MAPPING THE ROLE OF CHIFFON AND GCN5 HISTONE ACETYLTRANSFERASE COMPLEX IN DROSOPHILA

by

Eliana F. Torres-Zelada

A Dissertation

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Doctor of Philosophy

Department of Biochemistry

West Lafayette, Indiana

December 2021

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Vikki M. Weake

Department of Biochemistry

Dr. Ouriana Andrisani

Department of Basic Medical Sciences

Dr. Brian Dilkes

Department of Biochemistry

Dr. Mark Hall

Department of Biochemistry

Approved by:

Dr. Andrew Mesecar

To my parents Eliana Zelada Mázmela and Luis Torres Cabrera

ACKNOWLEDGMENTS

Foremost, I would like to express my sincere gratitude to my PhD advisor, Dr. Vikki Weake. Thank you for allowing me to be part of your lab and for your constant encouragement and mentoring during the last six years, as well as for always being such an enthusiastic and positive person and mentor. You guided my research in a way that allowed me think critically. This guidance, invaluable advice, continuous support, and opportunity allowed me to grow as a scientist and I am very grateful for this training as I move forward in my scientific career. You have been an amazing advisor and I wish to be like you one day, a strong woman in science. I would also like to thank my committee members, Dr. Ouriana Andrisani, Dr, Brian Dilkes, and Dr. Mark Hall, for your advice, feedback, and guidance through my degree.

Next, I would like to thank the all the former and current Weake lab members. I have learned a lot from everyone, and I am thankful for their constructive feedbacks and continuous support all the time. Special thanks to Dr. Hana Hall, who was such a prime part of my scientific career. You played a key role during the most challenging moments in this journey. I also want to thank Ben Anderson and Hannah Blum, two extraordinary undergraduates that have contributed significantly to my research.

I am extremely grateful to my family for their continuous and unparalleled love, help, and support. I am forever indebted to my parents for giving me the opportunities and experiences that have made me who I am. They encouraged me to explore new directions and challenges in my life and seek my own destiny. This journey would not have been possible if not for them.

Flying across the ocean from Peru to pursue PhD studies in the US, the people and community that are around me and share my cultural background, traditions, and personal goals have been very important for both my mental health and personal growth. I would like to express my appreciation and love to all my Lafayette friends who helped me go through difficult, frustrating moments, and who have always been there through the ups and downs of every aspect of my life during my stay at Purdue. It is you that make Lafayette a place full of sweet memories.

I would also like to say thank you to the love of my life, partner, and biggest supporter, José Rivas Padilla. Thank you for believing in me, even in the moments that I did not; for being an incredible supporter; and for your patience and company.

Finally, I would like to thank all my funding sources and supporters: The Bird Stair Fellowship, the Purdue Research Foundation (PRF) Grant, and the Bilsland Dissertation Fellowship.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

aa	amino acid
ADA	Ada2/Gcn5/Ada3 transcription activator
APC	anaphase promoting complex
ATAC	Ada2a-containing
bp	base pair
Cdc7	cell division cycle 7
CHAT	Chiffon Histone Acetyltransferase
ChIP-seq	chromatin immunoprecipitation sequencing
DEGs	differential expressed genes
Dbf4	dumbbell former 4 protein
DNA	deoxyribonucleic acid
DDK	Dbf4-dependent kinase
DDO	double-dropout media (-leu -trp)
FDR	false discovery rate
FL	full-length
Gcn5	general control nonrepressed 5
GO	gene ontology
HAT	histone acetyltransferase
H3K9ac	acetylated Lys-9 in histone H3
H3K14ac	acetylated Lys-14 in histone H3
H3K18ac	acetylated Lys-18 in histone H3
H3K23ac	acetylated Lys-23 in histone H3
H4K5ac	acetylated Lys-5 in histone H4
H4K12ac	acetylated Lys-12 in histone H4
H4K16ac	acetylated Lys-16 in histone H4
IP	immunoprecipitation
Mcm2-7	minichromosomal maintenance complex 2-7
MudPIT	multidimensional protein identification technology
PCA	principal component analysis

Pol II	RNA polymerase II
Pre-ORC	pre-origin recognition complex
QDO	quadruple-dropout media (-leu -trp -his -ade)
RNA-seq	RNA sequencing
rpm	revolutions per minute
RRPM	Reference adjusted Reads Per Million
SAGA	Spt-Ada-Gcn5 acetyltransferase
TSS	transcription start site
Y2H	yeast-two hybrid
YPD	yeast extract-peptone dextrose

ABSTRACT

The histone acetyltransferase (HAT) Gcn5 was first characterized in yeast, and is conserved throughout eukaryotes where it functions as part of large multi-subunit transcriptional coactivator complexes. *Drosophila melanogaster* contains four Gcn5-containing complexes: the Spt-Ada-Gcn5 Acetyltransferase (SAGA), Ada2a-containing (ATAC), Ada2/Gcn5/Ada3 transcription activator (ADA), and Chiffon Histone Acetyltransferase (CHAT) complexes. Each of these Gcn5 complexes is nucleated by unique Ada2 homologs (Ada2a or Ada2b) or splice isoforms (Ada2b-PB or Ada2b-PA) that share conserved N-terminal domains, and differ only in their C-terminal domains. Whereas the SAGA and ADA complexes are also present in the yeast *Saccharomyces cerevisiae*, ATAC has only been identified in other metazoa such as humans, and the CHAT complex appears to be unique to insects. In Chapter 1 I highlight key studies in fruit flies that have provided insight into the essential roles played by Gcn5 in the development of multicellular organisms. I outline the composition and activity of the four different Gcn5-containing complexes in *Drosophila*, describing key subunits that convey differences to these complexes in terms of their targeting, activity, and biological roles.

In Chapter 2, I describe the first identification and characterization of the insect-specific Gcn5-containing complex, CHAT. *Drosophila* has two splice isoforms of Ada2b paralog. Mass spectrometry proteomic studies reveal that only the Ada2b-PB isoform is present in SAGA; in contrast, the Ada2b-PA isoform associates with the Gcn5 HAT core subunits and Chiffon, the fly ortholog of Dbf4, forming the CHAT complex. I present our findings that CHAT is essential for both histone acetylation and fly viability through its CHAT-specific subunit, Chiffon. Chiffon is required for the specialized form of DNA replication, endoreplication, however it is not required for mitotic replication. Our mass spectrometry and genetics studies demonstrate that Chiffon interacts in a mutually exclusive manner with Cdc7 and Gcn5. Whereas the N-terminal domain of Chiffon interacts with Cdc7 (cell division cycle 7) forming the DDK complex, the C-terminal domain binds to Gcn5 nucleating the formation of CHAT. This studies also demonstrate that both complexes function independently in DNA replication and histone acetylation, respectively. Expression of the C-terminal domain, which partially rescues histone acetylation, also restores fly viability, suggesting that the essential function of Chiffon relates to its histone acetyltransferase activity rather than Cdc7 activation. I present data that reveal that both Chiffon FL that contain

premature stop codons on either position 174 (Chiffon FL^{WF24}) or 376 (Chiffon FL^{*}), can restore fly viability, and surprisingly a ~48kDa product can be detected when both Chiffon FL^{WF24} and Chiffon FL^{*} were immunoprecipitated via their C-terminal FLAG epitopes tags. This data supports the hypothesis that c*hiffon* is a dicistronic gene that encodes two distinct polypeptides from alternative translation start sites, generating separate DDK and CHAT complexes.

In Chapter 3, I explore the role of CHAT in regulating gene expression in *Drosophila* embryos. I show that the *Drosophila* Gcn5-containing complexes, SAGA/ADA and CHAT, have largely redundant roles in embryonic gene expression. However, when comparing RNA-sequencing (RNA-seq) of *chiffon* mutant and *ada2b* mutant embryos, the studies show that there is a little overlap between the genes disrupted when the Ada2b or Chiffon subunits of CHAT are disrupted. Moreover, our findings show that Chiffon is required for global H3K14ac in embryos far beyond that deposited by other Gcn5/Ada2b-containing complexes. This data suggests a model in which CHAT functions as a pioneer coactivator complex during embryogenesis that is necessary for the recruitment or activity of other HAT complexes. Finally, immunostaining studies revealed that there is a temporal switch between the expression of the *chiffon* gene products during a short window during early embryogenesis, in which the polypeptide product that nucleates formation of CHAT is expressed just prior to the wave of RNA polymerase II recruitment to chromatin in the early embryo. These studies provide further insight into the biological function of CHAT, and suggest that CHAT could have a key role in triggering the wave of early histone acetylation that stimulates zygotic genome activation.

In Chapter 4, I describe additional experiments that have not yet been submitted for publications that provide more information about the CHAT complex. These data include yeast two-hybrid studies examining the conservation in CHAT subunit interactions in other insects, and a yeast two-hybrid screening to identify novel interactors for *Drosophila* Cdc7. I also give more details about Chiffon, examine if its levels fluctuate during the cell cycle at either the mRNA/= or protein level, and test if it is a substrate of the Anaphase promoting complex (APC). Finally, I provide details of new reagents that we have generated, and describe how these can be used to provide more insight on the role of CHAT in gene expression.

STATEMENT OF PUBLISHED WORK

Chapter 1 is a pre-copyedited review that form part of the *Special Issue* in *BBA Gene Regulatory Mechanisms*, *Gcn5: the quintessential HAT*, edited by Vikki Weake. The article was written by Eliana F. Torres-Zelada with editorial contribution from Vikki Weake. The version of record,

Eliana F. Torres-Zelada, Vikki M. Weake (2020). The Gcn5 complexes in *Drosophila* as a model for metazoan. BBA Gene Regulatory Mechanisms, is available online at https://www.sciencedirect.com/science/article/abs/pii/S1874939920301930?via%3Dihub.

Chapter 2 is a pre-copyedited of an article published in *Journal of Cell Science*. The version of record, E.F. Torres-Zelada, R.E. Stephenson, A. Alpsoy. B.D. Anderson, S.K. Swanson, L. Florens, E.C. Dykhuizen, M.P. Washburn, V.M. Weake (2019). The *Drosophila* Dbf4 ortholog Chiffon forms a complex with Gcn5 that is necessary for histone acetylation and viability. *Journal of Cell Science*, is available online at https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6362396/. Chapter 3 is a pre-copyedited manuscript which has been submitted for publication. The article was written by Eliana F. Torres-Zelada with editorial contribution from Vikki Weake.

CHAPTER 1. INTRODUCTION: THE GCN5 COMPLEXES IN *DROSOPHILA* AS A MODEL FOR METAZOA

1.1 Introduction

Chromatin provides a barrier to processes that require access to the underlying DNA such as transcription and replication (Brownell and Allis 1996; Li et al. 2007). The nucleosome is the repeating unit of chromatin, and is composed of a heterotetramer of histones H3 and H4 flanked by two histone H2A/H2B heterodimers (Brownell and Allis 1996; Li et al. 2007). Histones can be post-translationally modified, predominantly on the N-terminal tails of the histone proteins (Brownell and Allis 1996; Li et al. 2007). These histone marks provide binding sites for other proteins that "read" these post-translational modifications, and can also potentially alter the nucleosome structure (Brownell and Allis 1996; Allis and Jenuwein 2016). One of the first and most well studied histone modifications is acetylation, whereby an acetyl group is added to lysine residues often on the N-terminal tails of histones H3 and H4 (Brownell and Allis 1996). Histone acetylation is generally associated with increased DNA accessibility because it stimulates chromatin remodeling (Gregory et al. 2001). Thus, histone acetylation usually correlates with, and contributes to active transcription (Brownell and Allis 1996; Wolffe 1994). Gcn5 was the first nuclear histone acetyltransferase (HAT) identified, first in *Tetrahymena thermophila* as described in (Brownell and Allis 2021), and subsequently in the yeast *Saccharomyces cerevisiae* as outlined by (Grant et al. 2021). Gcn5 has since been characterized in a wide range of eukaryotes (Koutelou et al. 2021; Grasser et al. 2021), and here we focus on how studies in the model insect species *Drosophila melanogaster* have provided insight into the expanded biological roles for Gcn5 in multicellular eukaryotes.

The fruit fly *Drosophila* has been used as a model organism extensively for genetic studies and developmental biology (Jennings 2011). The *Drosophila* genome shares 60% homology with humans, and about 75% of the genes responsible for human diseases have homologs in flies (Wangler et al. 2015; Yamamoto et al. 2014). Moreover, *Drosophila* possess homologs of nearly all of the key factors involved in chromatin modification and transcription, making the fruit fly a powerful model organism for studying chromatin biology (Karnay and Elefant 2017). In contrast to mammals, which often possess multiple paralogs of histone modifying enzymes, *Drosophila* usually encodes only a single gene for different histone modifying enzymes, providing a simpler

genetic system in which to dissect biological functions of various chromatin-based processes (Saravanan et al. 2019). The *Drosophila* life cycle takes place within 10 days under standard laboratory conditions, beginning with the hatching of an egg into a larval stage, followed by several larval molts, formation of a pupa, metamorphosis (transformation from an immature, larval form to the adult fly), and finally eclosion (emergence from the pupal case) into an adult fly (Jennings 2011). In this review, we provide a historical perspective on the identification of the *Drosophila* Gcn5 complexes. First, we describe the composition of the Gcn5 complexes in *Drosophila* in comparison to the orthologous complexes in *S. cerevisiae* and human cells. Next, we outline the essential subunits for fly development based on studies in mutant flies, and describe a subset of representative mutant phenotypes that highlight specialized functions for each Gcn5 complex in development. Finally, we describe the genome-wide localization patterns and biochemical roles for each Gcn5 complex in gene expression and other chromatin-based processes, where this has been defined. Last, we briefly discuss Gcn5 complexes in other insect species, and provide an overview of outstanding questions for future studies.

1.2 Gcn5 and its associated protein partners are conserved in *Drosophila.*

Immediately after the discovery that the *Tetrahymena* p55 HAT corresponded to *S. cerevisiae* Gcn5, it became clear that Gcn5 was conserved throughout eukaryotes: "*it seems likely that the yeast 55-kDa polypeptide is conserved across a wide range of eukaryotes*" (Brownell and Allis 1995; Brownell et al. 1996). Indeed, only three years later, the Allis group identified Gcn5 in the model insect species, *Drosophila* (Smith et al. 1998). In contrast to humans, who possess two Gcn5 paralogs (PCAF and Gcn5) (Candau et al. 1996; Yang et al. 1996), there is only one Gcn5 homolog in *Drosophila:* Gcn5 (FBgn0020388, CG4107). The gene encoding Gcn5 was historically named *pcaf* in flies (Carré et al. 2005), but it has since been renamed *gcn5* on FlyBase and in much of the recent literature; we refer to this gene as *gcn5* throughout this review. Gcn5 shares the domains that are common to all Gcn5 homologs including its HAT catalytic domain (469 - 634aa), Gcn5-N-Acetyltransferase (GNAT) domain (514 - 598aa), and bromodomain (717 - 795aa) (Smith et al. 1998), which binds acetylated lysine (Fig. 1.1) (Dhalluin et al. 1999). However, *Drosophila* Gcn5 shares higher similarity with both of the human Gcn5 paralogs than with *S. cerevisiae* Gcn5 (yGcn5). Moreover, Gcn5 contains a conserved N-terminal domain that is only found in metazoan Gcn5 homologs like human Gcn5 and PCAF (Smith et al. 1998). It has

been suggested that this N-terminal PCAF domain in human PCAF has E3 ubiquitin ligase activity (Linares et al. 2007), but this activity has not been demonstrated for human or *Drosophila* Gcn5 .

In all organisms, Gcn5 associates with other proteins that are critical for both its activity and targeting. Although both *S. cerevisiae* and *Drosophila* Gcn5 can acetylate free histone H3 *in vitro*, they are unable to acetylate nucleosomal substrates on their own (Grant et al. 1997; Smith et al. 1998). This lack of nucleosomal acetyltransferase activity is in contrast to human PCAF, which has been shown to acetylate nucleosomal substrates *in vitro* (Yang et al. 1996), and shares substantial homology with *Drosophila* Gcn5 (Fig. 1.1). In *S. cerevisiae*, Gcn5 associates tightly with two other proteins, Ada2 and Ada3, forming an heterotrimeric complex *in vitro* (Georgakopoulos et al. 1995; Horiuchi et al. 1995). A third Gcn5-interacting protein, Sgf29, was later identified in *S. cerevisiae* (Sanders et al. 2002; Lee et al. 2011). Together, Gcn5, Ada2, Ada3, and Sgf29 constitute the core Gcn5 HAT module that is sufficient for nucleosomal histone acetylation (Grant et al. 1997) (Espinola-Lopez and Tan 2021) *Drosophila*, like *S. cerevisiae*, has single homologs of Ada3 (FBgn0030891, CG7098) and Sgf29 (FBgn0050390, CG30390), which were readily identified by sequence comparisons with the *S. cerevisiae* proteins (Table 1.1) (Guelman et al. 2006a; Kusch et al. 2003; Muratoglu et al. 2003). In contrast, there are two paralogs of Ada2 in *Drosophila*: Ada2a (FBgn0263738, CG43663) and Ada2b (FBgn0037555, CG9638) (Kusch et al. 2003; Muratoglu et al. 2003; Qi et al. 2004). Both Ada2a and Ada2b share a similar domain structure to *S. cerevisiae* Ada2, possessing conserved ZZ and SANT domains (Fig. 1.2) (Kusch et al. 2003; Muratoglu et al. 2003; Qi et al. 2004). Ada2a also contains a C-terminal SWIRM domain that is present in *S. cerevisiae* Ada2 and human Ada2a and Ada2b (Gamper et al. 2009). In addition, there are two splice isoforms of Ada2b resulting from alternative usage of splice acceptor sites in the third exon: Ada2b-PA encoding a 418-aa protein (FBppp0081303, also referred to as Ada2bS) and Ada2b-PB encoding a 555-aa protein (FBppp0099776, also referred to as Ada2bL) (Qi et al. 2004). These Ada2b splice isoforms are expressed at equivalent levels during different developmental stages in flies, and share both the ZZ and SANT domains, differing only in their C-terminal regions (Fig. 1.2) (Qi et al. 2004). The longer Ada2b-PB isoform contains the SWIRM domain in its unique C-terminal region, while the Ada2b-PA isoform lacks this domain (Fig. 1.2). In flies, like in *S. cerevisiae*, the nucleosomal HAT activity of Gcn5 requires interactions with either Ada2a or Ada2b, and Ada3 (Grant et al. 1997; Kusch et al. 2003; Smith et al. 1998). Both Ada2 paralogs, Ada2a and Ada2b, are conserved in other multicellular eukaryotes including

Arabidopsis and humans (Muratoglu et al. 2003), providing an early hint that Gcn5's association with other proteins might expand its biological role in multicellular organisms.

1.3 *Drosophila* Ada2 proteins nucleate formation of distinct Gcn5 complexes

Gcn5 resides within three different multi-subunit complexes in *S. cerevisiae*: the large, highly similar Spt-Ada-Gcn5 Acetyltransferase (SAGA) and SAGA-like (SLIK/SALSA) complexes, and the small Ada2/Gcn5/Ada3 transcription activator (ADA) complex (Grant et al. 1997; Eberharter et al. 1999; Sterner et al. 2002; Grant et al. 2021). The presence of multiple versions of Ada2 in flies suggested that Gcn5 might reside within additional complexes in *Drosophila*, and raised the question as to which version of Ada2 was present in each complex. During the first decade of the twenty first century, a series of studies led by the Boros and Workman groups revealed the existence of two large multi-subunit Gcn5 complexes in flies. In *Drosophila*, the Ada2b paralog (specifically the Ada2b-PB isoform) is present in the SAGA complex, similar to that found in *S. cerevisiae* (Kusch et al. 2003; Muratoglu et al. 2003; Weake et al. 2009). In contrast, Ada2a resides within a Gcn5 complex that is not present in *S. cerevisiae*, the Ada2a-containing (ATAC) complex (Guelman et al. 2006a). ATAC was first identified in *Drosophila*, and it has since been characterized in mammalian cells and appears to be widely conserved in multicellular eukaryotes (Spedale et al. 2012). More recently, an ADA-like complex was also identified in *Drosophila* (Soffers et al. 2019), together with an insect-specific Gcn5 complex that contains the shorter Ada2b-PA splice isoform (Torres-Zelada et al. 2019). With the exception of the small ADA complex, which contains only the core HAT subunits, the *Drosophila* Gcn5 complexes each possess additional protein subunits that contribute to their unique biological activities.

1.3.1 SAGA

The first of the *Drosophila* Gcn5 complexes to be identified, SAGA, is a large 2 MDa complex that contains 20 different protein subunits (Kusch et al. 2003; Muratoglu et al. 2003). SAGA has been well characterized in *S. cerevisiae* where its subunits were historically first organized into four major modules: the HAT module (Gcn5, Ada2, Ada3, Sgf29), a deubiquitination module (DUB; Ubp8, Sus1, Sgf11, and Sgf73), the TATA binding protein-

Associated Factor (TAF) module (Taf5, Taf6, Taf9, Taf10, and Taf12), and the Suppressor of Ty's (SPT) module (Ada1, Spt3, Spt7, Spt8, and Spt20), with Tra1 originally being classified as a Spt protein, although it was not identified in the original genetic screen (Winston et al. 1984) . More recent structural studies have resulted in a re-organization of the subunits in the TAF and SPT modules into a structural core (Taf5, Taf6, Taf9, Taf10, Taf12, Spt7, Ada1, and Spt20), a TATA binding protein (TBP) binding module (Spt3 and Spt8), and a transcription factor (TF) binding module consisting only of Tra1 (Lee et al. 2011; Helmlinger et al. 2011, 2021). The composition of the HAT and DUB modules remains unchanged from the original modular organization. Ubp8 within the DUB module provides SAGA with a second histone modifying activity, catalyzing deubiquitination of monoubiquitinated histone H2B (H2Bub1) (Daniel et al. 2004; Henry et al. 2003; Cornelio-Parra et al. 2021). *Drosophila* SAGA (dSAGA) contains orthologs of all *S. cerevisiae* SAGA subunits with the exception of Spt8. In fact, no ortholog of the *Spt8* gene is present in the genome of any metazoan organism (Helmlinger and Tora 2017). A variant of SAGA, termed SLIK/SALSA, has been purified in *S. cerevisiae*, and contains a C-terminal truncated version of Spt7 and lacks Spt8 (Sterner et al. 2002). Although dSAGA, like SLIK/SALSA, lacks Spt8, the *in vivo* existence of SLIK/SALSA remains controversial because Spt7 can be cleaved at its C-terminus by the Pep4 protease *in vitro*, resulting in removal of the Spt8-binding domain and subsequent loss of Spt8 (Spedale et al. 2010). Moreover, the C-terminal domain that is absent from the SLIK/SALSA Spt7 variant is conserved in metazoan Spt7 (Martinez et al. 2001; Antonova et al. 2019), while the N-terminal bromodomain appears to be unique to *S. cerevisiae* Spt7 (Demény et al. 2007). Below, we outline the composition of each module of dSAGA, and describe the subunits that differ from their *S. cerevisiae* counterparts.

Although some dSAGA subunits could be identified by sequence similarity with their *S. cerevisiae* SAGA counterparts, mass spectrometry of affinity-purified SAGA complexes revealed incorporation of novel subunits that were not predicted by sequence comparisons with the *S. cerevisiae* SAGA components. The HAT module in dSAGA contains Gcn5, Ada3, and Sgf29, which are shared between all of the *Drosophila* Gcn5 complexes (Fig. 1.3) (Guelman et al. 2006a; Kusch et al. 2003; Muratoglu et al. 2003; Torres-Zelada et al. 2019; Soffers et al. 2019). Although initially both splice isoforms of Ada2b were presumed to be part of the SAGA complex, mass spectrometry of affinity-purified SAGA complexes demonstrated that only the longer Ada2b-PB splice isoform is part of the dSAGA HAT module (Weake et al. 2009; Torres-Zelada et al. 2019).

Orthologs of all four DUB module subunits are also present in flies. The histone deubiquitinase Ubp8 in *S. cerevisiae* corresponds to Nonstop in flies (FBgn0013717, CG4166), which was originally named for the axon targeting defect observed in *nonstop* mutants during neuronal development (Poeck et al. 2001). Both Nonstop and Sgf11 (FBgn0036804, CG13379) are necessary for deubiquitination of H2Bub1 in flies (Weake et al. 2008). The last two DUB module subunits in flies are Ataxin 7 (FBgn0031420, CG9866, the ySgf73 ortholog) and E(y)2 (FBgn0000617, CG6474, the ySus1 ortholog) (Table 1.1) (Georgieva et al. 2001; Mohan et al. 2014a; Rodriguez-Navarro et al. 2004). In flies, like *S. cerevisiae*, several of the structural core subunits are shared between the transcription coactivator complex Transcription Factor II D (TFIID) and SAGA, namely E(y)1 (Taf9, FBgn0000617, CG6474), Taf10b (FBgn0026324, CG3069), and Taf12 (FBgn0011290, CG17358) (Helmlinger and Tora 2017; Weake et al. 2009). However, other structural core subunits are unique to dSAGA and are not present in *Drosophila* TFIID. For example, the TAF5-like Wda (will decrease acetylation, FBgn0039067, CG4448), and TAF6-like Saf6 (SAGA factor-like TAF6, FBgn0031281, CG3883), are specialized TAF paralogs that are present in SAGA but not in TFIID (Weake et al. 2009; Guelman et al. 2006b). Similar specialization of TAF proteins has occurred in other metazoan organisms with incorporation of Taf5-like and Taf6-like subunits in mammalian SAGA (Helmlinger and Tora 2017; Timmers 2021). Other SAGA subunits in flies are much more conserved, although with considerable variation in some dSAGA subunits at the sequence level compared to their *S. cerevisiae* counterparts. For example, although Tra1 (Nipped-A, FBgn0053554, CG33554), Ada1 (FBgn0051866, CG31866), Spt3 (FBgn0037981, CG3169), and Spt7 (FBgn0030874, CG6506) were readily identified by sequence comparison with *S. cerevisiae* (Guelman et al. 2006b), Spt20 (FBgn0036374, CG17689) was not identified in flies until mass spectrometry of purified SAGA revealed the presence of this subunit (Weake et al. 2009). Tra1/Nipped-A is the largest SAGA subunit (411 kDa in flies) and is shared with another transcriptional coactivator complex, the Tat interactive complex 60 kDa (TIP60; also known as the Nucleosome Acetyltransferase of H4 (NuA4) complex), which also possesses HAT activity (Doyon et al. 2004; Helmlinger and Tora 2017). In contrast to the SAGA HAT module, which preferentially acetylates histone H3, TIP60/NuA4 acetylates histone H4 and H2A.Z (Allard et al. 1999; Babiarz et al. 2006). Last, SAGA contains two spliceosomal proteins, Sf3b3 (FBgn0035162, CG13900) and Sf3b5 (FBgn0040534, CG11985), that are not present in *S. cerevisiae* SAGA despite the existence of *S.*

cerevisiae homologs corresponding to these proteins (Rse1 and Ysf3) (Stegeman et al. 2016). Sf3b3 and Sf3b5 are shared with the Sf3b complex, a component of the U2 small nuclear ribonucleoprotein (snRNP), which recognizes the branch point sequence to facilitate spliceosome assembly (Gozani et al. 1996; Herold et al. 2009). Both of these spliceosomal proteins are also present in hSAGA (Table 1.1) (Helmlinger and Tora 2017; Martinez et al. 2001; Vermeulen et al. 2010). Overall, *Drosophila* SAGA resembles the human SAGA complex more closely than either of the *S. cerevisiae* SAGA or SLIK/SALSA complexes, possessing similar specialized Taf-like proteins and containing the two additional spliceosomal proteins. The presence of these additional subunits in the metazoan SAGA complex suggests that SAGA may have gained more specialized roles in gene expression in animals compared to *S. cerevisiae*.

Several recent studies have investigated the structure of SAGA, and have provided insight into how each SAGA subunit integrates into the complex as a whole. These studies are described in more depth in (Helmlinger et al. 2021), but are briefly described here to provide context for understanding the organization of *Drosophila* SAGA. In *S. cerevisiae*, Cryogenic Electron Microscopy (cryoEM) data revealed a existence of a central module containing the structural core and the TBP binding module subunits that forms flexible connections to the HAT and DUB modules, while the large Tra1 subunit exists as a separate module that can bind the activation domain of transcription factors (Wang et al. 2020a; Papai et al. 2020; Díaz-Santín et al. 2017; Sharov et al. 2017). In *S. cerevisiae*, the HAT module is anchored to SAGA by Ada3 binding to Taf6, and HAT module subunits are lost from SAGA when it was purified from Ada3 or Ada2 mutant *S. cerevisiae* (Lee et al. 2011; Papai et al. 2020; Han et al. 2014). Similarly, Sgf73 (Ataxin 7) anchors the DUB module to the *S. cerevisiae* SAGA complex, and DUB subunits are lost from SAGA purified from Sgf73 mutant *S. cerevisiae* (Kohler et al. 2008; Lee et al. 2009). The DUB module requires Sgf73 for activity in *S. cerevisiae* (Lim et al. 2013), but in flies and plants, an enzymatically active DUB module can exist in the absence of the Sgf73 ortholog Ataxin 7 (Mohan et al. 2014a; Nassrallah et al. 2018). Notably, there is no ortholog of Sgf73/Ataxin 7 in *Arabidopsis*, suggesting that the DUB module may function independent of SAGA as the major H2Bub1 deubiquitinase (Nassrallah et al. 2018; Grasser et al. 2021). In human cells, the protease Caspase 7 has been shown to cleave ATXN7, which could potentially release a free DUB module from SAGA (Guyenet et al. 2015). This mechanism may also exist in *Drosophila*, although it has not yet been demonstrated. Thus, an open question remains as to whether the biological functions

attributed to the DUB module subunit in flies (see section 1.6) are due to its role in SAGA or represent its independent activity. These questions are discussed further by Mohan et al. in (Cornelio-Parra et al. 2021)

1.3.2 ADA

Recently, the Workman group have also identified an ADA-like complex in flies (Soffers et al. 2019)In S. *cerevisiae*, ADA contains the HAT module and two additional proteins, ADA HAT component 1 and 2 (Ahc1 and Ahc2) (Lee et al. 2011; Eberharter et al. 1999). Early biochemical studies suggested that an ADA complex might also exist in flies because a small Ada2b-containing complex was detected by glycerol gradients of Ada2b-containing complexes (Muratoglu et al. 2003). Indeed, recently Soffers *et al*. showed that there is an ADA complex in flies, which like SAGA contains the Ada2b-PB splice isoform (Soffers et al. 2019). In contrast to *S. cerevisiae* ADA, the *Drosophila* ADA complex does not contain subunits corresponding to *S. cerevisiae* Ahc1/2, which do not have sequence homologs in flies or humans (Table 1.2) (Soffers et al. 2019). Thus, the ADA complex in flies does not possess any unique subunits that can be used to genetically distinguish it from SAGA.

1.3.3 ATAC

In addition to SAGA and ADA, flies also have an additional 820 kDa multi-subunit Gcn5 complex that is nucleated by the Ada2 paralog Ada2a: ATAC. Size-exclusion chromatography of the *Drosophila* Gcn5 complexes provided an early hint that Ada2a and Ada2b resided in distinct complexes (Kusch et al. 2003; Muratoglu et al. 2003). Indeed, three years after the identification of the Ada2a paralog, the 13 subunit ATAC complex was first characterized in flies, providing the foundation for studies on this Gcn5 complex in other organisms (Guelman et al. 2006a). ATAC shares the core HAT module subunits (Gcn5, Ada3, and Sgf29) with SAGA. In addition to the HAT module subunits, nine ATAC-specific subunits exist in flies. Six of these ATAC subunits are also present in the mammalian ATAC complex: Atac1 (FBgn0031876, CG9200, human ZZZ3 ortholog), Atac2 (FBgn0032691, CG10414, the human CRBP2 ortholog), D12 (FBgn0027490, CG13400, the human YEATS2 ortholog), Mocs2B (FBgn0039280, CG10238, equivalent to both the human hMoaE and hMBIP proteins), NC2β (FBgn0028926, CG4185, the human NC2B

ortholog), and Wds (FBgn0040066, CG17437, the human WDR5 ortholog) (Table 1.3) (Suganuma et al. 2010, 2008). The human ortholog of Chrac-14 (FBgn0043002, CG13399) has been detected in some human ATAC purifications (Wang et al. 2008), but was absent from others (Guelman et al. 2009). In contrast, two of the *Drosophila* ATAC subunits, Atac3 (FBgn0052343, CG32343) and Hcf (FBgn0039904, CG1710), appear to be specific to the fly ATAC complex and have not been detected in human ATAC (Guelman et al. 2006a; Wang et al. 2008; Nagy et al. 2010). Like SAGA, *Drosophila* ATAC contains a second histone modifying activity. The Atac2 subunit of ATAC contains a HAT domain, and *Drosophila* Atac2 possesses HAT activity toward histone H4 and H2A *in* v*itro* and *in vivo* (Suganuma et al. 2008). However, the human counterpart for Atac2 (CRBP2) does not possess detectable HAT activity toward histone H4, suggesting that Gcn5 is the only active HAT within the human ATAC complex (Guelman et al. 2009; Nagy et al. 2010). Thus, *Drosophila* ATAC contains two distinct acetyltransferase enzymes: Gcn5 and Atac2 (Suganuma et al. 2008). Less is known about the modular organization and structure of the ATAC complex compared with SAGA. However, ATAC contains several histone-fold domain proteins, NC2β, D12 and Chrac-14, which may play a structural role in ATAC similar to that involving the structural core subunits in SAGA. While Chrac-14 and NC2β fail to form heterodimers, human YEATS2 (the *Drosophila* D12 ortholog) and NC2β interact via their histone-fold domains (Suganuma et al. 2008; Wang et al. 2008). In addition, both *Drosophila* Chrac-14 and NC2β have the ability to form homodimers (Suganuma et al. 2008). Wds also contains seven WD repeats, and this motif is often involved in protein–protein interactions (Fig. 1.3). In humans, YEATS2 (the *Drosophila* D12 ortholog) and Atac2 play a role in the integrity of the ATAC complex (Guelman et al. 2009; Wang et al. 2008), suggesting that these subunits, together with Wds, Chrac-14 and NC2β, may play a central role in structural organization within the ATAC complex. Like SAGA, several ATAC subunits are shared with other chromatin modifying complexes. For example Chrac-14, Hcf, and Wds are also subunits of the COMPASS-like methyltransferase complexes, which are responsible for the bulk of di- and tri-methylation at histone H3K4 in *Drosophila* (Mohan et al. 2011).

1.3.4 CHAT

Last, *Drosophila* possess a unique Gcn5 complex that appears to be specific to insects: the Chiffon Histone Acetyltransferase (CHAT) complex. Whereas the Ada2b-PB splice isoform is

present in SAGA, the shorter Ada2b-PA splice isoform is not part of the SAGA, ADA or ATAC complexes (Weake et al. 2009). Instead, Ada2b-PA nucleates formation of a fourth Gcn5 complex in flies that contains the shared HAT module subunits (Gcn5, Ada3, and Sgf29) together with a fifth protein, Chiffon (FBgn0000307, CG5813) (Table 1.4, Fig. 1.3). (Torres-Zelada et al. 2019). Chiffon is the *Drosophila* homolog of Dbf4, which binds and activates the Cdc7 kinase, forming the Dbf4-dependent kinase (DDK) complex (Tower 2004). DDK phosphorylates the Mcm2-7 helicase, activating the initial step in DNA replication (Pasero et al. 1999; Labib 2010; Stephenson et al. 2015). In contrast to SAGA and ATAC, the CHAT complex is unlikely to exist in *S. cerevisiae* or humans, because Dbf4 does not co-immunoprecipitate with Gcn5 in either of these organism (Torres-Zelada et al. 2019). Moreover, Chiffon interacts with directly with Gcn5 via its C-terminal domain, and this region of the protein is not conserved outside of insects (Torres-Zelada et al. 2019).

1.4 Substrate specificity of the Gcn5 complexes

In general, the *Drosophila* Gcn5 complexes preferentially acetylate histone H3 *in vitro* and *in vivo* exhibiting the highest activity on K9 and K14 of both recombinant histone H3 peptides and nucleosomal substrates (Guelman et al. 2006a; Pankotai et al. 2005; Torres-Zelada et al. 2019; Soffers et al. 2019). Although SAGA, ADA, and CHAT show this characteristic HAT activity toward histone H3, the presence of the second HAT in *Drosophila* ATAC expands its activity toward both histones H3 and H4 (Guelman et al. 2006a; Suganuma et al. 2008). In fact, *Drosophila* ATAC shows strong specificity for histone H4 in nucleosomal substrates *in vitro* (Suganuma et al. 2008). Moreover, mutations in Atac2 result in reduced global levels of acetylated H4K16 in fly embryos, and *ada2a* mutations decrease levels of acetylated H4K5, H4K12, and H4K16 in polytene chromosomes (Ciurciu et al. 2008, 2006; Suganuma et al. 2008). Depletion of Atac2 or Gcn5 from *Drosophila* cells by RNAi revealed that Gcn5 selectively acetylates histone H3, whereas Atac2 has a narrow but not absolute substrate preference for lysines on both H3 and H4 (Feller et al. 2015). Other HATs have been shown to work together to deposit particular combinations of acetyl marks on chromatin; for example, CBP, MGEA5, and NAA10 act together to acetylate H4 on both K5 and K8 (Feller et al. 2015). Similarly, Atac2's preference for different lysine residues on histones H3 and H4 was modulated by the pre-existing acetylation pattern on those histones (Feller et al. 2015).These data suggest that both Gcn5 and Atac2 contribute to the

expanded HAT activity of the ATAC complex, which is likely influenced *in vivo* by the activity of other HATs.

Gcn5 specificity may be altered by its interaction with each Ada2 paralog because rescue experiments with hybrid Ada2 proteins showed that combining the unique C-terminal domain of Ada2a and Ada2b with the N-terminal domain of the other Ada2 paralog was sufficient to rescue the respective mutants and restore histone acetylation patterns (E. Vamos and Boros 2012). Notably, the two Ada2b splice isoforms also only differ in their C-terminal domains (Fig. 1.2). Thus, the divergent C-terminal domains of the different Ada2 paralogs and splice isoforms in *Drosophila* likely contribute to both the formation of the different Gcn5 complexes and to the differences in HAT specificity of each complex.

In addition to histones, Gcn5 acetylates a number of non-histone targets in flies, which expand the biological functions of the Gcn5 complexes. For example, *Drosophila* Gcn5 acetylates the chromatin remodeling ATPase subunit Imitation SWI (Fbgn0011604, CG8625, Iswi) at K753 both *in vivo* and in *vitro* (Ferreira et al. 2007). This region in Iswi (747 – 756aa) is similar to the N-terminal domain of histone H3, suggesting that Gcn5 may recognize Iswi in a similar fashion to histone H3 (Ferreira et al. 2007). Iswi is part of two nucleosome remodeling complexes in *Drosophila*: Nucleosome remodeling factor (NURF), and the Chromatin accessibility (CHRAC) complex (Corona and Tamkun 2004). However, the acetylated form of Iswi is only found in NURF, and is not present in the CHRAC complex (Ferreira et al. 2007). Notably, as discussed in more detail in section 1.7, mutations in the NURF subunit *iswi* or the ATAC subunit *ada2a* show similar phenotypes, and there is a genetic interaction between Ada2a and Iswi in flies (Carré et al. 2008). These data suggest that in *Drosophila*, ATAC might target Iswi as a substrate for acetylation by Gcn5, although this has not been tested. In addition to Iswi, *Drosophila* Gcn5 has been shown to acetylate Transcription factor EB (TFEB; FBgn0263112, CG43369), the ortholog of Mtif in flies (Wang et al. 2020b). Gcn5 acetylates K445 and K450 in Mtif, inhibiting autophagy and lysosomal biogenesis (Wang et al. 2020b). *Drosophila* Gcn5 also acetylates the Cyclin A associated protein Adenomatous polyposis coli 2, Apc2 (FBgn0026598, CG6193) (Liu et al. 2017). Acetylation of Apc2 promotes ubiquitination and degradation of Cyclin A, resulting in its turnover, which regulates the maintenance (both self-renewal and differentiation) of *Drosophila* germline stem cells (Liu et al. 2017). More details about acetylation of non-histone substrates by Gcn5 across a variety of organisms including *Drosophila* are described in (Downey 2021)

1.5 Gcn5 is essential for development in flies

Although Gcn5 is not essential in for proliferation in *S. cerevisiae*, loss of one of the human Gcn5 paralogs, Gcn5 (KAT2A), results in embryonic lethality (Xu et al. 2000; Howe et al. 2001). Thus, the Gcn5 complexes appear to have an essential role in development in multicellular eukaryotes. To characterize the function of Gcn5 in *Drosophila*, Antoniewski and colleagues generated several different null *gcn5* alleles (Table 1.5). Loss of *gcn5* blocks two critical stages in *Drosophila* development: oogenesis (egg development) and metamorphosis. In flies lacking Gcn5, oogenesis is arrested at stage 5 and 6, and zygotic *gcn5* mutants die during the late third instar larval stage (Fig. 1.4) (Carré et al. 2005). Moreover, adults with hypomorphic *gcn5* alleles show malformation of appendages such as abnormal elongated metathoracic twisted legs, and also exhibit a reduction in wing size and defects in wing-vein patterning, together with defects in cuticle formation (Carré et al. 2005). In addition, null *gcn5* mutants fail to form a puparium, one of the initial steps in metamorphosis, potentially due to defects in expression of genes that respond to the insect hormone ecdysone (Carré et al. 2005). Notably, *gcn5* mutants also exhibit severely reduced imaginal discs, suggesting that Gcn5 is required for cell proliferation in flies. Consistent with a potential role in cell proliferation, *gcn5* mutant imaginal discs showed a higher number of cells in S-phase, significantly more cells undergoing mitosis, and higher levels of apoptosis (Carré et al. 2005). Mutations in another shared HAT module subunit, *ada3*, result in similar phenotypes to those observed in *gcn5* mutants, with reduced size of imaginal discs and defects in oogenesis (Grau et al. 2008). The small imaginal discs in the *ada3* mutant led to the original name *diskette* (Grau et al. 2008), although this gene has since been renamed *Ada3* on FlyBase. *ada3* mutants also exhibit abnormal structure of polytene chromosomes; in particular showing changes in the banding pattern of the male X chromosome (Grau et al. 2008).

The severe developmental defects in *gcn5* mutants are likely to result from the combinatorial loss of all four *Drosophila* Gcn5 complexes. However, the identification and analysis of mutants that specifically disrupt each of the four Gcn5 complexes in flies suggests that at least three of these Gcn5 complexes are essential for development in flies. For example, mutations in *ada2a* or *ada2b* both result in developmental lethality and oogenesis arrest (Fig. 1.4) (Pankotai et al. 2005; Qi et al. 2004). Further, mutations that disrupt the SAGA-specific subunits *nonstop*, *sgf11*, *wda*, *taf10b*, and *saf6*, or the CHAT-specific *chiffon* and *ada2b-PA* subunits, also result in larval lethality (Table 1.5, Fig. 1.4) (Weake et al. 2009, 2008; Guelman et al. 2006b; Pahi et al. 2017;

Torres-Zelada et al. 2019). Thus, SAGA, ATAC, and CHAT are essential for fly development. Unfortunately, ADA function cannot be separated genetically from SAGA in flies because both complexes share the Ada2b-PB isoform, and ADA contains no unique subunits in flies(Soffers et al. 2019). It should be noted that it remains unclear as to why mutations that disrupt different subunits of Gcn5 complexes result in lethality at different developmental stages (Fig. 1.4). Some mutants may exhibit more severe defects and earlier lethality due to their function in complexes outside the Gcn5 complexes, such as *sf3b5*, which is present in both SAGA and the U2 snRNP (Herold et al. 2009). In addition, there may be a different amount of maternally supplied gene product that allows some Gcn5 complex mutants to survive to a later developmental stage. Germline mutants in several SAGA-specific mutants either fail to complete oogenesis, or cannot progress through embryogenesis (Fig. 1.4), supporting the idea that maternally supplied gene product is required for these zygotic mutants to progress to a later stage in development. However, the level or stability of maternally supplied gene product for different Gcn5 complex subunits has not been examined in flies. Overall, the characterization of mutants that specifically disrupt SAGA, ATAC, or CHAT provides some insight into the different roles of these complexes, and we outline specific biological functions of each complex in the following sections beginning with SAGA.

1.6 SAGA is critical for developmental processes defined by its modules

SAGA promotes transcription through both its catalytic activities and via interactions with the transcription machinery (Koutelou et al. 2010). In *Drosophila*, SAGA colocalizes extensively with RNA polymerase II (Pol II) and is present at the both the promoter-proximal pause site of lowly expressed or highly regulated genes, and on the gene body of actively transcribed genes (Li et al. 2017; Weake et al. 2011). Although SAGA colocalizes with Pol II at most actively transcribed genes, gene expression profiling studies of SAGA mutants originally suggested that different SAGA modules might be required for transcription of particular subsets of genes (Weake and Workman 2012). For example, only a subset of the genes bound by SAGA in embryonic muscle were downregulated in *sgf11* mutants, and these genes showed enriched expression in muscle and functions related to muscle development, suggesting a potential role for the SAGA DUB module in expression of tissue-specific genes (Li et al. 2017; Weake et al. 2011; Weake and Workman 2012). However, in human cells SAGA acetylates histone H3K9 and deubiquitinates H2Bub1 on all expressed genes (Bonnet et al. 2014), suggesting a much broader role in regulating

transcription. This broader role in transcription is consistent with the extensive colocalization of SAGA with Pol II in flies and in human cells (Weake et al. 2011; Bonnet et al. 2014). Since many of the early gene expression studies on *Drosophila* mutants used microarray analysis approaches that may not have been able to detect global changes in transcription (Table 1.6), it is possible that a much larger group of genes requires SAGA for proper expression in flies. In addition, studies in *S. cerevisiae* suggest that global changes in transcription can be buffered by changes in mRNA stability (Haimovich et al. 2013; Rodríguez-Molina et al. 2016; Baptista et al. 2017), and most gene expression studies in flies have examined steady-state mRNA levels. Thus, the genes identified in the expression profiling experiments in SAGA mutants may represent those subsets of genes that are most sensitive to loss of particular SAGA activities.

Despite the caveat that the gene expression profiling of SAGA mutants in flies may underestimate the number of genes regulated by SAGA, these studies have provided important insight into key developmental processes that require SAGA. Importantly, mutants that disrupt different modules of SAGA show different effects on gene expression, and exhibit specific developmental phenotypes. For example, mutations in *ada2b* disrupt oogenesis, whereas oogenesis progresses normally in *ataxin 7* or *nonstop* mutants (Li et al. 2017). Moreover, genes involved in DNA replication, eggshell formation, and chromosome organization were significantly downregulated in *ada2b* oocytes, but did not change in *ataxin 7* or *nonstop* mutants (Table 1.6) (Li et al. 2017). While the early zygotic genes are expressed properly in embryos that lack the maternal contribution for *ataxin 7* and *nonstop*, these embryos show later defects in cellularization and nuclear anchoring (Li et al. 2017), suggesting that maternally contributed SAGA is required for proper development during embryogenesis. Interpreting these phenotypes is complicated by the recent finding that *ada2b* encodes two splice isoforms, only one of which is in the SAGA complex (Weake et al. 2009; Torres-Zelada et al. 2019); thus, *ada2b* mutants disrupt all three of the SAGA, ADA and CHAT complexes, making it difficult to distinguish as to which complex is required for oogenesis in flies (Fig. 1.4).

The disruption in eye development caused by mutations in SAGA's DUB module provides a second example of how different activities of SAGA control development in flies. Although mutations that disrupt the DUB module such as *sgf11* and *nonstop* are lethal during the late larval/early pupal stage of development (Fig. 1.4), these mutants show characteristic defects in eye development in the late larval stage just prior to their death (Poeck et al. 2001; Weake et al. 2008;

Martin et al. 1995). During the third larval instar, photoreceptor neurons in the developing eye imaginal disc project their axons to specific regions of the developing brain (Martin et al. 1995). The SAGA subunit *nonstop* was first identified in a screen for genes involved in this photoreceptor axon targeting process (Martin et al. 1995). Mutations in *nonstop* result in a failure of photoreceptor axons to project to their correct target layer in the developing brain, the lamina, instead mistargeting into the deeper medulla region (Poeck et al. 2001). This axon targeting defect is caused by loss of *nonstop* or *sgf11* within the glial cells that mark the target layer in the lamina (Poeck et al. 2001). Transcriptome profiling of these glial cells from *nonstop* and *sgf11* larval brains identified genes involved in axon guidance (Table 1.6) (Ma et al. 2016). Moreover, RNAi knockdown or loss of function mutants in one of these DUB-regulated genes in glia, *multiplexin* (FBgn0260660, CG42543, Mp), resulted in axon targeting defects that were similar to those observed in *sgf11* mutants, arguing that at least some of these DUB-regulated genes in glia control axon targeting (Ma et al. 2016). Since *ada2b* mutants also show axon mistargeting phenotypes, albeit substantially weaker than those observed in *nonstop* or *sgf11*, the DUB module likely controls expression of these genes as part of the SAGA complex (Ma et al. 2016; Weake et al. 2008). However, in flies the DUB module can bind to chromatin independently of SAGA's HAT or structural core subunits (Li et al. 2017), and loss of *ataxin 7* results in decreased H2Bub1 levels due to promiscuous binding of the DUB module (Mohan et al. 2014a). Genes involved in locomotion, organ morphogenesis, and eye and neuronal development were highly regulated by the DUB module (Li et al. 2017), suggesting that it remains possible that the DUB module could control some aspects of eye development independent of SAGA.

