RESEARCH ARTICLE

Spt-Ada-Gcn5-Acetyltransferase (SAGA) Complex in Plants: Genome Wide Identification, Evolutionary Conservation and Functional Determination

Rakesh Srivastava¹,², Krishan Mohan Rai¹, Bindu Pandey¹, Sudhir P. Singh³, Samir V. Sawant¹*

¹ Plant Molecular Biology & Genetic Engineering Laboratory, Council of Scientific and Industrial Research, National Botanical Research Institute (CSIR-NBRI), Lucknow, Uttar Pradesh, India, ² School of Biotechnology, Faculty of Science, Banaras Hindu University, Varanasi, Uttar Pradesh, India, ³ National Agri-Food Biotechnology Institute (DBT), Mohali, Panjab, India

* samirsawant@nbri.res.in

Abstract

The recruitment of RNA polymerase II on a promoter is assisted by the assembly of basal transcriptional machinery in eukaryotes. The Spt-Ada-Gcn5-Acetyltransferase (SAGA) complex plays an important role in transcription regulation in eukaryotes. However, even in the advent of genome sequencing of various plants, SAGA complex has been poorly defined for their components and roles in plant development and physiological functions. Computational analysis of Arabidopsis thaliana and Oryza sativa genomes for SAGA complex resulted in the identification of 17 to 18 potential candidates for SAGA subunits. We have further classified the SAGA complex based on the conserved domains. Phylogenetic analysis revealed that the SAGA complex proteins are evolutionary conserved between plants, yeast and mammals. Functional annotation showed that they participate not only in chromatin remodeling and gene regulation, but also in different biological processes, which could be indirect and possibly mediated via the regulation of gene expression. The in silico expression analysis of the SAGA components in Arabidopsis and O. sativa clearly indicates that its components have a distinct expression profile at different developmental stages. The co-expression analysis of the SAGA components suggests that many of these subunits co-express at different developmental stages, during hormonal interaction and in response to stress conditions. Quantitative real-time PCR analysis of SAGA component genes further confirmed their expression in different plant tissues and stresses. The expression of representative salt, heat and light inducible genes were affected in mutant lines of SAGA subunits in Arabidopsis. Altogether, the present study reveals expedient evidences of involvement of the SAGA complex in plant gene regulation and stress responses.
Introduction

The regulation of gene expression is accomplished by the coordinated action of multiple events to ensure a perfect synchrony of cellular activities from chromatin modification to mRNA formation [1–4]. Gene regulation in eukaryotes requires association of pre-initiation complex (PIC), transcription factors and activators at promoters [1, 5, 6]. One well-known mechanism for transcriptional activation suggests that activator proteins interact with promoter to recruit the components of transcriptional machineries and co-activators such as Transcription Factor II D (TFIID) complex, SAGA and mediator complexes [7–9]. The SAGA complex, a group of multi-protein complex, is important to induce the transcription of a subset of RNA polymerase II-dependent genes [10–12]. Indeed, the SAGA complex is a perfect archetype for multi-subunit histone modifying complexes and co-activator which regulates transcription by RNA polymerase II [13–15]. The first member of the SAGA complex family was isolated in budding yeast Saccharomyces cerevisiae [16]. The 1.8 megadalton S. cerevisiae SAGA complex is composed of 20 conserved proteins and contains different classes of transcriptional co-activator proteins such as SPT (Suppressor of Ty insertions), ADA (alteration/deficiency in activation), GCN5 (general control non-depressive), TAF (TBP-associated factors) proteins and DUBm (deubiquitylation module) [17]. These proteins are organized into different functional and structural sub-modules and thereby executing several cellular functions: nucleosomal histone acetyltransferase (HAT), histone deubiquitinylation, TATA-binding protein (TBP) binding and activator binding [10, 13, 18].

Interestingly, the SAGA complex is engaged in several transcription regulatory processes, for instance, facilitating recruitment of the RNA polymerase II, transcription elongation, promoting nucleosome eviction and replication-coupled nucleosome assembly [15, 19–21]. In addition, the SAGA complex is associated with nuclear export of transcribed mRNA, co-transcriptional spliceosome assembly and transcriptional silencing at telomere region [22–25]. Albeit, extensive evidences about the SAGA complex coding proteins in human, S. cerevisiae and other metazoan species are present, the knowledge about plants SAGA complex still needs to be elucidated. However, functions of few individual proteins of the SAGA complex are reported in plants, which have been shown to be involved in light signaling, stress response and histone modification [26–31].

The present study aims at determining the genes encoding subunits of the SAGA complex across the plant species, mainly in the Arabidopsis thaliana and Oryza sativa, using in silico approaches, and exploring their potential roles in the gene regulation. We have highlighted the functional annotation, co-expression profiles and possible interactome among different proteins of the plant SAGA complex. To better understand the plant SAGA complex, we investigated its roles in regulating the light and stress-induced gene expressions in Arabidopsis.

Materials and Methods

Plant materials and treatment methods

Arabidopsis ecotype Columbia-0 (Col-0) seeds were grown on 0.5 x Murashige and Skoog (MS) medium, kept for 48 hr at 4°C, and then shifted for growth at 20°C ± 1 underneath white light (16 hr light/ 8 hr dark at 100–120 μmol·m⁻²·s⁻¹). For the stress treatments, three-week-old Arabidopsis excised leaves were used. Excised leaves were placed either in 0.5 x MS medium (mock treatment) or in 0.5 x MS medium with 150 mM NaCl solution for 24 hr. For the heat/ high temperature treatment, excised Arabidopsis Col-0 leaves were kept in 0.5 x MS medium and transferred into a 37°C incubator for 2 hr, whereas the control samples were kept at 22°C. The following mutants were used in the present study: gcn5/Salk_030913c; sgf29b
(Salk_128344c); chr5 (Salk_020296c); tralα (Salk_087015c); tafl2β (Salk_132293c); sgf11 (Salk_088988c) and taf13 (Salk_024774c), which were acquired from the Arabidopsis Biological Resource Center [32].

Plant’s genome database search for identification of SAGA complex
National Centre of Biotechnology Information (NCBI); TAIR (The Arabidopsis Information Resource) and RAP (Rice Genome Annotation Project) databases were used for the screening of the SAGA complex in Arabidopsis, O. sativa and other plant genomes. Protein sequences of S. cerevisiae and human SAGA complex components (Table 1, S1 and S2 Tables) were used as queries to execute a BLASTP program against the protein sequences of Arabidopsis and O. sativa.

Alignment and phylogenetic analysis
Clustal-X version 1.83 software program was used for multiple sequence alignment of the protein sequences [33]. The aligned sequences were further used as input to create phylogenetic trees with the Neighbor-Joining method using a Jones-Taylor-Thornton (JTT) model. Bootstrapping was performed, involving 1000 replicates, to represent the evolutionary history of the group analyzed. The evolutionary distance was computed in MEGA 6.06 version [34].

Domain analysis and chromosomal localization
The domain analysis was performed by CDD (Conserved Domain Database) and Pfam (protein families database) with an e-value ≤ 1.0. Chromosome Map Tool database was used to define the position of the SAGA complex genes on Arabidopsis chromosomes [35]. "Paralogous in Arabidopsis" were used for determining the gene duplications and their existence of duplicated segments on chromosome with parameters set to a threshold above 6 per block for paired proteins [36].

Conserved motif analysis
The cis-regulatory elements/motifs were analyzed in 1000 bp upstream from the transcription start site (TSS) by using web based database Plant cis-acting regulatory DNA elements (PLACE) and Plant Cis-Acting Regulatory Elements (PlantCARE) databases and portals [37, 38].

In silico microarray expression and protein interactome analysis
Microarray experiments data from Genevestigator database and analysis toolbox were employed to determine the gene expression profile of Arabidopsis and O. sativa SAGA complex genes in different tissue [39]. The cDNA signatures from Massively Parallel Signature Sequencing (MPSS) were used to count the number of corresponding mRNA molecules produced by each gene of Arabidopsis and O. sativa SAGA complex [40]. A protein-protein interaction network, for the prediction of functional associations within SAGA complex proteins, was prepared using the STRING database with a confidence threshold score of 0.6. [41]. The network was showed in the ‘evidence’ view, whereby lines linking proteins signify the category of evidence used in anticipating the association or interaction.

Functional annotation and co-expression analysis
Functional annotation and Gene Ontology analysis were performed from TAIR and agriGO [35, 42]. Co-expression analysis for gene pairs and co-expressed gene network analysis for each SAGA gene was acquired from ATTED-II (The Arabidopsis trans-factor and cis-element prediction database) version c4.1 [43].
RNA extraction and Real-time PCR analysis

Total RNA was extracted from the flowers, leaves, roots, seedlings, stems and siliques as well as from treated leaves by Sigma’s Spectrum plant total RNA isolation kit. The integrity of RNA, after DNase I treatment, was confirmed by agarose gel electrophoresis. Two microgram of total RNA was used as a template for first-strand cDNA synthesis using the Superscript-II RT kit (Invitrogen). Real-time PCR (qRT-PCR) gene expression analysis was performed and detected by using an ABI’s 7500 Fast Real-time PCR machine [44]. Gene specific forward and reverse primers were designed by using ABI’s-Primer express v2.0 software (S3 Table). The transcripts were normalized using Ubiquitin-10 (Ubq10, At4g05320) transcripts that work as internal control. The relative expression level of target genes was analysed by ΔΔCt method.

Table 1. SAGA complex classification in Arabidopsis and O. Sativa.

| SAGA Subunits | Yeast | Human | Arabidopsis thaliana | Orzya sativa | Functions |
|---------------|-------|-------|----------------------|-------------|-----------|
|               | Name  | Locus | Name                 | Locus       | In plants |
| ADAs          | Ada1  | ADA1  | AtADA1a At2g14850    | OsADA1a Os12g39090 | -         |
|               | Ada1b | ADAs  | AtADA1b At5g67410    | OsADA1b Os3g55450 | -         |
|               | Ada2  | ADA2  | AtADA2b At4g16420    | OsADA2b Os3g53960 | Response to auxin and cytokinin; Pleiotropic effects in development; Abiotic stress [28, 99] |
|               | Ada3  | ADA3  | AtADA3 At4g29790     | OsADA3 Os3g528300 | -         |
|               | Gcn5  | GCN5  | AtGcn5 At3g54610     | OsGCN5 Os10g28040 | HAT activity; Pleiotropic effects in development; Abiotic stress [28, 45, 99] |
| DUBm          | Ubp8  | USP22 | AtUBP22 At5g10790    | OsUBP22 Os4g55360 | -         |
|               | Sgf11 | ATXN7L3 | AtSGF11 At5g58575 | OsSGF11 Os5g28370 | -         |
|               | Sus1  | ENY2  | AtENY2 At3g27100     | OsSUS1 Os1g69110 | -         |
|               | Sgf73 | ATXN7 | AtSGF73 ND           | OsSGF73 ND     | -         |
| SPT           | Spt7  | SPT    | AtSPT7 At1g32750     | OsSPT7 Os6g43790 | -         |
|               | Spt8  | ND    | AtSPT8 ND            | OsSPT8 ND     | -         |
|               | Spt20 (Ada5) | SPT20 | AtSPT20 At1g72390 | OsSPT20 Os1g02860 | Photoperiodic flowering regulation [101, 102] |
| TAFs          | Taf5  | TAF5L | AtTAF5 At5g25150     | OsTAF5 Os6g44030 | Plant viability; Male gametogenesis; Pollen tube development [90] |
|               | Taf6  | TAF6L | AtTAF6 At1g04950     | OsTAF6 Os1g32750 | Plant viability; Pollen tube growth [83] |
|               | Taf9  | TAF9  | AtTAF9 At1g54360     | OsTAF9 Os3g29470 | -         |
|               | Taf9b | TAF9b | AtTAF9b Os7g42150    | OsTAF9b Os3g5450 | -         |
|               | Taf10 | TAF10 | AtTAF10 At4g31720    | OsTAF10 Os9g26180 | Salt tolerance during seed germination [103] |
|               | Taf12 | TAF12 | AtTAF12 At3g10070    | OsTAF12 Os1g63940 | -         |
|               | Chd1  | ND    | AtCHR5 At2g13370     | OsCHD1 OsJ_25446 | Embryo development; Seed maturation [106] |
|               | Sgf29 | STAF36 | AtSGF29a At3g27460 | OsSGF29 Os12g19350 | Flowering initiation, Auxiliary role in salt stress [61] |
|               | Tra1  | TRRAP | AtTRA1a At2g17930    | OsTRA1 Os7g45064 | -         |
|               | Tra1b | ATB1  | At4g36080            | OsTRA1 Os7g45064 | -         |

ND: Not detected.

doi:10.1371/journal.pone.0134709.t001
Results

Identification and classification of SAGA complex subunits in plants

The SAGA complex is a multiple subunit protein complex and is highly conserved among human, S. cerevisiae and Drosophila [13, 17]. The putative SAGA genes were identified in Arabidopsis and O. sativa genomes using protein sequences of S. cerevisiae and human SAGA genes as queries against the protein databases of Arabidopsis and O. sativa (NCBI, TAIR and RAP) (S1 and S2 Tables).

