

REVIEW

Open Access



Conservation and diversity of the eukaryotic SAGA coactivator complex across kingdoms

Ying-Jiun C. Chen^{1,2} and Sharon Y. R. Dent^{1,2*}

Abstract

The SAGA complex is an evolutionarily conserved transcriptional coactivator that regulates gene expression through its histone acetyltransferase and deubiquitylase activities, recognition of specific histone modifications, and interactions with transcription factors. Multiple lines of evidence indicate the existence of distinct variants of SAGA among organisms as well as within a species, permitting diverse functions to dynamically regulate cellular pathways. Our co-expression analysis of genes encoding human SAGA components showed enrichment in reproductive organs, brain tissues and the skeletal muscle, which corresponds to their established roles in developmental programs, emerging roles in neurodegenerative diseases, and understudied functions in specific cell types. SAGA subunits modulate growth, development and response to various stresses from yeast to plants and metazoans. In metazoans, SAGA further participates in the regulation of differentiation and maturation of both innate and adaptive immune cells, and is associated with initiation and progression of diseases including a broad range of cancers. The evolutionary conservation of SAGA highlights its indispensable role in eukaryotic life, thus deciphering the mechanisms of action of SAGA is key to understanding fundamental biological processes throughout evolution. To illuminate the diversity and conservation of this essential complex, here we discuss variations in composition, essentiality and co-expression of component genes, and its prominent functions across Fungi, Plantae and Animalia kingdoms.

Keywords: SAGA complex, Transcription, Coactivator, Epigenetics, GCN5, Development, Stress, Disease, Cancer, Evolution

Background

The groundbreaking characterization of general control nonrepressed protein 5 (Gcn5) as the first histone acetyltransferase (HAT) linked to transcription from studies in *Tetrahymena* and yeast led to the discovery of the Spt-Ada-Gcn5 acetyltransferase (SAGA) complex [1–3]. SAGA is a multi-module assembly that is best known as a coactivator for gene transcription, and it functions in transcriptional initiation and elongation through enzymatic activities for acetylation and deubiquitylation, ability to bind specific histone modifications, and direct interactions with transcription factors (TFs) such as

TATA-binding protein (TBP) [3–6]. GCN5 is the lysine acetyltransferase (KAT) within SAGA, targeting lysines in histones H3 and H2B, and is particularly required for acetylation of histone H3 at lysine 9 (H3K9ac) in mammalian cells [3, 7]. Ubiquitin protease 8 (Ubp8)/ubiquitin-specific peptidase 22 (USP22) is the deubiquitylase (DUB) subunit responsible for catalyzing removal of ubiquitin from monoubiquitinated H2A and H2B [8, 9]. Both GCN5 and Ubp8/USP22 enzymes also target non-histone substrates in and out of the nucleus [10, 11]. GCN5 and adjacent subunit SAGA complex associated factor 29 (SGF29) bind acetylated lysine residues on H3/H4 tails and di/tri-methylation on H3K4 (H3K4me2/3), respectively [12, 13]. Additionally, SAGA components are implicated in non-transcriptional events including maintenance of genome integrity (e.g., GCN5 and USP22),

*Correspondence: stroth@mdanderson.org

¹ Department of Epigenetics and Molecular Carcinogenesis, The University of Texas MD Anderson Cancer Center, Smithville, TX, USA
Full list of author information is available at the end of the article



© The Author(s) 2021. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

mRNA surveillance (e.g., Sgf73) and export (e.g., Sus1) [14–16].

Yeast Gcn5 and its interacting partner, transcriptional adaptor 2 (Ada2), were found to be part of two complexes designated ADA and SAGA, with respective molecular masses of 0.8 and 1.8 megadalton [3]. Affinity purification and small-scale immunoprecipitation (IP) experiments have demonstrated the conserved existence and composition of SAGA in Arabidopsis, Drosophila and human cells [17–20]. In this review, we highlight differences in complex composition, co-expression patterns, and significant functions of SAGA across all three kingdoms to better understand the conservation and diversity of this essential complex.

SAGA composition across kingdoms

Complex modules

SAGA consists of multiple modules: HAT (renamed KAT to reflect its activity towards lysines in non-histone substrates) [21], DUB, TF-binding and core structural modules (Additional file 1: Table S1). Whereas the above modules are found in all eukaryotes, currently an additional splicing module containing splicing factor 3B subunit 3 (SF3B3) and SF3B5 is evident only in metazoans. The core module docks TBP to SAGA via suppressor of ty 3 (Spt3) and Spt8 subunits, which recruits TBP to gene promoters for the formation of the pre-initiation complex (PIC) [22, 23]. A feature of SAGA is that many components and modules are shared with other complexes. In the case of human SAGA, the KAT module is also present in the ADA2A-containing (ATAC) complex, ataxin 7-like 3 (ATXN7L3) and enhancer of yellow 2 (ENY2) of the DUB module interact with other USPs (USP27 and USP51x) to form other DUB complexes, core TATA-binding protein associated factor (TAF) proteins are shared with the TFIID complex, transformation/transcription domain associated protein (TRRAP) is also a subunit of the TAT-interactive protein 60-kDa (TIP60) complex, and the splicing module is common between SAGA and SF3B complexes [24, 25].

The KAT modules of ATAC and SAGA differ in that they contain ADA2A and ADA2B, respectively [26]. Paralogous *ADA2A* and *ADA2B* genes are encoded solely in higher eukaryotic genomes, and whether the GCN5-containing KAT module is incorporated into ATAC or SAGA is determined by its assembly with an either ADA2A or ADA2B protein [27]. In Drosophila, the KAT subunits separately form the ADA complex, while the yeast ADA complex contains two additional subunits, ADA HAT complex component 1 (Ahc1) and Ahc2 [28, 29]. Results of chromatin IP (ChIP) or affinity purification in humans, Drosophila and Arabidopsis, along with the dissociation of DUB module subunits from SAGA by the

proteasome regulatory particle demonstrated in yeast, have all implied that the DUB module may also exist independently of SAGA [30–32]. These findings, together with its inherently large composition, suggest that distinct versions of SAGA may be present to execute diverse functions.

Components important for complex structure

Biochemical and genetic experiments have demonstrated that components of the core structural module, but not of KAT nor DUB modules, are required for the overall structural integrity of SAGA [3, 4]. Deletion of yeast KAT or DUB module subunits only interfered with the assembly of the corresponding module. Recent structural studies by high resolution cryo-electron microscopy revealed that the connections of the core module to the two enzymatic modules are considerably flexible in both yeast and human SAGA complexes [22, 23, 33]. The carboxyl-terminal part of Sgf73 is embedded in the core module, thereby tethering the DUB module to the remainder of SAGA. Interestingly, Arabidopsis lacks an orthologue of Sgf73/ATXN7, and how the plant DUB module interacts with other parts of SAGA is currently unclear [20]. The KAT module has no parts embedded in the core structure and binds to the core surface via Ada3 [22].

Variations in composition

SAGA composition has remained largely conserved throughout eukaryotic evolution, however duplicated homologues that increase structural and functional complexity have been found in higher eukaryotes. Arabidopsis has more duplicates of SAGA components compared to yeast and metazoans, namely TRA1A/B, histone acetyltransferase of the TAFII250 family 1/2 (HAF1/2), TAF6/TAF6B, TAF12/TAF12B, ADA1A/B and SGF29A/B (Additional file 1: Table S1). Plant genomes evolve at higher rates and gene duplication events occur more often in plants than in most other eukaryotes [34]. As a result, a duplicate copy is found for the majority of annotated genes in plant genomes. Affinity purification experiments found that in the cases of TRA1A/B, TAF6/TAF6B, TAF12/TAF12B, ADA1A/B and SGF29A/B, both paralogues interacted with other SAGA subunits [20]. It is currently undetermined whether these paralogues are functionally commutable, or whether specific combinations have unique functions. Specifically in vertebrates, GCN5 has a paralogue named p300/CBP-associated factor (PCAF; Fig. 1 and Additional file 1: Table S1). GCN5 and PCAF are mutually exclusive in the KAT module, resulting in the formation of discrete versions of SAGA as well as ATAC complexes [17, 26].