Third, analysis of mutations that disrupt the structural core and spliceosomal modules of SAGA suggests that like in *S. cerevisiae*, *Drosophila* SAGA can act as a transcriptional coactivator independent of its HAT or DUB activities (Koutelou et al. 2010). In *S. cerevisiae*, Tra1 recruits SAGA to promoters through interactions with transcription factors (Brown et al. 2001), allowing Spt3 and Spt8 to interact directly with component of the transcription machinery such as TBP (Sterner et al. 1999). In flies, mutations in the structural core subunit *Saf6* result in defective expression of SAGA-regulated genes without altering global levels of acetylated histone H3 or H2Bub1 (Weake et al. 2009). Likewise, mutations in the *sf3b5* spliceosomal SAGA subunit result in decreased expression of SAGA-regulated genes independent of changes in histone acetylation (Stegeman et al. 2016). Analysis of the relative levels of spliced and unspliced transcripts for genes

that are downregulated in *sf3b5* mutants shows that the decreased mRNA levels in *sf3b5* mutants are not necessarily due to changes in splicing efficiency (Stegeman et al. 2016). However, unlike other SAGA mutants, *sf3b5* is required for cell viability in flies, most likely due to its role as part of the U2 snRNP (Stegeman et al. 2016; Herold et al. 2009). It is unclear how Sf3b5 regulates gene expression as part of SAGA, although it is possible that it may mediate transient interactions between the transcriptional and splicing machinery, which share a common spatial and temporal distribution during the coupled processes of transcription and splicing (Weake et al. 2011; Herzel et al. 2017; Nuño-Cabanes and Rodríguez-Navarro 2021) Together, these studies suggest that SAGA plays a fundamental role in fly development because it regulates the expression of genes that are required for processes such as oogenesis, metamorphosis, and neuronal development. However, fundamental questions remain as to whether the distinct roles of SAGA in particular developmental processes result from independent activity of particular modules or subunits. In addition, it is unclear as to whether SAGA has overlapping or distinct roles with the ADA and CHAT complexes that are also disrupted in *ada2b* mutants. *Drosophila* SAGA may also regulate a broader set of genes than indicated by past gene expression studies that have profiled steady-state mRNA levels and have not been able to detect global changes in active transcription. In *S. cerevisiae*, data suggests that SAGA regulates expression of all genes (Bonnet et al. 2014; Baptista et al. 2017), while in human cells, SAGA deubiquitinates H2Bub1 on the transcribed region of all expressed genes, suggesting a widespread role in transcription regulation (Bonnet et al. 2014).

1.7 ATAC is a double HAT complex required for development

The ATAC complex is exclusive to multicellular eukaryotes, suggesting a potential function unique to development in multicellular organisms. Mutations that disrupt subunits of ATAC show developmental lethality during the larval or pupal stages (Table 1.5, Fig. 1.4). For example, *ada2a* mutants die during the pupal stage, and Ada2a is also essential for oogenesis (Pankotai et al. 2005). In addition, mutant flies that lack *hcf, wds*, *atac3*, and *atac2* die during either the larval or pupal stage of development (Table 1.5) (Suganuma et al. 2008; Hollmann et al. 2002; Rodriguez-Jato et al. 2011; Dietzl et al. 2007). The developmental lethality of ATAC mutants may be due to defects in response to the insect hormone ecdysone, which triggers molting during the larval instars, and is also required for the larval–pupal transition at the onset of metamorphosis (Danielsen et al. 2013). Both ecdysone levels and binding of its receptor to polytene chromosomes are reduced in a*da2a*

*a*nd *ada3* mutants (Pankotai et al. 2010). Moreover, genes required for ecdysone biosynthesis are misregulated in third instar larvae lacking Ada2a and Ada3 (Table 1.6) (Pankotai et al. 2010). Thus, ATAC may be essential for viability in flies in part because it controls levels of hormones that trigger formation of the adult fly.

Histone acetyltransferases often act synergistically with nucleosome remodeling complexes to regulate chromatin structure and gene expression (Featherstone 2002). In flies, ATAC interacts genetically and biochemically with the chromatin remodeling complex, NURF (Carré et al. 2008). Mutations in the NURF subunit *iswi* or the ATAC subunit *ada2a* show similar defects in eye development, with both mutants exhibiting small and rough eyes (Carré et al. 2008). In addition, ATAC and NURF coregulate expression of a subset of genes including *Ultrabithorax* (*Ubx*), *engrailed* (*en*), and *heat-shock protein 70* (*hsp70*) (Carré et al. 2008). Moreover, ATAC and NURF are both necessary to maintain proper chromatin structure, particularly on the X chromosome in male flies (Carré et al. 2008). In flies, expression of genes on the single male X chromosome is doubled to equal that from the two female X chromosomes in a process termed dosage compensation (Georgiev et al. 2011). During this process, the Males absent on the first (Mof) HAT within the Male Specific Lethal (MSL) complex acetylates H4K16 on the male X chromosome (Georgiev et al. 2011). Mutations in ATAC and NURF subunits such as *ada2a*, *gcn5*, and *nurf301* show increased frequency of bloated X chromosomes in male flies (Carré et al. 2008), suggesting that ATAC and NURF maintain proper chromosomal structure of the dosage compensated male X chromosome. Although ATAC acetylates H4K16 (Suganuma et al. 2008), the bloated X chromosomes observed in *ada2a* and *gcn5* mutant males show similar levels of acetylated H4K16 compared to their wild-type counterparts (Carré et al. 2008). Moreover, X-linked genes are not preferentially misregulated in *ada2a* or *gcn5* mutants (Carré et al. 2008). Thus, ATAC and NURF may work together to maintain the chromosomal structure of the dosage compensated male X chromosome, rather than playing a specific role in expression of X-linked genes (Carré et al. 2008). Notably, H4K16 acetylation by Mof antagonizes activity of another related chromatin remodeler, Iswi, in flies (Corona et al. 2002), and negatively regulates interactions between Iswi and its nucleosomal substrate *in vitro* (Clapier et al. 2002). Thus, the MSL and ATAC complexes may function synergistically with the related NURF and ISWI chromatin remodelers to maintain the structure, acetylation, and expression levels of dosage compensated genes in *Drosophila*. It is possible that the cooperative activity between ATAC and

NURF could involve the direct acetylation of one of the NURF subunits, Iswi, by Gcn5 (see Section 1.4) (Ferreira et al. 2007), although this remains to be tested.

Drosophila ATAC contains three histone-fold domain proteins, D12, Chrac-14 and NC2β, leading to the question as to whether ATAC itself possessed nucleosome remodeling activity because histone-fold domains can bind DNA (Baxevanis et al. 1995), and Chrac-14, as part of the CHRAC complex, facilitates nucleosome sliding (Hartlepp et al. 2005). In addition, the human ortholog of Chrac-14, Chrac-17, enhances nucleosome sliding by the Iswi complex (Kukimoto et al. 2004). Purified ATAC does not show remodeling activity by itself on nucleosomal substrates *in vitro* (Suganuma et al. 2008). However, ATAC can stimulate nucleosome sliding by the chromatin remodelers Iswi or SWItch/Sucrose Non-Fermentable (SWI-SNF) *in vitro* (Suganuma et al. 2008). Similarly, recombinant Chrac-14 or NC2β also stimulated nucleosome remodeling by SWI/SNF (Suganuma et al. 2008), suggesting that the histone-fold domain proteins in ATAC contributes to its impact on chromatin remodeling. Notably, the inclusion of acetyl-CoA in these *in vitro* nucleosome sliding assays enhanced the effect of ATAC, suggesting that the HAT activity of ATAC also contributes to stimulation of chromatin remodeling by complexes such as Iswi or SWI-SNF (Suganuma et al. 2008).

In addition to its roles in chromosome structure and interaction with chromatin remodelers, ATAC has been implicated in cell proliferation. Mutations in *gcn5* and *ada3* are associated with reduced size of imaginal discs, which are a highly proliferative tissue, and *gcn5* mutants also show an increased number of cells in S phase (Carré et al. 2005; Grau et al. 2008). However, since Gcn5 and Ada3 are core components of all the Gcn5 complexes in flies, it was not clear whether all or only some of these Gcn5 complexes had roles in cell proliferation. Studies in mammalian cells suggest that ATAC is likely to be responsible for the defects in cell proliferation in *gcn5* and *ada3* mutants due to its role in progression through the G2/M phase of the cell cycle (Guelman et al. 2009). Knockdown of Atac2 in mouse cells and studies using an *Atac2* knockout mouse model showed that loss of Atac2 results in an increase in the number of apoptotic cells and in an accumulation of cells in G2/M (Guelman et al. 2009). In addition, Ada2a and Ada3 RNAi knockdown in mouse NIH3T3 cells leads to mitotic abnormalities such as centrosome multiplication and defective midbody formation, and ATAC subunits such as Ada2a and Yeats2 localize to the mitotic spindle (Orpinell et al. 2010). Interestingly, SAGA does not appear to share this role in mitosis because deletion of Spt20 does not cause mitotic abnormalities, and Spt20 does

not localize to chromatin during mitosis (Orpinell et al. 2010). Although ATAC acetylates H4K16, loss of Ada2a and Ada3 results in the opposite acetylation phenotype in mitotic cells with knockdown cells showing an increase in acetylated H3K14 levels due to an decrease in the activity of the histone deacetylase Sirtuin 2 (SIRT2) (Orpinell et al. 2010). While a role for *Drosophila* ATAC in mitosis has not yet been characterized, it is possible that ATAC shares this function in flies and may be responsible for the decreased cell proliferation observed in *gcn5* and *ada3* mutants.

Last, ATAC has been implicated in controlling the expression of genes in stress-induced signaling pathways. Gcn5 complexes have a well characterized role in stress response signaling mediated by mitogen-activated protein kinases (MAPK) (Spedale et al. 2012). Osmotic stress can activate MAPK cascades, resulting in eventual activation of the c-Jun-NH2-terminal kinase (JNK) (Yang et al. 2003). In *Drosophila* S2 cells, sorbitol treatment induces osmotic stress and results in JNK activation (Suganuma et al. 2010). Importantly, ATAC directly interacts with MAPKs via its MBIP/Mocs2B subunits in both humans and flies (Suganuma et al. 2010; Fukuyama et al. 2000). Moreover, JNK activation in response to osmotic stress is inhibited by the expression of the ATAC subunit Mocs2B in *Drosophila* S2 cells, and ATAC is required for the transcription of JNK target genes such as *chickadee* in these cells (Suganuma et al. 2010). Thus, ATAC appears to directly interact with MAPK signaling proteins to mediate induction of stress response genes in flies, likely through its Mocs2B subunit. This role in stress response for the ATAC complex is reminiscent of *S. cerevisiae* SAGA's function in the endoplasmic reticulum (ER) stress pathway (Welihinda et al. 2000). In mammals, knockdown of the shared SAGA and ATAC subunit Sgf29 results in impaired transcription of ER stress genes, such as *GRP78* (Schram et al. 2013). The ER stress response transcription factor ATF6 recruits both SAGA and ATAC to ER stress response genes (Sela et al. 2012), suggesting that both SAGA and ATAC are involved in induction of stress response genes in metazoan organisms. Analysis of SAGA and ATAC localization on *Drosophila* polytene chromosomes suggest that these Gcn5 complexes regulate distinct sets of stress response genes, depending on the type of stress involved (Nagy et al. 2010). For example, induction of phorbol ester-induced protein kinase C (PKC) pathway genes increased colocalization of ATAC and Pol II without affecting SAGA (Nagy et al. 2010), arguing for a specific role of ATAC in induction of PKC genes in response to stress.

1.8 CHAT is an insect-specific Gcn5 complex that contains a protein associated with DNA replication

Whereas the other Gcn5 complexes identified in *Drosophila* are also present in *S. cerevisiae* or humans, the CHAT complex appears to be specific to insects and has an unknown biological function. In addition to the HAT module subunits (Gcn5, Ada3, and Sgf29), CHAT contains the short Ada2b-PA splice isoform and Chiffon, the *Drosophila* ortholog of Dbf4. Chiffon, like other Dbf4 orthologs, binds and activates the cell cycle kinase Cdc7 forming the Dbf4-dependent kinase complex (DDK) (Schüpbach and Wieschaus 1989; Stephenson et al. 2015). The DDK complex phosphorylates the Mcm2-7 helicase, activating it to unwind DNA at origins of replication, thus initiating DNA replication (Labib 2010; Stephenson et al. 2015). Although Dbf4 is highly conserved and is present in most eukaryotes except for plants, Chiffon contains a long C-terminal extension that is specific to insects (Fig. 1.5) (Tower 2004; Torres-Zelada et al. 2019). The conserved N-terminal domain of Chiffon (1- 400aa) binds and activates Cdc7, while the insect-specific C-terminal domain of Chiffon (401 - 1695aa) is necessary and sufficient to bind Gcn5 and nucleate CHAT formation (Torres-Zelada et al. 2019). Dbf4 is an essential gene in *S. cerevisiae* because of its role in DNA replication, but surprisingly, *chiffon* mutants were originally reported to be viable in *Drosophila* (Landis and Tower 1999). The *chiffon* gene was first identified in a screen for female sterile mutants, and *chiffon* females lay eggs with a thin and fragile chorion (eggshell) that resembles the fabric of the same name (Schüpbach and Wieschaus 1989). More recent analysis has shown that indeed, the Cdc7-binding domain of Chiffon is dispensable for fly viability, but surprisingly, the Gcn5-binding domain of *chiffon* is essential for development(Torres-Zelada et al. 2019). In fact, *chiffon* alleles that contain premature stop codons either within, or directly after, the N-terminal Cdc7-binding domain (separating both N- and C-polypeptides) are viable because they still produce a C-terminal product that binds Gcn5 and nucleates CHAT formation (Torres-Zelada et al. 2019). Both domains are encoded by a single large exon in the *chiffon* gene with no evidence of alternative splicing, suggesting that alternative translation start sites and/or proteolytic cleavage may be required to produce these two independent Chiffon polypeptides. These data suggest that *chiffon* could be a dicistronic gene that can independently express two distinct polypeptides that contain either the Cdc7- or Gcn5-binding domains, resulting in DDK or CHAT formation, respectively. It remains unclear as to whether the N- and C-terminal Chiffon polypeptides are expressed at the same time, and little is known about

how this process is controlled *in vivo*. The unusual *chiffon* gene structure is somewhat reminiscent of the *ada2a* gene, which also encodes two polypeptides with distinct functions: Ada2a and one of the subunits of RNA polymerase II, Rpb4, in flies and in other insects due to alternative splicing (Muratoglu et al. 2003).

The CHAT complex exhibits *in vitro* and *in vivo* HAT activity toward histone H3, similar to SAGA and ADA (Torres-Zelada et al. 2019) Analysis of histone acetylation levels in somatic mosaics for *chiffon* null alleles showed that loss of Chiffon decreases levels of histone H3 acetylated at K9, K14, and K18, but not K23 (Torres-Zelada et al. 2019). Although histone acetylation correlates with, and contributes to a specialized form of DNA re-replication in follicle cells termed gene amplification (Aggarwal and Calvi 2004), CHAT-mediated histone acetylation is not required for this type of DNA replication (Torres-Zelada et al. 2019). In *chiffon* mutant cells that lack only its N-terminal Cdc7-binding domain, ovary follicle cells lack the characteristic bromodeoxyuridine (BrDU) foci indicative of chorion gene amplification (Torres-Zelada et al. 2019). However, these DDK-deficient mutant cells retain wild-type histone acetylation levels. In contrast, *chiffon* mutants that lack only its C-terminal domain that binds Gcn5 show decreased histone acetylation, but do not exhibit loss of the characteristic BrDU foci indicative of chorion gene amplification (Torres-Zelada et al. 2019). Similarly, *ada2b* mutant follicle cells show decreased histone acetylation but retain wild-type BrDU incorporation (Torres-Zelada et al. 2019). Together, these data suggest that despite the presence of the Dbf4 ortholog Chiffon, the CHAT complex is not required for DNA replication in flies (Torres-Zelada et al. 2019). What then could be the role of the CHAT complex in insects? Currently, CHAT, like SAGA, seems to be essential for both histone H3 acetylation and for development in flies. *chiffon* mutants show decreased histone H3 acetylation not only in ovary follicle cells, but also in other tissues such as imaginal discs (Torres-Zelada et al. 2019). Moreover, the decreased acetylation at histone H3K14 in *chiffon* mutant cells is similar to that observed in *ada2b* mutants, which lack both the CHAT and SAGA isoforms (Torres-Zelada et al. 2019). Since mutations in the SAGA-specific subunit, *wda*, also reduce acetylation at histone H3K9 in embryos (Guelman et al. 2006b), both SAGA and CHAT likely contribute to H3 acetylation in flies. However, expression of the CHAT-specific Ada2b-PA isoform, but not the SAGA/ADA-specific Ada2b-PB isoform, is sufficient to almost fully restore viability to *ada2b* mutants (Pankotai et al. 2013; Torres-Zelada et al. 2019). These data suggest that either CHAT might compensate for some of SAGA's essential functions during development,

or that the Ada2b-PA splice isoform can incorporate into SAGA if Ada2b-PB is absent (Torres-Zelada et al. 2019). It remains unclear whether CHAT is necessary for gene expression, and if so, whether CHAT regulates common or distinct gene targets compared to SAGA and the other Gcn5 complexes in flies.

1.9 Roles for Gcn5 complexes in other insects

The Gcn5 complexes have been best studied in the model insect *Drosophila melanogaster*, and no Gcn5 complexes have been described in other insects yet. However, other insect species, like *Drosophila*, possess a single Gcn5 ortholog with shared domain structure including the metazoan-specific N-terminal domain (Fig. 1.5A). Both Ada2a and Ada2b are also widely conserved throughout insects suggesting that the ADA, SAGA, and ATAC complexes are likely present in all insect species (Fig. 1.5B). Further, like in *Drosophila*, Ada2b in most insect species has two splice isoforms that share a common N-terminal domain, which includes the Zinc finger ZZ-type and SANT domain, and have the specific C-terminal regions corresponding to the *Drosophila* Ada2b-PA and Ada2b-PB splice isoforms (Fig. 1.5B). The presence of both Ada2b splice isoforms in other insect species supports the idea that the CHAT complex is likely conserved across insect species. In addition, the Chiffon C-terminal extension that directly binds Gcn5 *in vitro* is conserved in a wide range of insect species from beetles to ants (Fig. 1.5C) (Tower 2004; Landis and Tower 1999; Torres-Zelada et al. 2019). Currently, the biological function of the CHAT complex is unknown, but it is possible that this complex plays a specialized role in insects due to some unique aspect of their development or physiology.

1.10 Conclusion and future directions

During evolution there has been a divergence and diversification of the Gcn5 complexes. *Drosophila* has provided a powerful model in which to identify and characterize these novel Gcn5 complexes, and was the first multicellular organism shown to contain the ADA, ATAC and CHAT complexes (Guelman et al. 2006a; Torres-Zelada et al. 2019; Soffers et al. 2019). The expanded repertoire of Gcn5 complexes in flies and in other metazoan organisms appears to result from divergence of the Ada2 subunit. While *S. cerevisiae* only has one Ada2 ortholog, flies have at least three versions of Ada2: Ada2a and the two splice isoforms of Ada2b. The finding that alternative

splicing of *ada2b* can generate new diversity in HAT complexes (Torres-Zelada et al. 2019; Soffers et al. 2019) suggests that there may be other Gcn5 complexes in multicellular organisms that remain to be discovered. It is possible that other novel Gcn5 complexes, like CHAT, may be specific to particular groups of species where they play more specialized roles in developmental processes. In light of the fairly recent finding that *Drosophila* possess four Gcn5 complexes rather than just SAGA and ATAC, it may be necessary to re-interpret some of the conclusions from previous studies showing specific roles for SAGA, or particular modules of SAGA, in developmental processes. New genome-wide studies of Gcn5 complex localization patterns and gene expression profiling will require careful selection of subunits, and should utilize spike-in control approaches that can identify potential global changes in gene expression (Jiang et al. 2011).

Over the past 20 years following the identification of Gcn5 in *Drosophila* (Smith et al. 1998), much insight has been obtained into the structure and function of SAGA from studies in yeast, flies, humans and plants. We refer the reader to the article by Brian Strahl and Scott Briggs (Strahl and Briggs 2021) for an in-depth discussion of SAGA's function in transcription, and an outline of key unanswered questions that remain about its function. The exciting new cryo-EM studies of *S. cerevisiae* SAGA illustrate how the different modular parts of the complex function as a whole (Liu et al. 2019; Wang et al. 2020a; Papai et al. 2020; Helmlinger et al. 2021), and we look forward to seeing these same approaches applied to the metazoan SAGA and ATAC complexes to elucidate the architectural organization of both complexes. Such studies will provide insight into the similarity and differences between SAGA and ATAC, and show for example, how the two HATs in ATAC might modify histones within the same nucleosome, and how the spliceosomal proteins in metazoan SAGA integrate into the complex. These studies, coupled with functional analysis in model systems such as flies, may help us to understand why the metazoan Gcn5 complexes have diverged in composition from yeast and plants. Plants, like yeast, lack the ATAC complex and do not have the Sf3b3 and Sf3b5 spliceosomal subunits of SAGA (Grasser et al. 2021). What, then, is the unique role that the ATAC complex plays in metazoan? Why does metazoan SAGA contain the spliceosomal subunits, and what is their function in the complex?

Insects offer a number of advantages over mammalian models to answer these key questions because of their short generation time, and wealth of genetic resources. In addition, since the *Drosophila* SAGA and ATAC complexes largely resemble their mammalian counterparts in terms of composition, flies provide a strong model for the metazoan-specific functions of the Gcn5

complex. *Drosophila* also provide an appropriate biological model to ask questions about Gcn5 complexes that are relevant to human disease. For instance, the neurodegenerative disease Spinocerebellar ataxia type 7 (SCA7) results from polyglutamine expansions in the gene encoding the DUB subunit Ataxin 7 (Holmberg et al. 1998; David et al. 1997). Flies have been used as a model for SCA7 (Mohan et al. 2014b; Latouche et al. 2007), and other polyglutamine related neurogenerative diseases such as SCA2 (Elden et al. 2010). In humans, SCA7 disease manifests retinal and cerebellar degeneration, and macular dystrophy causing blindness (David et al. 1997). In *Drosophila*, loss of Ataxin 7 causes neural and retinal degeneration, and impaired movement (Mohan et al. 2014a). Interestingly, similar phenotypes are observed when exogenous polyglutamine-expanded human Ataxin 7 is expressed in *Drosophila* (Mohan et al. 2014a). Thus, *Drosophila* provides a good model organism to study the mechanism of diseases such as SCA7 and could be used to screen compounds suitable for ameliorating symptoms of this neurodegenerative disease (Mohan et al. 2014b; Burke et al. 2013).

Last, the finding that alternative splicing of *ada2b* can generate new diversity in HAT complexes (Torres-Zelada et al. 2019; Soffers et al. 2019) suggests that there may be other Gcn5 complexes in multicellular organisms that remain to be discovered. It is possible that other novel Gcn5 complexes, like CHAT, may be specific to particular groups of species where they play more specialized roles in developmental processes. *Drosophila* remains an outstanding model for studying function of the Gcn5 complexes, but recent advances in technology allow us to consider examining alternative species outside of traditional model organisms. Expanding the studies on Gcn5 complexes into non-traditional species, including potentially other insects may provide insight into the specialized function of this quintessential HAT in multicellular organisms.

Table 1.1 *Drosophila* SAGA subunits

	FlyBase ID	**Annotation Symbol**	**Gene name**	**Gene symbol**	***S. cerevisiae ortholog***	***H. sapiens* ortholog**	**DNA/Histone domain/Enzymatic Activity**
HAT module	FBgn0030891	CG7098	*transcriptional Adaptor 3 (diskette)*	*Ada3*	*ADA3*	*TADA3*	
	FBgn0020388	CG4107	*Gcn5 acetyltransferase (Pcaf)*	*Gcn5*	*GCN5*	*GCN5/PCAF (KAT2A/KAT2B)*	PCAF, GNAT domain, Bromodomain, Acetyltransferase (EC 2.3.1.48)
	FBgn0050390	CG30390	*SAGA-associated factor 29 kDa*	*Sgf29*	*SGF29*	*SGF29*	Tudor-like domain
	FBgn0037555	CG9638	*transcriptional Adaptor 2b*	*Ada2b (PB isoform)*	*ADA2*	*TADA2B*	Zinc finger ZZ-type, SANT Myb domain
DUB module	FBgn0013717	CG4166	*nonstop*	*Not*	*UBP8*	*USP22 (UBP22)*	Zinc finger-UBP-type, Ubiquitin protease
	FBgn0036804	CG13379	*SAGA associated factor 11 kDa*	*Sgf11*	*SGF11*	*ATXN7L3*	
	FBgn0031420	CG9866	*Ataxin 7*	*Atxn7*	*SGF73*	*ATXN7/ATXN7L 1/ATXN7L2*	SCA7 domain
	FBgn0000618	CG15191	*enhancer of yellow 2*	*e(y)2*	*SUS1*	*ENY2*	
Core structural module	FBgn0039067	CG4448	*will decrease acetylation*	*Wda*	*TAF5*	*TAF5L*	WD40 domain
	FBgn0030874	CG6506	*Spt7*	*Spt7*	*SPT7*	*SUPT7L (STAF65G)*	Histone-fold domain
	FBgn0036374	CG17689	*Spt20*	*Spt20*	*SPT20*	*SUPT20H*	
	FBgn0051865	CG31865	*transcriptional Adaptor 1*	*Ada1*	*ADA1*	*TADA1*	Histone-fold domain
	FBgn0031281	CG3883	*SAGA factor-like TAF6*	*Saf6*	*TAF6*	*TAF6L*	Histone-fold domain

Table 1.1 continued

	FlyBase ID	Annotation Symbol	Gene name	Gene symbol	*S. cerevisiae ortholog*	*H. sapiens* ortholog	DNA/Histone domain/Enzymatic Activity
	FBgn0000617	CG6474	*enhancer of yellow 1*	*e(y)1*	*TAF9*	*TAF9/TAF9b*	Histone-fold domain
	FBgn0026324	CG3069	*TBP-associated factor 10b*	*Taf10b*	*TAF10*	*TAF10*	Histone-fold domain
	FBgn0011290	CG17358	*TBP-associated factor 12*	*Taf12*	*TAF12*	*TAF12*	Histone-fold domain
TBP binding	FBgn0037981	CG3169	*Spt3*	*Spt3*	*SPT3*	*SUPT3H*	Histone-fold domain
TF-binding module	FBgn0053554	CG33554	*Nipped-A*	*Nipped-A*	*TRA1*	*TRRAP*	PIK-related pseudokinase
Splicing module	FBgn0035162	CG13900	*Splicing factor 3b subunit 3*	*Sf3b3*	-	*SF3B3*	Cleavage/polyadenylation specificity factor
	FBgn0040534	CG11985	*Splicing factor 3b subunit 5*	*Sf3b5*	-	*SF3B5*	

The 20 *Drosophila* SAGA subunits can be organized into HAT, DUB, Core Structural, TBP binding, TF-binding, and splicing modules. The FlyBase ID, Annotation symbol (CG ID number), full gene name, and abbreviated gene symbol are shown for each *Drosophila* subunit, together with the orthologs from *S. cerevisiae* and *H. sapiens* (if present). Paralogous subunits are separated with a "/" sign. Alternative gene names are listed in parentheses. The protein domain and enzymatic activity (E.C. number) are based on FlyBase definitions for each *Drosophila* subunit. Note that Spt7 contains a bromodomain only in *S. cerevisiae*, but not in the metazoan orthologs. In addition, Spt8 is only present in the *S. cerevisiae* SAGA complex and is not listed here.

Table 1.2 *Drosophila* ADA subunits

FlyBase ID	**Annotation Symbol**	**Gene name**	**Gene symbol**	***S. cerevisiae ortholog***	**DNA/Histone domain/Enzymatic Activity**
FBgn0030891	CG7098	*transcriptional Adaptor 3 (diskette)*	*Ada3*	*ADA3*	
FBgn0020388	CG4107	*Gcn5 acetyltransferase (Pcaf)*	*Gcn5*	*GCN5*	PCAF, GNAT domain, Bromodomain, Acetyltransferase (EC 2.3.1.48)
FBgn0050390	CG30390	*SAGA-associated factor 29 kDa*	*Sgf29*	*SGF29*	Tudor-like domain
FBgn0037555	CG9638	*transcriptional Adaptor 2b*	*Ada2b (PB isoform)*	*ADA2*	Zinc finger ZZ-type, SANT domain

The FlyBase ID, Annotation symbol (CG ID number), full gene name, and abbreviated gene symbol are shown for each *Drosophila* ADA subunit, together with the ortholog from *S. cerevisiae*. Alternative gene names are listed in parentheses. The ADA complex has not been yet characterized in human cells. The protein domain and enzymatic activity (E.C. number) are based on FlyBase definitions for each *Drosophila* subunit. Note that the *S. cerevisiae* ADA complex contains two additional subunits AHC1 and AHC2 that are not present in the *Drosophila* ADA complex.

Table 1.3 *Drosophila* ATAC subunits

FlyBase ID	Annotation Symbol	Gene name	Gene symbol	*H. sapiens* ortholog	DNA/Histone domain/Enzymatic Activity
FBgn0030891	CG7098	*transcriptional Adaptor 3 (diskette)*	*Ada3*	*TADA3*	
FBgn0020388	CG4107	*Gcn5 acetyltransferase (Pcaf)*	*Gcn5*	*GCN5*	PCAF, GNAT domain, Bromodomain, Acetyltransferase (EC 2.3.1.48)
FBgn0050390	CG30390	*SAGA-associated factor 29 kDa*	*Sgf29*	*SGF29*	Tudor-like domain
FBgn0263738	CG43663	*transcriptional Adaptor 2a*	*Ada2a*	*TADA2A*	Zinc finger ZZ-type, SANT domain, SWIRM domain
FBgn0039904	CG1710	*Host cell factor*	*Hcf*	-	
FBgn0040066	CG17437	*will die slowly*	*Wds*	*WDR5*	WD40 domain
FBgn0027490	CG13400	*D12*	*D12*	*YEATS2*	YEATS
FBgn0043002	CG13399	*Chromatin accessibility complex 14 kD-protein*	*Chrac-14*	-	Histone-fold domain
FBgn0052343	CG32343	*Ada2a-complex component 3*	*Atac3*	-	
FBgn0028926	CG4185	*Negative Cofactor 2β*	*NC2β*	*NC2β*	Histone-fold domain
FBgn0031876	CG9200	*Ada2a-containing complex component 1*	*Atac1*	*ZZZ3*	SANT domain
FBgn0032691	CG10414	*Ada2a-containing complex component 2*	*Atac2*	*CRBP2*	GNAT domain/ Acetyltransferase (EC 2.3.1.48)
FBgn0039280	CG10238	*Molybdenum cofactor synthesis 2B*	*Mocs2B (dMoaE, Mocs2)*	*MBIP*	Molybdopterin biosynthesis MoaE

The FlyBase ID, Annotation symbol (CG ID number), full gene name, and abbreviated gene symbol are shown for each *Drosophila* ATAC subunit, together with the ortholog from *H. sapiens*. Alternative gene names are listed in parentheses. The ATAC complex is not present in *S. cerevisiae*. The protein domain and enzymatic activity (E.C. number) are based on FlyBase definitions for each *Drosophila* subunit.

Table 1.4 *Drosophila* CHAT subunits

FlyBase ID	**Annotation Symbol**	**Gene name**	**Gene symbol**	**DNA/Histone domain/Enzymatic Activity**
FBgn0030891	CG7098	*transcriptional Adaptor 3 (diskette)*	*Ada3*	
FBgn0020388	CG4107	*Gcn5 acetyltransferase (Pcaf)*	*Gcn5*	PCAF, GNAT domain, Bromodomain, Acetyltransferase (EC 2.3.1.48)
FBgn0050390	CG30390	*SAGA-associated factor 29 kDa*	*Sgf29*	Tudor-like domain
FBgn0037555	CG9638	*transcriptional Adaptor 2b*	*Ada2b (PA isoform)*	Zinc finger ZZ-type, SANT domain
FBgn0000307	CG5813	*Chiffon*	*Chif*	Zinc finger DBF-type

The FlyBase ID, Annotation symbol (CG ID number), full gene name, and abbreviated gene symbol are shown for each *Drosophila* CHAT subunit. Alternative gene names are listed in parentheses. The CHAT complex is not present in *S. cerevisiae* or human cells. The protein domain and enzymatic activity (E.C. number) are based on FlyBase definitions for each *Drosophila* subunit.

Table 1.5 Phenotypes associated with mutant alleles that disrupt subunits of Gcn5 complexes in *Drosophila.*

	Gene	Mutant allele	Nature of allele	Viable/lethal?	Phenotype	Reference
SAGA/ADA/ATAC/CHAT	*Ada3*	*ada3*^{3}	Nonsense: 371*	Lethal - early pupa	Decreased acetylation at H3K9, K14, K12; failure metamorphosis; reduced imaginal disc size.	(Grau et al. 2008)
	Gcn5	*gcn5*^{E333st} *gcn5*^{C137T} *gcn5*^{Q186st}	Nonsense: E333* Missense: C137Y Nonsense: Q186*	Lethal - late pupa	Decreased acetylation at H3K9, H3K14; Defects cell proliferation; failure to form puparium; photoreceptor axon mistargeting during eye development; oogenesis arrested in stage 5 and 6. Reduced imaginal disc size.	(Weake et al. 2008; Carré et al. 2005)
SAGA/ADA/CHAT	*Ada2b (PA- and PB isoforms)*	*ada2b*^{1} *ada2b*^{2} *ada2b*^{842}	Null: 1077bp deletion Null: 2.77kb deletion Null: 800bp deletion	Lethal - early pupa	Decreased acetylation at H3K9 in embryos and polytene chromosomes and H3K14 in ovary follicle cells and imaginal discs; defects in oogenesis; photoreceptor axon mistargeting during eye development.	(Qi et al. 2004; Weake et al. 2008; Pankotai et al. 2005; Torres-Zelada et al. 2019)
SAGA	*Nonstop*	*not*^{2}	Null: 538bp deletion	Lethal - pupa	Increased H2Bub1; photoreceptor axon mistargeting during eye development.	(Poeck et al. 2001; Weake et al. 2008)
	Sgf11	*sgf11*^{e01308}	Null: 5.97kb del etion	Lethal - late larva/early pupa	Increased H2Bub1; photoreceptor axon mistargeting during eye development.	(Weake et al. 2008, 2011)
	Ataxin 7	*ataxin 7* ^{KG02020}	Null:	Lethal - late larva	Neural and retinal degeneration; reduced locomotion; cellularization defects.	(Mohan et al. 2014a; Li et al. 2017)
	e(y)2	*e(y)2*^{1}	Null: 167bp deletion	Viable	Short stocky body and separated wings; eyes with altered facets; low fertility.	(Georgieva et al. 2001; Georgiev et al. 2011)

	Gene	Mutant allele	Nature of allele	Viable/lethal?	Phenotype	Reference
	Nipped-A	*nipped-A^{NC186}*	Missense: V885D	Lethal - early pupa	Defects in Notch signaling.	(Gause et al. 2006)
	wda	*wda^{11}* *wda^{4}* *wda^{8}*	Null: 1510bp deletion Null: 857bp deletion Null: 864bp deletion	Lethal - second instar larva	Decreased acetylation at H3K9.	(Guelman et al. 2006b)
	Saf6	*saf6^{303}*	Null: 303bp deletion	Lethal - second instar larvae		(Weake et al. 2009)
	e(y)1	*e(y)1^{17}* *e(y)1^{190}*	Null: 79bp deletion Null: 339bp deletion	Lethal - larva	Dysregulation of ovary follicle cell development.	(Xie et al. 2014)
	Taf10b	*taf10^{d25}*	Null: 900bp deletion	Lethal - pupae	Decreased acetylation at H3K14; defects in DNA repair efficiency.	(Pahi et al. 2015, 2017)
	Sf3b5	*sf3b5 ^{EY12579}*	Transposable element insertion.	Lethal - second instar larva	Reduced cell viability in eyes.	(Stegeman et al. 2016)
ATAC	*Ada2a*	*ada2a^{189}*	Null: 720bp deletion	Lethal - pupa	Oogenesis arrested; altered structure of the polytene chromosomes; banding pattern is distorted.	(Pankotai et al. 2005)
	hcf	*hcf^{HR1}*	Null: 4348bp deletion	Lethal - pupa	Heterozygous females are Sterile; oogenesis arrested at stage 8; decreased pupae size.	(Rodriguez-Jato et al. 2011)
	wds	*wds^{G0251}* *wds^{j25}*	Not specified	Lethal - larva	Defects in wristles and wing veins; heterozygous male and female are unfertile.	(Hollmann et al. 2002)

Table 1.5 continued

	Gene	Mutant allele	Nature of allele	Viable/lethal?	Phenotype	Reference
	Chrac-14	*chrac-14*^{KG01051}	Not specified	Viable	Eclosion defective; flight defective; radiation sensitive.	(Mathew et al. 2014)
	Atac3	*atac3* ^{GD4326}	RNAi	Lethal - pupa		(Dietzl et al. 2007)
	Atac2	*atac2* ^{e03046}	Transposable element insertion.	Lethal - second instar larva	Decreased acetylation at H4K16.	(Suganuma et al. 2008)
CHAT	*Chiffon*	*chif*^{Dsred}	Null:5.3kb deletion	Lethal - third instar larva	Decreased acetylation at H3K9, H3K14, and H3K18 in ovary follicle cells and imaginal discs; gene amplification disrupted; thin embryo chorion and rough eyes for *chif*^{WF24}.	(Landis and Tower 1999; Stephenson et al. 2015; Torres-Zelada et al. 2019)
		chif^{ETEB3}	Null: 6kb deletion			
		chif^{WF24}	Missense: T521C	Viable		

Mutant alleles or RNAi knockdown that disrupt subunits that are shared or specific to the SAGA, ADA, ATAC and CHAT complexes result in the described lethality and phenotypes, as outlined in the listed references. Only mutant alleles/RNAi knockdown that have been described in the literature are listed in this table. *, designated amino acid is altered to a stop codon.

Table 1.6 Gene expression analysis for Gcn5 complexes in *Drosophila*

Complex	Gene	Approach	# genes identified	Differentially expressed genes-pathways/processes	Reference
SAGA/ATAC/ ADA/CHAT	*Gcn5*	Microarray; third instar larvae	~284 genes	Morphogenesis.	(Carré et al. 2008)
	Ada3	Microarray; third instar larvae	~5565 genes	Cuticle formation and ecdysone response.	(Pankotai et al. 2010)
SAGA/ADA/CHAT	*Ada2b (PA & PB isoforms)*	RNA-seq; ovaries	>1000 genes	DNA replication, eggshell formation, chromosome organization, and DNA repair.	(Li et al. 2017)
		Microarray; third instar larvae	~344 genes	Early ecdysone response genes: glue proteins.	(Weake et al. 2008)
		Microarray; third instar larvae	~580 genes	Ecdysone-induced genes, cuticle formation, and defense mechanisms.	(Pankotai et al. 2010; Zsindely et al. 2009)
SAGA	*Nonstop*	RNA-seq; embryos (stage 5)	>6000 genes	Cellularization, embryonic development, and tissue morphogenesis.	(Li et al. 2017)
		Microarray; third instar larvae	~987 genes	Early ecdysone-response genes, puparial adhesion, eclosion, signal transduction, and central nervous system remodeling	(Weake et al. 2008)
		RNA-seq; third instar larvae glia	~1802 genes	Axon guidance, protein folding, cell morphogenesis, axon guidance, synaptic transmission.	(Ma et al. 2016)

Table 1.6 continued

Complex	Gene	Approach	# genes identified	Differentially expressed genes-pathways/processes	Reference
	Sgf11	Microarray; embryonic muscle or neurons	~443 genes (muscle); ~390 genes (neuron)	Protein folding, nervous system development, mesoderm development, muscle development, and anatomical structure development.	(Weake et al. 2011)
		Microarray; third instar larvae	~618 genes	Early ecdysone response genes, puparial adhesion, eclosion, signal transduction, and central nervous system remodeling	(Weake et al. 2008)
		RNA-seq; third instar larvae glia	~1644 genes	Axon guidance, protein folding, cell morphogenesis, axon guidance synaptic transmission.	(Ma et al. 2016)
	Ataxin 7	RNA-seq; embryos (stage 5)	>6000 genes	Cellularization, embryonic development, and tissue morphogenesis.	(Li et al. 2017)
ATAC	*Ada2a*	*Microarrays; third instar larvae*	*~7306 genes*	*Cuticle formation and ecdysone pathway response.*	(Pankotai et al. 2010)

Gene expression studies have been performed on homozygous mutants that disrupt subunits of the Gcn5 complexes SAGA, ADA, ATAC and CHAT. The number of differentially expressed genes identified using microarray or RNA-seq analysis by each study is listed, together with the major gene ontology processes and/or signaling pathways identified in the associated reference.

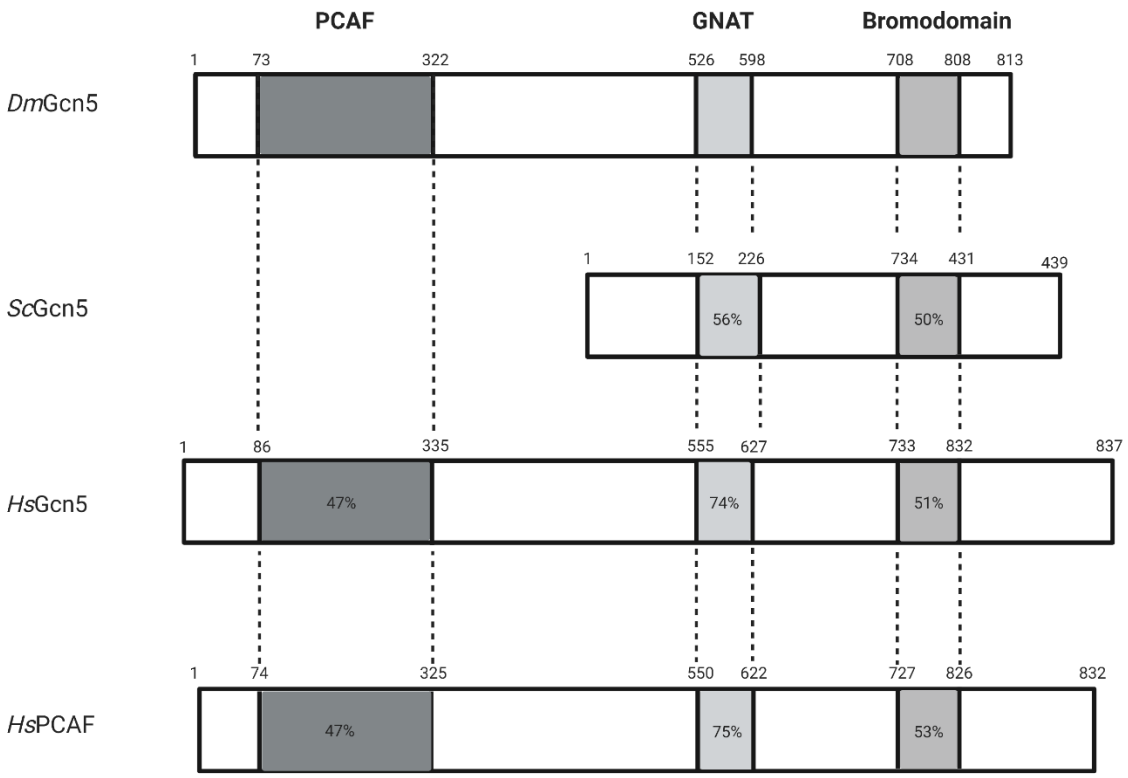

Figure 1.1 Schematic comparison of *Drosophila* Gcn5 orthologs.

Gcn5 amino acid sequences were aligned using Clustal Omega, and a schematic comparison of Gcn5 orthologs in *D. melanogaster*, *S. cerevisiae*, and *H. sapiens* was constructed. Accession numbers are as follow: *D. melanogaster* Gcn5, NP_648586.2; *S. cerevisiae* Gcn5, NP_011768.1; *H. sapiens* Gcn5, XP_006721880.1; *H. sapiens* PCAF, NP_003875.3. The highly conserved GNAT and Bromodomain, and the metazoan conserved PCAF domains are boxed in gray, and aligned in each ortholog as indicated by dotted lines. The amino acid positions for each domain are indicated by the numbers on top of each box. The percentage identity within the conserved domains in each Gcn5 ortholog relative to the corresponding domains in *Dm*Gcn5 is indicated by the % within each boxed domain.

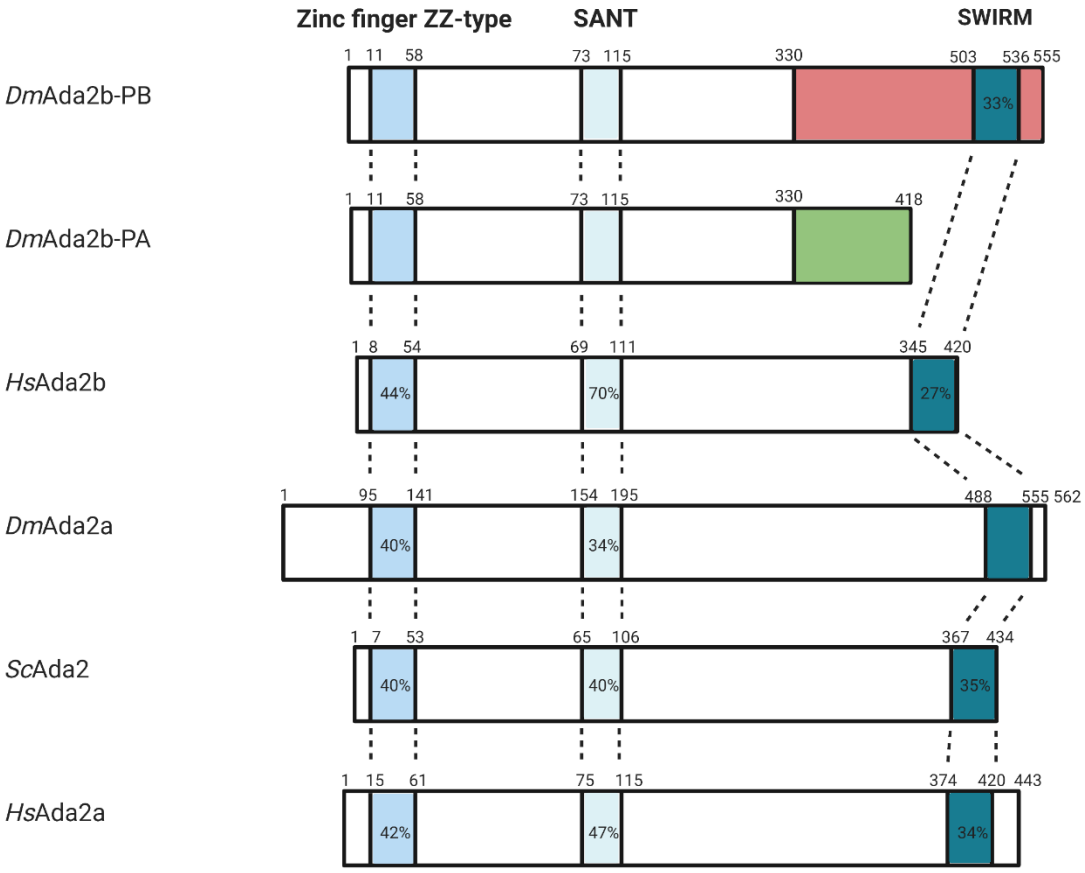

Figure 1.2 Schematic comparison of *Drosophila* Ada2a and Ada2b orthologs.

Ada2a and Ada2b amino acid sequences were aligned using Clustal Omega and a schematic comparison of *D. melanogaster* Ada2a and Ada2b (PA and PB isoforms) with *S. cerevisiae* Ada2 and *H. sapiens* Ada2a and Ada2b was constructed. Accession numbers are as follow: *D. melanogaster* Ada2a NP_001014636.1, Ada2b-PA NP_649773.1, Ada2b-PB NP_001027151.1; *H. sapiens* Ada2a NP_001159577.2, Ada2b NP_689506.2; *S. cerevisiae* Ada2 NP_010736.3. The conserved Zinc finger ZZ-type and SANT domains, and the SWIRM domains are boxed and aligned between the orthologs as indicated by dotted lines. The C-terminal specific domains for Ada2b-PA and Ada2b-PB are colored in green or orange, respectively. The amino acid positions for each domain are indicated by the numbers on top of each box. The percentage identity within the conserved domains in each Ada2a or Ada2b ortholog relative to the corresponding domains in *Dm*Ada2b or *Dm*Ada2a respectively is indicated by the % within each boxed domain. The % identity within the SWIRM domain is compared to *Dm*Ada2a. *Sc*Ada2 was aligned with *Dm*Ada2a.

