UNDERSTANDING HETEROCHROMATIN BIOLOGY THROUGH HISTONE MUTAGENESIS

Taylor Joel Richard Penke

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Curriculum in Genetics and Molecular Biology.

Chapel Hill
2017

Approved by:

Robert J. Duronio

A. Gregory Matera

Daniel J. McKay

Brian D. Strahl

Jeff J. Sekelsky

ABSTRACT

Taylor Joel Richard Penke: Understanding Heterochromatin Biology
through Histone Mutagenesis
(Under the direction of Robert J. Duronio)

The relatively large genomes of eukaryotic cells must be organized and compacted within the nucleus while maintaining DNA accessibility for essential processes such as transcription, replication, and DNA repair. This organization is accomplished in large part through the interaction of DNA with histone proteins to form a structure known as chromatin. Chromatin organization is regulated through epigenetic mechanisms, such as histone post-translational modifications (PTMs) or the differential incorporation of variant or canonical histones into chromatin. These processes are regulated differently throughout the genome, leading to functionally distinct chromatin environments. Regions where DNA is “open” or more accessible are collectively referred to as euchromatin, whereas “closed” or inaccessible regions are classified as heterochromatin. Proper heterochromatin formation is essential for regulating numerous cellular processes including, cell division, nuclear organization, gene expression, and DNA replication. A defining feature of heterochromatin is methylation of lysine nine on histone H3 (H3K9me), a histone PTM that recruits Heterochromatin Protein 1 (HP1). Although H3K9 methyltransferases and HP1 are necessary for proper heterochromatin structure, the specific contribution of H3K9 to heterochromatin function and animal development is unknown. Using our recently developed platform to engineer histone

genes in *Drosophila,* I generated H3K9R mutant flies, separating the functions of H3K9 and non-histone substrates of H3K9 methyltransferases. I observed that H3K9 plays an essential role in regulating the structure of pericentromeric heterochromatin and the repression of transposons, but not protein-coding gene expression. Furthermore, I generated a K9R mutation in the variant histone H3.3, revealing functional redundancies between variant H3.3K9 and canonical H3K9, though to differing extents in heterochromatin and euchromatin. Finally, I have used the H3K9R mutant as a tool to uncover general principles of genome regulation. Several previous studies have identified correlations between histone PTMs, transcription, and DNA replication; however, no causative relationship between these processes has been identified. The H3K9R mutant specifically disrupts pericentromeric heterochromatin providing a unique opportunity to determine the direct consequences of altered chromatin structure on replication. We demonstrated that changes in chromatin accessibility and most likely transcription are required but not sufficient for altered replication, influencing the framework through which we view genome regulation.

To my family both old and new

ACKNOWLEDGEMENTS

My journey through graduate school has benefitted from numerous individuals all of whom have my heartfelt thanks for making this experience personally and professionally fulfilling. First, I would like to thank Bob who has been an exemplary mentor and whose guidance has improved all aspects of my science. I am particularly grateful for the independence I was granted, his trust in my opinions and decisions, and his permission to pursue the scientific questions and skills I was interested in. The most telling aspect of Bob's mentorship is his desire to see his students succeed both in and outside of lab, often at the expense of his own interests. This quality was evident throughout my career at UNC and especially as I decided on my next career step.

I would also like to thank the members of my committee for their direction throughout many lab, histone group, and committee meetings. I thank Brian for sharing his passion for chromatin and a plethora of ideas never limited by perceived feasibility. Greg, thank you for instilling in me a meticulous attention to detail and being a great additional mentor. Jeff is one of the few Packers fans whose opinions I respect; I have appreciated the many suggestions you have given me. Finally, I am obliged to Dan for imparting many of his technical skills. Thank you for sharing many enjoyable scientific discussions and back of the fly box genetic schemes.

I am also grateful to have shared my graduate school experience with many wonderful lab and pod members. I have been blessed to work alongside many talented

scientists in the Matera, McKay, and Sekelsky labs. In particular, I would like to thank Mike Meers for establishing much of the scientific foundation that has facilitated what success I have had and for being a constant computational resource. Everyone in the Duronio lab has enriched my experience here, and although there are too many to name, I would like to thank Joy, Deirdre, and Este for going out of their way to welcome me to the lab and Mary for running the lab so efficiently. Finally, I would like to acknowledge Robin Armstrong as one of the most capable and enthusiastic scientists I have known. It has been a true pleasure to collaborate with you.

Outside of lab I have been supported by great friends both in Carolina and back in Minnesota. At the risk of leaving someone out, I won't name each of you, but I hope you all know I have cherished your friendship. As we continue to go separate ways, I hope we can remain in touch. Special thanks go out to Michael Bertucci for introducing me to the volleyball community at UNC and for remaining a close friend. I am also very grateful to have shared the adventure of moving to North Carolina with my former roommate and old friend Nathan Rodeberg. You made the transition to graduate school not only bearable but enjoyable, and I value the many experiences we have shared since.

Finally, I am most thankful for the support of my family. To David and Nanci, I am incredibly grateful for the loving and welcoming environment you have shared with me. Charlotte has been and continues to be a home away from home. To my father and mother, I thank you for instilling in me the work ethic and ambition that I have heavily relied on throughout graduate school. Knowing I always have someone to fall back on has given me the confidence to leave home and push myself personally and

professionally. To my sisters Dani and Grace, I hope I have been a good role-model even while living away. I now see in you both qualities that I myself look up to. To my wife and best friend Kelly, I thank you for encouraging me through every hardship and celebrating every triumph of graduate school. You have improved all aspects of my life, and I look forward to every adversity and accomplishment that the future brings us.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ac	Acetyl
ATRX	α-Thalassemia/mental retardation syndrome X-linked protein
BAC	Bacterial Artificial Chromosome
bp	Base pair
CID	Centromere identifier protein
ChIP	Chromatin immunoprecipitation
CUT&RUN	Cleavage under targets and release using nuclease
DAPI	4’,6-Diamidino-2-phenylindole
DAXX	Death domain-associated protein
DNA	Deoxyribonucleic acid
ENCODE	Encyclopedia of DNA elements
FAIRE	Formaldehyde assisted isolation of regulatory elements
GFP	Green fluorescent protein
H3K9	Ninth lysine of histone H3
H3K9R	Lysine to arginine mutation at the ninth lysine of histone H3
His4r	Histone H4 replacement
HisC	*Df(2L)HisC*^{ED1429}
HP1	Heterochromatin Protein 1
HWT	Histone wild-type
kb	Kilobase pair
me	Methyl
MNase	Micrococcal nuclease

PCNA	Proliferating cell nuclear antigen
PTM	Post-translational modification
RFP	Red fluorescent protein
RNA	Ribonucleic acid
RPKM	Reads per kilobase of transcript per million mapped reads
RPM	Reads per million
RT	Replication timing
SuUR	Suppressor of under-replication
YFP	Yellow fluorescent proteins

CHAPTER 1 – INTRODUCTION[1]

Chromatin regulation of DNA-dependent processes

The DNA from a single human cell laid end to end measures 2 meters long and must fit inside a nucleus with a diameter of six micrometers (Alberts et al. 2002). To accomplish this incredible feat, DNA in eukaryotes is organized and compacted through interaction with histone proteins in a structure defined as chromatin. The nucleosome is the fundamental unit of chromatin and consists of 147 base pairs of DNA wrapped around an octamer of histone proteins (Luger et al. 1997). Each octamer consists of two copies of the core histones: H2A, H2B, H3, and H4. In addition to the core histones, chromatin consists of the linker histone H1, as well as a host of other DNA-associated proteins.

Although DNA must be organized within the nucleus, many essential processes within the cell require access to DNA, necessitating a balance between DNA accessibility and compaction. Thus, chromatin influences many aspects of cell biology, including DNA replication, DNA repair, transcription, and RNA processing. For example, euchromatin, generally regarded as “open” chromatin, can be more accessible to proteins which initiate processes like DNA replication or transcription, and is therefore associated with activation of different processes. In contrast, heterochromatin, typically

[1] Part of this chapter originally appears as an article in Developmental Cell. The original citation is as follows: McKay DJ, Klusza S, Penke TJR, Meers MP, Curry KP, McDaniel SL, Malek PY, Cooper SW, Tatomer DC, Lieb JD, et al. 2015. Interrogating the Function of Metazoan Histones using Engineered Gene Clusters. *Dev Cell* **32**: 373–386.

described as “closed” chromatin, may physically prevent interaction of initiator proteins with DNA and is thus considered inhibitory to many processes. In addition to chromatin accessibility, the type of chromatin-associated proteins in a genomic region can influence the recruitment of other trans-acting factors. These factors may themselves regulate cellular processes and can also control chromatin architecture. Therefore, chromatin accessibility and the types of proteins associated with chromatin are interdependent, and coordination of both is required for proper regulation of DNA-dependent processes.

Several epigenetic mechanisms are employed to ensure robust coordination of chromatin structure and composition, including DNA methylation, histone post-translational modifications (PTMs), and the type of histone incorporated onto DNA; this work will focus on the latter two mechanisms. Histones are broadly categorized into two groups, replication-dependent and replication-independent. Replication-dependent or canonical histone genes are present in multiple copies in the genome, are synthesized during S-phase, and are deposited onto DNA by the histone chaperone CAF1 (Marzluff et al. 2002; Tagami et al. 2004; Verreault et al. 1996). In contrast, replication-independent or variant histones are single copy genes, are expressed throughout the cell cycle, and are incorporated onto DNA by HIRA or the ATRX-DAXX complex (Henikoff and Ahmad 2005; Tagami et al. 2004; Szenker et al. 2011). The presence of variant histones can have varying consequences including altered stability of the nucleosome (Jin and Felsenfeld 2007; Xu et al. 2010) or recruitment of different trans-acting proteins. For instance, CID, a histone H3 variant, establishes centromere identity and is required for the recruitment of kinetochore proteins (Blower and Karpen 2001;

Henikoff and Ahmad 2005; Mellone and Allshire 2003). H2AX and H2AZ, variant versions of the canonical H2A, play critical roles in DNA repair (Scully and Xie 2013; Price and Andrea 2014). The type of histones that compose a nucleosome and that are deposited onto DNA can therefore play an important role in regulating chromatin structure and function. Additionally, histones can be chemically modified at a variety of histone residues by attachment of several different chemical groups, including acetyl, methyl, phosphoryl, and many others (Strahl and Allis 2000). This process, referred to as histone post-translational modification (PTM), is an important means of chromatin regulation, which it can accomplish by two mechanisms. First, chemical modification of histones can adjust the charge of histones, potentially impacting chromatin accessibility. For example, the interaction between negatively charged DNA and positively charged histone proteins can be weakened by adding negatively charged modifications to histones to reduce their positive charge. Negatively charged chemical modifications such as acetyl groups are therefore associated with increased chromatin accessibility. Secondly, histone PTMs can serve as binding sites for proteins that specifically bind to modified histones. Recruitment of the initial protein can initiate binding of a host of other trans-acting factors that can alter chromatin structure and/or function. The suite of histone PTMs on a nucleosome (or range of nucleosomes) is frequently compared to a language instructing other proteins to establish both chromatin architecture and the collection of proteins bound to a particular region of chromatin (Strahl and Allis 2000). This language is referred to as the “histone code”.

Over the past twenty years a plethora of research has taken place investigating how the histone code is established and read. Massive collaborative efforts like the

ENCODE project have, as part of their mission, characterized the genomic distributions of many histone PTMs in humans as well as various model organisms (ENCODE Project Consortium et al. 2007; Ho et al. 2014; Roy et al. 2010). These studies have generated strong correlations between particular histone PTMs and different cellular processes, and, importantly, they have generated predictions for how histone PTMs could regulate fundamental processes such as DNA replication, transcription, and nuclear organization. Most of the functional studies aimed at testing these correlations have relied on mutation of the enzymes that add, remove, or bind to these modifications, also referred to as writers, erasers, and readers, respectively. These studies have been instrumental in elucidating the regulation and function of the histone code. However, the ability of histone modifying enzymes to target non-histone substrates as well as multiple histone residues has been increasingly recognized (Huang and Berger 2008; Sims and Reinberg 2008). Writer or eraser loss-of-function studies are therefore complicated by potential pleiotropic effects due to altered modification states on non-histone substrates in addition to the histone modification of interest. A more direct test of histone modification function is to mutate the target histone residue of interest to a non-modifiable amino acid. This approach has been successfully used in yeast for over twenty years to investigate histone residue function. In metazoans, however, histone gene engineering is complicated by the repetitive nature of histone genes. For example, histone genes in humans (~55 copies) are present at two genomic loci with essential genes located between histone repeats (Marzluff et al. 2002). Currently, histone mutagenesis in most metazoans would be extraordinarily difficult, even using CRISPR-Cas9 based genome engineering. In

Drosophila melanogaster, however, the histone genes are present at a single genomic locus (Lifton et al. 1978), permitting site specific deletion of the endogenous histone genes and replacement with an array of transgenic histone genes (Günesdogan et al. 2010; McKay et al. 2015). Comparing an array of wild-type histone genes to an array containing a point mutation at a specific histone residue enables functional interrogation of the residue. This histone replacement platform is described in detail below.

A histone gene replacement platform enables metazoan histone genetics

In *Drosophila* melanogaster each of the four core histones genes, H2A, H2B, H3, and H4, along with the linker histone H1, are organized in a 5 kb cluster which is tandemly repeated ~100 times (Lifton et al. 1978) (Figure 1.1). This entire locus was site specifically removed by Günesdogan et al. (2010) using the DrosDel system (Ryder et al. 2004) generating the ΔHisC allele. The histone deletion can be maintained in a heterozygous state using a balancer chromosome which prevents recombination on the chromosome containing the ΔHisC allele. Flies homozygous for the histone deletion are embryonic lethal ceasing development at cell cycle 15 after the maternal load of histones is depleted (Günesdogan et al. 2010; McKay et al. 2015). This lethality can be rescued by integration of one large (McKay et al. 2015) or multiple smaller (Günesdogan et al. 2010) arrays of transgenic histone clusters. Our platform, described in McKay et al. (2015), uses restriction digest cloning to multimerize twelve copies of the 5 kb histone repeat in a Bacterial Artificial Chromosome (BAC) (Figure 1.1). This construct is subsequently injected into *Drosophila* embryos and site specifically integrated into the genome at an attP site located on chromosome 3. The ectopic array of histone genes can then be used to rescue endogenous histone deletion mutants

(Figure 1.2). Although more difficult to integrate than the multiple smaller histone arrays used by (Günesdogan et al. 2010; Hödl and Basler 2012; Pengelly et al. 2013), the single histone array more faithfully mimics the continuous, repetitive nature of the endogenous array and provides a genetically more facile system (see Chapter 3). A detailed investigation of histone copy number and gene expression using this system can be found in Appendix I. These experiments establish the histone replacement platform as a suitable approach to engineering histone genes, as an ectopic array of 12 histone repeats can rescue both the viability, fertility, and histone protein level of endogenous histone deletion mutants. This foundation allows us to engineer various histone mutations to prevent histone PTM at any residue and study its specific contribution to chromatin structure and function. Using this platform, we thoroughly characterize the role of the ninth lysine of H3 (H3K9) in heterochromatin biology. Before an experimental description of H3K9 function, the reader might first benefit from a historical introduction into heterochromatin and its link to H3K9.

Figure 1.1 Schematic of the histone replacement platform. Top of figure indicates the endogenous histone locus on chromosome 2L. ΔHisC indicates molecular defined deletion of endogenous histone genes. Black arrowheads signify a 5 kb histone cluster composed of one copy of H1, H2B, H2A, H4, and H3. Tandem array of arrows indicates a histone cluster that has been multimerized in pMultiBAC, a Bacterial Artificial Chromosome. Components of pMultiBAC include a Multiple Cloning Site (MCS), an attB site for site-specific integration into the *Drosophila* genome, and a mini white cassette (w+) used for screening transgene integration events.

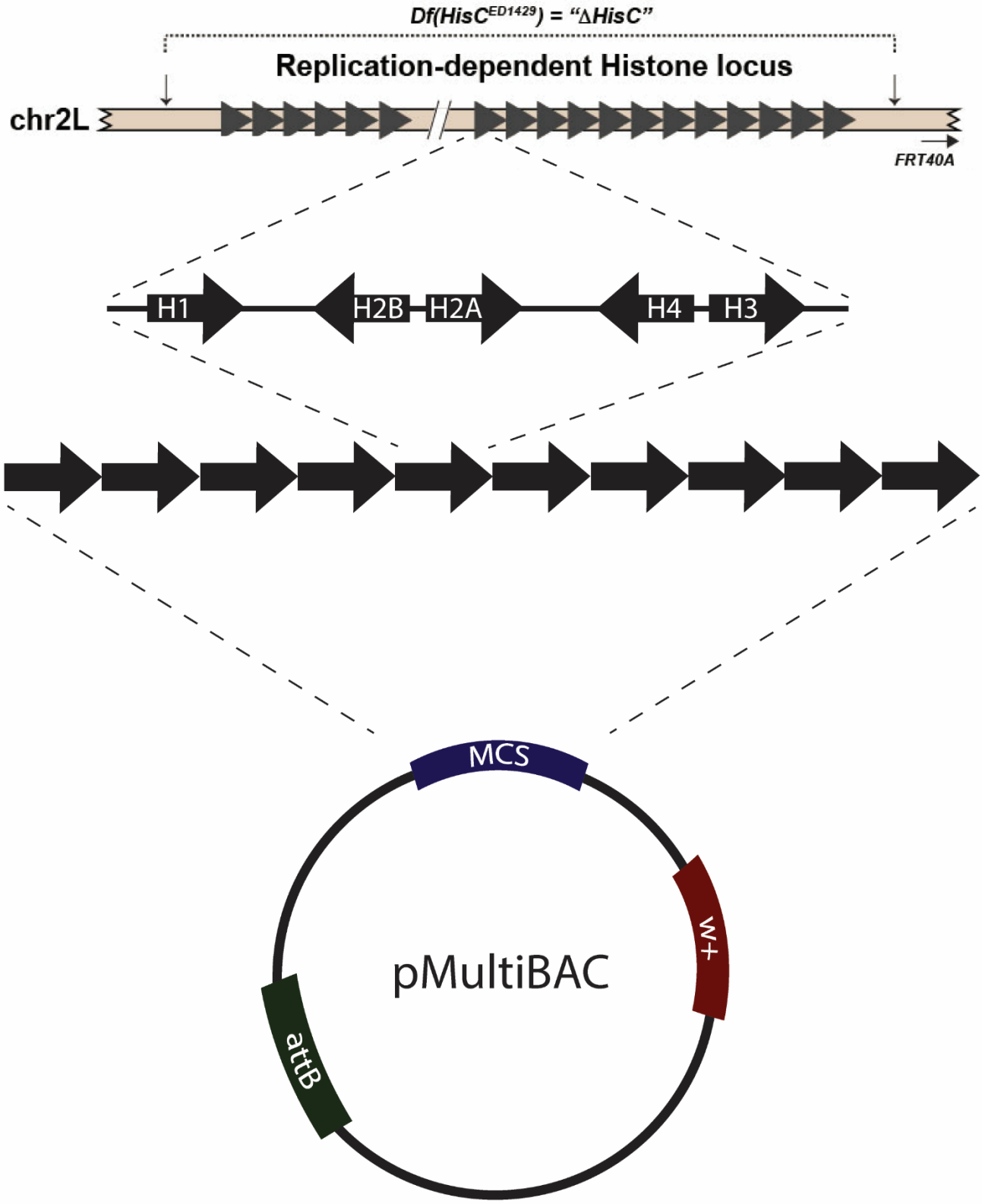
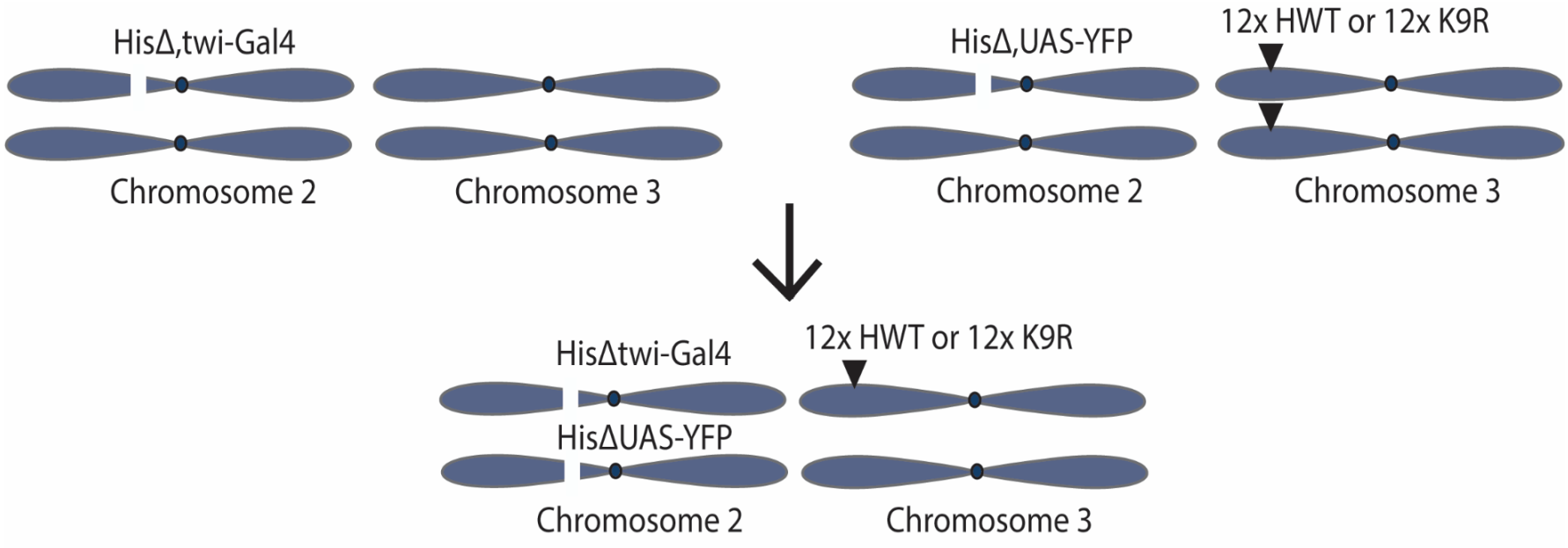

Df(HisC^ED1429) = "ΔHisC"
Replication-dependent Histone locus
chr2L
FRT40A
H1
H2B
H2A
H4
H3
MCS
pMultiBAC
w+
attB

Figure 1.2 Illustration of cross to generate 12x HWT/K9R animals. Animals heterozygous for the ΔHisC allele on chromosome 2L are mated; therefore, 25% of their progeny are homozygous for the endogenous histone deletion. The GAL4/UAS system was employed to identify homozygous ΔHisC progeny. Recombinant chromosomes were generated containing the ΔHisC and either twi-Gal4 or UAS-YFP (twi = mesoderm specific twist promoter; YFP = Yellow Fluorescent Protein). If both ΔHisC, twi-Gal4 and ΔHisC, UAS-YFP chromosomes are present in an animal, Gal4 can bind the UAS sequence and drive expression of YFP. Fluorescent animals can then be identified using an epifluorescent microscope. Additionally, one parent in this cross is homozygous for the ectopic rescue transgene ensuring all progeny receive one copy of the transgenic histone array. Histone wild-type (HWT) indicates wild-type array of histone genes while H3K9R indicates array where each lysine 9 of H3 contains a lysine-to-arginine point mutation.


HisΔ,twi-Gal4
Chromosome 2
Chromosome 3
HisΔ,UAS-YFP
12x HWT or 12x K9R
Chromosome 2
Chromosome 3
HisΔtwi-Gal4
12x HWT or 12x K9R
HisΔUAS-YFP
Chromosome 2
Chromosome 3

Historical perspective of heterochromatin

Heterochromatin was initially characterized by the German scientist Emil Heitz. Through development of a new in situ method for chromosome staining, Heitz observed that portions of chromosomes remained intensely staining throughout the cell cycle, naming these regions heterochromatin (Heitz 1928; Passarge 1979). He recognized "longitudinal differences" in chromosome structure, contrasting heterochromatin with what he called euchromatin, regions of chromosomes that unraveled after mitosis (Heitz 1929; Passarge 1979; Brown 1966). Heitz further concluded that regions of heterochromatin could become euchromatin and vice versa, foreshadowing the existence of different types of heterochromatin. In contrast to constitutive heterochromatin, regions of facultative heterochromatin may exist in distinct chromatin states in different cell types or across development. Before DNA was identified as the genetic material, Heitz also perceived an important relationship between gene activity and chromatin, arguing that heterochromatin contained lower gene density than euchromatin (Heitz and Bauer 1933; Passarge 1979; Brown 1966). Although he incorrectly considered heterochromatin genetically inert, he was correct that heterochromatin is generally a repressive transcriptional environment.

Since Heitz's work, heterochromatin was observed throughout the animal and plant kingdoms (Brown 1966; Swanson 1957). In particular, heterochromatin was found near centromeres and at the ends of chromosomes in many different species. Moreover, through analysis of the giant salivary gland chromosomes in *Drosophila*, as well as several additional species, specific regions of heterochromatin were found to coalesce into a structure termed the chromocenter (Brown 1966). In contrast, other

regions of heterochromatin were interspersed with euchromatin. This pattern was exemplified by the band and interband pattern of DNA-stained *Drosophila* polytene chromosomes.

In addition to structural characteristics of heterochromatin, functional roles were proposed and identified. Heterochromatin was observed to undergo low rates of meiotic crossing over, leading to the tendency of heterochromatic regions to be inherited together (Brown 1966; Hannah 1951; Roberts 1965). Contrastingly, somatic crossing over events were enriched in heterochromatin (Walen 1964). Although the mechanism of chromosome exchange was not understood at this point, heterochromatin was implicated in the regulation of these processes. Additionally, heterochromatic DNA was observed to be late replicating in S phase in a number of species (Taylor 1958, 1960a; Brown 1966) and underreplicated in polytene nuclei (Hsu et al. 1964).

Other investigations began to elucidate the role of chromatin in regulating gene expression through a number of findings. When genes located in euchromatin were moved adjacent to regions of heterochromatin through chromosomal inversions or transposition, these genes were repressed in a mosaic pattern, a phenomenon referred to a position-effect variegation (PEV) (Lewis 1950; Brown 1966; Muller 1930). Studies of PEV were among the first to suggest direct connections between heterochromatin and its now well-established role in gene repression. However, though Heitz originally considered heterochromatin inert, the fact that genes can be expressed from within heterochromatin was appreciated early on. For example, several genes were discovered within heterochromatin, and genes located on the heterochromatic Y-chromosome were identified as essential for male fertility (Hannah 1951; Khush et al.

1964). When genes that normally resided within heterochromatin were moved into euchromatic regions by chromosomal inversions they exhibited reduced expression, suggesting genes were adapted to express at proper levels within their original chromatin environment. Furthermore, early studies of “breakage-fusion-bridge” cycles by McClintock, and follow up work by Brink, suggested a connection between inappropriate heterochromatin structure during a particular developmental stage and mistimed expression of developmentally regulated genes. (McClinktock 1951; Brink 1964). McClintock’s observation that altered heterochromatin led to “mutable loci” correctly predicted the now well-known role of heterochromatin in the repression of transposon expression and their mobilization (McClinktock 1951).

Numerous other findings in subsequent years linked chromatin and transcription. Several publications reported that RNA metabolism is largely absent during the chromosome condensation that accompanies mitosis (Taylor 1960b; Prescott and Bender 1962; Konrad 1963). Moreover, the transcriptional differences between puffed and unpuffed bands in the chromosomes of diptera salivary glands were compared to the gene activity of euchromatin and heterochromatin (Brown 1966; Beermann 1963). Upon puffing the chromosomal bands enlarge and become indistinct analogous to the structural differences between the euchromatic interbands and the heterochromatic bands of polytene chromosomes. Puffs were identified as sites of highly active transcription drawing an additional connection between chromosome structure and function (Beermann 1963). At this time the role of histone proteins, their interaction with DNA, and their role in regulating gene expression was also under consideration (Huang et al. 1964; Allfrey et al. 1964).

Finally, comparison of heterochromatin formation in different sexes pointed towards a functional role for heterochromatin in regulating gene expression. In mealybugs the paternal set of chromosomes becomes heterochromatinized and is transcriptionally silent (Brown and Nur 1964; Brown 1966). Through analysis of feline neurons, Barr identified a heterochromatic component present in female nuclei but not males. This component, subsequently named the Barr body, was identified as one of the X chromosomes in female mammals that is randomly chosen to become heterochromatic and genetically inert (Brown 1966). Clear correlations between heterochromatin structure and the regulation of many DNA-dependent functions were established during this early seminal work; however, an understanding of how heterochromatin is established and mechanisms of its function were not yet elucidated.

Much of our understanding of how heterochromatin is established was derived from studies of PEV in *Drosophila*. A classic example of PEV, discovered by Muller in 1930, is a chromosomal inversion that juxtaposes the white gene next to pericentromeric heterochromatin (Figure 1.3) (Muller 1930). When near heterochromatin, the white gene, which is involved in red pigmentation of the eye, is silenced in a variegated manner. This variegated silencing results in a red and white mottled eye phenotype. Early work posited that variable spreading of heterochromatin into the white gene caused inconsistent gene repression (Demerec and Slizynska 1937). This hypothesis was supported by cytological examination of heterochromatin in animals containing the white inversion (Hartmann-Goldstein 1967); however, direct evidence that heterochromatin functioned to repress gene expression or an understanding of how it did so was lacking.

Fifty years after Muller's publication this ignorance began to be rectified. Analysis of environmental or genetics factors that altered PEV phenotypes uncovered many of the molecular regulators of PEV. The discoveries that reduced histone gene dosage (Moore et al. 1979; Khesin and Leibovitch 1978) and hyper-acetylation of histones (Mottus et al. 1980) suppressed PEV suggested histones regulated PEV. Additionally, several groups used this paradigm, as well as other examples of PEV, to screen for genetic enhancers and suppressors of PEV (Sinclair et al. 1983; Reuter and Wolff 1981; Locke et al. 1988; Henikoff 1990, 1979). Flies with the white inversion were randomly mutagenized, and mutations that enhanced (white eye) or suppressed (red eye) PEV were isolated and mapped. Many of these mutations were found in chromatin-associated proteins that localize to regions of heterochromatin. For example, the Elgin lab identified that mutation of the gene encoding Heterochromatin Protein 1a (HP1a) resulted in suppression of PEV, implicating HP1 in establishing heterochromatin (Eissenberg et al. 1990; James and Elgin 1986). Furthermore, altered chromatin structure, measure by micrococcal nuclease digests, was associated with PEV suggesting chromatin packaging regulates PEV (Wallrath and Elgin 1995). In total, about 150 genes involved in suppressing or enhancing PEV were identified through these screens (Schotta et al. 2003a).