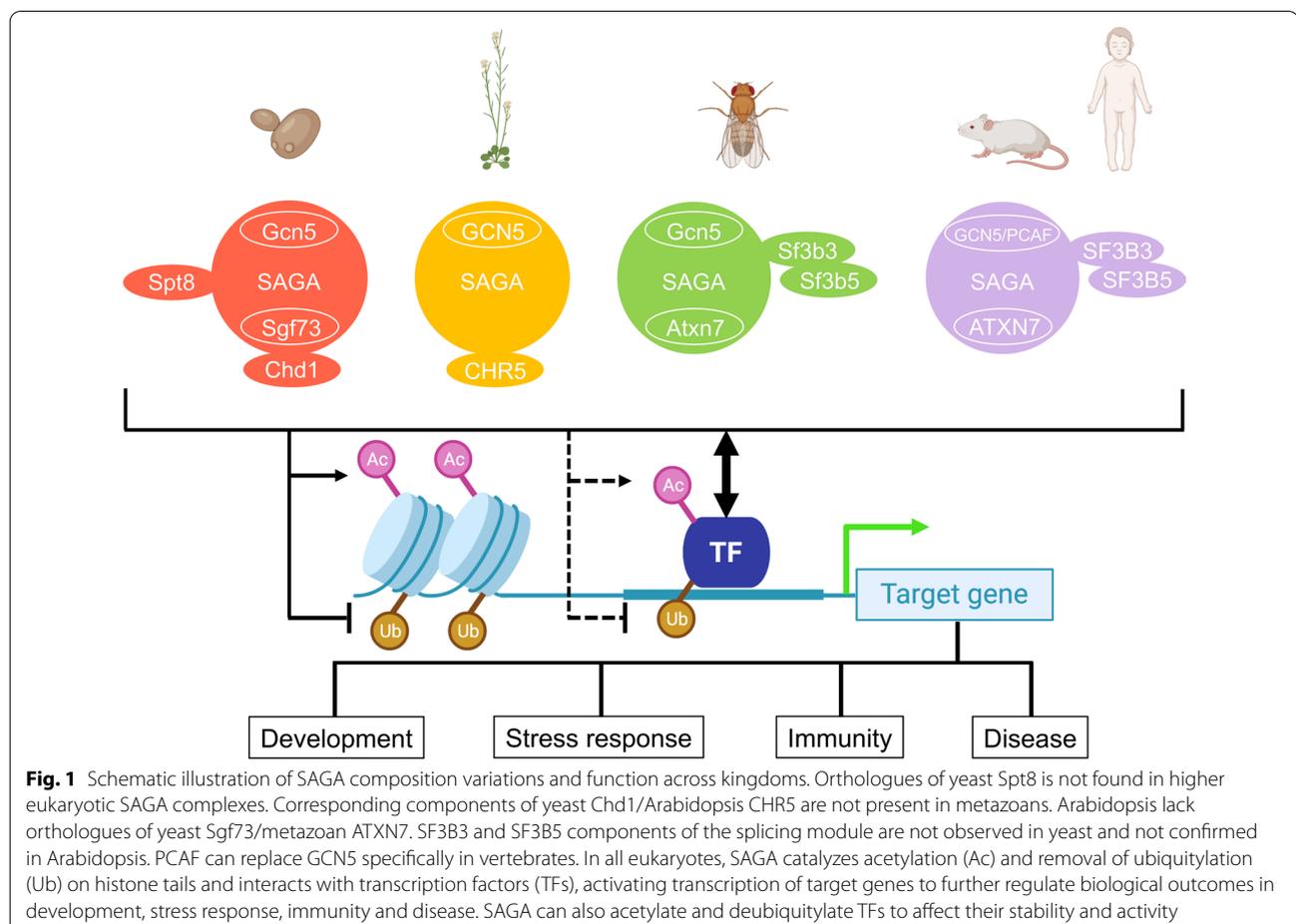
Orthologues of yeast Spt8 have not been identified in higher eukaryotic SAGA complexes (Fig. 1

and Additional file 1: Table S1). Considering that yeast SAGA-like (SLIK), alternatively called SAGA-altered Spt8-absent (SALSA), complex also lacks Spt8, multicellular SAGA complexes may be more similar to SLIK/SALSA [35, 36]. The SLIK/SALSA complex contains a carboxy-terminally truncated form of Spt7, resulting in the loss of the Spt8-interacting region [37]. However human SPT7 is homologous to the carboxy-terminal part of yeast Spt7, lacks a classical histone fold and bromodomain observed in its yeast orthologue, and instead harbors a bromo-associated domain that resembles a histone-like fold [38]. A component corresponding to yeast chromodomain-containing protein 1 (Chd1), an ATP-dependent chromatin remodeler, has solely been identified in Arabidopsis (CHD3; Fig. 1 and Additional file 1: Table S1) [20, 39]. Although orthologous CHD proteins exist in humans, none were detected to closely interact with SAGA members. As mentioned earlier, plant genomes do not appear to encode a Sgf73/ATXN7 orthologue found in yeast and metazoans. In metazoans, pre-mRNA splicing factors SF3B3 and SF3B5 were identified as SAGA subunits, but their functions in this

complex remain elusive [40]. Arabidopsis SF3B3/SIN3A-associated protein 130 (SAP130) protein (encoded by *SAP130A* and *SAP130B* genes) robustly co-purified with canonical SAGA components including ADA2B, SPT3 and TAF10 [20], suggesting that it may also be part of the SAGA complex.

Expression of SAGA component genes in tissues

Comprehensive expression studies of genes encoding SAGA components (hereafter referred to as SAGA component genes) in human tissues have not been reported, thus we generated a co-expression heatmap using the consensus dataset from the Human Protein Atlas (HPA), which combines HPA, GTEx and FANTOM5 datasets and includes RNA-sequencing data from 62 tissues and cell types (Fig. 2) [41–43]. The expression profiles show human SAGA component genes are generally ubiquitously expressed. In a majority of tissues, most SAGA component genes are expressed at comparable levels to the average expression of *TBP* (approximately 16 Normalized Expression). Interestingly, a pattern of enhanced expression in brain tissues such as the cerebral cortex and



cerebellum, the skeletal muscle and reproductive organs is observed for many SAGA component genes. The high expression of these genes in the cerebellum corresponds to the association of SAGA with a neurodegenerative disease, spinocerebellar ataxia type 7 (SCA7; further discussed in the following section). Catalytic subunit genes *GCN5*, *PCAF* and *USP22* particularly show increased expression in brain tissues. In fact, the top five tissues for *GCN5* expression are all those of the encephalon. Thalamus and amygdala are among the top five tissues for expression of both *GCN5* and *PCAF*. Correspondingly, *GCN5*, *PCAF*, as well as other SAGA members have critical developmental functions in the central nervous system (further discussed in the following section) [44, 45].

PCAF, *ADA3*, *SGF29*, *ADA2B*, *ENY2*, *USP22*, *TAF10* and *SF3B5* are all highly expressed in the skeletal muscle (Fig. 2), suggesting that the *PCAF*-containing version of SAGA may have a larger role in this tissue. Several studies have indicated the direct involvement of *PCAF* in muscle differentiation via disparate molecular pathways [46–48]. The KAT activity of *PCAF* on histones and DNA-binding myoblast determination protein (*MyoD*) activates the myogenic program, presumably in a multimeric complex that comprised *PCAF*, E1A binding protein p300/CREB-binding protein (p300/CBP) and *MyoD* [46, 47]. In committed muscle cell precursors, *PCAF* promotes histone deacetylase 2 (*HDAC2*) acetylation and recruitment to lamin A/C at the nuclear lamina, displacing *HDAC2* from *MyoD*, thus preventing deacetylation and inactivation of *MyoD* at the promoters of myogenic loci [48]. Developmental phenotypes were not observed in *Pcaf* null mice [49], suggesting redundancy between *GCN5* and *PCAF* in muscle tissues.

SF3B5 is expressed at extremely high levels, notably in pancreas, liver, bone marrow and blood cells (Fig. 2), which may reflect its functions in both SAGA and non-SAGA complexes. Interestingly, the expression of *ATXN7* splice variants are also abundant in the bone marrow and blood cells and can be grouped into two profiles, with variants 205, 207, 209, 210 significantly enriched in the thymus, other lymphoid tissues and bone marrow, and variants 212 to 214 more distributed in blood cells. Increased expression in blood cells, including T cells and others, is also observed for *SGF29*, *SPT20*, *TAF9* and *ADA1*. In agreement with these expression patterns, *SGF29*, *SPT20*, *ADA1* and other SAGA components were identified as positive modulators of forkhead box P3 (*FOXP3*) expression in regulatory T cells (Tregs; further discussed in the following section) [50, 51].

