PEG6000 treatment) with 16-, 6-, and 13.5-

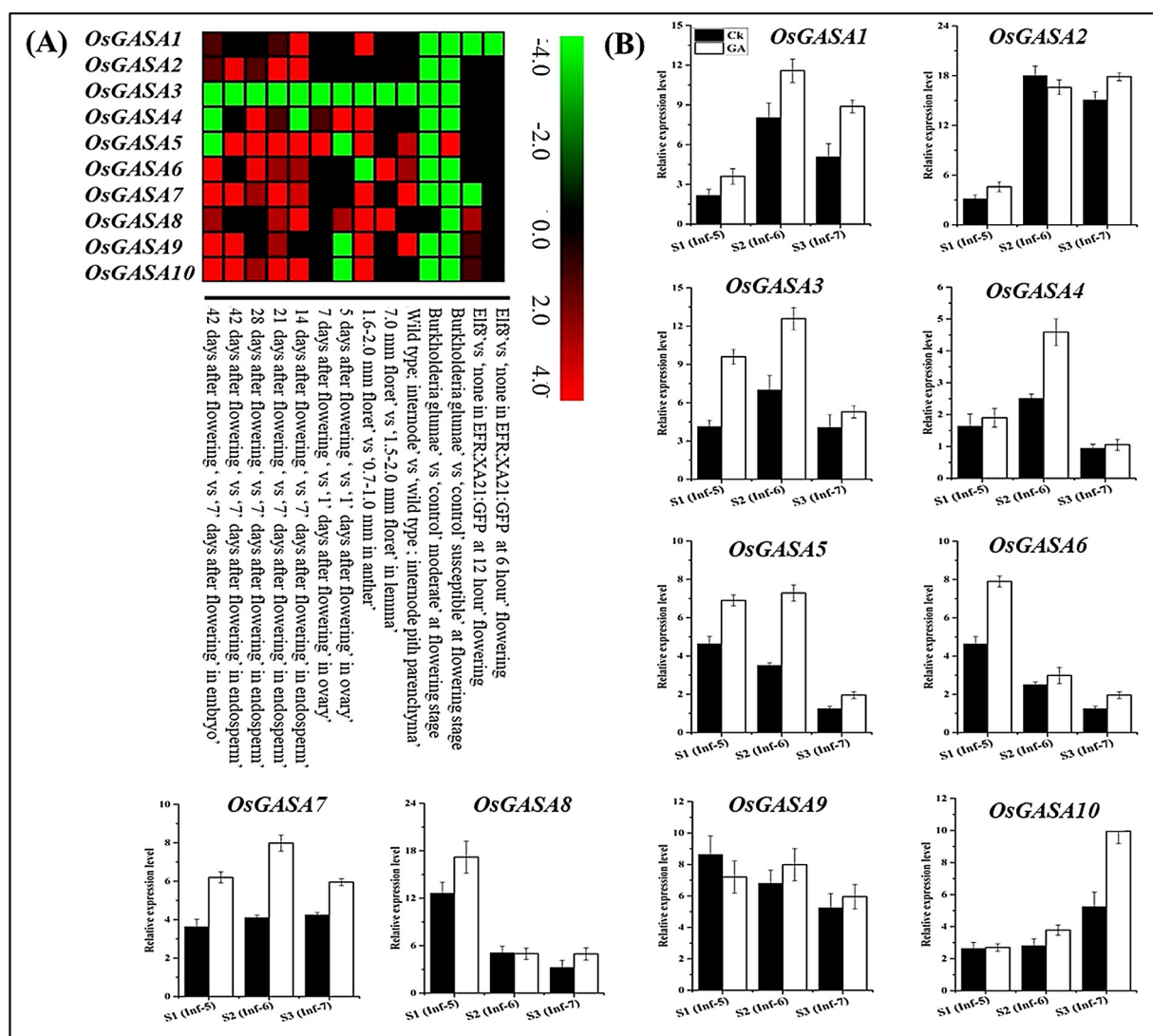


Fig. 5. The expression profiles of *GASA* genes family in rice at different panicle developmental stages. (A) The relative expression data of 10 *GASA* genes obtained in RNA-Seq analysis from GENEVESTIGATOR and ePF rice expression browser. (B) The qRT-PCR validation of rice *GASA* genes in three different developmental stages. Y-axes represent the relative expression levels at different stages. X-axis represents the three stages as S1: stage-1, inf-5 (formation of higher-order branches); S2: stage-2, inf-6 (differentiation of glumes); and S3: stage-3, inf-7 (differentiation of floral organs). Expression was measured in response to GA application. Error bars show standard deviations for three independent biological assays.

fold changes, respectively. Additionally, the *OsGASA3/6* was substantially induced by cold (4 °C) stress at 6 h and 12 h, respectively, and 11- and 5.5-fold changes were observed in the expression levels (Fig. 7A). The other time-point per stress treatments have not shown any obvious changes, as we considered the determinable/up-regulation of gene $\log_2$ FC > 2.0 in transcript level and gene/genes which does not meet the above criteria was considered no transcriptional changes or down-regulation.

We further analyzed the qRT-PCR based expression induced by ABA, GA, MeJA, and SA in various treatments (Fig. 7B). The results exposed the distinct response under these hormonal applications, whereas all the *GASA* genes showed a pronounced up-regulation under GA hormone, as the *OsGASA* genes are GA-regulated and the upstream transcriptional site also showed several copies of *cis*-regulatory elements (Additional File Fig. S3, Table. S3). All the *OsGASA* genes are highly induced by GA and ABA and significant fold changes were noticed when compared other hormonal applications. Notably, the *OsGASA* genes under MeJA treatments did not show any obvious changes; only slight transcriptional variation was noticed at several time-points (Fig. 7B). Though, *OsGASA10* expression was high under SA treatment, and a

gradual increase was observed, reaching the maximum expression (> 16.5-fold changes) at 24 h, whereas the other hormones did not show obvious transcriptional regulation (Fig. 7B). It should be noted that *GASA* genes exhibit an even more distinct differential response to external stimuli with obvious expression levels per stress/application. The accurate fold changes of some genes obtained by microarray or qRT-PCR were varied slightly; the variation tendencies of all the studied genes were identical, although the results presented here carry a minimal variation in the expression data and the highest consistency between the results obtained from using these two techniques. Finally, the key findings display a distinctive role of *OsGASA* genes in response to various environmental stresses as well as hormonal application in rice plant.

3.8. *OsGASA* expression against heavy-metal stress

Previously, the *GASA* genes were shown to be modulated only by hormonal regulation at the seedling stage (Zhang and Wang, 2008). Thus, here we extended our investigation of the *OsGASA* genes expression in response to heavy-metal stresses. The qRT-PCR analysis

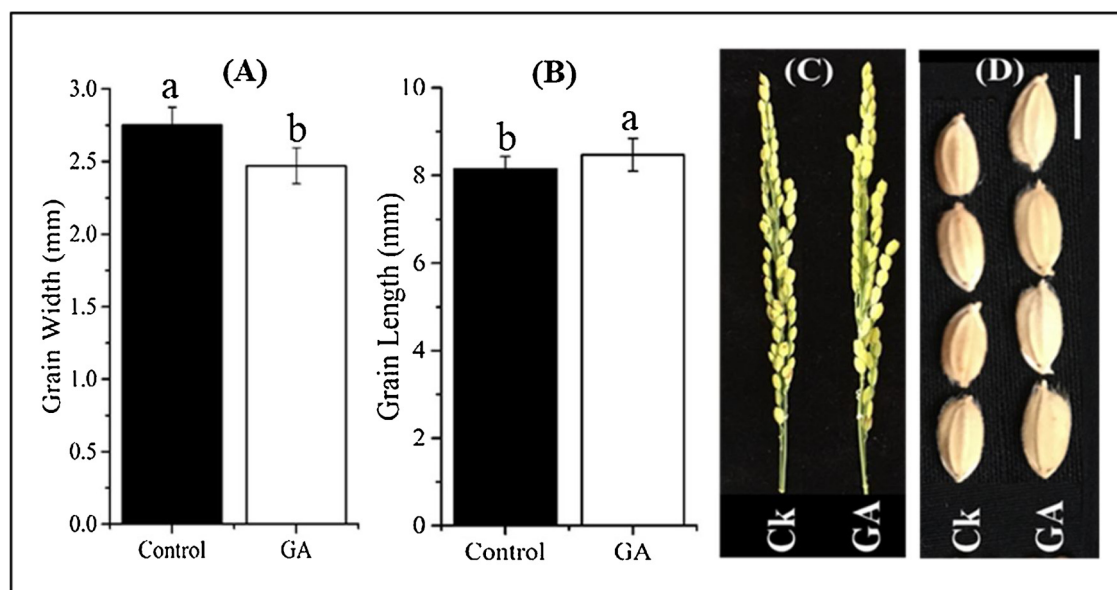


Fig. 6. The phenotypic differences of the grains compared to control and gibberellic acid (GA) treated rice panicles. (A) represents the grain width of the control and GA (100 μ M) treated panicles; (B) represents the grain length of the control and GA (100 μ M) treated panicles; (C) represents the panicle formation and grain morphology; (D) rice grain size/length and width of Ck-non-treated and treated with gibberellic acid (100 μ M), Bar = 5 mm. Ck indicates the non-treated panicle, and GA indicates the panicles treated with gibberellic acid (100 μ M) after differentiation of floral organs. The panicle samples ($n = 50$) were randomly selected at maturity. The data represent the mean $\pm$ SE of $n = 3$ independent determinations with two biological replicates. Different letters indicate the statistical significance ($P \leq 0.05$) analyzed by one-way ANOVA test.

revealed the transcriptional fluctuation with four (Fe, Cd, Ni, and Cr) different heavy metal stresses at various (0,3,6,12, and 24) time-points (Fig. 8). All the *OsGASA* genes were up-regulated by Cd stress, and the expression was 5 to 10-fold higher than when exposed to other metal stresses. With the exception of *OsGASA4/8*, they were up-regulated by Cr and Ni respectively, whereas the later exposure to these stresses employed irregular or abrupt changes in the expression of these genes (Fig. 8). The expression of *OsGASA6/8/9* was completely inhibited by Fe stress, while the other *GASA* genes were up-regulated by Fe metal stress (average 5-fold higher) at the 12 h time point. The *OsGASA10* gene was down-regulated by Cr stress at all analyzed time points and only up-regulated by Ni and Cd stresses (Fig. 8). Likewise, *OsGASA3* was strongly regulated (> 9 -fold) by Fe and Ni stresses. A dramatic expression of *OsGASA9* (> 3.5 -fold) was observed at 3 h (Ni, Cd, and Cr) and at 12 h under Cd, Fe, and Cr, which reflect the interaction behavior of other chemicals present in Hoagland's solution. More interestingly, all the metal stresses changed the transcription level at

various time points, thereby showing substantial expression levels, implying that these heavy metals are responsive to *OsGASA* genes and thus modulate their transcription regulation network in rice plants.

3.9. Gene ontology (GO) analysis

Next, we conducted the gene ontology (GO) analysis of the rice *GASA* (*OsGASA*) and *Arabidopsis GASA* (*AtGASA*) genes to further classify the functional divergence/relationship of the gene family and contribution in the process of plant growth as well as responses to external stimuli (Fig. 9A and B). In order to explore the hierarchical clustering relevance to biological functions, the *GASA* genes were projected for gene ontology to the slim analysis using Blast2GO tool. Further, we categorized the *in silico* functional analysis of *GASA* genes in the GO database. When comparing the GO of rice with *Arabidopsis*, we noticed that the *Arabidopsis GASA* genes are more valuable predictions related to various molecular and biological processes. The GO

Table 2

The *OsGASA* rice gene expression in specific tissues and responses to various biotic and abiotic stresses, where $\uparrow$ represents up-regulation; $\downarrow$ represents down-regulation.

Gene Name (MSU (LOC_Os ID))	<i>OsGASA</i> Gene Expression in Specific Tissues	<i>OsGASA</i> Gene Functions and Response to Stresses
<i>OsGASA1</i> (LOC_Os03g41060.1)	Heading stage (inflorescences, post, pre, endosperm) Seed after pollination	Gibberellin; Salt; drought; PEG; Cadmium $\uparrow$ Response to Cold; Heat; Potassium $\downarrow$
<i>OsGASA2</i> (LOC_Os03g55290.1)	Anthesis; Inflorescence emergence; Endosperm; Seed;	Response to Gibberellin; Drought; Anaerobic condition; Cadmium $\uparrow$ and NaCl, Heat $\downarrow$
<i>OsGASA3</i> (LOC_Os04g39110.1)	Inflorescence; Plant egg cells; Plant embryo; Pistil; Seed after pollination	Response to Cold; Salt; Anaerobic condition $\uparrow$; Cold; Heat $\downarrow$
<i>OsGASA4</i> (LOC_Os05g31280.1)	Early inflorescences; Pistil; Seeds after pollination	Response to salt; biotic stress and Gibberellin; SA $\uparrow$, Drought $\downarrow$
<i>OsGASA5</i> (LOC_Os05g35690.1)	Anthesis; Endosperm; Seed; Pre-flowering	Response to Drought $\uparrow$, biotic stress $\downarrow$, (Developmental regulation) $\uparrow$
<i>OsGASA6</i> (LOC_Os06g15620.1)	Anthesis; Early inflorescences; Pistil; Seeds after pollination	Response to GA; BR; Heat; $\uparrow$ and Cold, biotic stress; $\downarrow$ (Developmental regulation); $\uparrow$
<i>OsGASA7</i> (LOC_Os06g51320.1)	Flower induction stage; Seed after 5 days pollination	Cold; Cadmium and Sensitive to biotic stress; Gibberellin $\uparrow$, Heat; Drought; $\downarrow$
<i>OsGASA8</i> (LOC_Os07g40240.1)	Seed; Pistil; Early inflorescences	Response to abiotic stress; Drought; Salt; Cold $\uparrow$, Heat; Potassium and biotic stress; $\downarrow$
<i>OsGASA9</i> (LOC_Os09g24840.1)	Pistil; Early inflorescences	Response to Cold; $\uparrow$, Drought; Chromium; $\downarrow$
<i>OsGASA10</i> (LOC_Os10g02625.1)	Plant egg cell; Pistil	Response to salt; Cadmium; GA $\uparrow$, Drought; Heat; $\downarrow$