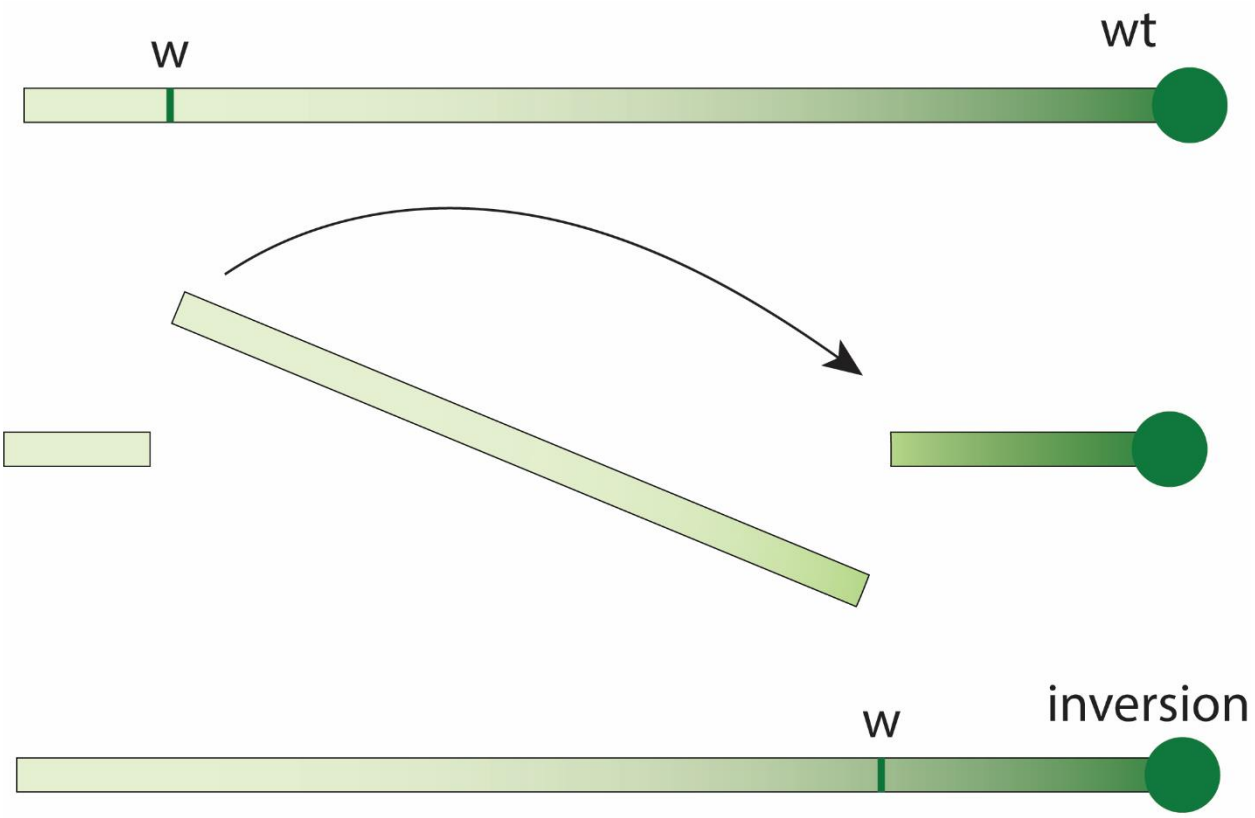
Many of the identified genes encoded proteins that post-translationally modify histones. A key example is Su(var)3-9, which was identified as an important contributor to heterochromatin gene silencing (Tschiersch et al. 1994). The human homolog of Su(var)3-9 (SUV39H1) was found to methylate the ninth lysine of histone H3 (H3K9) (Rea et al. 2000). In *Drosophila,* Su(var)3-9 along with two other H3K9

methyltransferases, Setdb1 and dG9a, are responsible for establishing mono-, di-, and tri-methylation of H3K9 (Schotta et al. 2002; Ebert et al. 2004; Brower-Toland et al. 2009). Methylated H3K9 serves as a binding site for HP1 (Bannister et al. 2001; Lachner et al. 2001; Nakayama et al. 2001) which can in turn recruit the Su(var)3-9 methyltransferase (Schotta et al. 2002; Eskeland et al. 2007). Methylation of H3K9 on a neighboring nucleosome can reinitiate the cycle providing a mechanism for the formation and spreading of heterochromatin.

H3K9me function has largely been investigated by mutation (Elgin and Reuter 2013; Wallrath et al. 2014) or sequestration (Herz et al. 2014) of the H3K9 methyltransferases or loss of the H3K9me reader HP1. These studies have provided strong evidence for H3K9me function in a number of important cellular functions including: centromere identity, kinetochore formation, and chromosome segregation (Allshire et al. 1995; Bernard et al. 2001; Mellone and Allshire 2003; Durand-Dubief and Ekwall 2008); nuclear organization (Csink and Henikoff 1996; Dernburg et al. 1996; Wit and Laat 2012); transcriptional silencing and activation (Huisinga et al. 2006; Yasuhara and Wakimoto 2006; Ebert et al. 2006); regulation of small RNAs (Grewal and Jia 2007; Pal-Bhadra et al. 2004); recombination (Jia et al. 2004); and genome integrity (Peng and Karpen 2008, 2009). Although these investigations have been fundamental to our understanding of heterochromatin biology, they are limited in three ways. First, most organisms have multiple H3K9 methyltransferases and removal of all enzymes can be genetically complicated. Second, HP1a is known to bind chromatin independently of H3K9 methylation (Figueiredo et al. 2012; Raffa et al. 2011); therefore, observed phenotypes in HP1 loss-of-function experiments may be due to H3K9-independent roles

for HP1. Finally, H3K9 methyltransferases can modify other histone residues as well as non-histone substrates, complicating analyses of writer mutation or sequestration experiments. A solution to all of these issues would be to mutate H3K9 to a non-modifiable residue. The remaining chapters describe the generation of an H3K9R mutant in *Drosophila* using the histone replacement platform and a thorough analysis of H3K9's contribution to chromatin regulated processes.

Figure 1.3 Position-effect variegation of white gene results in mottled eye. Top chromosome represents wild-type chromosome with white gene (w) in euchromatic region of chromosome arm. X-ray induced inversion of chromosome arm results in juxtaposition near pericentromeric heterochromatin indicated by darker color. Left image of *Drosophila* eye represents wild-type coloring. Right three eyes contain the white inversion and mottled eye phenotype.


w
wt
w
inversion


wt
inversion

CHAPTER 2 – DIRECT INTERROGATION OF THE ROLE OF H3K9 IN METAZOAN HETEROCHROMATIN FUNCTION[2]

Introduction

The eukaryotic genome is organized and compacted within the nucleus through interaction with histones and other proteins to form chromatin. Cytological observation originally divided chromatin into two subgroups, euchromatin and heterochromatin. Euchromatin is gene rich, transcriptionally active, and usually described as open or accessible. Conversely, heterochromatin is gene poor and generally considered transcriptionally repressive and inaccessible. There are two major types of heterochromatin, facultative and constitutive, that each have important cellular functions. The accessibility of facultative heterochromatin is regulated in order to control gene expression (Grewal and Jia 2007). By contrast, constitutive heterochromatin remains condensed throughout the cell cycle and functions in the repression of inappropriate recombination, transposons, and developmentally important genes (Grewal and Jia 2007; Peng and Karpen 2008; Eissenberg and Elgin 2014; Yasuhara and Wakimoto 2006). Despite its generally repressive role, a heterochromatic configuration is also required for expression of certain genes that reside within heterochromatin. Additionally, constitutive heterochromatin is implicated in the structural

[2] This chapter previously appeared as an article in Genes & Development. The original citation is as follows: Penke TJR, McKay DJ, Strahl BD, Gregory Matera A, Duronio RJ. 2016. Direct interrogation of the role of H3K9 in metazoan heterochromatin function. *Genes Dev* **30**: 1866–1880.

integrity of centromeres and therefore the promotion of faithful chromosome segregation during cell division (Kellum and Alberts 1995; Bernard et al. 2001). Given these critical functions, heterochromatin is generally considered essential for development.

A large body of evidence from fission yeast to humans has concluded that heterochromatin formation and function is accomplished by post-translational modification (PTM) of histones, particularly di- and tri-methylation (me2/me3) of H3K9. H3K9me2/me3 serves as a binding site for HP1 (Bannister et al., 2001; Lachner et al., 2001; Nakayama, 2001), which is proposed to mediate chromatin condensation via multimerization of HP1 molecules on nearby nucleosomes (Canzio et al. 2011; Azzaz et al. 2014). HP1 multimers also serve as a scaffold to recruit various other chromatin remodelers associated with condensed chromatin, including H3K9 methyltransferases (Grewal and Jia 2007; Elgin and Reuter 2013). This compacted state is thought to prevent or limit access of proteins to DNA to achieve the repressive functions of heterochromatin. In support of this model, tethering HP1 is sufficient to render chromatin less accessible to nucleases (Danzer and Wallrath 2004). Thus, methylation of H3K9 is thought to serve as the foundation for constitutive heterochromatin formation and a repressive chromatin environment.

Determination of H3K9me's role in heterochromatin formation has primarily relied on mutations of the enzymes that methylate H3K9. Although these studies have greatly enhanced our understanding of heterochromatin biology, they are limited in two ways. First, animals contain multiple H3K9 methyltransferases, and at least three different, partially redundant, enzymes methylate H3K9 in *Drosophila* (Elgin and Reuter 2013). Analyzing phenotypes in single or double mutants is therefore complicated by genetic

compensation. Second, enzymes that catalyze histone PTMs often have numerous non-histone substrates, confounding analysis of the biological contribution of a given PTM (Sims and Reinberg 2008; Huang and Berger 2008; Biggar and Li 2014; Zhang et al. 2015). For example, the fission yeast H3K9 methyltransferase Clr4 modifies the non-histone substrate Mlo3 to facilitate centromeric siRNA production, a process linked to heterochromatin formation (Zhang et al. 2011; Gerace et al. 2010). Preventing methylation of Mlo3 leads to a reduction in centromeric siRNAs, whereas preventing methylation of H3K9 via a K9R mutation does not. These findings underscore the importance of examining mutations in histone residues that leave other functions of histone-modifying enzymes intact. Until recently, the repetitive nature of histone genes in metazoans has prevented histone gene engineering. Consequently, despite being one of the most well-studied histone modifications, we do not know which aspects of heterochromatin structure and function require H3K9.

Recently, we and others have developed a method for functional replacement of replication-dependent histone genes in *Drosophila* (Günesdogan et al. 2010; McKay et al. 2015). To address the role of H3K9 in heterochromatin biology, we engineered flies that express only H3K9R mutant replication-dependent histones. In contrast to expectations, we find that H3K9R mutants express a relatively normal protein-coding transcriptome and can complete development, albeit with greatly reduced frequency compared to controls. However, we find that nucleosomes and HP1a are depleted from pericentromeric heterochromatin in H3K9R mutants, with HP1a redistributing along the largely euchromatic chromosome arms. We also find that most transposon families are

de-repressed, resulting in their mobilization. We propose that unrestricted transposition contributes to the reduced viability of H3K9R mutants.

Materials and Methods

Culture conditions

All stocks were maintained on standard corn media. For cross scheme see Figure 2.2A. 50 Δ*HisC, twi-GAL4/CyO* females and 20 Δ*HisC, UAS-2xEYFP/CyO ; HWT/HWT or K9R/K9R* males were placed in a cage at 25C and allowed to lay eggs on a grape juice agar plate. To measure the completion of embryogenesis, GFP positive eggs from a 4 hour collection were moved to a separate plate and aged 24 hours prior to counting hatching. The number of hatched eggs (observed) and the total number of eggs scored are indicated in Table 2.1. For all other developmental assays overnight collections were used. To measure the completion of development from egg hatching to adult eclosion, ~50 GFP positive larvae were moved to a corn media vial 48 hours after egg laying to separate Δ*HisC, UAS-2xEYFP* / Δ*HisC, twi-GAL4* mutants from their siblings. For each group of ~50 larvae the number that pupated and the number that eclosed as adults was determined and summed, as was the total number of larvae scored. An identical procedure was used to follow development of randomly selected *yw* (i.e. essentially wild type) progeny. Expected values for Chi-squared tests are based on observed value of *yw* animals. For genomic or molecular analyses of salivary gland, wing disc, or whole larvae samples, culture vials were cleared of wandering third instar larvae, and after 4-6 hours newly wandering larvae were selected. HWT and K9R wandering third instar larvae contained similarly sized wing discs as measured by

expression of the Wingless morphogen (data not shown). K9R larvae took 1-2 days longer than HWT to reach wandering third instar stage. In addition, culturing K9R mutants in isolation is necessary for development to the third larval instar stage because the survival of K9R animals is severely reduced when co-cultured with heterozygous siblings.

Immunofluorescence

Mitotic recombination experiments and salivary gland polytene chromosome preparations were performed as previously described (McKay et al. 2015; Cai et al. 2010). Salivary glands from HWT and K9R genotypes were squashed and stained on the same slide to control for variations in individual preparations. Classifications in Figure 2.3B were performed blindly before identifying genotype. Only chromocenters that could be unambiguously identified by the convergence of chromosome arms were scored. K9R mutant nuclei were identified by lack of anti-K9me2 (Abcam 1220) antibody staining. K9me3 and HP1a were stained with anti-K9me3 (Active Motif 39161) and anti-HP1a (DSHB C1A9) respectively. Images shown are single confocal images taken at a constant gain on a Leica TCS SP5 AOBS UV/spectral confocal laser-scanning system mounted on an inverted DM IRE2 microscope. *Gypsy* mobilization assay was performed with adaptations from Li et al. (2013). *w, ovo ; ΔHisC,UAS-2xEYFP/CyO, Act-GFP* females were crossed to either *ΔHisC, twi-GAL4/CyO, Act-GFP*; *HWT/HWT* or *K9R/K9R* males. Embryos were collected on grape juice plates overnight and aged 36-48 hours. 50 GFP negative larvae were moved to a corn media vial and allowed to wander or pupate. Gal80 from the *w, ovo* chromosome represses fluorescence in *ΔHisC, twi-GAL4/ΔHisC, UAS-2xEYFP* progeny but does not repress the distinct

fluorescent pattern from the *CyO, Act-GFP* chromosome because *Act-GFP* expression is not Gal4 dependent. The number of larvae or pupae with 0, 1, 2, or ≥3 YFP positive clones was then counted to assay cells in which Gal80 repression of *twi-GAL4* driven YFP is disrupted.

Sample Preparation for Sequencing

FAIRE-seq and RNA-seq samples from 3rd instar imaginal discs were prepared as previously described (McKay and Lieb 2013). HP1a ChIP-seq samples from whole 3rd instar larvae were prepared essentially as described (Soruco et al., 2013), with an equal amount of 3rd instar larval *Drosophila virilis* chromatin added to each HWT and K9R replicate before immunoprecipitation. Libraries were prepared with the Tru Seq DNA Kit (FAIRE), Total RNA TruSeq-Stranded Ribo-Zero Gold Kit (RNA), and the ThruPLEX DNA-seq Kit (ChIP). Sequencing was performed on an Illumina HiSeq2500.

Sequence Data Analysis

FAIRE-seq, and ChIP-seq samples were aligned to the dm6 reference genome (Release 6.04) using Bowtie2 default parameters (Langmead and Salzberg 2012). Analysis of only uniquely mapping reads (MAPQ≥10) provided similar results to those reported here. FAIRE peaks were called using MACS2 with a shift size of 110bp and a stringency cutoff of 0.01 (Zhang et al. 2008). FAIRE-seq signal at FAIRE peaks was normalized to sequencing depth. Differential signal analysis was performed with edgeR (Robinson et al. 2009). HP1a ChIP-seq signal within 1kb windows was normalized to the number of uniquely mapping *D. virilis* reads. For ChIP analysis windows were used instead of peaks due to broad redistribution of HP1a that precluded peak calling along chromosome arms. RNA-seq reads were aligned to the dm6 reference genome using

TopHat and assembled into transcripts with Cufflinks (Trapnell et al. 2014). DESeq2 was used for differential expression analysis for both RNA and HP1a ChIP datsets (Love et al. 2014b). Data was visualized using the Integrative Genomics Viewer (Robinson et al. 2011). The following modENCODE 3rd instar larval ChIP-seq data sets were used: K9ac=GSE48510, K27me3=GSE49490, K9me2=GSE47260, K9me3=GSE47258, HP1a=GSE47243. Peaks were called using the above parameters except broad peak calling was implemented for K9me2, K9me3, and HP1a datasets. Analysis of transposons and piRNA clusters was performed with piPipes using the dm3 reference genome (Han et al. 2015).

Results

H3K9 is important, but not essential, for completion of Drosophila development

To test the role of H3K9 in *Drosophila* development and chromatin architecture we generated lysine to arginine mutations at H3K9 using the histone replacement system we recently developed (McKay et al. 2015). In brief, the replication-dependent histones genes in *Drosophila melanogaster* reside at a single locus that contains ~100 copies of a tandemly repeated gene cluster *(HisC)*. Each histone repeat unit contains one copy of a gene encoding histone H1, H2A, H2B, H3, and H4. This arrangement enables functional rescue of a *HisC* deletion with a single BAC-based transgene containing a synthetic tandem array of 12 histone repeat units (Figure 1.2). *HisC* deletion animals containing a histone wild-type or H3K9R transgenic array are hereafter referred to as HWT or K9R, respectively. Importantly, *Drosophila* has two partially redundant histone H3 variants, H3.3A and H3.3B, encoded by single copy genes

located outside the *HisC* gene cluster (Henikoff and Ahmad 2005). H3.3 is enriched at active regions of the genome, contains modifications associated with euchromatin (Hake et al. 2006), and in flies is depleted from heterochromatin (Ahmad and Henikoff 2002). We therefore used the K9R mutant to investigate heterochromatin biology, and focused our analyses on functions of replication-dependent H3K9.

We first characterized the developmental consequences of the K9R mutation. Surprisingly, we found that K9R flies occasionally survive to adulthood (~2%), demonstrating replication-dependent H3K9 is not absolutely essential for completion of development (Table 2.1). However, most K9R mutant progeny are developmentally delayed by 1-2 days and display a broad lethal phase, dying during larval development or pupation. This developmental delay prompted us to examine the proliferative capabilities of K9R cells by generating mitotic clones via the FLP-FRT system (Xu and Rubin 1993). Clones of *HisC* deletion cells containing an HWT transgene were equivalent in size to control clones containing endogenous histones (Figure 2.1A,B). In contrast, K9R clones were approximately 2/3 the size of control clones, suggesting a modest growth defect. Western blots of whole 3rd instar HWT and K9R larvae revealed similar amounts of H3 (Figure 2.2B) but substantially decreased H3K9me2 and me3 in K9R mutants (~20 fold relative to HWT) (Fig. 2.1C). However, some residual H3K9me2/me3 signal remains in K9R mutants, perhaps representing H3K9 methylation of H3.3A and H3.3B. These results indicate that replication dependent H3K9 is important but not absolutely necessary for *Drosophila* development.

Genotype		Hatch				Pupate			Eclose		
	No. eggs	No. hatch	% hatch	p	No. larvae	No. pupate	% pupate	p	No. eclose	% eclose [b]	p
yw	300	252	84.0	1	150	125	83.3	1	113	75.3	1
His ; HWT	230	197	85.6	0.494	175	101	57.7	$< 10^{-4}$	84	48.0	$< 10^{-4}$
His ; K9R	300	246	82.0	0.345	160	56	35.0	$< 10^{-4}$	4 [a]	2.5	$< 10^{-4}$

Table 2.1: H3K9R flies can complete development. Shown are the percentage of hatched embryos from the total number (No.) of eggs counted and the percentage of pupae or adults developed from the total number of larvae used for cultures. P value for chi-square test calculated using *yw* observed values as expected values.

a. Genotype of H3K9R adults confirmed by DNA sequencing.
b. The physical manipulation of larvae used to culture the desired genotype independently of their siblings likely reduced the proportion of *yw* and *His; HWT* individuals that developed to adulthood compared to McKay et al. (2015).

Figure 2.1: H3K9R mutant cells proliferate with severely reduced H3K9me2 and me3 in pericentric heterochromatin. A) Twin spot analysis of FLP-FRT induced mitotic clones of *HisC* deletion wing disc cells rescued with an HWT transgene (top) or a K9R transgene (bottom). Wing discs were stained with DAPI to mark nuclei, anti-H3K9me2 to identify K9R cells, and anti-GFP to identify twin spots, (red lines). *HisC* deletion cells lack GFP and control sister clones are homozygous *HisC*+ and express 2X GFP. Scale bar = 1000μm. White boxes indicate magnified regions where the scale bar = 20μm. B) Quantification of twin spot clone area. Each dot represents the area of the experimental (HWT or K9R) clone divided by the area of the control twin spot clone. ** = $p<0.005$. C) Western blot analysis of total cellular histone isolated from whole third instar larvae. D) Polytene chromosome preparations from third instar larval salivary glands stained with DAPI and anti-H3K9me antibodies. Scale bar = 20μm. Magnified images show H3K9me3 staining at chromocenter (white) and chromosome arms (yellow). Scale bar = 10μm. Note that H3K9me3 signal at the chromocenter is overexposed to reveal staining on chromosome arms.

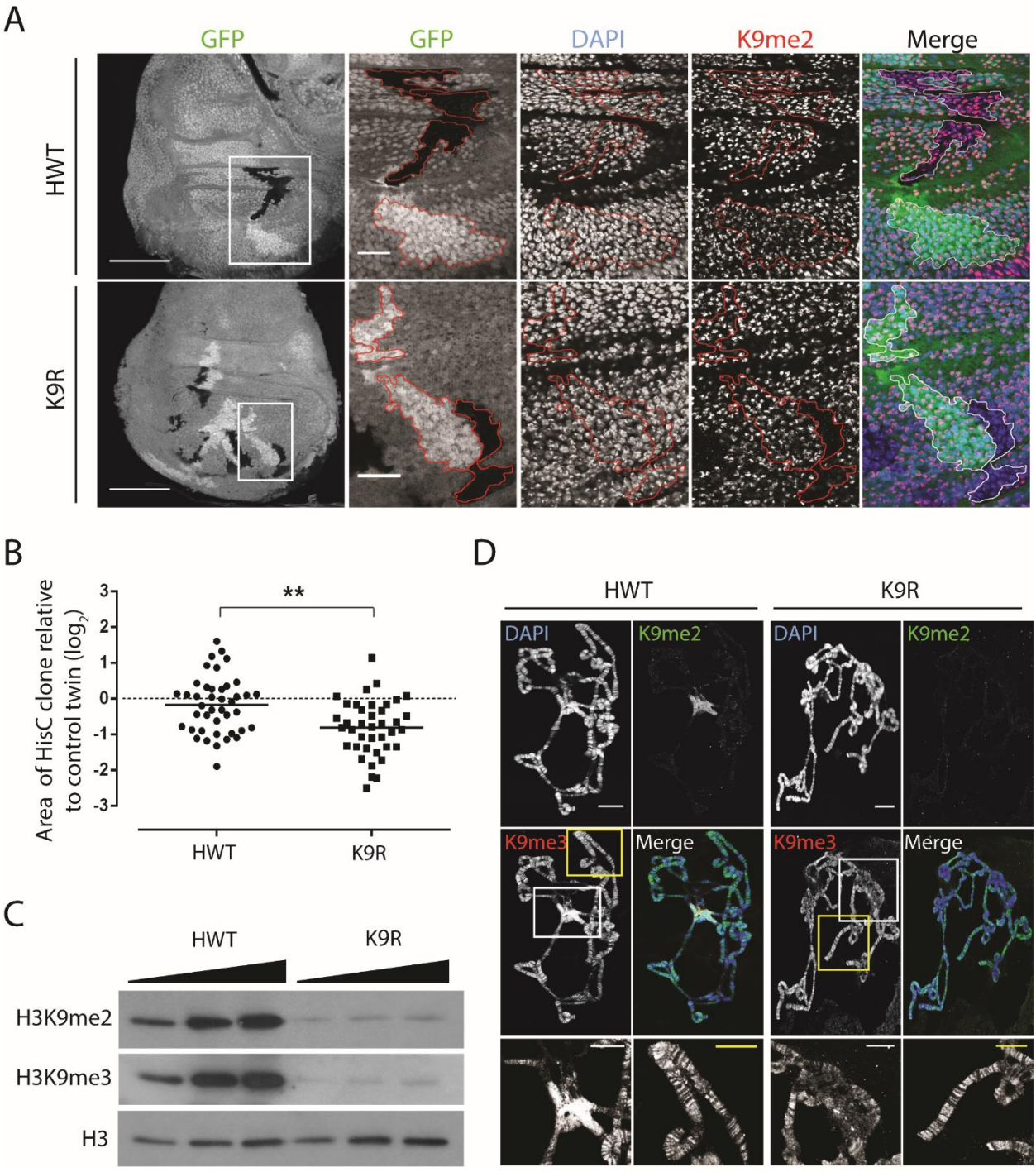
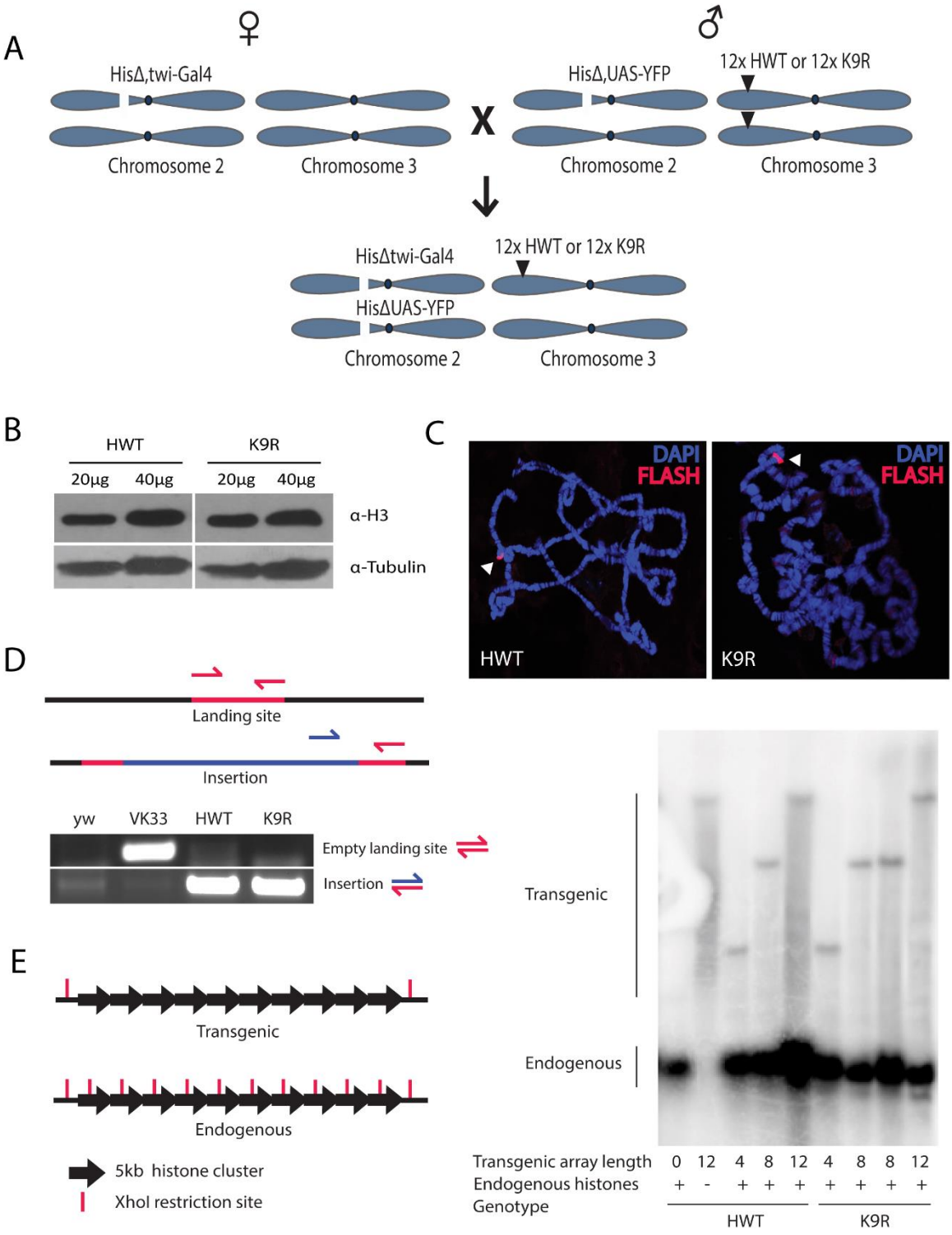

A
GFP
GFP
DAPI
K9me2
Merge
HWT
K9R
B
Area of HisC clone relative to control twin (log2)
**
HWT
K9R
C
HWT
K9R
H3K9me2
H3K9me3
H3
D
HWT
K9R
DAPI
K9me2
K9me3
Merge

Figure 2.2: Generation and verification of H3K9R mutant genotype. A) Progeny lacking endogenous histone genes were generated by crossing parents heterozygous for the *HisC* deletion and identified by GFP expression using the Gal4-UAS system. We analyzed *HisC* deletion mutants containing a transgenic array of 12 HWT or 12 H3K9R histone clusters integrated at the same VK33 *attP* site on chromosome 3. B) Anti-histone H3 western blot of total cell protein extracted from whole third instar larvae. C) Third instar larval polytene chromosome preparation stained with DAPI and anti-FLASH antibodies to mark the histone locus body (HLB), a nuclear body that recruits factors involved in replication-dependent histone mRNA biosynthesis. Both HWT and K9R transgenes form an HLB (arrowheads) at the expected location on chromosome 3L. D) PCR verification of transgene insertion into the VK33 *attP* transgene landing site. yw contains no VK33 landing site, and VK33 is the empty landing site. Arrows indicate primer locations used for PCR. E) Southern blot verification of histone cluster copy number in various HWT and K9R transgenes. We eliminated a XhoI site from the transgenic *H2A* gene arrays that is present in the H2A gene of the endogenous histone locus. XhoI digestion collapses the endogenous histone gene array to 5kb fragment but leaves the transgenic array intact. A 500bp sequence from *H2A* was used as a probe.


A
♀
♂
HisΔ,twi-Gal4
HisΔ,UAS-YFP
12x HWT or 12x K9R
X
Chromosome 2
Chromosome 3
Chromosome 2
Chromosome 3
HisΔtwi-Gal4
12x HWT or 12x K9R
HisΔUAS-YFP
Chromosome 2
Chromosome 3
B
HWT
K9R
20µg
40µg
20µg
40µg
α-H3
α-Tubulin
C
DAPI
FLASH
HWT
DAPI
FLASH
K9R
D
Landing site
Insertion
yw
VK33
HWT
K9R
Empty landing site
Insertion
E
Transgenic
Endogenous
5kb histone cluster
XhoI restriction site
Transgenic
Endogenous
Transgenic array length 0 12 4 8 12 4 8 8 12
Endogenous histones + - + + + + + + +
Genotype
HWT
K9R

H3K9 regulates chromocenter organization and nucleosome occupancy

Given the association of H3K9 methylation with heterochromatin formation, we next examined K9R mutants for defects in chromatin structure by examining salivary gland polytene chromosomes. Polytene chromosomes are generated by endoreduplication, wherein cells become polyploid through repeated rounds of S phase without cell division. The pericentromeric heterochromatin of salivary gland cells is under-replicated relative to the euchromatin and coalesces into a structure known as the chromocenter. The polytene chromosome arms in both HWT and K9R displayed typical banded structures. In contrast, the chromocenter of K9R nuclei was highly disrupted compared to HWT. Chromosome arms frequently failed to meet in an organized structure, and the DNA appeared less condensed (Figure 2.3A). We quantified this phenotype by blindly binning chromocenters into three categories: “organized”, “moderately disorganized”, and “severely disorganized”. Whereas 74% of HWT nuclei appeared organized, only 4% of K9R nuclei did, with 24% and 72% appearing moderately or severely disorganized, respectively (Figure 2.3B). This cytological defect corresponds with a loss of H3K9me. Although the chromocenter of HWT polytene chromosomes has strong H3K9me2 and -me3 signal, K9R chromocenters have severely diminished signal (Figure 2.1D). In contrast, along the largely euchromatic arms, H3K9me3 signal is unchanged in K9R mutants. Because replication-independent H3.3 variants are typically found at transcriptionally active areas, we surmise that methylation of H3.3A and H3.3B might be the source of H3K9me3 along chromosome arms in K9R mutants.