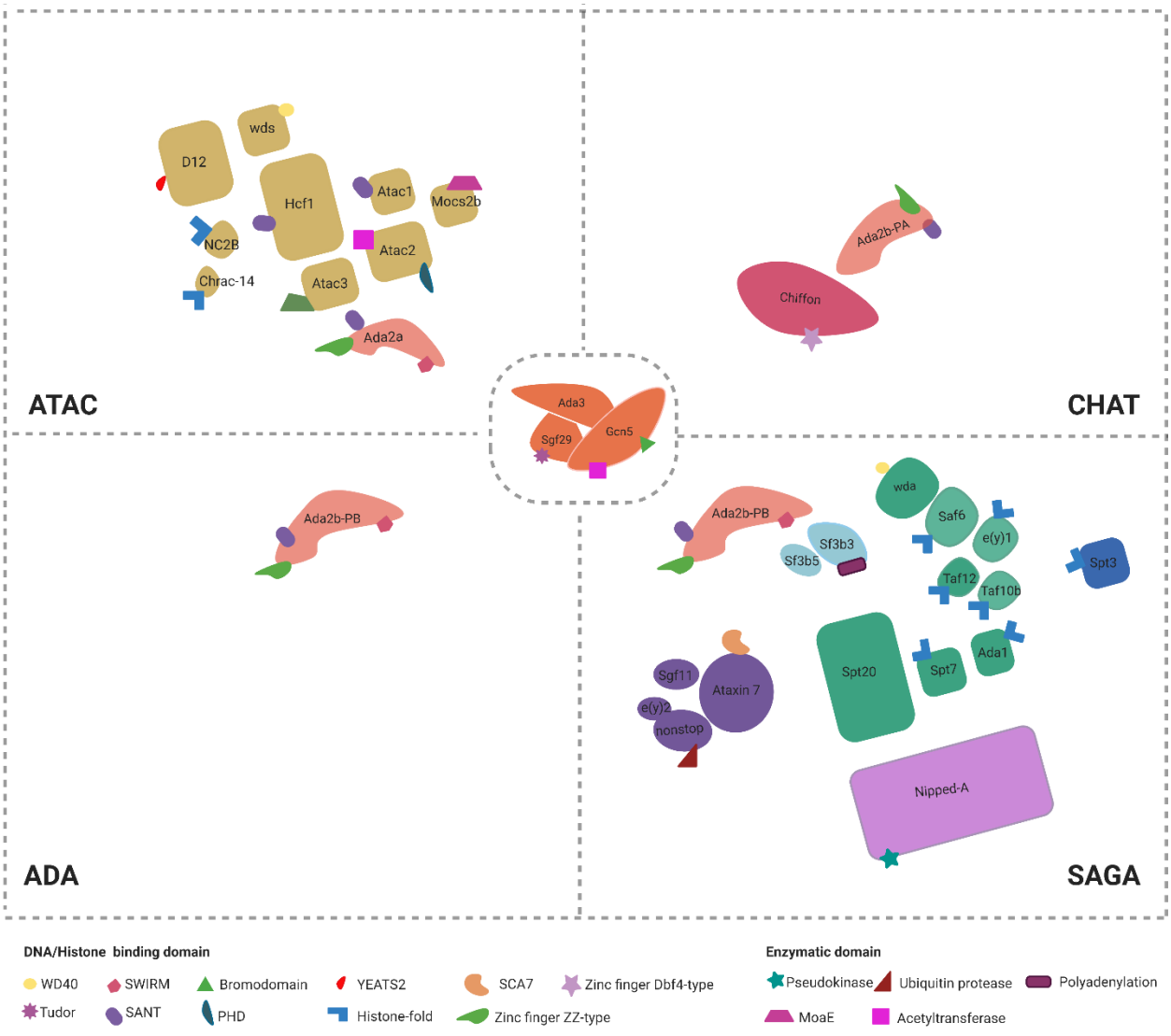

Figure 1.3 Schematic illustration of the subunit composition of SAGA, ATAC, ADA, and CHAT.

The subunits in the four Gcn5 complexes are shown in the four sections, with shared subunits of the core Gcn5 HAT module indicated in the central box. The area of each subunit is proportional to its relative molecular mass. Subunits are colored by complex, or by modules for SAGA, and the yeast Ada2 orthologs that nucleate formation of each complex shown in orange. Domains present in individual subunits are shown in the key below the figure.

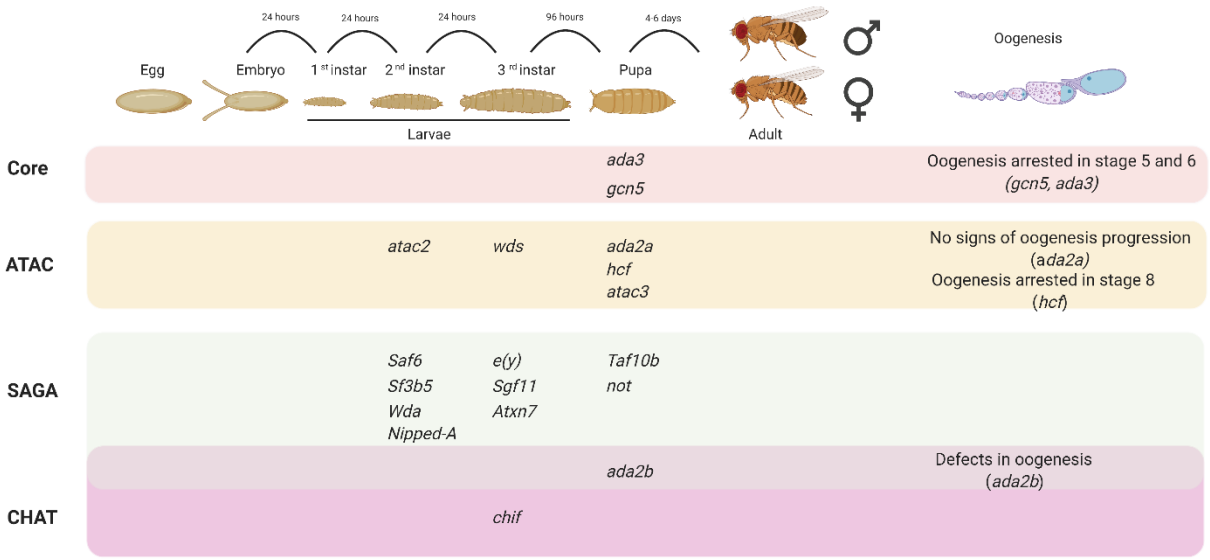

Figure 1.4 The *Drosophila* Gcn5 complexes are essential for fly development.

The life cycle of *Drosophila* comprises four successive stages, namely, egg, larva, pupa, and adult. Twenty-four hours after a female fly lays her eggs, larvae hatch. Larvae then undergo molting stages known as instars (three instar stages), during which the head, mouth, cuticle, spiracles, and hooks are shed. After ninety-six hours, the third instar larva encapsulates itself, forming a pupa. Metamorphosis takes place during the pupal stage, giving rise to all the structures in the adult fly. Oogenesis takes place within the ovary of female flies, and consists of 14 stages prior to deposition of the fertilized egg. The mutants shown disrupt subunits in the SAGA, ADA, ATAC or CHAT complexes, and result in lethality at the indicated developmental stage of the *Drosophila* life cycle. Mutations that have been shown to impact oogenesis are also indicated, but this has not been tested for all the mutant alleles shown. The *ada2b* mutant allele disrupts all three of the SAGA, ADA and CHAT complexes. The mutant alleles shown in this figure correspond to those listed in Table 1.5.

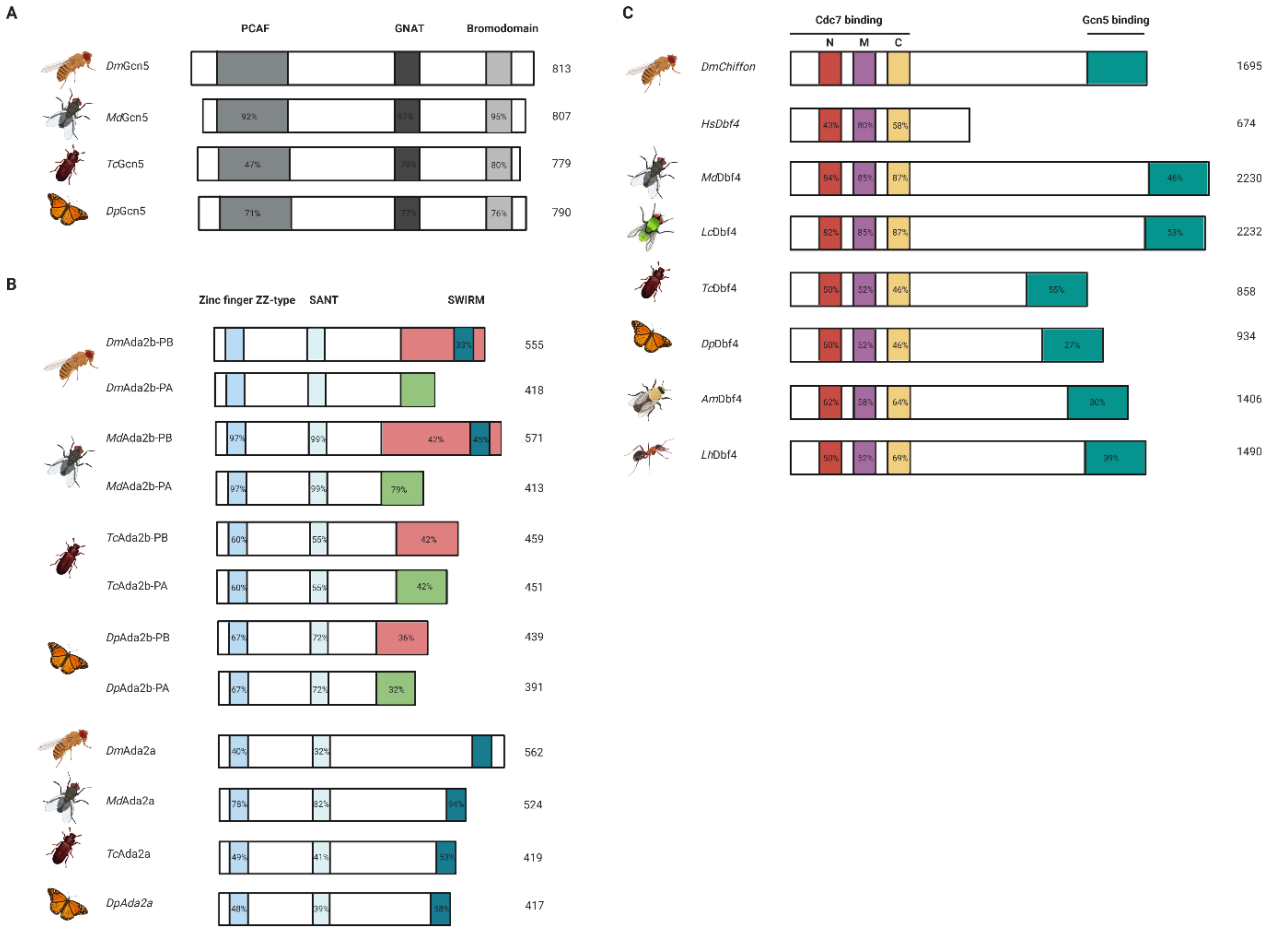

Figure 1.5 Insect Gcn5, Ada2, and Chiffon share regions of conservation with *Drosophila*.

Figure 1.5 continued

Insect Gcn5, Ada2, and Chiffon homologs were aligned using Clustal Omega. The insect species described in this figure are: Diptera, *D. melanogaster, Musca domestica* (House fly), *Lucilia cuprina* (Australian sheep blowfly); Coleoptera, *Tribolium castaneum* (Red flour beetle); Lepidoptera, *Danaus plexippus* (Monarch butterfly); Hymenoptera, *Apis Mellifera* (Western honey bee) and *Linepithema humile* (Argentine ant). A representative illustration of each insect is shown next to each aligned protein. (A) Accession numbers for Gcn5 homologs from the following insect species were used to generate this alignment: *D. melanogaster* NP_648586.2; *M. domestica* XP_005181707.1*; T. castaneum* XP_015835856.1; *D. plexippus* DPOGS216125. The GNAT, Bromodomain, and PCAF domains are boxed in gray. The percentage identity within the conserved domains in each Gcn5 ortholog relative to the corresponding domains in *Dm*Gcn5 is indicated by the % within each boxed domain. (B) Accession numbers for Ada2 homologs from the following insect species were used to generate this alignment: *D. melanogaster* Ada2b-PB NP_001027151.1, Ada2b-PA NP_6497731, Ada2a NP_001014636.1; *M. domestica* Ada2b-PA XP_005185291.1, Ada2b-PB XP_005186290.1, Ada2a XP_019894005.1; *T. castaneum* Ada2b-PA A0A139WFG5, Ada2b-PB XP_008195462, Ada2a XP_015835543.1, *D. plexippus* Ada2b-PA XP_032521398.1, Ada2b-PB XP_032521398.1, Ada2a XP_032528769.1. The Zinc finger ZZ-type, SANT, and SWIRM domains are boxed. The C-terminal specific domains for Ada2b-PA and Ada2b-PB are colored in green or orange, respectively. The percentage identity within the conserved domains in each Ada2a or Ada2b ortholog relative to the corresponding domains in *Dm*Ada2b or *Dm*Ada2a, respectively, is indicated by the % within each boxed domain. The % identity within the SWIRM domain is compared to *Dm*Ada2a. (C) Accession numbers for Dbf4/Chiffon homologs from the following species were used to generate this alignment: *D. melanogaster* AAD48779.1; *M. domestica* XP_019893793.1; *L. cuprina* A0A0L0CBC7; *T. castaneum* XM_008199666.2; *D. plexippus* OWR45390.1; *A. mellifera* XP_016770645.1; *L. humile* XP_012229084; *H. sapiens* NP_006707. The highly conserved region that interacts with Cdc7 (N, M, C domains) and the insect-specific Gcn5-binding domain are boxed. The percentage identity within the conserved domains in each Dbf4 ortholog relative to the corresponding domain in *Dm*Chiffon is indicated by the % within each boxed domain.

CHAPTER 2. THE *DROSOPHILA* DBF4 ORTHOLOG CHIFFON FORMS A COMPLEX WITH GCN5 THAT IS NECESSARY FOR HISTON ACETYLATION AND VIABILITY

Declaration of collaborative work
The work described in this chapter was the collaborative effort of Eliana F. Torres-Zelada, Robert Stephenson, Aktan Alpsoy, Emily Dykhuizen, Benjamin D. Anderson, Laurence Florens, and Michael P. Washburn, under the supervision of Dr. Vikki Weake. Robert Stephenson performed some of the immunostaining and western blot analysis, Aktan Alpsoy performed the immunoprecipitation analysis in human cells, Selene K. Swanson, Laurence Florens, and Michael P. Washburn performed mass spectrometry analysis. All other experiments were performed by Eliana F. Torres-Zelada or Dr. Vikki Weake.

2.1 Introduction

Chromatin modifications impact both transcription and cell cycle events such as DNA replication (Li et al. 2007; Ma et al. 2015). In particular, histone acetylation contributes to transcription, and correlates with the timing of the initial step in DNA replication, origin firing. The histone acetyltransferase (HAT) Gcn5 stimulates transcription by generating a permissive chromatin environment that facilitates chromatin remodeling by complexes such as SWI/SNF (Hassan et al. 2002; Weake and Workman 2010). Gcn5 also stimulates origin firing when tethered to a late-firing origin in yeast (Vogelauer et al. 2002) and enhances the rate of DNA synthesis from a chromatin template *in vitro* (Kurat et al. 2017). In *Saccharomyces cerevisiae*, Gcn5's function in transcription is mediated predominantly through the Spt-Ada-Gcn5 acetyltransferase (SAGA) complex (Grant et al. 1997). During evolution, there has been an expansion in the diversity of Gcn5-containing complexes (Spedale et al. 2012). In metazoans, including *Drosophila*, there are two homologs of the Gcn5-binding protein Ada2, Ada2a and Ada2b, which nucleate formation of the Ada2a-containing (ATAC) or SAGA transcription coactivator complexes respectively (Kusch et al. 2003; Wang et al. 2008). All Gcn5 complexes share Sgf29 and Ada3 subunits, which together with their respective Ada2 homolog, enable nucleosomal HAT activity (Balasubramanian et al. 2002; Grant et al. 1997; Spedale et al. 2012).The SAGA and ATAC complexes activate transcription during development and in response to signaling pathways or external stimuli (Spedale et al. 2012). In addition, ATAC has roles in cell cycle progression via acetylation of cyclin A, which promotes progression through mitosis (Orpinell et al. 2010).

In *Drosophila,* Ada2b has two splice isoforms that differ in their C-terminal regions but share a common N-terminal region (Pankotai et al. 2013; Qi et al. 2004). The short Ada2b-PA isoform binds Gcn5, and is necessary for histone H3 acetylation *in vivo* (Pankotai et al., 2013; Qi et al., 2004). Both Ada2b isoforms are required to fully complement *ada2b* mutations, although expression of the short Ada2b-PA isoform alone can support development and partially restore histone H3 acetylation (Pankotai et al. 2013). Although it has been assumed that the short Ada2b-PA isoform functions as part of the SAGA complex, our previous studies showed that the long Ada2b-PB splice isoform is associated with *Drosophila* SAGA (Weake et al. 2009). In this study, we describe a novel Gcn5-containing complex nucleated by the short Ada2b-PA isoform that contains the cell cycle regulatory protein, Chiffon: the Chiffon Histone Acetyltransferase, CHAT complex. Chiffon is the *Drosophila* ortholog of Dbf4 (Landis and Tower 1999), and Chiffon binds and activates the cell cycle kinase Cdc7 (Stephenson et al. 2015). In yeast, Dbf4 and Cdc7 form the Dbf4-Dependent Kinase (DDK) complex that phosphorylates the Mcm helicase to initiate DNA replication (Lei et al. 1997; Weinreich and Stillman 1999). While Dbf4 is essential for DNA replication in most organisms, previous studies found that *chiffon* null mutants were viable (Landis and Tower 1999). Here, we show that indeed the Cdc7-binding activity of Chiffon is dispensable in flies. However, the C-terminal insect-specific domain of Chiffon that nucleates formation of the CHAT complex is required in flies for both histone H3 acetylation and viability, but not for DNA replication. Our data demonstrate that the DNA replication and histone acetylation activities of Chiffon can be genetically separated, raising the question of why these two activities are encoded by the same gene. One possibility, although not tested in this study, is that the DDK and CHAT complexes are encoded as part of the same gene to coordinate their expression and/or levels during the cell cycle or development. This could provide a mechanism to coordinate histone acetylation with DNA replication, potentially during particular developmental stages in flies.

2.2 Identification of a novel Chiffon-Gcn5 complex in *Drosophila*

Drosophila Ada2b has two splice isoforms that differ in their C-terminal regions but share a common N-terminal region containing the zinc finger-like ZZ, SANT and two of the three previously described ADA box domains (Pankotai et al. 2013; Qi et al. 2004) (Fig. 2.1A). Similar to published observations (Pankotai et al. 2013), we observed that expression of Ada2b-PA alone partially rescued adult viability in *ada2b* null mutants, although both Ada2b isoforms were

required to fully complement lethality of the null *ada2b*^{1}/*ada2b*^{842} allele combination (Table 2.1). Expression of single-copy transgenes for both Ada2b-PA and Ada2b-PB restored viability to the expected one third of *ada2b*^{1}/*ada2b*^{842} progeny (30.9±4.2%, χ^2=0.23) compared with 20.7±5.2% with expression of Ada2b-PA alone. In contrast, expression of Ada2b-PB alone did not restore viability to *ada2b*^{1}/*ada2b*^{842} progeny. These data suggested that the Ada2b isoforms have nonredundant functions in flies, and that Ada2b-PA alone can partially support development. Thus, we sought to determine whether these Ada2b isoforms were both required in SAGA, or alternatively, like Ada2a, each Ada2b isoform nucleated formation of a distinct Gcn5-containing complex. To distinguish between these alternatives, we purified the Ada2b-PA and Ada2b-PB isoforms from cultured S2 cells using tandem FLAG-HA affinity chromatography and examined the co-purifying proteins using multidimensional protein identification technology (MudPIT). Using this approach, the Ada2b isoforms could be distinguished by peptide spectra that mapped to their unique C-terminal regions. Ada2b-PB co-purified all other 19 SAGA subunits (Stegeman et al. 2016), but did not co-purify any peptide spectra specific to the short Ada2b-PA isoform (Fig. 2.1B). Similarly, SAGA-specific purifications using bait proteins such as Spt3 and Spt20 contained peptide spectra specific to Ada2b-PB, but not Ada2b-PA. Instead, Ada2b-PA co-purified Gcn5, Ada3 and Sgf29, but not Ada2b-PB or other SAGA subunits. Ada2b-PA also did not co-purify ATAC-specific subunits such as Atac1, Atac2 or D12 (Table A1). Epitope tagging of Ada2b-PA did not disrupt its interaction with SAGA because similar results were observed with Ada2b isoforms tagged at their shared N- or unique C-termini. These data suggest that Ada2b-PA associates with Gcn5, Sgf29 and Ada3 in a complex that is distinct from either ATAC or SAGA.

To identify other proteins in this Ada2b-PA complex, we examined the MudPIT data to find proteins that co-purified specifically with Ada2b-PA, but not with SAGA-specific subunits. A single protein, Chiffon (CG5813; FBgn0000307), co-purified with Ada2b-PA or Sgf29, but not with other SAGA subunits or with the negative controls (Fig. 2.1B; Table A1). Moreover, reciprocal purifications of C-terminally tagged Chiffon co-purified Gcn5, Ada3, Sgf29 and Ada2b-PA, but not Ada2b-PB. Chiffon is the *Drosophila* ortholog of Dbf4, which binds and activates the cell cycle kinase Cdc7 to phosphorylate the Mcm helicase, initiating DNA replication (Landis and Tower 1999; Stephenson et al. 2015). The Chiffon-purified or Ada2b-PA-purified complexes exhibited similar levels and specificity of *in vitro* histone acetyltransferase activity to *Drosophila* SAGA purified via Ada2b-PB, with predominant activity on histone H3 in core

histones (Fig. 2.1C) and ability to acetylate histone H3 tail peptides (Fig. 2.1D). Thus, we conclude that Chiffon is a bona fide subunit of a novel histone acetyltransferase complex containing Ada2b-PA, Gcn5, Sgf29 and Ada3 that we named the CHAT complex.

2.3 Most CHAT complexes do not contain Cdc7

Because Chiffon is the regulatory subunit of the cell cycle kinase Cdc7 (Stephenson et al. 2015), we next asked whether Cdc7 was present in the CHAT complex. Only seven peptide spectra were identified for Cdc7 using C-terminally tagged Chiffon as bait, which co-purified 89–158 peptide spectra for each of the other subunits of the CHAT complex: Ada2b-PA, Gcn5, Sgf29 and Ada3 (Fig. 2.1B). Thus, we next asked whether Chiffon did, in fact, bind Cdc7 *in vivo*. We previously showed that the N-terminal domain of Chiffon [1–400 amino acids (aa)] is sufficient to bind and stimulate Cdc7 kinase activity *in vitro* (Stephenson et al. 2015). Indeed, Cdc7 and Chiffon interact *in vivo* because N-terminally tagged Chiffon co-purified 93 peptide spectra for Cdc7, and Cdc7 reciprocally co-purified 176 peptide spectra for Chiffon. In contrast, N-terminally tagged Chiffon or Cdc7 co-purified fewer than 13 peptide spectra for other components of the CHAT complex such as Gcn5. There are two possibilities for the mutually exclusive binding of Cdc7 and CHAT subunits with Chiffon; first, Cdc7 blocks binding of CHAT subunits to Chiffon; or second, *chiffon* encodes two separate polypeptides that interact with Cdc7 or CHAT independently. Our data support the latter possibility because most peptide spectra for C-terminally tagged Chiffon map to its C-terminal region, whereas most peptide spectra for N-terminally tagged Chiffon map to its N-terminal region (Fig. 2.1E). These data suggest that very little full-length Chiffon exists in asynchronous cultured cells. However, a small fraction of Chiffon might interact with both Cdc7 and CHAT because we identified a few peptide spectra corresponding to Cdc7 in CHAT purifications (Ada2b-PA and Chiffon-C). Similarly, we also observed a few peptide spectra for CHAT subunits in Cdc7 or Chiffon-N purifications. Thus, we conclude that although a small fraction of Chiffon protein might interact with Cdc7 and CHAT simultaneously, most Chiffon interacts separately with either Cdc7 or CHAT, likely as two independent Chiffon polypeptides.

2.4 The insect-specific C-terminal domain of Chiffon directly binds Gcn5

To test whether the N- and C-terminal domains of Chiffon could interact independently with Cdc7 and CHAT subunits, as suggested by the mass spectrometry data, we used a yeast two-hybrid approach to screen for interactions between different domains of Chiffon and each CHAT subunit (Fig. A.1). Using this approach, we identified a strong reciprocal interaction between the N-terminal Chiffon domain (1–400 aa) and Cdc7 (Fig. 2.2A). This is consistent with our MudPIT data and with the previous observation that the N-terminal 400 aa of Chiffon are sufficient to bind Cdc7 *in vitro* (Stephenson et al. 2015). We also observed a weak unidirectional interaction between Gcn5 and the C-terminal Chiffon domain (1243–1695 aa) (Fig. 2.2A). We were able to co-immunoprecipitate the recombinant C-terminal region (1400–1695 aa) of Chiffon and Gcn5 under low-salt (150 mM NaCl) conditions *in vitro* (Fig. 2.2B), further suggesting that Chiffon and Gcn5 interact directly, albeit weakly.

The C-terminal domain of Chiffon in *Drosophila* and other insects is much longer than that of other Dbf4 homologs and is not present in yeast or vertebrate Dbf4 (Tower 2004) (Fig. 2.2C). Moreover, the C-terminal region of Chiffon that binds Gcn5 shares several highly conserved regions with other insects (Fig. A.2). Because this insect-specific C-terminal domain interacted with Gcn5, we predicted that yeast or mammalian Dbf4 would be unlikely to interact with Gcn5. Indeed, we did not observe any peptides for Ada2, Gcn5, Sgf29 or Ada3 in TAP-purified Dbf4 from yeast cells (Fig. 2.3A). Moreover, human DBF4A and DBF4B co-immunoprecipitated CDC7, but not GCN5 (also known as KAT2A) or its paralog, PCAF (also known as KAT2B), from human embryonic kidney (HEK) 293T cells (Fig. 2.3B). Thus, the insect-specific C-terminal domain of Chiffon interacts directly with Gcn5, while the conserved N-terminal domain of Chiffon binds Cdc7.

Because Gcn5 is a component of all three of the SAGA, ATAC and CHAT complexes in flies, and because Gcn5 binds the C-terminal domain of Chiffon, we wondered why Chiffon did not associate with either the SAGA or ATAC complexes (Fig. 2.1B; Table A1). To answer this question, we examined the interaction of each Ada2b isoform with all SAGA subunits except Nipped-A (also known as Tra1) using yeast two-hybrid analysis (Fig. A.3). Using this approach, we found that Ada2b-PB, but not Ada2b-PA, auto-activated when fused to the Gal4 DNA-binding domain. This suggests that Ada2b-PB, but not Ada2b-PA, might interact with yeast transcriptional coactivators like SAGA to activate expression of the reporter genes in this assay. We also observed

that Ada2b-PA interacted with the CHAT subunits Gcn5 and Ada3, and surprisingly also with the SAGA-specific subunit Spt7 (Fig. A.3A). Further, Ada2b-PB fused to the Gal4 activating domain interacted with two additional SAGA-specific subunits that did not interact with Ada2b-PA: Spt3 and TAF12 (Fig A3B). These data suggest a model in which the unique C-terminal region of the Ada2b-PB isoform binds SAGA through Spt3 and TAF12, enhancing binding of Spt7 to the Ada2b-PB N-terminal, which precludes Gcn5 binding to Chiffon (Fig. A.3C). The Ada2b-PA isoform lacks the C-terminal region necessary for binding Spt3 and TAF12, preventing stable binding of Spt7 to the N-terminal of Ada2b-PA, and instead enabling Gcn5 to bind Chiffon. This model is partially based on the observation that Ada2b-PA did not interact with Spt7 in our mass spectrometry data, even though it is capable of binding Spt7 by yeast two-hybrid assay. We further suggest that the Ada2b-PB C-terminal domain might also be capable of interacting with yeast SAGA, potentially via yeast Spt3 and TAF12. Although Ada2b-PA and Ada2b-PB interacted in one direction by yeast two-hybrid assay (Fig. A.3B), our MudPIT data indicate that the Ada2b isoforms are not present in the same complex *in vivo* (Fig. 2.1B; Table A1). We conclude that the unique C-terminal regions of the Ada2b isoforms control protein–protein interactions that determine formation of the SAGA or CHAT complexes, and that the extended C-terminal domain in Ada2b-PB is necessary for SAGA formation (Fig. 2.3C).

2.5 Chiffon is necessary for histone H3 acetylation *in vivo*.

Since the CHAT complex exhibits *in vitro* histone acetyltransferase activity against histone H3, we next asked whether Chiffon was necessary for proper histone H3 acetylation *in vivo*. To do this, we used *Drosophila* ovary follicle cells in which specific regions of the genome undergo repeated bidirectional replication initiation to increase DNA copy number (Spradling and Mahowald 1980). Chiffon is necessary for gene amplification in these cells (Zhang and Tower 2004; Landis and Tower 1999; Stephenson et al. 2015). We generated somatic mosaics for the $chiffon^{ETBE3}$ null allele in ovaries using the FLP/FRT system and examined levels of different histone H3 acetyl marks by immunostaining. Notably, $chiffon^{ETBE3}$ mutant cells showed decreased levels of histone H3 acetylated at lysine 14 (H3K14ac) relative to the adjacent GFP-positive cells (Fig. 2.4A). We also observed decreased levels of H3K9ac and H3K18ac, but not H3K23ac, in $chiffon^{ETBE3}$ mutant cells (Fig. 2.4B). H3K18ac levels were only modestly reduced in $chiffon^{ETBE3}$ clones, consistent with data showing that p300/CBP (encoded by *nejire*) is the

major histone acetyltransferase for H3K18 in *Drosophila* and in mammalian cells (Jin et al. 2011; Tie et al. 2009). H3K14ac levels within *chiffon*^{ETBE3}, but not control *FRT40A*, clones were reduced to ~50% of the surrounding tissue (Fig. 2.4A, C). The nuclei in some *chiffon*^{ETBE3} clones appeared slightly more condensed using 4′,6-diamidino-2-phenylindole (DAPI) staining, suggesting that the reduced histone acetylation could be due to decreased DNA content in these cells. However, we and others have previously shown that *chiffon* is not essential for endoreplication, which determines the ploidy of follicle cells (Zhang and Tower 2004; Landis and Tower 1999; Stephenson et al. 2015). Moreover, some H3 acetyl marks such as H3K23ac were not reduced in *chiffon*^{ETBE3} clones (Fig. 2.4B). Together, these data demonstrate that Chiffon is required for full levels of histone H3 acetylation at lysines 9, 14 and 18 *in vivo*, suggesting that the CHAT complex contributes to bulk levels of histone H3 acetylation in flies. Interestingly, *ada2b*^{1} clones showed only slightly lower levels of H3K14ac (not significant) when compared with *chiffon*^{ETBE3} clones, despite the fact that the *ada2b*^{1} allele removes both Ada2b isoforms (Fig. 2.4A, C). This suggests that the CHAT complex, rather than SAGA, might contribute to the majority of histone H3 acetylation in ovary follicle cells. To test whether CHAT was also required in other cell types for histone H3 acetylation, we compared H3K14ac levels in *chiffon*^{ETBE3} and *ada2b*^{1} clones from imaginal discs. Similar to ovary follicle cells, most *chiffon*^{ETBE3} and *ada2b*^{1} clones from imaginal discs showed decreased levels of H3K14ac (Fig. A.4). Although some large *chiffon*^{ETBE3} clones appeared to have little or no H3K14ac staining, other *chiffon*^{ETBE3} clones showed only moderate decreases in H3K14ac more similar to those observed in ovary follicle cells. Notably, some *chiffon*^{ETBE3} clones also appeared to contain fewer nuclei, suggesting that Chiffon might also contribute to cell number, potentially through DNA replication, in this cell type. While some *ada2b*^{1} clones showed only slight decreases in H3K14ac, other clones showed similar levels of H3K14ac to those observed in *ada2b*^{1} ovary follicle clones. Previous studies showed that mutations in the SAGA-specific subunit, *wda*, strongly reduced H3K9ac levels in *Drosophila* embryos (Guelman et al. 2006b), suggesting that SAGA is necessary for full levels of histone H3 acetylation in embryos. It remains unclear whether SAGA and CHAT have overlapping or specialized functions with regard to histone H3 acetylation in *Drosophila*. However, our observation that the CHAT-specific Ada2b-PA isoform is sufficient to partially restore viability to *ada2b* null flies suggests that CHAT might compensate for some of SAGA's essential functions during development. Overall, these data indicate that *chiffon* is

required for histone H3 acetylation *in vivo*, and that the CHAT complex contributes to histone H3 acetylation in several tissues in flies.

2.6 CHAT-mediated histone acetylation is not required for gene amplification

Histone acetylation correlates with and contributes to localized replication at the amplified follicle cell origins (Aggarwal and Calvi 2004; Liu et al. 2012; McConnell et al. 2012). Moreover, mutations in *chiffon* eliminate gene amplification in follicle cells (Zhang and Tower 2004; Landis and Tower 1999; Stephenson et al. 2015). Therefore, we asked whether CHAT-mediated histone acetylation was also necessary for gene amplification. As observed previously, $chiffon^{ETBE3}$ clones lack the characteristic 5-bromo-2-deoxyuridine (BrdU) foci indicative of chorion gene re-replication that are present in the wild-type cells adjacent to the clone or in the *FRT40A* control clone (Fig. 2.4A, D). To test whether CHAT-mediated histone acetylation was required for gene amplification, we examined $ada2b^{1}$ somatic ovary mosaics, which exhibit decreased levels of histone H3 acetylation, similar to that observed in $chiffon^{ETBE3}$ mutant cells (Fig. 2.4A, C). In contrast to $chiffon^{ETBE3}$ clones that lack detectable gene amplification, we observed multiple $ada2b^{1}$ clones undergoing gene amplification (Fig. 2.4A, D). We note that there were more pyknotic nuclei in $ada2b^{1}$ clones, suggesting that loss of both Ada2b isoforms increased cell death in follicle cells, potentially due to pleiotropic effects resulting from loss of both the SAGA and CHAT complexes. Supporting this, loss of *ada2b* in the germline cells of female flies results in arrested oogenesis and increased apoptosis, suggesting that proper histone acetylation is necessary for other aspects of egg development (Li et al. 2017). Despite this, these data suggest that Ada2b-PA, which is necessary for histone acetyltransferase activity of the CHAT complex, is not required for gene amplification.

Because our MudPIT data and binding studies suggested that the N- and C-terminal domains of Chiffon interacted independently with Cdc7 and the CHAT complex, respectively, we wondered whether expression of these domains would restore either gene amplification or histone acetylation in *chiffon* mutants. To test this, we generated flies expressing either full-length Chiffon (1–1695 aa, Chiffon-FL), or its N-terminal (1–375 aa, Chiffon-N) or C-terminal (401–1695 aa, Chiffon-C) domains. Each Chiffon construct was expressed under the control of *chiffon* genomic regulatory elements from transgenes inserted in the third chromosome *attP2* site (Fig. 2.5, Materials and Methods). Both H3K14ac levels and gene amplification were restored in $chiffon^{ETBE3}$ mutant

clones by expression of a single copy of full-length Chiffon (Fig. 2.4A, C, D). In contrast, the N-terminal Chiffon transgene rescued gene amplification, but not histone acetylation, in *chiffon*^{ETBE3} clones. Further, the C-terminal Chiffon transgene partially rescued histone acetylation, but did not restore gene amplification, in *chiffon*^{ETBE3} clones. To our surprise, a full-length Chiffon transgene that contained a stop codon at position 376 (Chiffon-FL*), separating the N-terminal Cdc7-binding domain from the C-terminal Gcn5-binding domain, fully restored both gene amplification and histone acetylation in *chiffon*^{ETBE3} clones. As histone acetyltransferases function redundantly to stimulate follicle cell gene amplification (McConnell et al. 2012), we cannot exclude the possibility that CHAT functions redundantly with other histone acetyltransferases to stimulate origin activity. Indeed, although bulk H3K14ac was reduced in *chiffon*^{ETBE3} clones expressing the N-terminal Chiffon transgene, we observed residual H3K14ac foci that co-localized with the BrdU foci in half of the images (five of the ten images) analyzed for acetylation in this genotype. We also observed H3K14ac foci that co-localized with the BrdU foci in some of the *ada2b* clones (three of the nine images), but these were much fainter than those present in *chiffon*^{ETBE3} clones expressing the N-terminal Chiffon transgene. This suggests that other histone acetyltransferases target the amplified follicle cell origins in the absence of CHAT, likely including SAGA. Thus, we conclude that the histone acetyltransferase activity of the CHAT complex alone is not essential for the specialized gene amplification form of DNA replication that occurs in follicle cells.

2.7 CHAT-mediated histone acetylation is essential for viability in flies

To our surprise, a premature stop codon in one of the *chiffon* transgenes (Chiffon-FL*), that should have truncated the protein prior to the Gcn5-binding region, fully rescued histone acetylation in *chiffon* mutant cells. These data implied that there could be an internal translation start site within the single, large exon in the *chiffon* gene (Fig. 2.5). Indeed, we identified a potential consensus initiation codon sequence 393 aa from the end of the *chiffon* coding region that would be expected to generate a ~43 kDa polypeptide. Although *chiffon* has been reported to be dispensable for viability in flies, these conclusions were based largely on an allele containing a nonsense mutation at position 174, *chiffon*^{WF24} (Landis and Tower 1999). This mutation disrupts the Cdc7-binding domain and results in viable flies with phenotypes indicative of partially disrupted DNA replication, such as rough eyes and female infertility. Although these data

suggested that the Cdc7-binding activity of Chiffon was not essential for viability in flies, we wondered whether this was also the case for the CHAT complex. Because Ada2b-PA was sufficient to partially restore viability to *ada2b* mutants (Table 2.1), we hypothesized that the CHAT complex is essential for development in flies.

To test this, we used CRISPR-Cas9 technology to generate a new null *chiffon* allele in which the entire *chiffon* coding region was replaced with a visible eye marker, 3xP3-DsRed (Fig. 2.5, $chiffon^{DsRed}$). We then crossed these $chiffon^{DsRed}$ flies with the $chiffon^{ETBE3}$ null allele generated by Landis and Tower 1999, or the *Df(2L)RA5* deficiency that spans the *chiffon* gene and removes several adjacent genes. Lethality in the $chiffon^{ETBE3}$ flies was previously attributed to a secondary mutation in the nearby *cactus* gene, which is also missing in the *Df(2L)RA5* deficiency. However, we found that combinations of any of these three *chiffon* alleles resulted in complete adult lethality (Table 2.2). We then expressed single-copy *chiffon* rescue transgenes expressing full-length Chiffon (Chiffon-FL) with or without the $chiffon^{WF24}$ mutation (174Q>X). If the *chiffon* rescue transgene fully restored Chiffon function, we would expect to observe one third of adult progeny lacking the balancer chromosome. Moreover, we would expect that female adult progeny with restored Chiffon function would be fertile due to restoration of Chiffon activity in ovary follicle cells. Indeed, expression of the wild-type full-length Chiffon transgene fully restored both viability and female fertility in all three allele combinations (Table 2.2). Moreover, similar to Landis and Tower (1999), the Chiffon-FL^{WF24} transgene fully restored viability, but not female fertility, in the $chiffon^{DsRed}/chiffon^{ETBE3}$ progeny; similar results were observed with the other *chiffon* allele combinations. Thus, the $chiffon^{WF24}$ mutation, which disrupts the Cdc7-binding domain of Chiffon, eliminates Chiffon function with respect to female fertility, but does not disrupt Chiffon's role in adult viability. Next, we asked whether expression of the Chiffon-C domain, which partially rescued histone acetylation in *chiffon* clones but did not restore gene amplification, could restore adult viability. Indeed, consistent with the observations for the Chiffon-FL^{WF24} transgene, expression of the Chiffon-C domain restored viability, although not to the same extent as Chiffon-FL, but the resulting females were infertile. In contrast, expression of the Chiffon-N transgene did not restore viability, even though this transgene did rescue gene amplification in *chiffon* clones (Fig. 2.4A, D). Further supporting the possibility that *chiffon* contains an internal translation start site, a full-length Chiffon transgene that contained a stop codon at position 376 (Chiffon-FL*), separating the N-terminal Cdc7-binding domain from

the C-terminal insect-specific region, fully complemented both viability and fertility in *chiffon* mutants. Thus, we asked whether the Cdc7- and Gcn5-binding domains of Chiffon could function *in trans*. To test this, we expressed single copies of the Chiffon-N and Chiffon-C in combination, and found that this fully restored both viability and female fertility to *chiffon* mutants (Table 2.2). These data demonstrate that Chiffon, like Ada2b-PA, is essential for viability in flies. Moreover, the essential function of Chiffon relates to its histone acetyltransferase activity rather than Cdc7 activation.

Our genetic observations support the possibility that *chiffon* is a dicistronic gene that encodes two distinct polypeptides; although this type of gene structure is relatively rare in *Drosophila*, there are several examples of dicistronic genes in flies, including *stoned A* and *stoned B*, and *Adh* (Andrew et al. 1996; Brogna and Ashburner 1997; Komonyi et al. 2009). The coding sequence for the 1695-aa Chiffon protein encoded by the RD or RB transcripts lies within a single exon (Fig. 2.5), and northern blot analysis previously identified a single 6.5 kb *chiffon* transcript (Landis and Tower 1999), suggesting that alternative splicing is unlikely to account for our observations. To test whether a C-terminal product was generated from either of the Chiffon transgenes that contained premature stop codons, we immunoprecipitated the Chiffon-FL* and Chiffon-FL^{WF24} proteins via their C-terminal FLAG epitope tags, and performed western blotting with anti-FLAG antibodies. We observed a ~48 kDa product that was recognized by anti-FLAG antibodies in both the immunoprecipitations from Chiffon-FL* and Chiffon-FL^{WF24} lysates, but not those from untagged embryo lysates (Fig. 2.6A). Further, both Chiffon-FL* and Chiffon-FL^{WF24} co-immunoprecipitated Gcn5, suggesting that the C-terminal product expressed by these transgenes interacted with the CHAT complex. Thus, the Chiffon C-terminal domain nucleates CHAT formation, can be expressed from an alternative translation start site in the *chiffon* gene and is essential for viability in *Drosophila* (Fig. 2.6B).

2.8 Discussion

Here, we show that the *Drosophila Dbf4* ortholog *chiffon* is a dicistronic gene that encodes two distinct polypeptides from alternative translation start sites. Chiffon's two activities can be separated genetically; its N-terminal domain binds Cdc7 and its C-terminal domain binds the histone acetyltransferase Gcn5 (Fig. 2.6B). The interaction between Chiffon and Gcn5 forms the CHAT complex that is required for histone H3 acetylation and viability in flies. Thus, in addition

to the Gcn5-containing SAGA and ATAC complexes, flies contain a third Gcn5-containing complex: CHAT. The CHAT complex is not present in yeast or human cells, and is likely to be specific to insects because it is nucleated by Chiffon's insect-specific C-terminal domain. Our mass spectrometry data suggest that most Chiffon interacts in a mutually exclusive manner with Cdc7 and CHAT, and transgenes that separate the N- and C-terminal domains of *chiffon* fully restore both functions. Thus, our data demonstrate that the DDK and CHAT complexes function independently in DNA replication and histone acetylation, respectively.

What might be the function of this CHAT complex in flies? Gcn5 and another histone acetyltransferase, Esa1, stimulate DNA replication in yeast *in vitro* (Kurat et al. 2017). In addition, several histone acetyltransferases work together to stimulate follicle cell amplification in *Drosophila* (McConnell et al. 2012). However, our work argues against a role for the CHAT complex in DNA replication; although we cannot exclude the possibility that the CHAT complex functions redundantly with other histone acetyltransferases to stimulate DNA replication, CHAT is not essential for gene amplification in follicle cells. Because SAGA is required for proper gene expression in flies, and because the CHAT-specific Ada2b-PA isoform can restore viability to *ada2b* mutants, we propose that CHAT, like the SAGA and ATAC complexes, regulates gene expression in flies. In other organisms, Dbf4 levels fluctuate throughout the cell cycle to control activity of Cdc7 (Cheng Liang et al. 1999; Oshiro Guy et al. 1999): Dbf4 protein levels correlate with Cdc7 activity and increase at the G1–S transition, peak in S phase, and then become low during G1 phase, when Dbf4 is degraded by the anaphase-promoting complex (Cheng Liang et al. 1999; Oshiro Guy et al. 1999). One possibility in flies is that Chiffon levels are also cell cycle regulated, and if so, CHAT complex expression could be controlled by Chiffon levels, potentially peaking in S phase. Thus, the DDK and CHAT functions of *chiffon* could have evolved as part of the same gene structure to coordinate DNA replication with expression of CHAT target genes during the cell cycle in insects.

Although Dbf4 did not interact with Gcn5 in yeast or in human cells, some observations support a potential role for Dbf4 in gene expression in these organisms. For example, the C-terminal domain of human DBF4 (also known as ASK) binds the chromatin-associated protein Lens epithelium-derived growth factor (LEDGF; also known as PSIP1), which is associated with the MLL histone H3 methyltransferase complex (Hughes et al. 2010; Yokoyama and Cleary 2008). Further, the C-terminus of yeast Dbf4 binds forkhead transcription factors (Fang et al. 2017). In

addition, DDK phosphorylates Thr45 of histone H3 in budding yeast (Baker et al. 2010), demonstrating a direct role for DDK complexes in chromatin modification. Thus, although the CHAT complex might be specific to insects, Dbf4 orthologs could have a more general role in gene expression in addition to their essential DNA replication activity.

One unusual feature of Chiffon in flies is that its Cdc7-binding activity is dispensable for viability. Loss of either Dbf4 or Cdc7 disrupts DNA replication and mitosis in organisms from yeast to mammals, leading to growth defects and/or cell death (Labib 2010). In flies, *Cdc7* is also an essential gene (Stephenson et al. 2015), and recent work showed that *Cdc7* is required for early embryonic nuclear cycles, consistent with its essential role in DNA replication (Seller and O'Farrell 2018). However, our data show that Chiffon's Cdc7-binding activity is not essential for DNA replication or viability in flies, although it is required for follicle cell gene amplification in the ovary. These conclusions are consistent with the previous findings of Landis and Tower (1999), and are in stark contrast to the absolute requirement of Cdc7 and Dbf4 for DNA replication and cell viability in organisms from yeast to vertebrates (Labib 2010; Landis and Tower 1999). Despite this, *chiffon* is essential for development in flies but this is due to a requirement for the CHAT complex, likely due to its role in histone acetylation. Thus, our studies raise the question of how *Drosophila* Cdc7 can function in the absence of its Dbf4 regulatory partner, because flies do not have any other detectable sequence homolog for Dbf4. Whereas budding yeast possesses only one homolog for Dbf4 and Cdc7, several organisms possess paralogs of DDK subunits with specialized functions in meiosis and development. In particular, the vertebrate Dbf4B paralog has specialized roles in early embryogenesis (Collart et al. 2017; Silva et al. 2006; Yoshizawa-Sugata et al. 2005; Montagnoli et al. 2002). If Chiffon, like Dbf4B, has a more specialized developmental role in DNA re-replication in ovary follicle cells, then our data suggest that there might be an alternative mechanism to regulate Cdc7 activity in flies.

2.9 Material and Methods

2.9.1 Affinity purification, MudPIT analysis and histone acetyltransferase assays

Tandem FLAG-HA affinity purification and MudPIT analysis was conducted from stable *Drosophila* S2 cell lines as described previously (Stegeman et al. 2016). TAP purifications from *S. cerevisiae* was performed as described previously (Lee et al. 2004). To estimate relative

protein levels, distributed normalized abundance factors (dNSAFs) were calculated for each nonredundant protein or protein group (Zhang et al. 2010). Briefly, shared spectral counts (sSpC) were distributed based on spectral counts unique to each protein (uSpC). Histone acetyltransferase assays were performed as previously described (Stegeman et al., 2016) using Flag-purified complexes and HeLa core histones or human histone H3 peptide (K5–K23) as substrate.