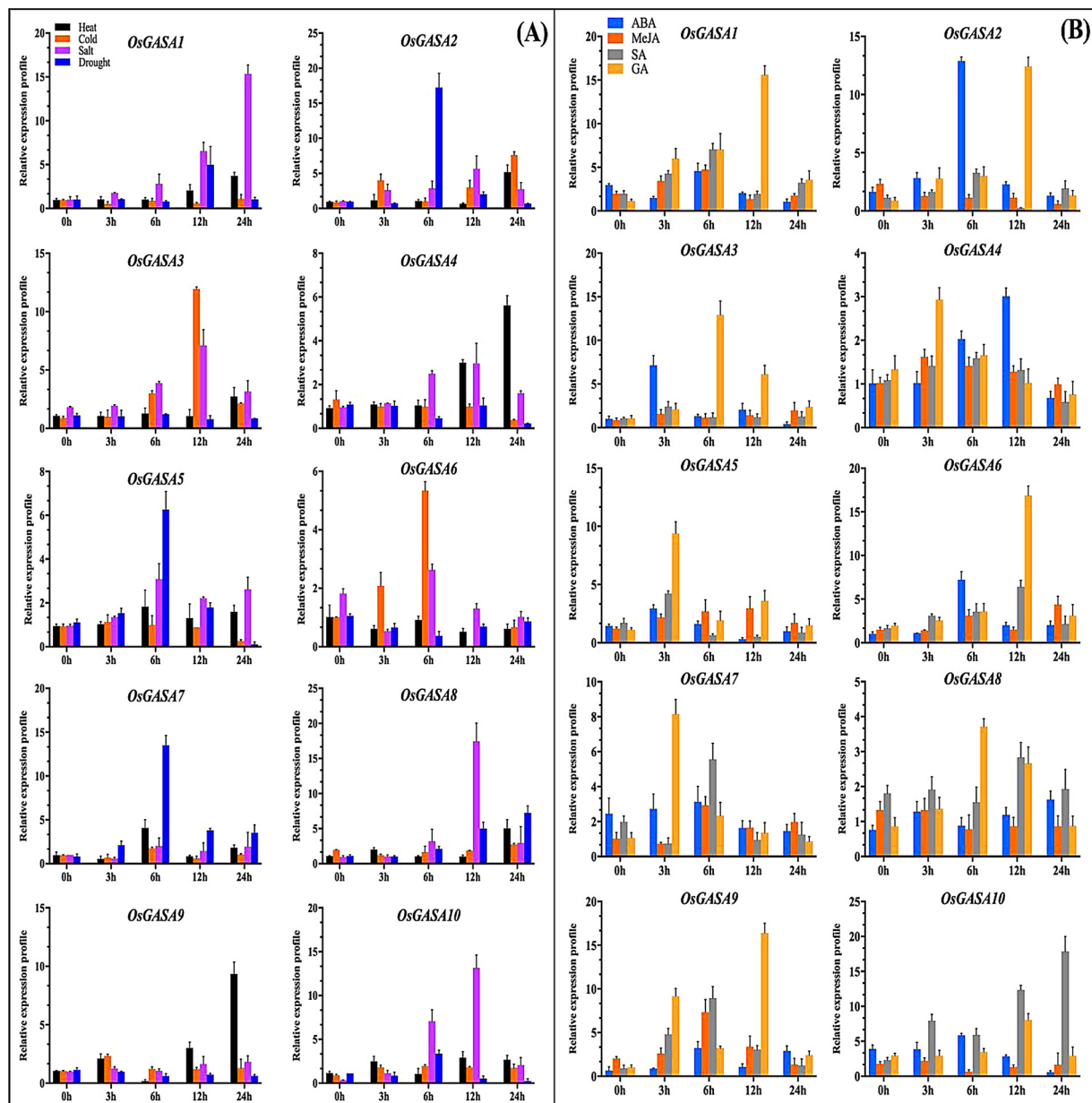


Fig. 7. Inducible expression patterns of rice GASA family genes under abiotic stresses and hormone applications. (A) represents the inducible expression obtained through qRT-PCR analysis under cold (4 °C), heat (40 °C), drought (40 °C), and salt (200 mM NaCl). (B) represents the inducible expression obtained through qRT-PCR analysis under ABA (100 μ M), SA (500 μ M), MeJA (100 μ M), and GA (100 μ M) hormone treatments. Two-week-old seedlings were used for the analysis. Further detail methodology for stress treatments/chemical preparation and plants growth condition were used according to our previous article (Muhammad et al., 2018). Data are means $\pm$ SD (n = 3) and are representative of similar results from three independent experiments.

terms were sub-characterized into three groups, i.e. biological process (BP), molecular function (MF), and cellular component (CC), following Ke et al. (2018). Our results demonstrated 24 functional groups in *Arabidopsis* and 10 in rice (Fig. 9A and B). The obtained results also indicate that the gene functions of the GASA family have substantially differentiated as they evolved. Nevertheless, the three groups significantly participate in the biological process (BP) (GO:0009719) as a GA-mediated signaling pathway and as a response to exogenous ABA, BR, salt and heat stresses, and protein phosphorylation. In cellular component (GO:0005623) as cell wall and extra cellular region and in molecular functions as ATP binding and protein binding. Interestingly, the GO groups of both rice and *Arabidopsis* were mostly identical and performed similar functions (Fig. 9A and B). It is obvious that there are fewer functional terms at molecular level. But, having more biological and cellular processes, these GO features of GASA genes highlight the

significance in hormone signal transduction and response to inducible stresses. Additionally, GASA genes may regulate metabolic pathways, which possibly contribute in the formation of the spikelet in rice plant. Hence, these genes are mostly responsive to GA stimuli and functionally involved in the binding of small molecules, so there is also possibility that these genes may participate in cellular activities, as most of the proteins are localized in the extra cellular region or cell wall. Although this aspect has not yet been clearly investigated, it can be only speculated. However, further research is needed to explore the exact localization of each protein and their particular expression during the panicle developmental process.

4. Discussion

Plants need to integrate external stimuli and internal signals in

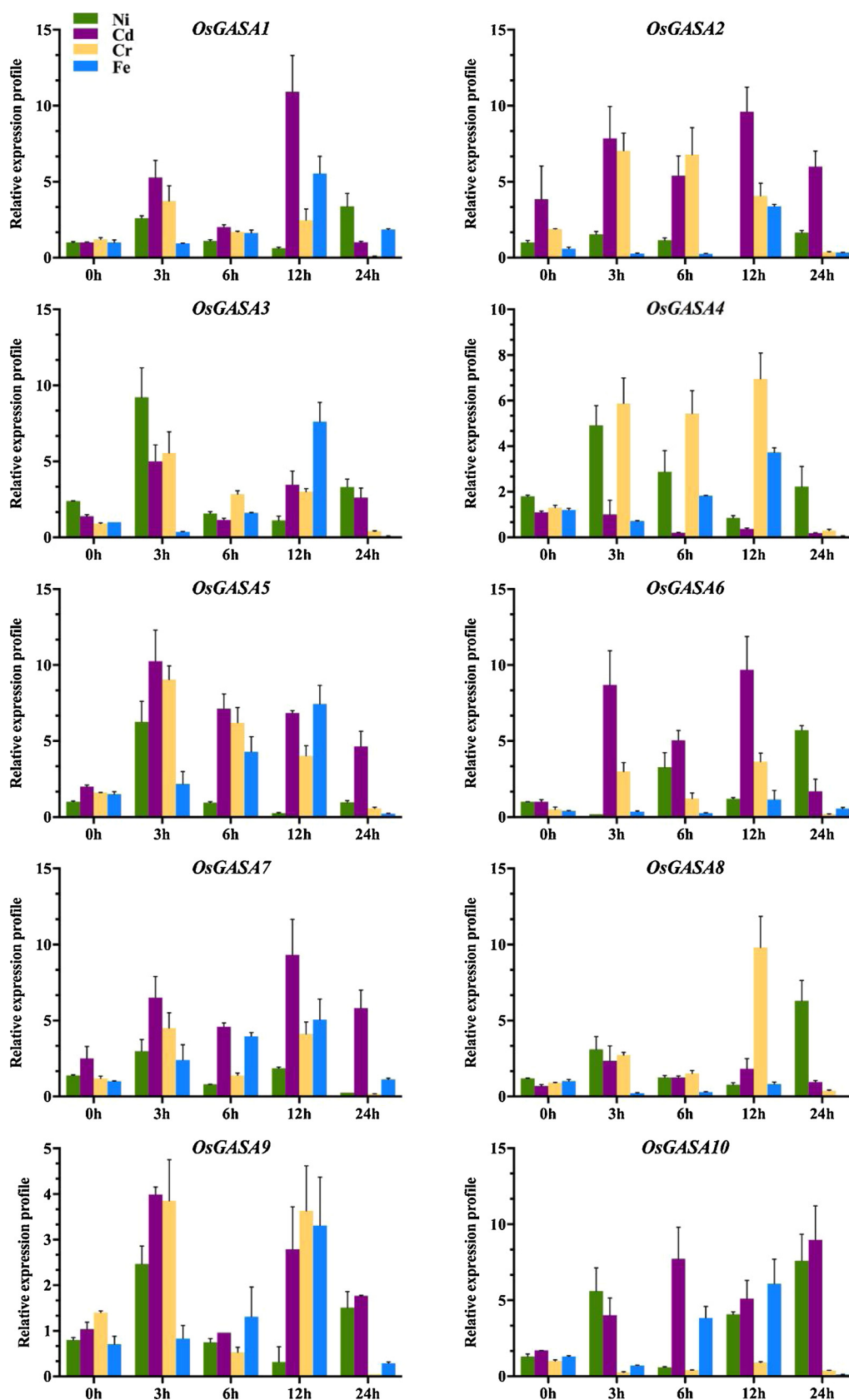


Fig. 8. Inducible expression patterns of GASA family genes in rice under heavy-metal stresses. Two-week-old seedlings were used for the analysis. Detailed methodology for stress treatments/chemical preparation and plant growth conditions is specified in our previous article (Muhammad et al., 2018). The data are means $\pm$ SD ($n = 3$) and indicate similar results from three independent experiments.

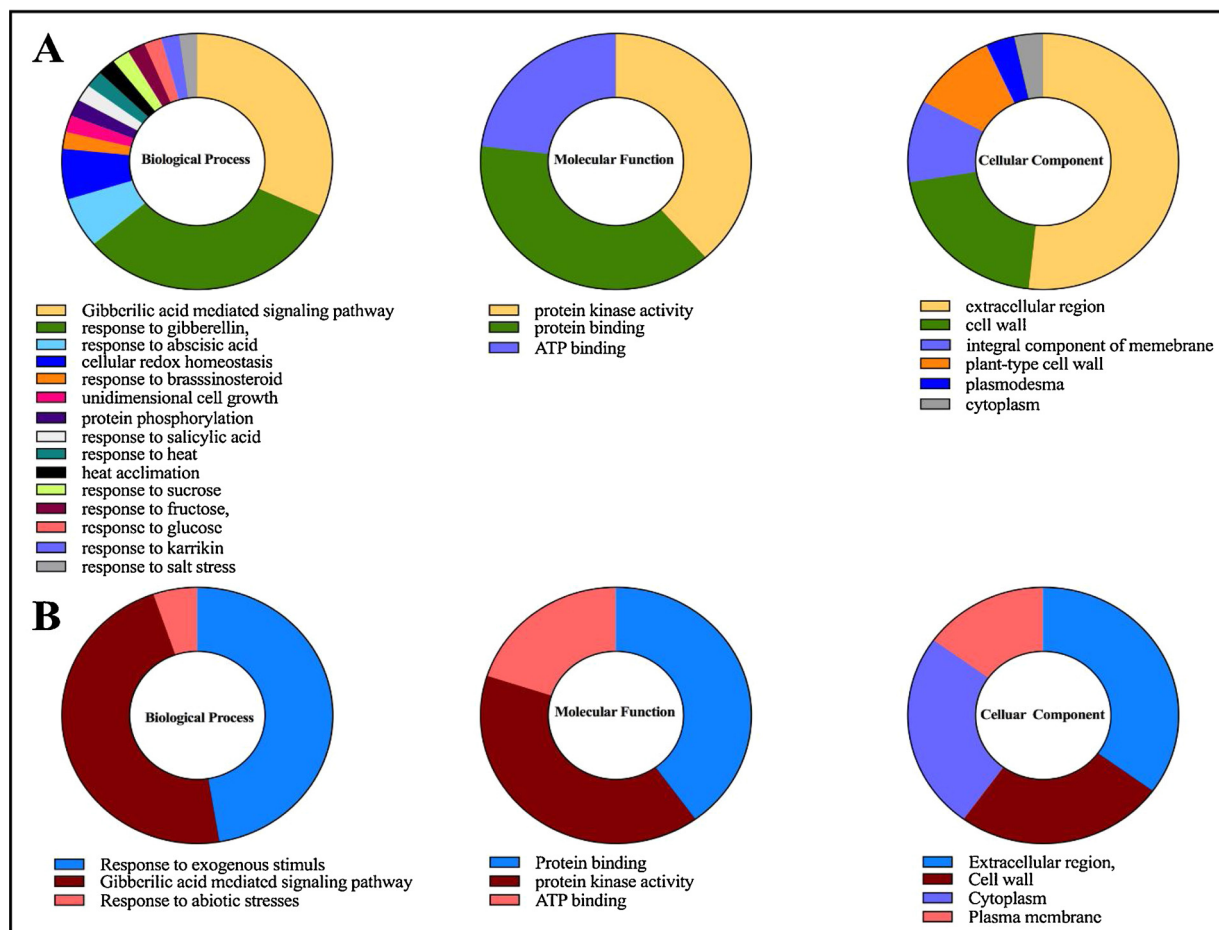


Fig. 9. The distribution of GO ontology analysis among the rice and *Arabidopsis* GASA proteins. The GO appeared in three different categories (biological processes, molecular functions, and cellular component). (A) represents the GO analysis of functionally involved *Arabidopsis* GASA genes, and (B) represents the GO analysis of rice GASA genes. The different colors in the graphical view indicate the cellular localization, biological involvement, and molecular function of the GASA protein family.

order to respond to environmental and endogenous cues through a complex network, and thus to adopt the accurate response and growth pattern of particular plant species. So far, our knowledge about these adverse environmental aspects and signaling cascades has expanded due to the identification of several gene families in the model plant *Arabidopsis*. Among these gene families, GASA (GA-stimulated *Arabidopsis*), a low-molecular weight protein, involved in numerous biological processes and molecular events. It significantly participate in multiple functions, consequently playing a momentous role in plant growth and developmental phases (Nahirnak et al., 2012a,b; Moyano-Canete et al., 2013; Meiyalaghan et al., 2014; Nahirnak et al., 2016; Fan et al., 2017). Keeping in mind the importance and remarkable contribution of the rice GASA gene family, we launched a systematical *in silico* approach to comprehensively investigate and subsequently follow the characterization in terms of the GASA gene family's specific role in panicle morphology and formation of floral biology, tissue specificity, and transcriptomic changes during stresses and exogenous application of hormones in rice plants.

4.1. Genome-wide identification and characteristics of GASA gene families in distantly related species

Recent advances in genome-scale sequencing technologies and frequently updated genomic databases have led to the development of new tools to identify and characterize novel genes (Nahirnak et al., 2016). Moreover, genomic comparisons between species, in which a newly identified gene is compared with orthologous genes in distantly related

species, are useful to generate information and may highlight new functional insights in crop plants (Lyons et al., 2008; Qu et al., 2016). In this context, we retrieved 119 genes in 10 different plant genomes of two major lineages (monocots and eudicots) (Fig. 1A and B, Additional File Table S1). Subsequent analysis of the identified GASA genes and their homologs conformed that they all encode a conserved GASA domain holding 12-cysteine residues (Fig. 2E) (Silverstein et al., 2007; Zhang and Wang, 2008; Fan et al., 2017). Additionally, the structural feature characteristics of all GASA family members consist of polypeptide molecules with an N-terminal 18–34 amino acid hydrophobic region that possibly serves as a cleavable signal sequence according to the SignalP.v4.0 software. The three-dimensional structure prediction approaches discovered that they are intricate in the formation of disulfide bonds, which are responsible for protein folding (Porto and Franco, 2013). Interestingly, the highly conserved 12-cysteine residues at the C-terminal were not changed during the course of evolution, thereby strongly implying the fundamental and important biochemical/biological functions (Wigoda et al., 2006; Rubinovich and Weiss, 2010, 2104). Despite the biological activities like they are responsive to GA stimulus, but as well other hormones and abiotic stresses. The GASA genes also act as an antioxidant, which may be due to the conserved cysteine region (Rubinovich and Weiss, 2010; Rubinovich et al., 2014). All the GASA gene families found in different species share identical physiochemical characteristics, such as maximum protein homology, length, weight, instability index, and conserved Cys residues (Fig. 3B, Additional File Table. S1). These outcomes are consistent with the fact that GASA genes encode small proteins (Berrocal-Lobo et al., 2002;

Nahirnak et al., 2016; Fan et al., 2017). However, we perceived some structural differences in exons/introns of GASA genes in similar groups, implying that exon or intron loss during long-term evolutionary processes may be the consequences of chromosomal rearrangements. In support of this statement (Xu et al., 2012; Guo et al., 2013), we noticed the same type of differences in *MdGASA* gene exon/intron distributions in similar groups, which were associated with chromosomal rearrangements during the evolutionary process. Therefore, in light of this evidence, it can be concluded that GASA proteins carry unique features, therefore occupying the specific position in the cell and plausibly may perform similar functions or biological activities at the cellular level.