Current models posit that the “closed” nature of heterochromatin prevents access of transcription factors to DNA. Our cytological observations of salivary gland

chromocenters suggested pericentric heterochromatin may be more “open” in K9R mutants. We hypothesized that disruption of heterochromatin organization in K9R mutants might allow chromatin remodelers or other factors to access DNA, leading to changes in nucleosome occupancy. We therefore interrogated nucleosome occupancy throughout the genome of K9R mutants using Formaldehyde Assisted Isolation of Regulatory Elements coupled with sequencing (FAIRE-seq). We performed FAIRE-seq on 3rd instar larval imaginal wing discs, a tissue consisting of an epithelial sheet of diploid cells that can be uniformly crosslinked. For both HWT and K9R, characteristic peaks of FAIRE signal occurred near transcription start sites as expected and correlated well with previously published FAIRE data (Figure 2.4, McKay & Lieb 2013). To compare open chromatin between the two genotypes, we identified peaks of FAIRE signal and quantified the number of reads overlapping each peak, normalizing to the average read depth across the genome. Strikingly, of the 24,025 total FAIRE peaks, 5,154 showed a significant increase in FAIRE signal in K9R mutants relative to HWT (Figure 2.3C, $p<0.01$). In contrast, 642 peaks exhibited increased signal in HWT relative to K9R, and these peaks were characterized by much smaller differences in signal. Importantly, increases in FAIRE signal in K9R mutants were not due to changes in DNA copy number (Figure 2.5A).

Figure 2.3: H3K9 regulates chromatin organization at pericentromeric heterochromatin. A) HWT and K9R salivary gland polytene chromosomes stained with DAPI and H3K9me2. Scale bar = 20µm. Bottom images show a magnified view (white squares) of the chromocenter (dashed lines). Scale bar =10µm. B) Quantification of chromocenter organization from *yw* (contains endogenous histones), HWT, and K9R. C) K9R/HWT ratio of normalized FAIRE signal from 3rd instar imaginal wing discs at FAIRE peaks called by MACS2 (CPM = counts per million). The X-axis indicates the average HWT and K9R signal at each peak. Darker colors in the heat map indicate a higher number of peaks. Red peaks are statistically significant as determined by edgeR ($p<0.01$). Lines indicate two-fold change. D) Genome browser shot of FAIRE peaks near the euchromatic gene *engrailed* and the heterochromatic gene *concertina*. Map= read mappability. E) K9R/HWT ratio of normalized FAIRE signal plotted versus genome coordinate of FAIRE peaks on chromosome 2 and 3. Green regions in the chromosome schematic indicate approximate locations of pericentromeric heterochromatin (Riddle et al. 2011; Hoskins et al. 2015). Blue indicates largely euchromatic regions. Loess regression line of modENCODE K9me3 ChIP signal shown in red. F) Boxplot of average ratio of FAIRE signal for FAIRE peaks assigned to one of nine chromatin states (Kharchenko et al. 2011). See also Figure 2.5C,D.

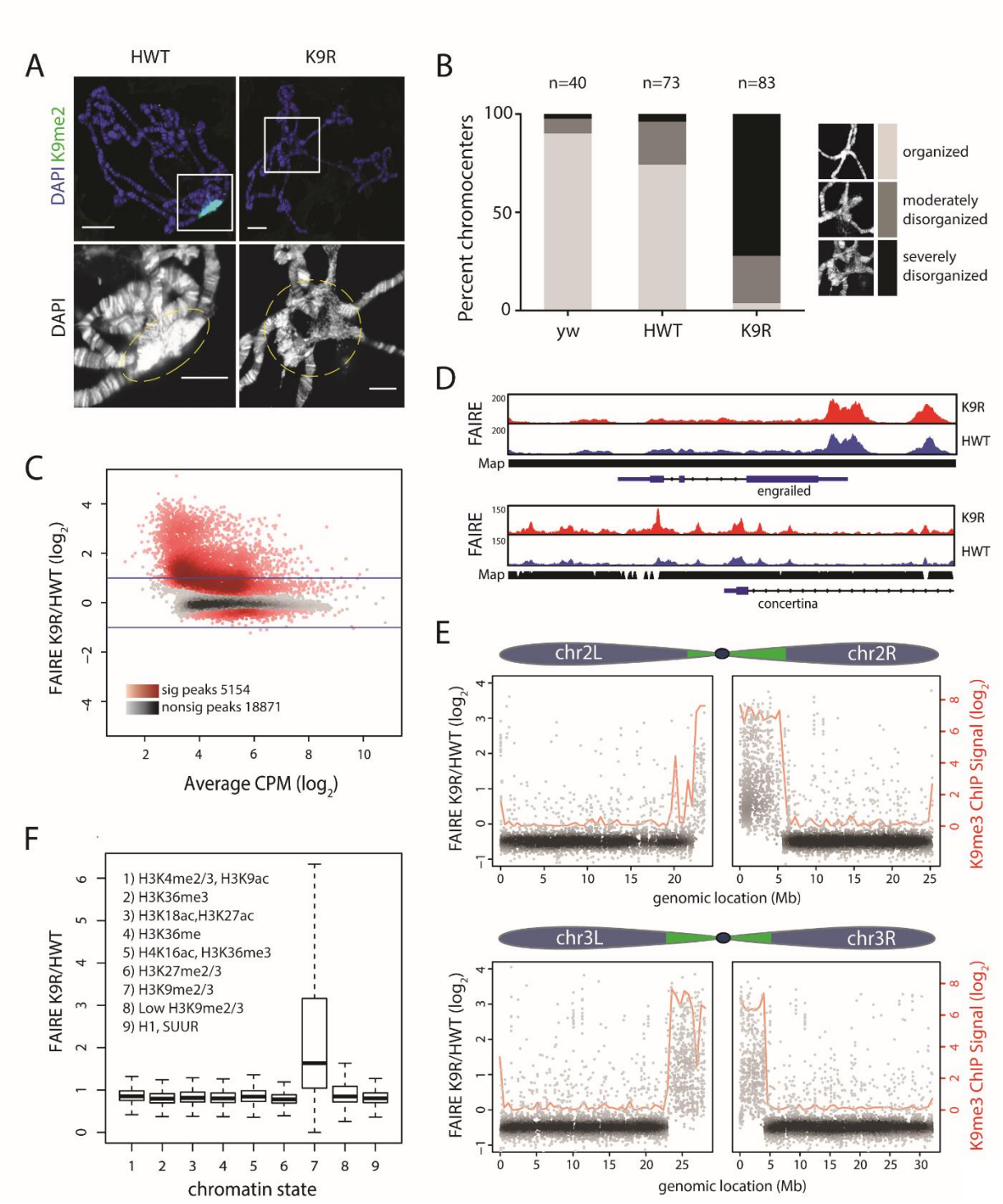
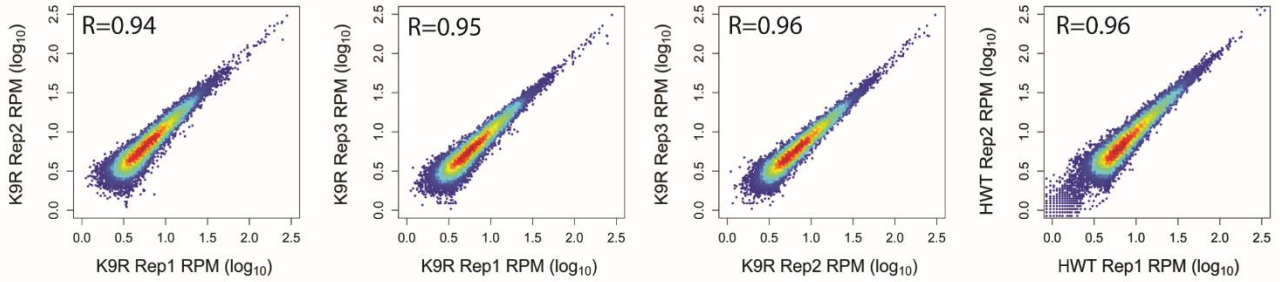
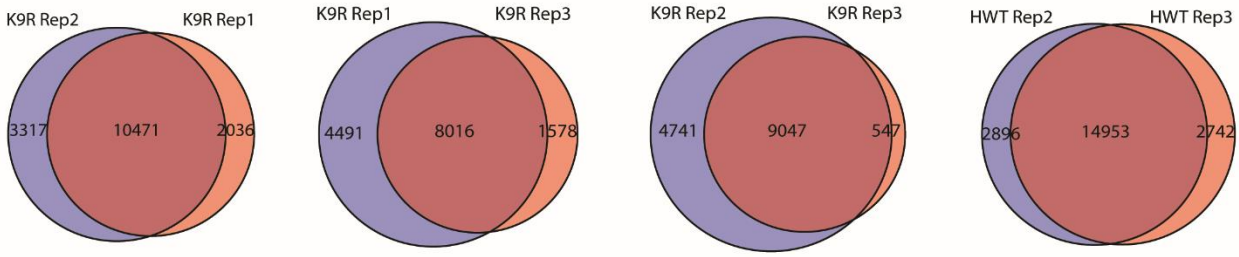

A
HWT
K9R
DAPI K9me2
DAPI
B
n=40
n=73
n=83
Percent chromocenters
100
50
0
yw
HWT
K9R
organized
moderately disorganized
severely disorganized
C
FAIRE K9R/HWT (log2)
sig peaks 5154
nonsig peaks 18871
Average CPM (log2)
D
FAIRE
K9R
HWT
Map
engrailed
concertina
E
chr2L
chr2R
chr3L
chr3R
FAIRE K9R/HWT (log2)
K9me3 ChIP Signal (log2)
genomic location (Mb)
F
1) H3K4me2/3, H3K9ac
2) H3K36me3
3) H3K18ac,H3K27ac
4) H3K36me
5) H4K16ac, H3K36me3
6) H3K27me2/3
7) H3K9me2/3
8) Low H3K9me2/3
9) H1, SUUR
FAIRE K9R/HWT
chromatin state

Figure 2.4: FAIRE-seq signal correlates between replicates and occurs at characteristic regions. A) FAIRE sequencing was performed in biological triplicate for HWT and K9R 3rd instar imaginal wing disc samples with three K9R and two HWT replicates meeting quality control standards. Normalized counts at each peak are shown in a heat scatter plot with red regions containing the highest number of peaks. R values shown are Pearson correlations. B) Venn diagrams showing peak overlap between replicates. C,D) Correlation analysis (C) and peak overlap (D) of HWT replicates and previously published FAIRE data in wing discs (McKay & Lieb 2013). E) Average HWT FAIRE signal at 5bp bins surrounding transcription start sites (TSS). Signal expressed as average reads per million (RPM). F) Percentage of total aligned reads in HWT and K9R replicates that uniquely mapped or could be mapped to multiple locations.

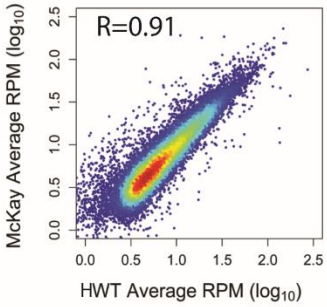
A



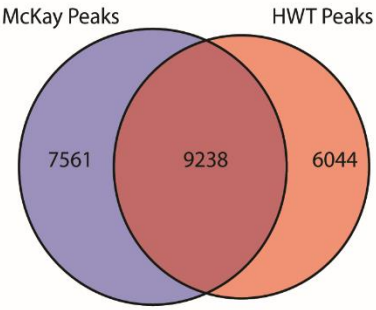
B



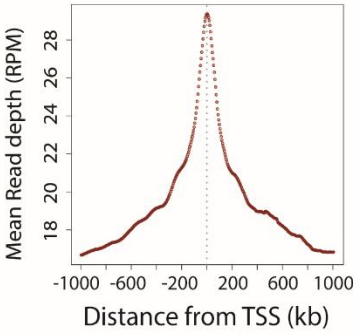
C



D



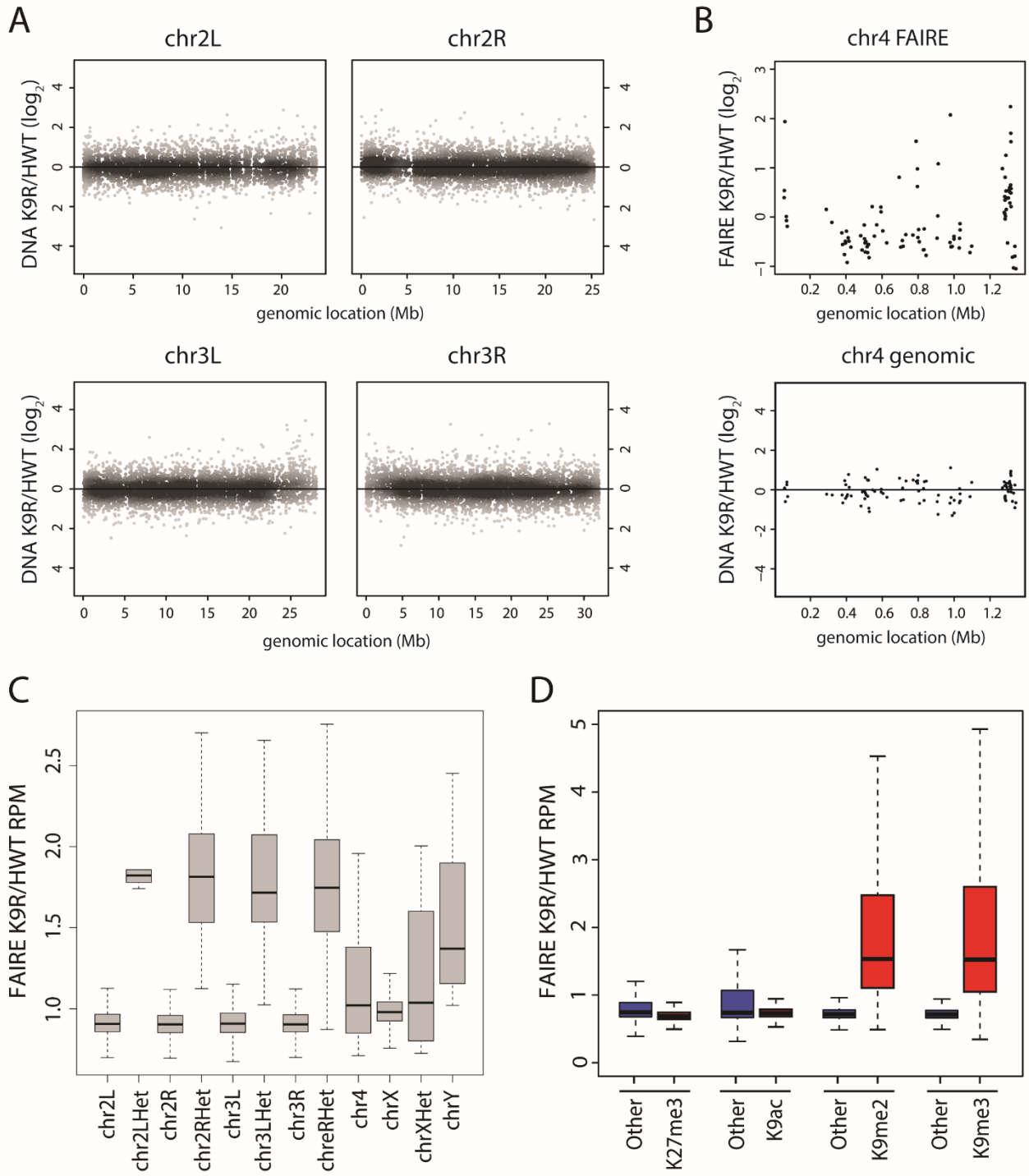
E



F

Replicate	Unique(%)	Multiple (%)
HWT 1	80.0	20.0
HWT 2	80.0	20.0
K9R 1	68.0	32.0
K9R 2	69.8	30.2
K9R 3	71.0	28.9

Figure 2.5: K9R mutants lack DNA copy number changes and have increased FAIRE signal at constitutive heterochromatin. A) Genomic DNA sequencing of HWT and K9R wing discs. Ratio of K9R to HWT read depth at merged FAIRE peak set (Figure 2.4) for chromosome 2 and 3. B) Ratio of K9R to HWT FAIRE-seq (top) and genomic DNA-seq (bottom) signal at the largely heterochromatic chromosome 4. C) FAIRE reads were aligned to the dm3 reference genome which contains scaffolds of repetitive heterochromatic regions separate from the largely euchromatic scaffolds. FAIRE signal expressed as Reads per Million (RPM) was used to calculate K9R/HWT ratio for each FAIRE peak. Boxplot shows average ratio for each dm3 scaffold. D) Comparison of HWT and K9R FAIRE signal at regions of K27me3, K9ac, K9me2, and K9me3 enrichment. Histone PTM enrichment was determined by peak calling from modENCODE 3rd instar larval ChIP-seq data (Celniker et al., 2009). A merged FAIRE peak data set from the two genotypes was used to separate FAIRE peaks that overlap a particular histone PTM peak from FAIRE peaks that did not overlap. Boxplots show the average K9R/HWT ratio of FAIRE signal for the two peak categories.


A
chr2L
chr2R
chr3L
chr3R
DNA K9R/HWT (log2)
genomic location (Mb)
B
chr4 FAIRE
FAIRE K9R/HWT (log2)
chr4 genomic
C
FAIRE K9R/HWT RPM
chr2L
chr2LHet
chr2R
chr2RHet
chr3L
chr3LHet
chr3R
chreRHet
chr4
chrX
chrXHet
chrY
D
Other
K27me3
K9ac
K9me2
K9me3

We next determined the genomic location of peaks with differences in FAIRE signal. In concordance with our results using polytene chromosomes, genome browsing suggested that increases in FAIRE signal occurred in heterochromatic regions. For example, the euchromatic *engrailed* locus showed similar FAIRE signatures between K9R and HWT, whereas the heterochromatic *concertina* locus displayed increased signal in K9R mutants (Figure 2.3D). Whole genome analysis demonstrated an increase in FAIRE signal near centromeres in K9R mutants, and that FAIRE signal was largely unchanged along chromosome arms (Figure 2.4E, 2.5B). The slight decrease in FAIRE signal along chromosome arms in K9R mutants is likely a byproduct of the normalization procedure due to the vast number of reads redistributed to heterochromatic regions in K9R samples (Figure 2.4F). These results demonstrate that H3K9 is important for establishing or maintaining nucleosome occupancy at pericentromeric heterochromatin.

Previously, Kharchenko et al. (2010) used various combinations of histone PTMs and chromatin proteins to define 9 distinct ‘chromatin states’ in *Drosophila* (Figure 2.3F). We therefore determined whether the changes in FAIRE signal we observed in K9R mutants correlated with a particular chromatin state. We compared the average ratio of signal in K9R samples to HWT controls for each of the 9 chromatin states. Regions of H3K9me2/me3, represented by chromatin state 7, on average had higher FAIRE signal in K9R samples (Figure 2.3F). By contrast, the FAIRE signal was similar between the two genotypes at all other chromatin states, showing that regions of nucleosome depletion in K9R mutants occur where H3K9me is normally present.

Taken together, our cytological and genome wide data support an essential role for H3K9 in chromatin organization at pericentromeric regions.

H3K9 is required for HP1a localization to pericentromeric heterochromatin

HP1 family proteins are characterized by a chromodomain, which binds to H3K9me2/me3, and a chromo-shadow domain, which multimerizes to mediate inter-nucleosomal interactions and recruit a variety of chromatin regulating proteins (Eissenberg and Elgin 2014). In fission yeast, H3K9A/R mutations disrupt localization of the HP1 homolog Swi6 (Mellone et al. 2003). Therefore, one possible explanation for the altered FAIRE signatures in H3K9R mutants is the inability to recruit HP1 due to lack of H3K9me2/me3. The best studied of the five *Drosophila* HP1 paralogs is HP1a, which primarily localizes to heterochromatin. We first examined HP1a localization via immunofluorescence in K9R salivary gland polytene chromosomes. Consistent with our earlier experiments, HP1a was severely depleted from the chromocenter in K9R mutants, but remained localized to telomeres and chromosome arms (Fig 2.6A). In fact, HP1a staining along chromosome arms in K9R mutants appeared stronger and was ectopically localized to regions that were undetectably stained in HWT (Figure 2.6A, yellow boxes). Because HP1a protein levels were comparable in HWT and K9R salivary glands (Figure 2.6B), this result suggests HP1a relocalizes from pericentromeric heterochromatin to chromosome arms in the absence of H3K9.

To further explore the effect of H3K9 loss on HP1, we analyzed HP1a association with the genome in K9R mutants by performing ChIP-seq for HP1a using nuclei isolated from whole 3rd instar larvae (Figure 2.7A,B). Given the HP1a immunofluorescence data in polytene chromosomes, we expected to detect

relocalization of reads from repetitive heterochromatic to euchromatic regions in K9R mutants. As seen in the FAIRE-seq data, given a fixed sequencing depth, a large increase in reads that map to one region of the genome (e.g., pericentric heterochromatin) results in a corresponding reduction in read coverage throughout the rest of the genome (Figure 2.3E). Consequently, normalizing to sequencing depth in this case misrepresents the relative signal between the two genotypes (Orlando et al. 2014). We therefore developed an alternative procedure that uses *Drosophila virilis* chromatin as an internal normalization control. The *D. melanogaster* HP1 antibody recognizes *virilis* HP1 (data not shown) and the *virilis* genome is sufficiently diverged from *melanogaster* to permit unambiguous mapping of high throughput sequencing reads. We added *virilis* chromatin to the *melanogaster* chromatin preparations prior to HP1a immunoprecipitation and normalized *melanogaster* sequencing reads to uniquely mapping *virilis* reads (Figure 2.7C). To focus on HP1a enriched regions, we quantified the total number of reads within 1kb windows across the genome and selected for analysis those windows with the highest number of reads (top 20%). Similar to results from polytene chromosomes, HP1a was significantly depleted at pericentromeric heterochromatin in K9R mutants but showed higher signal along chromosome arms (Figure 2.6C, 2.7E). For example, 99% of pericentromeric windows with significantly different HP1a signal showed more signal in HWT than in K9R samples. By contrast, 66% of significantly different windows in chromosome arms showed more signal in K9R than in HWT samples.

We also carried out HP1a ChIP from wing discs; however, the signal-to-noise ratios from these experiments were too low to make confident conclusions, presumably due to low input DNA. Nevertheless, the ChIP-seq read patterns described above were also seen in the wing disc dataset (Figure 2.7E).

Figure 2.6: HP1a relocalizes from pericentromeres to chromosome arms in the absence of H3K9 A) HWT and K9R salivary gland polytene chromosomes stained with anti-H3K9me2 and anti-HP1a antibodies. Scale bar = 20 μm. Bottom panels show a magnified view of the chromocenter (white) and a chromosome arm (yellow) for each genotype. Arrows indicate telomeres. Scale bar = 5 μm. B) α-HP1a Western blot of 3ug, 6ug, and 12ug of whole cell extract from HWT and K9R salivary glands. C) K9R/HWT ratio of HP1a ChIP-seq signal from whole 3rd instar larvae within 1kb windows tiled across the five autosome arms. The top 20% of 1kb windows with the highest counts are shown (see Fig 2.7 for all windows). Pie charts show percentage of significantly altered windows on pericentromeres or chromosome arms as called by edgeR ($p<0.01$). D) Scatterplot of HP1a signal at FAIRE peaks with higher signal in K9R samples (top) or a random selection of FAIRE peaks that are not significantly different between HWT and K9R (bottom).

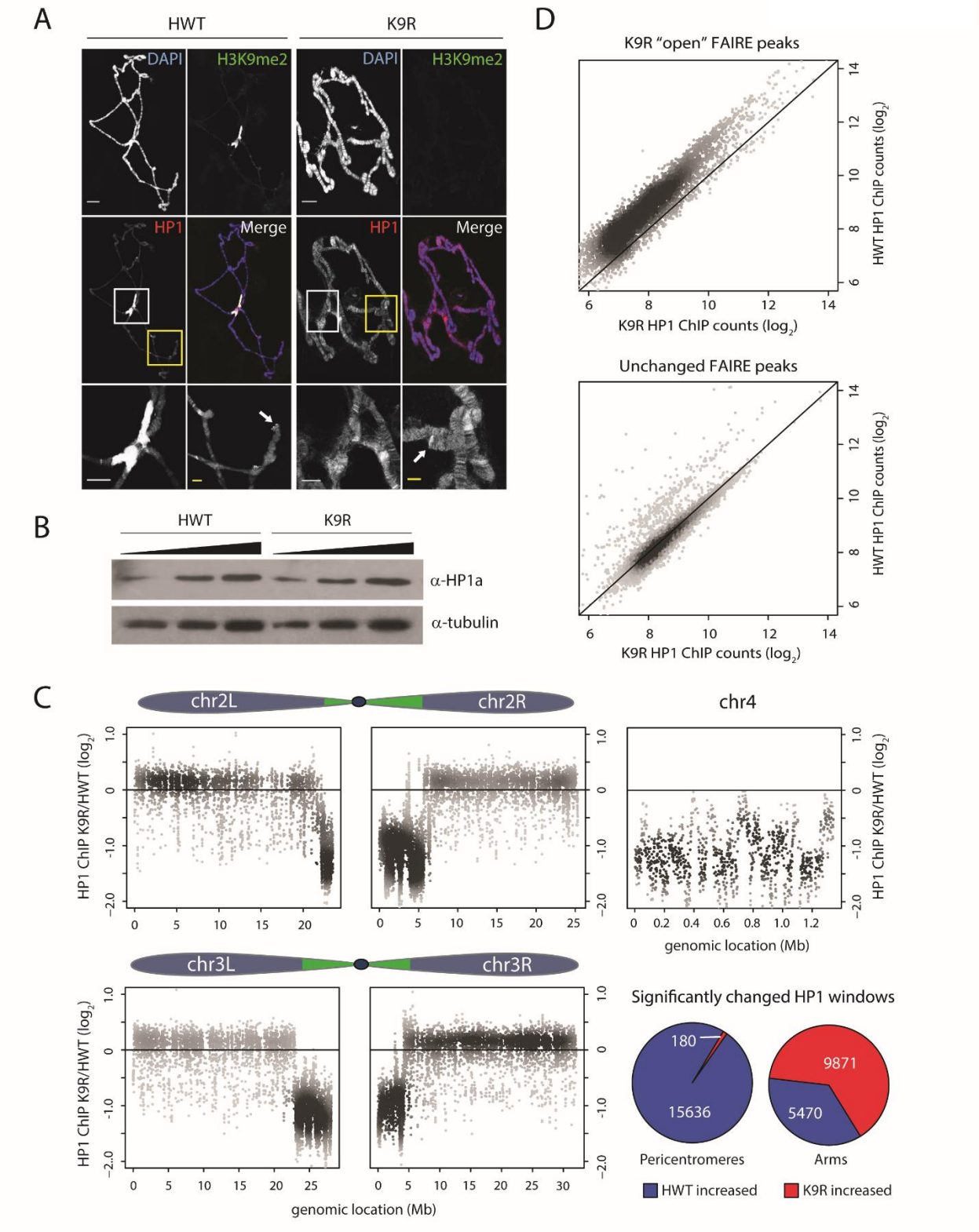
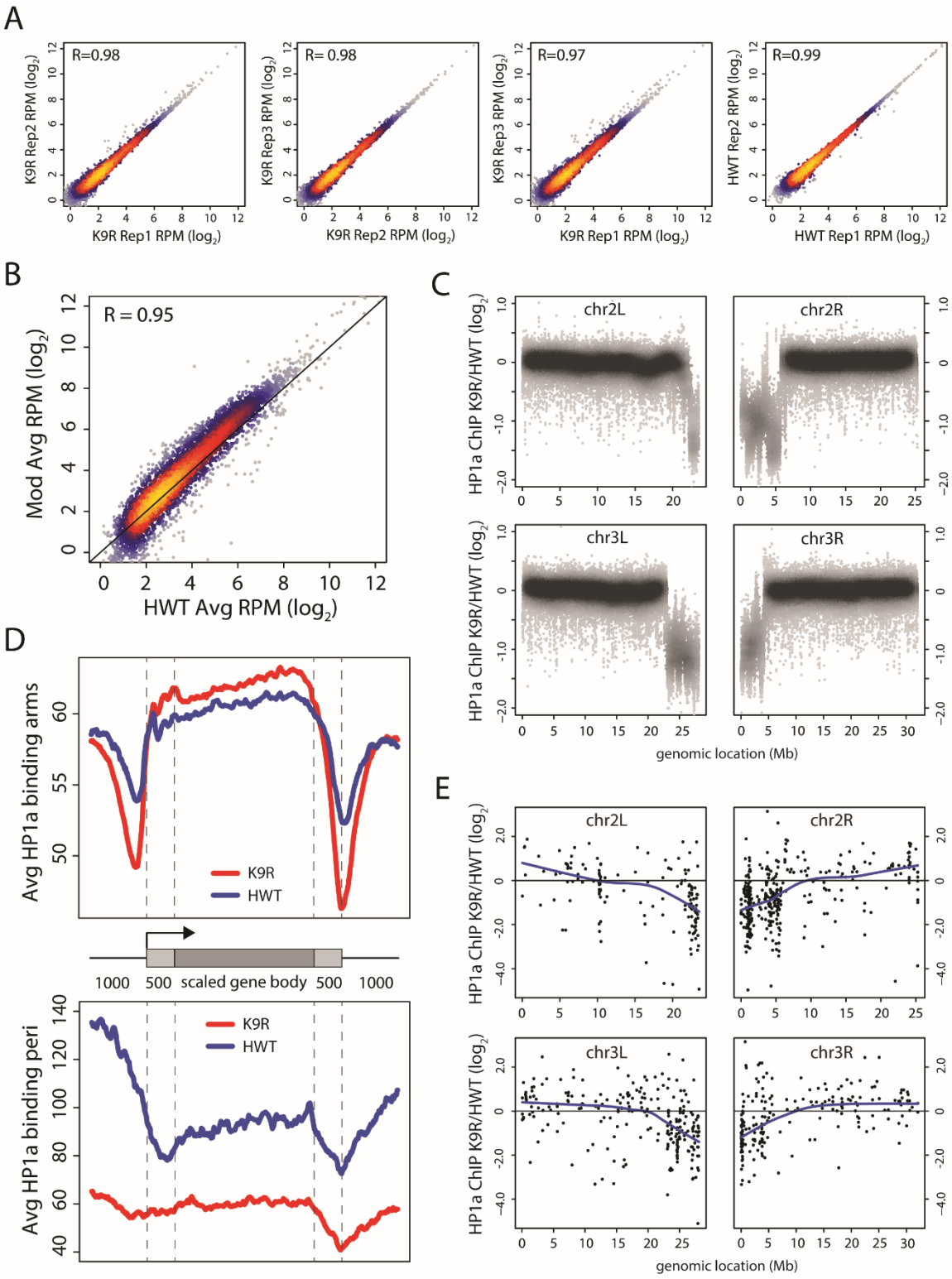

A
HWT
K9R
DAPI
H3K9me2
HP1
Merge
B
HWT
K9R
α-HP1a
α-tubulin
C
chr2L
chr2R
chr3L
chr3R
chr4
HP1 ChIP K9R/HWT (log2)
genomic location (Mb)
Significantly changed HP1 windows
180
15636
9871
5470
Pericentromeres
Arms
HWT increased
K9R increased
D
K9R "open" FAIRE peaks
Unchanged FAIRE peaks
HWT HP1 ChIP counts (log2)
K9R HP1 ChIP counts (log2)

Figure 2.7: HP1a relocalizes along chromosome arms in K9R mutants. A) Heatscatter plots of HP1a ChIP signal within 1 kb windows for two HWT and three K9R biological replicates. Signal is expressed as reads per million (RPM), and R values indicate Pearson correlation. B) Comparison of HWT HP1a ChIP-seq samples to modENCODE HP1a ChIP-seq samples from 3rd instar larvae. HP1a signal at 1kb windows was normalized to total number of aligned reads and expressed as RPM. R value indicates Pearson correlation. C) Ratio of K9R to HWT normalized HP1a signal plotted versus genomic location in megabases (Mb). Centering of HP1a fold change near zero on chromosome arms supports the *D. virilis* chromatin spike-in normalization procedure. D) Metagene analysis of HP1a signal across all genes on chromosome arms (top) and in pericentromeres (bottom). Schematic indicates 1000 bp upstream (left line) and 500 bp downstream (left light grey box) of the transcription start site (bent arrow) and 500 bp upstream (right light grey box) and 1000 bp downstream (right line) of the transcription termination site (right edge of light grey box). The rest of the coding regions (dark grey box) were scaled into an equal number of windows. E) HP1a ChIP-seq signal from wing discs at peaks called by MACS2. Blue line indicates loess regression line.