We identified four protein subunits in the ADA group of the SAGA complex, viz. ADA1, ADA2b, ADA3 and GCN5 (ADA4) (Table 1 and S2 Table). ADA2b (At4g16420) and GCN5 (At3g54610) have been previously studied in plants [7, 28, 30, 45, 46]; however, ADA1 and ADA3 proteins are yet to be characterized in plants. Two ADA1 proteins were identified each in Arabidopsis (At2g14850 and At5g67410) and O. sativa (Os12g39090 and Os03g55450) genome as homologs of S. cerevisiae and human ADA1 (Table 1 and S2 Table). Similarly, Arabidopsis (At4g29790) and O. sativa (Os01g73620) ADA3 were identified as homologs of S. cerevisiae and human ADA3 (Table 1 and S2 Table).

S. cerevisiae and human SAGA complex contain four proteins in DUBm group, which mainly participate in the histone deubiquitylation and mRNA export [47], however, we identified three out of four proteins in Arabidopsis and O. sativa (Table 1). Arabidopsis USP22 (At5g10790) and O. sativa UBP22 (Os04g55360) proteins were found as homologs of S. cerevisiae UBP8 and human USP22, respectively (S2 Table). Arabidopsis (At5g58575) and O. sativa (Os05g28370) SGF11 proteins were identified as homologs of human ATXN7L3 and S. cerevisiae SGF11 (Table 1 and S2 Table). Notably, S. cerevisiae SGF11 and human ATXN7L3 (S. cerevisiae SGF11 homolog) share low (15.3%) similarity between their protein sequences [48]. Arabidopsis (At3g27100) and O. sativa (Os01g69110) SUS1 showed significant homology with the corresponding S. cerevisiae SUS1 and human ENY2.

In S. cerevisiae and human, three to four proteins- SPT3, SPT7, SPT8 (not present in humans) and SPT20, have been reported in the SPT group of the SAGA complex (Table 1). Our study identified SPT3 and SPT20 proteins in Arabidopsis and O. sativa. Interestingly, the human SPT3 displays extensive sequence similarity to the histone fold motifs of TAF13 in its N-terminal region [49, 50]. We found conserved domain TAF13 in Arabidopsis At1g02680 and O. sativa Os01g23630 (Table 1 and S2 Table). The SPT20 domain was found to be conserved in Arabidopsis At1g72390 and O. sativa Os01g02860 proteins (Table 1 and S2 Table). In earlier studies, a low level of similarity was reported between SPT3 (30%) and SPT20 (32.5%) homologs of S. cerevisiae and human [51, 52]. The SPT7 protein contains Bromo-domain, a motif found in several transcription factors and co-activators, which is responsible for the acetylation of histones and transcriptional activation [53–55]. In Arabidopsis, 29 Bromo-domain-containing proteins are reported [56]. The BLAST analysis suggested that Arabidopsis At1g32750 (e-value 3e-07) and O. sativa Os06g43790/Os02g38980 (e-value 2e-07/1e-07 and protein similarity 29/25%, respectively) have the highest protein sequence similarity to S. cerevisiae SPT7 and particularly to its Bromo-domain region. However, human STAF65/SUPT7L (homolog of yeast Spt7) BLAST analysis revealed extremely low protein similarity and insignificant e-value of the search Spt7 homolog in Arabidopsis and rice genome. SPT8 protein of S. cerevisiae contains WD40 domain repeats and facilitates TBP interaction [8]. Arabidopsis and other plants encompass more than 200 putative WD40 domain containing proteins [57]. Arabidopsis At5g08390 and At5g23430 displayed protein similarity with corresponding S. cerevisiae SPT8. However, in plant genome, a large number of plant proteins comprising either Bromo-domain or WD40 domain, exhibited a substantial level of similarity with the Bromo-domain for SPT7 and the WD40 domain for SPT8, henceforth, further

Plants SAGA Complex
biochemical evidence is required to validate these subunits of the SAGA complex in the two plant species, *Arabidopsis* and *O. sativa*.

Interestingly, several TAFs subunits are shared by several complexes like TFIID, SAGA, SLIK (SAGA-like complex), and STAGA (SAGA altered, SPT8 absent) as earlier reported in *S. cerevisiae* and human [58]. Lago et al., 2004 explained about different TAFs and their conserved domain structures in *Arabidopsis* [59]. The TAF proteins in the SAGA complex include TAF5, TAF6, TAF9, TAF10 and TAF12. However, our genome-wide similarity search analysis identified two candidate proteins representing TAF12 in *O. sativa* (Table 1), unlike only one protein reported previously [59].

Apart from these four groups, some other components also present in the SAGA complex, such as CHD1 (chromo-domain helicase DNA binding protein 1), TRA1 (Transcription-associated protein 1) and SGF29 (SAGA-associated factor 29) (Table 1). The CHD subfamily-I *S. cerevisiae* CHD1 and human CHD2, share 45% protein similarity (S2 Table) [60]. BLAST searches identified *Arabidopsis* CHR5 (At2g13370) and *O. sativa* CHD (Os12g19350) as homologs of *S. cerevisiae* CHD1 and human CHD2 (Table 1 and S2 Table). Further, we also identified two proteins, At3g27460 and At5g40550 in *Arabidopsis* encoding SGF29, as reported recently [61] and one protein in *O. sativa* (Os12g19350) (Table 1 and S2 Table). TRA1 is a representative of a group of proteins that include DNA-dependent protein kinase catalytic subunit, ATM (Ataxia telangiectasia mutated) and TRRAP (transformation/transcription domain-associated protein), with the carboxyl-terminal regions related to phosphatidylinositol 3-kinases [62]. We identified two TRA1 protein orthologs in *Arabidopsis* (At2g17930 and At4g36080) and one in *O. sativa* (Os07g45064) with the corresponding *S. cerevisiae* TRA1 and human TRRAP (Table 1 and S2 Table). In some reports, RTG2 protein has been considered as a subunit of the SAGA complex [47], whereas sometimes it has been suggested as a variant of the SAGA complex, SLIK [1, 17, 63]. Further biochemical evidences are required to validate the presence of RTG2 in *Arabidopsis* and *O. sativa* SAGA/SLIK complex.

**Conserved domains in plant SAGA complex**

The protein domains of *Arabidopsis* and *O. sativa* SAGA subunits, identified with the corresponding domains of *S. cerevisiae* and human SAGA subunits, is presented in Fig 1 and S1 Fig. Intriguingly, numerous known structural features of protein domains in *Arabidopsis* and *O. sativa* SAGA complexes are common with *S. cerevisiae* and human SAGA complexes. These include HAT module, WD repeat domain, histone fold domains, DUBm, interaction and structural integrity protein domains. The protein similarity of each domain of the SAGA complex shows evolutionary conservation across the species (Table 2). The domains of plant SAGA components share moderate (30–50%) to the high (50% and above) similarity with their counterparts in *S. cerevisiae* and human excluding FAT-domain and chromo-domain (Table 2 and S1 Fig).

**Phylogenetic and evolutionary analysis of the SAGA complex family among different organisms**

To investigate the evolutionary association among *Arabidopsis* and *O. sativa* SAGA complex proteins, phylogenetic trees were made from the alignments of their full-length protein sequences together with SAGA complex proteins of mammals (*Homo sapiens, Mus musculus* and *Rattus norvegicus*), an arthropod (*Drosophila melanogaster*), fungi (*S. cerevisiae* and *Schizosaccharomyces pombe*), dicot plants (*A. thaliana, A. lyrata, Glycine max, Medicago truncatula, Populus trichocarpa, Ricinus communis, and Vitis vinifera*), monocot plants (*Brachypodium distachyon, O. sativa, Sorghum bicolor and Zea mays*), a tracheophyte (*Selaginella lepidophylla*), a bryophyte (*Physcomitrella patens*) and algae (*Chlamydomonas reinhardtii*).
Fig 1. Domain organization of representative SAGA complex proteins in *Arabidopsis* and *Oryza sativa*. The positions of conserved domains which are typical for SAGA are shown. Domain abbreviations are: BR: Bromodomain; NAT_SF: N-Acytransferase superfamily; ZZ_ADA: ZZ-type Zinc finger; SANT: SWI3, ADA2, N-CoR and TFIIIB DNA-binding domains; SAGA-TAD1: Transcriptional regulator of RNA pol II, SAGA, subunit; ADA3: ADA3 superfamily Histone acetyltransferases subunit 3; TRRAP: Transformation/Transcription domain-Associated Protein (TRRAP), pseudokinase domain; FAT: FRAP, ATM and TRRAP domain; FATc: FRAP, ATM, TRRAP C-terminal; TAF5-NTD2: N-terminal region of TATA Binding Protein (TBP) Associated Factor 5; WD-40: Trp-Asp (W-D) dipeptide 40 amino acid motifs; TFIIID-18: Transcription factor II D 18kDa subunit; PEPTIDASE C19D: Peptidase C19 contains ubiquitinyl hydrolases; Zn-UB: Zn-finger in ubiquitinyl-hydrolases; ENy2: enhancer of yellow; DEXDc: DEAD-like helicases superfamily; CHROMO: Chromatin organization modifier (chromo) domain; HELICc: Helicase superfamily c-terminal domain; SGF11: SaGa associated Factor 11; SWIRM: SWI3, RSC8 and MOIRA. The number represents the amino acids in domain and protein.