4.2. Diversity in phylogeny does not always reflect the functional differences

Previously, the GASA members in potato and other species were classified into three groups (Berrocal-Lobo et al., 2002; Furukawa et al., 2006; Zhang and Wang, 2008; Zimmermann et al., 2010; Meiyalaghan et al., 2014). Therefore, it is possible that the functional diversity, identification, characterization of novel genes, and completion of several genomic databases may cause an alteration in gene expansion and ultimately in the evolutionary process. However, the functional difference may not always reflect the phylogenetic diversity among the groups. According to our phylogenetic analysis, the *OsGASA4* and *OsGASA6* clustered into different subgroups (I and II) (Fig. 1A and B), although they possess similar physiological functions such as being induced by GA and repressed by ABA and SA. The same findings were reported by Qu et al. (2016). Huang et al. (2011) reported that *OsGASA1* and *OsGASA7* are responsible for rice grain elongation, while the *OsGASA7* homologs in *Triticum aestivum* and *T. Urutu* (*TuGASA7*) are also significantly associated with grain size and weight (Ling et al., 2013). Further study revealed that both are significantly involved in GA and BR pathways (Wang et al., 2009). Remarkably, these genes are clustered into different subgroups (III and IV) (Fig. 1A and B) in our tree. Likewise, the rice *OsGASA3* and *OsGASA4* are involved in cell division and proliferation in tissues, suggesting their possible role in rice panicle development (Furukawa et al., 2006; Lee et al., 2017). Moyano-Canete et al. (2013) showed that *FaGAST1* and *FaGAST2* exhibit differences in their expression profile and subfamilies but share a common physiological function. Moreover, potato GASA genes are grouped into dissimilar subgroups but implicated the same function in pathogen resistance (Segura et al., 1999; Berrocal-Lobo et al., 2002; Almasia et al., 2008; Faccio et al., 2011; Balaji and Smart, 2012; Rong et al., 2013; Mohan et al., 2014; Nahirnak et al., 2016). Therefore, considering the above facts and evidence, it can be suggested that clustering of the GASA genes into different subgroups in the phylogenetic tree may not associate with functional differences. The reason for clustering within the same group might be the close homology in genetic architecture and protein nature. Moreover, by searching several other aspects, we found that the close associates of *OsGASA* genes within subgroups possibly share a common ancestor, suggesting strong negative selection pressure in evolutionary history and thus in similar biological activities and functions. Hence, the above evidence and research findings clearly indicate that clustering in the same subgroup may have some homology in the evolutionary process and specific functional associations, while phylogenetic diversity does not always imply functional differences.

4.3. GASA proteins have obvious tissue specificity and play integral roles in the regulation of panicle developmental stages

We also elucidated the spatial regulation of the rice GASA gene family, as they have highly specific expression patterns with common characteristics reflecting functional collaboration in particular tissues. The qRT-PCR and microarray data exhibit slight differences in their expression level, but the results are highly consistent with the clustering

phylogenetic tree (Fig. 1A), although there are some differences in expression pattern obtained from both techniques that might be due to the experimental variation or distinct genotypes analyzed, e.g. the age of the plant and the sampling time. This statement was further supported by several researchers; they stated that the RNA-seq data and expression studies were diverse in potato GASA genes, which may be due to the variation in growth conditions and genotype (Segura et al., 1999; Berrocal-Lobo et al., 2002; Almasia et al., 2008; Meiyalaghan et al., 2014). Interestingly, all the *OsGASA* genes are highly regulated during pre/post flower emergence, particularly in anthers and pistil (Fig. 4). The tapetum appears to be a major site of GA production within the developing anther (Itoh et al., 2001; Kaneko et al., 2003; Hu et al., 2008). For example, the expression of the *Arabidopsis* GA-biosynthetic genes *AtGA3ox3* and *AtGA3ox4* reaches its maximum in the tapetum just prior to its degradation preceding pollen dehiscence (Hu et al., 2008). This finding is consistent with results from rice, signifying that the capacity for GA biosynthesis is reached relatively late in development (Chhun et al., 2007; Hirano et al., 2008). Almost all the *OsGASA* genes are down-regulated during the booting stage, though there is variable expression in the seedling stage. However, our results displayed that the members of subgroup II and IV (*OsGASA6/9* and *OsGASA1/2/3/8*) located on chromosome 3/4/6/7 are share similar expression patterns and gene structure (Fig. 2D). Furukawa et al. (2006) reported that *OsGASA1* and *OsGASA2* genes are highly expressed during panicle, where a mutation in *OsGASA1* gene leads to reduction in cell size and loss function as a GA hormone regulation (Wang et al., 2009; Lee et al., 2017). Likewise, in Thai jasmine rice, *OsGASA3* showed a response during the seedling stage and expressed more dominantly in the inflorescence than in other developmental stages (Jain et al., 2007; Boonpa et al., 2018). Several other researchers have also elaborated the contribution of rice GASA genes in various biological processes, such as cell elongation, shoot emergence, and flowering (Shi et al., 1992; Ben-Nissan et al., 2004; Wigoda et al., 2006; Alonso-Ramírez et al., 2009; Wang et al., 2009; Rubinovich and Weiss, 2010; Sun et al., 2013; Lee et al., 2015). Together, these facts provide sufficient support that rice GASA genes have obvious tissue specificity.

Earlier studies on GASA genes described the contribution in the regulation of flower induction in several other species, such as *Arabidopsis thaliana* (Zhang et al., 2009; Roxrud et al., 2007; Qu et al., 2016), strawberry (Fuente et al., 2006), potato (Berrocal-Lobo et al., 2002), *Gerbera hybrida* (Kotilainen et al., 1999), and *Petunia hybrida* (Ben-Nissan et al., 2004). Improvement of grain shape and size effectively contribute to the yield and quality of rice. In the past, traits associated with grain shape have been identified and cloned in rice plants (Luo et al., 2013; Si et al., 2016). Several studies have proved that some regulators have been positively associated with grain length, width, and shell size in rice (Song et al., 2007; Terao et al., 2010; Tabuchi et al., 2011; Wang et al., 2012). Previously, the *OsGASA* proteins have shown direct involvement in panicle development and grain size (Furukawa et al., 2006). Therefore, here we evaluated the expression level of *OsGASA* genes in three different stages under GA hormone regulation along with morphological changes in grain length and width. These findings revealed the differential expression of GASA genes, dominantly regulated in stage-2 and stage-3, whereas the regulation of stage-1 was considerably lower for most of the genes (Fig. 5B). The predominantly expressed genes especially related to the development of rice panicle and inflorescence architecture rapidly changed due to the critical period of panicle morphogenesis and consequently changes in whole floral biology (Meng et al., 2017). Other reasons are differences in the transcription level of a particular gene regulation, potentially caused by the variation in sample collection or the dissimilar age/tissue of individual plants (Ikeda et al., 2004; Itoh et al., 2005). Thus, there is a strong possibility that the regulation of these genes is important in the formation of the panicle developmental process, especially during transition stages. It has also been reported by numerous researchers that mRNA of *OsGASA* genes are stimulated by GA hormone, implying

important roles in many processes, e.g., seed germination, cell elongation, flower development, grain size, and width (Richards et al., 2001; Lee et al., 2015). More recently, a report disclosed that *OsGASA3* expression is regulated by phytohormones and thus influences cellular changes and panicle differentiation (Lee et al., 2017). Similarly, a comparative study performed in two rice lines (NIL-gw8 and NIL-gw8^{Amo}) during panicle formation in three different stages revealed that various types of genes are highly expressed during stage-2 and stage-3 (Ke et al., 2018), which is quite consistent with our obtained results. In this regard, it can be postulated that *OsGASA* genes can certainly promote the activities of other genes and participate in GA signaling pathways that can boost the expression of these genes during rice panicle differentiation and increase grain size/length. We also speculate a theoretical approach that the antagonistic interaction with hormonal regulation resulted in GA as the principal regulator of hormonal pathways and thus may modulate phenotypic differences during panicle development in rice. Another possible reason may be that the *OsGASA* genes influence the biosynthesis of secondary metabolites and thus may help in the regulation of rice panicle formation in developmental stages. Further studies are needed (forward and reverse genetics, hormonal pathway interactions) to explore this aspect. Thus, in light of all above-mentioned facts, it can be concluded that rice *GASA* genes play an integral role during panicle development and different transition stages.

4.4. *OsGASA* genes are not only induced by external stimuli but hormonal applications as well

External unfavorable environments (salinity, ion toxicity, drought, and temperature) considerably affect plant growth and development, where the transcriptional regulation directed by these abiotic stresses can cause general or specific effects in plants (Rengasamy, 2006; Stein et al., 2009; Cramer et al., 2011; Chinnusamy et al., 2007; Nakashima et al., 2009; Wang et al., 2013). Previously, it has been reported that the majority of *OsGASA* gene transcripts are responsive to abiotic stresses (Sun et al., 2013; Ko et al., 2007; Zhang and Wang, 2011), redox responses (Wigoda et al., 2006; Rubinovich and Weiss, 2010; Rubinovich et al., 2014), and hormones (Nambara et al., 2010; Nahirñak et al., 2012a,b). In this context, we performed the inducible expression pattern of *GASA* genes under abiotic, metal stresses and hormonal application. *In silico* analysis also revealed a number of potential *cis*-elements of *GASA* genes upstream of the transcription start sites, which are stress responsive and regulated by hormones (Additional File Fig. S2, Table. S3). In the present study, *OsGASA1/8/10* showed a response to salt stress, and *OsGASA2/5/7* were up-regulated by drought, *OsGASA3/6* by cold, and *OsGASA9/4* by heat stress. Ko et al. (2007) stated that *GASA4* is highly expressed under heat stress, and our results indicated the same pattern of expression for the *OsGASA4* gene (Fig. 7A). Moreover, the involvement of rice *GASA* genes in response to salt and drought stresses was also observed (Alonso-Ramírez et al., 2009; Sun et al., 2013). Likewise, Li et al. (2011) reported that *GsGASA1* expression responds to cold stress in *Arabidopsis*. Lee et al. (2017) demonstrated that the transcripts of the rice *OsGASA1* gene were induced by salt or ABA treatment; our results are highly consistent with their findings, with the expression level of the *OsGASA1* gene substantially induced under salt stress at certain time points (Fig. 7A). More recently, the homolog of *OsGASA1*, *TaGASR1*, was identified to be involved in tolerance to heat and oxidative stress in wheat (Zhang et al., 2017). The functional differences among the *GASA* family members are also described by other researchers (Wang et al., 2009; Zhang et al., 2009; Zhang and Wang, 2011; Sun et al., 2013; Rubinovich et al., 2014). The data presented herein combined with *in-silico* and experimentally performed analysis are much more consistent (Additional File Fig. S4 and Fig. 7A), implying a clear contribution of rice *GASA* genes towards response to abiotic stresses. However, experimental validation is required to confirm the functionality of the predicted regulation sites for each *GASA*

gene under specific stress environments.

The exogenous application of various hormones altered the expression of rice *GASA* genes (Fig. 7B). We noticed that all the *GASA* genes were significantly regulated by GA and ABA. Boonpa et al. (2018) discussed that *OsGASA7* is involved in gibberellin and brassinosteroid pathways. *OsGASA1* showed an antagonistic effect of GA and ABA hormonal regulation, and the same pattern of regulation for *OsGASA1* was also noticed by Lee et al. (2017). The *Arabidopsis GASA4/6* are up-regulated by GA and down-regulated by ABA in the seedling stage (Zhang and Wang, 2008). Similarly, the potato *GASA* genes are highly induced by ABA in leaves (Berrocal-Lobo et al., 2002). Our results show a complex expression level by hormonal regulation for some genes that is dependent on the length of the treatment. Nevertheless, it can be deduced that *OsGASA* genes are highly induced by hormones and thus play integral roles in the process of growth and development in rice plants.

The different expression level was also observed when young seedlings were subjected to heavy-metal stresses. For example, the noticeable expression of the *GASA* genes was recorded when exposed to Cd, followed by Cr metal stress, for all the genes except *OsGASA8* which is not induced by Cd stress (Fig. 8). An earlier report showed that the cysteine pattern in all *GASA* genes is similar to ATP binding Cassette 1 (ABCE1; Rubinovich and Weiss, 2010). Therefore, the latter carries redox-active iron-sulfur clusters linked to these cysteines (Barthelme et al., 2007). As the *GASA* proteins have a rich region of cysteines, we hypothesized that these genes may also contain iron-sulfur clusters and participate in the redox activity of the proteins, but so far this is not clearly evident. The redox regulation via iron atoms is found in numerous proteins, including ferredoxin and the metallo-protein superoxide dismutase (SOD) (Alscher et al., 2002; Miller, 2004). Rubinovich et al. (2014) reported that *Arabidopsis* cysteine-rich *GASA5* is a redox-active metallo-protein that suppresses gibberellin responses. If we carry our hypothesis of redox activity of *GASA* proteins, we may also conclude that heavy-metal stress enhances the biological activity of *GASA* genes. However, since these genes are newly identified and literature on this subject is limited, this conclusion may not be so reliable.

5. Conclusion

Rice *GASA* genes have been extensively involved in various aspects of plant growth and development. The identification of a greater number of these genes in multi gene families implies the importance of the diversification of these proteins in the plant kingdom. We found that *GASA* genes are highly specific and reliant on the plant growth and stage-specific expression pattern under exogenous GA application, indicating that *GASA* genes participate in regulatory mechanisms and interplay hormonal cross-talks. The diverse expression pattern of these genes under external stimuli suggests a clear contribution in the stress tolerance response. Nevertheless, we enlighten the various aspects and potential importance of *GASA* genes, which will further strengthen the fundamental background of these genes for new researchers. We hope it will also aid in understanding the molecular dissection and physiological response of these genes as well as shed light on other hormone-related rice gene families in rice plants and other cereal crops.

Conflict of interest

The authors declare that they have no conflict of interest.

Funding

This work was supported by the National Natural Science Foundation of China (grant no. 31770204 and 31270299) and the Natural Science Foundation of Shaanxi Province of China (grant no. 382016JM3007).

Author contributions

K.-M.C., W.-Q.L., and W.-T.L. conceptualized the research work; I.M. conducted the experimental work and bioinformatics analysis; X.-Q. J helped with qRT-PCR analysis; A.S and M.A helped with interpretation of data; X.-Y. W., M.-R.Z., and R.S contributed to data curation and preparation of biological material; I.M wrote the manuscript; and K.-M.C critically revised the manuscript. All authors have read and approved the final version of the manuscript.

Acknowledgements

The authors extend their sincere thanks to the editors of this journal and the anonymous reviewers for their valuable comments and suggestions that have significantly improved the manuscript. We are also grateful to Xiao-Yong Wei, Meng-Ru Zhou, and Qiong-Qiong Chen for their help during the experimental period.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:<https://doi.org/10.1016/j.jplph.2019.02.005>.