A
R=0.98
K9R Rep2 RPM (log2)
K9R Rep1 RPM (log2)
R= 0.98
K9R Rep3 RPM (log2)
K9R Rep2 RPM (log2)
R=0.97
K9R Rep3 RPM (log2)
K9R Rep1 RPM (log2)
R=0.99
HWT Rep2 RPM (log2)
HWT Rep1 RPM (log2)
B
R = 0.95
Mod Avg RPM (log2)
HWT Avg RPM (log2)
C
chr2L
chr2R
chr3L
chr3R
HP1a ChIP K9R/HWT (log2)
genomic location (Mb)
D
Avg HP1a binding arms
K9R
HWT
1000
500
scaled gene body
500
1000
Avg HP1a binding peri
E
chr2L
chr2R
chr3L
chr3R
HP1a ChIP K9R/HWT (log2)
genomic location (Mb)

To examine how HP1a redistribution to chromosome arms occurs, we asked if any specific genomic feature correlated with increased HP1a in K9R mutants. Using modENCODE data, we did not find a correlation between relocalized HP1 signal and any particular histone PTM. For example, we did not observe an increase in HP1 signal at regions of H3K27me3, despite the similarity in peptide sequence surrounding the H3K9 and H3K27 residues. Rather, metagene analysis demonstrated that HP1a is enriched across gene bodies along chromosome arms (Figure 2.7D). In addition to the pericentromeric regions, we observed a significant decrease in HP1a binding to the largely heterochromatic fourth chromosome (Figure 2.6C). These data indicate that H3K9 is necessary for recruitment of HP1a to pericentromeres and chromosome four, and in the absence of these strong binding sites, HP1a relocalizes to chromosome arms.

Interestingly, both FAIRE and HP1a signals are altered at the pericentromeric heterochromatin of K9R mutants relative to HWT. To determine whether this correlation might have a functional basis, we examined HP1a signal at FAIRE peaks that have greater signal in K9R relative to HWT samples. In K9R mutants, we found a decrease in HP1a signal at those regions of the genome with increased FAIRE signal (Figure 2.6D). In contrast, HP1a signal at a random selection of peaks was equivalent between K9R and HWT. This result suggests that HP1a functions to promote nucleosome occupancy or to maintain a “closed” chromatin environment. However, loss of HP1a from a subset of pericentric peaks (Figure 2.6D) and the fourth chromosome (Figure 2.6B) resulted in few changes in nucleosome occupancy, indicating that HP1 depletion does not always change nucleosome occupancy.

H3K9 prevents accumulation of transcripts from heterochromatic regions of the genome

Analyses in a variety of organisms have indicated that HP1 and H3K9 methyltransferases are important regulators of transcription of both protein coding and noncoding RNAs (Grewal and Jia 2007; Elgin and Reuter 2013). We therefore investigated whether the changes in chromatin organization and HP1a localization we observed in K9R mutant animals impact the transcriptome. To examine the effects of H3K9R on gene regulation, we performed total RNA-seq in wing discs from HWT and K9R 3rd instar larvae and assembled a transcriptome using Cufflinks (Figure 2.8A,B; Trapnell et al., 2014). Of transcripts that were differentially expressed between the two genotypes, 999 were higher in K9R discs relative to HWT compared to 175 that were lower (Figure 2.9A, $p<0.05$). These data suggest H3K9 generally functions to repress gene expression.

Figure 2.8: Protein-coding genes have similar expression levels in HWT and K9R genotypes. A) Comparison of RNA-seq data from three biological replicates of HWT and K9R wing discs. Scatterplots show FPKM (Fragments Per Kilobase of transcript per Million mapped reads) calculated by Cufflinks at each gene. R value represents Pearson correlation. B) Comparison of an average of HWT RNA-seq samples to wing disc RNA-seq from McKay & Lieb 2013. R value represents Pearson correlation. C) K9R/HWT ratio of expression for 46 genes located in heterochromatin on chromosome 2 that were identified using data from Corradini et al. 2007. Only genes with transcripts included in the dm6 reference transcriptome were included in the analysis. Red indicates the two genes, *CG30440* and *chitinase 3 (Cht3),* with significantly different expression between K9R and HWT samples as determine by DESeq2 (p value<0.05). *Concertina (cta), light (lt), rolled (rl)* are also indicated. D) Scatterplots of K9R/HWT expression ratios for transcripts that do not overlap (top) or overlap (bottom) H3K9ac enriched regions. Only 4% of transcripts from genes containing H3K9ac are differentially expressed (210/5260) compared to 14% of transcripts that are not associated with H3K9ac (950/6563). Moreover, H3K9ac is not associated with changes in FAIRE signal (see Figure S3D). Red dots indicate statistical significance as determined by DESeq2 ($p<0.05$).

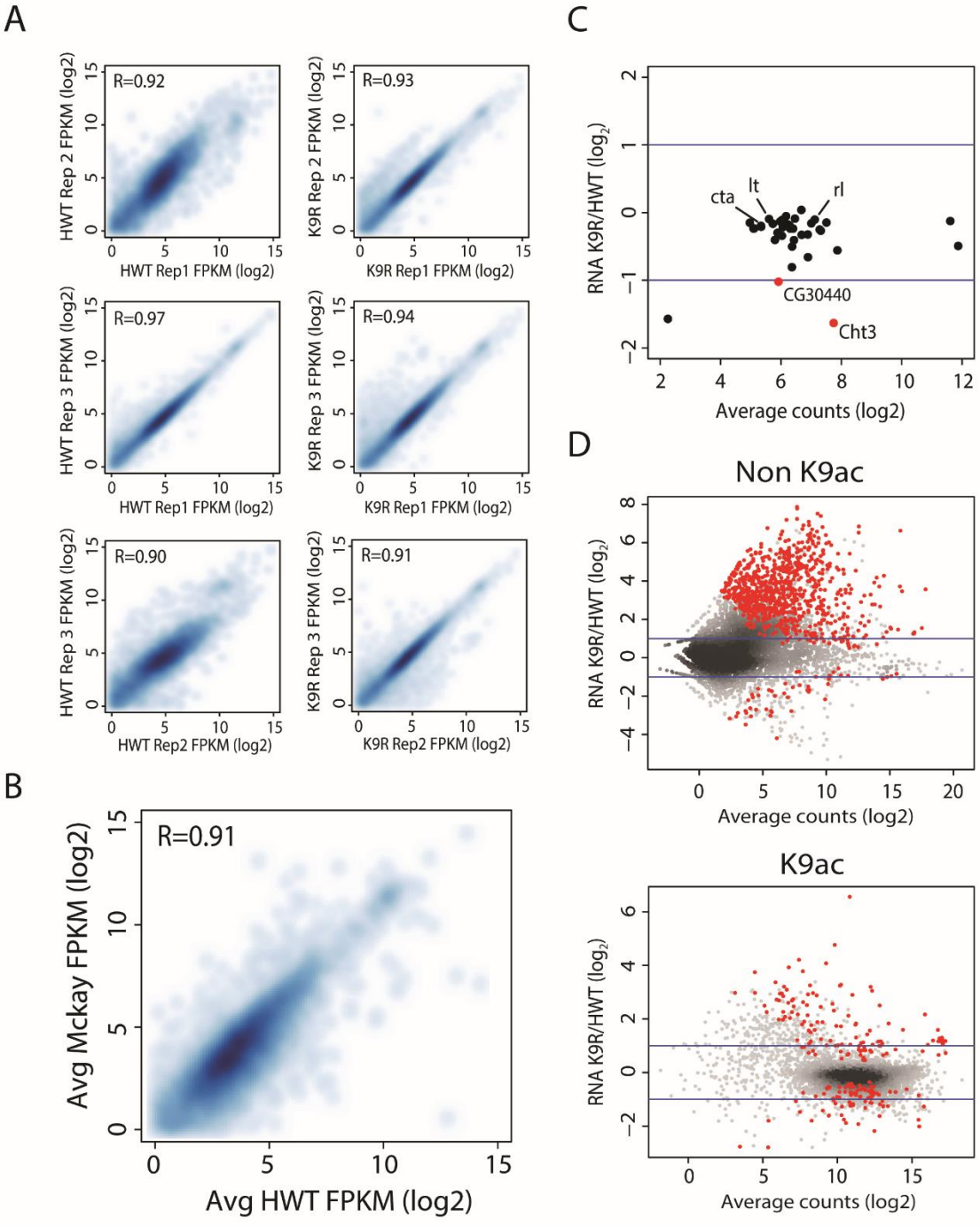
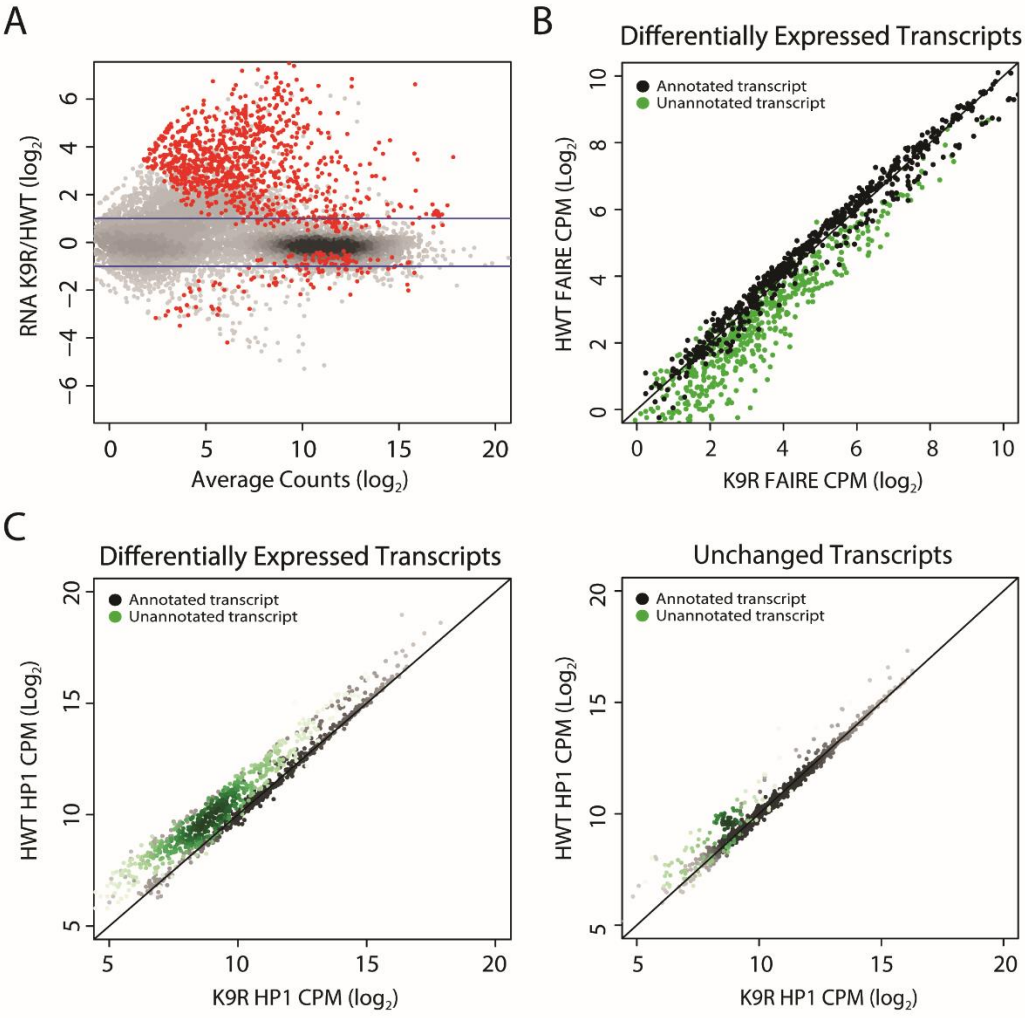

A
R=0.92
HWT Rep 2 FPKM (log2)
HWT Rep1 FPKM (log2)
R=0.93
K9R Rep 2 FPKM (log2)
K9R Rep1 FPKM (log2)
R=0.97
HWT Rep 3 FPKM (log2)
HWT Rep1 FPKM (log2)
R=0.94
K9R Rep 3 FPKM (log2)
K9R Rep1 FPKM (log2)
R=0.90
HWT Rep 3 FPKM (log2)
HWT Rep2 FPKM (log2)
R=0.91
K9R Rep 3 FPKM (log2)
K9R Rep2 FPKM (log2)
B
R=0.91
Avg Mckay FPKM (log2)
Avg HWT FPKM (log2)
C
cta
lt
rl
CG30440
Cht3
RNA K9R/HWT (log2)
Average counts (log2)
D
Non K9ac
RNA K9R/HWT (log2)
Average counts (log2)
K9ac
RNA K9R/HWT (log2)
Average counts (log2)

Figure 2.9: Differential expression of transcripts between HWT and K9R correlates with changes in FAIRE signal and HP1a localization A) K9R/HWT ratio of RNA-seq signal from 3rd instar imaginal wing discs. Statistically different ratios identified using DESeq2 are indicated in red ($p<0.05$). Blue lines indicate two-fold change. B) Normalized FAIRE signal at all transcripts differentially expressed between HWT and K9R (red dots in A). Signal expressed as counts per million (CPM). Black dots indicate transcripts annotated in the RefSeq reference transcriptome. Green dots here and in panel C indicate transcripts identified in Cufflinks transcriptome assembly but not in the reference transcriptome. C) Scatterplot of normalized HP1a ChIP signal within transcripts differentially expressed between K9R and HWT (left) or a random selection of transcripts that are not significantly different between the two genotypes (right).

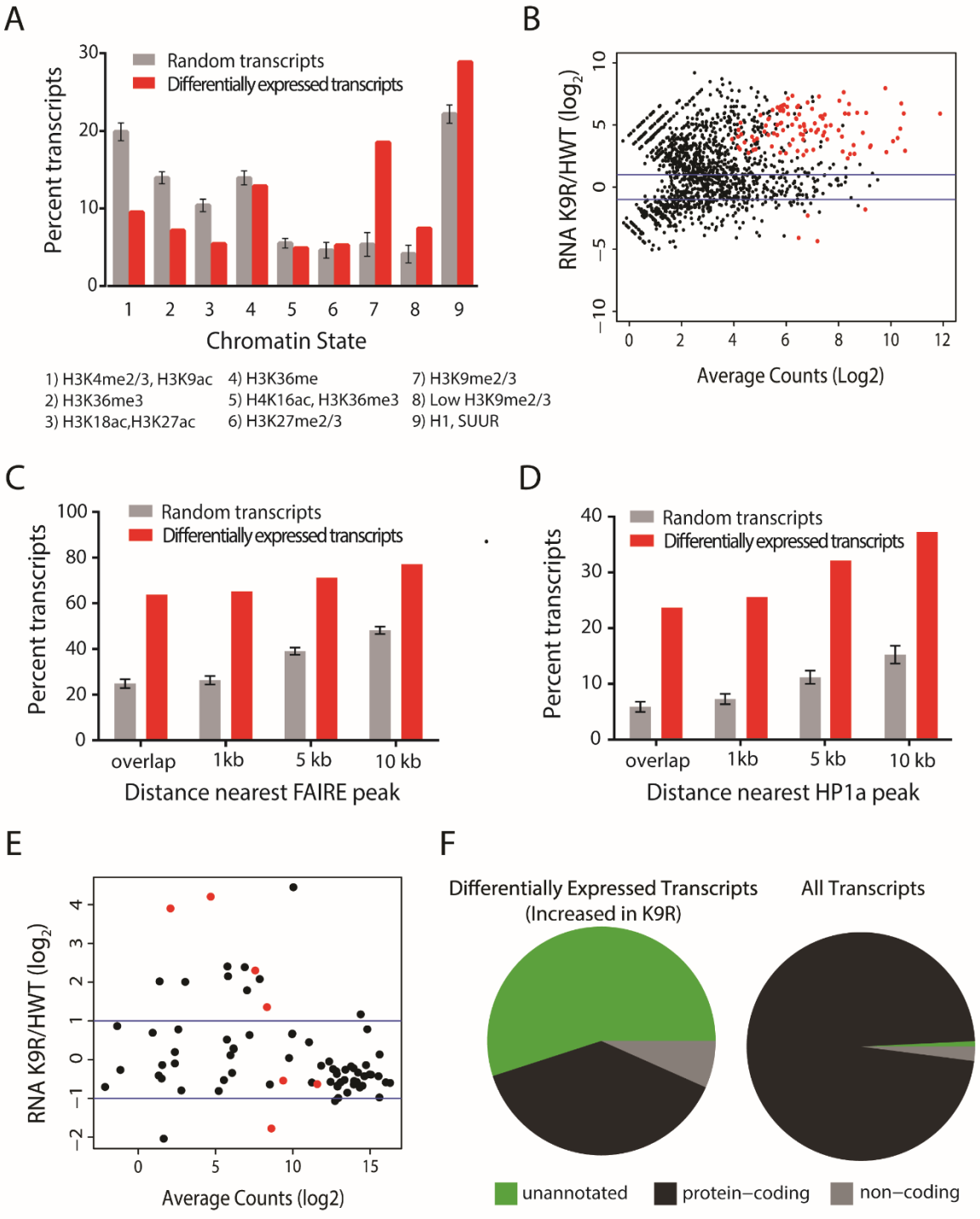

A
RNA K9R/HWT (log2)
Average Counts (log2)
B
Differentially Expressed Transcripts
Annotated transcript
Unannotated transcript
HWT FAIRE CPM (Log2)
K9R FAIRE CPM (log2)
C
Differentially Expressed Transcripts
Annotated transcript
Unannotated transcript
HWT HP1 CPM (Log2)
K9R HP1 CPM (log2)
Unchanged Transcripts
Annotated transcript
Unannotated transcript
HWT HP1 CPM (Log2)
K9R HP1 CPM (log2)

The majority of elevated transcripts in K9R were unannotated in the reference transcriptome but identified in our transcriptome assembly (Figure 2.10F). Surprisingly, significantly fewer than expected protein-coding genes were differentially expressed (Fisher's exact test, $p<10^{-15}$). Some protein-coding genes reside in heterochromatin and require a heterochromatic environment to be properly expressed (Yasuhara and Wakimoto 2006; Lu et al. 2000; Corradini et al. 2007). However, we found only 2 of 46 heterochromatic genes on chromosome 2L (Figure 2.8C) and a small subset of genes on the fourth chromosome (Figure 2.10E) that exhibited altered expression in K9R mutants. The majority of these genes, including the 2 whose expression changed, exhibited increased FAIRE signal and decreased HP1a binding in K9R mutants, demonstrating that these particular changes to chromatin are not sufficient to cause gene expression changes in wing discs. In addition to blocking methylation, the H3K9R mutation also prevents H3K9 acetylation, a modification found near transcription start sites and associated with gene activation (Wang et al. 2008; Kharchenko et al. 2011). However, we found that H3K9ac is not a predictor of altered gene expression in K9R mutants (Figure 2.8D). These data suggest H3K9, in contrast to HP1, is not essential for expression of most protein-coding genes (Grewal and Jia 2007; Elgin and Reuter 2013).

As with our FAIRE-seq and HP1a ChIP-seq data, transcripts differentially expressed between HWT and K9R were preferentially found in regions of heterochromatin (Figure 2.10A). To examine this correlation more closely, we compared our FAIRE-seq and RNA-seq datasets to establish the relationship between nucleosome occupancy and changes in gene expression. A scatter plot of HWT vs K9R FAIRE-seq data identified two classes of transcripts, those with equal FAIRE signal in

the two genotypes and those with higher FAIRE signal in K9R samples. The majority of the second class consisted of unannotated transcripts (Figure 2.9B). These transcripts were enriched in K9R samples (Figure 2.10B) and were composed of simple or interspersed repeats.

Figure 2.10: Changes in gene expression correlate with changes in FAIRE and HP1a signal. A) Transcripts with significantly different expression between HWT and K9R as called by DESeq2 ($p<0.05$) were separated into one of nine chromatin states based on classification in Kharchenko et al. 2010. The percentage of transcripts that fall into each category is compared to the average of 25 iterations of randomly selected transcripts. Error bars indicate standard deviation. B) K9R/HWT ratio of RNA-seq signal for transcripts identified in Cufflinks transcriptome assembly but not in reference transcriptome. Red indicates statistical significance called by DESeq2 ($p<0.05$). C) Shown in red is the percent of differentially expressed transcripts that overlap or are within 1 kb, 5 kb, or 10 kb of a statistically different FAIRE peak between HWT and K9R as determined by edgeR ($p<0.01$). 25 random selections of transcripts were put through the same pipeline and the average percent of transcripts within the various distances are shown in grey. Error bars indicate standard deviation. D) Similar analysis to C was performed with HP1a ChIP-seq peaks. E) K9R/HWT ratio of RNA-seq signal for genes on chromosome four. Red dots indicate statistical significance called by DESeq2 ($p<0.05$). F) Pie charts showing fraction of transcripts that were annotated as protein-coding (black) or non-coding (grey) or were unannotated in reference transcriptome but identified in transcriptome assembly (green). Left pie chart shows all differentially expressed transcripts increased in K9R relative to HWT, and right pie chart shows all transcripts in the newly assembled transcriptome.


A
Random transcripts
Differentially expressed transcripts
Percent transcripts
Chromatin State
1) H3K4me2/3, H3K9ac
2) H3K36me3
3) H3K18ac,H3K27ac
4) H3K36me
5) H4K16ac, H3K36me3
6) H3K27me2/3
7) H3K9me2/3
8) Low H3K9me2/3
9) H1, SUUR
B
RNA K9R/HWT (log2)
Average Counts (Log2)
C
Percent transcripts
overlap
1kb
5 kb
10 kb
Distance nearest FAIRE peak
D
Percent transcripts
Distance nearest HP1a peak
E
RNA K9R/HWT (log2)
Average Counts (log2)
F
Differentially Expressed Transcripts
(Increased in K9R)
All Transcripts
unannotated
protein–coding
non–coding

Similarly, in K9R mutants, HP1a was depleted at the majority of differentially expressed transcripts (Figure 2.9C). This result is specific, as HP1a signal was essentially equivalent at a random selection of transcripts that do not significantly differ between K9R and HWT (Figure 2.9C). Moreover, differentially expressed transcripts were more likely to be near distinct HP1a and FAIRE regions than were a random selection of transcripts (Figure 2.10C,D). The specificity of these transcripts to heterochromatic regions and their correlation to altered nucleosome occupancy and HP1a localization suggests H3K9 plays a direct role in their regulation.

H3K9 represses transposon activation and mobilization

The repetitive nature of the unannotated transcripts identified in the preceding analyses prompted us to look more specifically at transposons. In *Drosophila*, over 130 different transposon families have been identified, with most families consisting of multiple insertions throughout the genome (Rahman et al. 2015). This repetition complicates mapping reads and consequently the analysis of high-throughput sequencing data. We therefore used the piPipes program, which circumvents these difficulties by mapping reads to transposon families instead of unique transposon insertions in the genome (Han et al. 2015). In both FAIRE- and RNA-seq datasets, transposons showed higher signal in K9R samples compared to HWT, whereas control euchromatic regions or coding RNAs were largely equivalent between K9R and HWT (Figure 2.11A,B). Transcripts from piRNA clusters can be processed into small piRNAs or endo-siRNAs that function with Argonaute proteins to repress transposons (Brennecke et al. 2007; Czech et al. 2008). We therefore examined FAIRE and RNA signatures at piRNA clusters. Similar to transposon analyses, both datasets showed

increased signal in K9R samples relative to HWT (Figure 2.11A-C). We next examined HP1a localization at transposon families and piRNA clusters using a similar strategy. Whereas transposons and piRNA clusters from HWT samples showed HP1a enrichment, K9R samples lacked detectable signal above input (Figure 2.11D,E, 2.12A). Transposon activation in K9R mutants is therefore correlated with an increase in FAIRE signal and a decrease in HP1a localization at both transposons and piRNA clusters.

Figure 2.11: H3K9 represses transposon and piRNA cluster expression A) Scatterplot of FAIRE-seq signal (RPKM: reads per kilobase per million) at individual transposon families (red) or piRNA clusters (blue). H3K4me2/me3 enriched promoter regions (grey) are shown for comparison. B) Scatterplot of RNA-seq signal at transposons (red) and piRNA clusters (blue). A random selection of protein coding RNAs was selected for comparison (grey). C) Genome browser shot of FAIRE and RNA signal at the 42AB piRNA cluster. FAIRE signal is shown for both uniquely and multiple mapping reads. RNA signal contains uniquely mapping reads only. Highlighted areas indicate regions of increased FAIRE and RNA signal in K9R samples. D-E) Scatterplot of HP1a ChIP-seq signal at transposon families (D) or piRNA clusters (E) comparing input and immunoprecipitated (IP) samples.

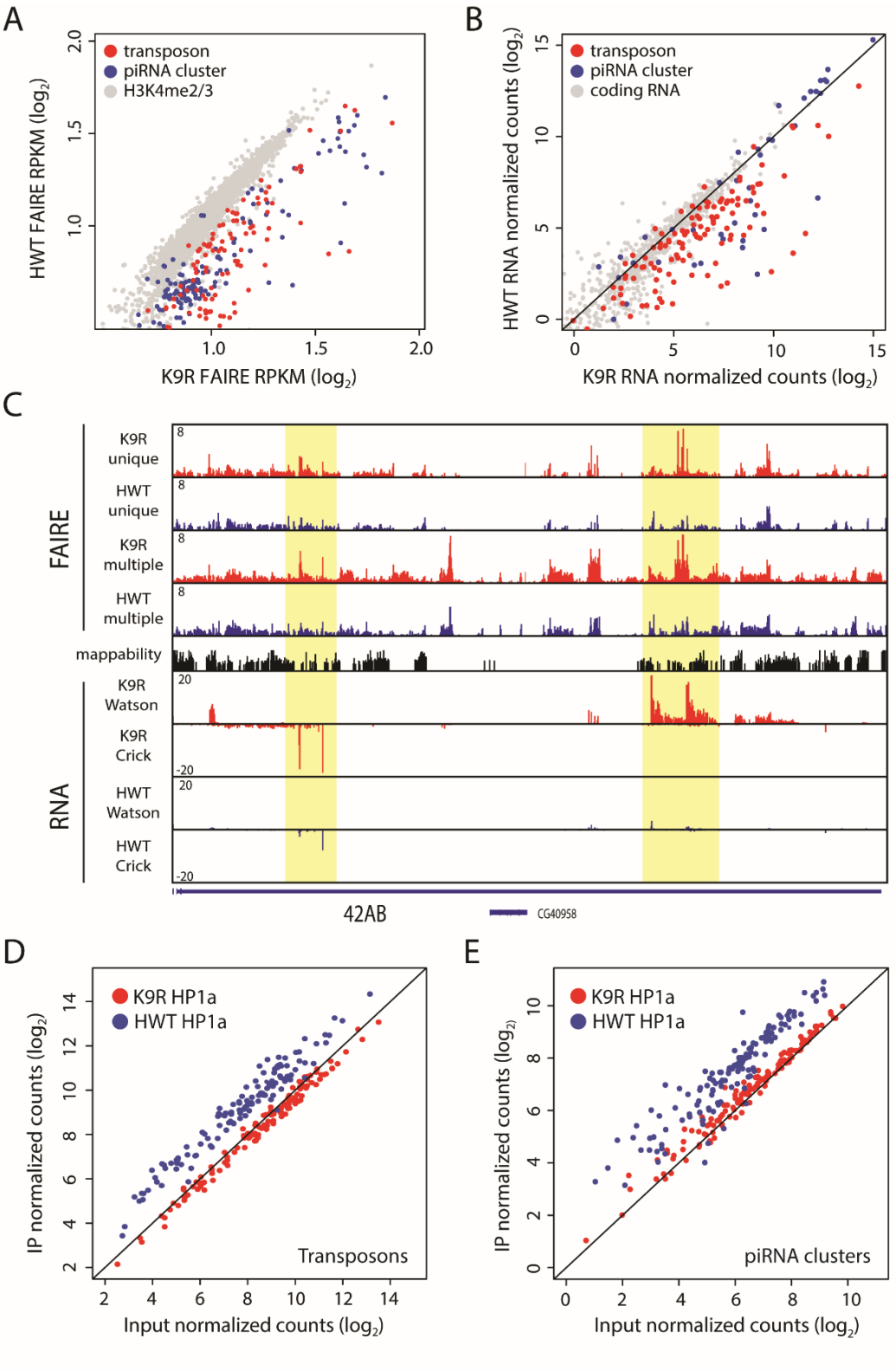
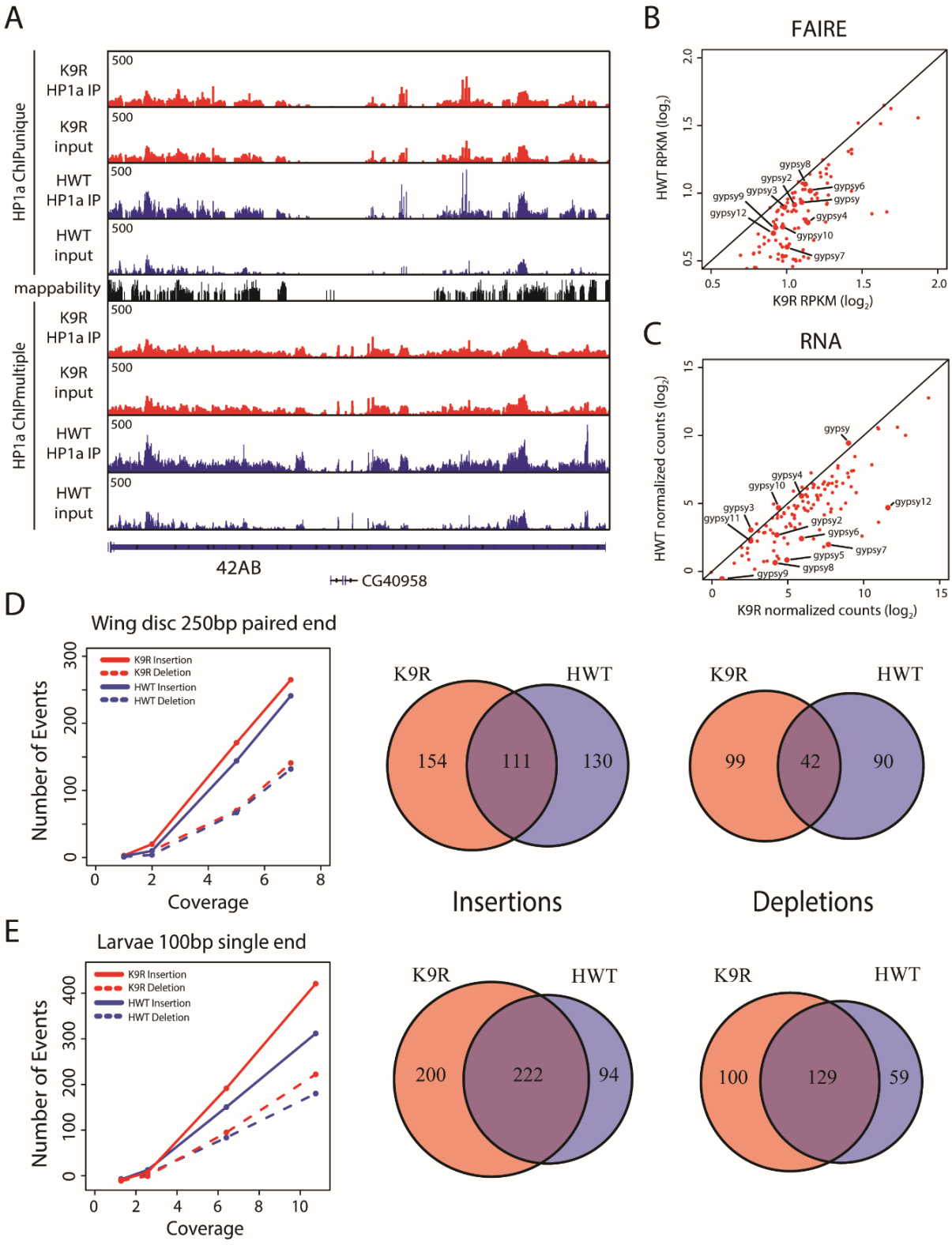

A
transposon
piRNA cluster
H3K4me2/3
HWT FAIRE RPKM (log2)
K9R FAIRE RPKM (log2)
B
transposon
piRNA cluster
coding RNA
HWT RNA normalized counts (log2)
K9R RNA normalized counts (log2)
C
FAIRE
K9R unique
HWT unique
K9R multiple
HWT multiple
mappability
RNA
K9R Watson
K9R Crick
HWT Watson
HWT Crick
42AB
CG40958
D
K9R HP1a
HWT HP1a
IP normalized counts (log2)
Input normalized counts (log2)
Transposons
E
K9R HP1a
HWT HP1a
IP normalized counts (log2)
Input normalized counts (log2)
piRNA clusters

Figure 2.12: Transposons are activated and mobilized in K9R mutants. A) Genome browser shot of sequencing reads mapping to the 42AB piRNA cluster from anti-HP1a immunoprecipitated (IP) chromatin samples compared to input samples. Top rows show uniquely mapping reads (MAPQ>10) while bottom rows show multiple mapping reads randomly distributed to one of a maximum of 10 possible assignments. HP1a ChIP-seq reads are enriched in the HWT sample (blue) but not the K9R sample (red). Signal represented as reads per million (RPM). B,C) Transposon FAIRE-seq and RNA-seq data from Figure 2.11A and B with various *gypsy* transposon families labeled. D,E) Whole genome sequencing data from wing disc (D) or larvae (E) was analyzed for transposon insertion or depletion events using TIDAL (Rahman et al. 2015). To examine the effect of genome coverage on detection we randomly subsampled reads from each genotype. Scatterplot shows the number of insertions (solid line) or depletions (dashed line) at various fold coverages (x axis) in K9R (red) and HWT (blue) samples. Venn diagrams indicate the number of insertions or depletions unique to or shared by each genotype.