doi:10.1371/journal.pone.0134709.g001
In order to evaluate the molecular evolutionary relationship and conservation among SAGA protein components in different organisms, we aligned the different SAGA subunits and constructed a phylogenetic tree for each group. The phylogenetic tree analysis inferred immense conservation among the SAGA protein domains in *S. cerevisiae*, mammals, *Arabidopsis*, *O. sativa*, algae, bryophyte and *Drosophila* (Figs 2 and 3; Table 3 and S2 Fig). In the case of ADA group, three clades were exhibited for each SAGA subunit. The first and second clades comprised GCN5 (ADA5) and ADA1 proteins, respectively, while the third clade further divided into subgroups - ADA3 and ADA2b (Fig 2). In the phylogenetic

| SAGA Subunits | Domain | Arabidopsis thaliana | Human | S. cerevisiae | O. sativa |
|---------------|--------|---------------------|-------|--------------|----------|
| ADA1a         | ADA1a  | 33 40               | 36    | 37           | 36       |
| ADA1b         | ADA1b  | 35 39               | 36    | 41           | -        |
| ADA2b         | ZZ_ADA2 | 47 64               | 57    | 69           | 51       |
|               | SANT   | 69 68               | 64    | 70           | 73       |
|               | SWIRM  | 63 60               | 58    | 54           | 50       |
| ADA3          | ADA3   | 45 46               | 39    | 44           | 45       |
| GCN5(ADA4)    | BROMO  | 72 63               | 72    | 59           | 63       |
|               | NAT_SF | 71 80               | 72    | 81           | 74       |
| SPT3          | TAF13  | 57 60               | 50    | 50           | 63       |
| SPT20(ADA5)   | SPT20  | 48 33               | 42    | 34           | 34       |
| TAF5          | TAF5_NTD2 | 62 59              | 62    | 63           | 63       |
|               | WD40   | 50 51               | 49    | 51           | 66       |
| TAF6          | TAF6   | 62 58               | 61    | 60           | 57       |
| TAF6b         | TAF6   | 57 54               | 80    | -            | -        |
| TAF9          | TAF9   | 73 64               | 62    | 53           | 60       |
| TAF9b         | TAF9   | -                  | 68    | 63           | 54       | -        |
| TAF10         | TAF10  | 69 55               | 66    | 56           | 53       |
| TAF12         | TAF12  | 54 53               | 83    | 66           | 69       |
| TAF12b        | TAF12  | 76 68               | 82    | 64           | -        |
| TRA1a         | FAT    | 56 57               | 20    | 17           | 59       |
|               | TRRAP  | 59 58               | 56    | 57           | 57       |
|               | FATc   | -                  | 94    | -            | 52       |
| TRA1b         | FAT    | 55 48               | -     | -            | -        |
|               | TRRAP  | 58 57               | 91    | -            | -        |
|               | FATc   | -                  | 97    | -            | -        |
| SGF29a        | SGF29  | 49 48               | 52    | 47           | 36       |
| SGF29b        | SGF29  | 48 47               | 84    | -            | -        |
| UBP8          | PEPTIDASE-C19 | 60 49         | 60    | 47           | 52       |
| SGT11         | SGT11  | 70 76               | 69    | 73           | 61       |
| SUS1          | ENY2   | 76 76               | 81    | 81           | 100      |
| CHD1/CHR5     | CHROMO | 19/35 18/34 72      | 21/33 | 14/32 18/16 |
|               | DEXDc  | 78 70               | 82    | 75           | 70       |
|               | HELICc | 79 87               | 91    | 77           | 73       |

*a*For *Arabidopsis* first protein then second protein domain similarity percentage with *O. sativa* first protein and second protein domain given.

*b*For two domains of Chromo present in *S. cerevisiae* and human.

Table 2. Domain similarities in SAGA complex protein among Arabidopsis, human, O. sativa and S. cerevisiae.

doi:10.1371/journal.pone.0134709.t002

and *Ostreococcus lucimarinus*). In order to evaluate the molecular evolutionary relationship and conservation among SAGA protein components in different organisms, we aligned the different SAGA subunits and constructed a phylogenetic tree for each group. The phylogenetic tree analysis inferred immense conservation among the SAGA protein domains in *S. cerevisiae*, mammals, *Arabidopsis*, *O. sativa*, algae, bryophyte and *Drosophila* (Figs 2 and 3; Table 3 and S2 Fig). In the case of ADA group, three clades were exhibited for each SAGA subunit. The first and second clades comprised GCN5 (ADA5) and ADA1 proteins, respectively, while the third clade further divided into subgroups - ADA3 and ADA2b (Fig 2). In the phylogenetic
analysis of ADA proteins from various organisms fall in a similar clade, excluding SpADA3 and DmADA3, which were close to the ADA1 clade (Fig 2). Similar to ADA group, other groups made several clades based on their similar protein domain specific phylogenetic tree analyses (Fig 3). In the phylogenetic tree of TAFs group, CrTAF5 and OlTAF12 proteins were present in different clades (S2 Fig). The phylogenetic tree constructed from plant SAGA proteins revealed that these proteins diverge into monocots and dicots (Figs 2 and 3). Based on the phylogenetic tree analysis, most protein domains in the SAGA subunits were remained extremely conserved in *S. cerevisiae*, mammals, *Arabidopsis*, *O. sativa*, algae, lycopsida, bryophyta and *Drosophila* during the course of evolution (Table 3).

The phylogenetic trees were also constructed using the representative domain sequences of each protein of the SAGA complex of *Arabidopsis*, *D. melanogaster*, mammals (*H. sapiens* and...
Fig 3. Phylogenetic relationship of DUBm, SPTs and other subunits of the SAGA complex. Different subunits of the SAGA complex protein sequences were used from At, Arabidopsis thaliana (red circle); Dm, Drosophila melanogaster; Hs, Homo sapiens (red square); Os, Oryza sativa (red diamond shape); Mm, Mus musculus; Rs, Rattus norvegicus; Sc, Saccharomyces cerevisiae (red triangle); Sp, Schizosaccharomyces pombe; Zm, Zea mays; Rc, Ricinus communis; Pt, Populus trichocarpa; Vv, Vitis vinifera; Al, Arabidopsis lyrata; Mt, Medicago truncatula; Bd, Brachypodium distachyon; Sb, Sorgum bicolor; Sm, Selaginella moellendorffii; Pp, Physcomitrella patens; Cr, Chlamydomonas reinhardtii and Ol, Ostreococcus lucimarinus. Phylogeny reconstruction was analysed by neighbour-joining statistical method. Test of phylogeny was analysed by the bootstrap method (1,000 replicates). (A) DUBm protein; (B) SPTs protein; (C) CHDs protein; (D) SGF29 protein; (E) TRA1 protein.

doi:10.1371/journal.pone.0134709.g003
M. musculus), O. sativa and S. cerevisiae (S3 Fig). The analysis of phylogenetic tree from different domain of ADA protein groups of the SAGA complex showed two clades, the first clade comprised of ADA3, GCN5, SWRIM-ADA2b and SANT-ADA2b that represents HAT modules of the SAGA complex. However, the ZZ-ADA2b domain, which is involved in interaction with GCN5, presented with ADA1 domains in the second clade (S3A Fig). The phylogenetic tree analysis of full length protein sequences indicated that ADA1 forms a different group from other ADA protein groups (Fig 2). The apparent reason behind the presence of two ZZ-ADA2b and ADA1 domains in one group might be that both protein domains are involved in protein-protein interactions. The phylogenetic tree constructed from each domain of the DUBm SAGA complex subunits suggested that SGF11 and Peptidase C19D-UBP domains were present in the same clade (S3C Fig), which is against the result obtained in the phylogenetic tree with full length protein (Fig 3A). TAFs domain grouped according to their domain features, for example, histone-fold domain containing TAF6, TAF9 and TAF12 domain were present in the same clade, in which TAF6 and TAF9 contains similar histone-fold domain (S3G Fig) [59]. TAF5 and TAF10 proteins were present in the same sub-group (S2 Fig) and their domains (NTD-TAF5 and TAF10) also showed a close relation (S3G Fig).

Notably, several paralogs were found for SAGA complex components in selected plants, mainly in G. max and P. trichocarpa (Table 3). Moreover, a variation was observed in the total number of SAGA complex components among dicots, as compared to monocots (Table 3). The variation in the number of the SAGA complex subunits suggested that these components could have been executed to accomplish the distinct and specialized roles in plants.

**Chromosomal distribution and functional annotation of plant SAGA complex**

The Arabidopsis Genome Initiative provides the opportunity to identify the instances of chromosomal block duplication in the genome [64]. We intended to investigate, whether proteins encoding for the SAGA complex are associated with chromosomal block duplication in Arabidopsis. We used TAIR chromosome map viewer and Paralogons in Arabidopsis for the localization of the SAGA components across the five chromosomes (S4 Fig). Most of these SAGA complex proteins were in the duplicated segmental regions of Arabidopsis chromosome [36]. Moreover, we also identified that some of the Arabidopsis and O. sativa SAGA subunits were found in more than one copy such as ADA1, TAF6, TAF9, SFG29 and TRA1. Thus, it seems that some of the SAGA proteins were duplicated during evolution.

The functional characterization analysis showed that Arabidopsis and O. sativa SAGA complex components play a key role in gene expression, transcription initiation, complex assembly and several metabolic and cellular processes (Fig 4). Gene Ontology predicted that plant SAGA complex components also participate in a transcription regulator activity, binding, catalytic activity as well as in the development of cell and organelle parts (S4 Table). Recent studies suggest participation of some of the plant SAGA complex subunits, for example Arabidopsis ADA2B, SFG29 and GCN5, in the light- [29], cold- [28, 65] and salt-induced [61] gene expression, flower development [66], histone acetylation [30, 45]. The functional characterization analysis also indicated that their involvement in auxin, cytokinin, ethylene and jasmonic acid mediated signaling pathways (S4 Table). In sum, functional and GO analysis predicated the involvement of the plant SAGA complex not only in chromatin remodeling, but also in abiotic and biotic processes.

**Protein-protein interactome analysis of Arabidopsis SAGA complex**

To examine interactions among Arabidopsis SAGA complex components, we mapped the SAGA proteins over STRING interactome, a database of known and predicted protein
Table 3. Putative SAGA complex genes in higher and lower plant organisms.

| Gene Symbol | Dicots | Monocots | Lycopsida | Bryophyte | Green alga |
|-------------|--------|----------|-----------|-----------|------------|
|             |        |          |           |           |            |
| SPT3        | XP_002889412 | XP_002859431 | XP_00230312 | XP_002275398 | XP_002515305 |
|             | XP_003597970 | -         | -         | -         | -          |
| SPT20       | XP_003289424 | XP_002887433 | XP_003627948 | XP_00294116 | XP_00272317 |
|             | XP_003291956 | -         | -         | -         | -          |
| ADA1        | XP_003598846 | XP_002898652 | XP_00230802 | XP_00276962 | XP_00257493 |
|             | XP_003598928 | XP_002877664 | XP_003689860 | XP_00230306 | XP_00250562 |
|             | XP_003582523 | -         | -         | -         | -          |
| ADA2b       | XP_003582030 | XP_002872330 | XP_00331901 | XP_002653494 | XP_002515336 |
|             | XP_003585686 | XP_002869325 | XP_00331900 | -         | -          |
|             | XP_003564525 | XP_002869187 | XP_003046832 | -         | -          |
| ADA3        | XP_003502678 | XP_002869409 | XP_00311946 | XP_00265703 | XP_002505305 |
|             | XP_003391858 | -         | -         | -         | -          |
| GCC5        | XP_003525080 | XP_002876352 | XP_00336892 | XP_00305812 | XP_00275146 |
|             | XP_003543477 | -         | -         | -         | -          |
| UBPI2       | XP_003530212 | XP_002871411 | XP_00230685 | XP_00268376 | XP_002515408 |
|             | XP_003546482 | XP_003069464 | XP_002334922 | -         | -          |
|             | XP_003549730 | -         | -         | -         | -          |
| SGFI1       | XP_003419022 | XP_002864591 | XP_00354973 | XP_003526995 | XP_003632167 |
|             | XP_003425630 | -         | -         | -         | -          |
| ENY2        | XP_003549747 | XP_002870718 | XP_00350888 | XP_00269535 | XP_00250517 |
|             | XP_003546339 | -         | -         | -         | -          |
| CHR5        | XP_003519517 | XP_002885672 | XP_003617298 | XP_00231358 | XP_00275100 |
|             | XP_003503602 | XP_003060162 | -         | -         | -          |
| TAF5        | XP_003061821 | XP_002871411 | XP_003036301 | XP_00360762 | XP_003518453 |
|             | XP_003523495 | -         | -         | -         | -          |
|             | XP_003543936 | -         | -         | -         | -          |
| TAF6        | XP_003516646 | XP_002882254 | XP_003021904 | XP_00230500 | XP_00228804 |
|             | XP_003517377 | XP_003000186 | XP_00289948 | XP_00276969 | XP_002305290 |
|             | XP_003562877 | XP_002881385 | -         | -         | -          |
| TAF9        | XP_003538385 | XP_002894479 | XP_003508543 | XP_00268587 | XP_002378391 |
|             | XP_003586666 | XP_00301803 | XP_003365985 | -         | -          |
| TAF10       | XP_003549834 | XP_002887280 | XP_003434302 | XP_00236715 | XP_002513799 |
|             | XP_003524747 | -         | -         | -         | -          |
| TAF12       | XP_003550294 | XP_002887775 | XP_003507615 | XP_002306150 | XP_00267105 |
|             | XP_003543789 | -         | -         | -         | -          |
| TAL1        | XP_003517177 | XP_002870266 | XP_003612164 | XP_002327756 | XP_00361895 |
|             | XP_003573630 | XP_002886517 | -         | -         | -          |
| SFG29       | XP_003577008 | XP_002870694 | ND         | XP_002363060 | XP_002524900 |
|             | XP_003574048 | XP_002877265 | -         | -         | -          |
|             | XP_003574072 | -         | -         | -         | -          |