References

- Almasia, N.I., Bazzini, A.A., Hopp, H.E., Vazquez-Rovere, C., 2008. Overexpression of snakin-1 gene enhances resistance to *Rhizoctonia solani* and *Erwinia carotovora* in transgenic potato plants. *Mol. Plant Pathol.* 9, 329–338.
- Alonso-Ramírez, A., Rodríguez, D., Reyes, D., Jiménez, J.A., Nicolás, G., López-Climent, M., Gómez-Cadenas, A., Nicolás, C., 2009. Evidence for a role of gibberellins in salicylic acid-modulated early plant responses to abiotic stress in Arabidopsis seeds. *Plant Physiol.* 150, 1335–1344.
- Alscher, R.G., Erturk, N., Heath, L.S., 2002. Role of super oxide dismutases (SODs) in controlling oxidative stress in plants. *J. Exp. Bot.* 53, 1331–1341.
- Artimo, P., Jonnalagedda, M., Arnold, K., Baratin, D., Csardi, G., De Castro, E., Duvaud, S., Flegel, V., Fortier, A., Gasteiger, E., 2012. ExPASy: SIB bioinformatics resource portal. *Nucleic Acids Res.* 40, W597–W603.
- Aubert, D., Chevillard, M., Dorne, A.M., Aulard, G., Herzog, M., 1998. Expression patterns of GASA genes in Arabidopsis thaliana: the GASA4 gene is up-regulated by gibberellins in meristematic regions. *Plant Mol. Biol.* 36, 871–883.
- Bailey, T.L., Williams, N., Misleh, C., Li, W.W., 2006. MEME: discovering and analyzing DNA and protein sequence motifs. *Nucleic Acids Res.* 34, W369–W373 (Web Server issue).
- Balaji, V., Smart, C.D., 2012. Over-expression of snakin-2 and extensin-like protein genes restricts pathogen invasiveness and enhances tolerance to *Clavibacter michiganensis* subsp. *michiganensis* in transgenic tomato (*Solanum lycopersicum*). *Trans. Res.* 21, 23–37.
- Barthelme, D., Scheele, U., Dinkelaker, S., Janoschka, A., Macmillan, F., Albers, S.V., Driessen, A.J., Stagni, M.S., Bill, E., Meyer-Klaucke, W., 2007. Structural organization of essential iron-sulfur clusters in the evolutionarily highly conserved ATP-binding cassette protein ABCE1. *J. Biol. Chem.* 282, 14598–14607.
- Ben-Nissan, G., Lee, J.Y., Borohov, A., Weiss, D., 2004. GIP, a Petunia hybrida GA induced cysteine-rich protein: a possible role in shoot elongation and transition to flowering. *Plant J.* 37, 229–238.
- Berrol-Lobo, M., Segura, A., Moreno, M., Lopez, G., Garcia-Olmedo, F., Molina, A., 2002. Snakin-2, an antimicrobial peptide from potato whose gene is locally induced by wounding and responds to pathogen infection. *Plant Physiol.* 128, 951–961.
- Boonpa, K., Tantong, S., Weerawanich, K., Panpet, P., Pringsulaka, O., Yingchutrakul, Y., Roytrakul, S., Sirikantaramas, S., 2018. Heterologous expression and antimicrobial activity of OsGASR3 from rice (*Oryza sativa* L.). *J. Plant Physiol.* 224–225, 95–102.
- Candido, E.S., Porto, W.F., Amaro, D.S., Viana, J.C., Dias, S.C., Franco, O.L., 2011. Structural and functional insights into plant bactericidal peptides. In: Méndez-Vilas, A. (Ed.), *Science Against Microbial Pathogens: Communicating Current Research and Technological Advances*. Formatex, pp. 951–960.
- Chhun, T., Aya, K., Asano, K., Yamamoto, E., Morinaka, Y., Watanabe, M., Kitano, H., Ashikari, M., Matsuoka, M., Ueguchi-Tanaka, M., 2007. Gibberellin regulates pollen viability and pollen tube growth in rice. *Plant Cell* 19, 3876–3888.
- Chinnusamy, V., Zhu, J., Zhu, J.K., 2007. Cold stress regulation of gene expression in plants. *Trends Plant Sci.* 12, 444–451.
- Conesa, A., Gotz, S., 2008. Blast2GO: a comprehensive suite for functional analysis in plant genomics. *Int. J. Plant Genomics* 21 (18), 3674–3676.
- Cramer, G.R., Urano, K., Delrot, S., Pezzotti, M., Shinozaki, K., 2011. Effects of abiotic stress on plants: a systems biology perspective. *BMC Plant Biol.* 11, 163.
- Dash, S., Van Hemert, J., Hong, L., Wise, R.P., Dickerson, J.A., 2012. PLEXdb: gene expression resources for plants and plant pathogens. *Nucleic Acids Res.* 40, D1194–D1201 (Database issue).
- Dong, L., Wang, F., Liu, T., Dong, Z., Li, A., Jing, R., Mao, L., Li, Y., Liu, X., Zhang, K., Wang, D., 2014. Natural variation of TaGASR7-1 A1 affects grain length in common wheat under multiple cultivation conditions. *Mol. Breed.* 34, 937–947.
- Faccio, P.C., Vazquez-Rovere, E., Hopp, G., Gonzalez, C., Decima-Oneto, E., Favret, A., Diaz Paleo Franzone, P., 2011. Increased tolerance to wheat powdery mildew by heterologous constitutive expression of the *Solanum chacoense* snakin-1 gene. *Czech J. Genet. Plant Breed.* 47, 135–141.
- Fan, S., Zhang, D., Zhang, L., Gao, C., Xin, M., Tahir, M.M., Li, Y., Ma, J., Han, M., 2017. Comprehensive analysis of GASA family members in the *Malus domestica* genome: identification, characterization, and their expressions in response to apple flower induction. *BMC Gen.* 18, 1–19.
- Fuente, J.I., Amaya, I., Castillejo, C., Sanchez-Sevilla, J.F., Quesada, M.A., Botella, M.A., Valpuesta, V., 2006. The strawberry gene FaGAST affects plant growth through inhibition of cell elongation. *J. Exp. Bot.* 57, 2401–2411.
- Furukawa, T., Sakaguchi, N., Shimada, H., 2006. Two OsGASR genes, rice GAST homologue genes that are abundant in proliferating tissues, show different expression patterns in developing panicles. *Genes Genet. Syst.* 81, 171–180.
- Guo, R.R., Xu, X.Z., Carole, B., Li, X.Q., Gao, M., Zheng, Y., Wang, X.P., 2013. Genome-wide identification, evolutionary and expression analysis of the aspartic protease gene superfamily in grape. *BMC Gen.* 14.
- Guo, M., Liu, J.H., Ma, X., Zhai, Y.F., Gong, Z.H., Lu, M.H., 2016. Genome-wide analysis of the Hsp70 family genes in pepper (*Capsicum annuum* L.) and functional identification of CaHsp70-2 involvement in heat stress. *Plant Sci.* 252, 246–256.
- Hirano, K., Ueguchi-Tanaka, M., Matsuoka, M., 2008. GID1-mediated gibberellin signaling in plants. *Trends Plant Sci.* 13, 192–199.
- Horton, P., Park, K.J., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C.J., Nakai, K., 2007. WoLF PSORT: protein localization predictor. *Nucleic Acids Res.* 35, 585–587.
- Hruz, T., Laule, O., Szabo, G., Wessendorp, F., Bleuler, S., Oertle, L., Widmayer, P., Gruissem, W., Zimmermann, P., 2008. Genevestigator V3: a reference expression database for the meta-analysis of transcriptomes. *Adv. Bioinf.* 420747.
- Hu, J.H., Mitchum, M.G., Barnaby, N., 2008. Potential sites of bioactive gibberellin production during reproductive growth in Arabidopsis. *Plant Cell* 20, 320–336.
- Huang, X., Zhao, Y., Wei, X., Li, C., Wang, A., Zhao, Q., Li, W., Guo, Y., Deng, L., Zhu, C., Fan, D., Lu, Y., Weng, Q., Liu, K., Zhou, T., Jing, Y., Si, L., Dong, G., Huang, T., Lu, T., Feng, Q., Qian, Q., Li, J., Han, B., 2011. Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. *Nat. Genet.* 44, 32–39.
- Huang, X., Zhao, Y., Wei, X., Li, C., Wang, A., Zhao, Q., Li, W., Guo, Y., Deng, L., Zhu, C., Fan, D., Lu, Y., Weng, Q., Liu, K., Zhou, T., Jing, Y., Si, L., Dong, G., Huang, T., Lu, T., Feng, Q., Qian, Q., Li, J., Han, B., 2012. Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. *Nat. Genet.* 44, 32–39.
- Ikeda, K., Sunohara, H., Nagato, Y., 2004. Developmental course of inflorescence and spikelet in rice. *Breed. Sci.* 54 (2), 147–156.
- Itoh, H., Ueguchi-Tanaka, M., Sentoku, N., Kitano, H., Matsuoka, M., Kobayashi, M., 2001. Cloning and functional analysis of two gibberellin 3 β -hydroxylase genes that are differently expressed during the growth of rice. *Proc. Natl. Acad. Sci. U. S. A.* 98, 8909–8914.
- Itoh, J., Nonomura, K., Ikeda, K., Yamaki, S., Inukai, Y., Yamagishi, H., Kitano, H., Nagato, Y., 2005. Rice plant development: from zygote to spikelet. *Plant Cell Physiol.* 46 (1), 23–47.
- Jain, M., Nijhawan, A., Arora, R., Agarwal, P., Ray, S., Sharma, P., Kapoor, S., Tyagi, A.K., Khurana, J.P., 2007. F-box proteins in rice genome-wide analysis, classification, temporal and spatial gene expression during panicle and seed development, and regulation by light and stress. *Plant Physiol.* 143, 1467–1483.
- Kaneko, M., Itoh, H., Inukai, Y., Sakamoto, T., Ueguchi-Tanaka, M., Ashikari, M., Matsuoka, M., 2003. Where do gibberellin biosynthesis and gibberellin signaling occur in rice plants? *Plant J.* 35, 104–115.
- Ke, S., Liu, X.J., Luan, X., Yang, W., Zhu, H., Liu, G., Zhang, G., Wang, S., 2018. Genome-wide transcriptome profiling provides insights into panicle development of rice (*Oryza sativa* L.). *Gene* 675, 285–300.
- Ko, C.B., Woo, Y.M., Lee, D.J., Lee, M.C., Kim, C.S., 2007. Enhanced tolerance to heat stress in transgenic plants expressing the GASA4 gene. *Plant Physiol. Biochem.* 45, 722–728.
- Kotilainen, M., Helariutta, Y., Mehto, M., Pollanen, E., Albert, V.A., Elomaa, P., Teeri, T.H., 1999. GEG participates in the regulation of cell and organ shape during corolla and carpel development in gerbera hybrida. *Plant Cell* 11, 1093–1104.
- Lee, T.H., Tang, H., Wang, X., Paterson, A.H., 2012. PGDD: a database of gene and genome duplication in plants. *Nucleic Acids Res.* 41, D1152–D1158.
- Lee, S.C., Han, S.K., Kim, S.R., 2015. Salt- and ABA-inducible OsGASR1 is involved in salt tolerance. *J. Plant Biol.* 58, 96–101.
- Lee, S.C., Kim, S.J., Han, S.K., An, G., Kim, S.R., 2017. A gibberellin stimulated-transcript, OsGARS1, controls seedling growth and α -amylase expression in rice. *J. Plant Physiol.* 214, 116–122.
- Lescot, M., Dehais, P., Thijs, G., Marchal, K., Moreau, Y., Van de Peer, Y., Rouze, P., Rombauts, S., 2002. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Res.* 30, 325–327.
- Leticia, I., Bork, P., 2016. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* 44, W242–W245.
- Li, K.L., Bai, X., Li, Y., Cai, H., Ji, W., Tang, L.L., Wen, Y.D., Zhu, Y.M., 2011. GsGASA1 mediated root growth inhibition in response to chronic cold stress is marked by the accumulation of DELLAs. *J. Plant Physiol.* 168, 2153–2160.
- Ling, H.Q., Zhao, S., Liu, D., Wang, J., Sun, H., Zhang, C., Fan, H., Li, D., Dong, L., Tao, Y., Gao, C., Wu, H., Li, Y., Cui, Y., Guo, X., Zheng, S., Wang, B., Yu, K., Liang, Q., Yang, W., Lou, X., Chen, J., Feng, M., Jian, J., Zhang, X., Luo, G., Jiang, Y., Liu, J., Wang, Z., Sha, Y., Zhang, B., Wu, H., Tang, D., Shen, Q., Xue, P., Zou, S., Wang, X., Liu, X.,