Enrichment of SAGA component transcripts in reproductive tissues is conserved among higher eukaryotes. Additionally, numerous component genes are abundantly expressed in *Arabidopsis* and *Drosophila* embryos, but

expression data in mammalian embryos are lacking [52–54]. In *Arabidopsis*, SAGA component genes show an increased expression pattern in the embryo, endosperm and maternal seed coat that support the growth and development of the embryo [52]. Accordingly, *TAF13* was shown to be essential for embryogenesis, possibly via interaction with *Medea* (*MEA*) and *Swinger* (*SWN*), two subunits of the plant polycomb repressive complex 2 (*PRC2*) [55]. The *taf13* mutant displays embryo arrest at the 8–16 cell stage and over-proliferation in the endosperm, resembling phenotypes of plants with *PRC2* mutations. In *Drosophila*, expression of nearly all SAGA component genes, including *Gcn5*, *Usp22/Nonstop*, *Ada1*, *Ada2b*, *Ada3*, *Sgf11*, *Sgf29*, *Spt3*, *Spt7*, *Spt20*, *Taf5/Wda*, *Taf6/Saf6* and *Taf9/E(y)1*, is enriched in the ovary and during the earliest stages of embryogenesis [53, 54]. In humans, SAGA component genes have a pattern of upregulated expression in the ovary, testis and breast of the reproductive system (Fig. 2). These expression data are consistent with the reported pivotal function of *GCN5* in mammalian spermiogenesis and fruitfly oogenesis [56, 57]. Mice deficient of *Gcn5* show aberrant chromatin dynamics and increased nucleosome retention in the sperm, ultimately leading to male fertility defects [57]. Taken together, the enhanced expression of SAGA component genes in reproductive tissues of higher eukaryotes is accordant with their requirement for successful reproduction.

Prominent roles of SAGA across kingdoms

Development

SAGA KAT and DUB module components that are essential for growth and development in metazoans, of which mutations cause lethality, are contrarily often not essential in yeast and plants (Additional file 1: Table S1). Differences in essentiality of homologous genes may reflect specific SAGA functions in particular metazoan or mammalian tissues or processes, such as the requirement of *USP22* for the formation of the mouse placenta [58]. At the molecular level, essentiality of genes may be altered by sequence changes that could potentially cause conformational differences in the encoded proteins, or the emergence of redundant genes. For instance, a large N-terminal extension, named the *PCAF* domain, exclusively exists in the metazoan *GCN5* and is important for nucleosome binding [59]. Albeit non-essential, deletion of yeast *GCN5* causes pleiotropic phenotypes including defective cell cycle progression and shortened replicative lifespan [60, 61]. Hypomorphic mutants of *Arabidopsis GCN5* also display pleiotropic defects such as dwarfism and aberrant root, leaf and floral development, but a null mutant has yet to be reported [62]. In *Drosophila*, *Gcn5* hypomorphic alleles cause abnormal appendages

and absence of abdominal cuticle deposition, while null alleles hamper onset of oogenesis and metamorphosis [56]. *Gcn5* null mice die soon after gastrulation, reaching approximately one-third of embryo development [49]. Deletion of PCAF, which is 71% identical to GCN5, in mice does not result in overt developmental defects before or after birth. However, embryos deficient for both *Gcn5* and *Pcaf* die earlier than those null for *Gcn5*, indicating partial redundancy between these KATs during early embryogenesis.

Tra1 is essential in budding yeast (*Saccharomyces cerevisiae*) but not in fission yeast (*Schizosaccharomyces pombe*), while deletions of the genes encoding the Taf proteins, which also function in the TFIID complex, cause lethality in both yeast species [63–65]. Depletion of other SAGA core components that are crucial for structural integrity, including Spt7, Spt20 and Ada1, results in severe growth phenotypes, whereas mutations in components not critical for complex integrity, including *Gcn5*, *Ada2*, *Ada3*, *Spt3* and *Spt8*, display moderately slow growth under various nutritional and stress conditions [4]. Deletion of *UBP8* has a positive effect for extending lifespan of yeast cells [66]. In contrast to yeast as was the case for GCN5, both fruitfly and mouse *ADA3*, *TRRAP*/*Tra1*, *ATXN7L3*/*Sgf11* and *USP22*/*Nonstop* are essential for early development [67–73]. Lethality of *Trrap* or *Ada3* null mutations in mice is linked to cell cycle progression defects [67, 72]. Death of mouse *Usp22* mutants is due at least in part to failure in placenta formation, which is related to inhibition of transforming growth factor beta (TGF- β) and receptor tyrosine kinase (RTK) signaling pathways that control vascularization [58].

Both SAGA-specific *Ada2b* and ATAC-specific *Ada2a* are essential in *Drosophila*, but at different developmental phases; *Ada2a* and *Ada2b* null mutants are larva and pupa lethal, respectively, indicating loss of *Ada2a* is slightly more severe than loss of *Ada2b* [74]. Arabidopsis *ada2b* null mutants exhibit pleiotropic phenotypes mirroring *gcn5* mutants but are not lethal, whereas *ada2a* null mutants do not exhibit pronounced phenotypes [62]. Of note, sequence identities indicate that Arabidopsis *ADA2A* is more closely related to human *ADA2B* while Arabidopsis *ADA2B* is more similar to human *ADA2A*. *ADA2A* and *ADA2B* share 45% sequence identity in Arabidopsis, 17% in *Drosophila*, and 27% in humans. In accordance with their functions in different complexes, the two paralogues are predictably non-redundant, as *Ada2a* could not rescue *Ada2b* mutant flies, and vice versa [75]. Rescue experiments using chimeric proteins indicate the C-terminal region determines the incorporation into SAGA or ATAC complexes. The N-terminus of *Ada2b* fused with C-terminus of *Ada2a* could partially rescue *Ada2a* mutants and restore ATAC-mediated

H4K5ac and H4K12ac levels; the N-terminus of *Ada2a* fused with C-terminus of *Ada2b* could fully rescue *Ada2b* mutants and restore SAGA-targeted H3K9ac and H3K14ac levels. Related to this finding, a SWI3-RSC8-MOIRA (SWIRM) domain lies within the C-terminus of *Ada2a* but appeared absent in *Drosophila* *Ada2b*. The SWIRM domain is annotated in both of the human paralogues, therefore it remains to be ascertained whether the C-terminal region consistently determines complex-specific incorporation in humans.

Stress response

Yeast SAGA members contribute to activation of various sets of stress-regulated genes under conditions such as glucose, osmotic and ER stresses [76–78]. For instance, SAGA is required for the recruitment of Mediator and switch/sucrose non-fermentable (SWI/SNF) complexes to promoters of genes activated by low glucose, as demonstrated by comparison of ChIP patterns in wild type and *gcn5* or *ada1* deletion strains [78]. Deletion of *GCN5*, *ADA2*, *ADA3*, or *SPT20* results in partially or completely defective unfolded protein response following ER stress due to failure of transcriptional induction of response programs [76]. Concordantly, human SAGA is recruited to ER stress-regulated genes possibly via interaction between GCN5 and the nuclear factor Y (NFY) that binds to ER stress-response element sequences in promoters [30].

In Arabidopsis, GCN5 and *ADA2B* have been widely studied in abiotic stress response [79]. Under heat stress, GCN5 is enriched at promoters of genes encoding heat stress TFs, heat shock factor A3 (HSFA3) and UV hypersensitive 6 (UVH6), where it promotes H3K9ac and H3K14ac that are associated with activation of these genes, thus preserving thermotolerance [80]. Upon cold exposure, GCN5, *ADA2B* and C-repeat binding factor 1 (CBF1) activate cold-responsive genes and consequently increase freezing tolerance [62]. GCN5 and *ADA2B* also mediate H3K9ac and H3K14ac to facilitate expression of salt stress-responsive genes, including genes required for maintenance of cell wall integrity and salt tolerance [81, 82]. The involvement of SAGA components in drought response was manifested in the black cottonwood tree *Populus trichocarpa* [83]. GCN5 and *ADA2B* promote H3K9ac, RNA polymerase II enrichment, and expression of drought-responsive genes upon recruitment by a drought-induced TF, abscisic acid-responsive element binding protein 1 (AREB1). Collectively, GCN5 and *ADA2B* regulate plant adaptation to abiotic stresses by interaction with stress-induced TFs and transcriptional activation of stress-responsive genes.

SAGA functions in stress response have been less characterized in metazoans, however they are

repeatedly linked to disease states (discussed below). Commonly in mammals, physiological stress and disorders are closely associated. One example is the acetylation and inactivation of peroxisome proliferators gamma coactivator 1 (PGC-1) family of transcriptional coactivators including PGC-1 α and PGC-1 β , which are conferred by GCN5 and mediate glucose metabolism [84, 85]. Dysregulation of these pathways generates metabolic stress and further contributes to pathogenesis and disease progression.