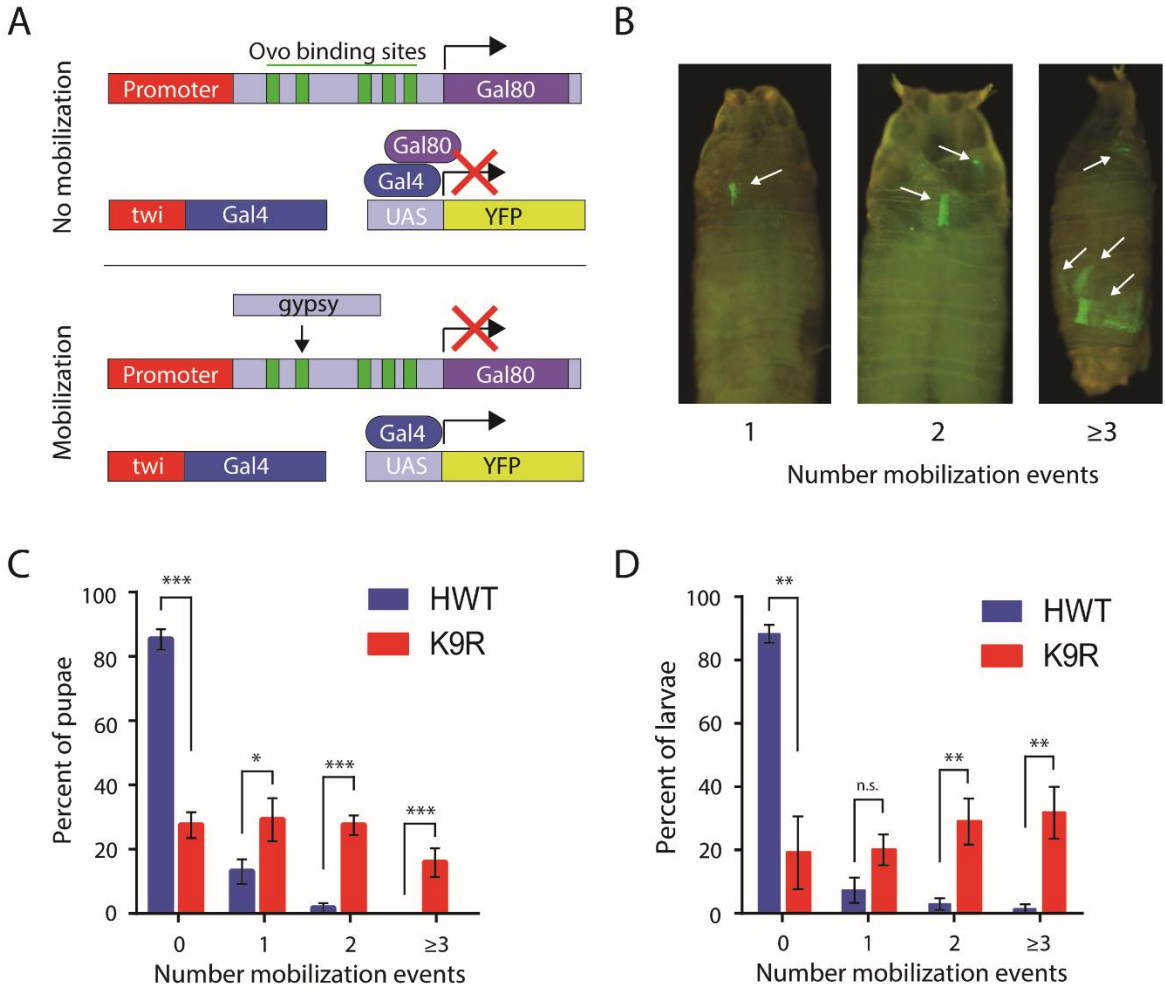

A
HP1a ChIPunique
K9R HP1a IP
K9R input
HWT HP1a IP
HWT input
mappability
HP1a ChIPmultiple
K9R HP1a IP
K9R input
HWT HP1a IP
HWT input
500
42AB
CG40958
B
FAIRE
HWT RPKM (log2)
K9R RPKM (log2)
gypsy8
gypsy2
gypsy6
gypsy3
gypsy9
gypsy
gypsy12
gypsy4
gypsy10
gypsy7
C
RNA
HWT normalized counts (log2)
K9R normalized counts (log2)
gypsy
gypsy4
gypsy10
gypsy12
gypsy3
gypsy11
gypsy2
gypsy6
gypsy7
gypsy5
gypsy8
gypsy9
D
Wing disc 250bp paired end
Number of Events
Coverage
K9R Insertion
K9R Deletion
HWT Insertion
HWT Deletion
K9R
HWT
154
111
130
Insertions
99
42
90
Depletions
E
Larvae 100bp single end
Number of Events
Coverage
K9R Insertion
K9R Deletion
HWT Insertion
HWT Deletion
K9R
HWT
200
222
94
100
129
59

Finally, the de-repression of transposons in K9R mutants suggests transposons may be activated, resulting in higher frequency of mobilization. To test this hypothesis, we took advantage of a gypsy-TRAP reporter that activates GFP expression upon de novo integration of *gypsy* transposons into a reporter transgene (Figure 2.13A, Li et al. 2013). Cell proliferation subsequent to de novo *gypsy* integration results in clones of GFP+ cells. GFP expression has previously been shown to coincide with molecularly confirmed *gypsy* integration events (Li et al. 2013). We therefore counted the number of GFP+ clones in pupae (Figure 2.13C) or larvae (Figure 2.13D) to assess the relative number of transposition events between K9R and HWT. The fraction of pupae that contained a transposition event in K9R animals was 75% compared to only 15% in HWT (Figure 2.13C). Similar results were obtained in larvae (Figure 2.13D). Additionally, significantly more K9R animals had multiple transposition events. These results demonstrate that *gypsy* transposons mobilize at an elevated frequency in K9R animals relative to controls. (Herz et al. 2014) also observed activation of a *gypsy-lacZ* reporter after expression of H3K9M, which likely inhibits H3K9 methyltransferases via sequestration (2014).

We also searched for evidence of transposon insertion and depletion events in our whole genome sequencing datasets using the program TIDAL, which uses split read analysis to identify junction reads that span both transposon and unique sequences (Rahman et al. 2015). We examined input DNA from both our wing disc FAIRE experiment and our whole larvae HP1a ChIP experiment. In wing discs, TIDAL detected a similar number of insertion and depletion events in K9R and HWT samples (154 vs 130 insertions and 99 vs 90 depletions, respectively; Figure 2.12D). In whole larvae,

however, we detected a two-fold increase in insertion and depletion events in K9R samples relative to HWT (200 vs 94 insertions and 100 vs 59 depletions, respectively; Figure 2.12E). Combined, these data support a model in which the absence of H3K9 prevents recruitment of HP1a and the formation of a repressive chromatin environment, leading to the activation and mobilization of transposons.

Figure 2.13: H3K9 represses gypsy transposon mobilization A) Schematic of the *gypsy* transposon mobilization assay. Ovo binding site (green bars) -dependent *gypsy* insertion disrupts Gal80 resulting in Gal4-directed expression of yellow fluorescent protein (YFP) (Li et al. 2013). B) Pupae showing 1, 2, or ≥3 YFP-positive clones representing *gypsy* mobilization events, which are limited to the mesoderm due to the twi-Gal4 driver and thus likely underestimate the total number of mobilization events in each animal. C,D) Histogram of the average number of pupae (C; n=200 for each genotype in 6 independent experiments) or larvae (D; n=100 for each genotype in 3 independent experiments) with 0, 1, 2, and 3 or more mobilization events. Error bars represent standard deviation (* $p<0.05$, ** $p<0.005$, *** $p<0.0005$).


A
No mobilization
Ovo binding sites
Promoter
Gal80
Gal80
Gal4
twi
Gal4
UAS
YFP
Mobilization
gypsy
Promoter
Gal80
Gal4
twi
Gal4
UAS
YFP
B
1
2
≥3
Number mobilization events
C
Percent of pupae
HWT
K9R
***
*
***
***
0
1
2
≥3
Number mobilization events
D
Percent of larvae
HWT
K9R
**
n.s.
**
**
0
1
2
≥3
Number mobilization events

Discussion

In diverse organisms from fission yeast to humans, recruitment of HP1 by H3K9me2/me3 serves as a paradigm for histone PTM-mediated chromatin regulation. Here, we have performed a set of genetic, cytological, and whole genome sequencing experiments in *Drosophila* to directly interrogate the role of H3K9 in animal development.

H3K9 and metazoan development

H3K9me function is typically inferred through H3K9 methyltransferase loss-of-function studies, including by mutation or sequestration of the enzymes (Herz et al. 2014). However, such studies are unable to directly test the contribution of H3K9 to heterochromatin biology due to complexities and redundancies among H3K9 methyltransferases, as well as their ability to methylate non-histone substrates. Although the three *Drosophila* methyltransferases, Su(var)3-9, dG9a, and SetDB1/Eggless, generally function at different regions of the genome, the enzymes are partially redundant and show complex genetic interactions. *Su(var)3-9* and *dG9a* single mutants are viable and fertile, but double mutants have reduced viability (Schotta et al. 2003b; Mis et al. 2006). Interestingly, *SetDB1* single mutants display reduced viability and fertility, and a loss of Su(var)3-9 in this background counterintuitively increases viability (Brower-Toland et al. 2009; Seum et al. 2007). To our knowledge, triple mutants have not been studied, and would be technically challenging to create. These complexities have hindered our understanding of the true function of H3K9 methylation. Although the K9R mutation that we have engineered in histone H3 cannot specifically

test the role of K9 methylation, K9R mutant flies nonetheless provide an important new resource for studying metazoan heterochromatin biology.

Given that many previous studies have concluded that H3K9me is essential for heterochromatin establishment and function, preventing this modification should have severe developmental consequences. Indeed, the viability of H3K9R mutants is greatly reduced. However, we were surprised to find that H3K9 is not absolutely essential for the completion of *Drosophila* development. Notably, *C. elegans* H3K9 methyltransferase mutants that lack detectable H3K9 methylation are both viable and fertile (Towbin et al. 2012). One possibility is that modification of *Drosophila* H3.3K9 compensates for the absence of H3K9. The converse may also be true as H3.3K9R mutants are viable in *Drosophila* (Sakai et al. 2009). However, our data indicate that H3.3 does not compensate for H3 in pericentric heterochromatin, as HP1 and nucleosomes are depleted from pericentric heterochromatin and we were unable to detect H3K9me2 or me3 signal at the chromocenter of polytene chromosomes. These data are in concordance with previous results showing that H3.3 is depleted from heterochromatin in wild-type flies (Ahmad and Henikoff 2002). In other organisms, there may be more redundancy between H3.3 and H3 at pericentric heterochromatin. In mouse embryonic stem cells, H3.3 is present at telomeres as well as pericentric repeats (Lewis et al. 2010; Goldberg et al. 2010) and is necessary for silencing a subset of transposable elements (Elsässer et al. 2015).

H3K9 regulation of chromatin architecture and composition

Recent studies have illuminated the diversity of heterochromatin that exists throughout the genome (Haynes et al. 2007; Riddle et al. 2011; Wang et al. 2014). These different types of heterochromatin may rely on H3K9me to differing extents. Indeed, in *Drosophila* embryos, H3K9me plays less of a role in heterochromatin formation at the 359bp repeat than at other heterochromatic regions (Yuan and Farrell 2016). Additionally, HP1 levels at certain promoters (Figueiredo et al. 2012) and recruitment to telomeres (reviewed in Raffa et al., 2011) are reported to be independent of H3K9me. Accordingly, we detected HP1a association with telomeres in K9R mutants (Fig. 3A).

Our cytological data from polytene chromosomes and genome-wide analyses of diploid cells both indicate a critical role for H3K9 in regulating chromatin architecture of the pericentromeres and to a lesser extent the fourth chromosome. Multimerization of HP1 molecules on neighboring or inter-strand nucleosomes has been proposed to mediate higher-order chromatin folding (Canzio et al. 2011; Azzaz et al. 2014). Loss of HP1a may therefore lead to unfolding of condensed structures and result in a more open chromatin environment. Such decondensation might explain the disorganized chromocenter we observed in polytene chromosome spreads and could allow nucleosome remodelers or transcription factors access to previously inaccessible chromatin, leading to the changes in nucleosome occupancy we measured by FAIRE. HP1 has also been implicated in regulating replication timing (Quivy et al. 2008; Schwaiger et al. 2010). Pericentromeric regions in polyploid salivary glands are under-replicated and disruption of this process might also contribute to the disorganized chromocenters we observed in K9R polytene chromosomes.

In both cytological and ChIP-seq experiments the loss of HP1a from pericentromeric regions in K9R mutants was accompanied by an increase in HP1a along chromosome arms. Previous studies demonstrated that HP1a normally associates with numerous euchromatic sites on chromosome arms (De Wit et al. 2007) in an RNA-dependent manner (Piacentini et al. 2003). Additionally, other HP1 isoforms, including *Drosophila* HP1c and mammalian HP1ɣ, are primarily euchromatic and found at transcriptionally active domains (Kwon and Workman 2011). An intriguing possibility is that H3K9 mediated binding of HP1a to pericentromeric regions may prevent its spread to inappropriate areas such as regions of HP1c enrichment. Along these lines, the Y chromosome's capacity to act as a suppressor of variegation has been attributed to its potential function as a sink for a limited pool of HP1 (Dorer and Henikoff 1994).

Our metagene analysis revealed that the increase of HP1a on chromosome arms occurs along gene bodies (Figure 2.7D). The presence of HP1 at some of these sites is required for proper expression of genes within these regions (Cryderman et al. 2005) and may be involved in RNA processing or transcriptional elongation (Piacentini et al. 2009; Vakoc et al. 2005). Although we identified a handful of cases in K9R mutants where ectopic HP1a signal correlated with antisense transcription in the middle of a gene, the majority of genes showed no change in expression levels (not shown).

H3K9 regulation of gene expression and transposons

One of our most striking observations was an increase in transposon-derived RNA in K9R mutants that is associated with an increase in open chromatin and loss of HP1 at almost all transposon families. Since transposons make up 15-20% of the

Drosophila melanogaster genome, loss of H3K9 has a substantial impact on global chromatin organization (Kaminker et al. 2002). HP1 is thought to limit RNA polymerase occupancy by facilitating formation of closed chromatin structures or recruiting additional silencing factors (Grewal and Jia 2007; Danzer and Wallrath 2004). The absence of HP1 and increased DNA accessibility could therefore permit inappropriate transcription. Active transposons can lead to illegitimate recombination, DNA breaks, and the disruption of genes (Slotkin and Martienssen 2007; Levin and Moran 2011). Organisms therefore have mechanisms that repress transposons, including the piRNA pathway predominantly active in germline cells, and the endo-siRNA pathway, which functions primarily in non-gonadal somatic cells (Brennecke et al. 2007; Czech et al. 2008; Kawamura et al. 2008; Ghildiyal and Zamore 2008). In both pathways, small RNAs direct Argonaute family proteins to silence targets (Mani and Juliano 2013; Senti and Brennecke 2010; Saito and Siomi 2010). These small RNAs are in part derived from piRNA clusters, which are composed of numerous inactive fragments of transposons that can be hundreds of kilobases long. In K9R mutant wing discs, transposon activation is correlated with an increase in the levels of piRNA clusters. Moreover, HP1a association with both transposons and piRNA clusters is reduced in K9R mutants. Because the endo-siRNA pathway is predominantly active in non-gonadal somatic cells, we speculate that the absence of H3K9 disrupts this pathway. Previous studies have demonstrated mutations in H3K9 methyltransferases and HP1 result in transposon activation (Brower-Toland et al. 2009; Lundberg et al. 2013). Increased piRNA levels could therefore be a response to this activation. Interestingly, the germline requires the H3K9 methyltransferase SetDB1 and the HP1 family protein Rhino for

piRNA cluster transcription (Rangan et al. 2011; Klattenhoff et al. 2009; Mohn et al. 2014). Other mechanisms for endo-siRNA production likely exist in non-gonadal cells as Rhino is not expressed outside the germline. Our results suggest neither methylation of H3K9 nor HP1a are necessary for piRNA expression in somatic cells. Alternatively, H3K9 and HP1a may be required for correct processing of piRNA cluster transcripts into endo-siRNAs. Preventing this processing may lead to a buildup of piRNA cluster transcripts and a decrease in endo-siRNAs available for transposon silencing.

We were surprised to find that protein-coding gene expression was similar between HWT and K9R mutants. In particular, expression of most protein-coding genes residing within heterochromatin on chromosome 2L was unchanged, suggesting H3K9 does not play a critical role in regulating expression of these genes. However, because these data were derived from wing discs only, H3K9 may regulate heterochromatin protein coding genes in other tissues. Expression of protein-coding genes associated with H3K9ac was also relatively unchanged by mutating H3K9. One possibility is that the major function of H3K9ac is to prevent H3K9me and subsequent gene repression. Thus, in K9R mutants the absence of H3K9ac may be inconsequential in the absence of H3K9me. Alternatively, other acetylation marks may compensate for the loss of H3K9ac, including H3.3K9ac.

Our studies demonstrate an important role for H3K9 in the regulation of pericentromeric chromatin architecture and the maintenance of transposon repression. Future analyses of other heterochromatin associated processes including mitosis, DNA replication, and DNA damage in K9R mutants will be of great interest. Given the conservation of H3K9 methyltransferases and HP1 proteins, our observations should be informative to studies of heterochromatin biology in other eukaryotes.

CHAPTER 3- FUNCTIONAL REDUNDANCY OF VARIANT AND CANONICAL HISTONE H3 LYSINE 9 MODIFICATION IN *DROSOPHILA*[3]

Introduction

DNA interacts with histones and other proteins to establish chromatin environments that affect all DNA-dependent processes. The establishment of chromatin environments is accomplished through multiple mechanisms that collectively comprise the bulk of epigenetic regulation found in eukaryotes. In particular, post-translational modification (PTM) of histones influences DNA/histone interactions and also provides binding sites for recruitment of chromatin modulators that influence gene expression, DNA replication and repair, and chromosome segregation during cell division (Wallrath et al. 2014). In addition to histone PTMs, epigenetic regulation is modulated by the type of histone protein deposited onto DNA. There are two major categories of histone proteins: the canonical histones and the closely related histone variants (Talbert and Henikoff 2010, 2017). These two histone categories are distinguished by the timing of their expression during the cell cycle and their mechanism of deposition onto DNA. Canonical histones are encoded by multiple genes (e.g., ~55 in humans and ~500 in flies), organized into clusters that are highly expressed during S-phase of the cell cycle, and are deposited onto DNA by the histone chaperone CAF-1 in a replication-coupled

[3] This chapter previously appeared as an article in Genetics. The original citation is as follows:
Penke, T.J., McKay, D.J., Strahl, B.D., Matera, A.G. and Duronio, R.J., 2017. Functional redundancy of variant and canonical histone H3 lysine 9 modification in Drosophila. *Genetics.* Accepted.

manner (Marzluff et al. 2002; Tagami et al. 2004; Verreault et al. 1996). In contrast, variant histones are typically encoded by one or two genes, are expressed throughout the cell cycle, and can be deposited onto DNA independently of replication by histone chaperones other than CAF-1 (Henikoff and Ahmad 2005; Tagami et al. 2004; Szenker et al. 2011). Variant histones are often deposited at specific genomic locations and have functions that can differ from canonical histones. For example, two histone H2A variants, H2AX and H2A.Z, play critical roles in DNA repair (Scully and Xie 2013; Price and Andrea 2014), and the histone H3 variant CENP-A localizes to centromeres and is essential for kinetochore formation (Blower and Karpen 2001; Henikoff and Ahmad 2005; Mellone and Allshire 2003).

The major histone H3 variant in animal genomes is H3.3, which in both mice and *Drosophila* is encoded by two different genes (*H3.3A* and *H3.3B*) that produce identical proteins. Variant histone H3.3 differs from canonical H3.2 and H3.1 by only four or five amino acids, respectively (Szenker et al. 2011). In each case, three of these different amino acids are located in the globular domain of H3.3 and are necessary and sufficient for interaction with the replication-independent chaperones HIRA and ATRX-DAXX (Tagami et al. 2004; Goldberg et al. 2010; Ahmad and Henikoff 2002; Lewis et al. 2010). In H3.2, the only replication-dependent histone in *Drosophila*, the fourth amino acid difference occurs at position 31 in the unstructured N-terminal tail (Szenker et al. 2011). Histones H3.2 and H3.1 (collectively hereafter referred to as H3) along with H3.3 are some of the most conserved proteins in all eukaryotes (Malik and Henikoff 2003). The conservation of amino acid differences between H3 and H3.3 during evolution strongly suggests that these proteins perform distinct functions. Indeed, H3.3 and H3

are deposited in different genomic regions in a variety of species (Mito et al. 2005; Schwartz and Ahmad 2005; Tamura et al. 2009; Jin et al. 2011; Kraushaar et al. 2013; Allis and Wiggins 1984). H3.3 is also enriched for different histone PTMs than H3 (Hake et al. 2006; McKittrick et al. 2004), and H3.3 containing nucleosomes can be less stable than those with H3 (Jin and Felsenfeld 2007; Xu et al. 2010). Although the epigenetic PTM signature on variant and canonical H3 histones is distinct, the degree to which particular histone PTMs found on both H3 and H3.3 can compensate for one another is not fully understood. Here, we explore the common and distinct functions of variant and canonical H3K9 function during *Drosophila* development.

H3.3 is associated with transcriptionally active regions of the genome with high nucleosome turnover, consistent with H3.3 being enriched in “activating” histone PTMs and depleted in “repressing” histone PTMs (Hake et al. 2006; McKittrick et al. 2004). One of the histone PTMs enriched on H3.3 relative to H3 is acetylation of lysine nine (K9ac), a mark associated with accessible chromatin (Hake et al. 2006; McKittrick et al. 2004). Previous studies have identified K9ac at promoters of genes and in regions of high transcriptional activity (Kharchenko et al. 2011; Bernstein et al. 2005; Liang et al. 2004; Roh et al. 2005). Additionally, mutation of H3K9 acetyltransferases results in compromised transcriptional output, suggesting K9ac contributes to or is a consequence of gene expression activation (Wang et al. 1998; Georgakopoulos and Thireos 1992; Kuo et al. 1998). Importantly, H3K9 acetyltransferases target other histone residues and have non-histone substrates as well (Glozak et al. 2005; Spange et al. 2009), indicating that one cannot deduce the function of K9ac solely by mutation of H3K9 acetyltransferases. For example, whereas mutation of the H3K9

acetyltransferase Rtt109 in budding yeast results in sensitivity to DNA-damaging agents, H3K9R mutants, which cannot be acetylated by Rtt109 are insensitive to DNA-damaging agents (Fillingham et al. 2008). Direct investigation of K9ac function in vivo therefore requires mutation of H3K9 itself. Previously, we used a *Drosophila* histone gene replacement platform (McKay et al. 2015) to generate a canonical $H3^{K9R}$ mutant, and found no significant changes in gene expression at regions of the genome enriched in K9ac (Penke et al. 2016). This observation raises the possibility that H3.3K9ac functions in gene regulation and can compensate for the absence of H3K9ac.

H3.3 is also found at transcriptionally inactive, heterochromatic regions of the genome (Goldberg et al. 2010; Lewis et al. 2010; Wong et al. 2010). Heterochromatin is enriched in H3K9 di- and tri-methylation (me2/me3), modifications that recruit Heterochromatin Protein 1 (HP1) and are essential for heterochromatin function (Bannister et al. 2001; Lachner et al. 2001; Nakayama et al. 2001; Penke et al. 2016). DNA within heterochromatin is composed of repeated sequence elements, many of which are transcriptionally silent and consist of immobile transposons or transposon remnants. Using mutations in the chaperones that deposit H3.3 onto DNA, it was recently demonstrated that H3.3 is essential for repression of endogenous retroviral elements and that H3.3 can be methylated at lysine nine (Elsässer et al. 2015). H3.3K9me3 is also important for heterochromatin formation at mouse telomeres (Udugama et al. 2015). These studies did not assess the contribution of canonical H3K9 because strategies for mutating all replication-dependent H3 genes in mammalian cells have not been developed. We recently showed in *Drosophila* that mutation of canonical H3K9 causes defects in heterochromatin formation and transposon repression (Penke

et al. 2016), similar to phenotypes observed in *C. elegans* in the absence of H3K9 methyltransferases (Zeller et al. 2016). In addition, we detected low levels of K9me2/me3 in $H3^{K9R}$ mutants. Combined, these data suggest methylated H3.3K9 plays a role in heterochromatin formation and can compensate for the absence of canonical H3K9. However, the extent of functional overlap between variant and canonical H3K9 and the intriguing possibility that identical modifications on variant or canonical histones have distinct functions has yet to be fully investigated.

In order to better understand the functions of H3 and H3.3 and to compare the functions of the variant and canonical H3K9 residues, we used CRISPR-Cas9 to generate a variant K9R substitution mutation ($H3.3^{K9R}$) in *Drosophila* and combined this with our previously described canonical $H3^{K9R}$ mutant (Penke et al. 2016). By comparing the individual mutant phenotypes of $H3^{K9R}$ and $H3.3^{K9R}$ to the combined $H3.3^{K9R}$ $H3^{K9R}$ mutants using a variety of genomic and cell biological assays, we demonstrate that variant and canonical versions of H3K9 can compensate for each other, although to substantially different extents in euchromatin versus heterochromatin. H3K9 plays a more substantial role than H3.3K9 in heterochromatin formation and in the repression of transposons, whereas they compensate for each other in controlling euchromatic gene expression, particularly in regions enriched in the activating modification, K9ac.

Materials and Methods

Generation of K9R mutant genotypes: Variant $H3.3^{K9R}$ mutants generated by the cross scheme illustrated in Figure 3.1A were selected by the absence of GFP fluorescence and/or the presence of straight wings. 1st instar larvae from the variant and canonical

H3.3^{K9R} H3.2^{K9R} cross described in Figure 3.1B were selected based on the presence of GFP fluorescence. Only larvae that receive the H3^{HWT} or H3^{K9R} transgene will survive embryogenesis as this transgene provides the only source of canonical histone genes. In table 2 rows one and two indicate progeny from the cross *yw ; H3.3^{2x1} / CyO, twiGFP x yw ; Df(2L)BSC110 / CyO, twiGFP*. Rows three and four indicate progeny from the cross *H3.3B^{K9R} ; H3.3^{2x1} / CyO, twiGFP x H3.3B^{K9R} ; Df(2L)BSC110 / CyO, twiGFP*. In these crosses, the expected ratio of heterozygous to homozygous *H3.3A^{2x1}* animals is 2:1, as *CyO, twiGFP/CyO, twiGFP* animals do not eclose as adults.

CRISPR-Cas9 Mutagenesis and Transgene Integration**:** A single gRNA targeting *H3.3B* near the K9 residue was inserted into pCFD3 and co-injected with a 2 kb homologous repair template containing the H3.3BK9R substitution. Constructs were injected into embryos expressing Cas9 from the *nanos* promoter (*nanos-cas9*; Kondo and Ueda 2013). Recovered *H3.3B^{K9R}* alleles were subsequently crossed into *H3.3A* null backgrounds. To generate *H3.3B* rescue constructs, a 5 kb genomic sequence containing the entire wild-type *H3.3B* transcription unit was PCR amplified from genomic DNA of *nanos-cas9* flies and cloned into pATTB. Gibson assembly (Gibson et al. 2009) using primers containing K9R or K9Q substitutions was used to generate mutated versions of *H3.3B*, and all three constructs were integrated into the 86FB *attP* landing site by ΦC31-mediated recombination.

Immunofluorescence: Salivary gland preparations stained using anti-H3K9me2, anti-H3K9me3, anti-H3K9ac, or anti-HP1a were performed as previously described (Cai et al. 2010). 1st instar larval brains were prepared similar to imaginal wing disc preparations described in Estella et al. (2008).

Western Blots: ImageJ densitometry analysis was used to determine K9me2, K9ac, or H3 band intensity. Histone modification signal was normalized to corresponding H3 loading control signal. Normalized signal from different titrations of the same genotype were averaged and consequent values were set relative to WT value. This process was completed for two biological replicates for both K9me2 and K9ac.

Sample Preparation and Sequence Data Analysis: FAIRE-seq and RNA-seq samples were prepared from wandering 3rd instar imaginal wing discs as previously described (McKay and Lieb 2013). Sequencing reads were aligned to the dm6 (6.04) reference genome using Bowtie2 (FAIRE) and Tophat (RNA) default parameters (Langmead and Salzberg 2012; Trapnell et al. 2014). FAIRE peaks were called with MACS2 using a shift size of 110bp and a stringency cutoff of 0.01 (Zhang et al. 2008). Transcripts were assembled with Cufflinks (Trapnell et al. 2014). Bedtools was used to determine read coverage at peaks and transcripts (Quinlan and Hall 2010) and DESeq2 was used to determine statistical significance ($p<0.05$) (Love et al. 2014b). The following modENCODE 3rd instar larval ChIP-seq data sets were used: K9me2=GSE47260, and K9me3=GSE47258. K9ac ChIP-seq data from imaginal wings discs was generated by Pérez-Lluch et al. (GSM1363590, 2015).

Chromatin state analysis was performed using data from Kharchenko et al (2010), which assigns small regions of the genome into one of nine different chromatin state. FAIRE peaks were classified as one or more chromatin states based on overlap with regions defined by Kharchenko et al. (2010). Of all the peaks in a particular chromatin state, we determined the percentage of peaks that had significantly different

FAIRE signal in mutant compared to WT samples. RNA chromatin state analysis was performed in a similar fashion.