- Wang, F., Yang, Y., An, X., Dong, Z., Zhang, K., Zhang, X., Luo, M.C., Dvorak, J., Tong, Y., Wang, J., Yang, H., Li, Z., Wang, D., Zhang, A., Wang, J., 2013. Draft genome of the wheat A-genome progenitor *Triticum urartu*. *Nature* 496, 87–90.
- Liu, J.P., Van Eck, J., Cong, B., Tanksley, S.D., 2002. A new class of regulatory genes underlying the cause of pear-shaped tomato fruit. *Proc. Natl. Acad. Sci. U. S. A.* 99 (20), 13302–13306.
- Luo, J., Liu, H., Zhou, T., Gu, B., Huang, X., Shangguan, Y., Zhu, J., Li, Y., Zhao, Y., Wang, Y., Zhao, Q., Wang, A., Wang, Z., Sang, T., Wang, Z., Han, B., 2013. An-1 encodes a basic helix-loop-helix protein that regulates awn development, grain size, and grain number in rice. *Plant Cell* 25 (9), 3360–3376.
- Meiyalaghan, S., Thomson, S.J., Fiers, M.W., Barrell, P.J., Latimer, J.M., Mohan, S., Jones, E.E., Conner, A.J., Jacobs, J.M., 2014. Structure and expression of *GSL1* and *GSL2* genes encoding gibberellin stimulated-like proteins in diploid and highly heterozygous tetraploid potato reveals their highly conserved and essential status. *BMC Gen.* 15, 2.
- Meng, Q., Li, X., Zhu, W., Yang, L., Liang, W., Dreni, L., Zhang, D., 2017. Regulatory network and genetic interactions established by *OsMADS34* in rice inflorescence and spikelet morphogenesis. *J. Integr. Plant Biol.* 59 (9), 693–707.
- Miller, A.F., 2004. Superoxide dismutases: active sites that save, but a protein that kills. *Curr. Opin. Chem. Biol.* 8, 162–168.
- Mohan, S., Meiyalaghan, S., Latimer, J.M., Gatehouse, M.L., Monaghan, K.S., Vanga, B.R., Pitman, A.R., Jones, E.E., Conner, A.J., Jacobs, J.M., 2014. *GSL2* over-expression confers resistance to *Pectobacterium atrosepticum* in potato. *Theor. Appl. Genet.* 127, 677–689.
- Moyano-Canete, M.L., Bellido, N., Garcia-Caparrós, L., Medina-Puche, F., Amil-Ruiz, J.A., Gonzalez-Reyes, J.L., Cabarello, J., Munoz-Blanco, R., Blanco-Portales, 2013. *FaGAST2*, a strawberry ripening-related gene, acts together with *FaGAST1* to determine cell size of the fruit receptacle. *Plant Cell Physiol.* 54, 218–236.
- Muhammad, I., Jing, X.Q., Shalmani, A., Ali, M., Yi, S., Gan, P.F., Li, W.Q., Liu, W.T., Chen, K.M., 2018. Comparative in silico analysis of Ferric Reduction Oxidase (*FRO*) genes expression patterns in response to abiotic stresses, metal and hormone applications. *Molecules.* 23, 1163.
- Mutasa-Göttgens, E., Hedden, P., 2009. Gibberellin as a factor in floral regulatory networks. *J. Exp. Bot.* 60, 1979–1989.
- Nahirnak, V., 2015. Análisis funcional de *snakin-1* en plantas transgénicas de *Solanum tuberosum*.
- Nahirnak, V., Almasia, N.I., Fernandez, P.V., Hopp, H.E., Estevez, J.M., Carrari, F., Vazquez-Rovere, C., 2012a. Potato *snakin-1* gene silencing affects cell division, primary metabolism, and cell wall composition. *Plant Physiol.* 158, 252–263.
- Nahirnak, V., Almasia, N.I., Hopp, H.E., Vazquez-Rovere, C., 2012b. *Snakin/GASA* proteins: involvement in hormone crosstalk and redox homeostasis. *Plant Signal. Behav.* 7, 1004–1008.
- Nahirnak, V., Rivarola, M., de-Urreta, M.G., Paniego, N., Hopp, H.E., Almasia, N.I., Vazquez-Rovere, C., 2016. Genome-wide analysis of the *Snakin/GASA* gene family in *Solanum tuberosum* cv. Kennebec. *Am. J. Potato Res.* 93, 172–188.
- Nakashima, K., Ito, Y., Yamaguchi-Shinozaki, K., 2009. Transcriptional regulatory networks in response to abiotic stresses in *Arabidopsis* and grasses. *Plant Physiol.* 149, 88–95.
- Nambara, E.M., Okamoto, K., Tatematsu, R.Y., Seo, M., Kamiya, Y., 2010. Absciscic acid and the control of seed dormancy and germination. *Seed Sci. Res.* 20, 55–67.
- Porto, W.F., Franco, O.L., 2013. Theoretical structural insights into the *snakin/GASA* family. *Peptides* 44, 163–167.
- Qu, J., Kang, S.G., Hah, C., Jang, J.C., 2016. Molecular and cellular characterization of *GA*-stimulated transcripts *GASA4* and *GASA6* in *Arabidopsis thaliana*. *Plant Sci.* 246, 1–10.
- Rengasamy, P., 2006. World salinization with emphasis on Australia. *J. Exp. Bot.* 57, 1017–1023.
- Richards, D.E., King, K.E., Aitali, T., Harberd, N.P., 2001. How gibberellin regulates plant growth and development: a molecular genetic analysis of gibberellin signaling. *Annu. Rev. Plant Physiol.* 52, 67–88.
- Rong, W., Qi, L., Wang, J., Du, L., Xu, H., Wang, A., Zhang, Z., 2013. Expression of a potato antimicrobial peptide *SN1* increases resistance to take-all pathogen *Gaeumannomyces graminis var. tritici* in transgenic wheat. *Funct. Integr. Genom.* 13, 403–409.
- Roxrud, I., Lid, S.E., Fletcher, J.C., Schmidt, E.D., Opsahl-Sorteborg, H.G., 2007. *GASA4*, one of the 14-member *Arabidopsis GASA* family of small polypeptides, regulates flowering and seed development. *Plant Cell Physiol.* 48, 471–480.
- Rubio-Vicovich, L., Weiss, D., 2010. The *Arabidopsis* cysteine-rich protein *GASA4* promotes *GA* responses and exhibits redox activity in bacteria and in planta. *Plant J.* 64, 1018–1027.
- Rubio-Vicovich, L., Ruthstein, S., Weiss, D., 2014. The *Arabidopsis* cysteine-rich *GASA5* is a redox-active metalloprotein that suppresses gibberellin responses. *Mol. Plant* 7, 244–247.
- Schultz, J., Milpetz, F., Bork, P., Ponting, C.P., 1998. SMART, a simple modular architecture research tool: identification of signaling domains. *Proc. Natl. Acad. Sci. U. S. A.* 95, 5857–5864.
- Segura, A., Moreno, M., Madueno, F., Molina, A., Olmedo, G., 1999. *Snakin-1*, a peptide from potato that is active against plant pathogens. *Mol. Plant Microbe Interact.* 12, 16–23.
- Shi, L., Gast, R.T., Gopalraj, M., Olszewski, N.E., 1992. Characterization of a shoot-specific, *GA3*- and *ABA*-regulated gene from tomato. *Plant J.* 2, 153–159.
- Si, L., Chen, J., Huang, X., Gong, H., Luo, J., Hou, Q., Zhou, T., Lu, T., Zhu, J., Shangguan, Y., Chen, E., Gong, C., Zhao, Q., Jing, Y., Zhao, Y., Li, Y., Cui, L., Fan, D., Lu, Y., Weng, Q., Wang, Y., Zhan, Q., Liu, K., Wei, X., An, K., An, G., Han, B., 2016. *OsSPL13* controls grain size in cultivated rice. *Nat. Genet.* 48 (4), 447–456.
- Silverstein, K.A., Moskal Jr., W.A., Wu, H.C., Underwood, B.A., Graham, M.A., Town, C.D., VandenBosch, K.A., 2007. Small cysteine-rich peptides resembling antimicrobial peptides have been under-predicted in plants. *Plant J.* 51, 262–280.
- Song, X., Huang, W., Shi, M., Zhu, M., Lin, H., 2007. A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase. *Nat. Genet.* 39 (5), 623–630.
- Stein, R., Duarte, G., Spohr, M., Lopes, S., Fett, J., 2009. Distinct physiological responses of two rice cultivars subjected to iron toxicity under field conditions. *Ann. Appl. Biol.* 154, 269–277.
- Sun, S., Wang, H., Yu, H., Zhong, C., Zhang, X., Peng, J., Wang, X., 2013. *GASA14* regulates leaf expansion and abiotic stress resistance by modulating reactive oxygen species accumulation. *J. Exp. Bot.* 64, 1637–1647.
- Tabuchi, H., Zhang, Y., Hattori, S., Omae, M., Shimizu-Sato, S., Oikawa, T., Qian, Q., Nishimura, M., Kitano, H., Xie, H., Fang, X., Yoshida, H., Kyoizuka, J., Chen, F., Sato, Y., 2011. *LAX PANICLE2* of rice encodes a novel nuclear protein and regulates the formation of axillary meristems. *Plant Cell* 23 (9), 3276–3287.
- Taylor, B.H., Scheuring, C.F., 1994. A molecular marker for lateral root initiation: the *RSI-1* gene of tomato (*Lycopersicon esculentum* Mill.) is activated in early lateral root primordia. *Mol. Gen. Genet.* 243, 148–157.
- Terao, T., Nagata, K., Morino, K., Hirose, T., 2010. A gene controlling the number of primary rachis branches also controls the vascular bundle formation and hence is responsible to increase the harvest index and grain yield in rice. *Theor. Appl. Genet.* 120 (5), 875–893.
- Wang, L., Wang, Z., Xu, Y., Joo, S.H., Kim, S.K., Xue, Z., Xu, Z., Chong, K., 2009. *OSGSR1* is involved in crosstalk between gibberellins and brassinosteroids in rice. *Plant J.* 57, 498–510.
- Wang, S., Wu, K., Yuan, Q., Liu, X., Liu, Z., Lin, X., Zeng, R., Zhu, H., Dong, G., Qian, Q., Zhang, G., Fu, X., 2012. Control of grain size, shape and quality by *OSPL16* in rice. *Nat. Genet.* 44 (8), 950–954.
- Wang, G.F., Li, W.Q., Li, W.Y., Wu, G.L., Zhou, C.Y., Chen, K.M., 2013. Characterization of rice *NADPH* oxidase genes and their expression under various environmental conditions. *Int. J. Mol. Sci.* 14, 9440–9458.
- Wigoda, N., Ben-Nissan, G., Granot, D., Schwartz, A., Weiss, D., 2006. The gibberellin-induced, cysteine-rich protein *GIP2* from *Petunia hybrida* exhibits in planta anti-oxidant activity. *Plant J.* 48, 796–805.
- Xu, G.X., Guo, C.C., Shan, H.Y., Kong, H.Z., 2012. Divergence of duplicate genes in exon-intron structure. *Proc. Natl. Acad. Sci. U. S. A.* 109 (4), 1187–1192.
- Yamaguchi, S., 2008. Gibberellin metabolism and its regulation. *Annu. Rev. Plant Biol.* 59, 225–251.
- Zhang, S., Wang, X., 2008. Expression pattern of *GASA*, downstream genes of *DELLA*, in *Arabidopsis*. *Chin. Sci. Bull.* 53, 3839–3846.
- Zhang, S., Wang, X., 2011. Overexpression of *GASA5* increases the sensitivity of *Arabidopsis* to heat stress. *J. Plant Physiol.* 168, 2093–2101.
- Zhang, S., Yang, C., Peng, J., Sun, S., Wang, X., 2009. *GASA5*, a regulator of flowering time and stem growth in *Arabidopsis thaliana*. *Plant Mol. Biol.* 69, 745–759.
- Zhang, S., Zhang, D., Fan, S., Du, L., Shen, Y., Xing, L., Li, Y., Ma, J., Han, M., 2016. Effect of exogenous *GA3* and its inhibitor paclobutrazol on floral formation, endogenous hormones, and flowering-associated genes in ‘Fuji’ apple (*Malus domestica* Borkh.). *Plant Physiol. Biochem.* 107, 178–186.
- Zhang, L., Geng, X., Zhang, H., Zhou, G., Zhao, A., Wang, F., Zhao, Y., Tian, X., Hu, Z., Xin, M., Yao, Y., Ni, Z., Sun, Q., Peng, H., 2017. Isolation and characterization of heat-responsive gene *TaGASR1* from wheat (*Triticum aestivum* L.). *J. Plant Biol.* 60, 57–65.
- Zimmermann, R., Sakai, H., Hochholdinger, F., 2010. The gibberellin acid stimulated-like gene family in maize and its role in lateral root development. *Plant Physiol.* 152, 356–365.

Additional Figures and Tables

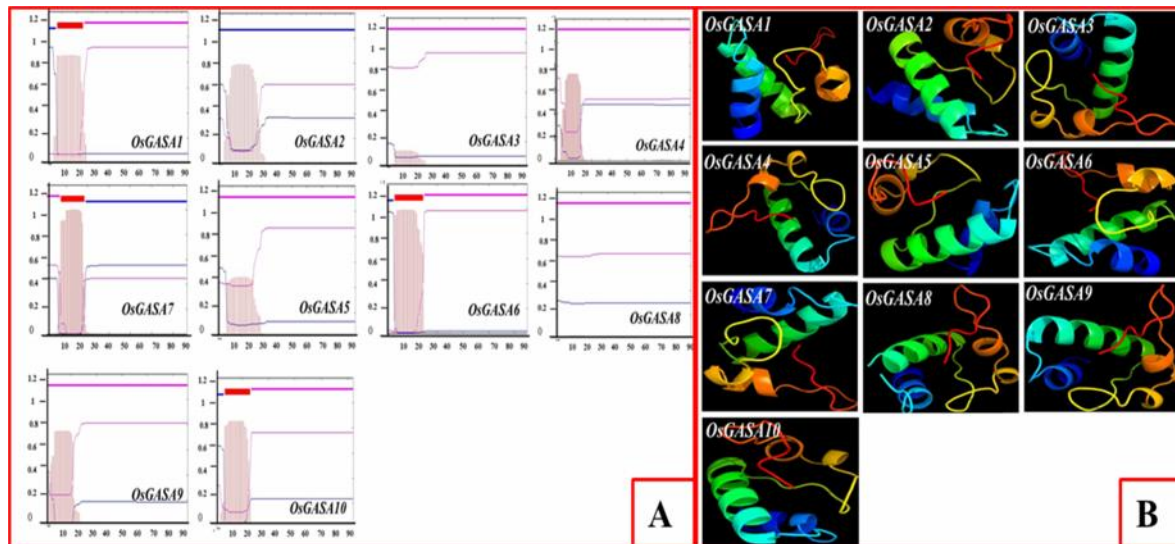


Figure S1. A. Transmembrane topology analysis of the GASA rice gene. The red peaks indicate the predicted transmembrane helices. B. The predicted three-dimensional tertiary structures of OsGASA proteins.

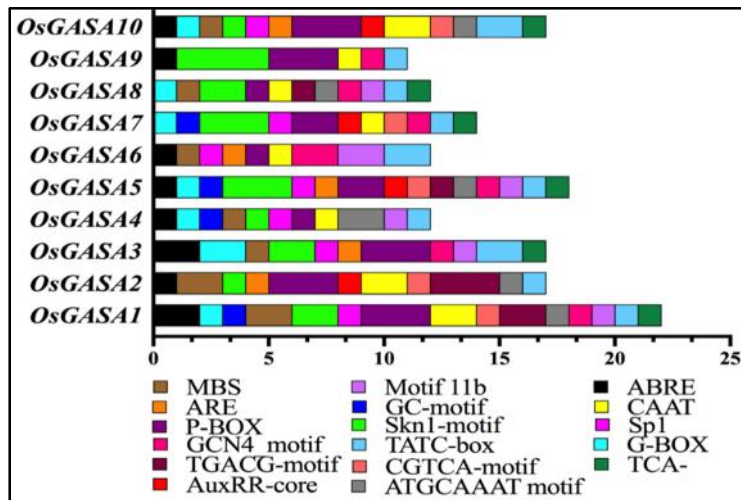


Figure S2. The *cis*-elements responding to hormone and abiotic stress identified in the sense strands of GASA promoter region. The different *cis*-elements are presented in different colors.

Figure S3. Amino acid sequence alignment analysis of GASA proteins from 10 species; *Oryza sativa*, *Zea mays*, *Brachypodium distachyon*, *Sorghum bicolor*, *Solanum tuberosum*, and *Setaria italic* in monocots, and *Arabidopsis thaliana*, *Cucumis sativus*, *Solanum Lycopersicon*, and *Capsicum annuum* in eudicots. Positions of the alignments where more than 50% of the sequences are identical are shown with black boxes. The conserved cysteine region is indicated below the alignment.