2.9.2 Fly stocks and genetics

Genotypes for flies used in this study are described in Table A2. The *chiffon*^{ETBE3} (Landis and Tower 1999) and *ada2b*^{1} (Qi et al. 2004) null alleles were used for somatic mosaic analysis. The null *ada2b*^{1} and *ada2b*^{842} (Pankotai et al. 2013) alleles that disrupt both Ada2b isoforms were used to assess adult survival. Ada2b rescue transgenes contain genomic *ada2b* enhancer sequences that begin −1878 bp from the transcription start site and extend +1782 bp to the end of the second exon. The alternative exon 3 and 4 sequences for each Ada2b isoform are fused directly to the 3′ end of exon 2. Constructs were generated in the pCa4B vector with the addition of the *Adh* 3′ UTR and polyadenylation signal sequences from the *pRmaHa3* vector. Transgenic flies were generated using the phiC31 site-specific integration system in the attP40 site on chromosome 2*L*. Chiffon rescue transgenes contain genomic *chiffon* enhancer sequences that span −3480 bp relative to the translation start site of the *chiffon-RD* transcript, and include the *chiffon* 3′ UTR sequences that extend 1056 bp past the stop codon of the *chiffon-RD* transcript. Chiffon domain constructs encode the indicated number of amino acids relative to 1695 aa full-length Chiffon based on the *chiffon-RD* transcript. Chiffon constructs were N- and C-terminally epitope tagged with 2xHA and FLAG, respectively. Transgenic flies were generated in the attP2 site on chromosome 3*L*. The *chiffon*^{DsRed} allele was generated using CRISPR-Cas9 technology (Gratz et al. 2014). The following guide RNAs were used to target the *chiffon* 5088 bp exon for replacement: 5′-GGAGGGAAACTTTATAGGAGTGG-3′ and 5′-GATGATGATTAGATGACACAGGG-3′. Flanking regions immediately upstream and downstream of the *chiffon* coding region (*chiffon-RD*) were cloned into the flyCRISPR vector pHD-DsRed-attP and used as a template for homologous recombination. Flies expressing DsRed were selected, and the insertion position of the 3xP3-DsRed-attP cassette was confirmed by PCR and sequencing. The *chiffon*^{ETBE3} allele was also confirmed by PCR and sequencing. The genomic positions of the regions deleted in

each *chiffon* allele are as follows: *chiffon*^{DsRed} chr2*L*, 16344356–16349852; *chiffon*^{ETBE3} chr2*L*, 16344400–16351631.

2.9.3 Immunohistochemistry

Somatic clones were induced in egg chambers, and ovaries were dissected from adult females at 3 days posteclosion, labeled with BrdU, fixed, and immunostained with anti-BrdU (#555627, BD Pharmingen, mouse, 1:20) and either anti-H3K14ac (07-353, Millipore, rabbit, 1:100), anti-H3K9ac (ab10812, Abcam, rabbit, 1:500), anti-H3K18ac (ab1191, Abcam, rabbit, 1:400) or anti-H3K23ac (ab47813, Abcam, rabbit, 1:700) antibodies, followed by Alexa Fluor 568- and Alexa Fluor 633-conjugated secondary antibodies (Life Technologies). They were then imaged as described previously (Stephenson et al. 2015). H3K14ac levels were quantified for 10–30 nuclei in each clone relative to a similar number of nuclei from the surrounding wild-type region of the tissue (GFP positive). Acetylation levels were determined as average sum intensity values for nuclear-localized fluorescence using NIS-Elements Analysis software. Acetylation levels were quantified for four individual frames from a *z*-stack image of each egg chamber. These four frames were selected based on those images that contained the brightest H3K14ac signal in the wild-type region (GFP positive) of the egg chamber. Somatic clones were induced in imaginal discs by heat shock for 30 min at 37°C 72 h after egg laying. Imaginal discs were dissected from wandering third-instar larvae and immunostained with anti-H3K14ac.

2.9.4 Phylogenetic analysis

The following protein sequences were aligned using Clustal Omega (Sievers et al. 2011) and used to generate a neighbor-joining phylogenetic tree, which was plotted using phytools in R (Revell 2012): *Bos taurus*, XP_024836692.1 and XP_015324178.1; *Canis lupus familiaris*, XP_532451.2 and XP_022278602.1; *Homo sapiens*, NP_006707.1 and NP_663696.1; *Gallus gallus*, XP_004939326.1 and XP_004948536.1; *Xenopus laevis*, ABB16337.1 and BAC76421.1; *Mus musculus*, NP_001177646.1; *Ceratitis capitata*, XP_004521831.1; *Lucilia cuprina*, XP_023301579.1; *Drosophila melanogaster*, AAD48779.1; *Camponotus floridanus*, EFN62957.1; *Pogonomyrmex barbatus*, XP_011633258.1; *Linepithema humile*, XP_012229084.1; *Apis mellifera*, XP_016770645.1; *Apis florea*, XP_003693265.1; *Tribolium*

castaneum, XP_008197891.1; *Schizosaccharomyces pombe*, CAA19117.1 and CAB39799.1; *Aspergillus nidulans* AAD01519.1; *S. cerevisiae*, NP_010337.3; *Eremothecium gossypii*, NP_986462.1; *Kluyveromyces lactis*, XP_455609.1.

2.9.5 Yeast two-hybrid assay

Yeast two-hybrid analysis was performed with the Matchmaker Gold Yeast two-hybrid system as per the manufacturer's instructions (Clontech Laboratories). Three independent transformed colonies were replica plated on the different selective media for each interaction tested.

2.9.6 Co-immunoprecipitation and western blotting analysis

Recombinant proteins (500 ng) were incubated with glutathione-sepharose (16100, Thermo Fisher Scientific) in the following buffer: 50 mM Tris-HCl pH 8.0, 150 mM NaCl, 0.1% NP-40. Embryo lysates (4 mg protein) were immunoprecipitated using FLAG-agarose in the following buffer: 50 mM Tris-HCl pH 8.0, 300 mM NaCl, 0.5% NP-40, 10% glycerol. C-terminally FLAG-tagged DBF4A, DBF4B or pENTER empty vector control (CG801040, CH874659, P100001, Vigene Biosciences, Rockville, MD) were transiently transfected into HEK293 T cells, and nuclear lysates (1 mg protein) were immunoprecipitated using Flag M2 antibodies and Protein-G Dynabeads in the following buffer: 50 mM Tris-HCl pH 8.0, 150 mM NaCl and 0.2% IPEGAL with PMSF, aprotinin, leupeptin and pepstatin. The following antibodies were used for western blot analysis: anti-GST (PC53, Millipore, rabbit, 1:1000), anti-His-HRP (MA1-21315, Invitrogen, mouse, 1:1000), anti-FLAG-HRP (A8592, Sigma-Aldrich, mouse, 1:5000), anti-*Drosophila* Gcn5 (rabbit, 1:3000) (Kusch et al., 2003), anti-Flag M2 (Sigma-Aldrich; 1:1000), anti-human CDC7 (ab108382, Abcam, 1:1000), anti-human PCAF (ab12188, Abcam, 1:500) and anti-human GCN5 (ab153903, Abcam, 1:1000). HEK293T cells were obtained from the American Type Culture Collection and were tested for mycoplasma contamination using a MycoAlert Mycoplasma Detection Kit (Lonza).

Table 2.1 Flies carrying the indicated *ada2b* null alleles were crossed and the surviving adult progeny were scored for presence of the balancer chromosome (TM3).

Allele combination	**Transgene (hemizygous)**	**Balancer siblings**	**Rescue flies**	**Total flies**	**x^2**	**Mean ± s.d. (n ≥ 4 crosses)**	**Rescued female fertility**
$\underline{ada2b^{1}}$ / $ada2b^{842}$	No transgene	987	0	987	$2.47x10^{-109}$	0 ± 0%	ND
	Ada2b-PA	726	202	928	$7.77x10^{-14}$	20.7 ± 5.2%	Fertile
	Ada2b-PB	670	0	670	$7.83x10^{-75}$	0 ± 0%	ND
	Ada2b-PA+Ada2b-PB	498	226	724	**0.23**	30.9 ± 4.2%	Fertile

Adult progeny carried one copy of each indicated rescue transgene. We would expect one third of adult progeny to lack the balancer chromosome if $ada2b^{-/-}$ flies expressing any of the hemizygous transgenes were viable. The mean percentage of rescued flies +/- s.d. is presented for ≥ 4 independent crosses each with ≥ 100 scored flies; *p*-value, Chi-squared test. Non-significant values highlighted in bold. Fertility was examined for rescued females; ND, not determined.

Table 2.2 Flies carrying the indicated *chiffon* null alleles were crossed and the surviving adult progeny were scored for presence of the balancer chromosome (CyO).

Allele combination	**Transgene (hemizygous)**	**Balancer siblings**	**Rescue flies**	**Total flies**	**x^2**	**Mean ± s.d. (n ≥ 4 crosses)**	**Rescued female fertility**
chif^{DsRed}/*chif*^{ETBE3}	No transgene	1060	0	1060	$2.82x10^{-117}$	0 ± 0%	ND
	Chiffon-FL	618	300	918	**0.67**	33.0±3.5%	Fertile
	Chiffon-N	871	0	871	$1.03x10^{-96}$	0 ± 0%	ND
	Chiffon-C	1027	279	1306	$4.45x10^{-20}$	21.6 ± 5.1%	Infertile
	Chiffon-FL^{WF24}	1010	462	1472	**0.11**	31.5 ± 1.6%	Infertile
	Chiffon-FL*	1093	563	1656	**0.57**	34.0 ± 1.4%	Fertile
	Chiffon-N+ Chiffon-C	556	251	807	**0.18**	31.4 ± 2.4%	Fertile
chifDsRed/Df(2L)RA5	No transgene	1095	0	1095	$4.40x10^{-121}$	0 ± 0%	ND
	Chiffon-FL	1026	546	1572	**0.24**	34.7±1.4%	Fertile
	Chiffon-N	1103	0	1103	$5.94x10^{-122}$	0 ± 0%	ND
	Chiffon-C	1420	443	1863	$2.17x10^{-18}$	23.7 ± 4.4%	Infertile
	Chiffon-FL^{WF24}	575	262	837	**0.21**	31.9 ± 3.6%	Infertile
	Chiffon-FL*	852	402	1254	**0.34**	31.7 ± 3.5%	Fertile
chifETBE3/Df(2L)RA5	No transgene	1319	0	1319	1.92E-145	0 ± 0%	ND
	Chiffon-FL	455	262	717	**0.07**	36.7±3.9%	Fertile
	Chiffon-N	820	0	820	$3.67x10^{-91}$	0 ± 0%	ND
	Chiffon-C	1310	260	1570	$3.90x10^{-45}$	16.2 ± 5.4%	Infertile

Table 2.2 continued

	Chiffon-FL^{WF24}	544	236	780	**0.07**	30.2 ± 2.6%	Infertile
	Chiffon-FL*	1172	357	1529	$1.21x10^{-16}$	23.5 ± 2.3%	Fertile

Adult progeny carried one copy of each indicated rescue transgene. We would expect one third of adult progeny to lack the balancer chromosome if *chiffon*$^{-/-}$ flies expressing any of the hemizygous transgenes were viable. The mean percentage of rescued flies +/- s.d. is presented for ≥ 4 independent crosses each with ≥ 100 scored flies; *p*-value, Chi-squared test. Non-significant values highlighted in bold. Fertility was examined for rescued females; ND, not determin

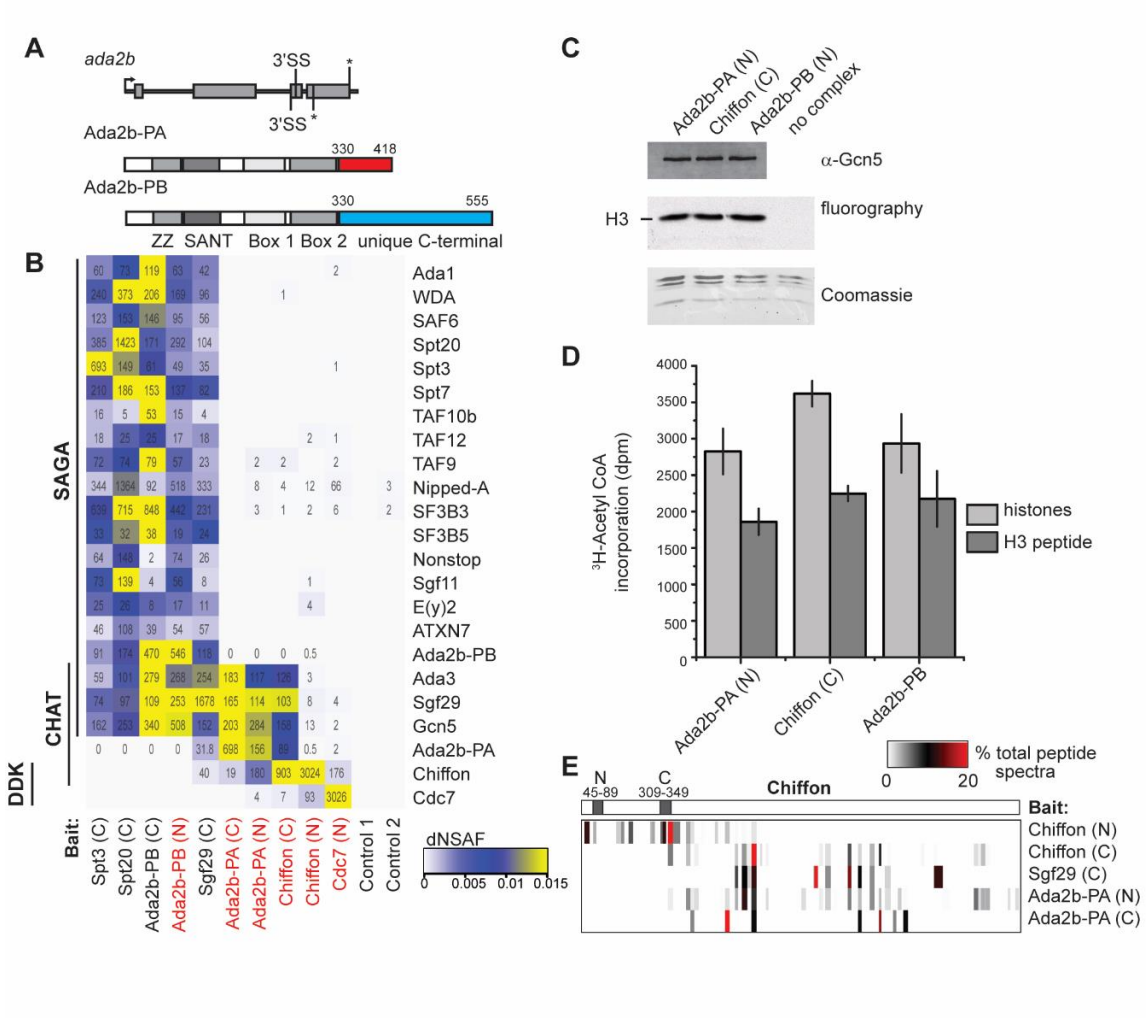

Figure 2.1 Identification of a novel Chiffon-Gcn5 complex in *Drosophila.*

(A) The *ada2b* gene encodes two splice isoforms, Ada2b-PA and Ada2b-PB, resulting from alternative use of 3′ splice acceptor sites in exon 3 (splice site, SS) that generate a frameshift following amino acid 330 (asterisks indicate stop codons). Ada2b isoforms differ only in their highlighted C-terminal regions (red/blue). (B) Heat map showing the relative spectral abundance of SAGA subunits, Chiffon and Cdc7 expressed as distributive normalized spectral abundance factor (dNSAF) in tandem FLAG-HA purifications from S2 cells using the indicated bait proteins (N/C epitope tag shown in brackets). Control 1, untagged; Control 2, CG6459 (nonspecific bait). Bait proteins new to this study are highlighted in red. The dNSAF scale represents abundance of subunits on a scale from yellow (high) to blue (low), with subunits that were not identified shown in white. dNSAF values used to generate the heat map are provided in Table A1. The numbers of spectra specific to each protein isoform (distributed spectra, dS) are shown in each box. Data for each bait protein represent the sum of two technical MudPIT experiments. (C) The histone acetyltransferase activity of FLAG-purified CHAT (via Ada2b-PA or Chiffon) or SAGA (Ada2b-PB) complexes containing equivalent amounts of Gcn5 as determined by western blotting with anti-Gcn5 antibody (left) were assayed using core histones as substrate. Incorporation of ^{3}H-acetyl CoA was assayed by fluorography (top right) and the migration of histone H3 was determined by Coomassie staining (bottom right). The negative control lane consists of histones and ^{3}H-acetyl CoA with no complex added. (D) Histone acetyltransferase activity of the indicated complexes was quantified by scintillation counting of ^{3}H-acetyl CoA incorporated into core histones or H3 tail peptides as in C. Mean±s.d. is shown for three independent histone acetyltransferase assays relative to no complex control. (E) Heat map showing the percentage of total spectra mapping to each region of full-length Chiffon (1695 aa) purifications using the indicated bait proteins as in B. The conserved Dbf4 N and C motifs in Chiffon are indicated by the gray shaded boxes. *Mass Spectrometry was performed by Selene K. Swanson, Laurence Florens, and Michael P. Washburn.

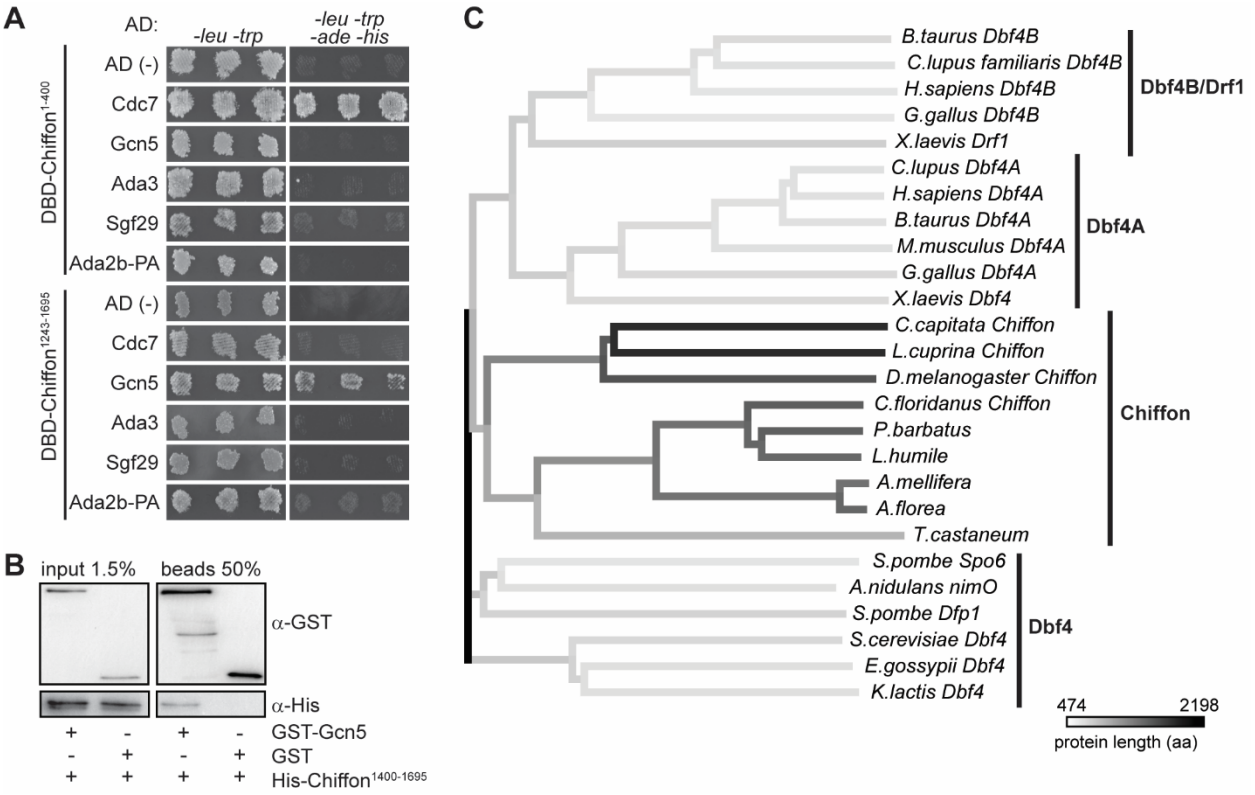

Figure 2.2 The insect-specific C-terminal domain of Chiffon directly binds Gcn5.

(A) Yeast two-hybrid assay was performed to test the pairwise interaction of each CHAT subunit with Chiffon. The Gal4 activating domain (AD) was fused to Cdc7, Gcn5, Ada3, Sgf29 or Ada2b-PA, and the Gal4 DNA-binding domain (DBD) was fused to either the N-terminal (1–400 aa) or C-terminal (1243–1695 aa) domains of Chiffon. Empty plasmids expressing only the AD or DBD were used to test for auto-activation of each protein. Three independent transformed yeast colonies were patched on media lacking leucine and tryptophan to test for the presence of the AD and DBD plasmids, and on media lacking leucine, tryptophan, adenine and histidine to test for interaction. (B) Glutathione-sepharose pull-down of recombinant GST–Gcn5 and the C-terminal domain of Chiffon (1400–1695 aa) tagged with His followed by western blotting with antibodies against GST and His. Representative data from three experiments are shown. (C) Phylogenetic tree constructed using neighbor-joining method showing Dbf4 homologs from fungi, insects and vertebrates based on Clustal-Omega multiple sequence alignment of full-length proteins. Shading represents protein length (aa, amino acids).

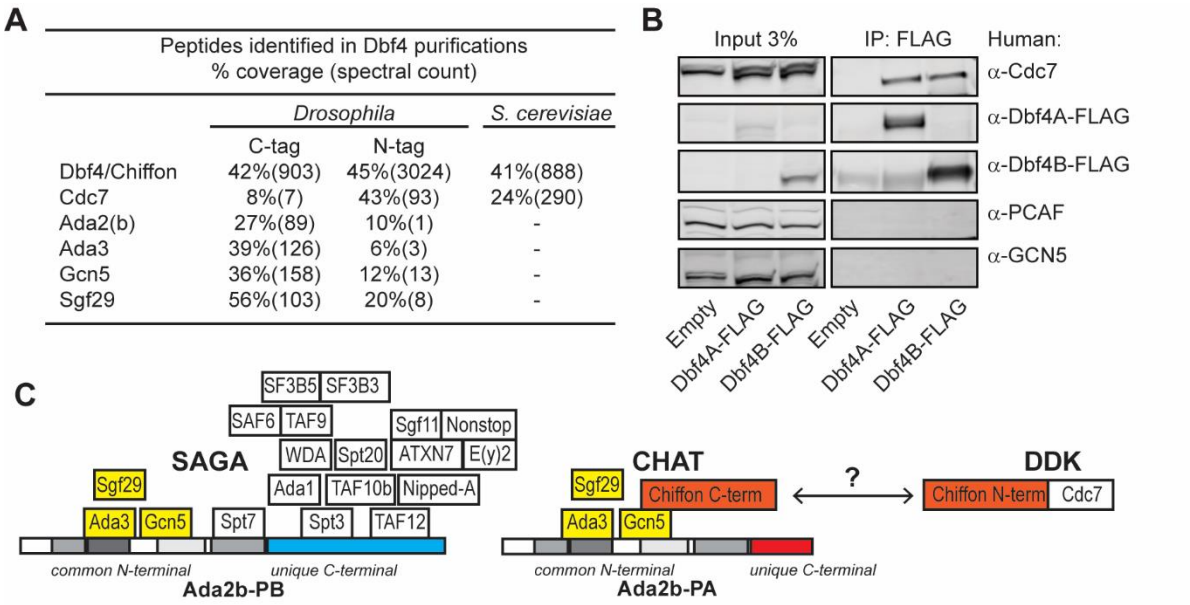

Figure 2.3 Dbf4 does not bind Gcn5 in yeast or humans

(A) Table showing proteins identified in Chiffon and Dbf4 purifications from *Drosophila melanogaster* (tandem FLAG-HA) or *Saccharomyces cerevisiae* (TAP-tagged). Sequence coverage (percentage) and number of spectra are shown for each protein. (B) FLAG-tagged human DBF4A or DBF4B were immunoprecipitated from HEK293T cell extracts using anti-FLAG antibodies, and analyzed by western blotting using the indicated antibodies. Control, empty vector. Representative data from three experiments are shown. (C) Schematic showing subunit composition of the SAGA, CHAT and DDK complexes. Interactions between subunits are based on the yeast two-hybrid analysis from Fig A1 and Fig A2, which suggest that Ada2b-PB binds Spt3 and TAF12 via its unique C-terminal domain to nucleate SAGA formation. In contrast, CHAT formation is nucleated by the binding of Chiffon's C-terminal to Gcn5, which precludes association of other SAGA subunits. Chiffon interacts with Cdc7 via its N-terminal domain to form the DDK complex, and the DDK and CHAT complexes appear to be largely separate *in vivo*. *Immunoprecipitation analysis in human cells was performed by Aktan Alpsoy and Emily Dykhuizen.

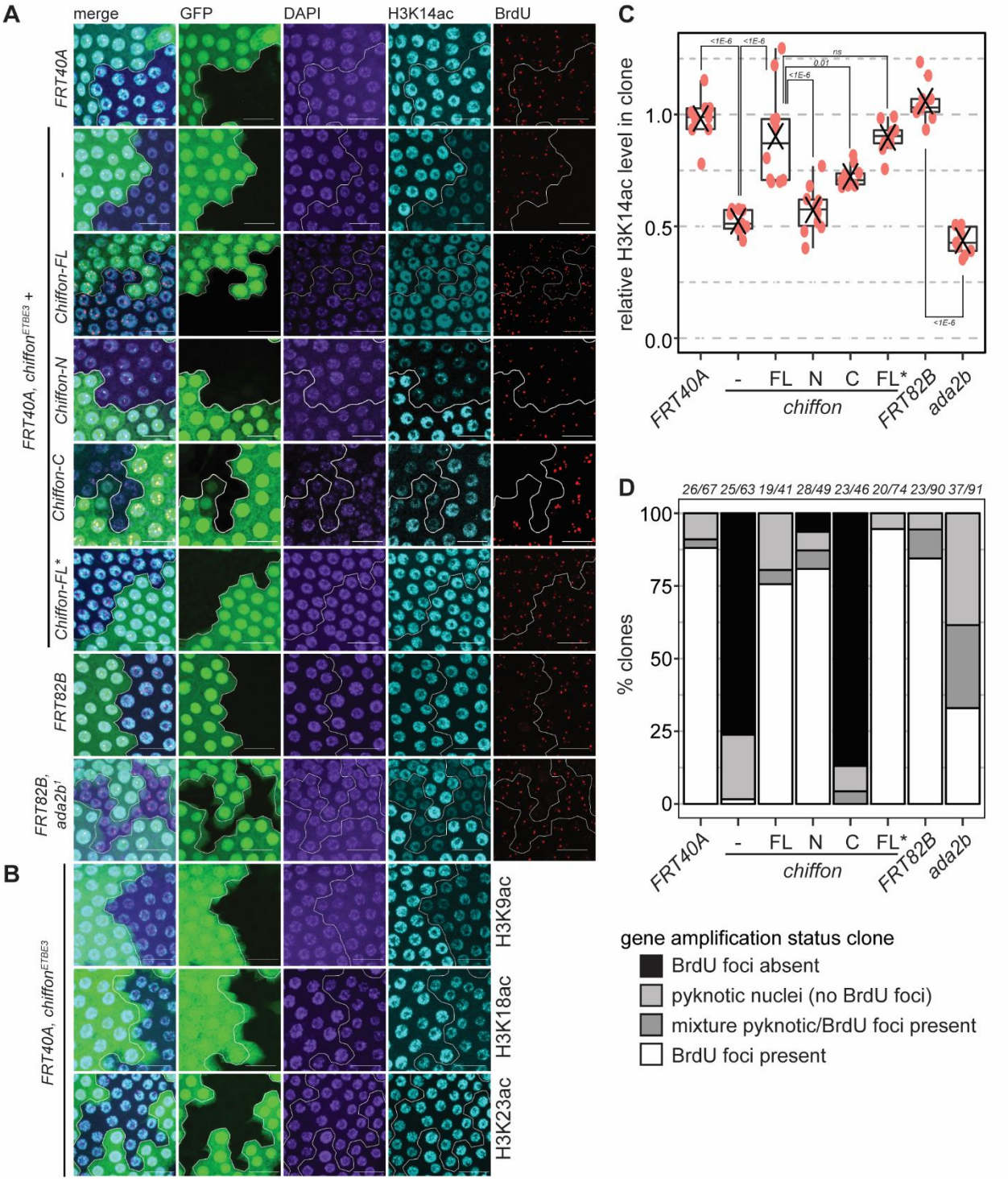

Figure 2.4 Chiffon is necessary for histone H3 acetylation *in vivo*

Figure 2.4 continued

(A) Mosaic egg chambers were generated using the FLP/FRT system for *chiffon*^{ETBE3} and *ada2b*^{1}, their respective controls, *FRT40A* and *FRT82B,* and for *chiffon*^{ETBE3} clones expressing single copies of the indicated Chiffon rescue transgenes. Maximum-intensity projection images showing BrdU incorporation, α-H3K14ac and DAPI staining from amplification-stage egg chamber follicle cells containing representative clones, marked by the absence of GFP and outlined in white. Scale bars: 20 μm. (B) Mosaic egg chambers for *chiffon*^{ETBE3} were examined for H3K9ac (*n*=5), H3K18ac (*n*=6) or H3K23ac (*n*=7) as in A. Representative images are shown for each histone modification. Scale bars: 20 μm. (C) Box plots showing relative H3K14ac levels in mutant clones versus GFP-positive control regions. 10–30 nuclei were quantified per region for 9–10 independent animals (red dots indicate clone analyzed from an individual animal; X, mean). *P*-values for the indicated comparisons were determined by ANOVA+Tukey-HSD; ns, not significant. (D) The percentage of clones undergoing gene amplification (BrdU-positive foci) in amplification-stage egg chambers from the indicated genotypes was determined. Several genotypes showed clones that were composed entirely or partially of pyknotic nuclei, which did not undergo gene amplification. The number of independent animals and clones examined for each genotype is shown above the plot (animals/clones).

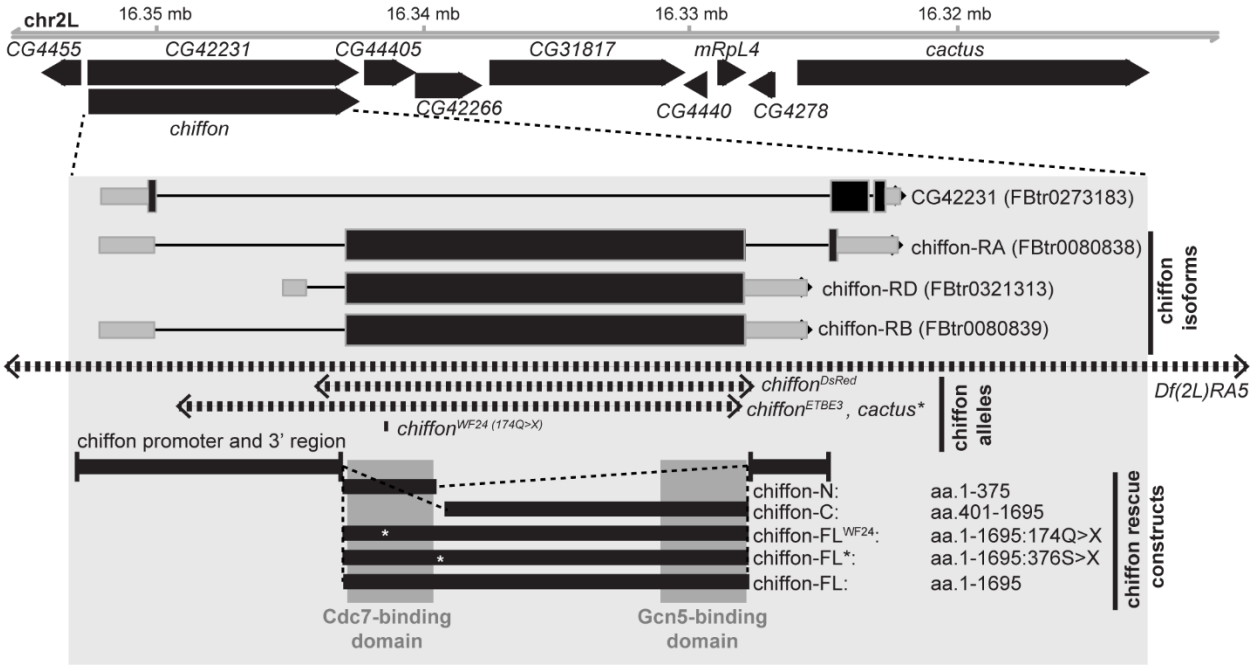

Figure 2.5 Map of *chiffon* gene structure

Schematic of *chiffon* locus showing nearby genes including *cactus*. The gene structure for *chiffon* is shown in the inset shaded box as coding sequences (black boxes), untranslated regions (gray boxes), and introns (lines). There are three annotated splice isoforms for *chiffon*: RA encodes a 1711-aa protein; RB and RD encode 1695-aa proteins from a single ~5 kb exon. An overlapping gene, *CG42231*, shares a promoter with *chiffon* but differs in its reading frame and encodes a separate polypeptide. The genomic regions deleted/mutated in each of the indicated *chiffon* alleles (*chiffon*^{DsRed}, *chiffon*^{ETBE3} and *chiffon*^{WF24}) are shown by the dashed line arrows. The *chiffon* rescue transgenes are shown by the black boxes at the base of the panel. Rescue constructs contain the indicated *chiffon* 5′ and 3′ regulatory regions (black boxes) and the *chiffon* coding sequences. The conserved Dbf4 N-terminal domain and C-terminal insect-specific Gcn5-binding domain are indicated by the shaded boxes overlaying the rescue constructs, and the position of each nonsense mutation in the rescue constructs is indicated by an asterisk.

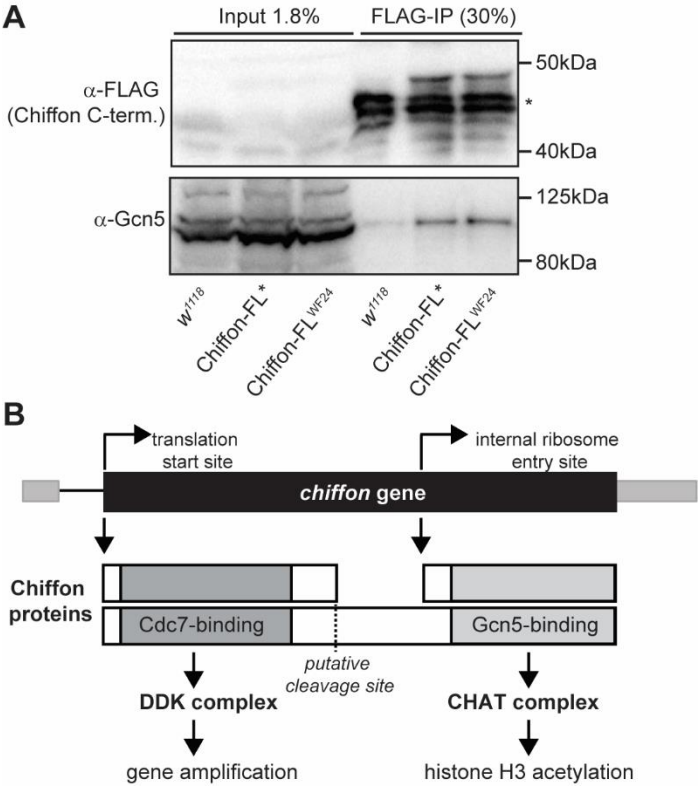

Figure 2.6 An internal translation start site in *chiffon* expresses a C-terminal product that binds Gcn5

(A) C-terminally FLAG-tagged Chiffon-FL* or Chiffon-FL^{WF24} transgenes, that contain premature stop codons at amino acids 376 or 174, respectively, were immunoprecipitated from embryo lysates using anti-FLAG antibodies. Co-immunoprecipitated proteins were analyzed by SDS-PAGE and western blotting with antibodies against FLAG (Chiffon) and Gcn5. The asterisk indicates nonspecific bands present in the w^{1118} control. Representative data from three experiments are shown. (B) Schematic illustrating the two polypeptides encoded by *chiffon*. The first start codon encodes full-length Chiffon (1695 aa) with the conserved Dbf4 Cdc7-binding domain in its N-terminal region. The N-terminal Chiffon product binds Cdc7, nucleates DDK formation and is necessary for gene amplification. An alternative internal ribosome entry site generates a C-terminal product containing the insect-specific Gcn5-binding domain that nucleates CHAT formation, and is essential for histone acetylation and development. Our data suggest that two mechanisms might control production of the alternative Chiffon products that nucleate DDK versus CHAT complex formation: (1) translational switching between cap-dependent and IRES-dependent start sites; and/or (2) proteolytic cleavage of full-length Chiffon. *Western blot was performed by Robert Stephenson.

CHAPTER 3. HAT ACTS UPSTREAM OF OTHER GCN5 COMPLEXES DURING EARLY *DROSOPHILA* EMBRYOGENESIS TRIGGERING GLOBAL HISTONE H3 ACETYLATON AND EXPRESSION OF DEVELOPMENTAL GENES

Declaration of collaborative work
The work described in this chapter was the collaborative effort of Eliana F. Torres-Zelada, Smitha George, and Hannah R. Blum, under the supervision of Dr. Vikki Weake. Smitha George performed both the embryo immunostaining and qRT-PCR analysis of Chiffon FL. All other experiments, RNA-seq, and ChIP-seq analysis were performed by Eliana F. Torres-Zelada under the supervision of Dr. Vikki Weake.

3.1 Introduction

Histone acetylation stimulates chromatin remodeling, thereby contributing to transcription activation. One of the best characterized histone acetyltransferases (HAT) is the highly conserved Gcn5, which functions as part of large multi-subunit transcriptional coactivator complexes to stimulate gene expression (Soffers and Workman 2020). *Drosophila* contains four Gcn5-containing complexes: SAGA (Kusch et al. 2003; Muratoglu et al. 2003), ATAC (Guelman et al. 2006a), CHAT (Torres-Zelada et al. 2019), and ADA (Soffers et al. 2019) . The formation of each of these complexes is determined by which Ada2 homolog incorporates into the complex: ATAC contains Ada2a, while SAGA and ADA contain the Ada2b-PB splice isoform, and CHAT contains the Ada2b-PA isoform. All Gcn5 complexes preferentially acetylate histone H3 with the highest activity on K9 and K14. However, ATAC also acetylates histone H4 due to the presence of a second HAT within the complex (Suganuma et al. 2008). Mutations in *ada2b* that disrupt SAGA, ADA, and CHAT alter gene expression (Li et al. 2017; Weake et al. 2008). However, there is remarkably little overlap between the genes regulated by *ada2b* and other SAGA subunits. Although this lack of overlap was previously attributed to SAGA's additional enzymatic activities, in light of findings that Ada2b splice isoforms nucleate formation of distinct Gcn5 complexes (Torres-Zelada et al. 2019), an alternative interpretation is that SAGA, ADA, and CHAT have distinct roles in gene expression. Previous data hint that CHAT is the major HAT required for development in flies because the CHAT-specific Ada2b-PA isoform can compensate for the loss of Ada2b-PB during development (Torres-Zelada et al. 2019). Here, we show that the majority of genes with disrupted expression in *ada2b* embryos are redundantly regulated by SAGA/ADA and

CHAT. Surprisingly, *chiffon* mutants that disrupt only CHAT cause different changes in gene expression compared with loss of *ada2b*, accompanied by a global loss of H3K14ac genome-wide in embryos. These data suggest that in addition to its HAT activity, CHAT has another role in gene expression mediated by its Chiffon subunit. We propose that CHAT functions as a pioneer coactivator complex during embryogenesis that is necessary for the later recruitment and/or activity of HAT complexes that activate gene expression programs essential for development.

3.2 Results and Discussion

3.2.1 SAGA/ADA and CHAT act redundantly to regulate gene expression in embryos.

Gene expression profiling of *ada2b* mutants has revealed widespread disruption of gene expression that was historically attributed to loss of SAGA activity (Torres-Zelada and Weake 2020). However, the recent finding that alternative splicing of *ada2b* can generate new diversity in Gcn5 complexes raises the question of whether these complexes have overlapping or distinct roles in regulating gene expression. To answer this question, we generated *ada2b* mutant embryos that express either the Ada2b-PA or Ada2b-PB isoforms, resulting in embryos that lack CHAT, or SAGA and ADA, respectively (Fig. 3.1A). SAGA and ADA cannot be distinguished genetically in flies through *ada2b* because Ada2b-PB is present in both complexes.

As wild-type (WT) controls, we used *ada2b* embryos that express both Ada2b-PA and Ada2b-PB, which fully restore viability (Torres-Zelada et al. 2019), as well as *elav-Gal4>GFP* embryos that do not carry the *ada2b* alleles (Fig. 3.1A). We observed a decrease in H3K14ac levels in the GFP-positive *ada2b* null embryos relative to their heterozygous (non-GFP) siblings by stage 13, indicating that loss of all Ada2b isoforms globally impacts histone H3 acetylation (Fig. 3.1B). Principal component analysis (PCA) revealed that the *ada2b* samples were distinct from the two WT controls (Fig. 3.1C). Interestingly, the *ada2b* embryos expressing either the Ada2b-PA isoform (ΔSAGA/ADA) or the Ada2b-PB isoform (ΔCHAT) grouped more closely to each other rather than to either the null or WT. These data suggested that SAGA/ADA and CHAT might have redundant or overlapping roles in gene expression during embryogenesis.

To identify genes that required SAGA/ADA or CHAT for expression, we first identified genes that were differentially expressed between *ada2b* and both WT controls. We identified 2111 differentially expressed genes (DEGs) between *ada2b* and WT. We conclude that these 2111

DEGs, corresponding to 22% of all expressed genes, represent the complete set of gene targets for the Ada2 subunits of SAGA, ADA, and CHAT in embryos. Next, we reasoned that there were three distinct possibilities for how these Gcn5 complexes could regulate gene expression: If SAGA/ADA and CHAT regulate expression of unique sets of genes, then some of the 2111 DEGs would be mis-regulated in the same direction and to the same extent (FC) in *ada2b* embryos expressing either the Ada2b-PA or Ada2b-PB isoforms (Fig. 3.1D). Alternatively, if SAGA/ADA and CHAT act cooperatively to regulate gene expression, then those DEGs would be mis-regulated in the same direction in both *ada2b* embryos expressing Ada2b-PA (*ada2b* + Ada2b-PA) and Ada2b-PB (*ada2b*+Ada2b-PB), but to a lesser extent than the null. Last, if SAGA/ADA and CHAT act redundantly, then the identified DEGs would not be mis-expressed in *ada2b* embryos expressing either splice isoform because SAGA/ADA and CHAT would compensate for the loss of each other at these genes.

When we clustered the 2111 DEGs identified in *ada2b* relative to WT, we found very few genes that were uniquely regulated by SAGA/ADA (*ada2b* + Ada2b-PA, 54 genes) while nearly a quarter were uniquely regulated by CHAT (*ada2b* + Ada2b-PB, 514 genes). Further, only 184 genes were cooperatively regulated by SAGA/ADA and CHAT, showing changes in expression upon loss of either isoform. In contrast, the majority of the DEGs (1359 genes) appear to be regulated redundantly by SAGA/ADA and CHAT, showing restored expression in the presence of either splice isoform (Fig. 3.1E). Examination of the normalized expression for these genes revealed that even at these redundantly regulated genes, loss of CHAT still had a stronger effect on gene expression relative to the loss of SAGA/ADA, particularly for downregulated genes. These data suggest that most genes that require SAGA, ADA, or CHAT for proper expression are regulated redundantly by these complexes, with a slightly stronger role for CHAT relative to the other complexes in embryos.

Gene Ontology (GO) analysis for these 1359 SAGA/ADA/CHAT-regulated genes showed enrichment for biological processes involved in development such as regionalization, negative regulation of cell communication, and positive regulation of transcription (Fig. 3.1F). Because Ada2b-PA is essential for fly viability but Ada2b-PB is not (Torres-Zelada et al. 2019), it is likely that either the 514 CHAT-regulated genes or the 184 SAGA/ADA/CHAT cooperatively-regulated genes represent those critical genes that when mis-expressed at this embryonic stage, cause lethality later in development. For this group of genes, the most enriched GO terms included gene

silencing, cellular macromolecule catabolic process, and posttranslational protein folding processes (Fig. 3.1G). Although Ada2b-PA restores adult viability to *ada2b* mutants, only 63% of the expected adults emerge, suggesting that SAGA, ADA, and CHAT act together to regulate the expression of genes that are essential for proper development. Overall, we conclude that SAGA/ADA and CHAT act redundantly at most genes in embryos, with a small proportion of genes being uniquely regulated by CHAT.

3.2.2 CHAT is necessary for global H3K14ac in embryos.

Since SAGA/ADA and CHAT shared overlapping roles in gene expression in embryos, we next asked how loss of CHAT affected histone acetylation in embryos. We previously showed that CHAT acetylates histone H3 with specificity for lysines 9, 14, and 18, with *chiffon* null ovary follicle cells showing a 50% decrease in H3K14ac levels (Torres-Zelada et al. 2019)We used a similar genetic approach to that used for *ada2b* to positively label embryos that contained two different *chiffon* null alleles with GFP (Fig. 3.2A). Using an antibody raised against the unique C-terminal region of Chiffon (1400 – 1695aa), we showed that *chiffon* embryos have a substantial decrease in Chiffon protein levels by stage 9 (Fig. 3.2B). Surprisingly, when we examined H3K14ac levels in *chiffon* embryos, we observed a stronger decrease in H3K14ac than that observed in *ada2b* mutants that disrupt SAGA, ADA, and CHAT (compare Fig.3.2B, 3.1B). When we quantified these data, we found that *chiffon* embryos showed 40% of the H3K14ac signal relative to their heterozygote siblings as compared with 70% in *ada2b* embryos (Fig. B.1). In contrast, *chiffon* and *ada2b* mutant ovary follicle cells both showed ~50% decreases in H3K14ac relative to their respective (Torres-Zelada et al. 2019). Because Ada2 subunits are essential for nucleosomal HAT activity (Grant et al. 1997), these data suggest that loss of the Chiffon subunit in CHAT has a stronger impact on histone acetylation in embryos that can be explained simply by loss of SAGA, ADA, and CHAT HAT activity.

To investigate how loss of *chiffon* affects H3K14ac genome-wide, we next performed H3K14ac ChIP-seq in *chiffon* stage 12 - 14 embryos. As a WT control, we performed ChIP-seq in *chiffon* embryos expressing the full-length Chiffon rescue transgene (Chiffon FL), which restores viability and H3K14ac levels. We examined H3K14ac levels relative to histone H3 to control for differences in nucleosomal occupancy, and sequenced input chromatin controls for each sample. Because we suspected that loss of *chiffon* would result in a global decrease in H3K14ac levels

based on our observations from embryo immunostaining results, we included spike-in *Saccharomyces cerevisiae* chromatin enabling us to normalize H3K14ac signal to this internal control (Fig. 3.2C). We then compared spike-in-normalized H3K14ac levels relative to histone H3 around the transcription start site (TSS) of all genes in *chiffon* embryos relative to WT (Fig. 3.2D). We observed a striking decrease in H3K14ac in *chiffon* embryos that was also readily observable in the individual biological replicates (Fig. B.2). Although H3K14ac is predominantly associated with promoters, a recent study has demonstrated peaks of H3K14ac in gene bodies, most likely due to the activity of the HAT Chameau rather than Gcn5 (Regadas et al. 2021). Strikingly, many of the H3K14ac peaks that are present at genes that lack canonical histone acetylation (and are thought to be deposited by Chameau) are also lost in *chiffon* embryos (Fig. B.2D). Together, these data argue that rather than specifically reducing H3K14ac levels only at CHAT-regulated genes, loss of the *chiffon* subunit in CHAT leads to a decrease in global levels of H3K14ac in embryos by affecting the activity and/or recruitment of other HATs.

3.3 Chiffon regulates expression of genes expressed in mid/late embryogenesis.

Because *chiffon* embryos showed a more substantial decrease in H3K14ac relative to *ada2b*, suggesting that the Chiffon subunit within CHAT functions upstream of other HATs, we next asked how loss of *chiffon* impacted gene expression in embryos. To do this, we performed RNA-seq in *chiffon* embryos at stage 12 - 14 as outlined in Figure. 3.3A. As WT controls, we used either *chiffon* embryos that express a single copy of a Chiffon FL transgene, or the parental *act-Gal4>GFP* embryos (Fig. 3.3A). We identified 996 genes that were differentially expressed between *chiffon* and both WT controls. In flies, *chiffon* encodes two polypeptides that have independent functions. The Chiffon N-terminal product is orthologous to Dbf4, which is a cyclin-like protein that binds and activates Cdc7 forming the Db4-dependent kinase (DDK) complex that initiates DNA replication (Landis and Tower 1999); we refer to this polypeptide as Chiffon-A. In contrast, the C-terminal domain of Chiffon, which is only conserved within insects, directly binds Gcn5 and nucleates formation of the CHAT complex; we refer to this polypeptide as Chiffon-B. Thus, the 996 DEGs identified in *chiffon* embryos could represent targets for either DDK or CHAT activity. Because expression of the ΔN-terminal transgene that restores CHAT function rescues both H3K14ac and adult viability in *chiffon* mutants (Torres-Zelada et al. 2019), we hypothesized that CHAT is necessary for histone H3 acetylation and gene expression in embryos. However,

because DDK phosphorylates histone H3T45 in yeast and mammalian cells (Baker et al. 2010), it is possible that DDK also contributes to gene expression. To distinguish between these possibilities, we performed RNA-seq on *chiffon* mutants that express single copies of each of the following transgenes: ΔN-terminal (ΔN; 401 – 1695aa) and FL with a stop codon at position 174, corresponding to the previously identified Chiffon FL^{WF24} allele (Fig. 3.3A) (Landis and Tower 1999)The ΔN and FL^{WF24} transgenes restore viability to *chiffon* mutants because they express Chiffon-B and rescue CHAT function; however the ΔN and FL^{WF24} constructs do not express Chiffon-A and do not restore DDK activity, hence the resulting adult females are sterile due to lack of gene amplification in ovary follicle cells (Torres-Zelada et al. 2019).