Immunity

SAGA components regulate the development of specific populations of immune cells. During granulocyte differentiation, GCN5 acts as a negative regulator through catalyzing acetylation of a TF, CCAAT enhancer binding protein alpha (C/EBP α), thereby inhibiting its DNA-binding ability [86]. On the other hand, GCN5 is required for T cell differentiation and full activation in immune response, wherein nuclear factor of activated T cells (NFAT) recruits GCN5 to catalyze H3K9ac at the *interleukin-2* (*IL-2*) promoter and to activate *IL-2* transcription upon antigen stimulation [87]. Acetylation of early response growth protein 2 (EGR2) by GCN5 activates the TF activity of EGR2 and stimulates the maturation of invariant natural killer T cells [88]. PCAF plays a larger role in facilitating production of induced Tregs than GCN5, which is achieved by PCAF-mediated IL-2 generation and mothers against decapentaplegic homologue 3 (SMAD3) phosphorylation in TGF- β signaling [89]. Double knockout of *Gcn5* and *Pcaf* in mice Tregs decreases Treg stability in peripheral lymphoid tissues and subsequently causes lethal autoimmunity. Recently, CRISPR screens performed by two groups simultaneously identified SAGA components, including USP22, ATXN7L3, ADA1, ADA2B, ADA3, SGF29, SPT20, TAF5L and TAF6L, as novel regulators promoting expression of FOXP3, a master TF governing the stability and suppressor activity of Tregs [50, 51]. USP22 maintains FOXP3 levels through both transcriptional and post-translational regulations: USP22 reduces H2B ubiquitination at the *FOXP3* locus and concurrently mediates deubiquitination of FOXP3 protein, thereby sustaining transcription and preventing protein degradation. Finally, in addition to direct regulation of immune cells, GCN5 and PCAF suppress TANK-binding kinase 1 (TBK1) activity in the cytoplasm by blocking phosphorylation events on TBK1, ultimately inhibiting innate immune signaling and interferon production upon viral infection [90]. Altogether, these studies demonstrate that SAGA components are critical for innate as well as adaptive immunity

and are potential therapeutic targets in the treatment of immune disorders such as autoimmune diseases.

Neural development and neurodegenerative diseases

SAGA members are involved in neural development of the central and peripheral nervous systems, and aberrations in a single component could lead to detrimental consequences. Mice with *Gcn5*-deficient neural stem and precursor cells have a microcephaly phenotype caused by reduction of brain mass, whereas mice homozygous for a catalytic-dead allele of *Gcn5* are defective in cranial neural tube closure, resulting in exencephaly [44, 91]. Although *Pcaf* null mice do not display developmental defects, they exhibit impaired memory and learning ability, as well as an overreactive emotional response to acute stress [45]. Knockdown of *Usp22* facilitates neural differentiation in the developing mouse brain [92]. USP22 is the predominant DUB that deubiquitinates and stabilizes hairy and enhancer of split 1 (HES1), a transcriptional repressor that maintains the undifferentiated state of neural stem cells. *Drosophila* *Usp22/Nonstop* and *Sgf11* regulate neural development in the visual system via H2B deubiquitination, however loss of their individual functions has no effects on global H3K9Ac [70]. KAT module mutants *Gcn5* and *Ada2b* display similar defects in axon targeting as those observed in *Nonstop* and *Sgf11* mutants, but are coupled with a clear reduction of global H3K9Ac levels. The flies bearing mutations interfering with DUB or KAT activities show overlapping and distinct transcriptional profiles, indicating the two enzymatic modules have both common and discrete targets.

Expansion of a polyglutamine tract in ATXN7 (PolyQ-ATXN7) results in SCA7, a disease that is characterized by cerebellar and retinal degeneration. PolyQ-ATXN7 can incorporate into the DUB module of SAGA, not altering DUB activity but forming a dominant-negative version of SAGA with inhibited coactivator activity [93, 94]. PolyQ-ATXN7 dramatically attenuates expression of rod-specific genes such as *rhodopsin* (*Rho*) and *G protein subunit alpha transducin 1* (*Gnat1*), leading to progressive loss of photoreceptor function in the SCA7 mouse model [95]. Deletion of *Gcn5* in SCA7 mice accelerates retinal degeneration and onset of ataxia, further suggesting a crucial role for SAGA in SCA7 pathogenesis [96].

Cancer

Genome-wide CRISPR-based screens have identified human SAGA component genes as dependencies for viability of different cancer cell types. GCN5 was isolated to be essential for acute myeloid leukemia (AML) cell proliferation; ADA2B and several other SAGA components were identified as selective dependencies essential for the

growth and survival of *MYCN*-amplified neuroblastoma cell lines [97, 98]. The large-scale Cancer Dependency Map (DepMap) [99, 100] revealed SAGA components are important for viability of a wide spectrum of cancer cell lines, including those of multiple myeloma, lung and breast cancer. Dependency scores of SAGA component genes exhibit a strong positive correlation with each other, indicating this group of genes function together as one regulatory complex. Ectopic expression of SAGA components can promote oncogenesis or tumor progression. GCN5 acetylates and reinforces stability of the oncogenic MYC TF that is overexpressed in a majority of human cancers, and in turn MYC facilitates expression of SAGA component genes in a positive feedback loop [101, 102]. Moreover, MYC directly interacts with GCN5 and TRRAP to recruit SAGA to its downstream targets for transcriptional activation [103]. GCN5 expression and function therefore accelerates growth of MYC-overexpressing AML, Burkitt's lymphoma (BL), hepatocellular carcinoma (HCC), colon and non-small cell lung cancer (NSCLC) in cellular systems and, shown in BL, HCC and NSCLC, in vivo mouse models [97, 104–109]. Overexpression of ADA3 enhances proliferation of breast cancer cells, which again is associated with MYC upregulation [110]. A microarray screen identified *USP22* as part of the 11-gene “death from cancer” signature driven by B cell-specific Moloney murine leukemia virus integration site 1 (BMI-1), a PRC1 member, in metastatic prostate tumors [111]. This BMI-1 pathway signature predicts disease recurrence, distant metastasis, and poor prognosis in 11 types of malignancies.

In addition to MYC, SAGA components also regulate cancer development through post-translational modifications of other non-histone substrates such as chromatin modifiers. PCAF acetylates enhancer of zeste homologue 2 (EZH2), a histone methyltransferase and PRC2 subunit, augmenting silencing of tumor suppressor genes and resulting in progression of lung adenocarcinoma [112]. Acetylation of intestine-specific homeobox (ISX) and bromodomain-containing protein 4 (BRD4) by PCAF promotes physical association and nuclear translocation of these proteins to drive epithelial–mesenchymal transition (EMT) through activation of EMT genes in lung cancer cells [113]. In prostate cancer, USP22 mediates the deubiquitination of xeroderma pigmentosum complementation group C (XPC), a nucleotide excision repair protein that detects DNA damage [114].

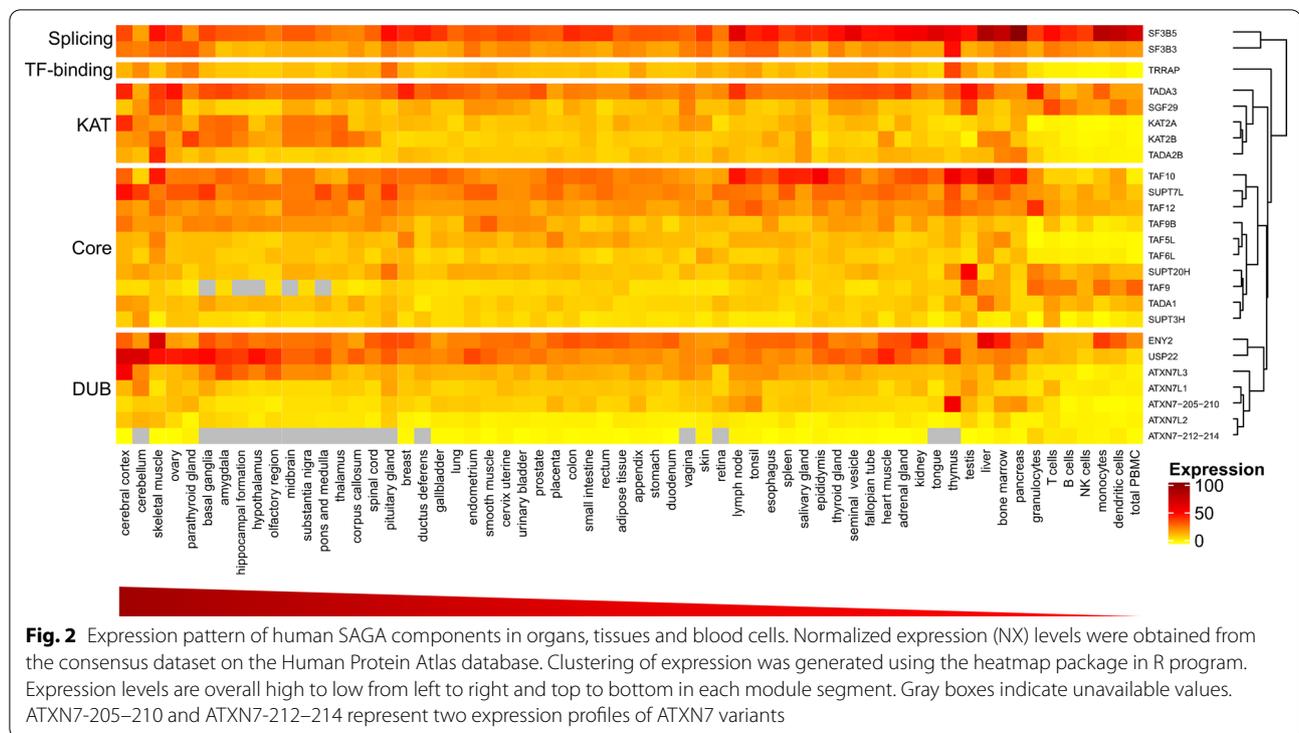
Conclusions and future perspectives

Unlike other GCN5-containing complexes such as the metazoan-specific ATAC complex, SAGA is uniquely conserved in all eukaryotes. The recently elucidated high-resolution structures of SAGA from two yeast genera