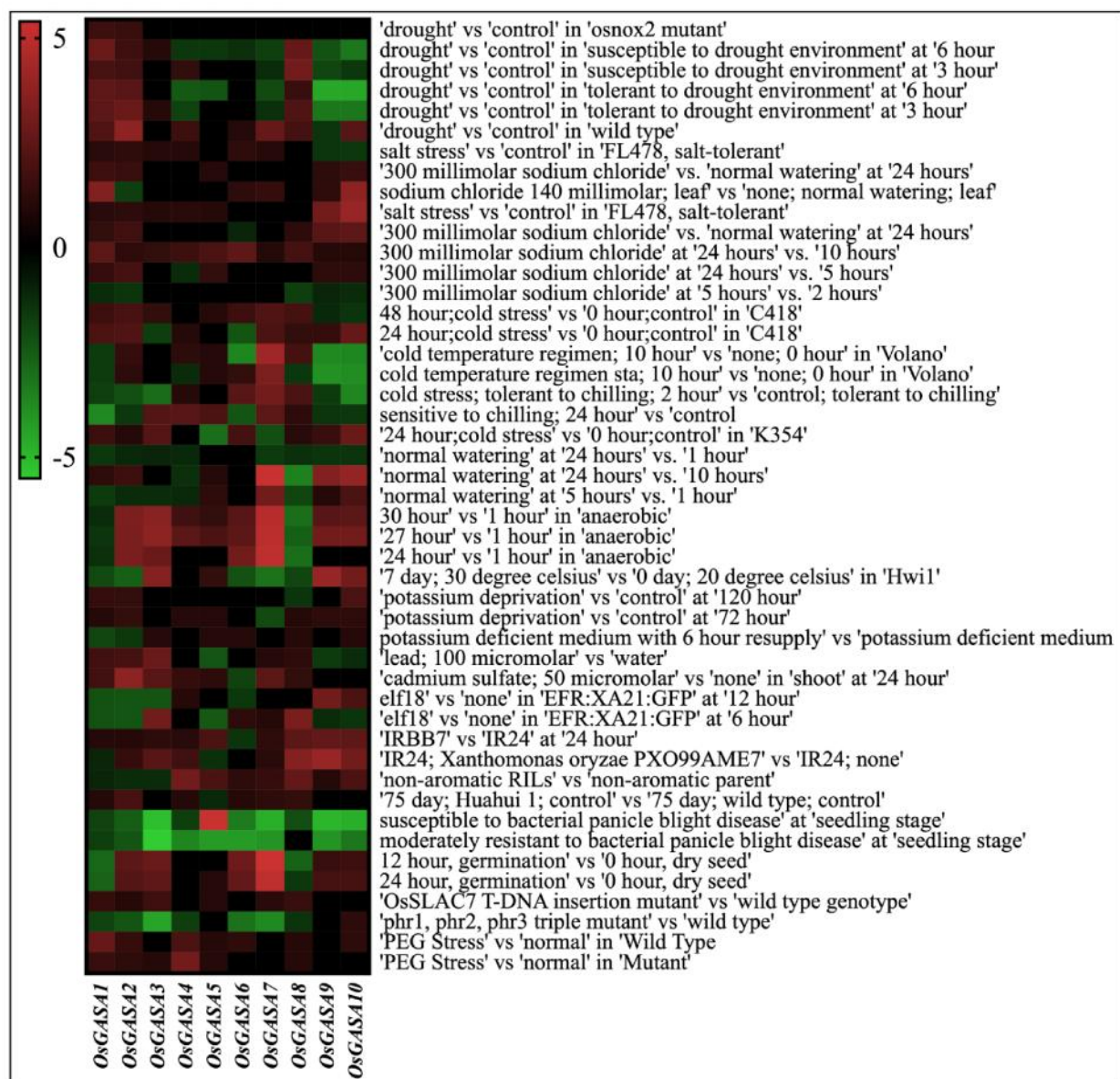


Figure S4. The inducible expression patterns of *OsGASA* genes under abiotic stresses. The inducible expression patterns of GASAs obtained from the database of Genevestigator and rice ePF expression browser, demonstrating different expression levels of *OsGASA* genes under various abiotic stresses. The results are presented in heatmaps in which the green/red coding boxes reflect the relative signal values and green to red represents increase in up-regulation of the gene expression level.

Table S1. Detailed information on physiochemical characteristics of GASA gene families

Translation ID	Amino Acid	M.W	pI	Composition of amino acid (%)	Instability index	Aliphatic index	GRAVY	N-Glyc (Asn) position	Localization predicted
Os03t0607200-01	94	9929.73	9.12	C(13.8), R,G(11.7),A,L(9.6)	49.93	62.34	-0.073	---	extra, chlo, vacu
Os03t0760800-01	93	9911.59	8.77	C(12.9), L(11.8),S(9.7)	36.78	66.13	-0.138	4	extra, chlo, cyto
Os04t0465300-01	105	11127.99	8.78	A,C(11.4),P(10.5),L(8.6)	48.28	66.10	-0.161	---	extra, mito, cyto, E.R
Os05t0376800-01	152	15922.36	8.98	A(13.8),P(11.8),S(9.9)	61.54	46.45	-0.189	53, 61, 62	extra, vacu, E.R, plas
Os05t0432200-00	92	9685.45	8.86	C(14.1),A(9.8),L(7.6)	40.14	64.78	0.043	---	extra, vacu
Os06t0266800-01	110	11308.32	9.28	G(17,3),C(10.9),L,K(9.1)	55.61	61.27	-0.053	---	extra, vacu
Os06t0729400-01	133	14401.77	9.06	C,S(9.8),A(9.0),L,P(7.5)	40.68	59.47	-0.177	---	extra, nucl, vacu
Os07t0592000-01	102	10678.64	8.78	C(11.8),A,P,V(9.8),G(8.8)	64.55	68.82	0.031	---	chlo
Os09t0414900-01	112	12149.15	9.04	C(10.7),L,K(9.8),A,G(8.9)	33.22	68.04	-0.210	---	extra, chlo, vacu, plas
Os10t0115550-00	71	8016.40	7.99	C(18.3),Y(8.5),E,K,P,V(7.0)	37.41	49.30	-0.090	---	extra, vacu
Bradi1g07890.1	93	10049.69	8.84	C,S(14.0),K,L(9.7),G(8.6)	55.77	51.40	-0.277	---	chlo, extra, cyto
Bradi1g07900.1	93	10165.78	8.76	C,S(12.9),K(9.7),L(8.6)	46.22	52.47	-0.242	---	extra, chlo, cyto
Bradi1g43850.1	94	9826.43	8.58	C(14.9),A(11.7),G,L(8.5)	34.10	67.55	0.017	---	extra, vacu
Bradi2g24320.1	94	9917.88	9.17	C(12.8),G,K(11.7),A(8.5)	38.54	67.55	0.127	---	extra, mito, vacu
Bradi4g30250.1	94	9826.43	8.58	C(14.9),A(11.7),G,L(8.5)	34.10	67.55	0.017	---	vacu, extra, chlo, nucl
Bradi1g07890.1	94	9917.88	9.17	C(12.8),G,K(11.7),L(9.6)	38.54	67.55	0.127	110	chlo, extra, cyto
Bradi1g07900.1	109	12024.98	8.86	C(11.0),G(10.1),L(9.2)	29.17	67.98	-0.199	---	extra, chlo, cyto
Csa1G045890	115	13049.31	9.19	R(13.0),C(11.3),G(10.4)	44.35	53.39	-0.414	---	extra, vacu, cyto,
Csa2G354910	85	9472.29	9.20	C(14.1),L(12.9),G(11.8)	34.97	64.24	-0.171	---	chlo, extra, nucl, cyto
Csa3G872170	102	10585.42	8.57	C(13.7),A(10.8),S,L(8.8)	60.59	76.57	0.249	---	extra, vacu,
Csa4G001520	422	47463.73	9.80	L(10.7),K(9.5),R(8.1)	36.06	91.21	-0.168	---	extra, cyto, plas
Csa6G493840	95	10253.05	9.05	S,C(12.6),K(11.6),L(10.5)	39.16	70.74	-0.075	---	extra, chlo, vacu
Csa2G368260	97	10621.64	9.37	L(14.4),C(12.4),R(10.3)	47.67	74.43	0.041	---	chlo, extra, mito, vacu,
Csa2G402110	106	11876.12	9.39	C(12.3),K(9.4),L(7.5)	51.08	68.11	-0.175	---	extra,
Csa3G826660	104	11458.74	9.64	C(11.5),A,G,P(9.6),K(8.7)	34.12	54.52	-0.183	---	plas, vacu, golgi
Csa3G829190	231	24240.85	9.77	P(27.7),V(10.0),A,K(7.8)	84.05	67.06	-0.263	---	extra, vacu, chlo,
Sb01g007360.1	93	9824.52	8.65	C(12.9),A,L(9.7),G(8.6)	37.35	67.20	-0.025	86	extra, vacu,
Sb01g026925.1	127	14151.30	8.87	C(11.0),S(10.2),L,K(7.9)	55.83	58.35	-0.425	12, 51, 55	nucl, extra, cyto
Sb02g024660.1	115	12536.58	8.95	G,L(11.3),C(10.4),K(8.7)	32.77	64.43	-0.321	---	extra, chlo, cyto, vacu,

Table S1. Continued...

Translation ID	Amino acid	M.W	pI	Composition (%)	Instability index	Aliphatic index	GRAVY	N-Glyc (Asn) position	Localization predicted
Sb06g032730	111	10945.6	8.47	A(21.6),C(11.7),S(10.8)	79.27	76.85	0.268	---	extra, vacu
Sb08g001840	128	13648.6	8.79	K(13.4),G(12.6),C(10.2)	43.09	53.78	-0.688	---	extra,
Sb09g018790	137	14453.0	9.41	P(16.1),A(12.4),C(10.2)	62.75	54.31	-0.160	46	chlo, extra, mito
Sb09g019270	97	10197.8	8.29	C(12.4),L,S(9.3),A,G(8.2)	38.45	69.38	0.039	---	extra, vacu, chlo
Sb09g021340	95	9883.50	8.66	C(12.6),A,G,S(9.5),L(8.4)	34.77	67.74	0.044	---	extra, vacu
Sb10g009640	114	11546.5	9.36	G(16.7),C(10.5),A,K(9.6)	57.25	58.25	-0.079	---	extra, chlo, vacu, cyto
Zm00001d030236_P001	105	11400.1	6.20	C(11.4),S(8.6),L,K,P(7.6)	51.53	64.10	-0.080	---	extra, chlo, nucl, cyto
Zm00001d033195_P001	93	10037.9	8.84	A,C(11.8),L,P(8.6),V(7.5)	48.39	79.78	0.211	---	chlo
Zm00001d033196_P001	93	9751.66	8.75	A(17.2),C(14.0),L(9.7)	56.83	80.00	0.351	---	chlo, mito, extra
Zm00001d033197_P001	144	15612.3	9.26	A(11.8),L(9.7),P(9.0)	57.99	73.96	-0.081	---	chlo, cyto, plas, nucl,
Zm00001d033201_P001	145	14979.44	9.13	P(15.2),A(13.1),C(10.3)	56.60	48.00	-0.191	---	chlo, cyto, plas, nucl,
Zm00001d033369_P001	100	10601.5	9.07	A(15.0),C(12.0),R,L(10.0)	55.21	77.20	0.145	---	chlo
Zm00001d001852_T001	143	15629.3	11.27	A(17.5),R(14.0),L(13.3)	81.66	104.48	0.047	---	chlo, mito, extra
Zm00001d003380_T00	115	11913.0	9.12	A(12.2),C(10.4),G,L(9.6)	37.12	78.96	0.093	110	extra, chlo, vacu, E.R
Zm00001d006750_T001	134	14345.2	8.64	C(13.4),P,S(9.7),G(8.2)	43.94	48.06	-0.418	---	extra, vacu
Zm00001d013222_T001	96	10163.9	8.91	C(12.5),A(10.4),L(9.4)	32.32	66.15	0.044	89	extra, vacu
Zm00001d014888_P001	200	22429.4	8.93	C(11.0),L(10.5),P(9.5)	53.93	71.20	-0.170	---	chlo, extra, nucl, vacu
Zm00001d038056_P001	96	10163.96	8.91	C(12.5),A(10.4),G,L(9.4)	32.32	66.15	0.044	53	extra, chlo, mito, vacu
Zm00001d010193_P001	93	9740.68	8.88	A(18.3),C(14.0),L(9.7)	56.02	81.08	0.363	---	extra, vacu, chlo
Zm00001d045685_P001	147	16339.07	11.79	R(16.3),P(12.2),L(10.9)	100.63	75.10	-0.443	81	chlo, mito, nucl, cyto
Zm00001d023290_T001	129	13564.68	9.14	G(16.3),K(13.2),C(10.1)	44.79	47.05	-0.572	---	extra, vacu
Si033027m	107	11539.65	9.44	K(12.1),C(11.2),A,G,L(9.3)	40.57	64.86	-0.226	---	extra, vacu
Si023795m	91	9636.40	9.03	C(13.2),A(9.9),G(8.8)	30.04	63.30	-0.013	---	extra, vacu
Si024382m	133	13658.83	9.16	A(16.5),P(13.5),C(10.5)	70.09	50.90	-0.111	47	chlo, extra
Si012350m	107	11539.65	9.44	K(12.1),C(11.2),A,G,L(9.3)	40.57	64.86	-0.226	---	extra, vacu
Si027353m	120	12866.96	9.16	K(13.3),C,G(10.8),A,N(8.3)	39.91	57.83	-0.492	---	extra, vacu
Si027818m	106	11725.56	9.10	C,G(10.4),K(9.4),L(7.5)	37.43	56.13	-0.497	---	extra, vacu
Si039591m	75	8459.80	8.22	C(17.3),K(10.7),P,S(8.0)	52.88	36.40	-0.441	---	extra, chlo, cyto, nucl

Gene ID	Amino Acid	M.W	pI	Composition (%)	Instability index	Aliphatic index	GRAVY	N-Glyc (Asn) position	Localization predicted
Solyc01g107370.2.1	89	10000.05	9.40	K(14.6),C(13.5),L(10.1)	20.11	60.34	-0.202	---	extra, vacu, chlo, nucl
Solyc01g111080.2.1	215	835535.43	9.68	L(10.7),S(9.1),I(7.8)	44.94	90.35	-0.194	---	extra, chlo, vacu
Solyc02g083870.2.1	102	11133.82	8.78	C(11.8),A,P,S(8.8),N,L,T(5.9)	52.29	53.63	-0.258	91, 96	extra, vacu
Solyc02g083880.2.1	103	10994.78	9.02	C(11.7),P,S(10.7),A,L,T(7.8)	62.76	60.68	-0.164	92, 97	extra, vacu
Solyc02g089350.2.1	112	12783.99	9.42	C,Q(10.7),K(9.8),N,G,P,V(6.2)	54.16	53.04	-0.556	---	extra, vacu, chlo,
Solyc03g113910.2.1	146	16052.99	9.17	L(11.6),C,V(8.2),A,K(7.5)	45.29	87.47	-0.038	141	extra, chlo, vacu, nucl
Solyc03g116060.2.1	104	11536.72	9.25	C(11.5),K(9.6),L,P(8.7)	38.15	60.10	-0.109	---	extra
Solyc04g017720.2.1	114	12645.17	9.17	P(13.2),L(11.4),C(10.5)	46.38	72.72	-0.141	44	extra, plas, vacu
Solyc04g078200.2.1	65	7185.44	8.96	C(18.5),K(16.9),S(9.2)	38.48	33.08	-0.697	---	nucl, mito, extra
Solyc05g006160.2.1	64	7034.36	9.16	C(18.8),G (10.9),R (7.8)	34.94	33.44	-0.411	59	mito, chlo, nucl, extra
Solyc06g007890.2.1	94	10637.62	8.79	C(14.9),P(10.6),G (6.4)	60.36	52.87	-0.410	---	extra, vacu,
Solyc06g007910.2.1	94	10321.48	9.35	C(14.9),K(10.6),P()7.4	48.35	78.72	0.127	45	extra, cyto, vacu, nucl
Solyc06g007920.1.1	83	8965.31	8.55	C(16.9),K(10.8)P(10.8)	47.11	29.40	-0.682	---	mito, chlo, extra, nucl
Solyc06g069790.2.1	104	11481.79	9.46	K(12.5),C(11.5),P(8.7)	41.24	64.71	-0.182	---	extra, chlo, vacu,
Solyc08g005620.2.1	89	9771.65	9.15	C(13.5),K(13.5),S(11.2)	44.05	63.60	0.022	---	extra
Solyc11g011210.1.1	96	10838.88	9.24	C(12.5),K(11.5),L(8.3)	36.26	56.88	-0.197	69	extra, chlo, nucl
Solyc11g017440.1.1	103	11586.01	9.75	K(12.6),C(11.7),I(7.8)	53.17	73.98	-0.164	7	extra, vacu
Solyc12g042500.1.1	88	9547.24	8.88	C(13.6),K(11.4),S(9.1)	37.50	64.32	-0.144	---	extra, mito, vacu,
Solyc12g089300.1.1	112	11925.80	8.00	C(10.7),G(8.9),L(8.9)	40.37	68.84	-0.080	---	extra, chlo, vacu, nucl
Sotub01g044530.1.1	89	9970.03	9.40	K(14.6),C(13.5),A(9.0)	17.29	61.46	-0.174	---	extra, chlo, nucl, vacu
Sotub01g048490.1.1	104	11037.80	8.91	C(11.5),A(11.5),L,S(10.6)	39.73	76.15	-0.031	---	extra, vacu
Sotub02g027220.1.1	102	11175.90	8.77	C(11.8),A(10.8),P,S(8.8)	54.49	53.73	-0.208	96	extra, vacu
Sotub02g027230.1.1	103	11202.09	9.02	C(11.7),P(9.7),L,S(8.7)	52.36	67.28	-0.091	92, 97	extra, vacu
Sotub02g033550.1.1	112	12713.89	9.42	Q(11.6),C(10.7),K(9.8)	54.16	53.04	-0.523	---	extra, vacu, chlo
Sotub03g027260.1.1	145	15869.89	9.27	L(12.4),A(8.3),C(8.3)	40.96	87.45	0.026	140	extra, vacu
Sotub03g029520.1.1	104	11522.69	9.25	K(12.5),C(11.5),P(8.7)	38.15	59.13	-0.105	---	extra, vacu
Sotub04g014730.1.1	114	12645.17	9.17	L(12.3),P(13.2),C(10.5),	44.69	72.72	-0.147	44	extra, E.R, plas
Sotub04g031180.1.1	88	9690.62	9.06	C(13.6),K(13.5),L(12.5)	29.59	75.34	-0.094	---	chlo, extra, vacu, cyto

Table S1. Continued...