Results

H3.3^{K9R} mutant animals are viable but sterile

In order to investigate the role of H3.3K9 in *Drosophila* development and compare it to the role of H3K9, we first generated an H3.3^{K9R} animal by introducing a K9R substitution at the endogenous *H3.3B* locus using CRISPR/Cas9 and then combining recovered *H3.3B^{K9R}* mutant alleles with a previously generated *H3.3A* null allele (H3.3A/B combined genotype denoted hereafter as *H3.3^{K9R}*; see Table 3.1 and Figure 3.1A for histone genotype nomenclature) (Sakai et al. 2009). These *H3.3^{K9R}* mutants, which contain the full complement of endogenous canonical *H3* genes, eclose as adults at the expected Mendelian ratios (Table 3.2) and appear morphologically normal. Therefore, canonical H3 can provide all of the H3K9 function during *Drosophila* development. This result is consistent with a previous study finding that flies without any H3.3 protein could be propagated as a stock if canonical H3.2 was expressed from a transgene using the *H3.3B* promoter (Hödl and Basler 2012). Our results are also in line with a previous report in which *H3.3A* and *H3.3B* null animals containing an *H3.3A^{K9R}* transgene were viable (Sakai et al. 2009). However, whereas these *H3.3A^{K9R}* transgenic animals were fertile (Sakai et al. 2009), we found that animals with an endogenous *H3.3B^{K9R}* mutation and the same *H3.3A* null allele used by Sakai et al. (2009) were sterile. The sterility of our *H3.3^{K9R}* animals was rescued in both males and females by a transgene containing the wild-type *H3.3B* gene ectopically integrated into

the genome, suggesting that the relative abundance of H3.3^{K9R} causes sterility (Figure 3.2). We conclude that H3.3K9 plays an essential role during gametogenesis and speculate that different amounts of H3.3^{K9R} histones from *H3.3A* or *H3.3B* promoters may account for the differences between our observations and those of Sakai et al. (2009).

	Canonical		Variant	
Genotype	**Endogenous**	**Transgenic**	**H3.3B**	**H3.3A**
WT	WT	-	WT	WT
H3.3B^{K9R}	WT	-	K9R	WT
H3.3A^{Null}	WT	-	WT	Δ
H3.3^{K9R}	WT	-	K9R	Δ
HWT	Δ	WT	WT	WT
H3^{K9R}	Δ	K9R	WT	WT
H3.3^{K9R} H3^{HWT}	Δ	WT	K9R	Δ
H3.3^{K9R} H3^{K9R}	Δ	K9R	K9R	Δ

Table 3.1: Genotype description of H3.3 and H3 K9R mutants. Top four rows indicate genotypes used for variant H3.3^{K9R} studies, whereas bottom four rows indicate genotypes used for combined variant and canonical H3.3^{K9R} H3^{K9R} experiments. Wild-type (WT), gene deletion (Δ), no transgenic array (-). See Figure 3.1 for full genotype description.

H3.3B	*H3.3A*	Observed	Expected	p	Fertile
WT	*Het*	535	535.3	n.s.	yes
WT	Δ	268	267.7	n.s.	yes
K9R	*Het*	400	438	**	yes
K9R	Δ	257[a]	219	**	No[b]

[a] The higher than expected number of observed *H3.3B*^{K9R} *H3.3A*^{Null} animals is presumably due to non-specific detrimental effects caused by the presence of a balancer chromosome in siblings with the *H3.3B*^{K9R} mutation and balancer-derived wild-type *H3.3A*.
[b] Both males and females.

Table 3.2: *H3.3*^{K9R} mutants are viable but sterile. Shown are the observed and Mendelian expected number of adults from two separate crosses. Animals heterozygous for the *H3.3A*^{2x1} deletion allele are indicated with (*Het*) while homozygous mutant animals are indicated with (Δ). Animals with wild-type *H3.3B* are indicated with (*WT*). *K9R* signifies the CRISPR-derived mutation of *H3.3B*, which is located on the X chromosome, and the indicated genotype is homozygous or hemizygous with Y chromosome. p value was calculated using Chi-square test (** $p<0.005$).

Figure 3.1: Crossing scheme to generate *H3.3^{K9R}* mutants and *H3.3^{K9R} H3^{K9R}* double mutants. A) Diagram of crosses used to generate H3.3^{K9R} mutants. B) Diagram of crosses used to generate H3.3^{K9R} H3^{K9R} double mutants. Stocks of intermediate genotypes could not be maintained. H3.3B^{K9R} 1 and 2 refer to independent CRISPR-Cas9-mediated substitution events. Df-H3.3A refers to deficiency uncovering H3.3A (Df(2L)BSC110). Boxed insets delineate full genotype for shorthand of all mutant flies used. All genotypes were confirmed through high-throughput sequencing data.

A

$$\frac{H3.3B^{K9R}\,(1)}{H3.3B^{K9R}\,(1)};\frac{CyO,twiGFP}{If}\ \text{x}\ \frac{H3.3B^{K9R}\,(1)}{Y};\frac{H3.3A^{2x1}}{CyO}\quad\Big|\quad\frac{H3.3B^{K9R}\,(2)}{H3.3B^{K9R}\,(2)};\frac{CyO,twiGFP}{If}\ \text{x}\ \frac{+}{Y};\frac{Df(2L)BSC110}{CyO}$$

$$\frac{H3.3B^{K9R}\,(1)}{H3.3B^{K9R}\,(1)};\frac{H3.3A^{2x1}}{CyO,twiGFP}\ \text{x}\ \frac{H3.3B^{K9R}\,(2)}{Y};\frac{Df(2L)BSC110}{CyO,twiGFP}$$

$$\frac{H3.3B^{K9R}\,(1)}{H3.3B^{K9R}\,(2)};\frac{H3.3A^{2x1}}{Df(2L)BSC110}$$

$WT = \frac{+}{+};\frac{+}{+}$

$H3.3B^{K9R} = \frac{H3.3B^{K9R}}{H3.3B^{K9R}};\frac{+}{+}$

$H3.3A^{Null} = \frac{+}{+};\frac{H3.3A^{2x1}}{Df(2L)BSC110}$

$H3.3^{K9R} = \frac{H3.3B^{K9R}}{H3.3B^{K9R}};\frac{H3.3A^{2x1}}{Df(2L)BSC110}$

B

$$\frac{H3.3B^{K9R}\,(2)}{H3.3B^{K9R}\,(2)};\frac{+}{+};\frac{+}{+}\ \text{x}\ \frac{+}{Y};\frac{H3.3A^{2x1},His\Delta,UAS2xYFP}{CyO};\frac{12x\ H3^{HWT}or\ 12x\ H3^{K9R}}{12x\ H3^{HWT}or\ 12x\ H3^{K9R}}$$

$$\frac{H3.3B^{K9R}\,(2)}{Y};\frac{H3.3A^{2x1},His\Delta,UAS2xYFP}{+};\frac{12x\ H3^{HWT}or\ 12x\ H3^{K9R}}{+}\ \text{x}\ \frac{H3.3B^{K9R}\,(1)}{H3.3B^{K9R}\,(1)};\frac{H3.3A^{2x1},His\Delta,twiGal4}{CyO};\frac{+}{+}$$

$$\frac{H3.3B^{K9R}\,(1)}{H3.3B^{K9R}\,(2)};\frac{H3.3A^{2x1},His\Delta,twiGal4}{H3.3A^{2x1},His\Delta,UAS2xYFP};\frac{12x\ H3^{HWT}or\ 12x\ H3^{K9R}}{+}$$

$HWT = \frac{+}{+};\frac{His\Delta,twiGal4}{His\Delta,UAS2xYFP};\frac{12x\ H3^{HWT}}{+}$

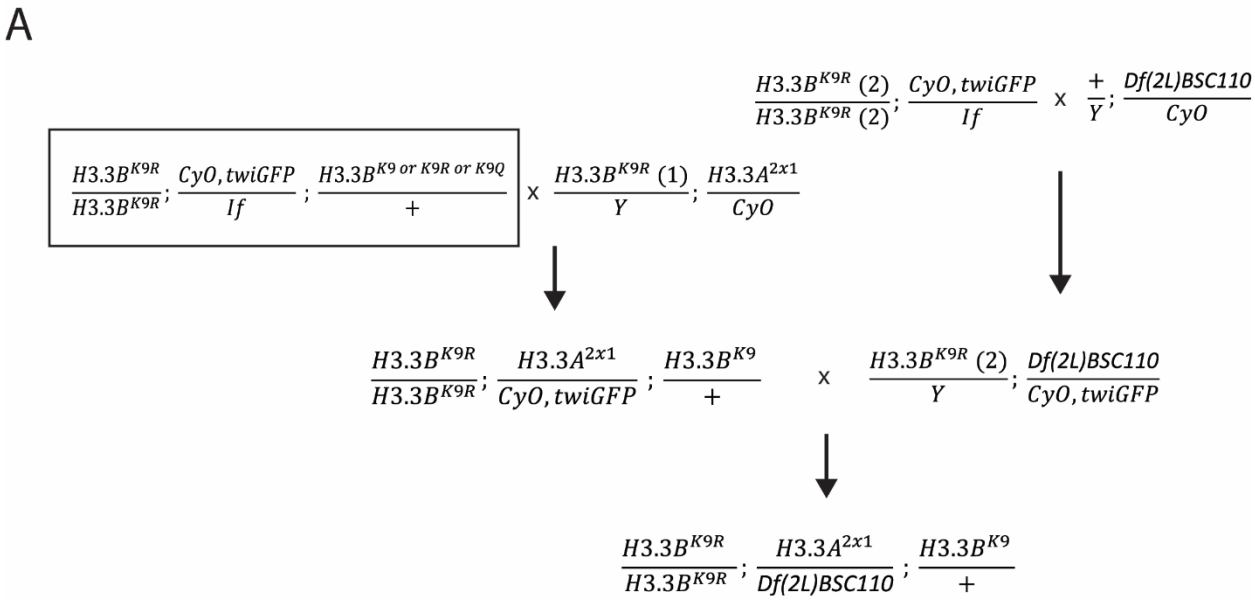
$H3^{K9R} = \frac{+}{+};\frac{His\Delta,twiGal4}{His\Delta,UAS2xYFP};\frac{12x\ H3^{K9R}}{+}$

$H3.3^{K9R}\,H3^{HWT} = \frac{H3.3B^{K9R}}{H3.3B^{K9R}};\frac{H3.3A^{2x1},His\Delta,twiGal4}{H3.3A^{2x1},His\Delta,UAS2xYFP};\frac{12x\ H3^{HWT}}{+}$

$H3.3^{K9R}\,H3^{K9R} = \frac{H3.3B^{K9R}}{H3.3B^{K9R}};\frac{H3.3A^{2x1},His\Delta,twiGal4}{H3.3A^{2x1},His\Delta,UAS2xYFP};\frac{12x\ H3^{K9R}}{+}$

Figure 3.2: Crossing scheme to generate *H3.3^{K9R}* mutants with H3.3B ectopically expressed transgenes. A) Diagram of crosses used to generate H3.3^{K9R} mutants with H3.3B^{K9}, H3.3B^{K9R}, or H3.3B^{K9Q} transgene. Box indicates fly genotype that was sterile in H3.3B^{K9R} and H3.3B^{K9Q} expressing animals.

A


H3.3B^K9R (2) / H3.3B^K9R (2); CyO, twiGFP / If x + / Y; Df(2L)BSC110 / CyO
H3.3B^K9R / H3.3B^K9R; CyO, twiGFP / If; H3.3B^K9 or K9R or K9Q / + x H3.3B^K9R (1) / Y; H3.3A^2x1 / CyO
H3.3B^K9R / H3.3B^K9R; H3.3A^2x1 / CyO, twiGFP; H3.3B^K9 / + x H3.3B^K9R (2) / Y; Df(2L)BSC110 / CyO, twiGFP
H3.3B^K9R / H3.3B^K9R; H3.3A^2x1 / Df(2L)BSC110; H3.3B^K9 / +

H3.3K9 and H3K9 have overlapping functions during development

We previously observed that canonical *H3*^{K9R} mutants could complete development, although 98% of these mutant animals died during larval or pupal stages (Penke et al. 2016). We considered the possibility that *H3*^{K9R} mutant animals progressed to late larval or pupal stages of development because of compensation by H3.3K9. We therefore tested if the *H3.3*^{K9R} genotype would advance the *H3*^{K9R} mutant stage of lethality by observing the development of animals in which the *H3.3*^{K9R} and *H3*^{K9R} mutant genotypes were combined (Table 3.1 and Figure 3.1B). The *H3*^{K9R} genotype was generated using our previously described histone replacement platform (McKay et al. 2015; Penke et al. 2016). Briefly, the endogenous array of ~100x canonical histone gene clusters was deleted and replaced with an ectopically located transgene encoding a BAC-based, 12x tandem array of canonical histone gene clusters in which the *H3* genes contain a K9R mutation (Figure 3.1B). A 12x tandem array of wild-type, canonical histone genes (denoted histone wild type or HWT, Figure 3.1B), which fully rescues deletion of the endogenous histone gene array, was used as a control (McKay et al. 2015). Similar to the *H3.3*^{K9R} mutants, *HWT* animals with the *H3.3*^{K9R} mutant genotype (denoted hereafter as *H3.3*^{K9R} *H3*^{HWT}; see Figure 3.1B) were viable (Table 3.3). However, only 34.6% of *H3.3*^{K9R} *H3*^{HWT} progeny eclosed as adults (Table 3.3) compared to essentially 100% of the *H3.3*^{K9R} genotype that contained the full complement of endogenous, wild-type *H3* genes (Table 3.2). This result suggests that in the presence of fewer total canonical *H3* gene copies, the *H3.3*^{K9R} mutation is more detrimental. Importantly, animals with the *H3.3*^{K9R} *H3*^{K9R} combined mutant genotype containing both the variant and canonical K9R mutation were 100% inviable,

dying with high penetrance at the 1st instar larval stage, much earlier than the majority of *H3^{K9R}* mutants. These results demonstrate that H3.3K9 can partially compensate for the absence of H3K9, indicating that H3.3K9 and H3K9 have some redundant functions.

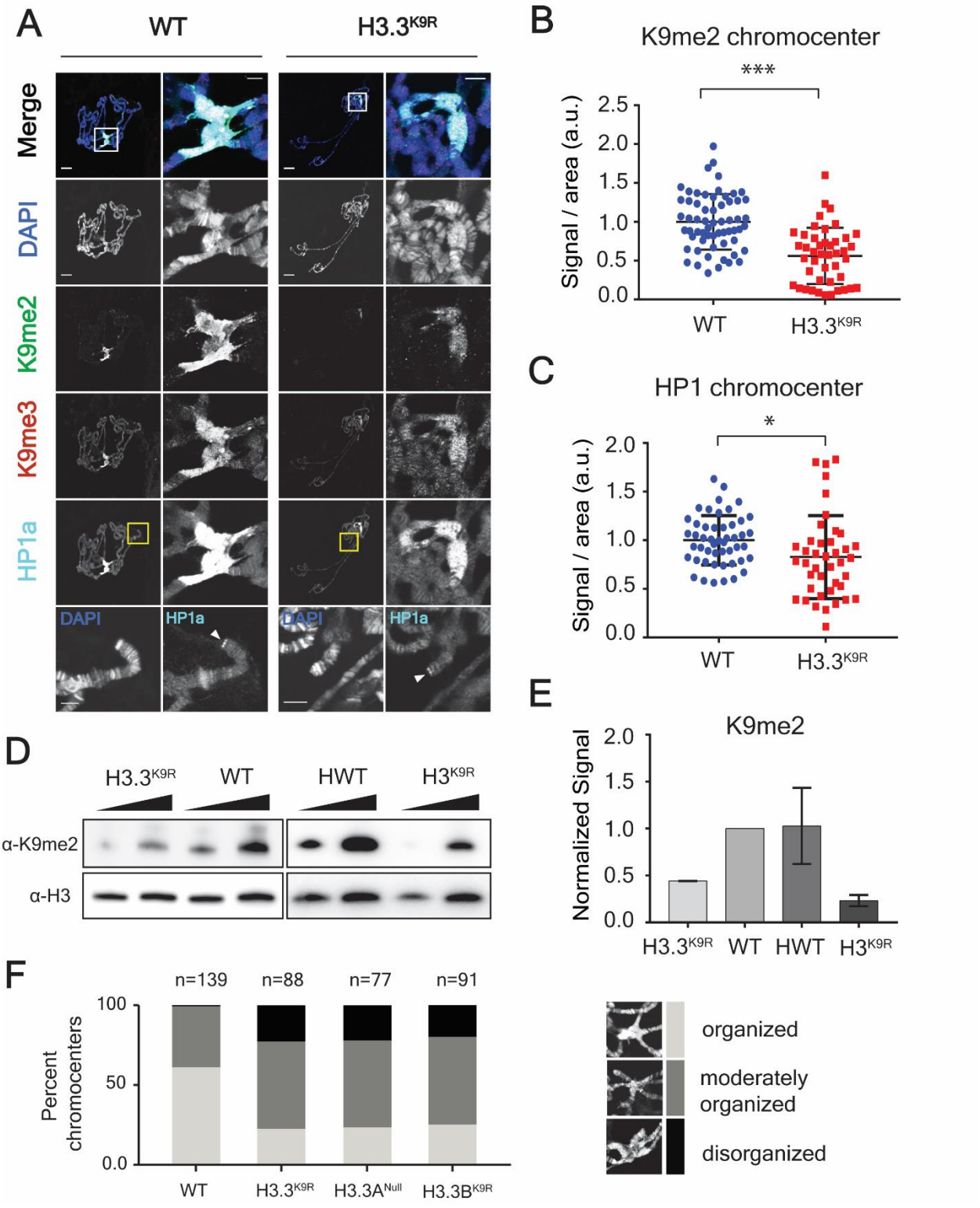
Table 3.3: *H3.3*^{K9R} and *H3*^{K9R} mutations are synthetically lethal. Shown are the percentage (%) of embryos that hatch and the percentage of larvae that pupate and become adults. Note that the pupation and eclosion values for each genotype were obtained from the same brood of animals (hence an identical number of animals analyzed), while the embryo hatching values were obtained from independent experiments. P value for chi-square test were calculated using *HWT* observed (Obs) values as expected values (* $p<0.05$, *** $p<0.0005$).

Genotype	Hatch				Pupate				Eclose			
	Obs	No.	%	p	Obs	No.	%	p	Obs	No.	%	p
HWT	389	450	86.4	-	98	140	70.0	-	88	140	62.9	-
$H3^{K9R}$	370	480	77.1	***	183	285	64.2	*	3	285	1.1	***
$H3.3^{K9R}$ $H3^{HWT}$	350	465	75.3	***	279	462	60.4	***	160	462	34.6	***
$H3.3^{K9R}$ $H3^{K9R}$	214	325	65.8	***	0	130	0.0	***	0	130	0.0	***

H3K9 PTMs are lost in animals lacking H3.3K9 and H3K9

We previously found that K9me2/me3 signal in *H3*^{K9R} mutant animals is substantially reduced but not absent. Thus, a possible reason why *H3.3*^{K9R} *H3*^{K9R} mutants have a more severe developmental defect than *H3*^{K9R} mutants is complete loss of K9me throughout the genome. We therefore assessed K9me2/me3 levels in *H3.3*^{K9R} and *H3.3*^{K9R} *H3*^{K9R} mutants by immunofluorescence. We first assessed K9me2/me3 levels in salivary gland polytene chromosomes of *H3.3*^{K9R} mutants, with the expectation that if H3.3K9 is methylated the signal will be reduced relative to controls. The salivary gland is a highly polyploid tissue (>1000C) and the alignment of chromatids in the polytene chromosomes results in easily visible structures that provide information about levels and genomic locations of histone PTMs using immunofluorescence. *H3.3*^{K9R} mutants had lower levels of both K9me2 and K9me3 compared to wild-type controls at the largely heterochromatic chromocenter, demonstrating that H3.3K9 is normally methylated in the pericentric heterochromatin of otherwise wild-type animals (Figure 3.3A, B). In support of this result, western blot analysis of salivary glands demonstrated that K9me2 levels were decreased in *H3.3*^{K9R} mutants compared to wild-type controls (Figure 3.3D, E).

Figure 3.3: K9me2/me3 and HP1a signal is decreased in H3.3^{K9R} mutants. A) 3rd instar larval salivary gland polytene chromosome spreads from wild-type (left) and H3.3^{K9R} mutants (right) stained with anti-K9me2, anti-K9me3, anti-HP1a, and DAPI to mark DNA. Right panel for each genotype shows enlarged chromocenter indicated by white boxes. Bottom panel shows magnified view of telomere indicated by yellow boxes. Scale bar = 20 microns (whole polytene) 5 microns (chromocenter/telomere). B, C) Immunofluorescent signal of K9me2 (B) or HP1a (C) at chromocenters in wild-type (WT) and H3.3^{K9R} mutants (a.u. = arbitrary units). Values were normalized to area of the chromocenter and set relative to the average WT value from matched slides. Significance was determined using t-test (* $p<0.05$, ** $p<0.005$, *** $p<0.0005$). D) Western blot of K9me2 from salivary glands with H3 used as loading control. E) K9me2 signal was quantified by densitometry and normalized to corresponding H3 loading control band. Normalized values were set relative to WT normalized signal. Error bars represent standard error of the mean from two independent biological replicates (see Materials and Methods). F) Quantification of chromocenter organization from *WT, H3.3^{K9R}, H3.3A^{Null}, H3.3B^{K9R}* mutants.


A
WT
H3.3K9R
Merge
DAPI
K9me2
K9me3
HP1a
DAPI
HP1a
DAPI
HP1a
B
K9me2 chromocenter
***
Signal / area (a.u.)
2.5
2.0
1.5
1.0
0.5
0.0
WT
H3.3K9R
C
HP1 chromocenter
*
Signal / area (a.u.)
2.0
1.5
1.0
0.5
0.0
WT
H3.3K9R
D
H3.3K9R
WT
HWT
H3K9R
α-K9me2
α-H3
E
K9me2
Normalized Signal
2.0
1.5
1.0
0.5
0.0
H3.3K9R
WT
HWT
H3K9R
F
n=139
n=88
n=77
n=91
Percent chromocenters
100
50
0.0
WT
H3.3K9R
H3.3ANull
H3.3BK9R
organized
moderately organized
disorganized

Because *H3.3*^{K9R} mutants exhibited reduced K9me2/me3 signal at the chromocenter, we next used immunofluorescence to examine localization of HP1a, which binds K9me2/me3. In line with reduced K9me2/me3 signal, HP1a signal at the chromocenter of H3.3^{K9R} mutants was reduced compared to wild-type controls (Figure 3.3A, C). Additionally, because HP1a and H3.3 both localize to telomeres (Goldberg et al. 2010; Lewis et al. 2010), we also examined HP1a signal at telomeres in *H3.3*^{K9R} mutants. We found that in the absence of H3.3K9 HP1a is still capable of localizing to telomeres (Figure 3.3A), as it is in *H3*^{K9R} mutants (Penke et al. 2016). These results suggest that the presence of either H3K9 or H3.3K9 at telomeres is sufficient for recruiting HP1a to telomeres.

Because *H3.3*^{K9R} *H3*^{K9R} mutants do not develop to the 3rd instar larval stage, we examined K9me2 levels in 1st instar larval brains. *H3*^{K9R} mutants and *H3.3*^{K9R} *H3*^{HWT} mutants each exhibited reduced K9me2 levels by immunofluorescence compared to *HWT* controls, consistent with the polytene chromosome data (Figure 3.4A). In contrast, the *H3.3*^{K9R} *H3*^{K9R} combined mutants had undetectable levels of K9me2 in the vast majority of cells (Figure 3.4A). These results provide further evidence that H3.3K9 is methylated and that the total amount of K9me is derived from both H3.3 and H3. Interestingly, a small number of cells in the *H3.3*^{K9R} *H3*^{K9R} 1st instar mutant brains retained low levels of K9me2 signal at the chromocenter (arrowheads, Figure 3.4). Cells with residual K9me2 express ELAV, a pan-neuronal marker, and lack expression of Deadpan and Prospero, markers of proliferating neuroblasts and ganglion mother cells, respectively (circles, Figure 3.5). These data indicate that cells with K9me2 positive chromocenters in *H3.3*^{K9R} *H3*^{K9R} mutant 1st instar larval brains are differentiated

neurons. We suspect that the K9me2 signal in these cells reflects maternally provided wild-type H3 protein remaining in the genomes of quiescent neurons that differentiated prior to having their maternal H3 fully replaced by zygotically expressed H3K9R mutant histones. A corollary to this conclusion is that the proliferating neuroblasts and their GMC daughters likely have progressed through a sufficient number of S phases such that replacement of maternal H3 with zygotic H3K9R eliminates detectable K9me2 signal.

Figure 3.4: K9me2/me3 signal is diminished in K9R mutants. A) 1st instar larval brains stained with anti-K9me2 and DAPI to mark DNA from *HWT*, *H3*^{K9R}, *H3.3*^{K9R}, and *H3.3*^{K9R} *H3*^{K9R} animals. Left panel shows max projection of 2 micrometer confocal sections through the entire brain. Right panel shows a magnified, single confocal section from the area indicated by the white boxes. Arrowheads indicate cells with residual K9me2 signal in *H3.3*^{K9R} *H3*^{K9R} animals. Scale bar = 50 microns (whole brain) 10 microns (enlarged image).

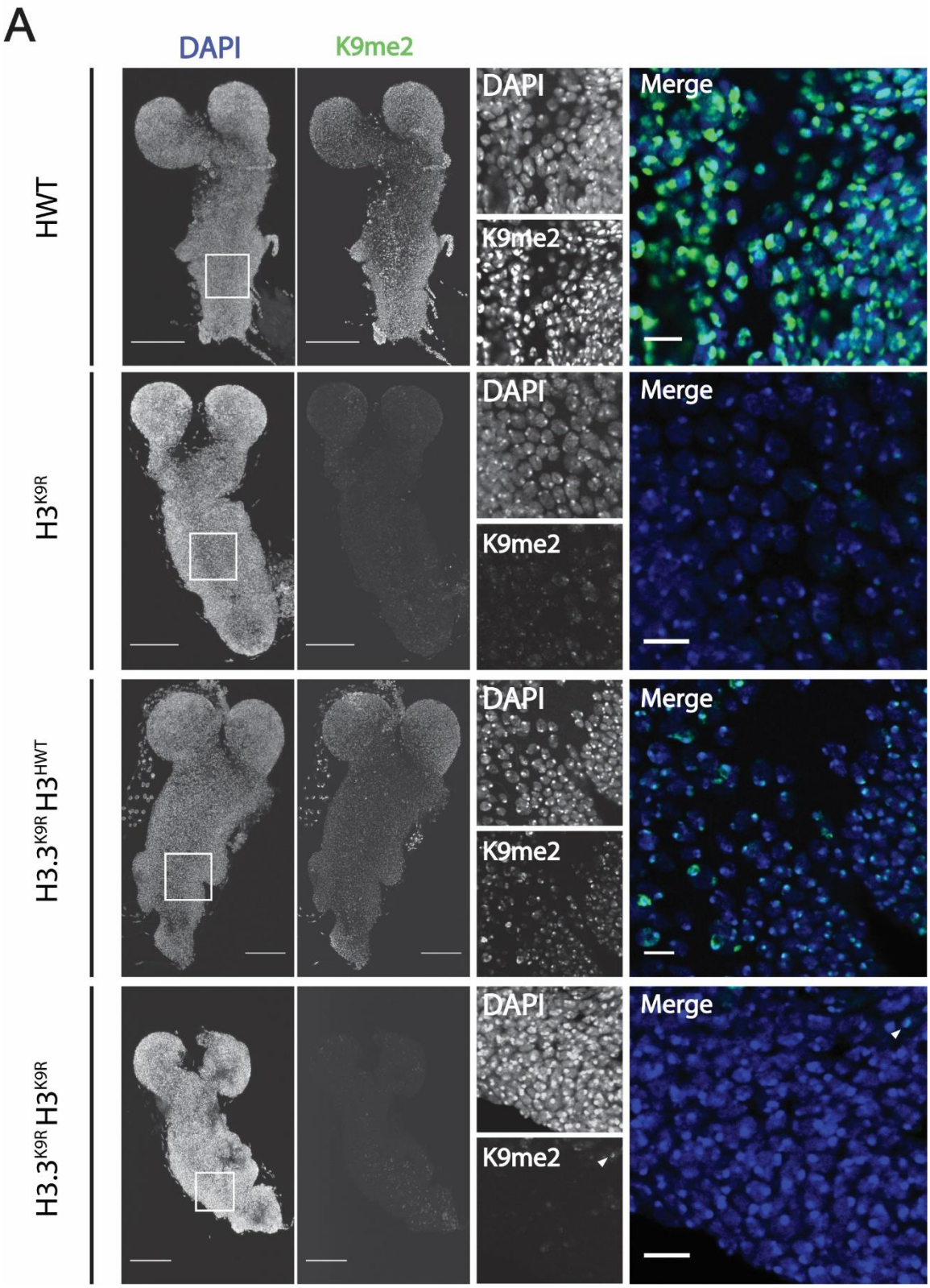
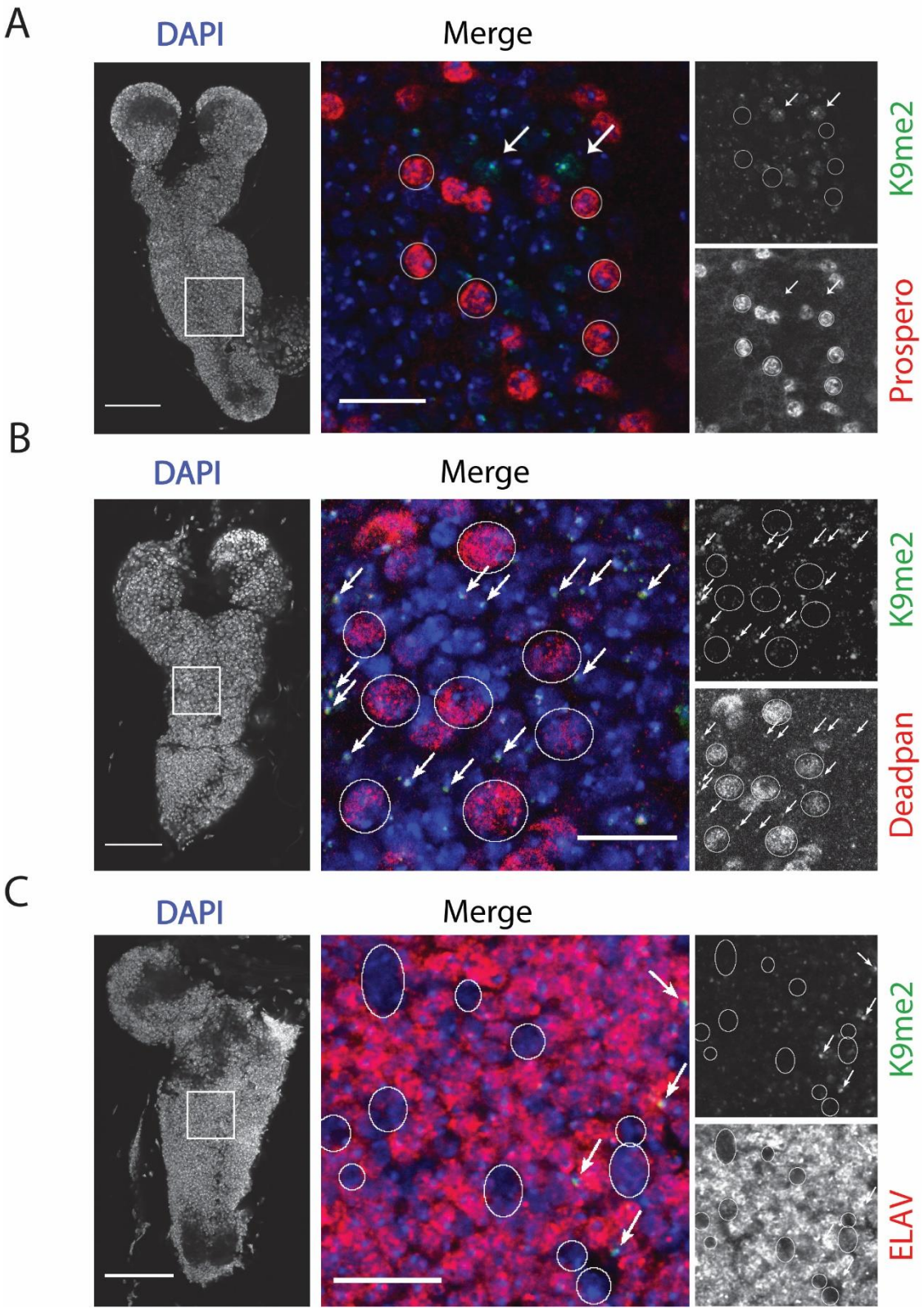
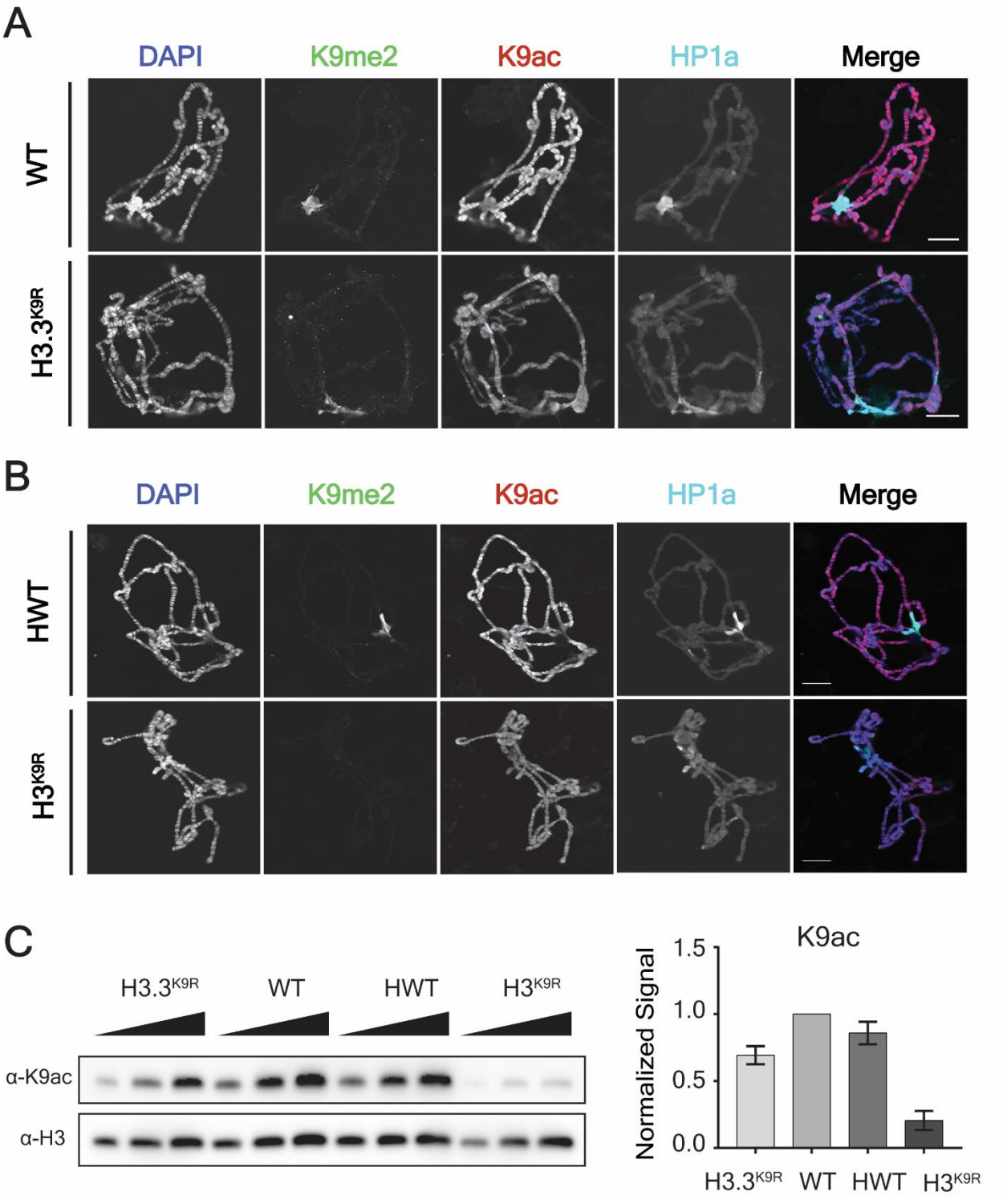