and from humans enlightened our understanding of its dynamic conformation in unicellular and higher eukaryotes [22, 23, 33]. Attributable to its intrinsic large composition and structural flexibility of the KAT and DUB modules in relation to the core module, each module may have independent roles in the process of transcriptional regulation, which is consistent with different impacts to gene expression profiles upon mutation of components belonging to separate modules [70]. Dissimilar mutant phenotypes also reflect different individual roles of subunits, such as cofactors in the enzymatic modules and core subunits involved in maintenance of complex integrity that are dispensable or absolutely essential [4]. In many cases, SAGA components also perform functions in other protein complexes, complicating interpretation of genetic phenotypes.

Our comparison of the compositions, co-expression patterns, and defined functions for SAGA suggests the existence of alternative versions of the complex across kingdoms as well as within a species (Figs. 1, 2). Some studies indicated that SAGA is a general coactivator required broadly for transcription [6, 115], whereas others conversely elucidated its specific targeting to gene promoters through interactions with particular transcription factors [30, 69, 100, 116]. The hypothesis that SAGA has tissue, developmental, physiological and pathological state-dependent compositions, interactions and functions would explain and support both findings. Although it is established that SAGA performs critical functions in development and stress response conservatively across Fungi, Plantae and Animalia kingdoms, and in immunity and disease in metazoans (Fig. 1), the underlying mechanisms of how distinct compositions assemble and functions are executed remain to be addressed. Our co-expression analysis of transcript levels indicates tissue-specific expression may provide one layer of regulation for SAGA composition, but protein quantification across tissues for all SAGA components is required to further validate this possibility. Comprehensive definition of SAGA interaction partners from different cell types and conditions would also generate important insights as to how SAGA is recruited to its targets in varying biological contexts. Employment of recent techniques including Assay for Transposase-Accessible Chromatin followed by sequencing (ATAC-seq) could aid in complementing the challenges of performing CHIP for certain SAGA members and greatly advance our knowledge of the impact of SAGA on chromatin organization and accessibility.

The discovery of Gcn5 as a gene-activating KAT more than two decades ago revolutionized our understanding of the connection between histone modification and transcriptional regulation [2]. However, we are still only at the



tip of the iceberg in exploring the potential of manipulating activities of GCN5 and other SAGA components to combat human ailments. In addition to directly impact tumor formation and progression, for example, targeting of SAGA components could affect response to immunotherapy given that SAGA functions in both innate and adaptive immune systems. Indeed, loss of PCAF in mice decreases lung adenocarcinoma tumor volume and augments anti-tumor immunity without triggering autoimmune side effects, potentiating clinical benefits from PCAF inhibition in combination with immunotherapy [89]. Data mining of publicly available resources unveiled the importance of poorly characterized SAGA component genes for viability of cancer cells (99, 100), thus clarifying these component functions may uncover new therapeutic targets for cancer treatment. Finally, the conserved role of SAGA in neural development among metazoans and its understudied role in muscle development or maintenance prompt further investigation of SAGA-mediated mechanisms in neurodegenerative and neuromuscular diseases. These research directions offer exciting prospects for future definition of SAGA functions in specific contexts.

Abbreviations

Gcn5: General control nonrepressed protein 5; HAT: Histone acetyltransferase; SAGA complex: Spt-Ada-Gcn5 acetyltransferase complex; TF: Transcription factor; TBP: TATA-binding protein; KAT: Lysine acetyltransferase; H3K9/14ac: Acetylation at lysine 9/14 in histone H3; Ubp8: Ubiquitin protease 8; USP22: Ubiquitin-specific peptidase 22; H3K4me2/3: Di/tri-methylation at lysine 4 in

histone H3; Ada: Transcriptional adaptor; IP: Immunoprecipitation; SF3B3/5: Splicing factor 3B subunit 3/5; Spt: Suppressor of ty; PIC: Pre-initiation complex; ATAC complex: ADA2A-containing complex; ATXN7(L3): Ataxin 7(-like 3); ENY2: Enhancer of yellow 2; TAF: TATA-binding protein associated factor; TRRAP: Transformation/transcription domain associated protein; TIP60: TAT-interactive protein 60-kDa; Ahc1/2: ADA HAT complex component 1/2; ChIP: Chromatin IP; HAF: Histone acetyltransferase of the TAFII250 family; PCAF: p300/CBP-associated factor; SLIK complex: SAGA-like complex; SALSA complex: SAGA-altered SPT8-absent complex; Chd: Chromodomain-containing protein; HPA: Human Protein Atlas; MyoD: Myoblast determination protein; p300/CBP: E1A binding protein p300/CREB-binding protein; HDAC2: Histone deacetylase 2; FOXP3: Forkhead box P3; Treg: Regulatory T cell; MEA: Medea; SWN: Swinger; PRC1/2: Polycomb repressive complex 1/2; TGF- β : Transforming growth factor beta; RTK: Receptor tyrosine kinase; SWIRM: SWI3-RSC8-MOIRA; SWI/SNF complex: Switch/sucrose non-fermentable complex; NFY: Nuclear factor Y; HSFA3: Heat shock factor A3; UVH6: UV hypersensitive 6; CBF1: C-Repeat binding factor 1; AREB1: Abscisic acid-responsive element binding protein 1; PGC-1: Peroxisome proliferators gamma coactivator 1; C/EBP α : CCAAT Enhancer binding protein alpha; NFAT: Nuclear factor of activated T cells; IL-2: Interleukin-2; EGR2: Early response growth factor 2; SMAD3: Mothers against decapentaplegic homologue 3; TBK1: TANK-binding kinase 1; HES1: Hairy and enhancer of split 1; PolyQ-ATXN7: Polyglutamine expanded ATXN7; SCA7: Spinocerebellar ataxia type 7; Rho: Rhodopsin; Gnat1: G protein subunit alpha transducin 1; AML: Acute myeloid leukemia; BL: Burkitt's lymphoma; HCC: Hepatocellular carcinoma colon cancer; NSCLC: Non-small cell lung cancer; BMI-1: B cell-specific Moloney murine leukemia virus integration site 1; EZH2: Enhancer of zeste homologue 2; ISX: Intestine-specific homeobox; BRD4: Bromodomain-containing protein 4; EMT: Epithelial-mesenchymal transition; XPC: Xeroderma pigmentosum complementation group C.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13072-021-00402-x>.

Additional file 1: Table S1. SAGA composition and essentiality of each component in yeast, Arabidopsis, Drosophila and mice. *Deletion of TRA1 is lethal in *S. cerevisiae* but viable in *S. pombe*. **Homozygotes die at 7 to 8 weeks. ND, not determined.

Acknowledgements

We thank Bhaskara Govinal Badiger for generation of the heatmap and critical reading of the manuscript. We also thank Evangelia Koutelou for useful comments on the manuscript.

Authors' contributions

YJCC wrote the manuscript and analyzed the co-expression data. SYRD contributed to writing and editing of the manuscript. All authors have read and approved the final manuscript.

Funding

This work was supported by NIH R35GM131678-01 and R01HD094400 grants to SYRD.

Availability of data and materials

Not applicable.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no conflict of interest.

Author details

¹Department of Epigenetics and Molecular Carcinogenesis, The University of Texas MD Anderson Cancer Center, Smithville, TX, USA. ²The Center for Cancer Epigenetics, The University of Texas MD Anderson Cancer Center, Houston, TX, USA.