Gene ID	Amino Acid	M.W	pI	Composition (%)	Instability index	Aliphatic index	GRAVY	N-Glyc (Asn) position	Localization predicted
Sotub05g012380.1.1	179	20138.13	8.96	C(7.3),S(8.3),T(8.2)	27.61	67.43	-0.245	54, 113	chlo, nucl, mito, cyto
Sotub05g010910.1.1	50	5458.58	9.25	C(20),R(10),G,K,P,T(8)	50.28	42.80	-0.304	45	cyto, chlo, nucl, mito
Sotub06g026790.1.1	103	11304.59	9.48	C(12.6),K(11.7),P,L(8.7)	36.85	64.47	-0.191	44	extra, chlo, vacu
Sotub08g008240.1.1	89	9756.59	9.16	C(13.5),K(12.4),S(11.2)	43.72	63.60	0.027	---	extra
Sotub11g008460.1.1	96	10865.95	9.32	C,K(12.5),L(9.4),P,S(7.3)	36.26	56.88	-0.237	7	extra, chlo, nucl, mito
Sotub11g013410.1.1	103	11497.90	9.72	K(13.6),C(11.7),A(9.7)	50.63	73.98	-0.200	---	extra, vacu
Sotub11g028480.1.1	100	11371.48	8.76	C(12),L,K(8),S,I(7)	55.24	74.10	0.125	12, 85	cyto, extra, chlo, nucl
Sotub12g015160.1.1	215	22618.77	8.99	L(10.7),P(20),G,V(8.4)	68.58	78.88	0.029	31	chlo, extra, vacu
Sotub12g021350.1.1	88	9549.17	8.76	C(13.6),K(10.2),G,S(8.0)	39.39	64.32	-0.185	---	extra, plas
Sotub12g021370.1.1	88	9665.35	8.58	C(14.8),K(10.2),L(8),S(6)	39.87	62.05	-0.128	---	extra,
Sotub12g027480.1.1	112	12009.80	6.01	C(11.6),G(8.2),L,T(8.0)	40.33	64.46	-0.074	45	extra, cyto, vacu
Capang01g000583	60	6826.10	9.01	C(20),K(13.3),P(11.7)	63.18	24.33	-0.680	---	mito, chlo, nucl
Capang01g005261	59	6298.36	9.49	C(20.3),R(13.6),T,G(10.2)	44.99	24.92	-0.620	---	chlo, extra, mito
Capang02g002323	102	11153.90	9.01	C(11.8),A(10.8),S(7.8)	46.06	49.90	-0.288	91, 96	extra, vacu
Capang02g002324	61	6697.66	8.81	C(19.7),S,N(9.8),R(8.2)	57.46	35.25	-0.534	50, 55	nucl,chlo,extra
Capang03g000718	61	6946.25	9.00	C(19.7),K(13.2),P(11.2)	66.92	22.30	-0.613	---	mito, chlo, nucl
Capang03g000959	145	15757.59	9.35	L(10.3),A(9),C,K(8.3)	41.45	81.38	-0.046	140	extra, vacu, chlo
Capang05g000890	89	9879.99	9.56	K(14.6),C(13.6),L(9.0)	21.69	65.73	-0.087	---	extra, vacu, chlo, cyto
Capang06g001579	61	7013.04	8.02	C(21.3),K(11.5),P,N(6.6)	64.02	22.30	-0.895	---	chlo, nucl, extra, mito
Capang06g002716	113	12553.20	9.46	C(12.4),K(10.6),L,V(7.1)	48.91	73.27	0.103	---	extra, cyto, mito, plas
Capang06g002717	59	6511.60	9.08	C(22),K(11.9),G(10.2)	54.14	19.83	-0.634	---	mito, extra, nucl
AT1G75750	98	10744.66	9.40	C(12.2),L(10.2),A,R(9.2)	43.38	79.69	-0.155	---	extra, golg, plas
AT4G09610	99	10531.44	8.98	C(13.1),L(10.1),S,T(8.1)	32.78	84.65	0.175	---	extra, vacu, mito
AT4G09600	99	10704.61	8.88	C(13.1),L(11.1),G(9.1)	45.34	82.73	0.154	---	extra, vacu
AT5G15230	106	11996.20	9.46	C,K(11.3),S,G(7.5)	57.75	54.25	-0.379	---	extra, chlo, vacu,
AT3G02885	97	10845.03	9.68	C(13.4),K(12.4),L(10.3)	38.56	58.35	-0.192	---	extra, chlo, nucl
AT1G74670	101	11341.38	9.01	C(12.9),K(9.9),L,T(7.9)	45.71	50.20	-0.251	---	extra, vacu
AT2G14900	108	11381.36	8.74	C,K,S(11.1),A(9.3),L(8.3)	33.70	74.17	-0.060	---	extra, vacu, mito

Table S1. Continued...

Gene ID	Amino acid	M.W	pI	Composition (%)	Instability index	Aliphatic index	GRAVY	N-Glyc (Asn) position	Localization predicted
AT2G39540	87	9440.02	8.63	C,S(13.8),K(10.3),G(8)	47.84	62.64	-0.082	---	extra, vacu, chlo
AT1G22690	119	12938.94	9.52	S(12.6),C(10.1),A(8.4)	46.47	61.51	-0.334	32,44	extra, vacu, nucl
AT5G59845	89	9746.47	8.97	C(13.5),S,K(11.2),D,P(5.6)	42.68	64.61	-0.208	---	extra, vacu, chlo, cyto
AT2G18420	94	10154.81	8.67	C(12.8),L(10.6),S(9.6)	35.87	75.74	0.013	---	extra, vacu, chlo
AT2G30810	106	11671.47	7.96	C(11.3),K(9.4),E(8.5)	47.57	53.40	-0.371	---	extra, mito, vacu,
AT5G14920	275	29138.45	9.98	P(29.5),T(14.9),V(9.5)	64.07	51.96	-0.525	---	extra, chlo, vacu
AT1G10588	90	9809.32	7.41	C,S(13.3),K(8.9),D(6.9)	58.35	57.33	-0.196	---	extra, vacu, chlo
AT3G10185	122	13561.74	9.08	P(12.3),C(9.8),Q(9),	59.71	59.10	-0.516	48	nucl, extra

The genome data of above species are downloaded from the public database, mentioned in Table-1. The original Locus Name or IDs are annotated in the database of each presented species. The physiochemical properties of the GASA proteins were analyzed with the online ExPASy proteomics server (Available online: <http://web.expasy.org/protparam/>). The N-glycosylation sites (Asn) of the GASA protein were searched on NetNGlyc (1.0 server) (Available online: <http://www.cbs.dtu.dk/services/NetNGlyc/>), and prediction of subcellular location were analyzed by (<http://www.genscript.com/wolf-psort.html>) and TargetP online tool (<http://www.cbs.dtu.dk/services/TargetP/>). The abbreviations are as follows; Sequenced IDs; MW: molecular weight (kDa), pI: isoelectric point, GRAVY; grand average of hydropathicity, predicted localization; extra-extracellular, vacu-vacuoles, chlo- chloroplast, cyto-cytoplasm, mito-mitochondria, nucl-nucleus, plas-plastids, and E.R-endoplasmic reticulum.

Table S2. GASA gene family members in ten plant species

Lineages	Species	Gene name	Translation ID	CDS	Sequence position	Chr	Strand
Monocots	<i>O. sativa</i>	<i>OsGASA1</i>	Os03t0607200-01	282	22816933..22817774	3	-
		<i>OsGASA2</i>	Os03t0760800-01	279	31464821..31465623	3	-
		<i>OsGASA3</i>	Os04t0465300-01	315	23242302..23243199	4	-
		<i>OsGASA4</i>	Os05t0376800-01	456	18183823..18185191	5	-
		<i>OsGASA5</i>	Os05t0432200-00	276	21198105..21198987	5	+
		<i>OsGASA6</i>	Os06t0266800-01	330	8848154..8848833	6	-
		<i>OsGASA7</i>	Os06t0729400-01	399	31082752..31083640	6	+
		<i>OsGASA8</i>	Os07t0592000-01	306	24125333..24127487	7	+
		<i>OsGASA9</i>	Os09t0414900-01	336	14828981..14830515	9	+
		<i>OsGASA10</i>	Os10t0115550-00	213	1005275..1005519	10	+
	<i>Z. maize</i>	<i>ZmGASA1</i>	Zm00001d030236_P001	387	115,306,970...115,307,682	1	+
		<i>ZmGASA2</i>	Zm00001d033195_P001	279	253,430,807...253,431,218	1	-
		<i>ZmGASA3</i>	Zm00001d033196_P001	279	253,431,912...253,432,689	1	-
		<i>ZmGASA4</i>	Zm00001d033197_P001	432	253,446,880...253,447,676	1	-
		<i>ZmGASA5</i>	Zm00001d033201_P001	336	253,760,332...253,761	1	-
		<i>ZmGASA6</i>	Zm00001d033369_P001	300	260,633,725...260,634,703	1	+
		<i>ZmGASA7</i>	Zm00001d001852_T001	429	1,946,288...1,947,091	2	+
		<i>ZmGASA8</i>	Zm00001d003380_T00	345	42,239,268...42,240,320	2	+
		<i>ZmGASA9</i>	Zm00001d006750_T001	402	215,851,864...215,852,764	2	+
		<i>ZmGASA10</i>	Zm00001d013222_T001	288	6,718,595...6,719,557	5	+
		<i>ZmGASA11</i>	Zm00001d014888_P001	600	67,030,219...67,031,109	5	+
		<i>ZmGASA12</i>	Zm00001d038056_P001	435	146,555,470...146,556,601	6	-
		<i>ZmGASA13</i>	Zm00001d010193_P001	279	103,471,796...103,474,708	8	-
		<i>ZmGASA14</i>	Zm00001d045685_P001	441	33,208,139...33,209,281	9	-
		<i>ZmGASA15</i>	Zm00001d023290_T001	387	2,230,214...2,231,452	10	-
	<i>S. bicolor</i>	<i>SbGASA1</i>	Sb01g007360.1	279	6314894...6315749	1	+
		<i>SbGASA2</i>	Sb01g026925.1	381	46068962...46069345	1	-
		<i>SbGASA3</i>	Sb02g024660.1	345	59156656...59157629	2	+
		<i>SbGASA4</i>	Sb06g032730.0	333	60833394...60833857	6	-
		<i>SbGASA5</i>	Sb08g001840.0	384	1868375...1869714	8	-

Table S2. Continued...

Lineages	Species	Gene name	Translation ID	CDS	Sequence position	Chr	Strand
Monocots	<i>S. bicolor</i>	<i>SbGASA6</i>	Sb09g018790.0	411	47123409...47124205	9	-
		<i>SbGASA7</i>	Sb09g019270.0	291	47987621...47988183	9	+
		<i>SbGASA8</i>	Sb09g021340.0	285	50822933...50823885	9	+
		<i>SbGASA9</i>	Sb10g009640.0	342	11093729...11094598	10	-
	<i>B. distachyon</i>	<i>BdGASA1</i>	Bradi1g07890.1	279	5577589...5578375	1	+
		<i>BdGASA2</i>	Bradi1g07900.1	279	5579684...5580382	1	+
		<i>BdGASA3</i>	Bradi1g43850.1	282	41693380...41694186	1	+
		<i>BdGASA4</i>	Bradi2g24320.1	282	22132424...22132909	2	-
		<i>BdGASA5</i>	Bradi4g30250.1	327	35956844...35957987	4	+
	<i>S. italica</i>	<i>SiGASA1</i>	Si033027m	318	30597652...30598541	2	+
		<i>SiGASA2</i>	Si023795m	331	19133565...19134441	3	-
		<i>SiGASA3</i>	Si024382m	399	23044542...23045220	3	+
		<i>SIGASA4</i>	Si012350m	321	34482291...34483149	7	+
		<i>SiGASA5</i>	Si027353m	360	417630...418386	8	-
		<i>SiGASA6</i>	Si027818m	356	426906...427662	8	-
		<i>SiGASA7</i>	Si039591m	225	33792543...33792770	9	-
	<i>S. tuberosum</i>	<i>StGASA1</i>	Sotub01g044530.1.1	267	83829846...83830464	1	+
		<i>StGASA2</i>	Sotub01g048490.1.1	312	87138433...87139159	1	+
		<i>StGASA3</i>	Sotub02g027220.1.1	306	39367647...39368165	2	-
		<i>StGASA4</i>	Sotub02g027230.1.1	309	39374988...39375460	2	-
		<i>StGASA5</i>	Sotub02g033550.1.1	336	45216587...45217307	2	+
		<i>StGASA6</i>	Sotub03g027260.1.1	435	53836215...53838805	3	+
		<i>StGASA7</i>	Sotub03g029520.1.1	312	55936845...55938367	3	+
		<i>StGASA8</i>	Sotub04g014730.1.1	342	11385801...11386719	4	+
		<i>StGASA9</i>	Sotub04g031180.1.1	264	67353590...67354360	4	+
		<i>StGASA10</i>	Sotub05g012380.1.1	537	7173243...7173429	5	+
		<i>StGASA11</i>	Sotub05g010910.1.1	150	5414317...5415765	5	+
		<i>StGASA12</i>	Sotub06g026790.1.1	309	51616097...51616691	6	+
		<i>StGASA13</i>	Sotub08g008240.1.1	267	3479481...3480115	8	+
		<i>StGASA14</i>	Sotub11g008460.1.1	288	1709362...1710314	11	-
		<i>StGASA15</i>	Sotub11g013410.1.1	309	8639998...8641198	11	-

Table S2. Continued...