A
DAPI
K9me2
HWT
H3K9R
H3.3K9R H3HWT
H3.3K9R H3K9R
DAPI
Merge
K9me2

Figure 3.5: Residual H3K9me2 signal is found in differentiated neurons in *H3.3^{K9R} H3^{K9R}* mutants. H3^{K9R} 1st instar brains were stained with DAPI in blue, anti-K9me2 in green, and either anti-Prospero (A), anti-Deadpan (B), or anti-ELAV (C) in red. Prospero marks ganglion mother cells, deadpan marks neuroblasts, and ELAV marks differentiated neurons. Arrows indicate K9me2 positive cells in which signal overlaps the chromocenter. Circles in A and B indicate prospero or deadpan positive cells respectively whereas circles in C indicate ELAV negative cells. K9me2 positive cells contain ELAV but neither prospero or deadpan. Single slice images of the whole brain are shown in the left panel (scale bar = 50 microns) and magnified views of individuals cells are shown in the right panels (scale bar = 10 microns).


A
DAPI
Merge
K9me2
Prospero
B
DAPI
Merge
K9me2
Deadpan
C
DAPI
Merge
K9me2
ELAV

We also found that levels of H3K9 acetylation were reduced in both the *H3.3*^{K9R} mutant and the *H3*^{K9R} mutant relative to controls, as determined both by immunofluorescence of salivary gland polytene chromosomes (Figure 3.6A, B) and by western blots of salivary gland extracts (Figure 3.6C). Because a substantial amount of K9ac is placed on H3.3, we considered the possibility that lack of K9ac was responsible for the fertility defects of *H3.3*^{K9R} mutants and the early lethality of *H3.3*^{K9R} *H3*^{K9R} mutants. To address this question, we integrated either an H3.3B^{K9}, an H3.3B^{K9R}, or an H3.3B^{K9Q} transgene into the same genomic position in order to determine if a K9Q acetyl mimic could restore function to *H3.3*^{K9R} mutants. Animals with only an *H3.3B*^{K9R} mutation at the endogenous locus (i.e., containing a wild-type H3.3A gene), and carrying either an H3.3B^{K9R} or H3.3B^{K9Q} transgene were sterile, precluding us from constructing the genotype to test if these transgenes could rescue the sterility of *H3.3*^{K9R} mutant adults (Figure 3.2). This result suggests that both the H3.3B^{K9R} and H3.3B^{K9Q} transgenes acted dominantly to compromise fertility. Furthermore, these data imply that H3.3B^{K9R} and H3.3B^{K9Q} histones are incorporated into chromatin.

Figure 3.6: K9ac signal is decreased in H3.3^{K9R} mutants. A,B) Polytene chromosome spreads from wild-type and *H3.3^{K9R}* mutants (A) or *HWT* and *H3^{K9R}* mutants (B) stained with anti-K9me2, anti-K9ac, anti-HP1a, and DAPI to mark DNA. Scale bar = 20 microns. C) Western blot of K9ac from salivary glands with H3 used as loading control. K9ac signal was quantified by densitometry and normalized to corresponding H3 loading control band. Normalized values were set relative to WT normalized signal. Error bars represent standard error of the mean from two independent biological replicates (see Materials and Methods).


A
DAPI
K9me2
K9ac
HP1a
Merge
WT
H3.3K9R
B
DAPI
K9me2
K9ac
HP1a
Merge
HWT
H3K9R
C
H3.3K9R
WT
HWT
H3K9R
α-K9ac
α-H3
K9ac
Normalized Signal
1.5
1.0
0.5
0.0
H3.3K9R
WT
HWT
H3K9R

H3.3K9 regulates chromatin organization at the chromocenter, telomeres, and transposons

We next asked if the reduction of K9me2/me3 in *H3.3*^{K9R} mutants affected chromatin organization by cytological examination of salivary gland polytene chromosomes. We examined the structure of the chromocenter in polytene chromosome spreads by binning chromocenters into three categories: “organized”, “moderately organized”, and “disorganized” (Figure 3.3F). We categorized chromocenters from four genotypes: wild-type (WT; i.e., with the endogenous canonical histone genes), an *H3.3A* null mutant (*H3.3A*^{Null}), an *H3.3B* K9R substitution mutant (*H3.3B*^{K9R}), and the *H3.3B*^{K9R} / *H3.3A*^{Null} double mutant in which all H3.3 contains the K9R substitution (*H3.3*^{K9R}) (Table 3.1 and Figure 3.1A). Whereas the majority of wild-type chromocenters were organized (60% organized vs 40% moderately organized), both the *H3.3B*^{K9R} and the *H3.3A*^{Null} single mutants had increased percentages of moderately organized and disorganized chromocenters (Figure 3.3F). For example, ~22% of chromocenters in the various *H3.3* mutants were disorganized compared to less than 1% of wild-type chromocenters. These results indicate that H3.3 contributes to chromocenter structure. Interestingly, the *H3.3*^{K9R} double mutant had the same proportion of moderately organized and disorganized chromocenters as either single mutant. This result suggests that either reducing *H3.3* gene dose (i.e., the *H3.3A*^{Null} allele) or expressing K9R mutant H3.3 histones (i.e., the *H3.3B*^{K9R} mutation), can prevent normal H3.3 function at pericentric heterochromatin.

Given the disrupted chromocenter structure in *H3.3*^{K9R} mutants, we next examined chromatin structure genome wide by performing Formaldehyde Assisted Isolation of Regulatory Elements followed by whole genome sequencing (FAIRE-seq).

FAIRE-seq provides a measure of local nucleosome occupancy across the genome, revealing regions of “open” chromatin that are relatively depleted of nucleosomes (Simon et al. 2013). We previously found using this technique that regions of heterochromatin enriched in K9me, particularly pericentromeric heterochromatin, were more open in *H3*^{K9R} mutants relative to *HWT* controls (Penke et al. 2016). To determine if *H3.3*^{K9R} mutants had a similar phenotype we performed FAIRE-seq in triplicate on imaginal wing discs from wandering 3rd instar larvae in *WT*, *H3.3A*^{Null}, *H3.3B*^{K9R}, and *H3.3*^{K9R} genotypes. Sequencing reads were aligned to the genome and peaks were called on each of the three replicates and combined into a merged peak set. Called peaks were consistent across replicates and read coverage across peaks was highly correlated ($R \geq 0.96$) (Figure 3.7A, B). Additionally, wild-type FAIRE data was consistent with previously generated data from wing discs (McKay and Lieb 2013) (Figure 3.7D). *H3.3A*^{Null}*, H3.3B*^{K9R}*, and H3.3*^{K9R} mutants each had a similar percentage of peaks with significantly altered FAIRE signal when compared to wild-type: 8.8%, 6.5%, and 7.9% respectively (Figure 3.8A-C). Moreover, significantly changed peaks across the three mutants exhibited a high degree of overlap. Of the 2,660 significantly changed peaks across all mutants, 21% were shared among all three and 52% by at least two mutants (Figure 3.9A). FAIRE signal at significantly changed peaks also displayed similar fold changes in mutants compared to wild-type and were not exacerbated in the double mutant compared to either single mutant (Figure 3.9B). These data suggest H3.3A and H3.3BK9 both function to regulate chromatin architecture.

Figure 3.7: Imaginal wing disc FAIRE signal is highly consistent across replicates and with previously generated FAIRE data. A) Scatterplots comparing normalized FAIRE signal at merged set of FAIRE peaks for all replicates of a particular genotype. Signal expressed as $\log_2$ reads per million (RPM). R value indicates Pearson Correlation. B) Venn Diagram showing overlap of MACS2 called peaks demonstrates peaks were highly similar across replicates. C) Average FAIRE signal at 5bp bins surrounding transcription start sites (TSS). Signal expressed as average reads per million (RPM). D) Correlation analysis of FAIRE signal at merged set of FAIRE peaks from average WT and previously published FAIRE data in wing discs (McKay & Lieb 2013).

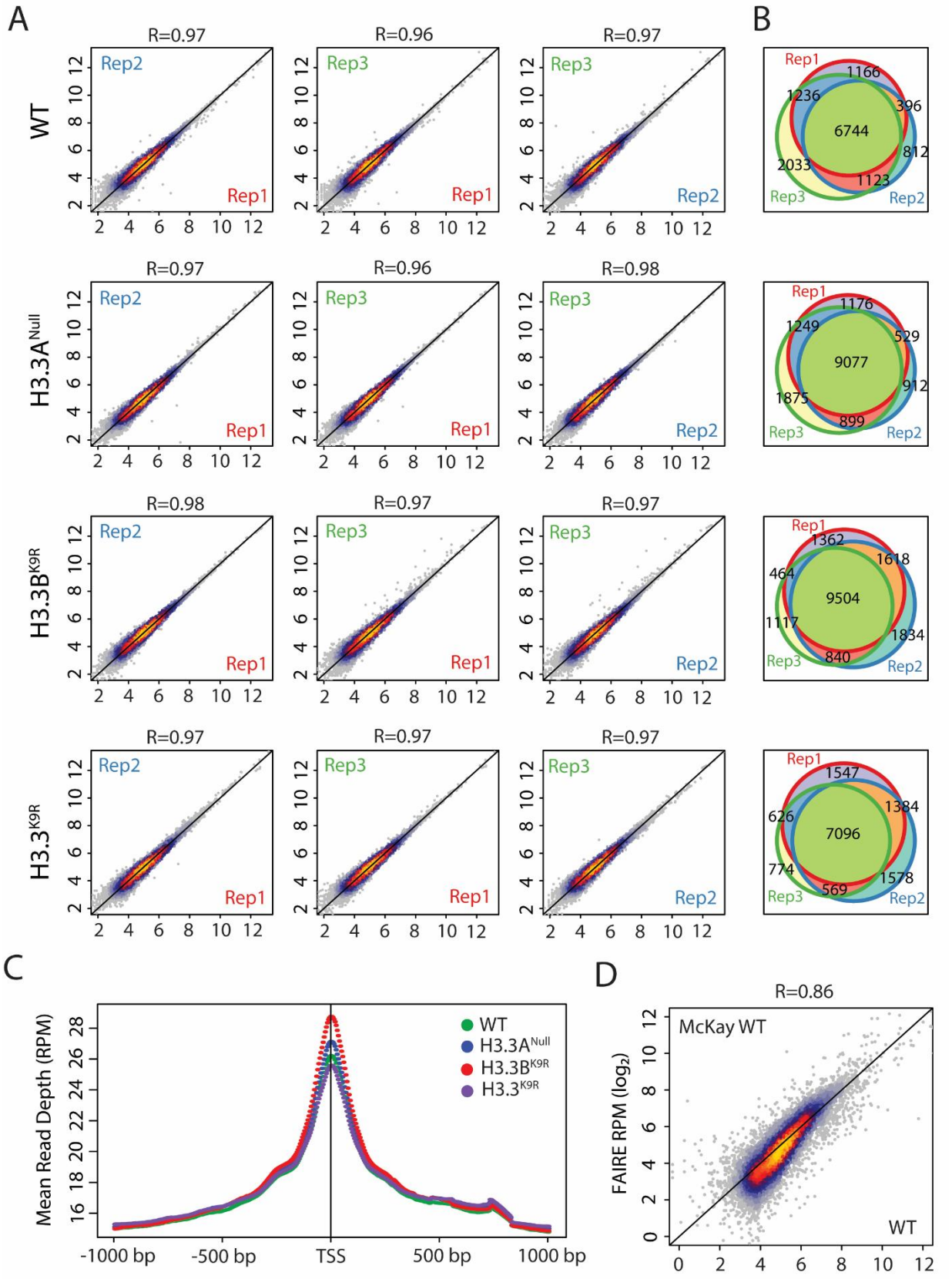
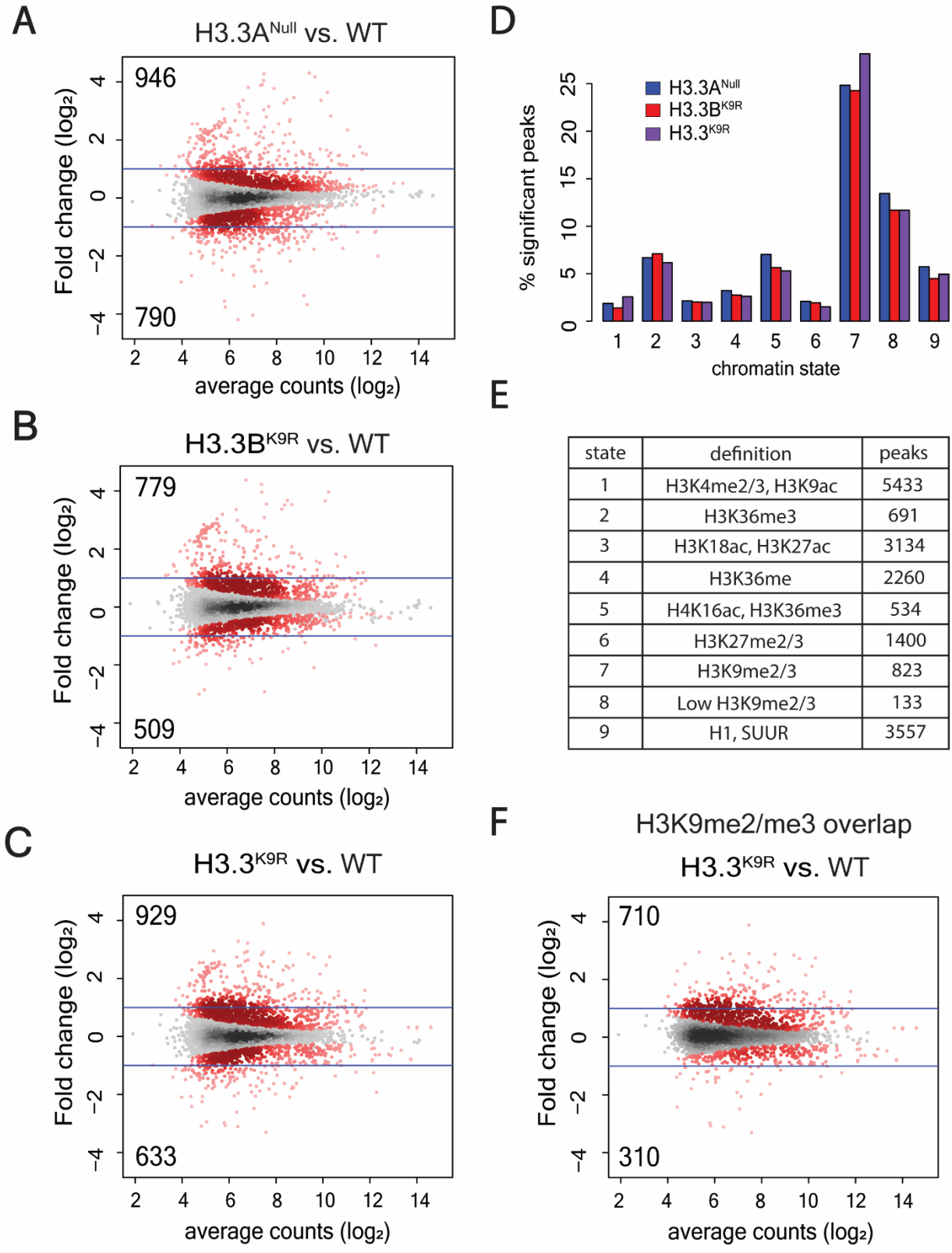
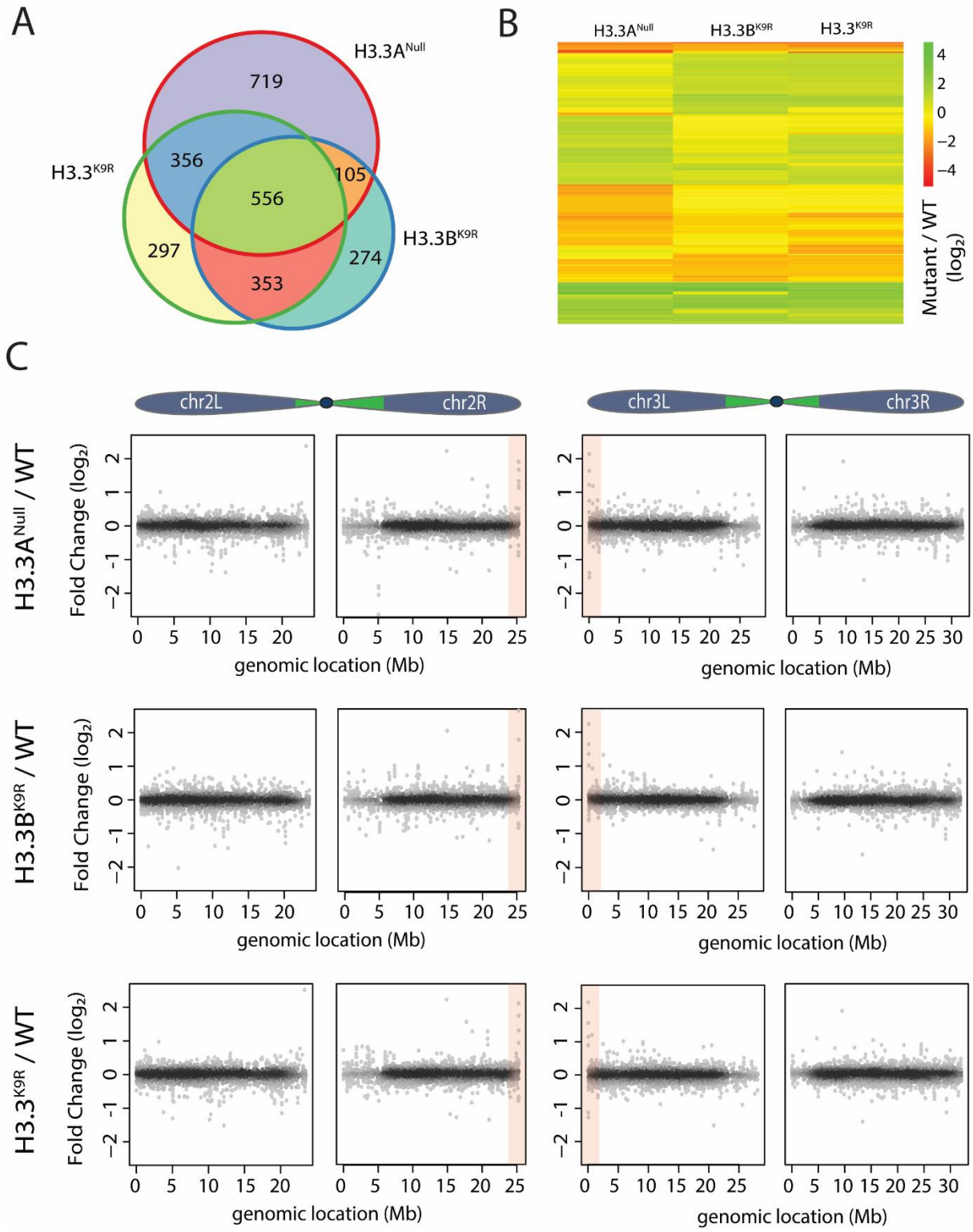

A
B
C
D
WT
H3.3A^Null
H3.3B^K9R
H3.3^K9R
R=0.97
R=0.96
R=0.98
R=0.97
R=0.97
R=0.97
R=0.98
R=0.97
R=0.97
R=0.97
R=0.97
R=0.97
Rep1
Rep2
Rep3
1166
1236
396
6744
812
2033
1123
1176
1249
529
9077
912
1875
899
1362
1618
464
9504
1117
1834
840
1547
1384
626
7096
774
1578
569
Mean Read Depth (RPM)
-1000 bp
-500 bp
TSS
500 bp
1000 bp
WT
H3.3A^Null
H3.3B^K9R
H3.3^K9R
R=0.86
McKay WT
FAIRE RPM (log2)
WT

Figure 3.8: H3.3K9 regulates chromatin architecture in regions of K9me. A-C) Mutant: WT ratio of *H3.3A*^{null} (A), *H3.3B*^{K9R} (B), or *H3.3*^{K9R} (C) FAIRE signal from 3rd instar imaginal wing discs at 19,738 FAIRE peaks called by MACS2. Red dots indicate significantly different peaks ($p<0.05$), and insets indicate the number of significantly increased (top) or decreased (bottom) peaks. Average counts signify average normalized reads that overlap a peak in mutant and WT samples. D) Percentage of peaks in a particular chromatin state that have significantly different FAIRE signal in mutants versus WT (top). Bottom panel shows a summary of histone modifications or proteins that define a chromatin state and the number of FAIRE peaks assigned to a given chromatin state. E) Boxplot of FAIRE enrichment over input at 126 transposon families (* indicates $p < 0.05$ and *** indicates $p <0.0005$). F) Plot from C showing only those peaks that overlap an K9me2 or K9me3 peak from modENCODE ChIP-seq data.



state	definition	peaks
1	H3K4me2/3, H3K9ac	5433
2	H3K36me3	691
3	H3K18ac, H3K27ac	3134
4	H3K36me	2260
5	H4K16ac, H3K36me3	534
6	H3K27me2/3	1400
7	H3K9me2/3	823
8	Low H3K9me2/3	133
9	H1, SUUR	3557

Figure 3.9: Regions of significantly changed FAIRE signal are similar across *H3.3* mutants. A) Venn diagram showing overlap of FAIRE peaks with significantly changed FAIRE signal between *H3.3* mutant and *WT* samples. Significance cutoff set at adjusted p value of 0.05 as determined by DESeq2 package (Love et al. 2014a). B) Heatmap showing *H3.3* mutant over *WT* fold change (log_2) of FAIRE signal at all significantly different FAIRE peaks shown in A. Fold changes are not exacerbated in $H3.3^{K9R}$ double mutants compared to either $H3.3B^{K9R}$ or *H3.3A* single mutant alone. C) Ratio of H3.3 mutant over *WT* FAIRE signal plotted versus genome coordinate of FAIRE peaks on chromosome 2 and 3. Light red boxes highlight telomeric areas with changes in FAIRE signal. Blue areas of the chromosome diagram indicate largely euchromatic regions, and green areas of the chromosome diagram represent approximate locations of pericentromeric heterochromatin (Hoskins et al. 2015; Riddle et al. 2011). Mb = megabase.

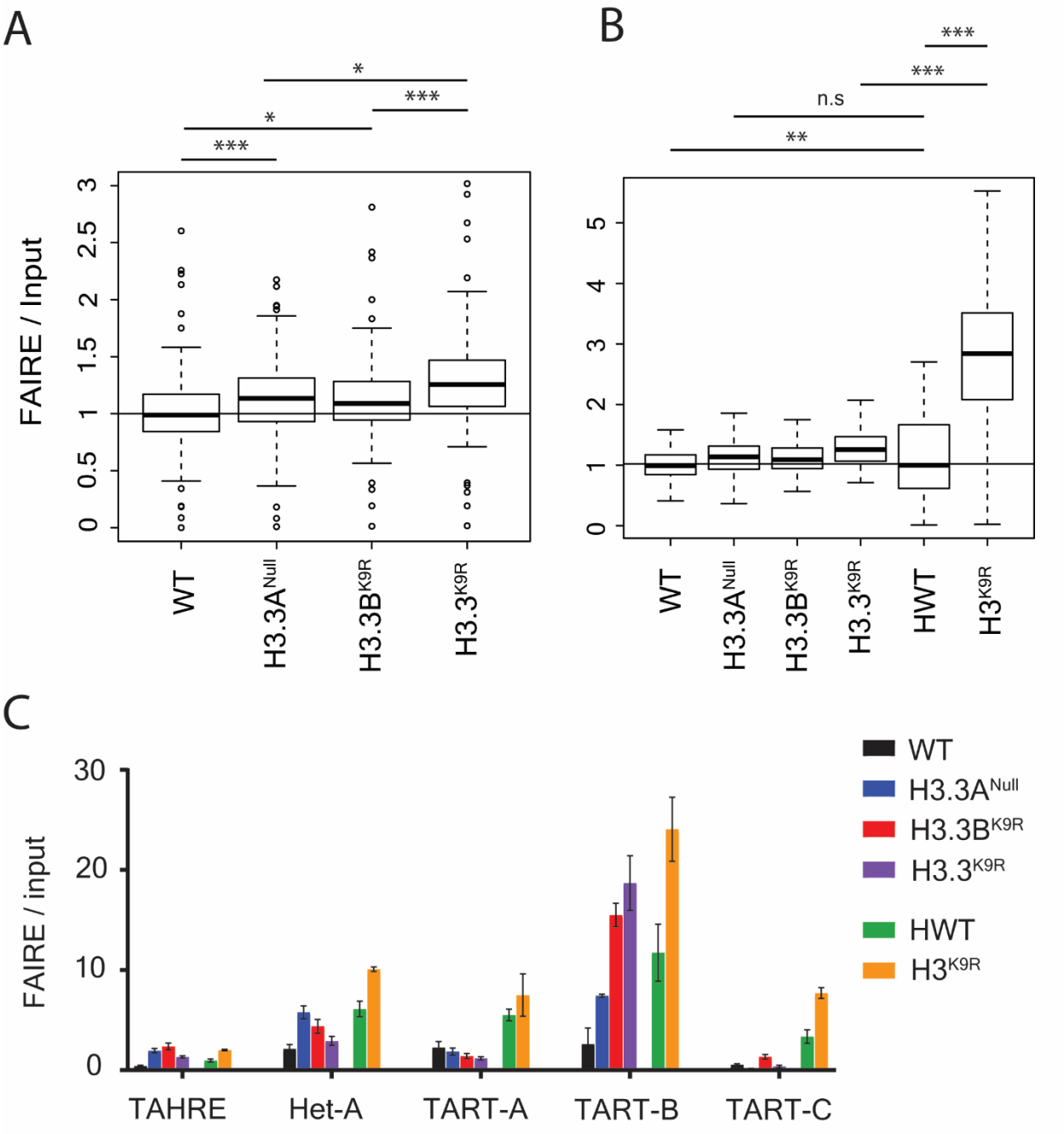

A
H3.3A^Null
H3.3^K9R
H3.3B^K9R
719
356
105
556
297
274
353
B
H3.3A^Null
H3.3B^K9R
H3.3^K9R
4
2
0
−2
−4
Mutant / WT (log2)
C
chr2L
chr2R
chr3L
chr3R
H3.3A^Null / WT
H3.3B^K9R / WT
H3.3^K9R / WT
Fold Change (log2)
genomic location (Mb)

We next asked if the changes in FAIRE signal we observed in *H3.3* mutants were characterized by a particular chromatin signature. We assigned each called FAIRE peak to one of nine different chromatin states characterized by different combinations of histone PTMs as defined by Kharchenko et al. (2010). We then calculated the percentage of FAIRE peaks that changed between an *H3.3* mutant and wild-type within each chromatin state. Regions of K9me2/me3 showed the highest percentage of changes in FAIRE signal in $H3.3A^{Null}$, $H3.3B^{K9R}$, and the $H3.3^{K9R}$ mutant compared to wild-type, supporting the idea that H3.3K9 is methylated and plays a necessary role in regulating chromatin architecture (Figure 3.8D). Changes in FAIRE signal were also more likely to occur in regions of H3K36me3, a mark that is enriched along gene bodies that are themselves enriched for H3.3 (Bannister et al. 2005; Szenker et al. 2011). Finally, we used modENCODE K9me2 and K9me3 ChIP-seq data to complement the chromatin state analysis. Of the FAIRE peaks significantly increased or decreased in $H3.3^{K9R}$ mutants compared to wild-type, 76.4% and 49.0% respectively overlapped a K9me2 or K9me3 peak (Figure 3.8F). These results demonstrate that altered FAIRE signal in $H3.3^{K9R}$ mutants occurred in regions normally occupied by K9me.