*Gm: Glycine max; Al: Arabidopsis lyrata; Mt: Medicago truncatula; Pt: Populus trichocarpa; Vv: Vitis vinifera; Rc: Ricinus communis; Bd: Brachypodium distachyon; Sb: Sorgum bicolor; Zm: Zea mays; Sm: Selaginella moellendorffii; Pp: Physcomitrella patens; Cr, Chlamydomonas reinhardtii and Ol, Ostreococcus lucimarinus.
ND: Not detected.

doi:10.1371/journal.pone.0134709.s003
interactions [41]. The analysis of Arabidopsis SAGA component proteins revealed an interconnected sub-network of 131-hub proteins (confidence score 0.6, Fig 5 and S5 Table). These analyses suggested that many hub proteins create a network which behaves as a functional module within the complex. Moreover, the protein-protein interaction analysis of S. cerevisiae SAGA proteins using the STRING database displayed 190-protein interactions with a confidence score of 0.6 (S6 Table). Interestingly, most of these protein-protein interactions were similar in the SAGA proteins of Arabidopsis and S. cerevisiae (S5 and S6 Tables). The mutation and biochemical characterization studies in S. cerevisiae and mammals established that these interactions are essential for SAGA structure and its stability. For instance, any alteration in SPT7, SPT20, TAF5, TAF10, or TAF12 affects the SAGA composition and integrity [67–69].

Fig 4. Functional annotation of plant SAGA complex components. (A) Arabidopsis (B) O. sativa. Functional annotations were performed by TAIR and agriGO databases. The percentage (%) associated with each annotation indicates the percentage of segments annotated to that category.

doi:10.1371/journal.pone.0134709.g004
In silico expression analysis of the SAGA complex encoding genes

Gene expression profiles of the SAGA complex components can provide significant evidences for their potential functional roles. The functional annotation of SAGA components in Arabidopsis and O. sativa revealed their diverse roles in plant development (Fig 4). To further validate, we examined the expression profile of the SAGA complex components in different tissues using Genevestigator microarray database and its expression meta-analysis tool [54], and MPSS database [55]. The expression profile of the SAGA complex encoding genes was examined in 9 different plant organs of Arabidopsis and O. sativa (Fig 6). AtTaf10, AtGcn5 and AtChr5 were expressed at low levels in all the examined developmental stages. However, the expression of AtAda1a, AtTaf12b, AtTaf6b and AtTra1a was higher in the aforesaid developmental stages. In the case of O. sativa, SAGA subunit genes were found highly expressed in booting, seedling, milk, flower and stem elongation stages (Fig 6B). During germination, transcript accumulation was observed at higher levels for AtTaf6/6b, AtTra1a, AtAda2b, AtTaf1a, AtTaf12b and AtUbp22 genes in Arabidopsis, whereas, for OsAda2b, OsAda3, OsTaf5, OsTaf12/12b, OsTaf1a, OsUbp22, OsSus1, OsSgf11 and OsTaf9 genes in O. sativa. In the booting stage of O. sativa, OsTaf5, OsTaf13, OsTaf12/12b, OsGcn5, OsAda2b, OsAda1a, OsSgf29, OsAda3, and OsSgf11 were among the highly expressed genes, whereas, OsSus1, OsSgf29, OsUbp22, OsTaf6 and OsSgf11 were the genes that highly expressed during dough developmental stages. The meta-analysis displayed an enhanced expression of SAGA component genes in the endosperm (micropylar, peripheral and chalazal), seed coat, suspensor callus and primary cells of Arabidopsis (Fig 7A), whereas in callus, sperm cells, panicle, leaf, pistil, stigma, ovary and root tip of O. sativa (Fig 7B). The results suggest a diverse role of SAGA component genes being expressed throughout different developmental phases in distinct plant organs and tissues.
Data was extracted from the MPSS database library (17 and 20 bases), representing 12 and 13 different anatomical parts of *Arabidopsis* and *O. sativa*, respectively. These signatures uniquely recognize specific gene, which show a perfect match (100% identity over the tag length), and signify a quantitative estimation of expression of that gene. These MPSS tags further confirmed transcript abundance of SAGA protein encoding genes in different plant parts (S7 and S8 Tables). Transcript differences are generally presented by the total number of tags (TPM, transcripts per million), low expression if smaller than 25 TPM, moderate expression if 26 to 250 TPM, while highly expressed in case of >250 TPM. Based on these signatures/tags, five *Arabidopsis* genes viz., *AtAda3*, *AtAda1b*, *AtTaf12*, *AtTra1a* and *AtSgf29* were expressed at low levels, whereas *AtTaf10* expressed at a higher level in leaf, root, siliques and callus (S7 Table). Other *Arabidopsis* SAGA genes exhibited a moderate level of transcript accumulation. MPSS analysis in *O. sativa* showed that *OsAda2b*, *OsTaf10* and *OsTra1* expressed at higher levels (>250 TPM). The maximum transcript abundance was observed for *OsAda2b* in mature leaves and for *OsTaf10* in young leaves, ovary and mature stigma and callus, whereas *OsTra1* was significantly expressed in most of the plant parts, except germinating seed and stem. The SAGA genes, *OsGcn5*, *OsAda1a*, *OsSpt20*, *OsSgf11*, *OsTaf9b* and *OsTaf12* expressed at low levels, whereas others at moderate levels (S8 Table).

**Co-expression analysis for gene pairs and gene network analysis of the SAGA complex**

The expression profiles of SAGA components in *Arabidopsis* and *O. sativa* using Genevestigator and MPSS revealed that many of the components have distinct tissue-specific expressions.
Fig 7. **In silico** expression patterns of the SAGA complex genes in anatomical tissues. The expression patterns in different anatomical tissues for SAGA genes were obtained from the Genevestigator microarray database tool. (A) Arabidopsis anatomical tissues; (B) O. sativa anatomical tissues.

doi:10.1371/journal.pone.0134709.g007
### Table 4. Co-expression analysis of SAGA component genes in *Arabidopsis*.

| S.N. | LOCUS1 | LOCUS2 | Mutual Rank (MR)* |
|------|--------|--------|-------------------|
| 1    | At1g04950 | TAF6   | At4g16420 ADA2B   |
| 2    | At1g02680 | TAF13  | At4g16420 ADA2B   |
| 3    | At1g72390 | SPT20  | At4g16420 ADA2B   |
| 4    | At5g10790 | UBP22  | At5g40550 SGF29B  |
| 5    | At1g04950 | TAF6   | At2g17930 TRA1A   |
| 6    | At1g02680 | TAF13  | At2g17930 TRA1A   |
| 7    | At1g72390 | SPT20  | At4g29790 ADA3    |
| 8    | At1g32750 | TAF1B  | At4g16420 ADA2B   |
| 9    | At1g17440 | TAF12B | At4g29790 ADA3    |
| 10   | At1g02680 | TAF13  | At1g72390 SPT20   |
| 11   | At1g72390 | SPT20  | At2g17930 TRA1A   |
| 12   | At1g02680 | TAF13  | At1g72390 SPT20   |
| 13   | At1g17440 | TAF12B | At2g13370 CHR5    |
| 14   | At1g02680 | TAF13  | At1g54140 TAF9    |
| 15   | At1g02680 | TAF13  | At1g04950 TAF6    |
| 16   | At1g72390 | SPT20  | At2g17930 TRA1A   |
| 17   | At1g02680 | TAF13  | At1g72390 SPT20   |
| 18   | At1g32750 | TAF1B  | At4g16420 ADA2B   |
| 19   | At1g02680 | TAF13  | At1g72390 SPT20   |
| 20   | At1g72390 | SPT20  | At2g17930 TRA1A   |
| 21   | At1g02680 | TAF13  | At1g54140 TAF9    |
| 22   | At1g02680 | TAF13  | At1g72390 SPT20   |
| 23   | At1g72390 | SPT20  | At2g17930 TRA1A   |
| 24   | At1g02680 | TAF13  | At1g72390 SPT20   |
| 25   | At1g02680 | TAF13  | At1g72390 SPT20   |
| 26   | At1g02680 | TAF13  | At1g72390 SPT20   |
| 27   | At1g17440 | TAF12B | At2g13370 CHR5    |
| 28   | At1g32750 | TAF1B  | At4g16420 ADA2B   |
| 29   | At1g02680 | TAF13  | At1g72390 SPT20   |
| 30   | At1g32750 | TAF1B  | At4g16420 ADA2B   |
| 31   | At1g32750 | TAF1B  | At4g16420 ADA2B   |
| 32   | At1g02680 | TAF13  | At1g72390 SPT20   |
| 33   | At1g02680 | TAF13  | At1g72390 SPT20   |
| 34   | At1g02680 | TAF13  | At1g72390 SPT20   |
| 35   | At1g02680 | TAF13  | At1g72390 SPT20   |
| 36   | At1g02680 | TAF13  | At1g72390 SPT20   |
| 37   | At1g02680 | TAF13  | At1g72390 SPT20   |
| 38   | At1g02680 | TAF13  | At1g72390 SPT20   |
| 39   | At1g02680 | TAF13  | At1g72390 SPT20   |
| 40   | At1g02680 | TAF13  | At1g72390 SPT20   |

*MR values represent here <1000 shows significant co-expression data (in bold)* [43]

doi:10.1371/journal.pone.0134709.t004
We further examined whether these genes co-express during plant development or in any other physiological condition. The co-expression data for each SAGA component gene pairs were generated from ATTED-II database, which includes 1388 microarray experimental data [43]. The strength of co-expression for the interconnecting gene pairs was determined by Mutual Rank (MR) process using these microarray data. Forty-two significant co-expression patterns (Table 4) were obtained between SAGA components from 171 co-expressing gene pairs (S9 Table). These co-expression patterns were identified under different biotic, abiotic, hormonal and tissue conditions, for example, co-expression analysis of gene pairs data showed that Ada2b was strongly co-expressed with Taf6, Taf13 and Spt20 genes in all the developmental and environmental stress conditions. Likewise, Ubp22 co-expressed with Sgf29b and Taf1b with Tra6 and Taf13, at high MR values. The significant MR values for Taf1b, Spt20, Tra1, Taf9, Gcn5, Ada2b and Ada3 suggest their co-expression at the tissue level. The genes, Spt20, Ada2b, Taf12b, Taf6, Taf10 showed co-expression in abiotic stress conditions (Table 4). Under hormonal condition, Ada2b, Ada3, Chr5, Taf6, Taf10, Taf12b, Taf13 and Taf13.
Tra1α exhibited a substantial level of co-expression strength, whereas Spt20 was found to be co-expressed with Chr5 in biotic stress condition (Table 4).