PCA revealed that the *chiffon* samples were distinct from the two WT controls and the FL^{WF24} genotype, which grouped together (Fig. 3.3B). These data suggested that the FL^{WF24} embryos that lack DDK activity were most similar to WT embryos rather than *chiffon* mutants, indicating that the loss of CHAT activity is responsible for most of the differences in gene expression observed in the *chiffon* embryos. Moreover, examination of relative gene expression levels at the 996 DEGs in each genotype revealed that the ΔN and FL^{WF24} genotypes largely resembled the WT controls (Fig. 3.3C). These data indicate that all of the 996 DEGs identified in *chiffon* embryos are regulated by CHAT rather than DDK. Interestingly, the ΔN transgene rescue construct was separated along with the second PC from both the WT and FL^{WF24} samples (Fig. 3.3B), suggesting that although the ΔN and FL^{WF24} transgenes both restore CHAT function – their activity differs. The ΔN and FL^{WF24} transgenes produce the same ~48kDa Chiffon-B protein product that binds Gcn5 and nucleates CHAT formation (Torres-Zelada et al. 2019), but differ at the nucleic acid level because the ΔN transgene lacks the first 1200 bp of the *chiffon* coding region. In addition, whereas the FL^{WF24} transgene fully restores viability to *chiffon* mutants, only 66% of the expected adults emerged in *chiffon* mutants expressing the ΔN transgene, suggesting that these 1200 bp of *chiffon* might contribute to the proper expression of CHAT (Torres-Zelada et al. 2019). We observed significantly higher expression of the ΔN transgene relative to the other constructs (Fig. B.3A), suggesting that this 1200 bp region contains negative regulatory elements that control *chiffon* transcript expression. If so, the higher expression of the ΔN transgene relative to FL^{WF24} could result in a slight gain-of-function for CHAT activity in terms of gene expression. Supporting this, 26 genes were differentially expressed between the ΔN and FL^{WF24} genotypes showing an opposite direction compared to the *chiffon* null (Fig. 3.3C). Together, these data demonstrate that

the Chiffon subunit with CHAT is necessary for expression of 10% of expressed genes in *Drosophila* embryos, with no detectable contribution from the DDK complex to gene expression at this developmental stage.

GO term analysis revealed that the *chiffon* DEGs were enriched for biological processes including isoprenoid metabolic processes, regulation of neurotransmitters, and cell morphogenesis involved in differentiation (Fig. 3.3D). Moreover, the 427 downregulated genes (Fig. B.3B), which represent potential targets for transcription activation by CHAT, were enriched for GO terms including salivary gland morphogenesis, ecdysone biosynthetic process, and cell morphogenesis involved in gastrulation (Fig. B.3C). Based on the enrichment of terms involved in development, we wondered if CHAT regulates developmental genes that first initiate expression during late embryogenesis. To examine this, we compared CHAT regulated genes with published developmental gene clusters (Graveley et al. 2011). While only 57% of the CHAT-regulated genes (570/996) fall into any of these developmental expression clusters, of these genes, 40% (228/570) are strongly associated with early-to-mid embryogenesis while another 47% (271/570) are associated with late embryogenesis and larvae stage (Fig. B.3D). Overall, our studies demonstrated that Chiffon within the CHAT complex regulates the expression of genes induced during embryo development.

Surprisingly, the overlap between the DEGs identified in the *ada2b* and *chiffon* mutants is quite low: 213 genes (Fig. B.3E). GO categories for these 213 genes include posttranslational protein folding, positive regulation of cell cycle process, and polytene chromosome puffing (Fig. B.3F). Based on these observations, and the stronger decrease in H3K14ac in *chiffon* embryos relative to *ada2b*, we conclude that the Chiffon-B subunit within the CHAT complex regulates gene expression in part through recruiting Gcn5 to chromatin to acetylate histone H3. However, our data suggest that Chiffon-B has additional roles in gene expression that are distinct from Ada2b/Gcn5 within CHAT, potentially functioning as a transcription coactivator. Our data further suggest that Chiffon-B activity is necessary for the subsequent recruitment and/or activity of other HAT complexes that target histone H3.

3.4 A switch between expression of the DDK and CHAT Chiffon products during embryonic development triggers CHAT formation prior to cellularization.

Because *chiffon* encodes two independent polypeptides that nucleate DDK or CHAT complex formation, and because loss of DDK activity had little effect on gene expression in late stage embryos and was dispensable for viability in flies (Torres-Zelada et al. 2019), we wondered if the Chiffon-A product that binds Cdc7 was even expressed in embryos. In *Drosophila* embryos, the first 13 cell cycles are maternally programmed and occur synchronously with extremely short cycles that exhibit no gap phases. Cdc7 is essential for these early embryonic cell cycles, and its protein signal declines by nuclear cycle (NC) 14 when the mid-blastula transition (MBT) initiates (Seller and O'Farrell 2018). However, it was unclear whether Cdc7 requires Chiffon-A for its activity during early embryogenesis because Chiffon-A is entirely dispensable for adult viability, whereas Cdc7 is an essential gene (Stephenson et al. 2015). To assess the expression of Chiffon-A and -B, we performed immunostaining in embryos expressing Chiffon that was epitope-tagged with HA on its N-terminus, and FLAG on its C-terminus. We used anti-HA antibodies to detect Chiffon-A and an anti-Chiffon antibody raised against the C-terminal end of Chiffon to detect Chiffon-B within the same embryo (Fig. 3.4A). No background immunostaining signal was detected in untagged WT embryo (w^{1118}) immunostained for HA under the conditions used (Fig. B.4C). We observed HA signal corresponding to Chiffon-A expression from NC3 to NC14, in a pattern resembling the published expression pattern of Cdc7 (Seller and O'Farrell 2018) (Fig. 3.4B). However, consistent with the lack of gene expression defects in later stage embryos lacking DDK, we did not detect expression of Chiffon-A (HA) after NC11 with no detectable expression in later stage embryos (Fig. B.4A). In contrast to the early embryonic expression of Chiffon-A, we did not detect expression of the Chiffon-B (anti-Chiffon) until NC10/11, with continued expression detected throughout the later stages of embryogenesis. Notably, both Chiffon-A and -B were detected together only at NC10/11 (Fig. 3.4B), suggesting that Chiffon FL might be present only transiently, if at all, at these nuclear cycles. Even if full-length Chiffon does exist transiently during NC10/11, our previous studies suggest that the full-length protein does not have an essential role (Torres-Zelada et al. 2019)

Next, we asked if Chiffon-A was necessary for these early embryonic nuclear cell cycles. We previously showed that the DDK activity of Chiffon was not necessary for adult viability, but was essential for follicle cell amplification (Torres-Zelada et al. 2019). Because *Cdc7* is required

for DNA replication during the early cell cycles (Seller and O'Farrell 2018), we hypothesized that the DDK activity of *chiffon* is required maternally for embryo development. To test this, we generated germline mosaic clones in which *chiffon* was maternally depleted (Fig. 3.4C). Loss of *chiffon* results in a complete failure of embryos to hatch, suggesting that *chiffon* is required maternally for embryo development but is not essential for oogenesis (Fig. 3.4C). We could rescue this hatch defect by expressing either the FL or N-terminal transgenes, suggesting that restoring DDK activity is sufficient for Chiffon's function in these early nuclear cell cycles. In contrast, the Chiffon ΔN transgene did not restore embryo hatching (Fig. 3.4C), suggesting that maternal CHAT is not necessary for these early stages of embryonic development.

Our data suggest that there is a switch at NC10/11 between the expression of the Chiffon-A product that nucleates DDK formation and the Chiffon-B product that nucleates CHAT formation. We wondered, therefore, whether expression of this Chiffon-B product coincided with the recruitment of Ada2b-PA within CHAT to the nucleus. To test this, we used anti-FLAG antibodies to immunostain embryos expressing epitope-tagged Ada2b-PA. Similar to the nuclear localization of Chiffon-B starting at NC10/11, Ada2b-PA showed nuclear-localized staining beginning at NC10/11 (Fig. 3.4D). Intriguingly, Ada2b-PB also showed a similar pattern of immunostaining beginning at NC10/11, albeit with weaker signal intensity, suggesting that SAGA/ADA recruitment to the nucleus occurs during the same temporal window as CHAT recruitment. Altogether, our data show that Chiffon-B nucleates the formation of CHAT in the early nuclear cycles of *Drosophila* embryogenesis before cellularization. We propose that this early formation of CHAT triggers histone H3 acetylation, and is necessary for the subsequent recruitment and/or activity of other histone H3 HATs such as SAGA, ADA, or even Chameau. Notably, CHAT formation occurs just prior to the de novo large-scale recruitment of RNA polymerase II (NC13-14) that leads to activation of the zygotic genome (Chen et al. 2013), suggesting that the timing of CHAT formation could have a key role in activating this wave of early transcription.

What could be responsible for this switch in expression between Chiffon-A and Chiffon-B during early embryonic development? Chiffon-A and Chiffon-B are both encoded from a single, large ~5kb exon in the *chiffon* gene, and we did not observe any substantial differences in mRNA levels of the regions corresponding to these products in early or late stage embryos expressing a paternal copy of the FL Chiffon transgene (Fig. B.4B). These data are consistent with Northern blotting analysis showing the presence of a single 6.5 kb *chiffon* transcript in *Drosophila* embryos

(Landis and Tower 1999), and indicate that alternative splicing is unlikely to underly the switch between Chiffon-A and Chiffon-B during embryonic development. Translational control of maternally-deposited mRNAs plays a central role in early *Drosophila* development because the two waves of zygotic transcription do not begin until NC8 and NC14. Many RNA-binding proteins control translation of mRNAs during early *Drosophila* development; for example, the RNA-binding protein Bruno binds to specific Bruno response elements (BREs), inhibiting translation of these BRE-containing mRNAs (Chekulaeva et al. 2006; Hamm and Harrison 2018). Interestingly, there are 3 BREs in the Chiffon-A region of the *chiffon* mRNA that could negatively regulate the translation of Chiffon N region after NC11 (Fig. B.4D). Supporting the hypothesis that negative regulatory elements are present in this N-terminal region that spans ~1200 bp, the ΔN transgene that lacked this region showed partial gain-of-function effects with respect to CHAT gene expression activity. We propose that the unique dicistronic gene structure of *chiffon* allows it to act as a developmental switch to trigger the timing of zygotic genome activation at the same time that embryonic nuclear cycles start to slow, due in part to decreased DDK activity. Because the C-terminal extension of Chiffon is only conserved within insect orthologs of Dbf4, it is likely that other mechanisms play a role in this transition in vertebrates and other animals.

3.5 Material and Methods

3.5.1 Genetics

Flies were raised in 12:12 h light:dark cycle at 25ºC on standard fly food (Lewis 1960). Genotypes for flies used in this study are described in Table B1. For RNA-seq experiments, flies were generated carrying two different *chiffon* (*chiffon^{ETBE3} and chiffon^{DsRed}*) (Torres-Zelada et al. 2019) or *ada2b (ada2b^{1} and ada2b^{842}*) (Pankotai et al. 2013, 2005) null alleles on chromosome 2 (*chiffon*) or chromosome 3 (*ada2b*), respectively, as either *actin-Gal4* (or *elav-Gal4* for *ada2b*) or *UAS-10XGFP*. To identify homozygous *chiffon* or *ada2b* mutants, we crossed flies as outlined in Figures 3.1 and 3.3 and manually selected GFP-positive embryos that carry the two different null alleles. Rescue transgenes for Chiffon or Ada2b isoforms were expressed in *trans* under control of their respective genomic regulatory sequences, as previously described (Weake et al. 2011; Torres-Zelada et al. 2019). We note that the Chiffon C-terminal transgene referred to as Chiffon-C (Torres-Zelada et al. 2019) is referred to as ΔN in this study. For ChIP-seq experiments,

homozygous *chiffon* mutants or wild type control were selected as outlined in Figure 3.3. We generated an epitope-tagged full-length *chiffon* transgene that was HA-tagged on its N-terminal domain, and FLAG-tagged on its C-terminal domain, and expressed this as the sole copy of Chiffon *in trans* in flies carrying two *chiffon* null alleles. Expression of the FL-Chiffon transgene fully restored viability and fertility to flies carrying the two *chiffon* null alleles.

3.5.2 RNA-seq

Total RNA from 6 embryos per replicate was extracted using the Direct-zol RNA microprep kit (Zymo Research #R2060). Libraries were generated from 25 ng RNA using the Ovation RNA-seq system (NuGEN) with unique dual indices for multiplexing and *Drosophila*-specific ribo-depletion. DEGs (False Discovery Rate, FDR < 0.05, FC ≥ 0.5) were identified using EdgeR (v3.30.3). GO term analysis was performed with clusterProfiler (v 3.18.1) and TopGO (v2.44.0).

3.5.3 RNA-seq analysis

Reads were trimmed using Trimmomatic (v0.38). Quality trimmed reads were mapped to the *D. melanogaster* genome (BDGP6.99) using HISAT2 (v2.0). Counts were identified for each gene using Htseq-count (v0.11.1) with default parameters. Counts were normalized by replicate using RUV normalization (R package *RUVseq*, v1.26.0). Differentially expressed genes (DEGs) were identified using EdgeR (v3.30.3) filtering low count samples, removing rRNA genes because RNA-seq libraries were ribo-depleted, and removing the following features: “no feature”, “ambiguous”, “too low aQual”, “not aligned”, and “alignment not unique”.

3.5.4 Chromatin immunoprecipitation

S. cerevisiae chromatin was prepared as described previously (Cloutier et al. 2013). Chromatin was fixed and prepared as in (Zeitlinger et al. 2007). Briefly, after dechorionation with 50% bleach, *Drosophila* embryos were transferred to 3 mL glass vial with PBT (PBS with 0.1% Triton X-100). PBT was then replaced with 230 µl fixation solution (50 mM HEPES, 1 mM EDTA, 0.5 mM EGTA, 100 mM NaCl) plus %1.8 formaldehyde and 750 µl n-heptane. Embryos were shaken vigorously for 15 min at room temperature. Embryos were centrifuged for 1 min at 500 x

g at 4 °C and the supernatant was discarded. Fixation reaction was then quenched by addition of 1.5 mL PBT-glycine (PBT with 250 mM glycine), followed by vigorous shaking for 1 min at room temperature, and collection by centrifugation as above. Finally, embryos were washed twice with 1 mL PBT and then there were resuspended in 1 mL of PBS with 0.5% Triton X-100. 300 GFP-positive embryos (per replicate) were then manually collected using a dissecting microscope with fluorescence (Nightsea SFA), and snap-frozen in liquid nitrogen. Then, chromatin pellet was washed three times with buffer A1 (15 mM HEPES, pH 7.5, 15 mM NaCl, 60 mM KCl, 4 mM $MgCl_2$, 0.5% Triton X-100) and once with buffer A2 (15 mM HEPES, pH 7.5, 140 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 1% Triton X-100, 0.1% sodium deoxycholate, 0.1% SDS, 0.5% N-laurosylsarcosine) and finally chromatin was sonicated in 130 µL of buffer A2 in Covaris E220 with the following conditions: 14min, 2% duty cycle, 105 Watts and 200 c.p.b to obtain an average fragment size of ~400bp. After centrifugation at 14,000 rpm for 10 min at 4°C, soluble chromatin was diluted with buffer A2 (0.1 %SDS) and used for ChIP. ChIP was performed as described (Jauregui-Lozano et al. 2021) with the following modification: 1 µg of *Drosophila* chromatin was mixed with 50 ng (5%) *S. cerevisiae* chromatin (BY4741) as a spike-in control to enable us to normalize our signal to this internal control (Chen et al. 2013). Spike-in factors are reported in Table B3. Chromatin was incubated with 1 µg of each of the following antibodies: anti-acetylated H3-Lys14 (rabbit; 07353, Millipore), anti-H3 (rabbit; ab1791, Abcam) at 4ºC overnight with rotation. Immunoprecipitated protein-DNA complexes were incubated with 25 µL Dynabeads protein G (ThermoFisher #10004D) for 4 hours at 4°C. Protein-DNA complexes were eluted from the magnetic beads with Elution buffer (1X TE, 1% SDS, 250 mM NaCl), treated with RNAse A (ThermoFisher #EN0531) at 37°C for 30 min and with Proteinase K (ThermoFisher #AM2546) at 55°C overnight. DNA purification, quantification, and input fragment size determination were performed as previously described (Jauregui-Lozano et al. 2021). ChIP-seq libraries were generated from 1 ng input and 0.5 ng ChIP DNA using the Ovation Ultralow library system (NuGEN) with unique dual indices for multiplexing.

3.5.5 ChIP-seq Analysis

Reads were trimmed using Trimmomatic (v0.38) to filter out low quality reads and remove adapter contamination. Quality trimmed reads were mapped to the *D. melanogaster* genome (BDGP6.99) and *S. cerevsiae* (S288C) genome using Bowtie2 (v2.3.5.1) using -sensitive settings.

For H3K14ac signal, spike-in factors were calculated as described (Orlando et al. 2014) and used to generate normalized bigwig files using deepTools (v3.1.1) bamCoverage subpackage, generating Reference adjusted Reads Per Million (RRPM). Metaplots and genomic distribution heatmaps were made with deepTools (v3.1.1) subpackages: computeMatrix, plotHeatmap, and plotProfile.

3.5.6 Immunostaining

Embryos were immunostained as in (Rothwell and Sullivan 2007) using : H3K14ac (1:200, rabbit, 07353, Millipore), HA (1:200, rat, 11867423001, Roche), FLAG (1:200, rabbit; F7425, Sigma), *Drosophila* Chiffon-C (rabbit, 1:200), Alexa Fluor 488- and Alexa Fluor 568- conjugated secondary antibody (1:400; goat; Thermo Scientific); 1 µg/mL 4′,6-diamidino-2-phenylindole, DAPI.

3.5.7 qRT-PCR

Quantitative real time PCR (qRT-PCR) analysis for mRNA levels of Chiffon-A or Chiffon-B during early developmental stage was performed on RNA isolated from single embryos collected 0 – 3h AEL using Direct-zol RNA Micro-prep kit (Zymo Research, Cat. #R2062). Relative expression for each gene was normalized to *Rpl32*. Primers are listed in Table B2.

3.5.8 Germline clones

hsFLP/Y; ovoD1, FRT40A/Cy) males were crossed to females of the indicated genotype *e.g. chiffon*^{ETBE3}*, FRT40A/CyO* that were homozygous for rescue transgene on chromosome 3 (see Table B1 for genotypes). Progeny were heat-shocked for 2 hours on two subsequent days 3 – 4 days AEL, and non-CyO females were selected and crossed with WT (w^{1118}) males to assess fertility and embryo hatch rates. Non-production of eggs, hatched and non-hatched embryos were counted for individual female progeny (n). Unhatched embryos were defined as failure to produce first instar larvae >26h AEL.

3.5.9 Antibody production

The Chiffon-C polyclonal antibody was generated against His-tagged 1400 – 1695 C-terminal recombinant Chiffon protein expressed in *E. coli* injected into rabbits. The rabbit serum was affinity purified against GST-tagged recombinant Chiffon C-terminal domain, and used for immunostaining as described.

3.5.10 Data availability

RNA-seq data and ChIP-seq are accessible through GEO repository under series accession number GSE179065.

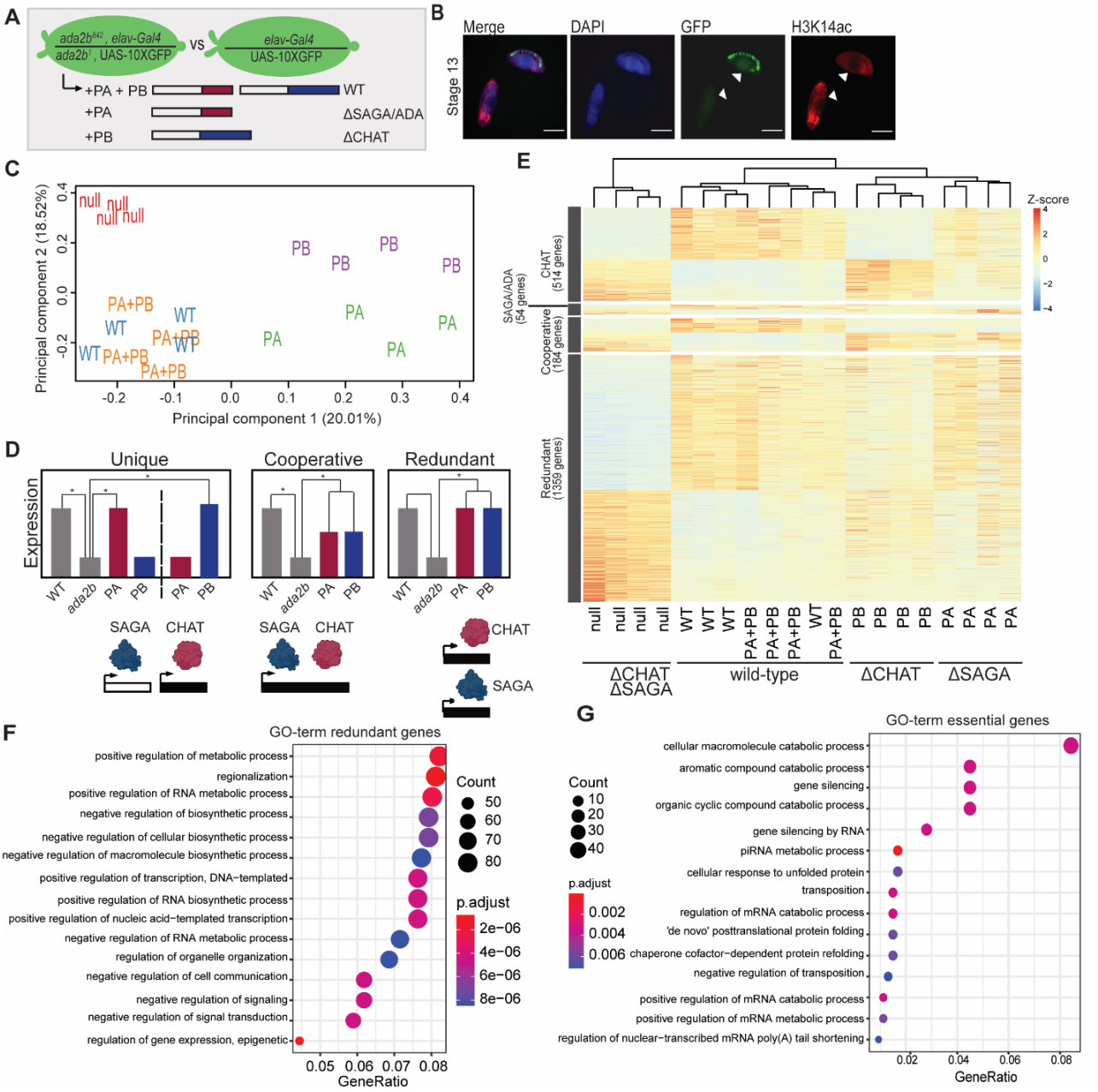

Figure 3.1 Ada2b splice isoforms act redundantly to regulate gene expression in embryos.

(A) Outline of the RNA-seq design. (B) Embryos were stained for DAPI and H3K14ac. Scale bars: 20 µm. (C) PCA of normalized counts for each sample. (D) Idealized bar plots demonstrating the criteria for Ada2b regulated genes. (E) Heatmap of RNA-seq expression z-scores computed for DEGs in *ada2b* versus PA+PB and WT. (F, G) GO terms for genes regulated redundantly by SAGA/ADA and CHAT (F) or requiring CHAT for unique and cooperative expression (G).

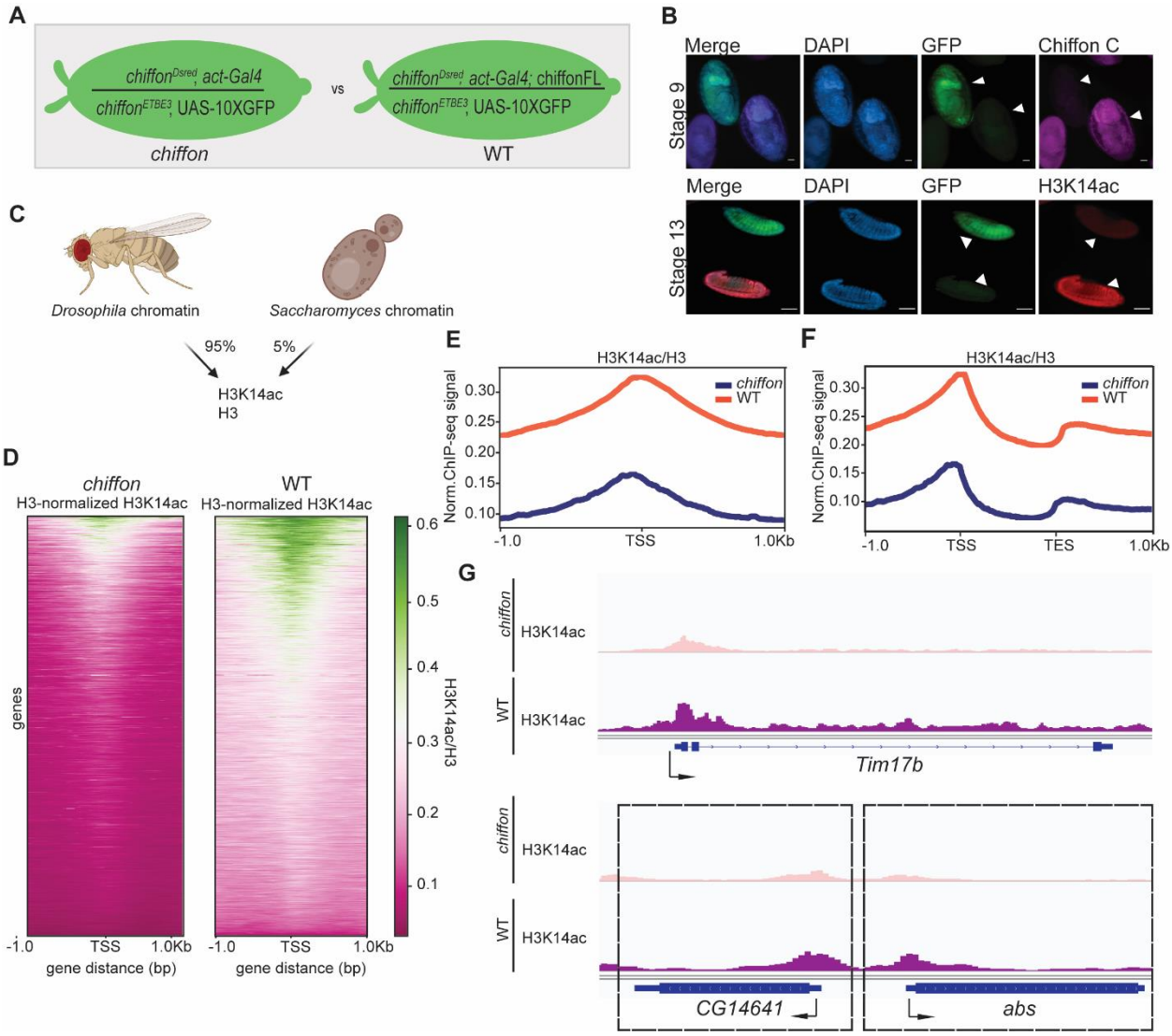

Figure 3.2 CHAT is necessary for global H3K14ac levels in embryos.

(A) Outline of the ChIP-seq design. (B) Embryos were stained for DAPI, H3K14ac, and the C-terminal region of Chiffon. Scale bars: 20 µm. (C) Schematic for spike-in normalization. (D) Heatmap showing RRPM-normalized H3K14ac ChIP-seq signal around the TSS of protein-coding genes in *chiffon* and WT embryos. (E) Metaplot of RRPM-normalized H3K14ac ChIP-seq signal around the TSS averaged for all protein-coding genes in *chiffon* and WT. (F) Metaplots of *chiffon* and WT RRPM-normalized H3K14ac ChIP-seq signal over gene bodies averaged for all protein-coding genes. (G) Genome browser snapshots showing RRPM-normalized H3K14ac signal at representative genes comparing *chiffon* and WT.

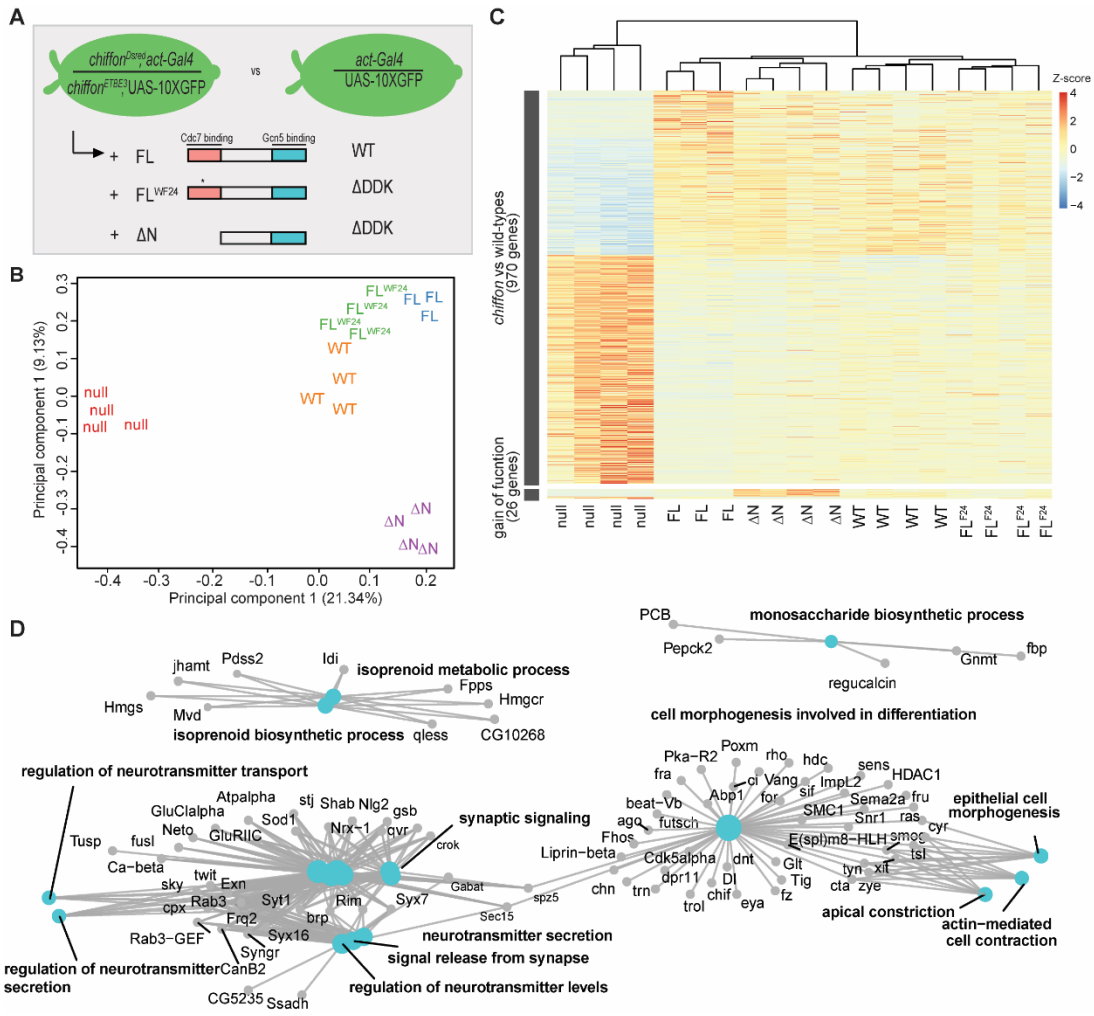

Figure 3.3 Chiffon regulates gene expression in embryos.

(A) Outline of the Chiffon RNA-seq design. (B) PCA of normalized counts for each sample. (C) Heatmap of RNA-seq expression z-scores computed for DEGs in *chiffon* versus FL and WT. (D) Gene Concept Network plot (Cnetplot) highlighting linkage of individual genes and associated functional categories of over-represented genes in *chiffon* embryos.

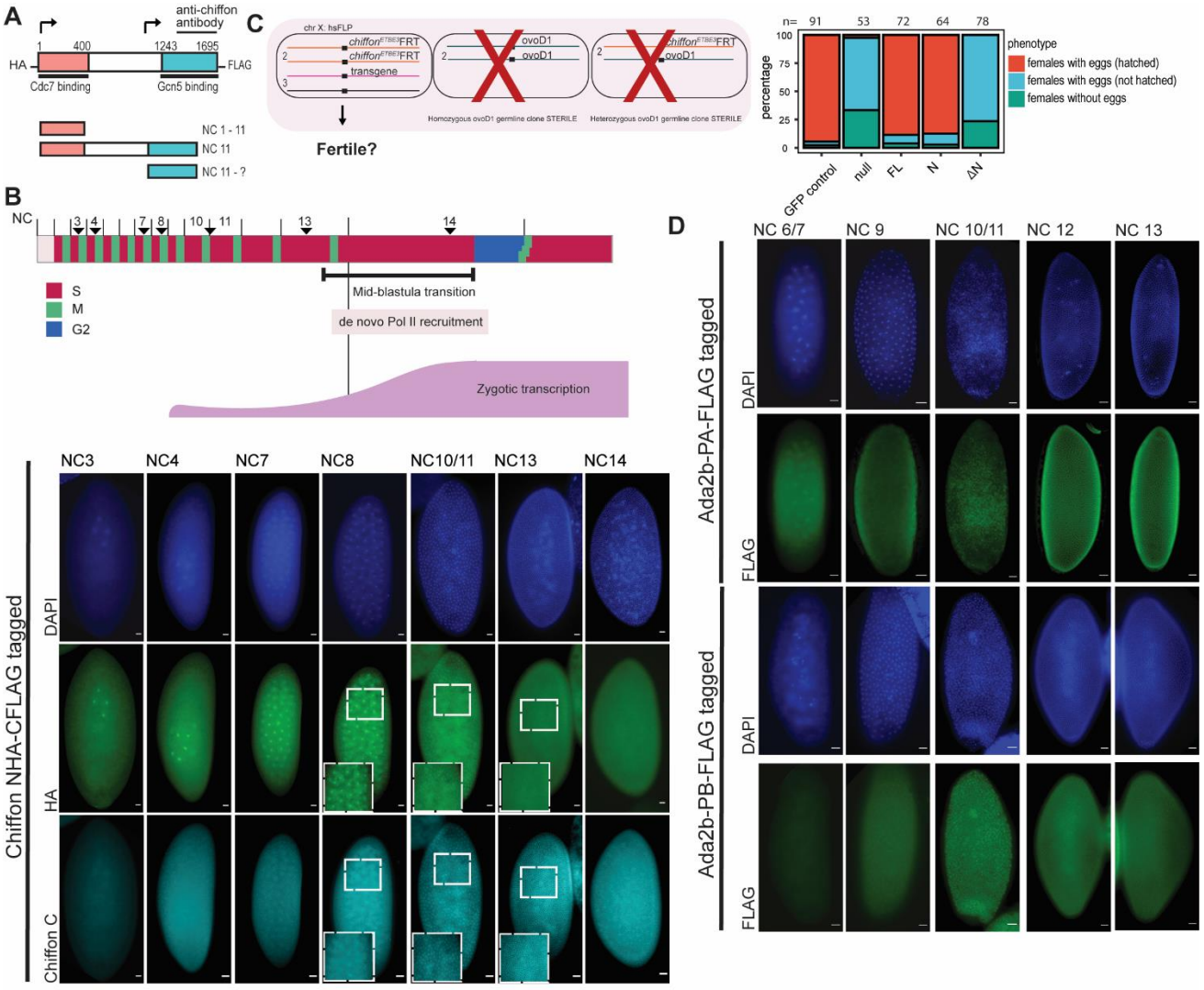

Figure 3.4 The DDK and CHAT complexes act sequentially during the early nuclear embryonic cycles.

(A) Schematic of the epitope-tagged FL Chiffon transgene. (B) Embryos were stained with HA, anti-Chiffon, and DAPI. Insets: enlarged views of highlighted areas. (C) Germline clones for *chiffon* were generated to assess if DDK or CHAT function was necessary for early embryonic development. Fertility was assessed in individual females (n shown above bars). (D) Embryos were stained with FLAG and DAPI to detect Ada2b-PA or Ada2b-PB. Scale bars: 20 µm (B, D). *Embryo immunostaining (B) was performed by Smitha George.

CHAPTER 4. ADDITIONAL OBSERVATIONS AND FUTURE DIRECTIONS

4.1 Introduction

DNA replication is restricted to occur only once per cell cycle by the strict temporal separation of origin licensing and firing. During origin licensing in G1, the pre-origin recognition complex (pre-ORC) assembles onto origins where it loads the Mcm2-7 DNA helicase. Origin firing and DNA replication initiation require phosphorylation of the Mcm2-7 helicase by the serine-threonine kinase Cdc7 (Tanaka and Araki 2010). Cdc7 is conserved from yeast to humans and is regulated by its association with Dbf4, forming the DDK complex (Dbf4-dependent kinase) (Oshiro Guy et al. 1999; Jackson et al. 1993). In *Saccharomyces cerevisiae*, Cdc7 transcript and protein levels remain constant throughout the cell cycle (Sclafani et al. 1988), yet its function is restricted at the G1/S transition when it is required to initiate DNA replication. Dbf4 protein expression correlates with Cdc7 kinase activity, peaking in S phase, and then declining at the beginning of G1 when Dbf4 is degraded by the Anaphase Promoting Complex (APC) (Oshiro Guy et al. 1999; Weinreich and Stillman 1999). Cdc7 and Dbf4 are essential in *S. cerevisiae* due to their requirement for the initiation of DNA replication during mitosis, and this requirement is due to their ability to phosphorylate the Mcm2-7 complex (Lei et al. 1997). DDK phosphorylates the inhibitor domain of Mcm4 and the absence of the inhibitor domain allows DDK-independent yeast growth (Sheu and Stillman 2010).

Drosophila has a single known Dbf4 homologue, Chiffon (Torres-Zelada et al. 2019) and two Cdc7 homologs (Stephenson et al. 2015). As in yeast, Chiffon binds Cdc7 and stimulates its kinase activity (Stephenson et al. 2015). However, Chiffon is not required for DNA replication in mitotic cells in *Drosophila*. First, *chiffon* mutant somatic clones can be obtained, indicating that Chiffon is not essential for cell proliferation (Stephenson et al. 2015). Second, Chiffon is not essential for endoreplication, although it is required for a specialized form of DNA replication known as gene amplification (Stephenson et al. 2015; Torres-Zelada et al. 2019). Third, the N-terminal domain of Chiffon, which is necessary and sufficient for Chiffon to bind Cdc7, is not required for viability in flies (Torres-Zelada et al. 2019). Together, these results indicate that either Cdc7 functions independently of Dbf4 in flies, or that there is an alternative Dbf4 subunit in *Drosophila* that has not yet been identified. Supporting the second possibility, *Drosophila* Cdc7

can only rescue viability of the yeast *cdc7* mutant when co-expressed with *Drosophila* Chiffon, indicating that Cdc7 alone is not sufficient for initiating DNA replication in yeast (Stephenson et al. 2015).

There has been a divergence and diversification of the Gcn5 complexes during evolution. The collection of Gcn5 complexes in flies and in other metazoan organisms appears to result from duplication of the Ada2 subunit. The finding that alternative splicing of *ada2b* can generate new diversity in HAT complexes (Torres-Zelada et al. 2019; Soffers et al. 2019) indicates that there may be other Gcn5-containing complexes in multicellular organisms that have not been yet discovered. However, there are some questions that remained to be elucidated. First, is CHAT restricted to insects, and if so, to a particular order of insects? Second, can Cdc7 be activated or targeted by another *Drosophila* protein in the absence of Chiffon? Third, is Chiffon necessary for the CHAT subunits (Gcn5, Ada2b-PA, Sgf29, and Ada3) to interact and for their activity on histones or nucleosomes? Last, does Chiffon have transcription coactivator activity and can it interact with the core transcriptional machinery such as TBP?. Here we discuss our preliminary observations addressing several of these research questions, and outline the progress we have made on these further studies to understand function of the CHAT complex and Chiffon in insects.

Some of these studies were conducted by undergraduate researchers in the Weake Laboratory under my supervision and guidance. Ben Anderson performed the dCdc7 screening and Hannah Blum performed the Y2H testing interaction of Gcn5- and Dbf4-orthologs in insects.

4.2 Chiffon is cell cycle regulated.

In most organisms, Dbf4 protein levels change with respect to the cell cycle and this variable expression determines the time that DDK activity occurs (Montagnoli et al. 2002; Cheng Liang et al. 1999; Oshiro Guy et al. 1999). To test if Chiffon protein and/or mRNA levels alter with respect to the cell cycle, we synchronized *Drosophila* Kc167 cells with hydroxyurea to arrest cells in G1, and then released these cells for 4 hours to obtain cells in S phase. We also treated Kc167 cells with DMSO to arrest cells in G2. DNA content was analyzed by Flow cytometry (FACS) (Fig. 4.1), and cell extracts were fractioned as described previously (Yoshizawa-Sugata et al. 2005) to examine chromatin-bound or cytoplasmic proteins. We used lamin as a loading control for chromatin-bound extract. We immunoblotted for Chiffon using an antibody raised to the C-terminal part of Chiffon (described in chapter 3). We did not observe a band corresponding to full-

length Chiffon (~190 kDa) in any samples, consistent with our previous observations that Chiffon protein is highly unstable and might not exist *in vivo* (see chapters 2 and 3 for details). Instead, we detected a strong band around 80 kDa, which was mostly present in the nuclear fraction, and was not detected using the pre-immune serum (Fig. C.1). This ~80 kDa Chiffon band showed increased abundance in the chromatin fraction in S phase relative to G1, G2, or asynchronous cells (Fig. 4.1, Fig. C.1). We also observed a strong increase in the levels of chromatin-bound histone H3K14ac in S phase. Since Gcn5 is bound to chromatin and it is maintained at similar levels throughout the cell cycle, but H3K14Ac levels only peak in S phase, these data suggest that Chiffon could be required to activate and/or recruit Gcn5 to acetylate histone H3 in S phase.

Previous studies show that the transcript levels of yeast Dbf4 are cell cycle regulated suggesting that cell cycle regulation of Dbf4 protein levels might be due to transcriptional regulation (Chapman and Johnston 1989). However, more recent genome-wide analysis and Northern blot analysis show that Dbf4 transcripts do not fluctuate during the cell cycle (Oshiro Guy et al. 1999; Weinreich and Stillman 1999). In contrast, in *Xenopus* and humans, Dbf4 related proteins are cell cycle regulated at both the protein and mRNA level (Kumagai et al. 1999; Yoshizawa-Sugata et al. 2005). Thus, we wondered whether Chiffon transcript levels were regulated through the cell cycle. To test this, RNA samples were collected at different cell cycles phases (G1, S, and G2) and checked for transcript levels using a set of primers for *chiffon* that correspond to the C-terminal region. Our data show that like Cdc7, Chiffon transcript levels do not change during the cell cycle at the steady-state mRNA level (Fig. 4.2). Since Chiffon transcript levels are constant during cell cycle, changes in transcription cannot account for the changes in Chiffon protein levels that we observed in Figure 4.1. Thus, we next tested the possibility that the fluctuation in Chiffon protein levels was due to Chiffon protein degradation by the APC complex. Indeed, our western blots using antibodies against Chiffon showed a degradation pattern that is common for Dbf4-orthologs (Fig. 4.1). To examine this possibility, we treated Kc167 cells with Nocodazole, which has been shown to inactivate the APC complex. Our results show no difference between cells treated with Nocodazole or the controls (Asynchronized, DMSO), indicating that Chiffon is unlikely to be a substrate of the APC complex in *Drosophila* (Fig. C.2). We conclude that more work needs to be done to elucidate the mechanism on how Chiffon is cell cycle regulated at the protein level, and if this is relevant to its biological function *in vivo*.

4.3 Identifying an alternative mechanism of Cdc7 activation in *Drosophila.*

In *Drosophila*, Cdc7 is an essential gene that is required for endocycling, however, the only Cdc7 regulatory subunit ortholog known so far, Chiffon, is not required for this specialized form of DNA replication (Stephenson et al. 2015). The only other Cdc7 ortholog in flies is CG5790 (FBgn0035677), which is expressed exclusively in the testis (Stephenson et al. 2015), so it would not be able to initiate DNA replication in other tissues. Therefore, we reasoned that there are two possible explanations; either Cdc7 can function independently of its Dbf4 regulatory subunit in flies, or flies possess an alternative protein that activates Cdc7. Because dCdc7 must be co-expressed with Chiffon to rescue the growth defect of yCdc7 (Stephenson et al. 2015), we concluded that dCdc7 cannot stimulate DNA replication initiation by its own, at least in yeast. Here, we sought to determine whether dCdc7 has an alternative mechanism of activation. Using mass spectrometry, we have identified 104 candidate Cdc7-interacting proteins that did not co-purify with Chiffon or negative controls (based on the rationale that Cdc7 activators should not co-purify with Chiffon). We prioritized candidates with the highest peptide abundance in Cdc7 purifications (Table 4.1). As an alternative approach, we screened for *Drosophila* proteins that interact with dCdc7 by a Y2H mating assay and identified two candidates. We then tested eight of the strongest candidates identified using these two approaches for interaction with dCdc7 by Y2H. We note that human DbfA was identified by Y2H (Kumagai et al. 1999), providing precedent for this strategy being appropriate to identify potential activators based on their ability to interact with Cdc7. Our Y2H screen identified only one protein: Maf1 (FBgn0267861, CG40196) that showed positive interaction with dCdc7 (Fig 4.3).

Maf1 is the ortholog of human MAF1, which is a negative regulator of RNA polymerase III. In *Drosophila*, Maf1 is inhibited by the conserved regulator of cell and tissue growth, the small G-protein R (Ras), promoting tRNA synthesis (Sriskanthadevan-Pirahas et al. 2018). Our data hint that insects could possess an alternative mechanism to activate dCdc7 or recruit it to chromatin to initiate DNA replication in mitotic cells that is different from that used by mammals, however other studies are needed to confirm whether Maf1 is able to activate dCdc7. To further investigate whether Maf1 is the regulatory subunit of dCdc7, we would test if Maf1 restores dCdc7-dependent growth to *cdc7* mutants yeast, as described in (Stephenson et al. 2015). To do this, we will express epitope-tagged (HA and FLAG, respectively) dCdc7 and Maf1 from galactose-inducible expression vectors. We will co-transform vectors into *S. cerevisiae* containing a temperature-

sensitive mutant of yCdc7 (Axelrod and Rine 1991). Yeast will be grown in selective liquid medium with 2% glucose (to repress protein expression), spotted as serial dilutions on medium containing 2% Galactose (to activate protein expression) and incubated at permissive (23°C) or restrictive (30°C) temperatures to assess growth, as in (Stephenson et al. 2015). Further, we will immunoprecipitate HA-tagged dCdc7 from yeast extracts and assess co-purification of FLAG-tagged Maf1 by immunoblot with anti-HA and FLAG antibodies. If Maf1 is the regulatory subunit of dCdc7, it would restore growth of yeast on galactose at restrictive temperature, plus it would co-immunoprecipitate with dCdc7. Our data suggest that since dMaf1 might bind Pol III, it could possibly recruit dCdc7 to promoters, in order to activate DNA replication. Importantly, unlike yeast, where replication origins are sequence-defined, origins in metazoan are not defined at the levels of DNA sequence (Ekundayo and Bleichert 2019). In contrast, in *Drosophila*, origins of replication are often associated with transcriptionally active regions near promoters (MacAlpine et al. 2004; Schübeler et al. 2002)

4.4 The Gcn5 HAT core complex can form without Chiffon.

As discussed in Chapter 1, the Gcn5 HAT core complex is formed by Gcn5, Ada3, Ada2 and Sgf29. Each of Gcn5-containing complexes present in metazoans is nucleated by the unique Ada2 homologs or splice isoforms that share conserved N-terminal domains, and differ only in their C-terminal domains. The Workman group has recently identified the ADA complex in flies, which is formed by the Gcn5 core subunits including the long splice isoform of Ada2b, Ada2b-PB (Soffers et al. 2019). The ADA complex in flies does not contain any other subunit as the Ada complex present in yeast (Soffers et al. 2019). However, whether the Gcn5 HAT core complex carrying the short splice isoform Ada2b-PA, is formed without Chiffon is not known. We wondered whether the short Gcn5 HAT core complex is able to be formed in the absence of Chiffon. To assess this, Gcn5, Ada3, Sgf29, and tagged-HisFLAG Ada2b-PA were co-expressed in insect cells as pairwise combinations of Ada2b-PA and any of the other subunit. Then, FLAG immunoprecipitation was performed in order to pull down Ada2b-PA, followed by a western blot using antibodies against Gcn5, Ada3, or Sgf29 as correspond per pairwise combination. This approach showed that Ada2b-PA interacts with each core subunits when are co-expressed together, except for Sgf29 (Fig. 4.4, lanes 1, 2, 3, 4). Furthermore, Ada2b-PA interacts with Sfg29 only in the presence of Ada3 (Fig. 4.4, lane 6). Interestingly, when Ada2b-PA is co-expressed with the

rest of the 3 subunits (Ada3, Sgf29, and Gcn5) we were able to detect, after immunoprecipitation, all 4 subunits that form the Gcn5 HAT core suggesting that as the Gcn5 core containing the Ada2b-PB, the one that contains the Ada2b-PA can also be formed, even in the absence of Chiffon. (Fig. 4.4, lane 7). We conclude that in the absence of Ada2b-PA (as in the *ada2b* mutant embryos), the Gcn5 core module of the CHAT complex is unlikely to form, whereas in the absence of Chiffon (as in *chiffon* embryos), Gcn5 can still assemble into the minimal module required for nucleosomal HAT activity. Thus, the differences in gene expression between *ada2b* and *chiffon* embryos may reflect loss of HAT activity or Chiffon function, respectively.