Lineages	Species	Gene name	Translation ID	CDS	Sequence position	Chr	Strand
Monocots	<i>S. tuberosum</i>	<i>StGASA16</i>	Sotub11g028480.1.1	300	43440602...43440904	11	+
		<i>StGASA17</i>	Sotub12g015160.1.1	645	20949488...20951410	12	+
		<i>StGASA18</i>	Sotub12g021350.1.1	264	49364847...49365647	12	-
		<i>StGASA19</i>	Sotub12g021370.1.1	264	49416269...49416897	12	-
		<i>StGASA20</i>	Sotub12g027480.1.1	336	57090866...57092366	12	-
Eudicots	<i>A. thaliana</i>	<i>AtGASA1</i>	AT1G75750	294	28,441,501...28,442,435	1	-
		<i>AtGASA2</i>	AT4G09610	297	6074770 ...6075645	4	-
		<i>AtGASA3</i>	AT4G09600	297	6072804...6073646	4	-
		<i>AtGASA4</i>	AT5G15230	318	4944816...4946299	5	+
		<i>AtGASA5</i>	AT3G02885	291	637921...639178	3	-
		<i>AtGASA6</i>	AT1G74670	303	28053030...28054149	1	+
		<i>AtGASA7</i>	AT2G14900	324	640413...6405352	2	+
		<i>AtGASA8</i>	AT2G39540	361	16500788...16501414	2	+
		<i>AtGASA9</i>	AT1G22690	357	8027126...8028126	1	+
		<i>AtGASA10</i>	AT5G59845	267	24111304...24112100	5	+
		<i>AtGASA11</i>	AT2G18420	282	7993801...7994554	2	+
		<i>AtGASA12</i>	AT2G30810	318	13127826...13128666	2	-
		<i>AtGASA13</i>	AT5G14920	825	4826141...4828384	5	+
		<i>AtGASA14</i>	AT1G10588	270	3501202...3501904	1	+
		<i>AtGASA15</i>	AT3G10185	366	3145526...3146283	3	+
	<i>C. sativus</i>	<i>CsaGASA1</i>	Csa1G045890	345	5192306 ... 5194012	1	+
		<i>CsaGASA2</i>	Csa2G354910	255	16625133 ... 16625763	2	+
		<i>CsaGASA3</i>	Csa3G872170	291	17758877 ... 17759342	2	-
		<i>CsaGASA4</i>	Csa4G001520	318	20623924 ... 20626301	2	-
		<i>CsaGASA5</i>	Csa6G493840	312	32796840 ... 32797911	3	-
		<i>CsaGASA6</i>	Csa2G368260	693	33173634 ... 33176140	3	-
		<i>CsaGASA7</i>	Csa2G402110	306	36426105 ... 36426899	3	-
		<i>CsaGASA8</i>	Csa3G826660	1200	287038 ... 290725	4	-
		<i>CsaGASA9</i>	Csa3G829190	295	23777116 ... 23777679	6	-

Table S2. Continued...

Lineages	Species	Gene name	Translation ID	CDS	Sequence position	Chr	Strand
<i>Eudicots</i>	<i>S. Lycopersicon</i>	<i>SIGASA1</i>	Solyc01g107370.2.1	267	86698147... 86698940	1	+
		<i>SIGASA2</i>	Solyc01g111080.2.1	645	89226964...89219287	1	-
		<i>SIGASA3</i>	Solyc02g083870.2.1	306	41710589...41710691	2	+
		<i>SIGASA4</i>	Solyc02g083880.2.1	309	41715236...41715339	2	+
		<i>SIGASA5</i>	Solyc02g089350.2.1	336	45776616...45776728	2	+
		<i>SIGASA6</i>	Solyc03g113910.2.1	438	58001941...58002087	3	+
		<i>SIGASA7</i>	Solyc03g116060.2.1	312	59625317...59625213	3	-
		<i>SIGASA8</i>	Solyc04g017720.2.1	342	7793774...7793888	4	+
		<i>SIGASA9</i>	Solyc04g078200.2.1	195	60598399...60598464	4	+
		<i>SIGASA10</i>	Solyc05g006160.2.1	191	847742...847678	5	-
		<i>SIGASA11</i>	Solyc06g007890.2.1	282	1784223...1784317	6	+
		<i>SIGASA12</i>	Solyc06g007910.2.1	282	1795940...1796034	6	+
		<i>SIGASA13</i>	Solyc06g007920.1.1	249	1799096...1799180	6	+
		<i>SIGASA14</i>	Solyc06g069790.2.1	312	39850860...39850756	6	-
		<i>SIGASA15</i>	Solyc08g005620.2.1	267	486408...486497	8	+
		<i>SIGASA16</i>	Solyc11g011210.1.1	288	4264776...4264680	11	-
		<i>SIGASA17</i>	Solyc11g017440.1.1	309	8384626...8384523	11	-
		<i>SIGASA18</i>	Solyc12g042500.1.1	264	43268699...43268611	12	-
		<i>SIGASA19</i>	Solyc12g089300.1.1	336	62879946...62879834	12	-
	<i>C. annuum</i>	<i>CaGASA1</i>	Capang01g000583	180	12745819-12745998	1	+
		<i>CaGASA2</i>	Capang01g005261	177	334144919-334145095	1	-
		<i>CaGASA3</i>	Capang02g002323	306	146299200-146299757	2	+
		<i>CaGASA4</i>	Capang02g002324	183	146314797-146314979	2	+
		<i>CaGASA5</i>	Capang03g000718	183	12172892-12173074	3	+
		<i>CaGASA6</i>	Capang03g000959	435	17127264-17128868	3	+
		<i>CaGASA7</i>	Capang05g000890	267	58494261-58494704	5	+
		<i>CaGASA8</i>	Capang06g001579	183	44135837-44136019	6	-
		<i>CaGASA9</i>	Capang06g002716	339	195819805-195821460	6	+
		<i>CaGASA10</i>	Capang06g002717	177	195827494-195827670	6	+

The data source of each gene was annotated on their corresponding databases and the links are mentioned in Table-1. Chr represents chromosome; Strand represents strand orientation (plus/minus).

	Site Name	Sequence	Function
Hormones	AuxRR-core	GGTCCAT	<i>Cis</i> -acting regulatory element involved in auxin responsiveness
	CGTCA-motif	CGTCA	<i>Cis</i> -acting regulatory element involved in the MeJA responsiveness
	CGTCA-motif	CGTCA	<i>Cis</i> -acting regulatory element involved in the MeJA responsiveness
	TATC-box	TATCCCA	<i>Cis</i> -acting element involved in gibberellin-responsiveness
	P-box	CCTTTTG	gibberellin-responsive element
	P-box	GCCTTTGAGT	gibberellin-responsive element
	TCA-element	TCAGAAGAGG	<i>Cis</i> -acting element involved in salicylic acid responsiveness
	TCA-element	CCATCTTTTT	<i>Cis</i> -acting element involved in salicylic acid responsiveness
	ABRE	CCTACGTGGC	<i>Cis</i> -acting element involved in the abscisic acid responsiveness
	ABRE	TACGTG	<i>Cis</i> -acting element involved in the abscisic acid responsiveness
	Motif 11b	CCGCCGCGCT	abscisic acid responsive element
Stress	G-Box	CACGTG	<i>Cis</i> -acting regulatory element involved in light responsiveness
	G-Box	CACGTG	<i>Cis</i> -acting regulatory element involved in light responsiveness
	GC-motif	CCCCCG	enhancer-like element involved in anoxic specific inducibility
	ARE	TGGTTT	<i>Cis</i> -acting regulatory element essential for the anaerobic induction
	ATGCAAAT motif	ATACAAAT	<i>Cis</i> -acting regulatory element associated to the TGAGTCA motif
	Skn-1_motif	GTCAT	<i>Cis</i> -acting regulatory element required for endosperm expression
	MBS	CAACTG	MYB binding site involved in drought-inducibility
	TATA-Box	TACAAAA	core promoter element around -30 of transcription start
	CAAT-box	CCAAT	common <i>cis</i> -acting element in promoter and enhancer regions
	Sp1	CC(G/A)CCC	light responsive element
	GCN4_motif	TGTGTCA	<i>cis</i> -regulatory element involved in endosperm expression

Table S3 Hormone and stress response motifs identified in the putative promoter sequences. Numbers of *cis*-regulatory elements identified along the putative promoter sequences of each rice GASA gene (PlantCARE) website (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>) (Lescot et al., 2002).

TableS4. The gene-specific primers used for RT-qPCR in this study

Primers used for RT-QPCR expression analysis			
Gene Name	Primer Sequences	Product size (bp)	Exon no
<i>OsGASA-1</i>	Forward:5' CGCTGCTAATGCTTCTCCTC 3'	98	1
	Reverse:5'CACTTGCTCCCGCAGAAC 3'		
<i>OsGASA-2</i>	Forward:5'ATGAAGCTCAACACCACCAC 3'	95	1
	Reverse:5'GCAGAAATCTGATCCAGCCA 3'		
<i>OsGASA-3</i>	Forward:5'CGAGCACTCGCACAAGAAG 3'	102	2
	Reverse:5' GGCAC TTGGTCATGTTGTTG 3'		
<i>OsGASA-4</i>	Forward:5'GAAGCCGTGCATGTTCTTCT 3'	99	1
	Reverse:5' CGCTTGGTCTTCCAATCGTT 3'		
<i>OsGASA-5</i>	Forward:5' AAGTCGCTCTTGTGCATCTC 3'	96	1
	Reverse:5' AATCGATGGCGTAAGCTTGT 3'		
<i>OsGASA-6</i>	Forward:5'GTCCTCCTCCTCTTTCCAT 3'	105	1
	Reverse:5' TCTTG TACTGCGTGTGGAG 3'		
<i>OsGASA-7</i>	Forward:5' AGTCCAAGACCAACAATGGC 3'	129	2
	Reverse:5' GGGAAGAAGAAGAAGCAGCA 3'		
<i>OsGASA-8</i>	Forward:5' ATGAGGGTTCCTCCTCTTCG 3'	88	1
	Reverse:5'CGACGGTGAGATCCTGGAA 3'		
<i>OsGASA-9</i>	Forward:5'CTCCTTGCCATCTCCATAGC 3'	82	1
	Reverse:5' GTCCTCCCTTGCTTACTTGG 3'		
<i>OsGASA-10</i>	Forward:5'TTTGTGGTTGAGGTGTCTGG 3'	112	1
	Reverse:5' CCGTATGTACCGGATGGAAC 3'		

OsActin1 (accession ID KC140126) was chosen as an internal control in rice.

Table S5. Frequency of amino acids in OsGASA proteins

Abbreviation	Amino acid	Amino acid frequency
A	Alanine	0.0939
C	Cysteine	0.114
D	Aspartic acid	0.0273
E	Glutamic acid	0.0246
F	Phenylalanine	0.021
G	Glycine	0.0902
H	Histidine	0.0164
I	Isoleucine	0.0146
K	Lysine	0.0675
L	Leucine	0.0793
M	Methionine	0.031
N	Asparagine	0.041
P	Proline	0.0793
Q	Glutamine	0.0273
R	Arginine	0.051
S	Serine	0.0802
T	Threonine	0.0529
V	Valine	0.051
W	Tryptophan	0.00729
Y	Tyrosine	0.0301

Abbreviation; denotation letter of amino acid, amino acid frequency; the ratio of frequently occurring amino acids. Cysteine occurred most frequently in GASA proteins, which is mentioned in red color. The data were obtained from meme-suite.org/opal.

Table S6. Expression changes of OsGASA genes in different tissues

Tissues	OsGASA1		OsGASA2		OsGASA3		OsGASA4		OsGASA5	
	Log ₂	FC	Log ₂	FC	Log ₂	FC	Log ₂	FC	Log ₂	FC
Rt	0.8338	1.000	0.7849	0.941	0.8663	1.038	0.1364	0.165	0.4114	0.496
St	1.1597	1.390	0.3010	0.361	2.3879	2.874	1.0772	1.265	2.2515	2.725
Lf	1.2298	1.474	2.3879	2.864	0.2986	0.368	3.8399	4.654	1.2792	1.543
FL	0.3549	0.425	2.7342	3.279	2.1890	2.634	3.5008	4.258	1.1576	1.394
Ls-1	1.3117	1.199	0.9818	1.193	1.6896	2.024	0.1375	0.168	1.4761	1.778
Ls-2	1.5271	1.831	1.2265	1.471	0.8647	1.035	1.7369	2.043	0.9351	1.121
Lb-1	0.3548	0.425	1.8999	2.278	3.2064	3.876	2.3219	2.854	1.1460	1.376
Lb-2	0.2341	0.280	1.9953	2.398	2.6601	3.154	0.9686	1.253	1.9716	2.375
Nd-1	1.0749	1.289	0.5069	0.608	0.2750	0.334	0.4719	0.567	0.3126	0.378
Nd-2	0.2132	0.254	0.1505	0.184	1.6896	2.026	0.4419	0.538	0.1505	0.186
Int-1	0.4213	0.505	1.3402	1.634	2.2515	2.743	0.3022	0.364	0.9406	1.145
Int-2	0.5234	0.627	1.4568	1.745	1.2321	1.492	1.2481	1.498	0.3421	0.421
Pre-inf	4.6878	5.625	5.4001	6.547	5.5616	6.674	2.8894	3.475	4.1328	4.965
Post-inf	3.7476	1.192	3.8318	4.595	3.7355	4.543	2.1113	2.543	1.7329	2.034

Tissues	OsGASA6		OsGASA7		OsGASA8		OsGASA9		OsGASA10	
	Log ₂	FC	Log ₂	FC	Log ₂	FC	Log ₂	FC	Log ₂	FC
Rt	3.2064	1.6266	2.4956	1.2660	0.9418	0.4774	2.7909	1.4588	-3.184	-1.616
St	0.9406	0.4771	1.7277	0.8776	0.6968	0.3543	0.2630	0.1335	-3.158	-1.573
Lf	1.9712	1.001	4.0669	2.0742	2.0738	1.0542	1.1440	0.5806	1.6378	0.8402
FL	1.8159	0.9212	2.6827	1.3642	1.7092	1.7093	2.8315	1.4365	0.2986	0.1516
Ls-1	2.2310	1.1276	1.2271	0.6226	0.2515	0.1276	1.9335	0.9811	-2.251	-1.1421
Ls-2	1.6757	0.8546	1.0703	0.5432	1.5563	0.7895	1.8956	0.9754	-2.257	-1.1651
Lb-1	0.6826	0.3462	1.9527	0.9907	2.0588	1.0455	2.1140	1.0786	-2.851	-1.4463
Lb-2	0.8809	0.4474	1.8571	0.9431	1.1844	0.6008	1.2108	0.6345	-1.231	-0.6235
Nd-1	0.4305	0.2184	0.8734	0.4453	2.1844	1.1087	1.9712	1.001	-2.235	-1.1334
Nd-2	1.6393	0.8342	0.9927	0.5043	1.9942	1.0112	2.0738	1.0543	2.6345	1.3367
Int-1	1.2105	0.6140	1.0157	0.5251	1.2142	0.6251	1.0772	0.5467	1.2078	0.6321
Int-2	0.9406	0.4775	1.8155	0.9312	1.5950	0.8091	0.8647	0.3278	0.1531	0.0778
Pre-inf	1.6259	0.8345	4.5558	2.1322	5.1866	2.6453	3.2989	1.6735	2.4936	1.2543
Post-inf	5.6217	2.8519	3.2509	1.6543	3.8308	1.9547	1.7739	0.8978	1.1460	0.5931

Expression analysis of 10 OsGASA genes in 14 different tissues were analyzed by qRT-PCR, and the expression levels were normalized with *OsActin1* (Gene ID: KC140126). Mean values were obtained from three independent PCR amplifications. The analyzed expression values for each GASA gene per tissue were converted into Log₂ ratio. For fold changes, (FC = log₂ final value - log₂ initial value) formula was used, where the initial value of OsGASA gene expression in root as control or base unit and the final value of OsGASA gene expression in various developmental phases. The values obtained from the differences of the OsGASA expression level in roots (log₂-initial value) and the OsGASA expression level in various developmental phases (log₂-final value) were considered as fold changes. The different tissues were denoted as: Rt-root, St-stem, and Lf-leaf. Booting stage:FL, flag leaf; Ls-1, leaf sheath-1; Ls-II, leaf sheath-II; Lb-I, leaf blade-I; Lb-II, Leaf blade-II; N-I,Node-I; N-II, Node-II; Int-I, Internode-I; Int-II, Internode-II. Flowering stage: Pre-Inf, Pre-inflorescences; Post-Inf, Post-inflorescence.