The co-expressed gene network analysis was done to identify the genes, which co-regulate with the SAGA complex (S5 Fig). Co-expressed gene network provides the evidence of highly interconnected expression modules of a subset of genes, which additionally show another layer of regulation, and consequently the complementary evidences to understand gene function network. A total of 181 proteins was found to be co-regulated with SAGA complex components (S10 Table). Approximately 36% of 181 proteins are recognized to be involved with regards to abiotic or biotic stimulus or stress, developmental processes, transcription regulation, signal transduction and other biological processes (S6 Fig). This analysis further indicates a potential role of the SAGA complex in regulating plant development and responses to various physiological stresses.

Expression analysis of the SAGA complex subunit during developmental stages and stress conditions

We performed a quantitative gene expression analysis of ten representative SAGA components by QRT-PCR in different Arabidopsis tissues: flowers, mature leaves, siliques, six-day-old seedlings, stems and roots (Fig 8A). The gene expression profile of the SAGA complex components was plotted with reference to the expression of ubiquitin. The genes of the SAGA complex, although expressing at a lower level compared with Ubiquitin, showed consistent expression in almost all the examined plant parts (Fig 8A). These indicate the involvement of the SAGA complex in gene regulation throughout the plant body. However, there were certain components that showed spatial preference, for example, the expression of Spt20 was relatively higher in root and leaf, whereas, Sgf11 in leaves and seedlings than other examined tissues.

The effect of high temperature and salt concentrations was also examined on the expression pattern of the SAGA components in Arabidopsis. The excised leaves of Arabidopsis were either exposed to a high temperature at 37°C for 2 hr or 150 mM NaCl for a period of 24 hr for high salt stress condition, and the gene expression of the SAGA components was compared with their respective controls. The gene expression of the most of selected components of the SAGA complex was induced under elevated salt concentration (Fig 8B) and high temperature (Fig 8C); however, the fold of induction varies for different components. Interestingly, Sgf29b expression was suppressed in salt treatment condition (Fig 8B). Thus, the qRT-PCR results suggested the significance of SAGA components gene expression in plants during abiotic stresses.

As discussed above, in silico co-expression analysis of the SAGA complex subunits suggested that these subunits were co-expressed in the tissues, hormones and stress conditions. Notably, the quantitative gene expression analysis of selected SAGA components further supported the co-expression analysis, such as Spt20 and Chr5, Taf13 and Tra1, Spt20 and Chr5 showed high co-expression with significant MR value in tissue; while Spt20 and Ubp22, Sgf11 and Tra1, Spt20 and Chr5, Gcn5 and Taf6 were co-expressed and considerable MR value in abiotic stress. The qRT-PCR analysis is in agreement with the in silico co-expression profile (Table 4, Fig 8 and S9 Table).

SAGA complex regulates expression of heat, salt and light-induced genes

SAGA complex facilitates the PIC assembly in the core promoter region of yeast and human genes [70–74]. Little is known about how SAGA complex facilitates gene regulation in plants. To address this, RNA was isolated from seven homozygous T-DNA SAGA subunit Arabidopsis mutants and wild type plants, grown under different conditions such as light/dark, high salt or...
heat stress (Fig 9 and S7 Fig). The gene expression of light induced (At1g67090 and At4g02770) [2, 75], salt induced (At2g40140 and At1g56600) [76, 77] and heat induced (At1g71000 and At5g12030) [78] genes was examined in these mutant in comparison to the wild type plants by qRT-PCR (Fig 9A). The expression of both the light activated genes was
considerably reduced in all the mutants, except in sgf11 for both the genes and in gcn5 for At4g02770, in which the relative expression values were not statistically significant (Fig 9B). In the case of salt stress, the expression level of both the salt induced genes declined in mutants as compared to the wild type, except At1g56600 in taf13, which was statistically not significant (Fig 9C). Under heat stress, expression of the heat activated genes was decreased in mutants, except At5g12030 in gcn5 and taf13 and At1g71000 in sgf11 mutant which were not statistically significant (Fig 9D). These results anticipated that SAGA complex plays significant roles in the transcription regulation of stress inducible genes.

Discussion

The SAGA complex has been previously shown to be associated with transcriptional regulation of ~10% RNA polymerase II-dependent S. cerevisiae genes, which contribute in response to DNA damage and other stress conditions such as heat, oxidation, and metabolic starvation [71, 72, 79]. A recent report indicates that SAGA complex regulates all active genes and present at their promoters and transcribed regions [80]. With the computational approach, we identified 18 putative SAGA complex subunits in Arabidopsis and O. sativa. The protein similarities among Arabidopsis and S. cerevisiae SAGA complex subunits are low (17%) to medium (51%), as observed between S. cerevisiae and human SAGA complex (15% to 56%; S2 Table). Since the SAGA complex is involved in the fine-tuning of gene expression, this could be one of the reasons for the poor protein similarities. Our results on in silico expression, GO analysis and qRT-PCR of plant SAGA complex representative genes suggested their role in various cellular, physiological and molecular processes. The previous reports on the functions of ADA2b, GCN5, TAF10, TAF6 and SGF29 in plants are in accordance with our study, suggesting conservation of the SAGA complex throughout evolution [28, 46, 61–83]. Thus, the presence of conserved domain is helpful in identifying most of the putative members of plant SAGA complexes in different plant organism databases. Beside the low level similarity in full protein sequence (S2 Table), most of the domains present in plant SAGA complex encoding genes were found conserved among different organisms (Figs 2 and 3; Table 2). The similarity between conserved domain’s amino acid sequences of Arabidopsis SAGA was observed higher, i.e. from 30% to 97% (Table 2 and S1 Fig). Notably, similar range of similarities was found between the key domains of the SAGA complex in S. cerevisiae and human (Table 2 and S1 Fig). On the basis of protein or conserved domain similarity and phylogenetic analysis, our results altogether suggested that plant SAGA complex was observed to be closer to the human than to the yeast SAGA complex (Figs 2 and 3; Table 2 and S2 Table).

Our analysis of protein alignment, phylogenetic tree and chromosomal distribution suggested that many plant SAGA complex representative genes might have duplicated during evolution (Figs 2 and 3; S2 Fig). For example, Taf6, Taf9, Taf12, Ada1, Tra1 and Sgf29 have been found duplicated in either O. sativa or Arabidopsis. Besides these genes, other SAGA subunit genes are also found duplicated in other lower and higher plant groups (Table 3). This duplication event may also lead to variability in the SAGA complex components in plants like Ada2a-containing (ATAC), SLIK/SALSA or STAGA [74], or sometimes shares subunits with other complexes like TFIID [68]. The protein interactome analysis suggested that Arabidopsis SAGA complex proteins interact with each other and thus further suggested their conservation in plants (Fig 5). The structural integrity of the SAGA complex is dependent on the protein-protein interactions as evident in our study, and also discussed in previous reports; such as TAF10 and TAF12 associate directly via their histone fold domains with SPT7 and ADA1, establishing SPT7-TAF10 and ADA1-TAF12 heterodimer, respectively [84, 85], whereas TAF5 interact with ADA1, ADA3 and SPT7 [69].
Our results suggested that the SAGA complex encoding genes expressed in most of the plant parts and playing an essential role in plant development. Previous reports in *Arabidopsis*, *gcn5* exhibit pleiotropic developmental abnormalities, such as abnormal meristem role, dwarfism, loss of apical dominance, defects in floral organ identity [28, 29, 31, 86–88]. An insertion of T-DNA elements in the *Arabidopsis Ada2b* produces a dwarf phenotype with defects in root and shoot development [28, 87, 89]. *Arabidopsis sgf29a−* shows a little delay in leaf and flower development [61]. Importantly, some reports on plant TAFs (TAF5, 6 and 10) indicated their indispensable role in plant development [81, 83, 90]. Notably, SAGA complex is also critically involved in developmental aspects and is indispensable for viability in metazoan [11]. Recently, ubiquitin protease activity of the SAGA complex showed significant regulation of the expression of the tissues specific genes and developmental processes in *Drosophila* [91]. In *Drosophila*, loss of SAGA subunit functions, such as ADA2b, SGF11 and Nonstop protein (homolog of ENY2), display photoreceptor axon targeting defects, whereas, GCN5 has an essential role in the development of eye and wing disc [92, 93]. While, mice TAF9b and GCN5 are required for the regulation of genes during neuronal and mesoderm development [94, 95]. These accumulating evidences indicate that the functions of the SAGA complex in higher organisms involve more sophisticated mechanisms in regulation of gene expression than unicellular counterpart like *S. cerevisiae*.

SAGA complex expedites the gene expression that anticipates to various environmental cues such as DNA damage and abiotic stress conditions [12]. Many reports, as discussed above, reveal that the SAGA complex is directly or indirectly contributing in various developmental and stress regulated processes, for example, arsenite stress conditions [52] osmotic stress [96] and ultraviolet induced [97]. The yeast SAGA complex also takes part in the up-regulation of several genes during environmental stress, including carbon starved condition [71]. Our results support the stress inducible expression of several SAGA components in *Arabidopsis*. Interestingly, the promoter sequence analysis of the SAGA components revealed several stress responsive cis-motifs (S11 Table), indicating their involvement in transcription regulation activities in response to stress. Nevertheless, further experimentation is needed to validate the involvement of these motifs in the regulation of the SAGA component genes. The expression analysis of the SAGA subunits supports its potential roles in response to environmental cues (Figs 6 and 7). These results are in accordance with the earlier published reports on plant ADA2b, GCN5 and SGF29a [26, 28, 46, 61]. *Arabidopsis ada2b-1−* mutant displays enhanced hypersensitivity to salt and abscisic acid stress than wild-type plants [26, 61]. Although, loss of SGF29a function displays salt stress tolerance, the gene expression level of stress-related genes markers such as COR78 (cold regulated 78), RAB18 (responsive to aba 18), and RD29b (responsive to desiccation 29b) are lower in sgf29a− mutant after 3 hr of NaCl treatment [61]. *Arabidopsis* HAT protein GCN5 and co-activator ADA2b proteins play significant roles in cold responses and loss of functions of these proteins showed a decline in the expression of several cold-regulated genes [27, 28, 98]. Altogether, the property of the SAGA complex in the regulation of stress genes is not only well maintained within plants, but also comparable to *S. cerevisiae* or human [71, 74].

In conclusion, we identified 18 subunits of the SAGA complex in *Arabidopsis* and *O. sativa*. The protein similarities at the level of conserved domain indicate that the SAGA complex is conserved in eukaryotes such as *S. cerevisiae*, plants and mammals. The expression analysis of the SAGA components indicates that the networks of SAGA complex are involved in various biological processes in plants, including development, physiology and response to environmental stresses via gene regulation. This study advances our understanding about SAGA components and their different functions in plants.
Supporting Information

S1 Fig. Domain similarity between human, S. cerevisiae, Arabidopsis and O. sativa SAGA complex encoding genes. Sequence alignments of the representative domains of each protein of the SAGA complex were done by using Clustal X.

S2 Fig. The phylogenetic relationship of the TAFs group of the SAGA complex. SAGA complex subunit protein sequences were used from At, A. thaliana (red circle); Dm, D. melanogaster; Hs, H. sapiens (red square); Os, O. sativa (red diamond shape); Mm, M. musculus; Rn, R. norvegicus; Sc, S. cerevisiae (red triangle); Sp, S. pombe; Zm, Z. mays; Rc, R. communis; Pt, P. trichocarpa; Vv, V. vinifera; Al, A. lyrata; Mt, M. truncatula; Bd, B. distachyon; Sh, S. bicolor; Sm, S. moellendorffii; Pp, P. patens; Cr, C. reinhardtii and Ol, O. lucimarinus. Phylogeny reconstruction was analyzed by neighbour-joining statistical method based on the JTT matrix-based model. Test of phylogeny was analyzed by the bootstrap method (1,000 replicates). Evolutionary analyses were conducted in MEGA 6.06.