4.5 The CHAT complex is specific to the Diptera order

The finding that alternative splicing of *ada2b* can generate new diversity in HAT complexes (Torres-Zelada et al. 2019; Soffers et al. 2019) opens the possibility that there may be other Gcn5 complexes in multicellular organisms that remain to be discovered. Our data (see chapter 1 and 2) suggests that the CHAT complex might be present in other insects besides *Drosophila*. Expanding the studies on Gcn5 complexes into non-traditional species, including other insects may provide insight into specialized functions of this HAT in multicellular organisms. To assess whether the CHAT complex is also present in other insect, we tested for the interaction of the conserved C-terminal extension (1243 – 1695aa) that directly binds Gcn5 interaction in flies (chapter 1 and 2) and Gcn5-orthologs in Diptera, Hemiptera, and Coleoptera orders by a Y2H assay. Using this approach, we found that Dbf4 and Gcn5 interacted only in the Diptera order (Fig. 4.5). Interestingly, we also found that the *Tribolium* Chiffon C-terminal extension auto-activated when fused to the Gal4 DNA-binding domain, suggesting that *Tribolium* Chiffon might interact with yeast transcriptional coactivators like SAGA to activate expression of the reporter genes in this assay. This finding provides some support for the hypothesis that the C-terminal region of Chiffon possesses intrinsic transcription coactivator activity.

4.6 Future Directions

In Chapter 3, we propose that CHAT functions as a pioneer coactivator complex during embryogenesis, due to its specific subunit Chiffon, that is necessary for the later recruitment of others HAT complexes that activate gene expression programs which are essential for

development. Interestingly, Chiffon possesses a DNA binding domain (493 – 505 aa) which is present in the C-terminal extension of Chiffon (Chiffon-B). This is an interesting observation as it suggests a potential role of the CHAT complex as a coactivator in flies. We propose that potentially Chiffon binds DNA, thereby recruiting the Gcn5 core subunits (Ada3, Sgf29, and Ada2b-PA), and subsequently leading to the recruitment of other HAT complexes. However, more work is needed to understand how CHAT regulates gene expression in *Drosophila*.

In Chapter 2 we showed that Chiffon binds to Gcn5 by the very end part of the C-terminal domain (1243 – 1695aa), however, we have not tested whether this short region of Chiffon is sufficient for fly viability. Dr. Smitha George has made transgenic flies expressing 2xNHA-3xCFLAG-tagged Chiffon C in order to elucidate whether this region is required for histone acetylation and fly viability, similar to the Chiffon ΔN truncation. What is more, we wonder whether the expression of this short Chiffon C resembles the expression of Chiffon-B during embryo development, or whether the expression timing changes.

In Chapter 3 we showed that CHAT regulates the expression of genes required for fly development, however, we would like to identify genes that are directly regulated by the CHAT complex. Our effort to characterize the genome-wide distribution of the CHAT complex have been hampered probably because our Chiffon-tagged transgenic flies contain only a single FLAG tag in the C-terminal part, making the chromatin immunoprecipitation of the CHAT complex technically challenging. For this reason, Dr. Smitha George also generated transgenic flies expressing 2xNHA-3xCFLAG-tagged Chiffon FL in order to perform ChIP-seq experiments with FLAG antibodies to characterize the genome-wide distribution of the CHAT complex. An intriguing question for future studies is whether genes bound to CHAT represent transcribed genes marked by H3K14ac. To do this, we will perform ChIP-seq against Pol II using an antibody that recognizes all forms of the carboxy-terminal domain (4H8) and examine H3K14ac relative to histone H3 as a control. Using this approach, we will have a better insight on the mechanism of how CHAT regulates gene expression. Finally, our immunostaining experiments, using an antibody raised against the unique C-terminal domain of Chiffon, showed that the expression of Chiffon-B product starts at NC10/11 (see chapter 3 for details). However, we had difficulties getting good embryo images using FLAG antibodies, therefore, the new transgenic flies expressing 3xCFLAG-tagged Chiffon FL might be a powerful reagent that we can use to examine the mechanism separating expression of the Chiffon-A and Chiffon-B products.

4.7 Material and Methods

4.7.1 Cell cycle analysis

Kc167 Drosophila cells were maintained in M3+BYPE + 20% FBS + 0.5% Penn-strep at 25^{o}C. When cells were in log-phase the day of experiment, cells were set to 1e6 cells/ml the day of the cell cycle arrest experiment. To obtain cells mostly in G1-phase, cells were incubated with 1 mM of Hydroxyurea (HU) for 24 hr at 25^{o}C. To obtain cells in G2-phase, cells were incubated with 3% of DMSO for 24 hr at 25^{o}C. To obtain cells in S-phase, incubate cells with 1 mM of Hydroxyurea (HU) for 24 hr at 25^{o}C, and after 24 hr, cells were harvested by centrifugation (2000 rpm) for 5 min at RT. Cell were washed in 2 mL PBS, followed by adding of fresh media and incubated 4 hours at 25^{o}C. 4 mL of arrested cells (8e6 cells/ml) were then stained with Propidium iodide (PI) for FACS-DNA content analysis.

4.7.2 Cell Fractionation

Cells were fractioned into cytoplasm, nuclear-soluble, and nuclear pellet fractions (chromatin) as described in (Yoshizawa-Sugata et al. 2005) with minor modifications: Briefly, 20 mL of cells (4e6 cells/mL) were harvested at 3000 rpm at 4^{o}C for 5 minutes, pellet was washed in 1 mL of cold PBS. Then, cells were washed in hypotonic buffer (20 mM HEPES-KOH, pH 7.6, 5 mM potassium acetate, 0.5 mM $MgCl_2$, 0.5 mM dithiothreitol) buffer supplemented with complete protease inhibitor mixture (Roche Applied Science) and centrifuged at 3000 rpm at 4^{o}C for 5 minutes. Pellet was homogenized in 200 ul of hypotonic buffer followed by centrifugation at 5,000 rpm for 3 min to separate supernatant (cytoplasm) and pellet (nuclei). The pellet was resuspended in 150 µL CSK buffer (20 mM HEPES-KOH, pH 7.6, 40 mM potassium glutamate, 1 mM $MgCl_2$, 1 mM EGTA, 300 mM sucrose, 1 mM dithiothreitol, 1 mM Na_3VO_4, and protease inhibitors) containing 0.5 % Triton X-100 and incubated on ice for 10 min followed by centrifugation at 3,000 rpm for 3 min to separate the supernatant (nuclear-soluble fractions) and pellet (chromatin-enriched nuclear-insoluble fractions). The pellet was washed (100 µL) once with CSK-Triton X-100 buffer and resuspended in the same buffer (150 µL).

4.7.3 qRT-PCR

After cells were arrested, cells were pelleted and washed with PBS at 1000 rpm for 5 min. Then, cells (pellet) were lysed in 600 µL of TRIzol reagent. RNA was purified using the Zymo-spin IC columns (Zymo Research #C1004-50). Relative expression for each gene was normalized to the geometric mean of two reference genes: *Rpl32* and *eIF1-A*. Primers are listed in Table C2.

4.7.4 APC inactivation experiment

Kc167 cells were incubated with Nocodazole (5 mg/mL in DMSO, to a final concentration of 2 ng/mL) or 3% of DMSO for 8hr. Cells then were harvested and protein lysates were prepared in 200 µL of NP-40 buffer (50 mM Tris-CL, pH 8, 150 mM NaCl, 1% NP-40, and protease inhibitors).

4.7.5 Yeast Strains and Plasmids:

The Matchmaker Gold Yeast Two-Hybrid System supplied the yeast strains used for the mating experiments, Y2HGold and Y187 (Clontech Laboratories). Prior to the mating assays, the yeast strains were transformed with several plasmids generated from the vectors also supplied in the Yeast Two-Hybrid System. Y2HGold was transformed with one of the following: pGBKT7, the vector containing just the binding domain; pGBKT7-T, a company control plasmid; or pGBKT7-Cdc7, Y187 was transformed with one of the following: pGADT7, the plasmid containing just the activating domain; pGADT7-lam, the company negative control plasmid; pGADT7-53, the company positive control plasmid; or pGADT7-ChiffonN, a plasmid containing the first 1200bp of Chiffon.

4.7.6 Yeast Matting

The mating assay was performed following the specifications in the Matchmaker Gold Yeast 2-Hybrid Manual, with minor modifications (Clontech Laboratories). Birefly, Y2HGold containing pGBKT7-Cdc7 was mated with the normalized Universal *Drosophila* Mate & Plate Library from Clontech. The library was constructed from mRNA extracted from embryos, larva, and adult *Drosophila*. The library was then normalized to equalize the abundance of the cDNA. The bait and prey strains were incubated overnight prior to mating at 30° C and at 225 rpm for 16

hours until an OD_{600} of .8 was reached. The bait strain was then combined with the prey library and incubated at 30°C and 50 rpm for 21 hours prior to plating. The mating was then centrifuged twice at 1,000g for 10 minutes, before resuspension in 10 mL of 0.5X YPDA/Kan (50 µg/mL kanamycin). 200 µL aliquots were spread plated onto 55 150 mm DDO/X/A (SD -trp/-leu supplemented with X-Alpha-gal) (250 ng/mL AbaA) plates. To determine the efficiency of the screen, 100 µL aliquots of 1/10, 1/100, 1/1000, and 1/10000 dilutions were spread plated onto 100 mm SD/-Trp, SD/-Leu, and DDO plates. These plates were incubated at 30°C for five days, and at 4°C for one day prior to analysis.

4.7.7 Patching and Replica Plating for the Yeast Matting

Plates from the yeast mating were analyzed for blue or large colonies. 300 colonies were selected, patched onto DDO/X/A (250 ng/mL AbaA), and incubated at 30°C for three days. After three days the colonies that showed a blue color were selected and, onto QDO/X (SD/-trp/-leu/-ade/-his supplemented with X-Alpha-gal) and incubated at 30°C for three days.

4.7.8 Yeast two-hybrid assay

Yeast two-hybrid analysis was performed with the Matchmaker Gold Yeast two-hybrid system as per the manufacturer's instructions (Clontech Laboratories). Three independent transformed colonies were replica plated on the different selective media for each interaction tested. See Table C1 to see plasmids used.

4.7.9 Cloning and Purification of Recombinant Gcn5 core complex from Insect Cells

Coding sequences for Ada2b-PA, Gcn5, Ada3, and Sgf29 were cloned into pBACPAK8 vectors with Ada2b-PA N-terminal His-FLAG epitope-tagged. Each combination (Gcn5 + Ada2b-PA, Ada3 + Ada2b=PA, Sgf29 + Ada2b-PA, Ada2b-PA + Gcn5 + Ada3, Ada2b-PA + Ada3 + Sgf29, and Ada2b-PA + Gcn5 + Ada3 + Sgf29) were transfected into Sf21 cells, and viral supernatants were generated as described previously (Clontech Laboratories 2007). Sf21 cells were infected with FLAG-Ada2b-PA with any other protein, as described in each combination. Briefly, cells were lysed in Lysis Buffer (50 mM HEPES, pH 7.5, 2 mM MgCl2, 10% glycerol, 0.2% Triton X-100, 20 mM imidazole, 500mM NaCl, supplemented with complete protease

inhibitor mixture (Roche Applied Science)), and the soluble supernatant was incubated with Ni-NTA-agarose, washed in lysis buffer, and eluted in Ni-NTA Elution Buffer I (50 mM HEPES, pH 7.5, 300 mM NaCl, 10% glycerol, 250 mM imidazole). Ni-NTA elutions were pooled and bound to anti-FLAG affinity gel beads (Sigma). Beads were washed twice with High Salt Wash Buffer (50 mM HEPES, pH 7.5, 10% glycerol, 300 mM NaCl) and once with Low Salt Wash Buffer (50 mM HEPES, pH 7.5, 150 mM NaCl, 10% glycerol), and protein-protein complex were eluted in Elution Buffer II (50 mM HEPES, pH 7.5, 150 mM NaCl, 10% glycerol, supplemented with complete protease inhibitor mixture (Roche Applied Science)) containing 0.5 mg/ml 3XFLAG peptide.

4.7.10 Western blot

The following antibodies were used for western blot analysis: anti-Gcn5 (rabbit, 1:1000), anti-Ada3 (rabbit, 1:3000), anti-Sgf29 (rabbit. 1:500), anti-FLAG M2-peroxidase (HRP) (A8592, Sigma, 1:5000), anti-lamin (ADL40, DSHB, 1:1000), anti-H3K14ac (07-353, Millipore, 1:1000), anti-H3 (ab1791, Abcam, 1:1000), anti-Chiffon (rabbit, 1:1000), anti-preimm (rabbit, 1:1000).

4.7.11 Cloning of Chiffon C- and Chiffon FL- 3XFLAG-tagged

Chiffon C (1243 – 1695 aa) and Chiffon FL (1 – 1695aa) were cloned into pCa4B vectors with N-terminal 2xHA and C-terminal 3xFLAG epitope tags.

Table 4.1 Candidate dCdc7-interacting proteins.

	Symbol	Gene name	Peptide spectra in dCdc7 IP	Interaction with dCdc7 by Y2H
Mass spectrometry candidates	chif	Chiffon	178	✓
	CG30069	CG30069	85	×
	CG7791	CG7791	83	×
	CG42674	CG42674	41	×
	upset	upSET	29	×
	CG12792	lethal(2)09851		
Y2H screening candidates	CG40196	Maf1		✓
	CG11242	TBCB		×

104 proteins were identified in dCdc7 IPs. The number of peptide spectra is shown for the top 6 candidates with the most peptide spectra. 2 more candidates were identified from the Y2H screening. ✓ corresponds to a positive interaction and × corresponds to a negative interaction.

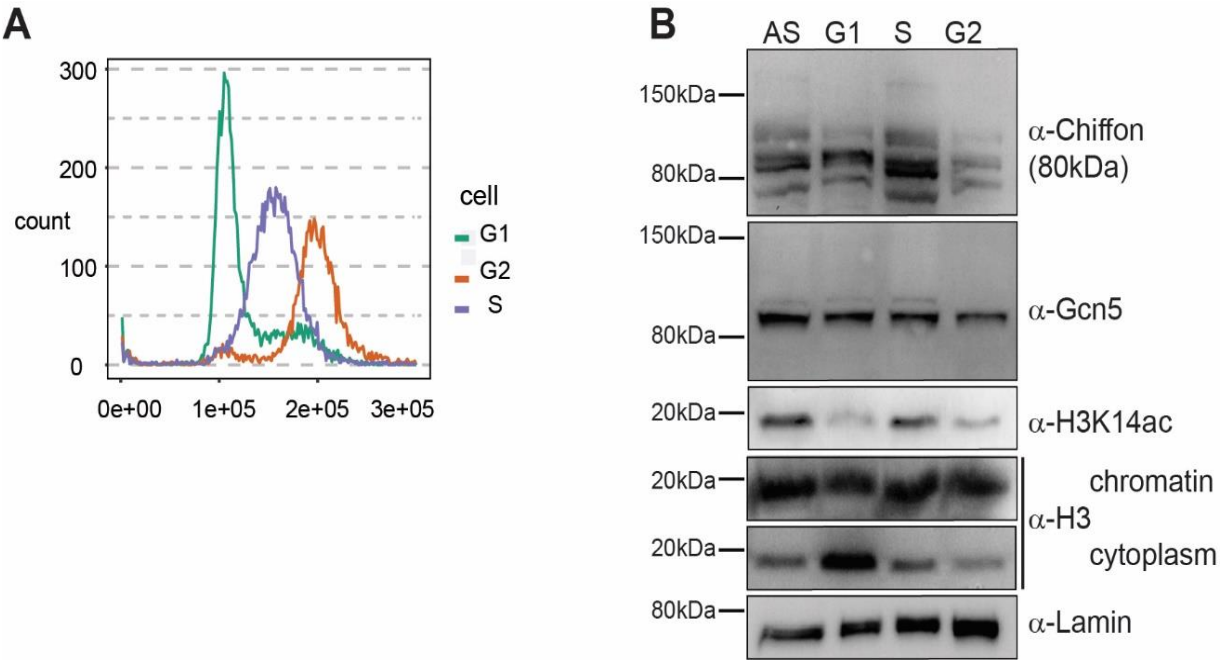

Figure 4.1 Chiffon protein is cell cycle regulated.

(A) Kc167 cells were synchronized by treatment with hydroxyurea and released for 4 hr to obtain S phase-enriched populations relative to asynchronous control. Cell cycle status was determined by flow cytometry analysis of DNA content. (B) Chromatin extracts were analyzed for presence of indicated proteins by western blotting. Solid bar indicates expected position of truncated Chiffon (80 kDa), H3K14ac (17 kDa), Gcn5 (90kDa), and lamin (74 kDa). AS accounts for asynchronized cells.

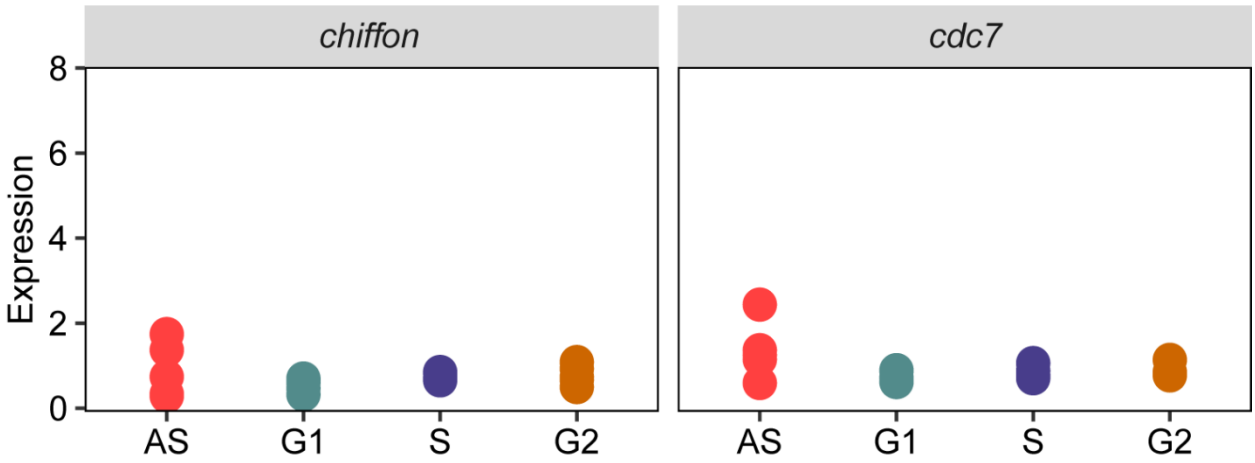

Figure 4.2 Chiffon mRNA levels are not cell cycle regulated

mRNA levels were analyzed and expression of the indicated genes (chiffon, cdc7) was assessed in five biological replicates of arrested Kc167 cells using qRT-PCR with primers specific for Chiffon or Cdc7. Transcript levels are shown relative to the geometric mean of *Rpl32 and eIf1-A*. AS accounts for asynchronized cells.

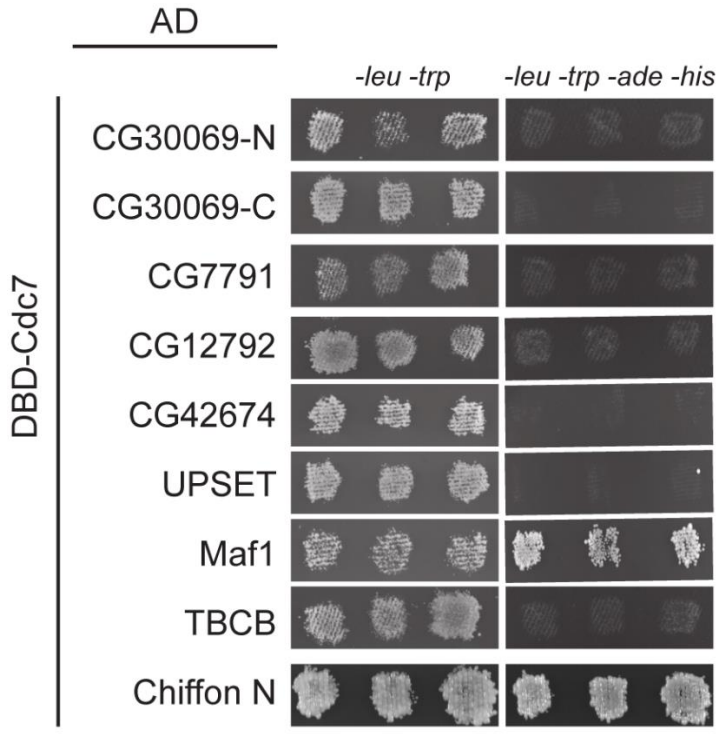

Figure 4.3 Maf1 is a possible regulatory subunit of dCdc7.

Yeast two-hybrid assay was performed to test the interaction between the indicated proteins. Three independent transformants were replica plated on media lacking leucine, tryptophan, adenine, and histidine to test for each interaction (right panel) relative to media lacking leucine and tryptophan alone to test for presence of the plasmids (left panel).

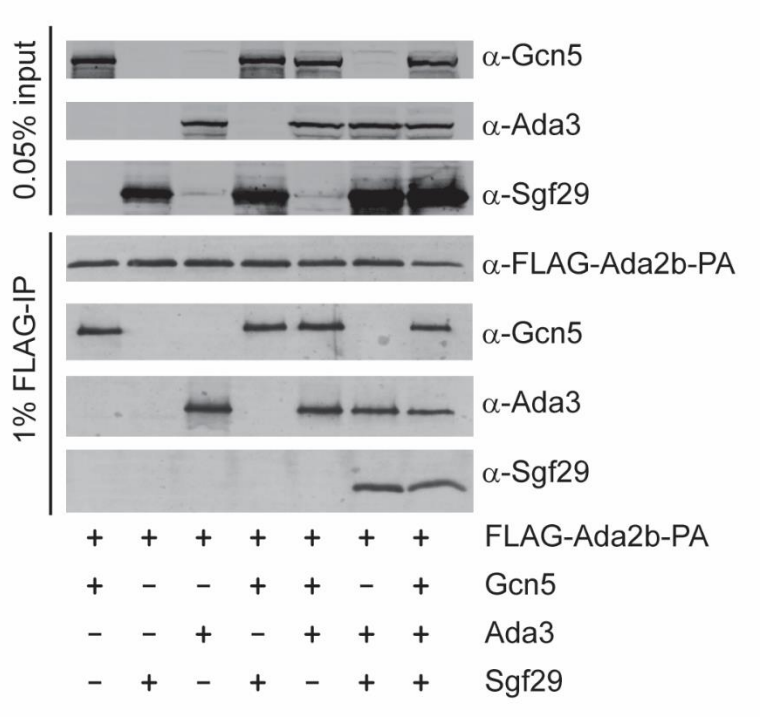

Figure 4.4 The Gcn5 core containing the Ada2b-PA can be formed in the absence of Chiffon.

Gcn5, Ada3, Sgf29, and N-HISFLAG-Ada2b-PA was co-expressed in Sf21 cells as pairwise combinations. Ni-NTA elutions were pooled and bound to anti-FLAG affinity gel beads. Protein-protein complexes were eluted and then both input and IP samples were probed with antibodies against FLAG, Gcn5, Ada3, and Sgf29.

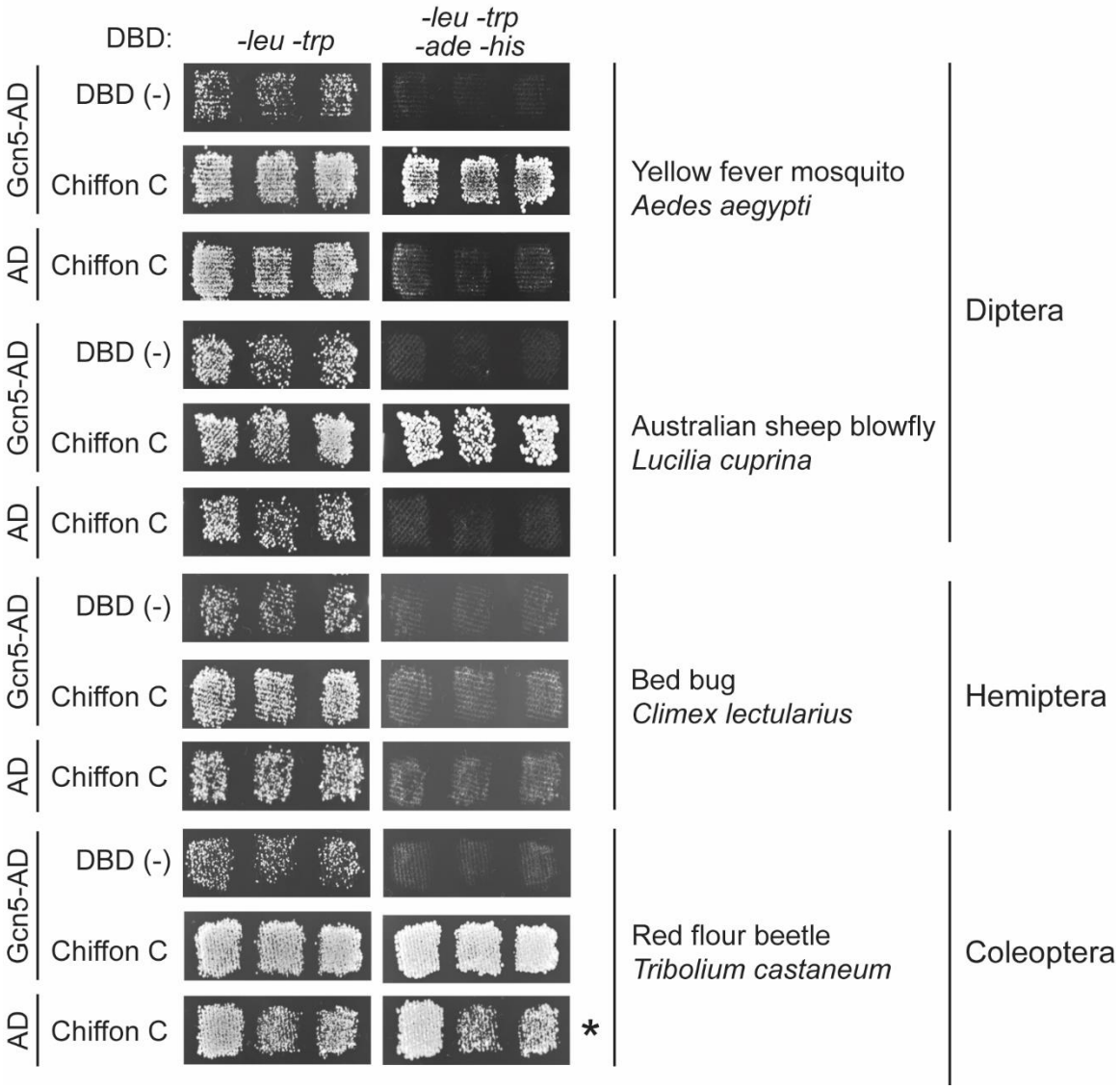

Figure 4. 5 CHAT is present in the Diptera order.

Yeast two-hybrid assay was performed to test the interaction between the indicated proteins for three different insects Orders: *Aedes aegypti* and *Lucilia cuprina* (Diptera), *Climex lectularius* (Hemiptera), and *Tribolium castaneum* (Coleoptera). Three independent transformants were replica plated on media lacking leucine, tryptophan, adenine, and histidine to test for each interaction (right panel) relative to media lacking leucine and tryptophan alone to test for presence of the plasmids (left panel). *Y2H was performed by Hannah Blum.

APPENDIX A. SUPPLEMENTAL MATERIAL CHAPTER 2

Table A1 Distributed normalized spectral abundance factor (dNSAF) values for MudPIT analysis

		Bait protein									
Subunit	**Complex**	**Spt3-C**	**Spt20-C**	**Ada2b-PB-C**	**Ada2b-PB-N**	**Sgf29-C**	**Ada2b-PA-C**	**Ada2b-PA-N**	**Chiffon-C**	**Chiffon-N**	**Cdc7-N**
Ada1	SAGA	0.003332	0.0073407	0.022048	0.0046685	0.0035263	0	0	0	0	0.000062
WDA	SAGA	0.005525	0.0155485	0.015822	0.0051914	0.0033412	0	0	0.0000535	0	0
SAF6	SAGA	0.002934	0.0066090	0.01162	0.0030241	0.0020197	0	0	0	0	0
Spt20	SAGA	0.003516	0.0235308	0.00521	0.0035582	0.0014358	0	0	0	0	0
Spt3	SAGA	0.030871	0.0120177	0.009065	0.0029124	0.0023570	0	0	0	0	0.0000250
Spt7	SAGA	0.010006	0.0160467	0.024321	0.0087099	0.0059066	0	0	0	0	0
TAF10b	SAGA	0.001874	0.0010606	0.020716	0.0023449	0.0007084	0	0	0	0	0
TAF12	SAGA	0.001924	0.0048393	0.008917	0.0024250	0.0029092	0	0	0	0.0001417	0.0000601
TAF9	SAGA	0.004430	0.0082443	0.016217	0.0046797	0.0021394	0	0.0002606	0.0002861	0	0.0000692
Nipped-A	SAGA	0.001552	0.0111466	0.001385	0.0031194	0.0022721	0	0.0000764	0.0000419	0.0000359	0.0001676
SF3B3	SAGA	0.008908	0.0180480	0.039439	0.0082218	0.0048684	0	0.0000885	0.0000324	0.0000184	0.0000470
SF3B5	SAGA	0.006641	0.0116600	0.025512	0.0051018	0.0073015	0	0	0	0	0
Nonstop	SAGA	0.002207	0.0092416	0.00023	0.0034051	0.0013555	0	0	0	0	0
Sgf11	SAGA	0.006371	0.0219648	0.001165	0.0065211	0.0010555	0	0	0	0.0000578	0
E(y)2	SAGA	0.004234	0.007973	0.00452	0.0038416	0.0028164	0	0	0	0.0004491	0
ATXN7	SAGA	0.000810	0.0034448	0.002292	0.0012693	0.0015180	0	0	0	0	0
Ada2b-PB	SAGA	0.002804	0.0097101	0.048326	0.0224538	0.0055065	0	0	0	0.0000102	0
Ada3	SAGA/CHAT/ATAC	0.001815	0.0056262	0.028636	0.0110014	0.0118136	0.0347687	0.0076228	0.0090130	0.0000611	0
Sgf29	SAGA/CHAT/ATAC	0.004380	0.0103954	0.021523	0.0199808	0.1501475	0.0603112	0.0142893	0.0141747	0.0003139	0.0001332
Gcn5	SAGA/CHAT/ATAC	0.003408	0.0096382	0.023865	0.0142614	0.0048347	0.0263765	0.0126541	0.0077293	0.0001813	0.0000236
Ada2b-PA	CHAT	0	0	0	0	0.0019684	0.1763971	0.0135192	0.0084681	0.0000135	0.0000460
Chiffon	CHAT	0	0	0	0	0.0006045	0.0011730	0.0038109	0.0209900	0.0200429	0.0009901
Cdc7	DDK	0	0	0	0	0	0	0.000207	0.0003977	0.0015066	0.0415410
Atac1	ATAC	0	0	0	0	0.0034766	0	0	0	0	0
Atac2	ATAC	0	0	0	0	0.0020988	0	0	0	0	0
D12	ATAC	0	0	0	0	0.00399154	0	0	0	0	0

Table A1 continued

		Bait protein				
Subunit	**Complex**	**Control 1**	**Control 2**	**Length (aa)**	**MW (Da)**	**pI**
Ada1	SAGA	0	0	308	35399	7.5
WDA	SAGA	0	0	743	83728	6.9
SAF6	SAGA	0	0	717	79289	6.3
Spt20	SAGA	0	0	1873	200987	9.5
Spt3	SAGA	0	0	384	43506	9
Spt7	SAGA	0	0	359	39460	5.4
TAF10b	SAGA	0	0	146	15784	6.8
TAF12	SAGA	0	0	160	17630	8.9
TAF9	SAGA	0	0	278	29314	9.3
Nipped-A	SAGA	0	4.72E-05	3790	435332	7.8
SF3B3	SAGA	0	9.72E-05	1227	136616	5.5
SF3B5	SAGA	0	0	85	9971	6.4
Nonstop	SAGA	0	0	496	56441	8
Sgf11	SAGA	0	0	196	21322	8.1
E(y)2	SAGA	0	0	101	11474	4.8
ATXN7	SAGA	0	0	971	104206	7
Ada2b-PB	SAGA	0	0	555	62031	7.1
Ada3	SAGA/CHAT/ATAC	0	0	556	59916	7.4
Sgf29	SAGA/CHAT/ATAC	0	0	289	32116	8.6
Gcn5	SAGA/CHAT/ATAC	0	0	813	92169	8.9
Ada2b-PA	CHAT	0	0	418	48316	8.8
Chiffon	CHAT	0	0	1711	189217	8.5
Cdc7	DDK	0	0	700	78665	9.2
Atac1	ATAC	0	0	356	40986	5.1
Atac2	ATAC	0	0	774	89175	6
D12	ATAC	0	0	969	111120	9.2

Relative spectral abundance of SAGA, ATAC, and CHAT subunits (rows) expressed as dNSAF (distributive normalized spectral abundance factor) in tandem FLAG-HA purification from S2 cells using the indicated proteins as bait proteins (columns), relative to control purifications from untagged S2 cells (control 1) or S2 cells expressing non-specific tagged protein (control 2) N or C epitope tagging indicated by N- or C- respectively.

Table A2 **.** Fly stocks used in chapter 2

Fly stock	**Description**	**Genotype**	**Source/Reference**
w^{1118}	Control untagged stock for embryo co-immunoprecipitations	*w*^{1118}	Bloomington BL3605
ada2b^{1}	Null *ada2b* allele removes both Ada2b splice isoforms	*ada2b*^{1}*/TM3, Sb*^{1} *Ser*^{1}	Mattias Mannervik (Qi et al., 2004)
ada2b^{842}	Null *ada2b* allele removes both Ada2b splice isoforms	*ada2b*^{842}*/TM3, Sb*^{1} *Ser*^{1}	Imre Boros (Pankotai et al., 2013)
ada2b^{1}*, FRT82B*	*ada2b, FRT82B* stock for somatic mosaic analysis	*ada2b*^{1}*, P{ry*$^{+t7.2}$*=neoFRT}82B/TM3, Sb*^{1} *Ser*^{1}	This study
Ada2b-PA	Rescue transgene for Ada2b-PA isoform	*yw; P{w*$^{+mC}$*=ADA2B-PA_ada2bEN}attP40*	(Weake et al., 2011)
Ada2b-PB	Rescue transgene for Ada2b-PA isoform	*yw; P{w*$^{+mC}$*=ADA2B-PB_ada2bEN}attP40*	This study
chiffon^{ETBE3}	Null *chiffon* allele that deletes *chiffon* coding region: chr2L: 16344400 - 16351631	*chif*^{ETBE3}*/CyO*	John Tower (Landis and Tower 1999)
chiffon^{ETBE3}*, FRT40A*	*chiffon, FRT40A* stock for somatic mosaic analysis	*chif*^{ETBE3}*, P{ry*$^{+t7.2}$*=neoFRT}40A/CyO*	(Stephenson et al. 2015)
Df(2L)RA5	Deficiency spans 35E1 – 36A1 including *chiffon* and *cactus* genes	*Df(2L)RA5/CyO*	Bloomington BL6915 (Landis and Tower 1999)
chiffon^{DsRed}	CRISPR-Cas9 allele replaces *chiffon* coding region with 3xP3-DsRed marker: chr2L: 16344356 - 16349852	*chif*$^{DsRed-attP}$*/CyO*	This study
Chiffon-FL	Rescue transgene for full-length Chiffon (1 – 1695aa)	*y w; P{w*$^{+mC}$*=Chiffon-FL*$^{1-1695}$*-NHACFLAG_chifp}attP2*	This study
Chiffon-C	Rescue transgene for C-terminal domain Chiffon (401 – 1695aa)	*y w; P{w*$^{+mC}$*=Chiffon*$^{401-1695}$*-NHACFLAG_chifp}attP2*	This study

Table A2 continued

Chiffon-N	Rescue transgene for N-terminal domain Chiffon (1 – 375aa)	*y w; P{w^{+mC}=Chiffon$^{1-400}$-NHACFLAG_ chifp}attP2*	This study
Chiffon-FL^{WF24}	Rescue transgene for full-length Chiffon (1 – 1695aa) containing nonsense mutation at 174aa (c.520C>T; p.174Q>X)	*y w; P{w^{+mC}=Chiffon-FL^{WF24}-NHACFLAG_ chifp}attP2*	This study
*Chiffon-FL**	Rescue transgene for full-length Chiffon (1 – 1695aa) containing nonsense mutation at 376aa (c.1127C>A; p.376S>X)	*y w; P{w^{+mC}=Chiffon-FL*-NHACFLAG_ chifp}attP2*	This study
hs-FLP; FRT40A, ubi-nlsGFP	Heat-shock FLP, FRT40A GFP stock for somatic mosaic analysis	*y^{1} w^{1118} P{ry$^{+t7.2}$=70FLP}3F;P{w^{+mC}=Ubi-GFP(S65T)nls}2L P{ry$^{+t7.2}$=neoFRT}40A*	This study
FRT40A; ey-FLP.N	FRT40A control stock for somatic mosaic analysis	*w^{1118}; P{ry$^{+t7.2}$=neoFRT}40A; P{ry$^{+t7.2}$=ey-FLP.N}6, ry^{506}*	Bloomington BL8212
ey-FLP;;FRT82B	FRT82B control stock for somatic mosaic analysis	*y^{d2} w^{1118} P{ry$^{+t7.2}$=ey-FLP.N}2 P{GMR-lacZ.C(38.1)}TPN1;;P{ry$^{+t7.2}$=neoFRT}82B*	Bloomington BL5619
hs-FLP;;FRT82B, ubi-nlsGFP	Heat-shock FLP, FRT28B GFP stock for somatic mosaic analysis	*y^{1} w^{1118} P{ry$^{+t7.2}$=70FLP}3F;;P{ry$^{+t7.2}$=neoFRT}82B P{w^{+mC}=Ubi-GFP(S65T)nls}3R/TM6C, Sb*	Scott Hawley

Abbreviations: c., cDNA; p., protein

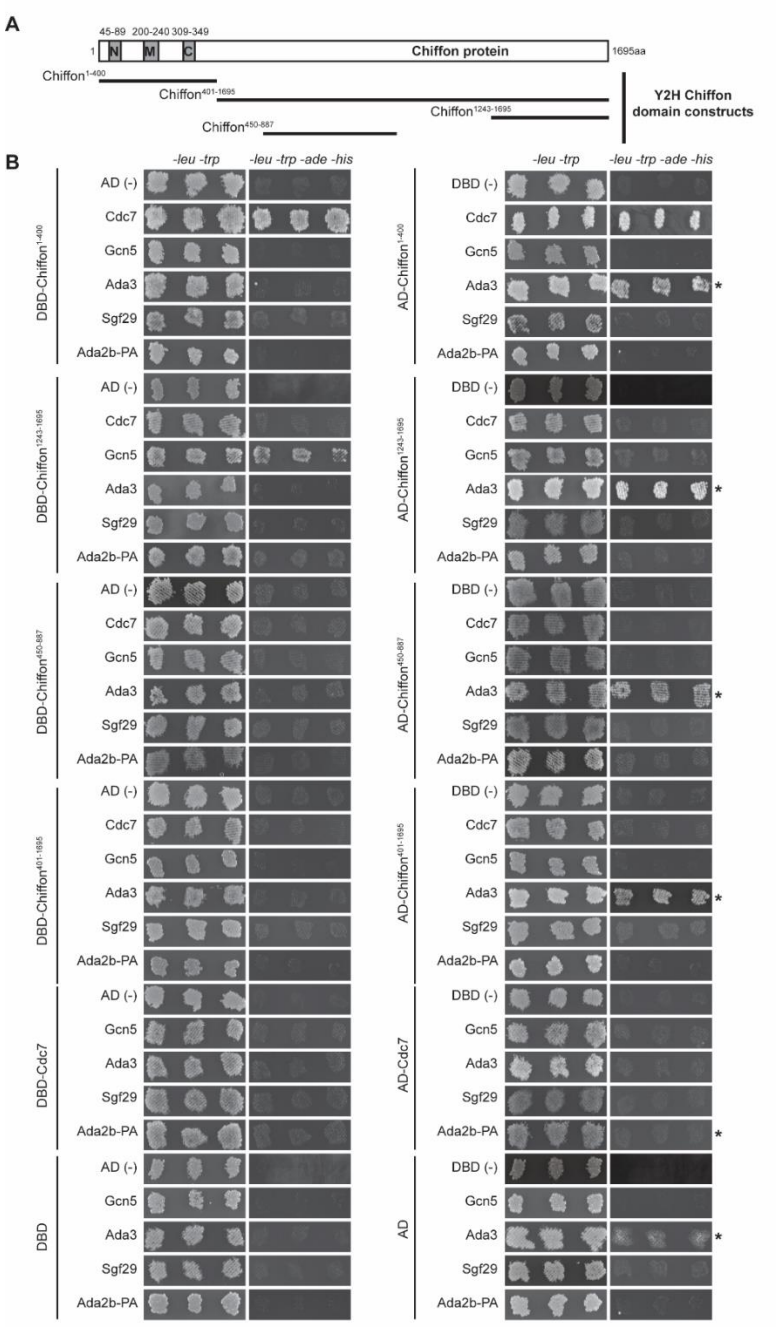

Figure A.1 The N-terminal domain of Chiffon interacts with Cdc7, and the C-terminal domain of Chiffon interacts with Gcn5.

(A) Schematic showing the Chiffon domains tested by yeast two-hybrid analysis. The conserved Dbf4 N and C motifs are indicated by shaded grey boxes. (B) Yeast two-hybrid assay was performed to test the interaction of the CHAT subunits Gcn5, Ada2b-PA, Sgf29, Ada3, and Cdc7 with different Chiffon domains. The Gal4 activating domain (AD) or the Gal4 DNA binding domain (DBD) were fused to the indicated proteins. Empty plasmids expressing only the AD or DBD were used to test for auto-activation of each protein (*, auto-activation). Cells were patched on media lacking leucine and tryptophan to test for presence of the AD and DBD plasmids, and on media lacking leucine, tryptophan, adenine and histidine to test for interaction. Three independent transformed yeast colonies were patched for each interaction tested.

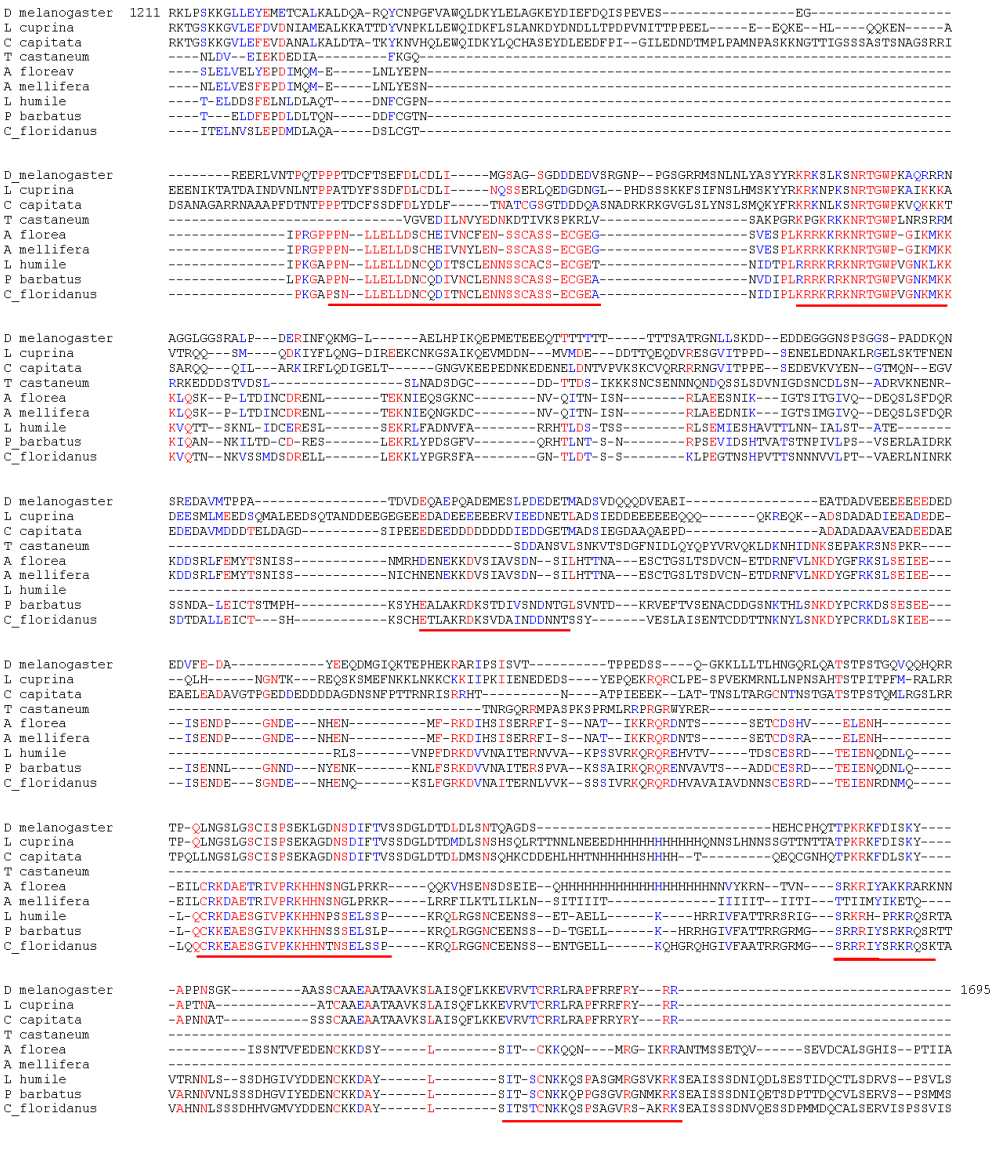

Figure A.2 The C-terminal, Gcn5-binding domain of Chiffon shares regions of conservation within insect species.

Insect Dbf4 homologs were aligned using Clustal Omega. The aligned region contains 1211 – 1695aa of *Drosophila* Chiffon, which includes the region that interacts with Gcn5 by yeast two-hybrid assay (1243 – 1695aa) and by co-immunoprecipitation using recombinant proteins (1400 – 1695aa). Regions of potential conservation within insects are underlined in red. Dbf4 homologs from the following insect species were used to generate this alignment: Diptera: *Drosophila melanogaster*, *Lucilia cuprina* (Australian sheep blowfly) and *Ceratitis capitata* (Mediterranean fruit fly). Coleoptera: *Tribolium castaneum* (Red flour beetle). Hymenoptera: *Apis mellifera* (Western honey bee), *Apis florea* (Dwarf honey bee), *Linepithema humile* (Argentine ant), *Pogonomyrmex barbatus* (Red harvester ant), *Camponotus floridanus* (Florida carpenter ant).

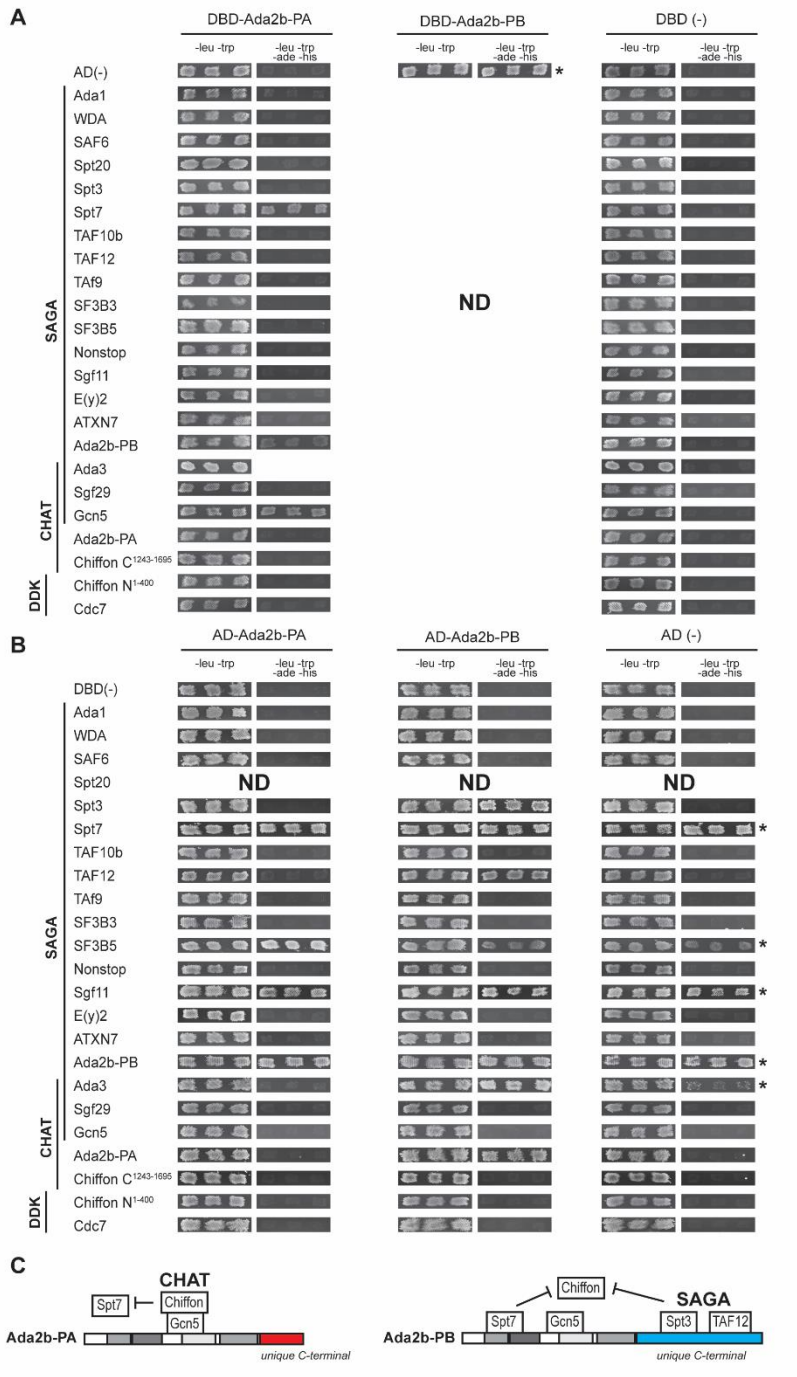

Figure A.3 Ada2b isoforms interact with overlapping and distinct SAGA subunits.