Received: 22 April 2021 Accepted: 3 June 2021

Published online: 10 June 2021

References

- Brownell JE, Allis CD. An activity gel assay detects a single, catalytically active histone acetyltransferase subunit in Tetrahymena macronuclei. *Proc Natl Acad Sci USA*. 1995;92:6364–8.
- Brownell JE, Zhou J, Ranalli T, Kobayashi R, Edmondson DG, Roth SY, et al. Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell*. 1996;84:843–51.
- Grant P, 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 Dev*. 1997;11:1640–50.
- 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. *Mol Cell Biol*. 1999;19:86–98.
- Daniel J, Torok M, Sun Z, Schieltz D, Allis C, Yates J III, et al. Deubiquitination of histone H2B by a yeast acetyltransferase complex regulates transcription. *J Biol Chem*. 2004;279:1867–71.
- 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 Dev*. 2014;28:1999–2012.
- Jin Q, Yu LR, Wang L, Zhang Z, Kasper LH, Lee JE, et al. Distinct roles of GCN5/PCAF-mediated H3K9ac and CBP/p300-mediated H3K18/27ac in nuclear receptor transactivation. *EMBO J*. 2011;30:249–62.
- Henry KW, Wyce A, Lo WS, Duggan LJ, Emre NC, Kao CF, et al. Transcriptional activation via sequential histone H2B ubiquitylation and deubiquitylation, mediated by SAGA-associated Ubp8. *Genes Dev*. 2003;17:2648–63.
- Zhang XY, Pfeiffer HK, Thorne AW, McMahon SB. USP22, an hSAGA subunit and potential cancer stem cell marker, reverses the polycomb-catalyzed ubiquitylation of histone H2A. *Cell Cycle*. 2008;7:1522–4.
- Melo-Cardenas J, Zhang Y, Zhang DD, Fang D. Ubiquitin-specific peptidase 22 functions and its involvement in disease. *Oncotarget*. 2016;7:44848–56.
- Downey M. Non-histone protein acetylation by the evolutionarily conserved GCN5 and PCAF acetyltransferases. *Biochim Biophys Acta Gene Regul Mech*. 2020;22:194608.
- Li S, Shogren-Knaak MA. The Gcn5 bromodomain of the SAGA complex facilitates cooperative and cross-tail acetylation of nucleosomes. *J Biol Chem*. 2009;284(14):9411–7.
- Bian C, Xu C, Ruan J, Lee KK, Burke TL, Tempel W, et al. Sgf29 binds histone H3K4me2/3 and is required for SAGA complex recruitment and histone H3 acetylation. *EMBO J*. 2011;30:2829–42.
- Rodríguez-Navarro S, Fischer T, Luo MJ, Antúnez O, Bretschneider S, Lechner J, et al. Sus1, a functional component of the SAGA histone acetylase complex and the nuclear pore-associated mRNA export machinery. *Cell*. 2004;116:75–86.
- Atanassov BS, Evrard YA, Multani AS, Zhang Z, Tora L, Devys D, et al. Gcn5 and SAGA regulate shelterin protein turnover and telomere maintenance. *Mol Cell*. 2009;35:352–64.
- Kim M, Choi Y, Kim H, Lee D. SAGA DUBm-mediated surveillance regulates prompt export of stress-inducible transcripts for proteostasis. *Nat Commun*. 2019;10:2458.
- Martinez E, Kundu TK, Fu J, Roeder RG. (1998) A human SPT3-TAFII31-GCN5-L acetylase complex distinct from transcription factor IID. *J Biol Chem*. 1998;273:23781–5.
- Guelman S, Suganuma T, Florens L, Weake V, Swanson SK, Washburn MP, et al. The essential gene wda encodes a WD40 repeat subunit of Drosophila SAGA required for histone H3 acetylation. *Mol Cell Biol*. 2006;26:7178–89.
- Nassrallah A, Rougée M, Bourbousse C, Drevensek S, Fonseca S, Iniesto E, et al. DET1-mediated degradation of a SAGA-like deubiquitination module controls H2Bub homeostasis. *Elife*. 2018;7:e37892.
- Pfab A, Bruckmann A, Nazet J, Merkl R, Grasser KD. The adaptor protein ENY2 is a component of the deubiquitination module of the Arabidopsis SAGA transcriptional co-activator complex but not of the TREX-2 complex. *J Mol Biol*. 2018;430:1479–94.
- Allis CD, Berger SL, Cote J, Dent S, Jenuwien T, Kouzarides T, et al. New nomenclature for chromatin-modifying enzymes. *Cell*. 2007;131:633–6.
- Papai G, Frechard A, Kolesnikova O, Crucifix C, Schultz P, Ben-Shem A. Structure of SAGA and mechanism of TBP deposition on gene promoters. *Nature*. 2020;577:711–6.
- Wang H, Dienemann C, Stützer A, Urlaub H, Cheung ACM, Cramer P. Structure of the transcription coactivator SAGA. *Nature*. 2020;577:717–20.
- Atanassov BS, Mohan RD, Lan X, Kuang X, Lu Y, Lin K, et al. ATXN7L3 and ENY2 coordinate activity of multiple H2B deubiquitinases important for cellular proliferation and tumor growth. *Mol Cell*. 2016;62:558–71.
- Helmlinger D, Tora L. Sharing the SAGA. *Trends Biochem Sci*. 2017;42:850–61.
- Wang YL, Faiola F, Xu M, Pan S, Martinez E. Human ATAC Is a GCN5/PCAF-containing acetylase complex with a novel NC2-like histone fold module that interacts with the TATA-binding protein. *J Biol Chem*. 2008;283:33808–15.
- Riss A, Scheer E, Joint M, Trowitzsch S, Berger I, Tora L. Subunits of ADA-two-A-containing (ATAC) or Spt-Ada-Gcn5-acetyltransferase (SAGA) coactivator complexes enhance the acetyltransferase activity of GCN5. *J Biol Chem*. 2015;290:28997–9009.
- Eberharter A, Sterner DE, Schieltz D, Hassan A, Yates JR 3rd, Berger SL, et al. The ADA complex is a distinct histone acetyltransferase complex in *Saccharomyces cerevisiae*. *Mol Cell Biol*. 1999;19:6621–31.