S3 Fig. Molecular phylogenetic analysis of domains of the SAGA complex components. Amino acid sequences of domains of the SAGA complex subunits were used from At, A. thaliana; Dm, D. melanogaster; Hs, H. sapiens; Os, O. sativa; Mm, M. musculus; Sc, S. cerevisiae. (A) Protein domain of ADAs group; (B) Protein domain of SPTs group; (C) Protein domain of DUBm group; (D) Protein domain of CHD subunit; (E) Protein domain of SGF29 subunit; (F) Protein domain of TRA1 subunit; (G) Protein domain of TAFs group.

S4 Fig. Chromosomal distribution of the SAGA subunit genes in the Arabidopsis genome. SAGA encoding genes are plotted on the five Arabidopsis chromosomes according to their sequence spots. The chromosome number is shown at the top of each chromosome and the centromeric regions by constriction on chromosome line bar. Each identical duplicated chromosomal segment is marked by same line colour. The scale is in mega bases (Mb).

S5 Fig. Co-expressed gene network analysis for Arabidopsis SAGA complex. The co-expressed gene networks are drawn based on their rank of correlation from ATTED-II database. Orange line displays conserved co-expressed which is inferred from the comparison with mammalian co-expression data provided from COXPRESdb; Red dotted line display protein-protein interaction information that is provided from TAIR and IntAct. The octagon shape indicates transcription factor genes. White circles shape indicates SAGA complex genes which were used to give input for generating gene network. Gray circle shape indicates other genes in co-expressed gene networks.

S6 Fig. Analysis of the biological process of co-expressed gene network. A biological process is analyzed by TAIR database using 181 co-expressed genes obtained from ATTED-II database.

S7 Fig. Characterization of Arabidopsis mutant lines. Characterization of Arabidopsis chr5, gcn5, sgf1T, sgf29b, taf12b, taf13 and tra1a T-DNA insertion homozygous mutants were done by qRT-PCR. RNA was isolated from homozygous T-DNA insertion mutants and Col-0 leaves or seedlings.
S1 Table. Query sequences from S. cerevisiae and human used to search Arabidopsis and O. sativa genome for SAGA gene families.
(PDF)

S2 Table. Protein similarities of SAGA encoding gene in Arabidopsis, O. sativa, human and S. cerevisiae.
(PDF)

S3 Table. List of primer used in qRT-PCR.
(PDF)

S4 Table. Gene Ontology list of Arabidopsis and O. sativa for SAGA complex gene.
(PDF)

S5 Table. List of Arabidopsis SAGA complex subunits known and predicted protein interactions from STRING database.
(PDF)

S6 Table. List of S. cerevisiae SAGA complex subunits known and predicted protein interactions from STRING database.
(PDF)

S7 Table. MPSS data for Arabidopsis SAGA complex encoding genes showing different tissue-specific abundance.
(PDF)

S8 Table. MPSS data for O. sativa SAGA complex encoding genes showing different tissue-specific abundance.
(PDF)

S9 Table. List of co-expression genes in Arabidopsis.
(PDF)

S10 Table. List of genes obtained from ATTED-II for Arabidopsis SAGA complex co-expressed gene network analysis.
(PDF)

S11 Table. Analysis of cis-regulatory element in 1000bp upstream promoter sequences from TSS in SAGA complex subunit genes using PlantCARE and PLACE database.
(PDF)

Acknowledgments
The authors are thankful to the CSIR, India for funding under the Institutional and CSIR-EMPOWER (OLP79) project. Rakesh Srivastava is obliged to the CSIR for fellowship support.

Author Contributions
Conceived and designed the experiments: RS SVS. Performed the experiments: RS KMR BP. Analyzed the data: RS SVS. Contributed reagents/materials/analysis tools: RS KMR BP SPS. Wrote the paper: RS KMR SPS SVS.
References

1. Thomas MC, Chiang CM. The general transcription machinery and general cofactors. Critical reviews in biochemistry and molecular biology. 2006; 41(3):105–78. Epub 2006/07/25. doi: 10.1080/10409230600648736 PMID: 16858867.

2. Srivastava R, Rai KM, Srivastava M, Kumar V, Pandey B, Singh SP, et al. Distinct role of core promoter architecture in regulation of light-mediated responses in plant genes. Mol Plant. 2014; 7(4):626–41. Epub 2013/11/02. doi: 10.1093/mp/sst146 PMID: 24177688.

3. Kiran K, Ansari SA, Srivastava R, Lodhi N, Chaturvedi CP, Sawant SV, et al. The TATA-box sequence in the basal promoter contributes to determining light-dependent gene expression in plants. Plant physiology. 2006; 142(1):364–76. doi: 10.1104/pp.106.084319 PMID: 16844831; PubMed Central PMCID: PMC1557599.

4. Lodhi N, Ranjan A, Singh M, Srivastava R, Singh SP, Chaturvedi CP, et al. Interactions between upstream and core promoter sequences determine gene expression and nucleosome positioning in tobacco PR-1a promoter. Biochim Biophys Acta. 2008; 1779(10):634–44. doi:10.1016/j.bbagrm.2008.07.010 PMID: 18723134.

5. Shandilya J, Roberts SG. The transcription cycle in eukaryotes: from productive initiation to RNA polymerase II recycling. Biochim Biophys Acta. 2012; 1819(5):391–400. Epub 2012/02/07. doi: 10.1016/j.bbagrm.2012.01.010 PMID: 22306664.

6. Srivastava R, Srivastava R, Singh UM. Understanding the patterns of gene expression during climate change. Climate Change Effect on Crop Productivity: CRC Press; 2014. p. 279–328.

7. Naar AM, Lemon BD, Tjian R. Transcriptional coactivator complexes. Annual review of biochemistry. 2001; 70:475–501. Epub 2001/06/08. doi:10.1146/annurev.biochem.70.1.475 PMID: 11395415.

8. Sermwittayawong D, Tan S. SAGA binds TBP via its Spt8 subunit in competition with DNA: implications for TBP recruitment. The EMBO journal. 2006; 25(16):3791–800. Epub 2006/08/05. doi: 10.1038/sj.emboj.7601265 PMID: 16888622; PubMed Central PMCID: PMC1553190.

9. Weake VM, Workman JL. Inducible gene expression: diverse regulatory mechanisms. Nature reviews Genetics. 2010; 11(6):426–37. doi: 10.1038/nrg2781 PMID: 20421872.

10. Samara NL, Wolberger C. A new chapter in the transcription SAGA. Current opinion in structural biology. 2011; 21(6):767–74. Epub 2011/10/22. doi: 10.1016/j.sbi.2011.09.004 PMID: 22014650; PubMed Central PMCID: PMC3232345.

11. Weake VM, Workman JL. SAGA function in tissue-specific gene expression. Trends in cell biology. 2012; 22(4):177–84. Epub 2011/12/27. doi: 10.1016/j.tcb.2011.11.005 PMID: 22196215; PubMed Central PMCID: PMC3322277.

12. Baker SP, Grant PA. The SAGA continues: expanding the cellular role of a transcriptional co-activator complex. Oncogene. 2007; 26(37):5329–40. Epub 2007/08/19. doi: 10.1038/sj.ong.1210603 PMID: 17694076; PubMed Central PMCID: PMC2746020.

13. Koutelou E, Hirsch CL, Dent SY. Multiple faces of the SAGA complex. Current opinion in cell biology. 2010; 22(3):374–82. Epub 2010/04/07. doi: 10.1016/jceb.2010.03.005 PMID: 20363118; PubMed Central PMCID: PMC2900470.

14. Sterner DE, Grant PA, Roberts SM, Duggan LJ, Belotserkovskaya R, Pacella LA, et al. Functional organization of the yeast SAGA complex: distinct components involved in structural integrity, nucleosome acetylation, and TATA-binding protein interaction. Molecular and cellular biology. 1999; 19(1):86–98. Epub 1998/12/22. PMID:9858534; PubMed Central PMCID: PMC383868.

15. Govind CK, Zhang F, Qiu H, Hofmeyer K, Hinnebusch AG. Gcn5 promotes acetylation, eviction, and methylation of nucleosomes in transcribed coding regions. Molecular cell. 2007; 25(1):31–42. Epub 2007/01/16. doi: 10.1016/j.molcel.2006.11.020 PMID: 17218269.

16. Grant PA, Duggan L, Cote J, Roberts SM, Brownell JE, Candau R, et al. Yeast Gcn5 functions in two multisubunit complexes to acetylate nucleosomal histones: characterization of an Ada complex and the SAGA (Spt/Ada) complex. Genes & development. 1997; 11(13):1640–50. Epub 1997/07/01. PMID: 9224714.

17. Daniel JA, Grant PA. Multi-tasking on chromatin with the SAGA coactivator complexes. Mutat Res. 2007; 618(1–2):135–48. Epub 2007/03/06. doi: 10.1016/j.mrfmmm.2006.09.008 PMID: 17337012; PubMed Central PMCID: PMC1892243.

18. Han Y, Luo J, Ranish J, Hahn S. Architecture of the Saccharomyces cerevisiae SAGA transcription coactivator complex. The EMBO journal. 2014. Epub 2014/09/14. doi: 10.15252/embj.201488638 PMID: 25216679.

19. Larschan E, Winston F. The S. cerevisiae SAGA complex functions in vivo as a coactivator for transcriptional activation by Gal4. Genes & development. 2001; 15(15):1946–56. Epub 2001/08/04. doi: 10.1101/gad.911501 PMID: 11485989; PubMed Central PMCID: PMC312753.
20. Bhaumik SR, Green MR. Differential requirement of SAGA components for recruitment of TATA-box-binding protein to promoters in vivo. Molecular and cellular biology. 2002; 22(21):7365–71. Epub 2002/10/09. PMID: 12370284; PubMed Central PMCID: PMC135674.

21. Burgess RJ, Zhou H, Han J, Zhang Z. A role for Gcn5 in replication-coupled nucleosome assembly. Molecular cell. 2010; 37(4):469–80. Epub 2010/03/02. doi: 10.1016/j.molcel.2010.01.020 PMID: 20188666; PubMed Central PMCID: PMC2954627.

22. Kohler A, Schneider M, Cabal GG, Nehrbass U, Hurt E. Yeast Ataxin-7 links histone deubiquitination with gene gating and mRNA export. Nature cell biology. 2008; 10(6):707–15. Epub 2008/05/20. doi: 10.1038/nbc2173 PMID: 18488019.

23. Garcia-Oliver E, Garcia-Moliner V, Rodriguez-Navarro S. mRNA export and gene expression: the SAGA-TREX-2 connection. Biochim Biophys Acta. 2012; 1819(6):555–65. Epub 2011/12/20. doi: 10.1016/j.bbapap.2011.11.011 PMID: 22178374.

24. Jacobson S, Pilius L. The SAGA subunit Ada2 functions in transcriptional silencing. Molecular and cellular biology. 2009; 29(22):6033–45. doi: 10.1128/MCB.00542-09 PMID: 19737915; PubMed Central PMCID: PMC2772568.

25. Gunderson FQ, Johnson TL. Acetylation by the transcriptional coactivator Gcn5 plays a novel role in co-transcriptional spliceosome assembly. PLoS genetics. 2009; 5(10):e1000682. doi: 10.1371/ journal.pgen.1000682 PMID: 19834536; PubMed Central PMCID: PMC2752994.