Figure A.3 continued

(A, B) Yeast two-hybrid assay was performed to test the interaction of Ada2b-PA or Ada2b-PB with the indicated SAGA subunits, Chiffon domains, and Cdc7. The Gal4 activating domain (AD) or the Gal4 DNA binding domain (DBD) were fused to the indicated proteins. Empty plasmids expressing only the AD or DBD were used to test for auto-activation of each protein (*, auto-activation). We did not test Ada2b-PB/DBD in combination with other SAGA subunits by yeast two-hybrid because Ada2b-PB auto-activated when fused to the DBD; in contrast, Ada2b-PA did not auto-activate. Cells were patched on media lacking leucine and tryptophan to test for presence of the AD and DBD plasmids, and on media lacking leucine, tryptophan, adenine and histidine to test for interaction. Three independent transformed yeast colonies were patched for each interaction tested. ND, not determined. (C) Model for differential binding of Ada2b isoforms with SAGA or CHAT. Both Ada2b isoforms bind Gcn5 and Ada3 and can interact with Spt7 by yeast two-hybrid; however, only Ada2b-PB associates with Spt7 *in vivo*. By yeast two-hybrid analysis, Ada2b-PB, but not Ada2b-PA, interacts with Spt3 and TAF12. We propose that Ada2b-PB binds Spt3 and TAF12 via its unique C-terminal domain (highlighted in blue), stabilizing association of Spt7 with its common N-terminal domain. Binding of Spt3, TAF12 and Spt7 to Ada2b-PB promotes formation of SAGA and prevents Gcn5 from binding Chiffon, potentially via steric clashes. In contrast, Ada2b-PA does not interact with Spt3 or TAF12 because it contains an alternative C-terminal domain that lacks the necessary binding regions (highlighted in red). Although Ada2b-PA is capable of binding Spt7 via its N-terminal region, this association is destabilized in the absence of Spt3 or TAF12. Instead, in the absence of Spt7 binding, Gcn5 interacts with the C-terminal of Chiffon and promotes CHAT complex formation. It is possible that Chiffon binding to Gcn5 might also prevent Ada2b-PA from binding Spt7. Notably, Ada2b-PB, but not Ada2b-PA, autoactivates expression of the reporter genes when fused to the DBD, suggesting that the unique C-terminal domain of Ada2b-PB may also interact with yeast SAGA subunits. Although by yeast two-hybrid Ada2b-PA and Ada2b-PB showed an interaction in one direction (panel B, AD-Ada2b-PB + DBD-Ada2b-PA), we never observed peptide spectra for Ada2b-PA isoforms in Ada2b-PB purifications (Table A1), suggesting that this interaction does not occur *in vivo*.

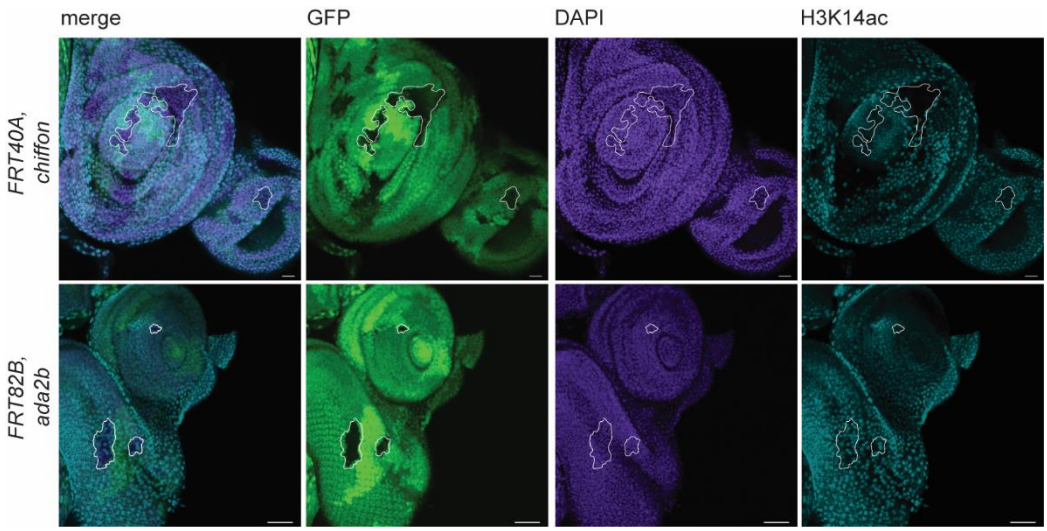

Figure A.4 Chiffon is necessary for histone H3 acetylation in imaginal discs

Mosaic imaginal discs were generated using the FLP/FRT system for *chiffon*^{ETBE3} and *ada2b*^{1}. Maximum intensity projection images showing α-H3K14ac and DAPI staining from imaginal discs containing representative clones, marked by the absence of GFP and outlined in white. Scale bars, 20 □m.

APPENDIX B. SUPPLEMENTAL MATERIAL CHAPTER 3

Table B1 Summary of *Drosophila* stocks used in chapter 3.

Fly Stock	Description	Genotype	Source/Reference
w^{1118}	Control untagged stock for immunostaining and germline clones	w^{1118}	Bloomington BL3605
chiffon^{DsRed}, *actin-Gal4*	*chiffon* null allele on chromosome 2 with actin-Gal4 for GFP sorting of *chiffon* mutant embryos	*P{w[+mC]=Act5C-GAL4}25FO1*, *chif*$^{[DsRed-attP]}$*/CyO*	This study
chiffon^{DsRed}*,actin-Gal4*; chiffonFL	*chiffon* null allele on chromosome 2 with actin-Gal4 plus a transgene expressing full length (1 – 1695aa) for GFP sorting.	*yw; P{w[+mC]=Act5C-GAL4}25FO1*, *chif*$^{[DsRed-attP]}$*/CyO; P{w+mC =ChiffonFL]-NHACFLAG_ chifp}attP2*	This study
chiffon^{DsRed}*,actin-Gal4*; chiffonFL^{WF24}	*chiffon* null allele on chromosome 2 with actin-Gal4 plus a transgene expressing FL^{WF24} (1 – 1695aa) containing a nonsense mutation at 174aa (c.520C>T; p.174Q>X) for GFP sorting.	*yw; P{w[+mC]=Act5C-GAL4}25FO1*, *chif*$^{[DsRed-attP]}$*/CyO; P{w+mC =Chiffon FL*^{WF24}*4-NHACFLAG_ chifp}attP2*	This study
chiffon^{DsRed}*,actin-Gal4*; chiffonΔN	*chiffon* null allele on chromosome 2 as actin-Gal4 plus a transgene expressing ΔN (401 – 1695aa) for GFP sorting.	*yw; P{w[+mC]=Act5C-GAL4}25FO1*, *chif*$^{[DsRed-attP]}$*/CyO; P{w+mC =ChiffonΔN-NHACFLAG_ chifp}attP2*	This study
chif^{ETBE3}*, 10XUAS:GFP*	*chiffon* null allele on chromosome 2 with 10XUAS:GFP for GFP sorting of *chiffon* null mutants embryos	*chif*$^{[ETBE3]}$, *P{w[+mC]=10XUAS-IVS-mCD8::GFP}attP40*	This study
10XUAS-IVS-mCD8:GFP	GFP control stock for GFP sorting of *chiffon* mutants	w^{*}*; P{y[+t7.7] w[+mC]=10XUAS-IVS-mCD8::GFP}attP40*	Bloomington 32186
Act5C-GAL4	Actin-Gal4 control stock for GFP sorting of *chiffon* mutants	y^{1} *w[*]; P{w[+mC]=Act5C-GAL4}25FO1/CyO, y[+]*	Bloomington 4414

Fly Stock	Description	Genotype	Source/Reference
ada2b^{1}, *10XUAS:GFP*	*ada2b* null allele on chromosome 3 with 10XUAS:GFP for GFP sorting of *ada2b* mutants embryos	*ada2b*$^{[1]}$, *P{w[+mC]=10XUAS-IVS-mCD8::GFP}su(Hw)attP1/TM3, Sb*$^{[1]}$	This study
ada2b^{842}, *elav-GAL4*	*ada2b* null allele on chromosome 3 with elav-GAL4 for GFP sorting of *ada2b* mutants embryos	*ada2b*$^{[842]}$, *w[*]; P{w[+mC]=elav-GAL4}LL7/TM3, Sb*$^{[1]}$	This study
ada2b^{1}, 10XUAS:GFP; ada2b-PA	*ada2b* null allele on chromosome 3 with 10XUAS:GFP plus a transgene expressing Ada2b-PA isoform for GFP sorting	*yw; P{w+mC =ADA2B-PA_ ada2bEN}attP40; ada2b*$^{[1]}$, *w*$^{+mC}$*=10XUAS-IVS-mCD8::GFP}su(Hw)attP1*	This study
ada2b^{1}, 10XUAS:GFP; ada2b-PB	*ada2b* null allele on chromosome 3 with 10XUAS:GFP plus a transgene expressing Ada2b-PB isoform for GFP sorting	*yw; P{w+mC =ADA2B-PB_ ada2bEN}attP40; ada2b[1], w[+mC]=10XUAS-IVS-mCD8::GFP}su(Hw)attP1ada2b*$^{[1]}$	This study
ada2b^{842}, elav4-GAL4; ada2b-PB	*Ada2b* null allele on chromosome 3 with elav-GAL4 plus a transgene expressing Ada2b-PA isoform for GFP sorting	*yw; P{w+mC =ADA2B-PB_ ada2bEN}attP40; ada2b[842], w[*]; P{w[+mC]=tubP-GAL4}LL7/TM3, Sb*$^{[1]}$	This study
10XUAS-IVS-mCD8::GFP	GFP control stock for GFP sorting of ada2b mutants	*w*^{1118}*; P{y[+t7.7] w[+mC]=10XUAS-IVS-mCD8::GFP}su(Hw)attP1*	Bloomington 32187
Elav4-GAL4	GFP control stock for GFP sorting of ada2b mutants	*w*^{1118}*; P{y[+t7.7] w[+mC]=GMR27E06-GAL4}attP2*	Bloomington 45530
Ada2b-PB-HF2	Ada2b-PB HA and FLAG tagged in the C-terminal for immunostaining	*yw; P{w+mC =ADA2B-PB-HF2_ada2bEN}attP40*	This study
Ada2b-PA-HF2	Ada2b-PB HA and FLAG tagged in the C-terminal for immunostaining	*yw; P{w+mC =ADA2B-PA-HF2_ada2bEN}attP40*	This study
chiffon^{ETBE3}*/chiffon*^{DsRed}*;* ChiffonFL-NHACFLAG	Chiffon FL HA tagged in the N-terminal and FLAG tagged in the DN-terminal rescue on *chiffon* null mutant background	*yw; chif*$^{[ETBE3]}$*, P{ry[+t7.2]=neoFRT}40A/ chiffon*$^{DsRed-attP}$*; P{w+mC =ChiffonFL-NHACFLAG_ chifp}attP2*	This study
chiffon ^{ETBE3}, FRT40A	*chiffon, FRT40A* stock for germline clone experiment	*chif*$^{[ETBE3]}$*, P{ry[+t7.2]=neoFRT}40A/CyO*	(Stephenson et al. 2015)
chiffon ^{ETBE3}, FRT40A; chiffonΔN (FLAG-HA)	*chiffon, FRT40A* stock plus transgene expressing ΔN (401 – 1695aa) for germline clone experiment	*yw; chif*$^{[ETBE3]}$*, P{ry[+t7.2]=neoFRT}40A/CyO; P{w+mC =ChiffonDN-NHACFLAG_ chifp}attP2*	This study

Table B1 continued

Fly Stock	Description	Genotype	Source/Reference
chiffon ^{ETBE3}, FRT40A; chiffonN[376S>X] (FLAG-HA)	*chiffon, FRT40A* stock plus transgene expressing N (1 – 400aa) for germline clone experiment	*yw; chif*$^{[ETBE3]}$, *P{ry[+t7.2]=neoFRT}40A/CyO; P{w+mC =ChiffonN[376S>X]-NHACFLAG_ chifp}attP2/MKRS*	This study
chiffon ^{ETBE3}, FRT40A; chiffonFL (FLAG-HA))	*chiffon, FRT40A* stock plus a transgene expressing full length (1 – 1695aa) for germline clone experiment	*yw; chif*$^{[ETBE3]}$, *P{ry[+t7.2]=neoFRT}40A/CyO; P{w+mC =ChiffonFL-NHACFLAG_ chifp}attP2/MKRS*	This study
FRT40A, ubi-nlsGFP	FRT40A control stock for germline clone experiment	*w*^{1118}*; P{w[+mC]=Ubi-GFP(S65T)nls}2L P{ry[+t7.2]=neoFRT}40A/CyO*	Bloomington 5629
hsFLP; Adv^{1}*/CyO*	Heat-shock FLP, Adv stock for germline clone experiment	*P{ry[+t7.2]=hsFLP}1, w[1118]; Adv[1]/CyO*	Bloomington 6
ovoD1, FRT40A	ovoD1 FRT40 stock used for germline clone analysis	*P{w[+mC]=ovoD1-18}2La, P{ry[+t7.2]=neoFRT}40A/Dp(?;2)bw[D], S[1] wg[Sp-1] Ms(2)M[1] bw[D]/CyO*	Bloomington 2121

Table B2 Primers sequence used in chapter 3

Gene		**Sequence**
Rpl32	Forward	5'- CTTCTGGTTTCCGGCAAGGTA-3'
	Reverse	5'- GGTCGAAATACGAGCCGCA-3'
nanos	Forward	5'-ACTTTGAGTCCGCCCATTAC-3'
	Reverse	5'-GTGGTACTGTCGCTGCATAA-3'
even skipped	Forward	5'-CCTGGTTGTGGACCTCTTG-3'
	Reverse	5'-TGCCGTTCAAGGAGTTATCC-3'
HA-*chiffon (Chiffon-A)*	Forward	5'-TCCTTACGACGTACCAGACTAT-3'
	Reverse	5'-CTGCTGTTGAGTGGCTAGTT-3'
FLAG-*chiffon (Chiffon-B)*	Forward	5'-ATCTCGCAGTTCCTGAAGAAG-3'
	Reverse	5'-TGTCATCATCGTCCTTGTAGTC-3'

Primers used for qRT-PCR analysis in single embryos. Forward and Reverse primer sequences per each set of primer are indicated.

Table B3 Spike-in factors for each sample

	Total reads	**Reads mapped to yeast**	**% reads mapped to yeast**	**Reads mapped to *Drosophila***	**% reads mapped to *Drosophila***	**=1/(reads mapped to yeast/ 1000000)**
chiffon Input	22166876	908841.916	4.1	20511010.36	92.53	1.100301
chiffon Input	8034356	258706.2632	3.22	7505695.375	93.42	3.865388
chiffon Input	24049771	759972.7636	3.16	22231608.31	92.44	1.315837
chiffon + FL Input	17523176	818332.3192	4.67	16089780.2	91.82	1.221997
chiffon + FL Input	17920568	768792.3672	4.29	16540684.26	92.3	1.300741
chiffon + FL Input	16724441	551906.553	3.3	15533660.8	92.88	1.811901
chiffon H3	25891602	284807.622	1.1	24291501	93.82	3.511142
chiffon H3	25469115	48391.3185	0.19	13114047.31	51.49	20.66486
chiffon H3	27381269	273812.69	1	25470056.42	93.02	3.652132
chiffon + FL H3	23477331	166689.0501	0.71	21880872.49	93.2	5.999194
chiffon + FL H3	27092846	357625.5672	1.32	25421217.4	93.83	2.796221
chiffon + FL H3	25378005	172570.434	0.68	23720821.27	93.47	5.794735
chiffon H3K14ac	25689421	9831341.417	38.27	11444637.06	44.55	0.101716
chiffon H3K14ac	17677254	4857709.399	27.48	10153814.7	57.44	0.205858
chiffon H3K14ac	12951478	3673039.161	28.36	7201021.768	55.6	0.272254
chiffon + FL H3K14ac	19620564	2974477.502	15.16	13234070.42	67.45	0.336193
chiffon + FL H3K14ac	22111727	3471541.139	15.7	16349410.94	73.94	0.288057
chiffon + FL H3K14ac	25864350	2260544.19	8.74	17820537.15	68.9	0.442371

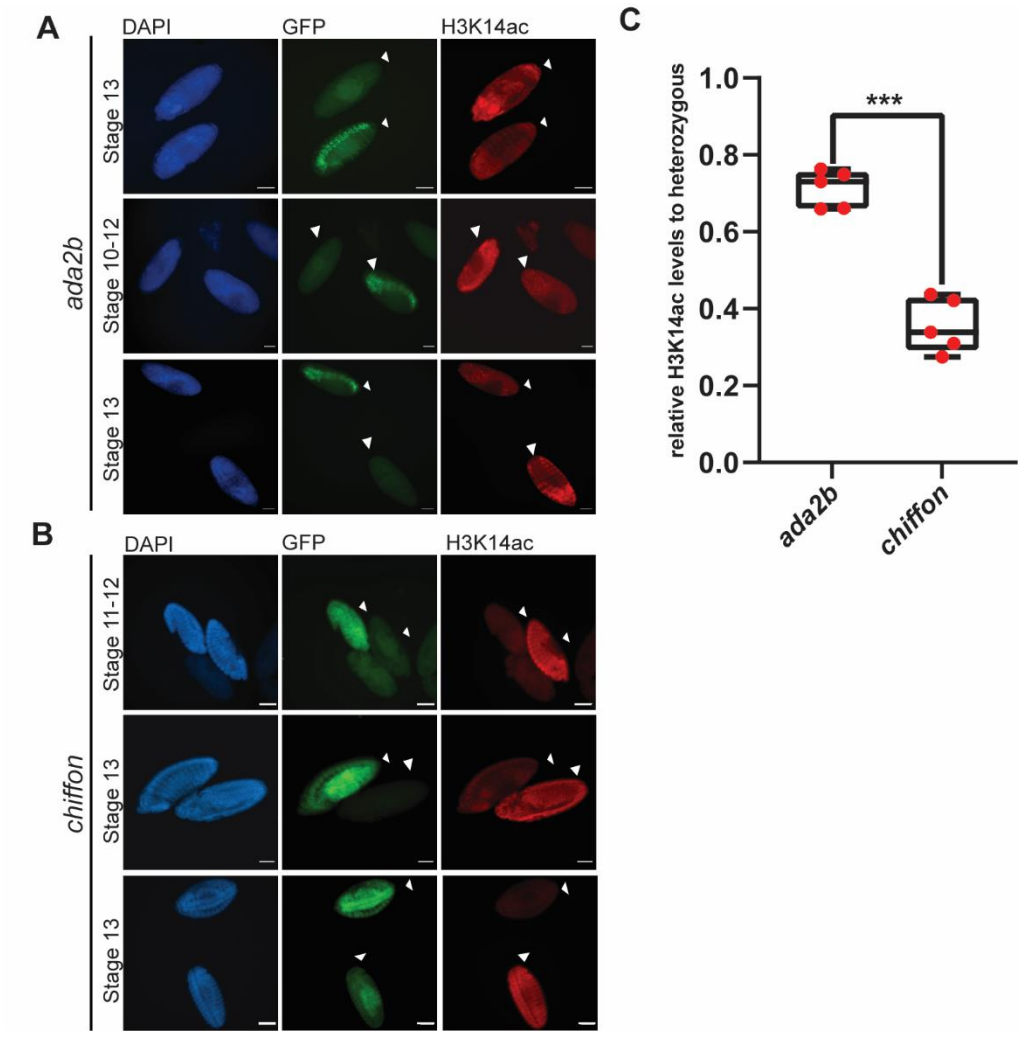

Figure B.1 Chiffon is required for H3K14ac in embryos.

(A) Stage 11 - 13 embryos were examined for DAPI and H3K14ac. GFP-positive embryos are *chiffon* null. Scale bars: 20 µm. (B) Box plots showing relative H3K14ac levels in GFP versus non-GFP (heterozygote siblings) embryos. 5 independent embryos were quantified. *p*-value (**, p = 0.0001) for the indicated comparison was determined by t-test.

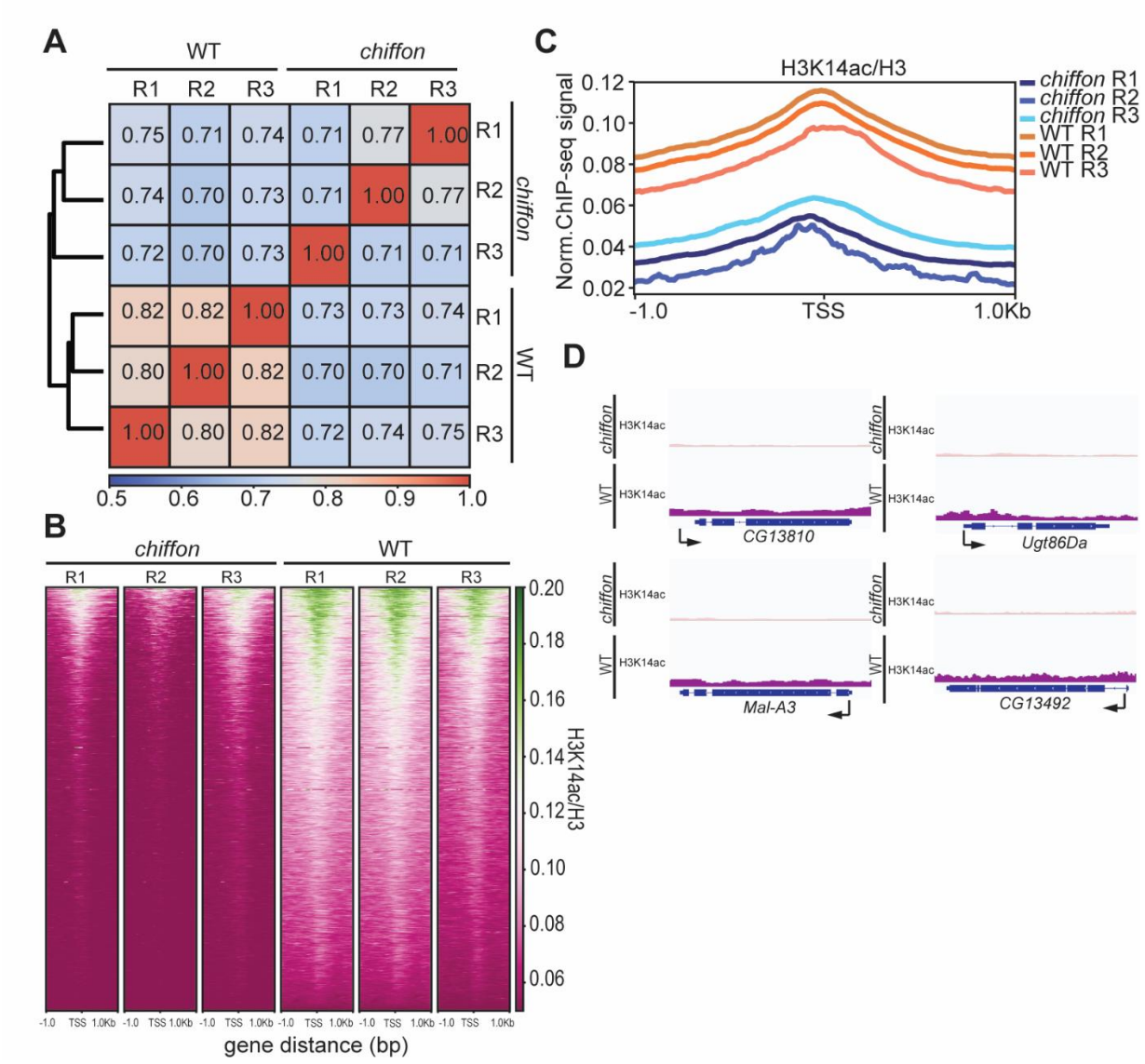

Figure B.2 CHAT is required for global H3K14ac levels in embryos.

(A) Spearman correlation heatmap of H3K14ac ChIP-seq data comparing *chiffon* and WT (*chiffon* + FL) samples. (B) Heatmaps showing RRPM-normalized H3K14ac ChIP-seq signal around the TSS of protein-coding genes comparing *chiffon* and WT. (C) Metaplot of RRPM-normalized H3K14ac ChIP-seq signal around the TSS averaged for all protein-coding genes in *chiffon* and WT samples for each replicate. (D) Genome browser snapshots showing RRPM-normalized H3K14ac signal for three representative genes containing unique H3K14ac peaks for *chiffon* and WT samples.

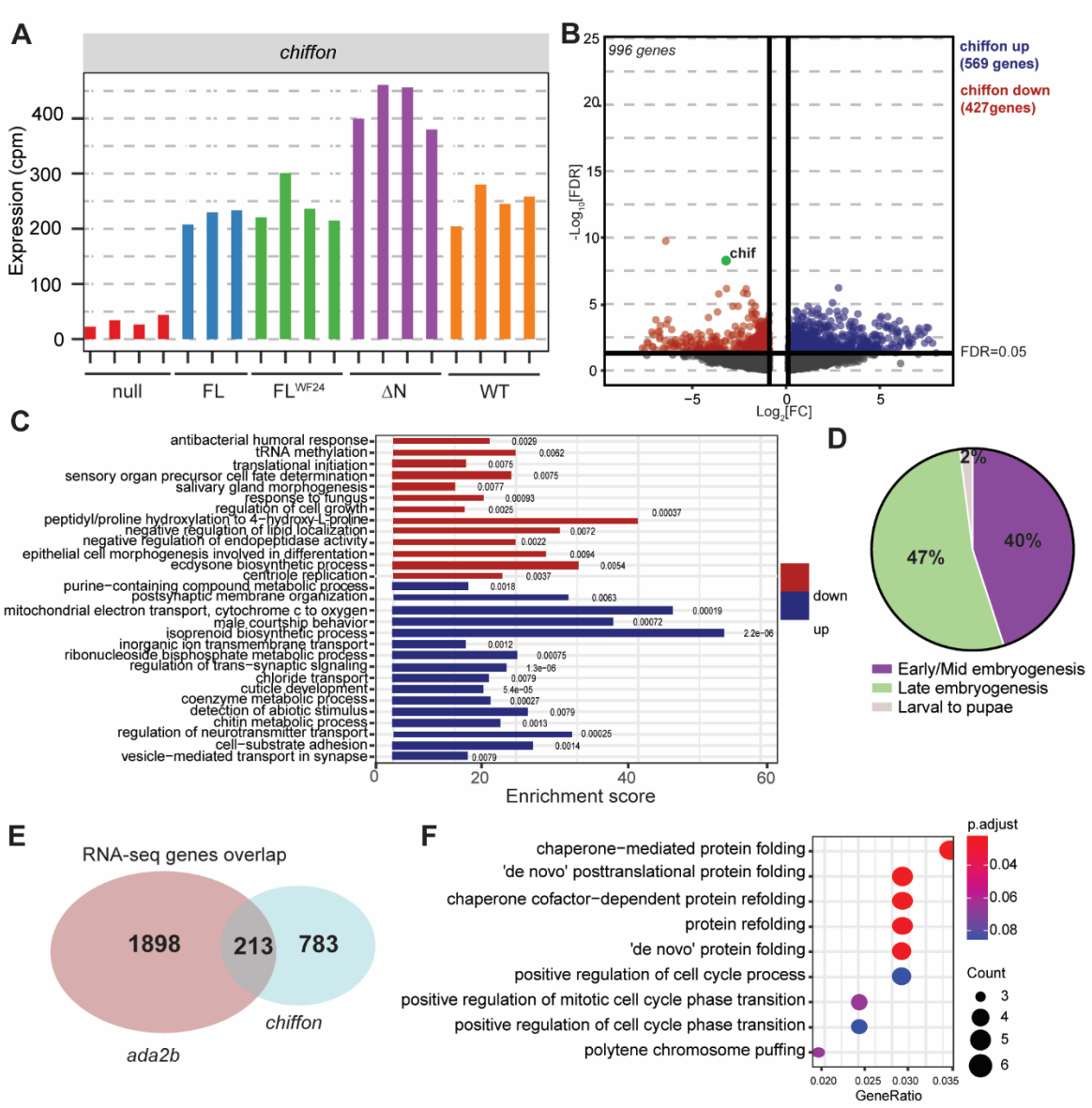

Figure B.3 Chiffon is required for expression of developmental genes.

(A) Bar plot showing counts per million (cpm) of replicate normalized counts for Chiffon across all genotypes and samples. (B) Volcano plot showing the fold change of differential expressed genes of CHAT regulated genes plotted as $\log_2$(fold change in counts per million reads, CPM) for each gene relative to its p-value ($-\log_{10}$[FDR]). The dashed line shows the statistical significance cut-off (FDR<0.05) used for DEG analysis. (C) Over-represented GO terms ($p < 0.01$, Fisher's exact test) were identified for 427 downregulated or 569 upregulated genes relative to all 9765 expressed genes using TopGO. (D) Proportion of CHAT-regulated genes in each developmental cluster. (E) Venn diagram indicating the overlapping genes detected in the *chiffon* and *ada2b* RNA-seq analysis. (F) GO terms analysis on *chiffon* and *ada2b* commonly regulated genes.

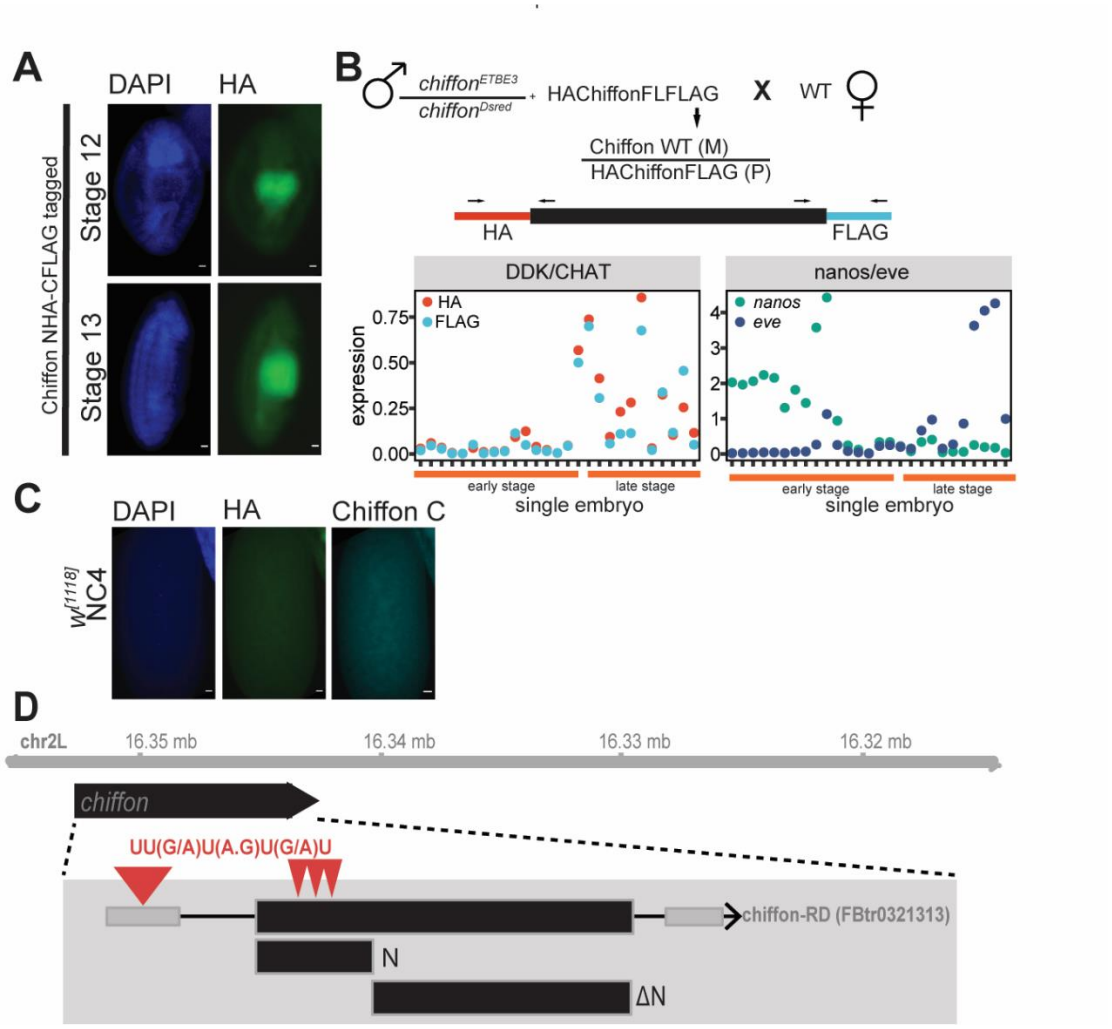

Figure B.4 The switch between Chiffon-A and B expression is not regulated at the mRNA level.

(A) Epitope-tagged Chiffon-FL stage 12 embryos were examined for DAPI and HA. Scale bars: 20 µm. (B) Male flies expressing the epitope-tagged Chiffon-FL transgene were crossed to untagged females (w^{1118}), and expression of the paternal *chiffon* gene was assessed in single embryos using qRT-PCR with primers specific for Chiffon-A (using HA forward primer and *chiffon* 5' reverse primer) or Chiffon-B (using a *chiffon* forward primer in the 3' region and a FLAG reverse primer). Transcript levels are shown relative to *Rpl32*, and embryos were ranked from earliest to latest developmental stages by examining expression of *nanos* (*nos*, early) and *even skipped* (*eve*, late). (C) Untagged WT embryo (w^{1118}) at NC4 immunostained for DAPI, HA, and Chiffon C-terminal antibody showing lack of background signal under the conditions used, and no expression of Chiffon-B in early embryos. (D) Schematic of *chiffon* gene showing Bruno response elements present in the first 1200bp (N) that is absent from ΔN region.

APPENDIX C. SUPPLEMENTAL MATERIAL CHAPTER 4

Table C1 Plasmids made for this project.

Plasmid	Gene	Insert	Experiment
pGBKT7-Chiffon N	chiffon	ChiffonN	Y2H
pGADT7-Chiffon N	chiffon	ChiffonN	Y2H
pGBKT7-Chiffon DN	chiffon	ChiffonΔN (401-1695aa)	Y2H
pGADT7-Chiffon DN	chiffon	ChiffonΔN (401-1695aa)	Y2H
pGBKT7-Chiffon DN2	chiffon	Chiffon C-terminal without DNA binding domain (600a-1695aa)	Y2H
pGBKT7-Chiffon DN2	chiffon	Chiffon C-terminal without DNA binding domain (600a-1695aa)	Y2H
pGADT7-Chiffon C2	chiffon	Chiffon C-terminal conserved motifs (450a-887aa)	Y2H
pGBKT7-Chiffon C2	chiffon	Chiffon C-terminal conserved motifs (450a-887aa)	Y2H
pGBKT7-Chiffon C	chiffon	Chiffon C-terminal conserved regions (1243-1695aa)	Y2H
pGADT7-Chiffon C	chiffon	Chiffon C-terminal conserved regions (1243-1695aa)	Y2H
pGBKT7-Ada2b-PA	Ada2b	Ada2b-PA	Y2H
pGADT7- Ada2b-PA	Ada2b	Ada2b-PA	Y2H
pGADT7-CG30069-C	CG30069	CG30069-C	Y2H
pGBKT7-CG30069-C	CG30069	CG30069-C	Y2H
pGADT7-CG30069-N	CG30069	CG30069-N	Y2H
pGBKT7-CG30069-N	CG30069	CG30069-CN	Y2H
pGADT7-CG7791	CG7791	CG7791	Y2H
pGBKT7-CG7791	CG7791	CG7791	Y2H
pGADT7-Ada2b N	Ada2b	Ada2b N-shared domain	Y2H
pGBKT7-Ada2b N	Ada2b	Ada2b N-shared domain	Y2H

Table C1 continued

Plasmid	**Gene**	**Insert**	**Experiment**
pGADT7-Ada2b-PB C	Ada2b	Ada2b-PB C-specific domain	Y2H
pGBKT7-Ada2b-PB C	Ada2b	Ada2b-PB C-specific domain	Y2H
pGADT7-Ada2b-PA C	Ada2b	Ada2b-PA C-specific domain	Y2H
pGBKT7-Ada2b-PA C	Ada2b	Ada2b-PA C-specific domain	Y2H
pGADT7-Maf1	Maf1	Maf1	Y2H
pGBKT7-Maf1	Maf1	Maf1	
pGADT7-TBCB	TBCB	TBCB	Y2H
pGBKT7-TBCB	TBCB	TBCB	Y2H
pAC-chifsgRNA-Cas9	chiffon	chifsgRNA	sgRNA to target N terminal domainn of Chiffon. (for GFP tagged)
pAC-chifCtermgsgRNA-Cas9	chiffon	chifCtermsgRNA	sgRNA to target C terminal domainn of Chiffon. (for GFP tagged)
pCA4B-Chif FL-2	chiffon	Chiffon FL	For transgenic flies
pCA4B-Chif FL^{WF24}-2	chiffon	Chiffon FL^{WF24}	For transgenic flies

Table C2 Primers used in chapter 4

Gene		**Sequence**
Rpl32	Forward	5'- CTTCTGGTTTCCGGCAAGGTA-3'
	Reverse	5'- GGTCGAAATACGAGCCGCA-3'
eIF1-A	Forward	5'-GCTGGGCAACGGTCGTCTGGAGGC-3'
	Reverse	5'-CGTCTTCAGGTTCCTGGCCTCGTCCGG-3'
chiffon	Forward	5'-TTAGGTGGCAGTCGAGCTTT-3'
	Reverse	5'-TCGCGGGAATTCTGTTTATC-3'
Cdc7	Forward	5'-AGCTGCAGTCATGGGTCTG-3'
	Reverse	5'-TAAAGGTTCCGTTGCCAATC-3'

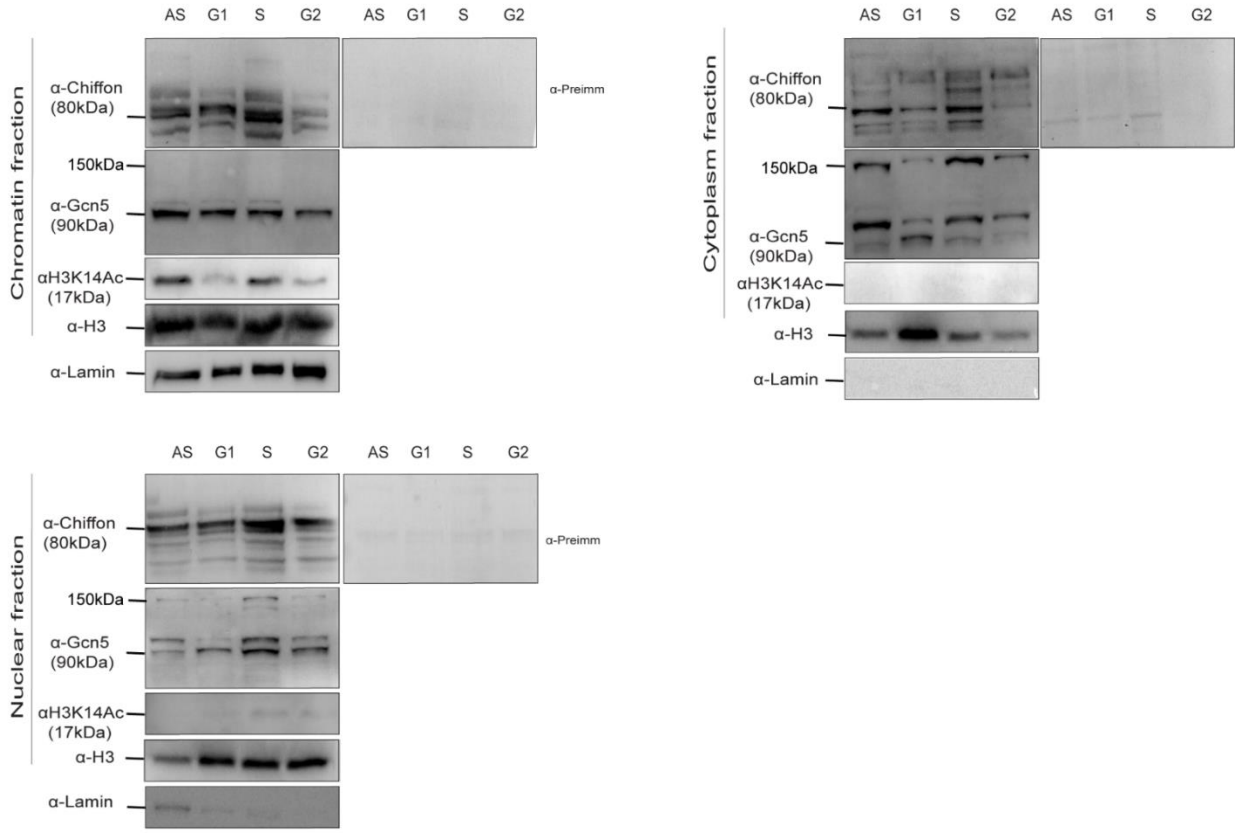

Figure C.1 Chromatin-bound Chiffon and H3K14ac are cell cycle regulated.

Chromatin, cytoplasm, and nuclear extracts were analyzed for presence of indicated proteins by western blotting. Solid bar indicates expected position of truncated Chiffon (80 kDa), H3k14ac (17 kDa), and Gcn5 (90 kDa). Pre-immune antibody was used to test for antibody specificity. As, asynchronous cells.

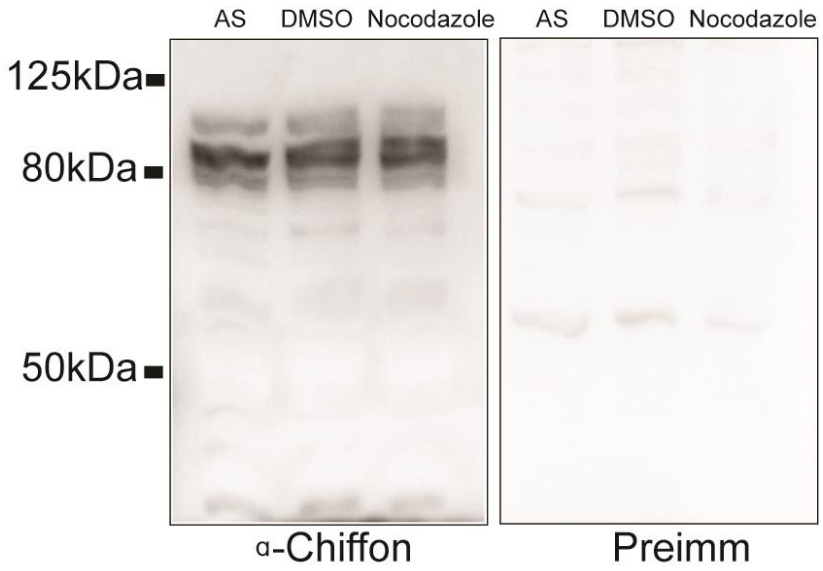

Figure C.2 Chiffon is not a substrate of the APC complex.

Kc167 cells treated with either DMSO or Nocodazole were analyzed by Western blot with antibodies against Chiffon. Pre-immune antibody was used to test for antibody specificity. As, asynchronous cells.

REFERENCES

Aggarwal BD, Calvi BR. 2004. Chromatin regulates origin activity in Drosophila follicle cells. *Nature* 430: 372–376.

Allard S, Utley RT, Savard J, Clarke A, Grant P, Brandl CJ, Pillus L, Workman JL, Cote J. 1999. NuA4, an essential transcription adaptor/histone H4 acetyltransferase complex containing Esa1p and the ATM-related cofactor Tra1p. *EMBO J* 18: 5108–5119.

Allis CD, Jenuwein T. 2016. The molecular hallmarks of epigenetic control. *Nature Reviews Genetics* 17: 487–500.

Andrew J, Smith M, Merakovsky J, Coulson M, Hannan F, Kelly LE. 1996. The stoned Locus of Drosophila melanogaster Produces a Dicistronic Transcript and Encodes Two Distinct Polypeptides. *Genetics* 143: 1699.

Antonova SV, Boeren J, Timmers HTM, Snel B. 2019. Epigenetics and transcription regulation during eukaryotic diversification: the saga of TFIID. *Genes & Development* 33: 888–902.

Axelrod A, Rine J. 1991. A role for CDC7 in repression of transcription at the silent mating-type locus HMR in Saccharomyces cerevisiae. *Mol Cell Biol* 11: 1080–1091.

Babiarz JE, Halley JE, Rine J. 2006. Telomeric heterochromatin boundaries require NuA4-dependent acetylation of histone variant H2A.Z in Saccharomyces cerevisiae. *Genes Dev* 20: 700–710.

Baker SP, Phillips J, Anderson S, Qiu Q, Shabanowitz J, Smith MM, Yates JR 3rd, Hunt DF, Grant PA. 2010. Histone H3 Thr 45 phosphorylation is a replication-associated post-translational modification in S. cerevisiae. *Nat Cell Biol* 12: 294–298.

Balasubramanian R, Pray-Grant MG, Selleck W, Grant PA, Tan S. 2002. Role of the Ada2 and Ada3 Transcriptional Coactivators in Histone Acetylation*. *Journal of Biological Chemistry* 277: 7989–7995.

Baptista T, Grünberg S, Minoungou N, Koster MJE, Timmers HTM, Hahn S, Devys D, Tora L. 2017. SAGA Is a General Cofactor for RNA Polymerase II Transcription. *Molecular Cell* 68: 130-143.e5.

Baxevanis AD, Arents G, Moudrianakis EN, Landsman D. 1995. A variety of DNA-binding and multimeric proteins contain the histone fold motif. *Nucleic Acids Res* 23: 2685–2691.

Bonnet J, Wang C-Y, Baptista T, Vincent SD, Hsiao W-C, Stierle M, Kao C-F, Tora L, Devys D. 2014. The SAGA coactivator complex acts on the whole transcribed genome and is required for RNA polymerase II transcription. *Genes Dev* 28: 1999–2012.

Brogna S, Ashburner M. 1997. The Adh-related gene of Drosophila melanogaster is expressed as a functional dicistronic messenger RNA: multigenic transcription in higher organisms. *The EMBO Journal* 16: 2023–2031.

Brown CE, Howe L, Sousa K, Alley SC, Carrozza MJ, Tan S, Workman JL. 2001. Recruitment of HAT Complexes by Direct Activator Interactions with the ATM-Related Tra1 Subunit. *Science* 292: 2333.

Brownell JE, Allis CD. 1995. An activity gel assay detects a single, catalytically active histone acetyltransferase subunit in Tetrahymena macronuclei. *Proc Natl Acad Sci USA* 92: 6364.

Brownell JE, Allis CD. 2021. HAT discovery: Heading toward an elusive goal with a key biological assist. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* 1864: 194605.

Brownell JE, Allis CD. 1996. Special HATs for special occasions: linking histone acetylation to chromatin assembly and gene activation. *Curr Opin Genet Dev* 6: 176–184.

Brownell JE, Zhou J, Ranalli T, Kobayashi R, Edmondson DG, Roth SY, Allis CD. 1996. Tetrahymena Histone Acetyltransferase A: A Homolog to Yeast Gcn5p Linking Histone Acetylation to Gene Activation. *Cell* 84: 843–851.

Burke TL, Miller JL, Grant PA. 2013. Direct inhibition of Gcn5 protein catalytic activity by polyglutamine-expanded ataxin-7. *J Biol Chem* 288: 34266–34275.

Candau R, Moore PA, Wang L, Barlev N, Ying CY, Rosen CA, Berger SL. 1996. Identification of human proteins functionally conserved with the yeast putative adaptors ADA2 and GCN5. *Mol Cell Biol* 16: 593–602.

Carré C, Ciurciu A, Komonyi O, Jacquier C, Fagegaltier D, Pidoux J, Tricoire H, Tora L, Boros IM, Antoniewski C. 2008. The Drosophila NURF remodelling and the ATAC histone acetylase complexes functionally interact and are required for global chromosome organization. *EMBO Rep* 9: 187–192.