29. Soffers JHM, Li X, Saraf A, Seidel CW, Florens L, Washburn MP, et al. Characterization of a metazoan ADA acetyltransferase complex. *Nucleic Acids Res.* 2019;47:3383–94.
30. 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. *Mol Cell Biol.* 2009;29:1649–60.
31. Lim S, Kwak J, Kim M, Lee D. Separation of a functional deubiquitylating module from the SAGA complex by the proteasome regulatory particle. *Nat Commun.* 2013;4:2641.
32. Li X, Seidel CW, Szerszen LT, Lange JJ, Workman JL, Abmayr SM. Enzymatic modules of the SAGA chromatin-modifying complex play distinct roles in *Drosophila* gene expression and development. *Genes Dev.* 2017;31:1588–600.
33. Herbst DA, Esbin MN, Louder RK, Dugast-Darzacq C, Dailey GM, Fang Q, et al. Structure of the human SAGA coactivator complex: The divergent architecture of human SAGA allows modular coordination of transcription activation and co-transcriptional splicing. *bioRxiv.* 2021. <https://doi.org/10.1101/2021.02.08.430339>.
34. Panchy N, Lehti-Shiu M, Shiu SH. Evolution of gene duplication in plants. *Plant Physiol.* 2016;171:2294–316.
35. Pray-Grant MG, Schieltz D, McMahon SJ, Wood JM, Kennedy EL, Cook RG, et al. The novel SLIK histone acetyltransferase complex functions in the yeast retrograde response pathway. *Mol Cell Biol.* 2002;22:8774–86.
36. Sterner DE, Belotserkovskaya R, Berger SL. SALSA, a variant of yeast SAGA, contains truncated Spt7, which correlates with activated transcription. *Proc Natl Acad Sci USA.* 2002;99:11622–7.
37. Wu PY, Winston F. Analysis of Spt7 function in the *Saccharomyces cerevisiae* SAGA coactivator complex. *Mol Cell Biol.* 2002;22:5367–79.
38. Martinez E, Palhan VB, Tjernberg A, Lyman ES, Gamper AM, Kundu TK, et al. Human STAGA complex is a chromatin-acetylating transcription coactivator that interacts with pre-mRNA splicing and DNA damage-binding factors in vivo. *Mol Cell Biol.* 2001;21:6782–95.
39. Pray-Grant MG, Daniel JA, Schieltz D, Yates JR 3rd, Grant PA. Chd1 chromodomain links histone H3 methylation with SAGA- and SLIK-dependent acetylation. *Nature.* 2005;433:434–8.
40. Sun C. The SF3b complex: splicing and beyond. *Cell Mol Life Sci.* 2020;77:3583–95.
41. Keen JC, Moore HM. The genotype-tissue expression (GTEx) project: linking clinical data with molecular analysis to advance personalized medicine. *J Pers Med.* 2015;5:22–9.
42. Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, et al. Proteomics. Tissue-based map of the human proteome. *Science.* 2015;347:1260419.
43. Yu NY, Hallström BM, Fagerberg L, Ponten F, Kawaji H, Carninci P, et al. Complementing tissue characterization by integrating transcriptome profiling from the Human Protein Atlas and from the FANTOM5 consortium. *Nucleic Acids Res.* 2015;43:6787–98.
44. Bu P, Evrard YA, Lozano G, Dent SY. Loss of Gcn5 acetyltransferase activity leads to neural tube closure defects and exencephaly in mouse embryos. *Mol Cell Biol.* 2007;27:3405–16.
45. Maurice T, Duclot F, Meunier J, Naert G, Givalois L, Meffre J, et al. Altered memory capacities and response to stress in p300/CBP-associated factor (PCAF) histone acetylase knockout mice. *Neuropsychopharmacology.* 2008;33:1584–602.
46. Puri PL, Sartorelli V, Yang XJ, Hamamori Y, Ogryzko VV, Howard BH, et al. Differential roles of p300 and PCAF acetyltransferases in muscle differentiation. *Mol Cell.* 1997;1:35–45.
47. Sartorelli V, Puri PL, Hamamori Y, Ogryzko V, Chung G, Nakatani Y, et al. Acetylation of MyoD directed by PCAF is necessary for the execution of the muscle program. *Mol Cell.* 1999;4:725–34.
48. Santi S, Cenni V, Capanni C, Lattanzi G, Mattioli E. PCAF Involvement in Lamin A/C-HDAC2 interplay during the early phase of muscle differentiation. *Cells.* 2020;9:1735.
49. 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. *Nat Genet.* 2000;26:229–32.
50. Cortez JT, Montauti E, Shifrut E, Gatchalian J, Zhang Y, Shaked O, et al. CRISPR screen in regulatory T cells reveals modulators of Foxp3. *Nature.* 2020;582:416–20.
51. Loo CS, Gatchalian J, Liang Y, Leblanc M, Xie M, Ho J, et al. A Genome-wide CRISPR screen reveals a role for the non-canonical nucleosome-remodeling BAF complex in Foxp3 expression and regulatory T cell function. *Immunity.* 2020;53:143–57.e8.
52. Srivastava R, Rai KM, Pandey B, Singh SP, Sawant SV. Spt-Ada-Gcn5-acetyltransferase (SAGA) complex in plants: genome wide identification, evolutionary conservation and functional determination. *PLoS ONE.* 2015;10:e0134709.
53. Robinson SW, Herzyk P, Dow JA, Leader DP. FlyAtlas: database of gene expression in the tissues of *Drosophila melanogaster*. *Nucleic Acids Res.* 2013;41:D744–50.
54. Larkin A, Marygold SJ, Antonazzo G, Attrill H, Dos Santos G, Garapati PV, Goodman JL, Gramates LS, Millburn G, Strelets VB, Tabone CJ, Thurmond J, FlyBase Consortium. FlyBase: updates to the *Drosophila melanogaster* knowledge base. *Nucleic Acids Res.* 2021;49:D899–907.
55. Lindner M, Simonini S, Kooiker M, Gagliardini V, Somssich M, Hohenstatt M, et al. TAF13 interacts with PRC2 members and is essential for Arabidopsis seed development. *Dev Biol.* 2013;379:28–37.
56. Carre C, Szymczak D, Pidoux J, Antoniewski C. The histone H3 acetylase dGcn5 is a key player in *Drosophila melanogaster* metamorphosis. *Mol Cell Biol.* 2005;25:8228–38.
57. Luense LJ, Donahue G, Lin-Shiao E, Rangel R, Weller AH, Bartolomei MS, Berger SL. Gcn5-mediated histone acetylation governs nucleosome dynamics in spermiogenesis. *Dev Cell.* 2019;51:745–758.e6.
58. Koutelou E, Wang L, Schibler AC, Chao HP, Kuang X, Lin K, et al. USP22 controls multiple signaling pathways that are essential for vasculature formation in the mouse placenta. *Development.* 2019. <https://doi.org/10.1242/dev.174037>.
59. Xu W, Edmondson DG, Roth SY. Mammalian GCN5 and P/CAF acetyltransferases have homologous amino-terminal domains important for recognition of nucleosomal substrates. *Mol Cell Biol.* 1998;18:5659–69.
60. Howe L, Auston D, Grant P, John S, Cook RG, Workman JL, et al. Histone H3 specific acetyltransferases are essential for cell cycle progression. *Genes Dev.* 2001;15:3144–54.
61. Kim S, Ohkuni K, Couplan E, Jazwinski SM. The histone acetyltransferase GCN5 modulates the retrograde response and genome stability determining yeast longevity. *Biogerontology.* 2004;5:305–16.
62. Vlachonasis KE, Thomashow MF, Triezenberg SJ. Disruption mutations of ADA2b and GCN5 transcriptional adaptor genes dramatically affect Arabidopsis growth, development, and gene expression. *Plant Cell.* 2003;15:626–38.
63. Yamamoto T, Poon D, Weil PA, Horikoshi M. Molecular genetic elucidation of the tripartite structure of the *Schizosaccharomyces pombe* 72 kDa TFIID subunit which contains a WD40 structural motif. *Genes Cells.* 1997;2:245–54.
64. Mitsuzawa H, Ishihama A. Identification of histone H4-like TAF in *Schizosaccharomyces pombe* as a protein that interacts with WD repeat-containing TAF. *Nucleic Acids Res.* 2002;30:1952–8.
65. Helmlinger D, Marguerat S, Villén J, Swaney DL, Gygi SP, Bähler J, et al. Tra1 has specific regulatory roles, rather than global functions, within the SAGA co-activator complex. *EMBO J.* 2011;30:2843–52.
66. McCormick MA, Mason AG, Guyenet SJ, Dang W, Garza RM, Ting MK, et al. The SAGA histone deubiquitinase module controls yeast replicative lifespan via Sir2 interaction. *Cell Rep.* 2014;8:477–86.
67. Herczeg Z, Hulla W, Gell D, Cuenin C, Leonart M, Jackson S, et al. Disruption of Trrap causes early embryonic lethality and defects in cell cycle progression. *Nat Genet.* 2001;29:206–11.
68. Gause M, Eissenberg JC, Macrae AF, Dorsett M, Misulovin Z, Dorsett D. Nipped-A, the Tra1/TRRAP subunit of the *Drosophila* SAGA and Tip60 complexes, has multiple roles in Notch signaling during wing development. *Mol Cell Biol.* 2006;26:2347–59.
69. Grau B, Popescu C, Torroja L, Ortuño-Sahagún D, Boros I, Ferrús A. Transcriptional adaptor ADA3 of *Drosophila melanogaster* is required for histone modification, position effect variegation, and transcription. *Mol Cell Biol.* 2008;28:376–85.
70. 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. *EMBO J.* 2008;27:394–405.
71. Lin Z, Yang H, Kong Q, Li J, Lee SM, Gao B, et al. USP22 antagonizes p53 transcriptional activation by deubiquitinating Sirt1 to suppress cell