26. Hark AT, Vlachonasios KE, Pavangadkar KA, Rao S, Gordon H, Adamakis ID, et al. Two Arabidopsis orthologs of the transcriptional coactivator ADA2 have distinct biological functions. Biochim Biophys Acta. 2009; 1789(2):117–24. Epub 2009/10/22. doi: 10.1016/j.bbadis.2009.09.003 PMID: 19296980.

27. Stockinger EJ, Mao Y, Regier MK, Triezenberg SJ, Thomashow MF. Transcriptional adaptor and histone acetyltransferase proteins in Arabidopsis and their interactions with CBF1, a transcriptional activator involved in cold-regulated gene expression. Nucleic acids research. 2001; 29(7):1524–33. Epub 2001/03/27. PMID: 11266554; PubMed Central PMCID: PMC312676.

28. Vlachonasios KE, Thomashow MF, Triezenberg SJ. Disruption mutations of ADA2b and GCN5 transcriptional adaptor genes dramatically affect Arabidopsis growth, development, and gene expression. The Plant cell. 2003; 15(3):626–38. Epub 2003/03/05. PMID: 12615937; PubMed Central PMCID: PMC150018.

29. Benhamed M, Bertrand C, Servet C, Zhou DX. Arabidopsis GCN5, HD1, and TAF1/HAF2 interact to regulate histone acetylation required for light-responsive gene expression. The Plant cell. 2006; 18(11):2893–903. Epub 2006/11/07. doi: 10.1105/tpc.106.043489 PMID: 17085686; PubMed Central PMCID: PMC16993931.

30. Benhamed M, Martin-Magniette ML, Taconnat L, Bitton F, Servet C, De Clercq R, et al. Genome-scale Arabidopsis promoter array identifies targets of the histone acetyltransferase GCN5. The Plant journal: for cell and molecular biology. 2008; 56(3):493–504. Epub 2008/07/06. doi: 10.1111/j.1365-313X.2008.03606.x PMID: 18644902.

31. Bertrand C, Bergounioux C, Domenichini S, Delarue M, Zhou DX. Arabidopsis histone acetyltransferase AtGCN5 regulates the floral meristem activity through the WUSCHEL/AGAMOUS pathway. The molecular cell. 2003; 12(7):903–14. Epub 2003/03/05. doi: 10.1016/S1090-2720(03)00082. Epub 2003/03/27. PMID: 12740375.

32. Alonso JM, Stepanova AN, Leisse TJ, Kim CJ, Chen H, Shinn P, et al. Genome-wide insertional mutagenesis of Arabidopsis thaliana. Science. 2003; 301(5633):638–41. Epub 2003/11/21. doi: 10.1126/science.1086391301.5633.638 [pii]; PMID: 12893945.

33. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic acids research. 1997; 25(24):4876–82. Epub 1998/02/28. doi: gka797 [pii]. PMID: 9396791.

34. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol. 2013; 30(12):2725–9. Epub 2013/07/17. doi: 10.1093/molbev/mst197 PMID: 24132122; PubMed Central PMCID: PMC3840312.

35. Lamesch P, Berardini TZ, Li D, Swarbreck D, Wilks C, Sasidharan R, et al. The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. Nucleic acids research. 2003; 31:153–7. Epub 2002/11/08. doi: 10.1093/nar/24.1.153 PMID: 12205164; PubMed Central PMCID: PMC63091.

36. Blank G, Hokamp K, Wolfe KH. A recent polyploidy superimposed on older large-scale duplications in the Arabidopsis genome. Genome Res. 2003; 13(2):137–44. Epub 2003/02/05. doi: 10.1101/gr.751803 PMID: 12562592; PubMed Central PMCID: PMC420368.

37. Lescot M, Dehais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, et al. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. Nucleic acids research. 2002; 30(1):325–7. Epub 2001/12/26. PMID: 11752327.
38. Higo K, Ugawa Y, Iwamoto M, Korenaga T. Plant cis-acting regulatory DNA elements (PLACE) database: 1999. Nucleic acids research. 1999; 27(1):297–300. Epub 1998/12/10. doi: gkc003 [pii]. PMID: 9847208.

39. Zimmermann P, Hirsch-Hoffmann M, Henning L, Grisssem W. GENEVESTIGATOR. Arabidopsis microarray database and analysis toolbox. Plant physiology. 2004; 136(1):2621–32. Epub 2004/09/18. doi: 10.1104/pp.104.04367136/1/2621 [pii]. PMID: 15375207; PubMed Central PMCID: PMC523327.

40. Meyers BC, Lee DK, Vu TH, Tej SS, Edberg SB, Matvienko M, et al. Arabidopsis MPSS. An online resource for quantitative expression analysis. Plant physiology. 2004; 135(2):801–13. Epub 2004/06/03. doi: 10.1104/pp.104.039495 PMID: 15173564; PubMed Central PMCID: PMC514116.

41. Franceschini A, Szklarczyk D, Frankild S, Kuhn M, Simonovic M, Roth A, et al. STRING v9.1: protein-protein interaction networks, with increased coverage and integration. Nucleic acids research. 2013; 41(D1):D808–D815. doi: 10.1093/nar/gks1094

42. Du Z, Zhou X, Ling Y, Zhang Z, Su Z. agrigO: a GO analysis toolkit for the agricultural community. Nucleic acids research. 2010; 38(Web Server issue):W64–70. Epub 2010/05/04. doi: 10.1093/nar/gkq310 PMID: 20435677; PubMed Central PMCID: PMC2896167.

43. Obayashi T, Hayashi S, Saeki M, Ohta H, Kinoshita K. ATTED-II provides coexpressed gene networks for Arabidopsis. Nucleic acids research. 2009; 37(Database issue):D987–91. Epub 2008/10/28. doi: 10.1093/nar/gkn807 PMID: 18953027; PubMed Central PMCID: PMC2686564.

44. Singh SP, Pandey T, Srivastava R, Verma PC, Singh PK, Tuli R, et al. BECLIN1 from Arabidopsis thaliana under the generic control of regulated expression systems, a strategy for developing male sterile plants. Plant Biotechnol J. 2010; 8(9):1055–22. doi: 10.1111/j.1467-7652.2010.00527.x PMID: 21050365.

45. Servet C, Conde e Silva N, Zhou DX. Histone acetyltransferase ATGCN5/HAG1 is a versatile regulator of developmental and inducible gene expression in Arabidopsis. Mol Plant. 2010; 3(4):670–7. Epub 2010/05/12. doi: sqs018 [pii] doi:10.1093/mp/sqs018 PMID: 20457643.

46. Watanabe K, Ohta H, Kinoshita K. Histone-Like Pairs. Mol Cell Biol. 2000; 20(1):340–6. doi:10.1128/MCB.01076-08 PMID: 10365964.

47. Dhalluin C, Carlson JE, Zeng L, He C, Aggarwal AK, Zhou MM. Structure and ligand of a histone acetyltransferase bromodomain. Nature. 1999; 399(6735):491–4. Epub 1999/07/15. doi: S0028-078X(98)95500-8 [pii] doi:10.1038/10640 PMID: 9787080.

48. Nagy Z, Riss A, Romier C, le GuezenneC X, Dongre AR, Orpinell M, et al. The human SPT20-containing SAGA complex plays a direct role in the regulation of endoplasmic reticulum stress-induced genes. Molecular and cellular biology. 2009; 29(6):1649–60. Epub 2008/12/31. doi: MCB.01076-08 [pii] doi:10.1128/MCB.01076-08 PMID: 19114550; PubMed Central PMCID: PMC2648226.

49. Yu J, Madison JM, Mundlos S, Winston F, Olsen BR. Characterization of a human homologue of the Saccharomyces cerevisiae transcription factor sp3 (SUPT3H). Genomics. 1998; 53(1):90–7. Epub 1998/10/27. doi: S0888-7543(98)95500-8 [pii] doi:10.1006/geno.1998.5500 PMID: 9787080.

50. Laubinger AG, Inouye C, Jain R, Tjian R. Bromodomains mediate an acetyl-histone encoded antisilencing function at heterochromatin boundaries. Molecular cell. 2009; 32(8):612–23. Epub 2009/05/04. doi: MCB.01076-08 [pii] doi:10.1093/mcb/mbp270 PMID: 20330952.

51. Yu J, Madison JM, Mundlos S, Winston F, Olsen BR. Characterization of a human homologue of the Saccharomyces cerevisiae transcription factor sp3 (SUPT3H). Genomics. 1998; 53(1):90–7. Epub 1998/10/27. doi: S0888-7543(98)95500-8 [pii] doi:10.1006/geno.1998.5500 PMID: 9787080.
59. Lago C, Clerici E, Mizzi L, Colombo L, Kater MM. TBP-associated factors in Arabidopsis. Gene. 2004; 205(1–2):231–41. Epub 2004/11/06. doi: S0378-1194(04)00515-3 [pii] doi: 10.1016/j.gene.2004.08.023 PMID: 15527982.

60. Marfella CG, Imbalzano AN. The Chd family of chromatin remodelers. Mutat Res. 2007; 618(1–60.

58. Mishra A, Puranik S, Prasad M. Structure and regulatory networks of WD40 protein in plants. J Plant Biochem Biotechnol. 2012; 21(1):32–9. doi: 10.1007/s13562-012-0134-1

61. Nagy Z, Tora L. Distinct GCN5/PCAF-containing complexes function as co-activators and are involved in transcription regulation and global histone acetylation. Oncogene. 2007; 26(37):5341–57. doi: 10.1038/sj.onc.1210604 PMID: 17694077.

62. Kaldis A, Tsementzi D, Tanriverdi O, Vlachonasios KE. Arabidopsis thaliana transcriptional co-activators ADA2b and SGF29a are implicated in salt stress responses. Planta. 2011; 233(4):749–62. Epub 2011/01/05. doi: 10.1007/s00425-010-1337-0 PMID: 21193996.

63. Saleh A, Schieltz D, Tanriverdi O, Vlachonasios KE. Arabidopsis thaliana transcriptional co-activators ADA2b and SGF29a are implicated in salt stress responses. Planta. 2011; 233(4):749–62. Epub 2011/01/05. doi: 10.1007/s00425-010-1337-0 PMID: 21193996.

64. Saleh A, Schieltz D, Tanriverdi O, Vlachonasios KE. Arabidopsis thaliana transcriptional co-activators ADA2b and SGF29a are implicated in salt stress responses. Planta. 2011; 233(4):749–62. Epub 2011/01/05. doi: 10.1007/s00425-010-1337-0 PMID: 21193996.

65. Saleh A, Schieltz D, Tanriverdi O, Vlachonasios KE. Arabidopsis thaliana transcriptional co-activators ADA2b and SGF29a are implicated in salt stress responses. Planta. 2011; 233(4):749–62. Epub 2011/01/05. doi: 10.1007/s00425-010-1337-0 PMID: 21193996.

66. Saleh A, Schieltz D, Tanriverdi O, Vlachonasios KE. Arabidopsis thaliana transcriptional co-activators ADA2b and SGF29a are implicated in salt stress responses. Planta. 2011; 233(4):749–62. Epub 2011/01/05. doi: 10.1007/s00425-010-1337-0 PMID: 21193996.

67. Saleh A, Schieltz D, Tanriverdi O, Vlachonasios KE. Arabidopsis thaliana transcriptional co-activators ADA2b and SGF29a are implicated in salt stress responses. Planta. 2011; 233(4):749–62. Epub 2011/01/05. doi: 10.1007/s00425-010-1337-0 PMID: 21193996.

68. Saleh A, Schieltz D, Tanriverdi O, Vlachonasios KE. Arabidopsis thaliana transcriptional co-activators ADA2b and SGF29a are implicated in salt stress responses. Planta. 2011; 233(4):749–62. Epub 2011/01/05. doi: 10.1007/s00425-010-1337-0 PMID: 21193996.