Carré C, Szymczak D, Pidoux J, Antoniewski C. 2005. The Histone H3 Acetylase dGcn5 Is a Key Player in Drosophila melanogaster Metamorphosis. *Mol Cell Biol* 25: 8228.

Chapman JW, Johnston LH. 1989. The yeast gene, DBF4, essential for entry into S phase is cell cycle regulated. *Exp Cell Res* 180: 419–428.

Chekulaeva M, Hentze MW, Ephrussi A. 2006. Bruno Acts as a Dual Repressor of oskar Translation, Promoting mRNA Oligomerization and Formation of Silencing Particles. *Cell* 124: 521–533.

Chen K, Johnston J, Shao W, Meier S, Staber C, Zeitlinger J. 2013. A global change in RNA polymerase II pausing during the Drosophila midblastula transition. *Elife* 2: e00861.

Cheng Liang, Collyer Tim, Hardy Christopher F. J. 1999. Cell Cycle Regulation of DNA Replication Initiator Factor Dbf4p. *Molecular and Cellular Biology* 19: 4270–4278.

Ciurciu A, Komonyi O, Boros IM. 2008. Loss of ATAC-specific acetylation of histone H4 at Lys12 reduces binding of JIL-1 to chromatin and phosphorylation of histone H3 at Ser10. *J Cell Sci* 121: 3366–3372.

Ciurciu A, Komonyi O, Pankotai T, Boros IM. 2006. The Drosophila histone acetyltransferase Gcn5 and transcriptional adaptor Ada2a are involved in nucleosomal histone H4 acetylation. *Mol Cell Biol* 26: 9413–9423.

Clapier CR, Nightingale KP, Becker PB. 2002. A critical epitope for substrate recognition by the nucleosome remodeling ATPase ISWI. *Nucleic Acids Res* 30: 649–655.

Clontech Laboratories I. 2007. BacPAK Baculovirus Expression System User Manual.

Clontech Laboratories I.. *Matchmaker® Gold Yeast Two-Hybrid System User Manual*.

Cloutier SC, Wang S, Ma WK, Petell CJ, Tran EJ. 2013. Long Noncoding RNAs Promote Transcriptional Poising of Inducible Genes. *PLOS Biology* 11: e1001715.

Collart C, Smith JC, Zegerman P. 2017. Chk1 Inhibition of the Replication Factor Drf1 Guarantees Cell-Cycle Elongation at the Xenopus laevis Mid-blastula Transition. *Developmental Cell* 42: 82-96.e3.

Cornelio-Parra DV, Goswami R, Costanzo K, Morales-Sosa P, Mohan RD. 2021. Function and regulation of the Spt-Ada-Gcn5-Acetyltransferase (SAGA) deubiquitinase module. *Biochim Biophys Acta Gene Regul Mech* 1864: 194630.

Corona DFV, Clapier CR, Becker PB, Tamkun JW. 2002. Modulation of ISWI function by site-specific histone acetylation. *EMBO Rep* 3: 242–247.

Corona DFV, Tamkun JW. 2004. Multiple roles for ISWI in transcription, chromosome organization and DNA replication. *Biochim Biophys Acta* 1677: 113–119.

Daniel JA, Torok MS, Sun Z-W, Schieltz D, Allis CD, Yates JR 3rd, Grant PA. 2004. Deubiquitination of histone H2B by a yeast acetyltransferase complex regulates transcription. *J Biol Chem* 279: 1867–1871.

Danielsen ET, Moeller ME, Rewitz KF. 2013. Chapter Two - Nutrient Signaling and Developmental Timing of Maturation. In *Current Topics in Developmental Biology* (eds. A.E. Rougvie and M.B. O'Connor), Vol. 105 of, pp. 37–67, Academic Press http://www.sciencedirect.com/science/article/pii/B9780123969682000026.

David G, Abbas N, Stevanin G, Dürr A, Yvert G, Cancel G, Weber C, Imbert G, Saudou F, Antoniou E, et al. 1997. Cloning of the SCA7 gene reveals a highly unstable CAG repeat expansion. *Nature Genetics* 17: 65–70.

Demény MA, Soutoglou E, Nagy Z, Scheer E, Jànoshàzi À, Richardot M, Argentini M, Kessler P, Tora L. 2007. Identification of a Small TAF Complex and Its Role in the Assembly of TAF-Containing Complexes. *PLOS ONE* 2: e316.

Dhalluin C, Carlson JE, Zeng L, He C, Aggarwal AK, Zhou M-M, Zhou M-M. 1999. Structure and ligand of a histone acetyltransferase bromodomain. *Nature* 399: 491–496.

Díaz-Santín LM, Lukoyanova N, Aciyan E, Cheung AC. 2017. Cryo-EM structure of the SAGA and NuA4 coactivator subunit Tra1 at 3.7 angstrom resolution ed. J.M. Berger. *eLife* 6: e28384.

Dietzl G, Chen D, Schnorrer F, Su K-C, Barinova Y, Fellner M, Gasser B, Kinsey K, Oppel S, Scheiblauer S, et al. 2007. A genome-wide transgenic RNAi library for conditional gene inactivation in Drosophila. *Nature* 448: 151–156.

Downey M. 2021. Non-histone protein acetylation by the evolutionarily conserved GCN5 and PCAF acetyltransferases. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* 1864: 194608.

Doyon Y, Selleck W, Lane WS, Tan S, Cote J. 2004. Structural and functional conservation of the NuA4 histone acetyltransferase complex from yeast to humans. *Mol Cell Biol* 24: 1884–1896.

E. Vamos E, Boros IM. 2012. The C-terminal domains of ADA2 proteins determine selective incorporation into GCN5-containing complexes that target histone H3 or H4 for acetylation. *FEBS Letters* 586: 3279–3286.

Eberharter A, Sterner DE, Schieltz D, Hassan A, Yates JR, Berger SL, Workman JL. 1999. The ADA Complex Is a Distinct Histone Acetyltransferase Complex in Saccharomyces cerevisiae. *Mol Cell Biol* 19: 6621.

Ekundayo B, Bleichert F. 2019. Origins of DNA replication. *PLOS Genetics* 15: e1008320.

Elden AC, Kim H-J, Hart MP, Chen-Plotkin AS, Johnson BS, Fang X, Armakola M, Geser F, Greene R, Lu MM, et al. 2010. Ataxin-2 intermediate-length polyglutamine expansions are associated with increased risk for ALS. *Nature* 466: 1069–1075.

Espinola-Lopez JM, Tan S. 2021. The Ada2/Ada3/Gcn5/Sgf29 histone acetyltransferase module. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* 1864: 194629.

Fang D, Lengronne A, Shi D, Forey R, Skrzypczak M, Ginalski K, Yan C, Wang X, Cao Q, Pasero P, et al. 2017. Dbf4 recruitment by forkhead transcription factors defines an upstream rate-limiting step in determining origin firing timing. *Genes & Development* 31: 2405–2415.

Featherstone M. 2002. Coactivators in transcription initiation: here are your orders. *Current Opinion in Genetics & Development* 12: 149–155.

Feller C, Forné I, Imhof A, Becker PB. 2015. Global and Specific Responses of the Histone Acetylome to Systematic Perturbation. *Molecular Cell* 57: 559–571.

Ferreira R, Eberharter A, Bonaldi T, Chioda M, Imhof A, Becker PB. 2007. Site-specific acetylation of ISWI by GCN5. *BMC Molecular Biology* 8: 73.

Fukuyama K, Yoshida M, Yamashita A, Deyama T, Baba M, Suzuki A, Mohri H, Ikezawa Z, Nakajima H, Hirai S, et al. 2000. MAPK upstream kinase (MUK)-binding inhibitory protein, a negative regulator of MUK/dual leucine zipper-bearing kinase/leucine zipper protein kinase. *J Biol Chem* 275: 21247–21254.

Gamper AM, Kim J, Roeder RG. 2009. The STAGA subunit ADA2b is an important regulator of human GCN5 catalysis. *Mol Cell Biol* 29: 266–280.

Gause M, Eissenberg JC, Macrae AF, Dorsett M, Misulovin Z, Dorsett D. 2006. 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* 26: 2347–2359.

Georgakopoulos T, Gounalaki N, Thireos G. 1995. Genetic evidence for the interaction of the yeast transcriptional co-activator proteins GCN5 and ADA2. *Mol Gen Genet* 246: 723–728.

Georgiev P, Chlamydas S, Akhtar A. 2011. Drosophila dosage compensation: males are from Mars, females are from Venus. *Fly (Austin)* 5: 147–154.

Georgieva S, Nabirochkina E, Dilworth FJ, Eickhoff H, Becker P, Tora L, Georgiev P, Soldatov A. 2001. The novel transcription factor e(y)2 interacts with TAF(II)40 and potentiates transcription activation on chromatin templates. *Mol Cell Biol* 21: 5223–5231.

Gozani O, Feld R, Reed R. 1996. Evidence that sequence-independent binding of highly conserved U2 snRNP proteins upstream of the branch site is required for assembly of spliceosomal complex A. *Genes Dev* 10: 233–243.

Grant PA, Duggan L, Côté J, Roberts SM, Brownell JE, Candau R, Ohba R, Owen-Hughes T, Allis CD, Winston F, et al. 1997. Yeast Gcn5 functions in two multisubunit complexes to acetylate nucleosomal histones: characterization of an Ada complex and the SAGA (Spt/Ada) complex. *Genes & Development* 11: 1640–1650.

Grant PA, Winston F, Berger SL. 2021. The biochemical and genetic discovery of the SAGA complex. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* 1864: 194669.

Grasser KD, Rubio V, Barneche F. 2021. Multifaceted activities of the plant SAGA complex. *Biochim Biophys Acta Gene Regul Mech* 1864: 194613.

Gratz SJ, Ukken FP, Rubinstein CD, Thiede G, Donohue LK, Cummings AM, O'Connor-Giles KM. 2014. Highly Specific and Efficient CRISPR/Cas9-Catalyzed Homology-Directed Repair in Drosophila. *Genetics* 196: 961–971.

Grau B, Popescu C, Torroja L, Ortuño-Sahagún D, Boros I, Ferrús A. 2008. Transcriptional adaptor ADA3 of Drosophila melanogaster is required for histone modification, position effect variegation, and transcription. *Mol Cell Biol* 28: 376–385.

Graveley BR, Brooks AN, Carlson JW, Duff MO, Landolin JM, Yang L, Artieri CG, van Baren MJ, Boley N, Booth BW, et al. 2011. The developmental transcriptome of Drosophila melanogaster. *Nature* 471: 473–479.

Gregory PD, Wagner K, Horz W. 2001. Histone acetylation and chromatin remodeling. *Exp Cell Res* 265: 195–202.

Guelman S, Kozuka K, Mao Y, Pham V, Solloway MJ, Wang J, Wu J, Lill JR, Zha J. 2009. The Double-Histone-Acetyltransferase Complex ATAC Is Essential for Mammalian Development. *Mol Cell Biol* 29: 1176.

Guelman S, Suganuma T, Florens L, Swanson SK, Kiesecker CL, Kusch T, Anderson S, Yates JR, Washburn MP, Abmayr SM, et al. 2006a. Host Cell Factor and an Uncharacterized SANT Domain Protein Are Stable Components of ATAC, a Novel dAda2A/dGcn5-Containing Histone Acetyltransferase Complex in Drosophila. *Mol Cell Biol* 26: 871.

Guelman S, Suganuma T, Florens L, Weake V, Swanson SK, Washburn MP, Abmayr SM, Workman JL. 2006b. The essential gene wda encodes a WD40 repeat subunit of Drosophila SAGA required for histone H3 acetylation. *Mol Cell Biol* 26: 7178–7189.

Guyenet SJ, Mookerjee SS, Lin A, Custer SK, Chen SF, Sopher BL, La Spada AR, Ellerby LM. 2015. Proteolytic cleavage of ataxin-7 promotes SCA7 retinal degeneration and neurological dysfunction. *Hum Mol Genet* 24: 3908–3917.

Haimovich G, Medina DA, Causse SZ, Garber M, Millán-Zambrano G, Barkai O, Chávez S, Pérez-Ortín JE, Darzacq X, Choder M. 2013. Gene Expression Is Circular: Factors for mRNA Degradation Also Foster mRNA Synthesis. *Cell* 153: 1000–1011.

Hamm DC, Harrison MM. 2018. Regulatory principles governing the maternal-to-zygotic transition: insights from Drosophila melanogaster. *Open Biology* 8: 180183.

Han Y, Luo J, Ranish J, Hahn S. 2014. Architecture of the Saccharomyces cerevisiae SAGA transcription coactivator complex. *The EMBO Journal* 33: 2534–2546.

Hartlepp KF, Fernández-Tornero C, Eberharter A, Grüne T, Müller CW, Becker PB. 2005. The histone fold subunits of Drosophila CHRAC facilitate nucleosome sliding through dynamic DNA interactions. *Mol Cell Biol* 25: 9886–9896.

Hassan AH, Prochasson P, Neely KE, Galasinski SC, Chandy M, Carrozza MJ, Workman JL. 2002. Function and Selectivity of Bromodomains in Anchoring Chromatin-Modifying Complexes to Promoter Nucleosomes. *Cell* 111: 369–379.

Helmlinger D, Marguerat S, Villén J, Swaney DL, Gygi SP, Bähler J, Winston F. 2011. Tra1 has specific regulatory roles, rather than global functions, within the SAGA co-activator complex. *EMBO J* 30: 2843–2852.

Helmlinger D, Papai G, Devys D, Tora L. 2021. What do the structures of GCN5-containing complexes teach us about their function? *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* 1864: 194614.

Helmlinger D, Tora L. 2017. Sharing the SAGA. *Trends Biochem Sci* 42: 850–861.

Henry KW, Wyce A, Lo W-S, Duggan LJ, Emre NCT, Kao C-F, Pillus L, Shilatifard A, Osley MA, Berger SL. 2003. Transcriptional activation via sequential histone H2B ubiquitylation and deubiquitylation, mediated by SAGA-associated Ubp8. *Genes Dev* 17: 2648–2663.

Herold N, Will CL, Wolf E, Kastner B, Urlaub H, Lührmann R. 2009. Conservation of the protein composition and electron microscopy structure of Drosophila melanogaster and human spliceosomal complexes. *Mol Cell Biol* 29: 281–301.

Herzel L, Ottoz DSM, Alpert T, Neugebauer KM. 2017. Splicing and transcription touch base: co-transcriptional spliceosome assembly and function. *Nature Reviews Molecular Cell Biology* 18: 637–650.

Hollmann M, Simmerl E, Schäfer U, Schäfer M. 2002. The essential Drosophila melanogaster gene wds (will die slowly) codes for a WD-repeat protein with seven repeats. *Molecular Genetics and Genomics* 268: 425–433.

Holmberg M, Duyckaerts C, Dürr A, Cancel G, Gourfinkel-An I, Damier P, Faucheux B, Trottier Y, Hirsch EC, Agid Y, et al. 1998. Spinocerebellar ataxia type 7 (SCA7): a neurodegenerative disorder with neuronal intranuclear inclusions. *Hum Mol Genet* 7: 913–918.

Horiuchi J, Silverman N, Marcus GA, Guarente L. 1995. ADA3, a putative transcriptional adaptor, consists of two separable domains and interacts with ADA2 and GCN5 in a trimeric complex. *Mol Cell Biol* 15: 1203–1209.

Howe L, Auston D, Grant P, John S, Cook RG, Workman JL, Pillus L. 2001. Histone H3 specific acetyltransferases are essential for cell cycle progression. *Genes Dev* 15: 3144–3154.

Hughes S, Jenkins V, Dar MJ, Engelman A, Cherepanov P. 2010. Transcriptional Co-activator LEDGF Interacts with Cdc7-Activator of S-phase Kinase (ASK) and Stimulates Its Enzymatic Activity *. *Journal of Biological Chemistry* 285: 541–554.

Jackson AL, Pahl PM, Harrison K, Rosamond J, Sclafani RA. 1993. Cell cycle regulation of the yeast Cdc7 protein kinase by association with the Dbf4 protein. *Mol Cell Biol* 13: 2899–2908.

Jauregui-Lozano J, Bakhle K, Weake VM. 2021. In vivo tissue-specific chromatin profiling in Drosophila melanogaster using GFP-tagged nuclei. *Genetics*. https://doi.org/10.1093/genetics/iyab079 (Accessed July 1, 2021).

Jennings BH. 2011. Drosophila – a versatile model in biology & medicine. *Materials Today* 14: 190–195.

Jiang L, Schlesinger F, Davis CA, Zhang Y, Li R, Salit M, Gingeras TR, Oliver B. 2011. Synthetic spike-in standards for RNA-seq experiments. *Genome Res* 21: 1543–1551.

Jin Q, Yu L-R, Wang L, Zhang Z, Kasper LH, Lee J-E, Wang C, Brindle PK, Dent SYR, Ge K. 2011. Distinct roles of GCN5/PCAF-mediated H3K9ac and CBP/p300-mediated H3K18/27ac in nuclear receptor transactivation. *The EMBO Journal* 30: 249–262.

Karnay AM, Elefant F. 2017. Chapter 14 - Drosophila Epigenetics. In *Handbook of Epigenetics (Second Edition)* (ed. T.O. Tollefsbol), pp. 205–229, Academic Press http://www.sciencedirect.com/science/article/pii/B9780128053881000146.

Kohler A, Schneider M, Cabal GG, Nehrbass U, Hurt E. 2008. Yeast Ataxin-7 links histone deubiquitination with gene gating and mRNA export. *Nat Cell Biol* 10: 707–715.

Komonyi O, Schauer T, Papai G, Deak P, Boros IM. 2009. A product of the bicistronic Drosophila melanogaster gene CG31241, which also encodes a trimethylguanosine synthase, plays a role in telomere protection. *Journal of Cell Science* 122: 769–774.

Koutelou E, Farria AT, Dent SYR. 2021. Complex functions of Gcn5 and Pcaf in development and disease. *Biochim Biophys Acta Gene Regul Mech* 1864: 194609.

Koutelou E, Hirsch CL, Dent SY. 2010. Multiple faces of the SAGA complex. *Current Opinion in Cell Biology* 22: 374–382.

Kukimoto I, Elderkin S, Grimaldi M, Oelgeschläger T, Varga-Weisz PD. 2004. The Histone-Fold Protein Complex CHRAC-15/17 Enhances Nucleosome Sliding and Assembly Mediated by ACF. *Molecular Cell* 13: 265–277.

Kumagai H, Sato N, Yamada M, Mahony D, Seghezzi W, Lees E, Arai K, Masai H. 1999. A novel growth- and cell cycle-regulated protein, ASK, activates human Cdc7-related kinase and is essential for G1/S transition in mammalian cells. *Mol Cell Biol* 19: 5083–5095.

Kurat CF, Yeeles JTP, Patel H, Early A, Diffley JFX. 2017. Chromatin Controls DNA Replication Origin Selection, Lagging-Strand Synthesis, and Replication Fork Rates. *Molecular Cell* 65: 117–130.

Kusch T, Guelman S, Abmayr SM, Workman JL. 2003. Two Drosophila Ada2 homologues function in different multiprotein complexes. *Mol Cell Biol* 23: 3305–3319.

Labib K. 2010. How do Cdc7 and cyclin-dependent kinases trigger the initiation of chromosome replication in eukaryotic cells? *Genes & Development* 24: 1208–1219.

Landis G, Tower J. 1999. The Drosophila chiffon gene is required for chorion gene amplification, and is related to the yeast Dbf4 regulator of DNA replication and cell cycle. *Development* 126: 4281.

Latouche M, Lasbleiz C, Martin E, Monnier V, Debeir T, Mouatt-Prigent A, Muriel M-P, Morel L, Ruberg M, Brice A, et al. 2007. A conditional pan-neuronal Drosophila model of spinocerebellar ataxia 7 with a reversible adult phenotype suitable for identifying modifier genes. *J Neurosci* 27: 2483–2492.

Lee KK, Prochasson P, Florens L, Swanson SK, Washburn MP, Workman JL. 2004. Proteomic analysis of chromatin-modifying complexes in Saccharomyces cerevisiae identifies novel subunits. *Biochemical Society Transactions* 32: 899–903.

Lee KK, Sardiu ME, Swanson SK, Gilmore JM, Torok M, Grant PA, Florens L, Workman JL, Washburn MP. 2011. Combinatorial depletion analysis to assemble the network architecture of the SAGA and ADA chromatin remodeling complexes. *Mol Syst Biol* 7: 503.

Lee KK, Swanson SK, Florens L, Washburn MP, Workman JL. 2009. Yeast Sgf73/Ataxin-7 serves to anchor the deubiquitination module into both SAGA and Slik(SALSA) HAT complexes. *Epigenetics Chromatin* 2: 2.

Lei M, Kawasaki Y, Young MR, Kihara M, Sugino A, Tye BK. 1997. Mcm2 is a target of regulation by Cdc7–Dbf4 during the initiation of DNA synthesis. *Genes & Development* 11: 3365–3374.

Lewis E. 1960. A new standard food medium. Drosophila Information Service 34: 117–118. 34: 117–118.

Li B, Carey M, Workman JL. 2007. The Role of Chromatin during Transcription. *Cell* 128: 707–719.

Li X, Seidel CW, Szerszen LT, Lange JJ, Workman JL, Abmayr SM. 2017. Enzymatic modules of the SAGA chromatin-modifying complex play distinct roles in Drosophila gene expression and development. *Genes Dev* 31: 1588–1600.

Lim S, Kwak J, Kim M, Lee D. 2013. Separation of a functional deubiquitylating module from the SAGA complex by the proteasome regulatory particle. *Nat Commun* 4: 2641.

Linares LK, Kiernan R, Triboulet R, Chable-Bessia C, Latreille D, Cuvier O, Lacroix M, Le Cam L, Coux O, Benkirane M. 2007. Intrinsic ubiquitination activity of PCAF controls the stability of the oncoprotein Hdm2. *Nature Cell Biology* 9: 331–338.

Liu G, Zheng X, Guan H, Cao Y, Qu H, Kang J, Ren X, Lei J, Dong M-Q, Li X, et al. 2019. Architecture of Saccharomyces cerevisiae SAGA complex. *Cell Discovery* 5: 25.

Liu J, McConnell K, Dixon M, Calvi BR. 2012. Analysis of model replication origins in Drosophila reveals new aspects of the chromatin landscape and its relationship to origin activity and the prereplicative complex. *MBoC* 23: 200–212.

Liu T, Wang Q, Li W, Mao F, Yue S, Liu S, Liu X, Xiao S, Xia L. 2017. Gcn5 determines the fate of Drosophila germline stem cells through degradation of Cyclin A. *The FASEB Journal* 31: 2185–2194.

Ma J, Brennan KJ, D'Aloia MR, Pascuzzi PE, Weake VM. 2016. Transcriptome Profiling Identifies Multiplexin as a Target of SAGA Deubiquitinase Activity in Glia Required for Precise Axon Guidance During Drosophila Visual Development. *G3: Genes|Genomes|Genetics* 6: 2435.

Ma Y, Kanakousaki K, Buttitta L. 2015. How the cell cycle impacts chromatin architecture and influences cell fate. *Frontiers in Genetics* 6: 19.

MacAlpine DM, Rodríguez HK, Bell SP. 2004. Coordination of replication and transcription along a Drosophila chromosome. *Genes & Development* 18: 3094–3105.

Martin KA, Poeck B, Roth H, Ebens AJ, Ballard LC, Zipursky SL. 1995. Mutations disrupting neuronal connectivity in the Drosophila visual system. *Neuron* 14: 229–240.

Martinez E, Palhan VB, Tjernberg A, Lymar ES, Gamper AM, Kundu TK, Chait BT, Roeder RG. 2001. 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* 21: 6782–6795.

Mathew V, Pauleau A-L, Steffen N, Bergner A, Becker PB, Erhardt S. 2014. The histone-fold protein CHRAC14 influences chromatin composition in response to DNA damage. *Cell Rep* 7: 321–330.

McConnell KH, Dixon M, Calvi BR. 2012. The histone acetyltransferases CBP and Chameau integrate developmental and DNA replication programs in Drosophila ovarian follicle cells. *Development* 139: 3880–3890.

Mohan M, Herz H-M, Smith ER, Zhang Y, Jackson J, Washburn MP, Florens L, Eissenberg JC, Shilatifard A. 2011. The COMPASS family of H3K4 methylases in Drosophila. *Mol Cell Biol* 31: 4310–4318.

Mohan RD, Dialynas G, Weake VM, Liu J, Martin-Brown S, Florens L, Washburn MP, Workman JL, Abmayr SM. 2014a. Loss of Drosophila Ataxin-7, a SAGA subunit, reduces H2B ubiquitination and leads to neural and retinal degeneration. *Genes Dev* 28: 259–272.

Mohan RD, Workman JL, Abmayr SM. 2014b. Drosophila models reveal novel insights into mechanisms underlying neurodegeneration. *Fly (Austin)* 8: 148–152.

Montagnoli A, Bosotti R, Villa F, Rialland M, Brotherton D, Mercurio C, Berthelsen J, Santocanale C. 2002. Drf1, a novel regulatory subunit for human Cdc7 kinase. *The EMBO Journal* 21: 3171–3181.

Muratoglu S, Georgieva S, Pápai G, Scheer E, Enünlü I, Komonyi O, Cserpán I, Lebedeva L, Nabirochkina E, Udvardy A, et al. 2003. Two Different Drosophila ADA2 Homologues Are Present in Distinct GCN5 Histone Acetyltransferase-Containing Complexes. *Mol Cell Biol* 23: 306.

Nagy Z, Riss A, Fujiyama S, Krebs A, Orpinell M, Jansen P, Cohen A, Stunnenberg HG, Kato S, Tora L. 2010. The metazoan ATAC and SAGA coactivator HAT complexes regulate different sets of inducible target genes. *Cell Mol Life Sci* 67: 611–628.

Nassrallah A, Rougée M, Bourbousse C, Drevensek S, Fonseca S, Iniesto E, Ait-Mohamed O, Deton-Cabanillas A-F, Zabulon G, Ahmed I, et al. 2018. DET1-mediated degradation of a SAGA-like deubiquitination module controls H2Bub homeostasis eds. C.S. Hardtke, R. Amasino, E. Kaiserli, and K. Grasser. *eLife* 7: e37892.

Nuño-Cabanes C, Rodríguez-Navarro S. 2021. The promiscuity of the SAGA complex subunits: Multifunctional or moonlighting proteins? *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* 1864: 194607.

Orlando DA, Chen MW, Brown VE, Solanki S, Choi YJ, Olson ER, Fritz CC, Bradner JE, Guenther MG. 2014. Quantitative ChIP-Seq normalization reveals global modulation of the epigenome. *Cell Rep* 9: 1163–1170.

Orpinell M, Fournier M, Riss A, Nagy Z, Krebs AR, Frontini M, Tora L. 2010. The ATAC acetyl transferase complex controls mitotic progression by targeting non-histone substrates. *The EMBO Journal* 29: 2381–2394.

Oshiro Guy, Owens Julia C., Shellman Yiqun, Sclafani Robert A., Li Joachim J. 1999. Cell Cycle Control of Cdc7p Kinase Activity through Regulation of Dbf4p Stability. *Molecular and Cellular Biology* 19: 4888–4896.

Pahi Z, Borsos BN, Vedelek B, Shidlovskii YV, Georgieva SG, Boros IM, Pankotai T. 2017. TAF10 and TAF10b partially redundant roles during Drosophila melanogaster morphogenesis. *Transcription* 8: 297–306.

Pahi Z, Kiss Z, Komonyi O, Borsos BN, Tora L, Boros IM, Pankotai T. 2015. dTAF10- and dTAF10b-Containing Complexes Are Required for Ecdysone-Driven Larval-Pupal Morphogenesis in Drosophila melanogaster. *PLoS One* 10: e0142226–e0142226.

Pankotai T, Komonyi O, Bodai L, Ujfaludi Z, Muratoglu S, Ciurciu A, Tora L, Szabad J, Boros I. 2005. The homologous Drosophila transcriptional adaptors ADA2a and ADA2b are both required for normal development but have different functions. *Mol Cell Biol* 25: 8215–8227.

Pankotai T, Popescu C, Martín D, Grau B, Zsindely N, Bodai L, Tora L, Ferrús A, Boros I. 2010. Genes of the Ecdysone Biosynthesis Pathway Are Regulated by the dATAC Histone Acetyltransferase Complex in Drosophila. *Mol Cell Biol* 30: 4254.

Pankotai T, Zsindely N, Vamos EE, Komonyi O, Bodai L, Boros IM. 2013. Functional characterization and gene expression profiling of Drosophila melanogastershort dADA2b isoform-containing dSAGA complexes. *BMC Genomics* 14: 44.

Papai G, Frechard A, Kolesnikova O, Crucifix C, Schultz P, Ben-Shem A. 2020. Structure of SAGA and mechanism of TBP deposition on gene promoters. *Nature* 577: 711–716.

Pasero P, Duncker BP, Schwob E, Gasser SM. 1999. A role for the Cdc7 kinase regulatory subunit Dbf4p in the formation of initiation-competent origins of replication. *Genes Dev* 13: 2159–2176.

Poeck B, Fischer S, Gunning D, Zipursky SL, Salecker I. 2001. Glial Cells Mediate Target Layer Selection of Retinal Axons in the Developing Visual System of Drosophila. *Neuron* 29: 99–113.

Qi D, Larsson J, Mannervik M. 2004. Drosophila Ada2b Is Required for Viability and Normal Histone H3 Acetylation. *Mol Cell Biol* 24: 8080.

Regadas I, Dahlberg O, Vaid R, Ho O, Belikov S, Dixit G, Deindl S, Wen J, Mannervik M. 2021. A unique histone 3 lysine 14 chromatin signature underlies tissue-specific gene regulation. *Molecular Cell* 81: 1766-1780.e10.

Revell LJ. 2012. phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution* 3: 217–223.

Rodriguez-Jato S, Busturia A, Herr W. 2011. Drosophila melanogaster dHCF interacts with both PcG and TrxG epigenetic regulators. *PLoS One* 6: e27479–e27479.

Rodríguez-Molina JB, Tseng SC, Simonett SP, Taunton J, Ansari AZ. 2016. Engineered Covalent Inactivation of TFIIH-Kinase Reveals an Elongation Checkpoint and Results in Widespread mRNA Stabilization. *Mol Cell* 63: 433–444.

Rodriguez-Navarro S, Fischer T, Luo M-J, Antunez O, Brettschneider S, Lechner J, Perez-Ortin JE, Reed R, Hurt E. 2004. Sus1, a functional component of the SAGA histone acetylase complex and the nuclear pore-associated mRNA export machinery. *Cell* 116: 75–86.

Rothwell WF, Sullivan W. 2007. Drosophila Embryo Dechorionation. *Cold Spring Harbor Protocols* 2007: pdb.prot4826.

Sanders SL, Jennings J, Canutescu A, Link AJ, Weil PA. 2002. Proteomics of the eukaryotic transcription machinery: identification of proteins associated with components of yeast TFIID by multidimensional mass spectrometry. *Mol Cell Biol* 22: 4723–4738.

Saravanan KA, Kumar H, Chhotaray S, Preethi AL, Talokar AJ, Natarajan A, Parida S, Bhushan B, Panigrahi M. 2019. Drosophila melanogaster: a promising model system for epigenetic research. *Biological Rhythm Research* 1–19.

Schram AW, Baas R, Jansen PWTC, Riss A, Tora L, Vermeulen M, Timmers HThM. 2013. A Dual Role for SAGA-Associated Factor 29 (SGF29) in ER Stress Survival by Coordination of Both Histone H3 Acetylation and Histone H3 Lysine-4 Trimethylation. *PLOS ONE* 8: e70035.

Schübeler D, Scalzo D, Kooperberg C, van Steensel B, Delrow J, Groudine M. 2002. Genome-wide DNA replication profile for Drosophila melanogaster: a link between transcription and replication timing. *Nature Genetics* 32: 438–442.

Schüpbach T, Wieschaus E. 1989. Female sterile mutations on the second chromosome of Drosophila melanogaster. I. Maternal effect mutations. *Genetics* 121: 101.

Sclafani RA, Patterson M, Rosamond J, Fangman WL. 1988. Differential regulation of the yeast CDC7 gene during mitosis and meiosis. *Mol Cell Biol* 8: 293–300.

Sela D, Chen L, Martin-Brown S, Washburn MP, Florens L, Conaway JW, Conaway RC. 2012. Endoplasmic reticulum stress-responsive transcription factor ATF6alpha directs recruitment of the Mediator of RNA polymerase II transcription and multiple histone acetyltransferase complexes. *J Biol Chem* 287: 23035–23045.

Seller CA, O'Farrell PH. 2018. Rif1 prolongs the embryonic S phase at the Drosophila mid-blastula transition. *PLOS Biology* 16: e2005687.

Sharov G, Voltz K, Durand A, Kolesnikova O, Papai G, Myasnikov AG, Dejaegere A, Ben Shem A, Schultz P. 2017. Structure of the transcription activator target Tra1 within the chromatin modifying complex SAGA. *Nature Communications* 8: 1556.

Sheu Y-J, Stillman B. 2010. The Dbf4–Cdc7 kinase promotes S phase by alleviating an inhibitory activity in Mcm4. *Nature* 463: 113–117.

Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, et al. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology* 7: 539.

Silva T, Bradley RH, Gao Y, Coue M. 2006. Xenopus CDC7/DRF1 Complex Is Required for the Initiation of DNA Replication *. *Journal of Biological Chemistry* 281: 11569–11576.

Smith ER, Belote JM, Schiltz RL, Yang XJ, Moore PA, Berger SL, Nakatani Y, Allis CD. 1998. Cloning of Drosophila GCN5: conserved features among metazoan GCN5 family members. *Nucleic Acids Res* 26: 2948–2954.

Soffers JHM, Li X, Saraf A, Seidel CW, Florens L, Washburn MP, Abmayr SM, Workman JL. 2019. Characterization of a metazoan ADA acetyltransferase complex. *Nucleic Acids Research* 47: 3383–3394.

Soffers JHM, Workman JL. 2020. The SAGA chromatin-modifying complex: the sum of its parts is greater than the whole. *Genes Dev* 34: 1287–1303.

Spedale G, Mischerikow N, Heck AJR, Timmers HTM, Pijnappel WWMP. 2010. Identification of Pep4p as the protease responsible for formation of the. *J Biol Chem* 285: 22793–22799.

Spedale G, Timmers HThM, Pijnappel WWMP. 2012. ATAC-king the complexity of SAGA during evolution. *Genes & Development* 26: 527–541.

Spradling AC, Mahowald AP. 1980. Amplification of genes for chorion proteins during oogenesis in Drosophila melanogaster. *Proc Natl Acad Sci USA* 77: 1096.

Sriskanthadevan-Pirahas S, Deshpande R, Lee B, Grewal SS. 2018. Ras/ERK-signalling promotes tRNA synthesis and growth via the RNA polymerase III repressor Maf1 in Drosophila. *PLOS Genetics* 14: e1007202.

Stegeman R, Spreacker PJ, Swanson SK, Stephenson R, Florens L, Washburn MP, Weake VM. 2016. The Spliceosomal Protein SF3B5 is a Novel Component of Drosophila SAGA that Functions in Gene Expression Independent of Splicing. *J Mol Biol* 428: 3632–3649.

Stephenson R, Hosler MR, Gavande NS, Ghosh AK, Weake VM. 2015. Characterization of a Drosophila Ortholog of the Cdc7 Kinase: A ROLE FOR Cdc7 IN ENDOREPLICATION INDEPENDENT OF CHIFFON. *Journal of Biological Chemistry* 290: 1332–1347.

Sterner DE, Belotserkovskaya R, Berger SL. 2002. SALSA, a variant of yeast SAGA, contains truncated Spt7, which correlates with activated transcription. *Proc Natl Acad Sci USA* 99: 11622.

Sterner DE, Grant PA, Roberts SM, Duggan LJ, Belotserkovskaya R, Pacella LA, Winston F, Workman JL, Berger SL. 1999. Functional organization of the yeast SAGA complex: distinct components involved in structural integrity, nucleosome acetylation, and TATA-binding protein interaction. *Mol Cell Biol* 19: 86–98.

Strahl BD, Briggs SD. 2021. The SAGA continues: The rise of cis- and trans-histone crosstalk pathways. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* 1864: 194600.

Suganuma T, Gutiérrez JL, Li B, Florens L, Swanson SK, Washburn MP, Abmayr SM, Workman JL. 2008. ATAC is a double histone acetyltransferase complex that stimulates nucleosome sliding. *Nature Structural & Molecular Biology* 15: 364–372.

Suganuma T, Mushegian A, Swanson SK, Abmayr SM, Florens L, Washburn MP, Workman JL. 2010. The ATAC Acetyltransferase Complex Coordinates MAP Kinases to Regulate JNK Target Genes. *Cell* 142: 726–736.

Tanaka S, Araki H. 2010. Regulation of the initiation step of DNA replication by cyclin-dependent kinases. *Chromosoma* 119: 565–574.

Tie F, Banerjee R, Stratton CA, Prasad-Sinha J, Stepanik V, Zlobin A, Diaz MO, Scacheri PC, Harte PJ. 2009. CBP-mediated acetylation of histone H3 lysine 27 antagonizes Drosophila Polycomb silencing. *Development* 136: 3131–3141.

Timmers HThM. 2021. SAGA and TFIID: Friends of TBP drifting apart. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* 1864: 194604.

Torres-Zelada EF, Stephenson RE, Alpsoy A, Anderson BD, Swanson SK, Florens L, Dykhuizen EC, Washburn MP, Weake VM. 2019. The Drosophila Dbf4 ortholog Chiffon forms a complex with Gcn5 that is necessary for histone acetylation and viability. *J Cell Sci* 132: jcs214072.

Torres-Zelada EF, Weake VM. 2020. The Gcn5 complexes in Drosophila as a model for metazoa. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* 194610.

Tower J. 2004. Developmental gene amplification and origin regulation. *Annu Rev Genet* 38: 273–304.

Vermeulen M, Eberl HC, Matarese F, Marks H, Denissov S, Butter F, Lee KK, Olsen JV, Hyman AA, Stunnenberg HG, et al. 2010. Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers. *Cell* 142: 967–980.

Vogelauer M, Rubbi L, Lucas I, Brewer BJ, Grunstein M. 2002. Histone Acetylation Regulates the Time of Replication Origin Firing. *Molecular Cell* 10: 1223–1233.

Wang H, Dienemann C, Stützer A, Urlaub H, Cheung ACM, Cramer P. 2020a. Structure of the transcription coactivator SAGA. *Nature* 577: 717–720.

Wang Y, Huang Y, Liu J, Zhang J, Xu M, You Z, Peng C, Gong Z, Liu W. 2020b. Acetyltransferase GCN5 regulates autophagy and lysosome biogenesis by targeting TFEB. *EMBO Rep* 21: e48335.

Wang Y-L, Faiola F, Xu M, Pan S, Martinez E. 2008. 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* 283: 33808–33815.

Wangler MF, Yamamoto S, Bellen HJ. 2015. Fruit Flies in Biomedical Research. *Genetics* 199: 639.

Weake VM, Dyer JO, Seidel C, Box A, Swanson SK, Peak A, Florens L, Washburn MP, Abmayr SM, Workman JL. 2011. Post-transcription initiation function of the ubiquitous SAGA complex in tissue-specific gene activation. *Genes Dev* 25: 1499–1509.

Weake VM, Lee KK, Guelman S, Lin C-H, Seidel C, Abmayr SM, Workman JL. 2008. SAGA-mediated H2B deubiquitination controls the development of neuronal connectivity in the Drosophila visual system. *EMBO J* 27: 394–405.

Weake VM, Swanson SK, Mushegian A, Florens L, Washburn MP, Abmayr SM, Workman JL. 2009. A novel histone fold domain-containing protein that replaces TAF6 in Drosophila SAGA is required for SAGA-dependent gene expression. *Genes Dev* 23: 2818–2823.

Weake VM, Workman JL. 2010. Inducible gene expression: diverse regulatory mechanisms. *Nature Reviews Genetics* 11: 426–437.

Weake VM, Workman JL. 2012. SAGA function in tissue-specific gene expression. *Trends Cell Biol* 22: 177–184.

Weinreich M, Stillman B. 1999. Cdc7p–Dbf4p kinase binds to chromatin during S phase and is regulated by both the APC and the RAD53 checkpoint pathway. *The EMBO Journal* 18: 5334–5346.

Welihinda A, Tirasophon W, Kaufman R. 2000. The transcriptional co-activator ADA5 is required for HAC1 mRNA processing in vivo. *The Journal of biological chemistry* 275: 3377–81.

Winston F, Chaleff DT, Valent B, Fink GR. 1984. Mutations affecting Ty-mediated expression of the HIS4 gene of Saccharomyces cerevisiae. *Genetics* 107: 179–197.

Wolffe AP. 1994. Nucleosome positioning and modification: chromatin structures that potentiate transcription. *Trends in Biochemical Sciences* 19: 240–244.

Xie G, Yu Z, Jia D, Jiao R, Deng W-M. 2014. E(y)1/TAF9 mediates the transcriptional output of Notch signaling in Drosophila. *J Cell Sci* 127: 3830.

Xu W, Edmondson DG, Evrard YA, Wakamiya M, Behringer RR, Roth SY. 2000. Loss of Gcn5l2 leads to increased apoptosis and mesodermal defects during mouse development. *Nat Genet* 26: 229–232.

Yamamoto S, Jaiswal M, Charng W-L, Gambin T, Karaca E, Mirzaa G, Wiszniewski W, Sandoval H, Haelterman NA, Xiong B, et al. 2014. A drosophila genetic resource of mutants to study mechanisms underlying human genetic diseases. *Cell* 159: 200–214.

Yang S-H, Sharrocks AD, Whitmarsh AJ. 2003. Transcriptional regulation by the MAP kinase signaling cascades. *Gene* 320: 3–21.

Yang XJ, Ogryzko VV, Nishikawa J, Howard BH, Nakatani Y. 1996. A p300/CBP-associated factor that competes with the adenoviral oncoprotein E1A. *Nature* 382: 319–324.

Yokoyama A, Cleary ML. 2008. Menin Critically Links MLL Proteins with LEDGF on Cancer-Associated Target Genes. *Cancer Cell* 14: 36–46.

Yoshizawa-Sugata N, Ishii A, Taniyama C, Matsui E, Arai K, Masai H. 2005. A Second Human Dbf4/ASK-related Protein, Drf1/ASKL1, Is Required for Efficient Progression of S and M Phases *. *Journal of Biological Chemistry* 280: 13062–13070.

Zeitlinger J, Zinzen RP, Stark A, Kellis M, Zhang H, Young RA, Levine M. 2007. Whole-genome ChIP–chip analysis of Dorsal, Twist, and Snail suggests integration of diverse patterning processes in the Drosophila embryo. *Genes & Development* 21: 385–390.

Zhang H, Tower J. 2004. Sequence requirements for function of the Drosophila chorion gene locus ACE3 replicator and ori-β origin elements. *Development* 131: 2089–2099.

Zhang Y, Wen Z, Washburn MP, Florens L. 2010. Refinements to Label Free Proteome Quantitation: How to Deal with Peptides Shared by Multiple Proteins. *Anal Chem* 82: 2272–2281.

Zsindely N, Pankotai T, Ujfaludi Z, Lakatos D, Komonyi O, Bodai L, Tora L, Boros IM. 2009. The loss of histone H3 lysine 9 acetylation due to dSAGA-specific dAda2b mutation influences the expression of only a small subset of genes. *Nucleic Acids Res* 37: 6665–6680.

VITA

Eliana F. Torres-Zelada

Education

Universidad Peruana Cayetano Heredia, Peru	Biology	B.S	2012
Purdue University, IN	Biochemistry	graduate student	2015-present

Employment and professional experience

Graduate research assistant, Department of Biochemistry, Purdue University
Advisor: Dr. Vikki Weake 04/2016-present

Intern, Instituto Técnológico de Chascomús, Argentina 10/2014 – 06/2015
PI: Dr. Pablo Strobl-Mazulla

Intern, Vienna Biocenter Summer School, Austria 06/2012 – 08/2012
PI: Dr. Wolfgang Busch

Research Assistant, Instituto de Medicina Tropical, Peru 01/2012 – 12/2014
PI: Dr. Dionicia Gamboa

Undergrad research assistant, Instituto de Medicina Tropical, Peru 03/2011-12/2011
PI: Dr. Dionicia Gamboa

Intern, Yale University, USA 01/2011 – 03/2011
PI: Dr. Martín García-Castro

Undergad research assistant, Peru 03/2009 – 05/2010
PI: Dr. José Espinoza

Honor and Awards

Arnold Kent Balls Award, Department of Biochemistry, Purdue University 04/2021

Bilsland Dissertation Fellowship Award, Department of Biochemistry, Purdue University, 04/2020

Purdue Research Foundation (PRF) Research Grant Award, Department of Biochemistry, Purdue University, 04/2019

Bird Stair Graduate Research Fellowship Award, Biochemistry Department, Purdue University, 04/2018

Graduate School's Lynn Fellowship, Purdue University 08/2015-08/2016

Vienna Biocenter Summer School Fellowship (Undergraduate) 2012

Research Experience for Peruvian Undergraduates Fellowship (Undergraduate) 2011

Consejo Nacional de Ciencia y Tecnologia (Concytec) Scholarship (Undergraduate) 2011

Conferences and presentations

Biochemistry Retreat, Purdue University. 08/2019. Talked entitled "Mapping the role of Chiffon as part of the CHAT complex in *Drosophila*".

Biochemistry Retreat, Purdue University. Talked entitled 08/2018. "The Dbf4 ortholog Chiffon forms a novel histone acetyltransferase complex with Gcn5 that is necessary for histone H3 acetylation in flies"

CIS Symposium. Purdue University. 05/2018. Talked entitled "*Chromatin in development and ageing"*

Mechanisms of Eukaryotic Transcription, Cold Spring Harbor, New York, 09/2019. Poster entitled "The Dbf4 ortholog Chiffon forms a novel histone acetyltransferase complex with Gcn5 that is necessary for histone H3 acetylation in flies)

Transcriptional regulation by Chromatin and RNA polymerase II symposium, Utah, 10/2018. Poster entitled "The Dbf4 ortholog Chiffon forms a novel histone acetyltransferase complex with Gcn5 that is necessary for histone H3 acetylation in flies."

The American Society of Tropical Medicine and Hygiene, Washington, 2013. Poster entitled "Population structure of *Plasmodium vivax* in an urban village of the Peruvian Amazon (San Juan-Iquitos)

Publications

Torres-Zelada, E. F., and Weake, V. M. (2020). The Gcn5 complexes in Drosophila as a model for metazoans. *Biochim. Biophys. Acta BBA - Gene regulatory mechanisms*, p 194610.

Torres-Zelada, E. F., Stephenson, R. E., Alpsoy, A., Anderson, B. D., Swanson, S. K., Florens, L., Dykhuizen, E. C., Washburn, M. P., and Weake, V. M. (2019). The Drosophila Dbf4 ortholog Chiffon forms a complex with Gcn5 that is necessary for histone acetylation and viability. *J. Cell Sci.* 16;132(2)

Bouzas, S. O., Marini, M. S., **Torres Zelada, E.,** Buzzi, A. L., Morales Vicente, D. A., and Strobl-Mazzulla, P. H. (2016) Epigenetic activation of Sox2 gene in the developing vertebrate neural plate. *Mol. Biol. Cell.* 27, 1921–1927

Delgado-Ratto, C., Gamboa, D., Soto-Calle, V. E., Van den Eede, P., **Torres, E.**, Sánchez-Martínez, L., Contreras-Mancilla, J., Rosanas-Urgell, A., Rodriguez Ferrucci, H., Llanos-Cuentas, A., Erhart, A., Van Geertruyden, J.-P., and D'Alessandro, U. (2016) Population Genetics of Plasmodium vivax in the Peruvian Amazon. *PLoS Negl. Trop. Dis***.** 10, e0004376

Manuscripts submitted/in Preparation

Torres-Zelada, E. F., George, Smitha, Blum, Hannah R., and Weake, Vikki M. (2021) CHAT acts upstream of other Gcn5 complexes during early *Drosophila* embryogenesis, triggering global histone H3 acetylation and expression of developmental genes. *Submitted Manuscript.*

George, Smitha, Blum, Hannah R., **Torres-Zelada, E. F**., and Weake, Vikki M (2021) Conserved role of Dbf4 ortholog Chiffon across selected dipteran insects. *Manuscript in preparation.*

Teaching Activities

Graduate Teaching Assistant, Department of Biochemistry, Purdue University

Biochemistry Lab 101 (BCHM 101), Dr. Vikki Weake 01/2019 – 05/2019

Graduate Teaching Assistant, Department of Biochemistry, Purdue University

Biochemistry Lab 101 (BCHM 101), Dr. Vikki Weake 01/2018 – 05/2018

Graduate Mentor, Department of Biochemistry, Purdue University for:

Ben Anderson (Biochemistry undergraduate, 08/2017 – 05/2019)

Ben Anderson (SURF undergraduate, 05/2017 – 07/2017)

Sollymar Pellot (SROP undergraduate, 05/2018 – 07/2018)

Jillian Cornell (PULSe rotation student, 10/2019 – 12/2019)

Hannah Blum (Biochemistry undergraduate, 08/2019 – present)

Service and Leadership

President, Peruvian Association at Purdue (PAP), Purdue University 04/2019 – 04/2020
Application committee, Research Experience for Peruvian Undergraduates 01/2017 – 01/2019