- apoptosis and is required for mouse embryonic development. *Mol Cell*. 2012;46:484–94.
72. Mohibi S, Gurumurthy CB, Nag A, Wang J, Mirza S, Mian Y, et al. Mammalian alteration/deficiency in activation 3 (Ada3) is essential for embryonic development and cell cycle progression. *J Biol Chem*. 2012;287:29442–56.
 73. Wang F, El-Saafin F, Ye T, Stierle M, Negroni L, Durik M, et al. Histone H2Bub1 deubiquitylation is essential for mouse development, but does not regulate global RNA polymerase II transcription. *Cell Death Differ*. 2021. <https://doi.org/10.1038/s41418-021-00759-2>.
 74. 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. *Mol Cell Biol*. 2005;25:8215–27.
 75. Vamos EE, Boros IM. The C-terminal domains of ADA2 proteins determine selective incorporation into GCN5-containing complexes that target histone H3 or H4 for acetylation. *FEBS Lett*. 2012;586:3279–86.
 76. Welihinda AA, Tirasophon W, Kaufman RJ. The cellular response to protein misfolding in the endoplasmic reticulum. *Gene Expr*. 1999;7:293–300.
 77. Proft M, Struhl K. Hog1 kinase converts the Sko1-Cyc8-Tup1 repressor complex into an activator that recruits SAGA and SWI/SNF in response to osmotic stress. *Mol Cell*. 2002;9:1307–17.
 78. Biddick RK, Law GL, Chin KK, Young ET. The transcriptional coactivators SAGA, SWI/SNF, and mediator make distinct contributions to activation of glucose-repressed genes. *J Biol Chem*. 2008;283:33101–9.
 79. Moraga F, Aquea F. Composition of the SAGA complex in plants and its role in controlling gene expression in response to abiotic stresses. *Front Plant Sci*. 2015;6:865.
 80. Hu Z, Song N, Zheng M, Liu X, Liu Z, Xing J, et al. Histone acetyltransferase GCN5 is essential for heat stress-responsive gene activation and thermotolerance in *Arabidopsis*. *Plant J*. 2015;84:1178–91.
 81. Kaldis A, Tsementzi D, Tanriverdi O, Vlachonasiou KE. *Arabidopsis thaliana* transcriptional co-activators ADA2b and SGF29a are implicated in salt stress responses. *Planta*. 2011;233:749–62.
 82. Zheng M, Liu X, Lin J, Liu X, Wang Z, Xin M, et al. Histone acetyltransferase GCN5 contributes to cell wall integrity and salt stress tolerance by altering the expression of cellulose synthesis genes. *Plant J*. 2019;97:587–602.
 83. Li S, Lin YJ, Wang P, Zhang B, Li M, Chen S, et al. The AREB1 transcription factor influences histone acetylation to regulate drought responses and tolerance in *Populus trichocarpa*. *Plant Cell*. 2019;31:663–86.
 84. Lerin C, Rodgers JT, Kalume DE, Kim SH, Pandey A, Puigserver P. GCN5 acetyltransferase complex controls glucose metabolism through transcriptional repression of PGC-1 α . *Cell Metab*. 2006;3:429–38.
 85. Kelly TJ, Lerin C, Haas W, Gygi SP, Puigserver P. GCN5-mediated transcriptional control of the metabolic coactivator PGC-1 β through lysine acetylation. *J Biol Chem*. 2009;284:19945–52.
 86. Bararia D, Kwok HS, Welner RS, Numata A, Sárosi MB, Yang H, et al. Acetylation of C/EBP α inhibits its granulopoietic function. *Nat Commun*. 2016;7:10968.
 87. Gao B, Kong Q, Zhang Y, Yun C, Dent SYR, Song J, et al. The Histone acetyltransferase Gcn5 positively regulates T cell activation. *J Immunol*. 2017;198:3927–38.
 88. Wang Y, Yun C, Gao B, Xu Y, Zhang Y, Wang Y, et al. The lysine acetyltransferase GCN5 is required for iNKT cell development through EGR2 acetylation. *Cell Rep*. 2017;20:600–12.
 89. Liu Y, Bao C, Wang L, Han R, Beier UH, Akimova T, et al. Complementary roles of GCN5 and PCAF in Foxp3+ T-regulatory cells. *Cancers*. 2019;11:554.
 90. Jin Q, Zhuang L, Lai B, Wang C, Li W, Dolan B, et al. Gcn5 and PCAF negatively regulate interferon- β production through HAT-independent inhibition of TBK1. *EMBO Rep*. 2014;15:1192–201.
 91. Martínez-Cerdeño V, Lemen JM, Chan V, Wey A, Lin W, Dent SR, et al. N-Myc and GCN5 regulate significantly overlapping transcriptional programs in neural stem cells. *PLoS ONE*. 2012;7:e39456.
 92. Kobayashi T, Iwamoto Y, Takashima K, Isomura A, Kosodo Y, Kawakami K, et al. Deubiquitinating enzymes regulate Hes1 stability and neuronal differentiation. *FEBS J*. 2015;282:2475–87.
 93. Palhan VB, Chen S, Peng GH, Tjernberg A, Gamper AM, Fan Y, et al. Polyglutamine-expanded ataxin-7 inhibits STAGA histone acetyltransferase activity to produce retinal degeneration. *Proc Natl Acad Sci USA*. 2005;102:8472–7.
 94. Lan X, Koutelou E, Schibler AC, Chen YC, Grant PA, Dent SY. Poly(Q) expansions in ATXN7 affect solubility but not activity of the SAGA deubiquitinating module. *Mol Cell Biol*. 2015;35:1777–87.
 95. Helmlinger D, Hardy S, Abou-Sleymane G, Eberlin A, Bowman AB, Gansmüller A, et al. Glutamine-expanded ataxin-7 alters TFTC/STAGA recruitment and chromatin structure leading to photoreceptor dysfunction. *PLoS Biol*. 2006;4:e67.
 96. Chen YC, Gatchel JR, Lewis RW, Mao CA, Grant PA, Zoghbi HY, et al. Gcn5 loss-of-function accelerates cerebellar and retinal degeneration in a SCA7 mouse model. *Hum Mol Genet*. 2012;21:394–405.
 97. Tzelepis K, Koike-Yusa H, De Braekeleer E, Li Y, Metzakopiani E, Dovey OM, et al. A CRISPR dropout screen identifies genetic vulnerabilities and therapeutic targets in acute myeloid leukemia. *Cell Rep*. 2016;17:1193–205.
 98. Durbin AD, Zimmerman MW, Dharia NV, Abraham BJ, Iniguez AB, Weichert-Leahey N, et al. Selective gene dependencies in MYCN-amplified neuroblastoma include the core transcriptional regulatory circuitry. *Nat Genet*. 2018;50:1240–6.
 99. Cancer Dependency Map (DepMap) at Broad Institute. 2018. <https://dep-map.org/portal/>. Accessed 10 Jan 2021.
 100. DepMap at Sanger Institute. 2019. <https://depmap.sanger.ac.uk/>. Accessed 10 Jan 2021.
 101. Patel JH, Du Y, Ard PG, Phillips C, Carella B, Chen CJ, et al. The c-MYC oncoprotein is a substrate of the acetyltransferases hGCN5/PCAF and TIP60. *Mol Cell Biol*. 2004;24:1082634.
 102. Hirsch CL, Coban Akdemir Z, Wang L, Jayakumaran G, Trcka D, Weiss A, et al. Myc and SAGA rewire an alternative splicing network during early somatic cell reprogramming. *Genes Dev*. 2015;29:803–16.
 103. McMahon SB, Wood MA, Cole MD. The essential cofactor TRRAP recruits the histone acetyltransferase hGCN5 to c-Myc. *Mol Cell Biol*. 2000;20:556–62.
 104. Chen L, Wei T, Si X, Wang Q, Li Y, Leng Y, et al. Lysine acetyltransferase GCN5 potentiates the growth of non-small cell lung cancer via promotion of E2F1, cyclin D1, and cyclin E1 expression. *J Biol Chem*. 2013;288:14510–21.
 105. Yin YW, Jin HJ, Zhao W, Gao B, Fang J, Wei J, et al. The histone acetyltransferase GCN5 expression is elevated and regulated by c-Myc and E2F1 transcription factors in human colon cancer. *Gene Expr*. 2015;16:187–96.
 106. Majaz S, Tong Z, Peng K, Wang W, Ren W, Li M, et al. Histone acetyltransferase GCN5 promotes human hepatocellular carcinoma progression by enhancing ALB1 expression. *Cell Biosci*. 2016;6:47.
 107. Mustachio LM, Roszik J, Farria AT, Guerra K, Dent SY. Repression of GCN5 expression or activity attenuates c-MYC expression in non-small cell lung cancer. *Am J Cancer Res*. 2019;9:1830–45.
 108. Farria AT, Mustachio LM, Akdemir ZHC, Dent SYR. GCN5 HAT inhibition reduces human Burkitt lymphoma cell survival through reduction of MYC target gene expression and impeding BCR signaling pathways. *Oncotarget*. 2019;10:5847–58.
 109. Farria AT, Plummer JB, Salinger AP, Shen J, Lin K, Lu Y, et al. Transcriptional Activation of MYC-Induced Genes by GCN5 Promotes B-cell Lymphomagenesis. *Cancer Res*. 2020;80:5543–53.
 110. Griffin NI, Sharma G, Zhao X, Mirza S, Srivastava S, Dave BJ, et al. ADA3 regulates normal and tumor mammary epithelial cell proliferation through c-MYC. *Breast Cancer Res*. 2016;18:113.
 111. Glinesky GV, Berezovska O, Glinskii AB. Microarray analysis identifies a death-from-cancer signature predicting therapy failure in patients with multiple types of cancer. *J Clin Invest*. 2005;115:1503–21.
 112. Wan J, Zhan J, Li S, Ma J, Xu W, Liu C, et al. PCAF-primed EZH2 acetylation regulates its stability and promotes lung adenocarcinoma progression. *Nucleic Acids Res*. 2015;43:3591–604.
 113. Wang LT, Liu KY, Jeng WY, Chiang CM, Chai CY, Chiou SS, et al. PCAF-mediated acetylation of ISX recruits BRD4 to promote epithelial-mesenchymal transition. *EMBO Rep*. 2020;21:e48795.
 114. McCann JJ, Vasilevska IA, Poudel Neupane N, Shafi AA, McNair C, Dylgjeri E, et al. USP22 functions as an oncogenic driver in prostate cancer by regulating cell proliferation and DNA repair. *Cancer Res*. 2020;80:430–43.
 115. Baptista T, Grünberg S, Minoungou N, Koster MJE, Timmers HTM, Hahn S, et al. SAGA is a general cofactor for RNA polymerase II transcription. *Mol Cell*. 2017;68:130–43.e5.
 116. 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 Dev*. 2011;25:1499–509.

Publisher's Note

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