69. Saleh A, Schieltz D, Tanriverdi O, Vlachonasios KE. Arabidopsis thaliana transcriptional co-activators ADA2b and SGF29a are implicated in salt stress responses. Planta. 2011; 233(4):749–62. Epub 2011/01/05. doi: 10.1007/s00425-010-1337-0 PMID: 21193996.

70. Saleh A, Schieltz D, Tanriverdi O, Vlachonasios KE. Arabidopsis thaliana transcriptional co-activators ADA2b and SGF29a are implicated in salt stress responses. Planta. 2011; 233(4):749–62. Epub 2011/01/05. doi: 10.1007/s00425-010-1337-0 PMID: 21193996.

71. Saleh A, Schieltz D, Tanriverdi O, Vlachonasios KE. Arabidopsis thaliana transcriptional co-activators ADA2b and SGF29a are implicated in salt stress responses. Planta. 2011; 233(4):749–62. Epub 2011/01/05. doi: 10.1007/s00425-010-1337-0 PMID: 21193996.

72. Saleh A, Schieltz D, Tanriverdi O, Vlachonasios KE. Arabidopsis thaliana transcriptional co-activators ADA2b and SGF29a are implicated in salt stress responses. Planta. 2011; 233(4):749–62. Epub 2011/01/05. doi: 10.1007/s00425-010-1337-0 PMID: 21193996.

73. Saleh A, Schieltz D, Tanriverdi O, Vlachonasios KE. Arabidopsis thaliana transcriptional co-activators ADA2b and SGF29a are implicated in salt stress responses. Planta. 2011; 233(4):749–62. Epub 2011/01/05. doi: 10.1007/s00425-010-1337-0 PMID: 21193996.

74. Saleh A, Schieltz D, Tanriverdi O, Vlachonasios KE. Arabidopsis thaliana transcriptional co-activators ADA2b and SGF29a are implicated in salt stress responses. Planta. 2011; 233(4):749–62. Epub 2011/01/05. doi: 10.1007/s00425-010-1337-0 PMID: 21193996.

75. Saleh A, Schieltz D, Tanriverdi O, Vlachonasios KE. Arabidopsis thaliana transcriptional co-activators ADA2b and SGF29a are implicated in salt stress responses. Planta. 2011; 233(4):749–62. Epub 2011/01/05. doi: 10.1007/s00425-010-1337-0 PMID: 21193996.
Gangloff YG, Werten S, Romier C, Carre L, Poch O, Moras D, et al. The human TFIID components

Cohen R, Schocken J, Kaldis A, Vlachonasios KE, Hark AT, McCain ER. The histone acetyltransfer-

Furumoto T, Tamada Y, Izumida A, Nakatani H, Hata S, Izui K. Abundant expression in vascular tis-

Tamada Y, Nakamori K, Nakatani H, Matsuda K, Hata S, Furumoto T, et al. Temporary expression of

Lago C, Clerici E, Dreni L, Horlow C, Caporali E, Colombo L, et al. The Arabidopsis TFIID factor

Gangloff YG, Sanders SL, Romier C, Kirschner D, Weil PA, Tora L, et al. Histone folds mediate selec-

Rizhsky L, Liang H, Shuman J, Shulaev V, Davletova S, Mittler R. When defense pathways collide.

Lee TI, Causton HC, Holstege FC, Shen WC, Hannett N, Jennings EG, et al. Redundant roles for the

Taji T, Ohsumi C, Iuchi S, Seki M, Kasuga M, Kobayashi M, et al. Important roles of drought- and cold-

Bonnet J, Wang CY, Baptista T, Vincent SD, Hsiao WC, Stierle M, et al. The SAGA coactivator com-

Long JA, Ohno C, Smith ZR, Meyerowitz EM. TOPLESS regulates apical embryonic fate in Arabidop-

Weake VM, Dyer JO, Seidel C, Box A, Swanson SK, Peak A, et al. Post-transcription initiation function of the ubiquitous SAGA complex in tissue-specific gene activation. Genes & development. 2011; 25 (14):1499–509. doi: 10.1103/gad.2042611 PMID: 21764853; PubMed Central PMCID: PMC3143940.

Weake VM, Lee KK, Guelman S, Lin CH, Seidel C, Abmayr SM, et al. SAGA-mediated H2B deubiqui-

66. doi:10.1007/pp.1033431 [pii]. PMID: 15047901.

Lee TI, Causton HC, Holstege FC, Shen WC, Hannett N, Jennings EG, et al. Redundant roles for the TFIID and SAGA complexes in global transcription. Nature. 2000; 405(6787):701–4. Epub 2000/06/23. doi: 10.1038/35015104 PMID: 10864329.

Bonnet J, Wang CY, Baptista T, Vincent SD, Hsiao WC, Stierle M, et al. The SAGA coactivator complex acts on the whole transcribed genome and is required for RNA polymerase II transcription. Genes & development. 2014; 28(18):1999–2012. doi: 10.1101/gad.250225.114 PMID: 25228644; PubMed Central PMCID: PMC4173158.

Furumoto T, Tamada Y, Izumida A, Nakatani H, Hata S, Izui K. Abundant expression in vascular tissue of plant TAF10, an orthologous gene for TATA box-binding protein-associated factor 10, in Flavera trinervia and abnormal morphology of Arabidopsis thaliana transformants on its overexpression. Plant Cell Physiol. 2005; 46(1):108–17. Epub 2005/01/22. doi: pmid006 [pii] doi: 10.1093/pcp/pco006 PMID: 15659449.

Tamada Y, Nakamori K, Nakatani H, Matsuda K, Hata S, Furumoto T, et al. Temporary expression of the TAF10 gene and its requirement for normal development of Arabidopsis thaliana. Plant Cell Physiol. 2007; 48(1):134–46. Epub 2006/12/07. doi: pco048 [pii] doi: 10.1093/pcp/pco048 PMID: 17148695.

Lago C, Clerici E, Dreni L, Horlow C, Caporali E, Colombo L, et al. The Arabidopsis TFIID factor AtTAF6 controls pollen tube growth. Dev Biol. 2005; 285(1):91–100. Epub 2005/07/26. doi: S0012-1606(05)00398-2 [pii] doi: 101016/j.ydbio.2005.06.006 PMID: 16039640.

Gangloff YG, Sanders SL, Romier C, Kirschner D, Wei A, Pota, L, et al. Histone folds mediate selective heterodimerization of yeast TAF(II)25 with TFIID components yTAF(II)47 and yTAF(II)65 and with SAGA component ySPT7. Molecular and cellular biology. 2001; 21(5):1841–53. doi: 10.1128/MCB.21.5.1841-1853.2001 PMID: 11238921; PubMed Central PMCID: PMC86751.

Gangloff YG, Wertens R, Romier C, Carre L, Poch O, Moras D, et al. The human TFIID components TAF(II)135 and TAF(II)20 and the yeast SAGA components ADA1 and TAF(II)68 heterodimerize to form histone-like pairs. Molecular and cellular biology. 2000; 20(1):340–51. PMID: 10594036; PubMed Central PMCID: PMC85089.

Cohen R, Schocken J, Kaldis A, Vlachonasios KE, Hark AT, McCain ER. The histone acetyltransferase GCN5 affects the inflorescence meristem and stamen development in Arabidopsis. Planta. 2009; 230(6):1207–21. Epub 2009/09/23. doi: 10.1007/s00425-009-1012-5 PMID: 19771450.

Kornet N, Scheres B. Members of the GCN5 histone acetyltransferase complex regulate PLETHORA-mediated root stem cell niche maintenance and transit amplifying cell proliferation in Arabidopsis. The Plant cell. 2009; 21(4):1070–9. Epub 2009/04/21. doi: tpc.108.065300 [pii] doi: 10.1105/tpc.108.065300 PMID: 19376833.

Long JA, Ohno C, Smith ZR, Meyerowitz EM. TOPLESS regulates apical embryonic fate in Arabidopsis. Science. 2006; 312(5779):1520–3. Epub 2006/06/10. doi: 312/5779/1520 [pii] doi: 10.1126/science.1123841 PMID: 16763149.

Sieberer T, Hauser MT, Seifert GJ, Luschneg C. PROPRZ1, a putative Arabidopsis transcriptional adaptor protein, mediates auxin and cytokinin signals in the control of cell proliferation. Curr Biol. 2003; 13(10):837–42. Epub 2003/05/16. doi: S0960982203003270 [pii] doi: 1227 PMID: 12747832.

Mougiou N, Poulos S, Kaldis A, Vlachonasios K. Arabidopsis thaliana TBP-associated factor 5 is essential for plant growth and development. Molecular Breeding. 2012; 30(1):355–66. doi: 10.1007/s11032-011-9626-2

Weake VM, Dyer JO, Seidel C, Box A, Swanson SK, Peak A, et al. Post-transcription initiation function of the ubiquitous SAGA complex in tissue-specific gene activation. Genes & development. 2011; 25 (14):1499–509. doi: 10.1103/gad.2042611 PMID: 21764853; PubMed Central PMCID: PMC3143940.

Weake VM, Lee KK, Guelman S, Lin CH, Seidel C, Abmayr SM, et al. SAGA-mediated H2B deubiquitination controls the development of neuronal connectivity in the Drosophila visual system. The EMBO journal. 2008; 27(2):394–405. doi: 10.1038/sj.emboj.7601966 PMID: 18188155; PubMed Central PMCID: PMC2234343.
93. Carre C, Szymczak D, Pidoux J, Antoniewski C. The histone H3 acetylase dGcn5 is a key player in Drosophila melanogaster metamorphosis. Molecular and cellular biology. 2005; 25(18):8228–38. doi: 10.1128/MCB.25.18.8228-8238.2005 PMID: 16135811; PubMed Central PMCID: PMC1234334.

94. Herrera FJ, Yamaguchi T, Roelink H, Tjian R. Core promoter factor TAF9B regulates neuronal gene expression. eLife. 2014; 3:e02559. doi: 10.7554/eLife.02559 PMID: 25006164; PubMed Central PMCID: PMC4083437.

95. Xu W, Edmondson DG, Evrard YA, Wakamiya M, Behringer RR, Roth SY. Loss of Gcn5l2 leads to increased apoptosis and mesodermal defects during mouse development. Nature genetics. 2000; 26(2):229–32. doi: 10.1038/79973 PMID: 11017084.

96. Zapater M, Sohrmann M, Peter M, Posas F, de Nadal E. Selective requirement for SAGA in Hog1-mediated gene expression depending on the severity of the external osmostress conditions. Molecular and cellular biology. 2007; 27(11):3900–10. Epub 2007/04/04. doi: MCB.00089-07 [pii] doi:10.1128/MCB.00089-07 PMID: 17403898.

97. Pankotai T, Komonyi O, Bodai L, Ujfaludi Z, Muratoglu S, Ciurciu A, et al. The homologous Drosophila transcriptional adaptors ADA2a and ADA2b are both required for normal development but have different functions. Molecular and cellular biology. 2005; 25(18):8215–27. Epub 2005/09/02. doi: 25/18/8215 [pii] doi:10.1128/MCB.25.18.8215-8227.2005 PMID: 16135810.

98. Endo M, Tanigawa Y, Murakami T, Araki T, Nagatani A. PHYTOCHROME-DEPENDENT LATE-FLOWERING accelerates flowering through physical interactions with phytochrome B and CONSTANS. Proc Natl Acad Sci U S A. 2013; 110(44):18017–22. doi: 10.1073/pnas.1310631110 PMID: 24127609; PubMed Central PMCID: PMC3816489.

99. Shen Y, Devic M, Lepiniec L, Zhou DX. Chromodomain, Helicase and DNA-binding CHD1 protein, CHR5, are involved in establishing active chromatin state of seed maturation genes. Plant Biotechnol J. 2015. doi: 10.1111/pbi.12315 PMID: 25581843.