We also observed increased FAIRE signal at telomeres in all three *H3.3* mutant genotypes, particularly on chromosomes 2R and 3L (Figure 3.9C), suggesting that H3.3 regulates telomeric chromatin architecture. In *Drosophila,* telomeres are composed of retrotransposons enriched in K9me2/me3 (Levis et al. 1993; Cenci et al. 2005). H3.3 plays a similar role in the mouse, in which H3.3 null mutant embryonic stem cells exhibit an increase in transcripts from transposons (Elsässer et al. 2015) and telomeres (Udugama et al. 2015). Additionally, we previously observed transposon activation and

mobilization in canonical *H3*^{K9R} mutants (Penke et al. 2016). For these reasons, we examined FAIRE signal at transposons in our *H3.3* mutants using the piPipes pipeline, which avoids ambiguity in aligning reads to repetitive transposons by mapping to transposon families (Han et al. 2015). Both *H3.3A*^{Null} and *H3.3B*^{K9R} mutants resulted in significantly increased FAIRE signal at transposons, and *H3.3*^{K9R} mutants had on average even higher increased FAIRE signal at transposons (Figure 3.10A, B). Moreover, FAIRE signal at some telomeric transposons, particularly TART-B, was increased in *H3.3* mutants (Figure 3.10C). However, the extent of increase in *H3.3*^{K9R} mutants was not as severe as previously observed for *H3*^{K9R} mutants (Penke et al. 2016) (Figure 3.10B). These results support a role for H3.3K9 in chromatin-mediated transposon repression, though to a lesser extent than H3K9.

Figure 3.10: Imaginal wing disc FAIRE signal of *H3.3* mutants is increased at telomeres and transposons. A) Boxplot of average FAIRE enrichment determined by piPipes pipeline across 126 transposon families (Han et al. 2015). Genomic DNA from *Drosophila* embryos used as input control. B) Boxplots in A shown alongside FAIRE enrichment for *HWT* and $H3^{K9R}$ mutants from a separate experiment (Penke et al. 2016). C) FAIRE enrichment of *H3.3* and $H3^{K9R}$ mutants at telomeric transposons. Error bars indicate standard deviation from three replicates for each genotype. Statistical significance determined by paired t-test ($p<0.05$ *, $p<0.005$ **, $p<0.0005$ ***, n.s. = not significant).


A
FAIRE / Input
0
0.5
1
1.5
2
2.5
3
*
***
*
***
WT
H3.3A^Null
H3.3B^K9R
H3.3^K9R
B
n.s
**
***
***
0
1
2
3
4
5
WT
H3.3A^Null
H3.3B^K9R
H3.3^K9R
HWT
H3^K9R
C
FAIRE / input
0
10
20
30
TAHRE
Het-A
TART-A
TART-B
TART-C
WT
H3.3A^Null
H3.3B^K9R
H3.3^K9R
HWT
H3^K9R

H3.3K9 and H3K9 functions overlap in regions of K9ac and partially in regions of K9me

To investigate the cause of lethality when both variant and canonical H3 histones contain the K9R mutation, we performed RNA-seq of 1st instar larvae from four genotypes: *HWT*, *H3^{K9R}*, *H3.3^{K9R} H3^{HWT}*, and *H3.3^{K9R} H3^{K9R}* (Table 3.1). Larvae of the correct genotype were identified by GFP fluorescence (see Materials and Methods). RNA sequencing reads were aligned to the genome using Tophat, transcript assembly was performed by Cufflinks, and DESeq2 was used for statistical analysis (Trapnell et al. 2014; Love et al. 2014a). Genotypes were verified by examination of reads mapping to variant and canonical histones, and correlation analysis demonstrated transcript abundance across all assembled transcripts was highly similar among replicates, and was also similar to previously generated data from wild-type 1st instar larvae (Figure 3.11) (Graveley et al. 2011). Additionally, histone expression was similar across all genotypes, suggesting that variation in histone levels do not underlie observed phenotypes (Figure 3.12A). In line with our previous analysis of *H3^{K9R}* RNA-seq data from imaginal wing discs (Penke et al. 2016), the majority of significantly changed transcripts in *H3^{K9R}* 1st instar samples was increased compared to *HWT* (247 increased vs 41 decreased), supporting a role for H3K9me in gene silencing (Figure 3.13A). *H3.3^{K9R} H3^{HWT}* samples had a similar number of significantly changed transcripts, and again most transcripts showed increased signal compared to *HWT* (203 vs 126), though fold changes were smaller than *H3^{K9R}* mutants (Figure 3.13B). By contrast, the *H3.3^{K9R} H3^{K9R}* combined mutant genotype caused a much more pronounced effect on gene expression compared to either the *H3.3^{K9R} H3^{HWT}* or the *H3^{K9R}* mutant genotypes (Figure 3.13C); 869 transcripts exhibited increased RNA signal and 1036 transcripts were decreased compared to *HWT* samples. The number of decreased transcripts in

H3.3^{K9R} *H3*^{K9R} animals compared to *HWT* was therefore about ten-fold higher than either the variant or canonical K9R mutant alone. Thus, similar to our viability analysis (Table 3.3), these RNA-seq results demonstrated that variant and canonical versions of H3K9 compensate for each other in the regulation of gene expression.

Figure 3.11: 1st instar larvae RNA signal of *H3.3^{K9R}* and *H3^{K9R}* mutants is highly consistent across replicates and with previously generated RNA data. A) Scatterplot of normalized RNA signal at transcripts assembled by cufflinks for all replicates of a particular genotype. RNA signal shown in $\log_2$ transformed transcripts per million (TPM). R indicates Pearson correlation. B) Scatterplot of normalized RNA signal comparing *HWT* and modENCODE RNA-seq data for 1st instar larvae (Graveley et al. 2011).

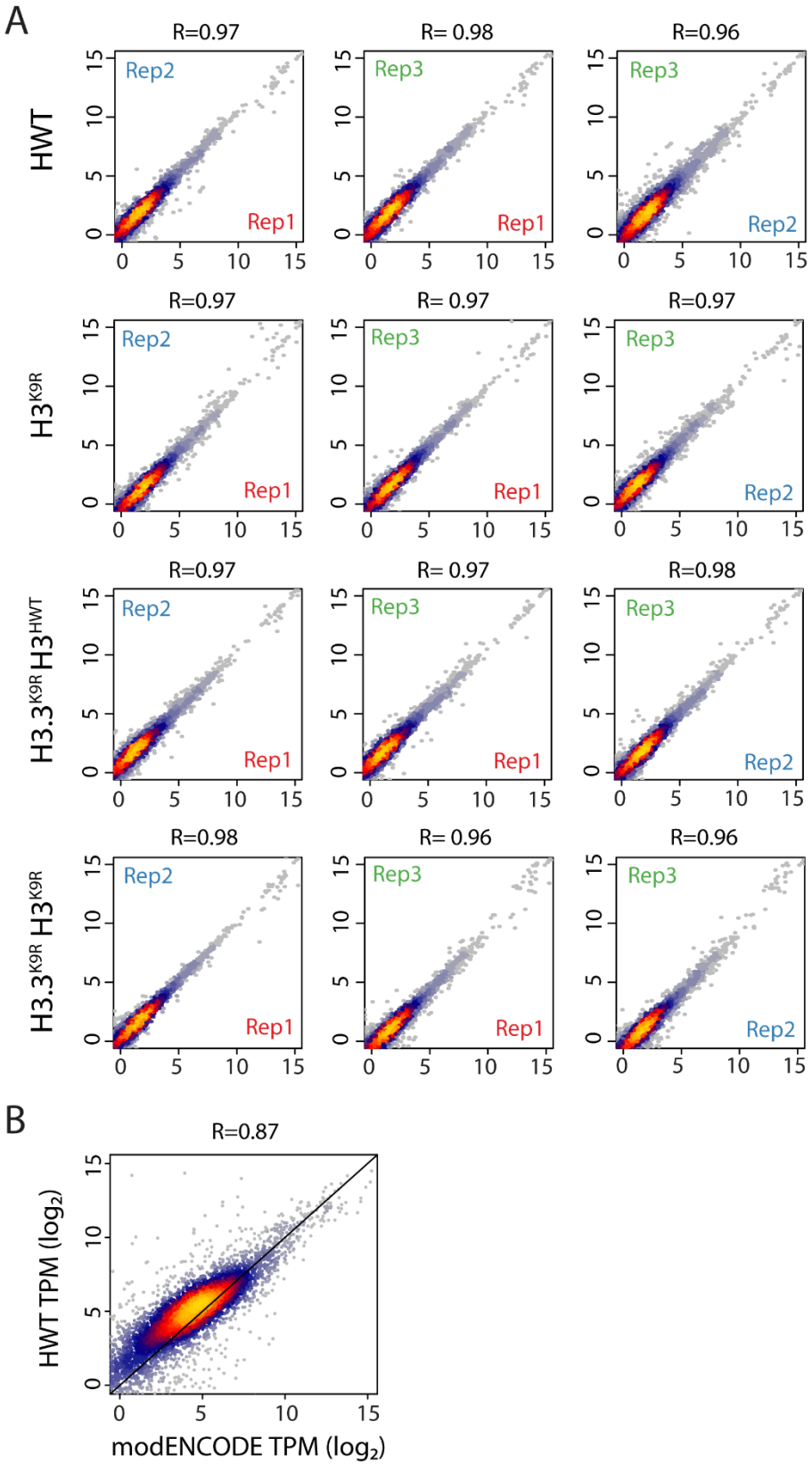

A
HWT
H3K9R
H3.3K9R H3HWT
H3.3K9R H3K9R
R=0.97
R= 0.98
R=0.96
R=0.97
R= 0.97
R=0.97
R=0.97
R= 0.97
R=0.98
R=0.98
R= 0.96
R=0.96
Rep1
Rep2
Rep3
0
5
10
15
B
R=0.87
HWT TPM (log2)
modENCODE TPM (log2)

Figure 3.12: RNA signal across transposon families is increased in *H3.3^{K9R}* and *H3^{K9R}* mutants. A) Histone steady-state RNA levels are similar between variant and canonical H3K9R mutants. Reads were mapped using Bowtie2 to a custom index file containing one copy of the histone repeat (one copy of each replication-dependent histone) and H3.3B. Shown are the RPKM values in the coding region of each individual histone or across the entire histone repeat. Reads were normalized to the number of reads that uniquely map to the entire genome. Error bars represent standard deviation of three independent replicates for each genotype. B-D) Normalized counts of RNA signal at transposon families determined by piPipes pipeline (Han et al. 2015). B) Boxplot showing average fold change of RNA signal at 126 transposon families in *H3.3^{K9R}* and *H3^{K9R}* mutants compared to *HWT*. Paired t-tests were used to analyze statistical differences across genotypes (* $p<0.05$, ** $p<0.005$, *** $p<0.0005$). C) Heatmap of FAIRE signal showing K9R mutant over *HWT* fold change at transposon families. D) RNA normalized counts at telomeric transposons in K9R mutants. Error bars indicate standard deviation derived from three replicates for each genotype.

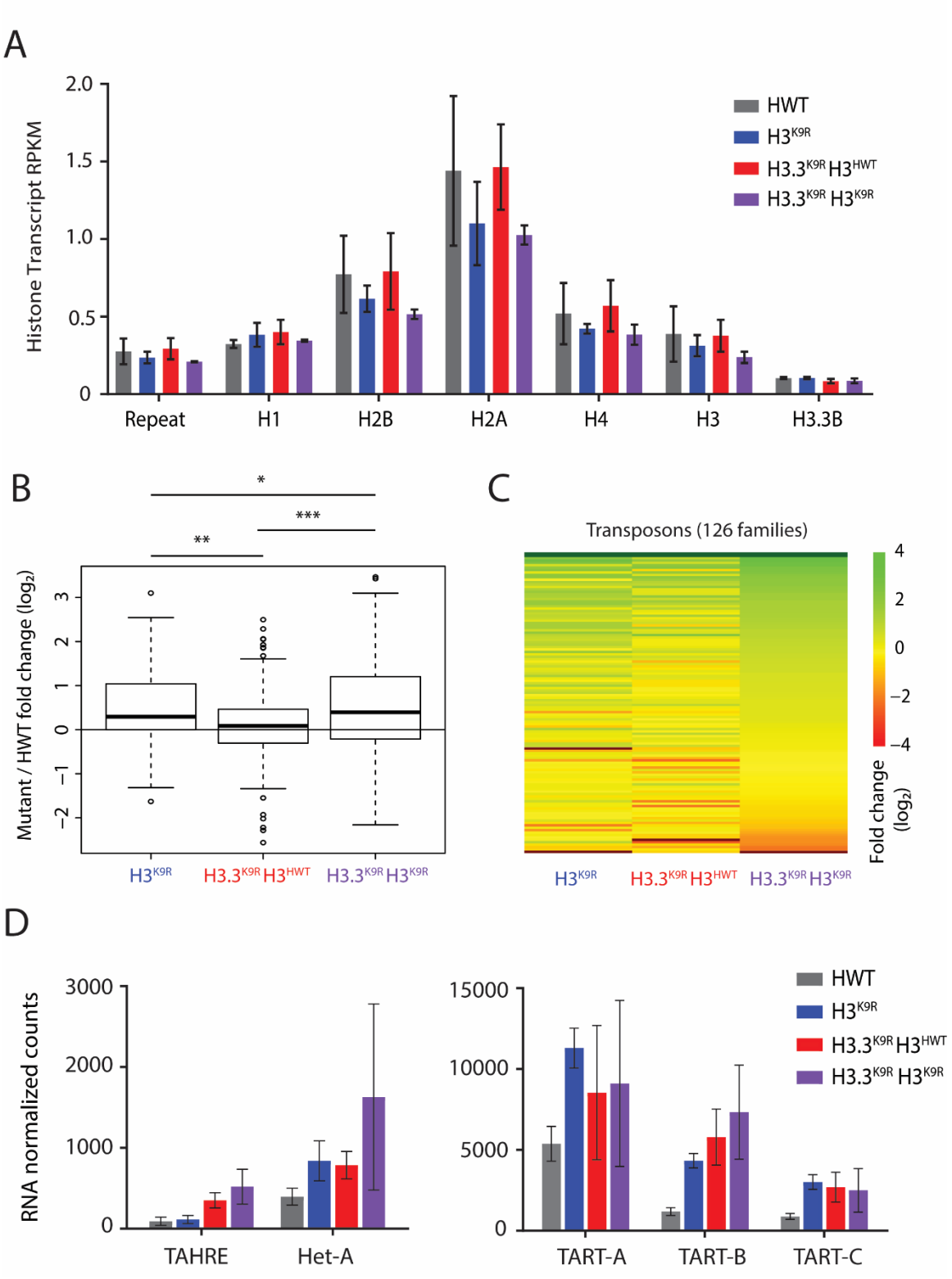
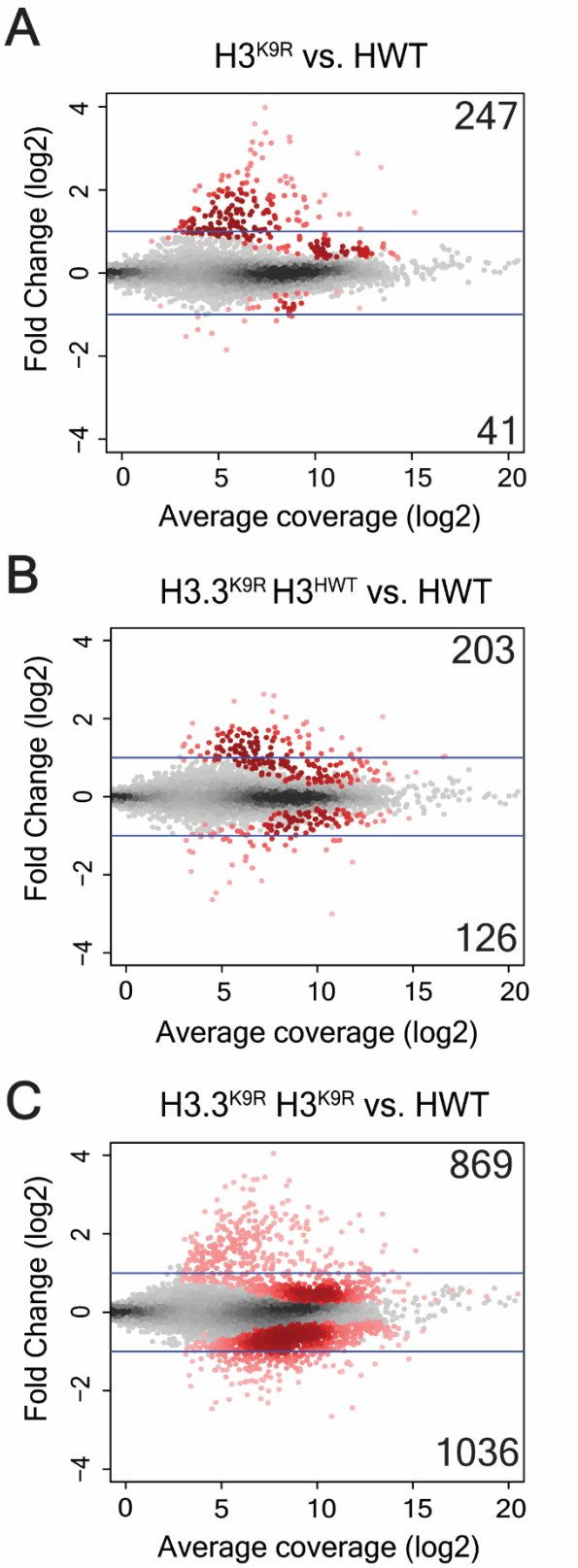

A
Histone Transcript RPKM
2.0
1.5
1.0
0.5
0
HWT
H3K9R
H3.3K9R H3HWT
H3.3K9R H3K9R
Repeat
H1
H2B
H2A
H4
H3
H3.3B
B
*
***
**
Mutant / HWT fold change (log2)
3
2
1
0
−1
−2
H3K9R
H3.3K9R H3HWT
H3.3K9R H3K9R
C
Transposons (126 families)
4
2
0
−2
−4
Fold change (log2)
H3K9R
H3.3K9R H3HWT
H3.3K9R H3K9R
D
RNA normalized counts
3000
2000
1000
0
TAHRE
Het-A
15000
10000
5000
0
TART-A
TART-B
TART-C
HWT
H3K9R
H3.3K9R H3HWT
H3.3K9R H3K9R

Figure 3.13: H3.3K9 and H3K9 redundantly regulate gene expression. A-C) Mutant: WT ratio of $H3^{K9R}$ (A), $H3.3^{K9R}$ (B), or $H3.3^{K9R}$ $H3^{K9R}$ (C) RNA signal from 1st instar larvae at 10,253 transcripts assembled by Cufflinks. The Y axis indicates the $\log_2$ transformation of mutant/control signal between the genotypes being compared (indicated at the top of each plot). Red dots indicate significantly different transcripts ($p<0.05$) and insets signify the number of significantly increased (top) or decreased (bottom) transcripts. Average coverage signifies the average number of normalized reads that overlap a transcript in mutant and $H3^{HWT}$ samples.


A
H3K9R vs. HWT
247
41
Fold Change (log2)
Average coverage (log2)
B
H3.3K9R H3HWT vs. HWT
203
126
Fold Change (log2)
Average coverage (log2)
C
H3.3K9R H3K9R vs. HWT
869
1036
Fold Change (log2)
Average coverage (log2)

Because we observed increases in FAIRE signal at transposons in *H3.3^{K9R}* mutants from wing disc samples, we examined RNA levels of transposon families in 1st instar larvae. Similar to our previous RNA-seq observations from *H3^{K9R}* mutant wing discs (Penke et al. 2016), RNA signal at transposons in *H3^{K9R}* 1st instar larvae was increased relative to the *HWT* control (Figure 3.12B, C). Although on average transposon levels were only slightly higher in *H3.3^{K9R} H3^{HWT}* mutants compared to *HWT*, transposon levels in *H3.3^{K9R} H3^{K9R}* combined mutants were significantly higher than either *H3.3^{K9R} H3^{HWT}* or *H3^{K9R}* mutants alone (Figure 3.12B, C). Moreover, telomeric transposons are generally increased in all K9R mutants compared to *HWT* controls (Figure 3.12D). Together these results support an overlapping role for H3.3K9 and H3K9 in regulating gene expression and transposon repression.

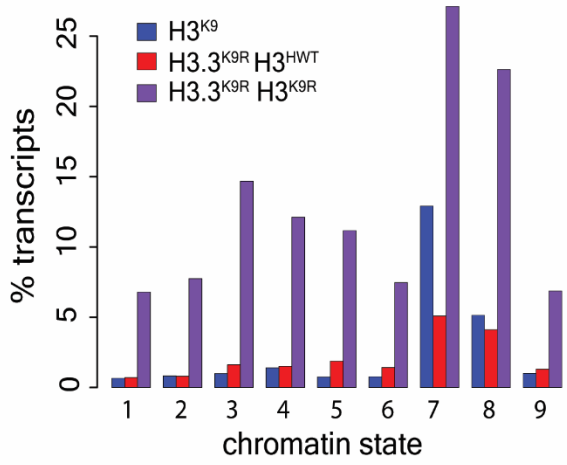
We next examined chromatin signatures of significantly altered transcripts to explore the mechanism of the observed gene expression changes. All transcripts were assigned to one or more chromatin states based on their overlap with genomic regions defined by Kharchenko et al. (2010). We then determined the percentage of transcripts within a given chromatin state that were either increased or decreased in K9R mutants relative to *HWT* controls (Figure 3.14 A-C). Transcripts in regions of K9me2/me3 (chromatin state 7 and 8) were the most likely to have significantly increased RNA levels in mutants compared to *HWT*. Although *H3.3^{K9R} H3^{K9R}* combined mutants had the highest percentage of chromatin state 7 transcripts that were significantly increased (~26%), *H3^{K9R}* mutants also displayed a high percentage (~13%) of change within chromatin state 7 (Figure 3.14A, D, Figure 3.15A). These results suggest that H3.3K9

contributes to gene repression in regions of K9me2/me3 but cannot completely compensate for the absence of H3K9.

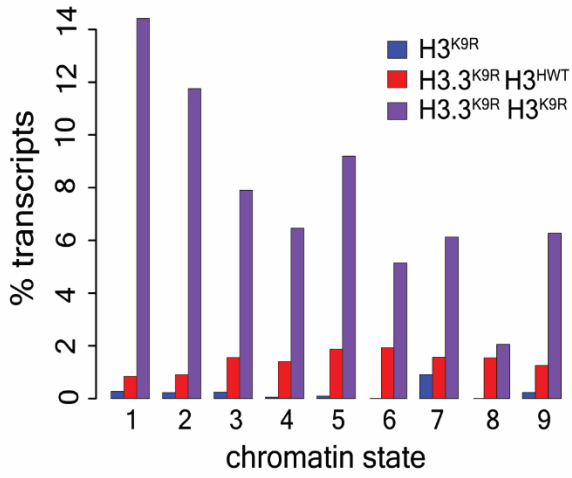
In contrast to upregulated transcripts, very few transcripts were significantly decreased in *H3.3*^{K9R} *H3*^{HWT} or *H3*^{K9R} mutants. However, the *H3.3*^{K9R} *H3*^{K9R} combined mutant displayed numerous significant decreases in transcript abundance. Interestingly, transcripts in chromatin state 1, characterized by K9ac, were most likely to be decreased (Figure 3.14B, E). Several other chromatin states showed elevated transcript changes, particularly in the *H3.3*^{K9R} *H3*^{K9R} combined mutant; however, in this analysis transcripts can be assigned to more than one chromatin state. Indeed, many transcripts in chromatin state 1 also overlap other chromatin states. We therefore performed a supplementary analysis that examined only transcripts that overlap a single chromatin state. This analysis demonstrated that transcripts solely in chromatin state 1 were much more likely to change in K9R mutants than those in other chromatin states (Figure 3.15A). Similar results were obtained using imaginal wing disc K9ac ChIP data from Pérez-Lluch et al. (2015). Whereas few transcripts that overlapped K9ac were significantly altered in either single mutant (68 in *H3*^{K9R} and 116 in *H3.3*^{K9R} *H3*^{HWT}mutants), 1195 K9ac associated transcripts exhibited changed expression levels in *H3.3*^{K9R} *H3*^{K9R} combined mutants (Figure 3.16). These data suggest that in regions of K9ac, H3.3 and H3 can completely compensate for each other. Additionally, these data provide strong evidence that K9ac facilitates gene expression.

Figure 3.14: H3.3K9 and H3K9 redundancy differs in heterochromatin and euchromatin. A, B) Percentage of transcripts in a chromatin state that have significantly increased (A) or decreased (B) RNA signal in mutants versus HWT. C) Table indicates the number of transcripts that overlap a particular chromatin state. D, E) Heatmaps showing fold change of K9R mutants over *HWT* at chromatin state 7 regions (D) and chromatin state 1 regions (E). Each row indicates a transcript that overlaps the indicated chromatin state.

A



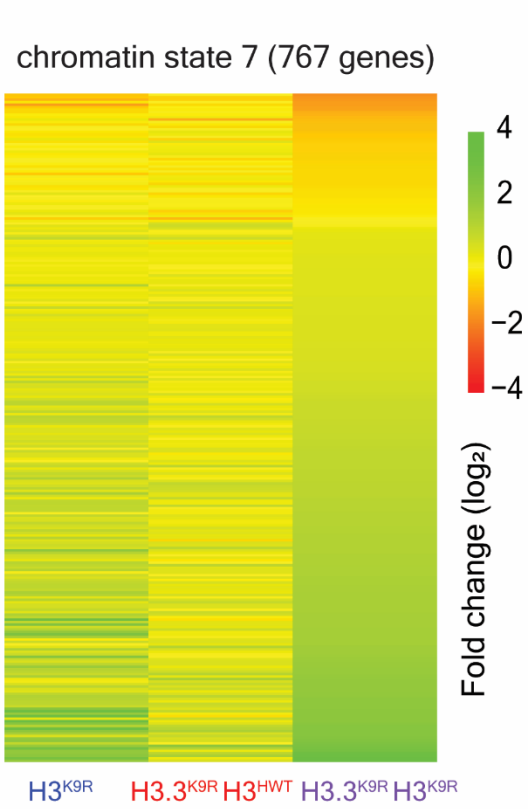
B



C

state	definition	genes
1	H3K4me2/3, H3K9ac	4848
2	H3K36me3	3641
3	H3K18ac, H3K27ac	1608
4	H3K36me	1857
5	H4K16ac, H3K36me3	1066
6	H3K27me2/3	777
7	H3K9me2/3	767
8	Low H3K9me2/3	389
9	H1, SUUR	4004

D



E

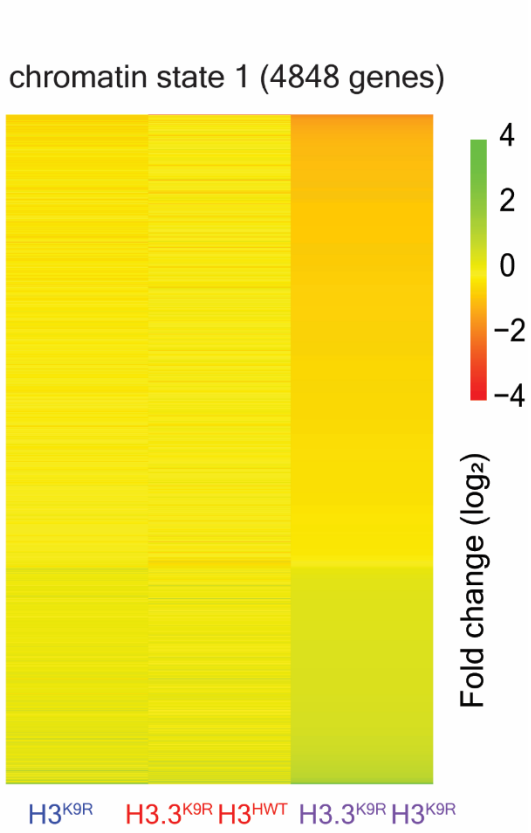


Figure 3.15: H3.3K9 compensates for H3K9 at regions of H3K9ac and partially at regions of H3K9me. A) Heatmap showing fold change of K9R mutants over *HWT* at regions of H3K9ac (left) and H3K9me (right). Each row indicates a transcript that overlaps the indicated chromatin state. B) Barplot similar to Figures 4D and 4E analyzing only those transcripts that overlap a single chromatin state.

A

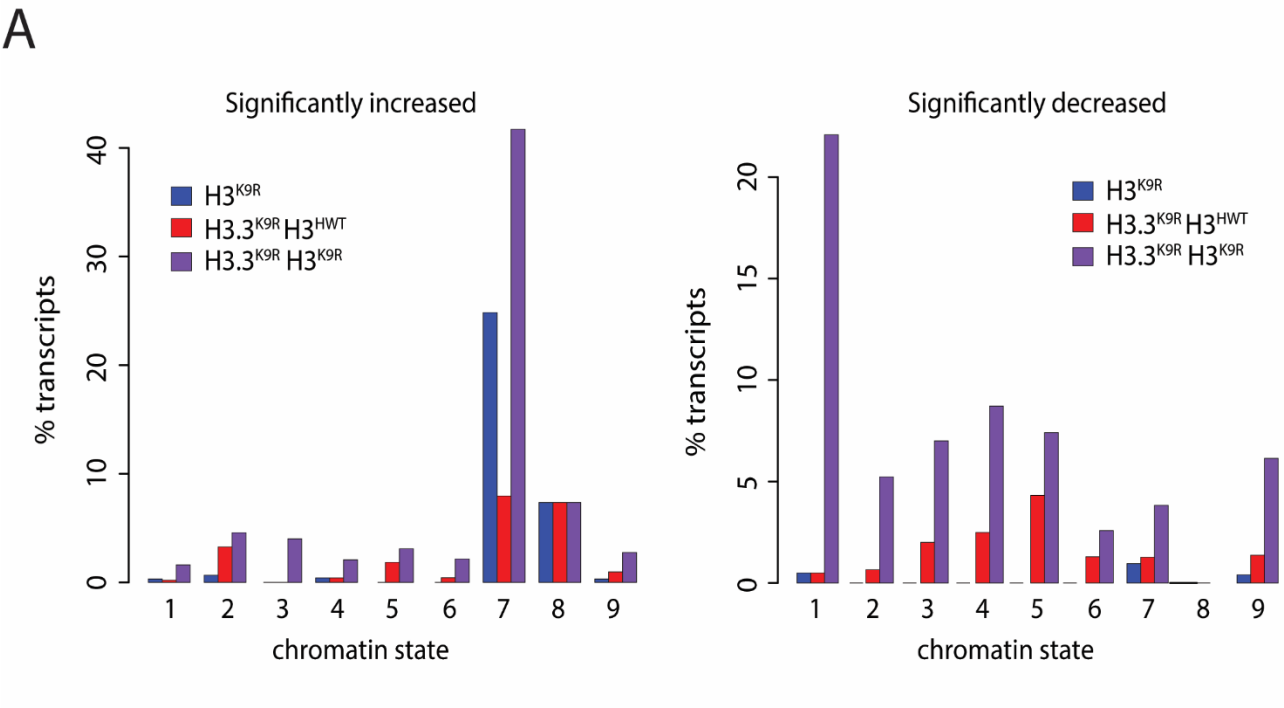
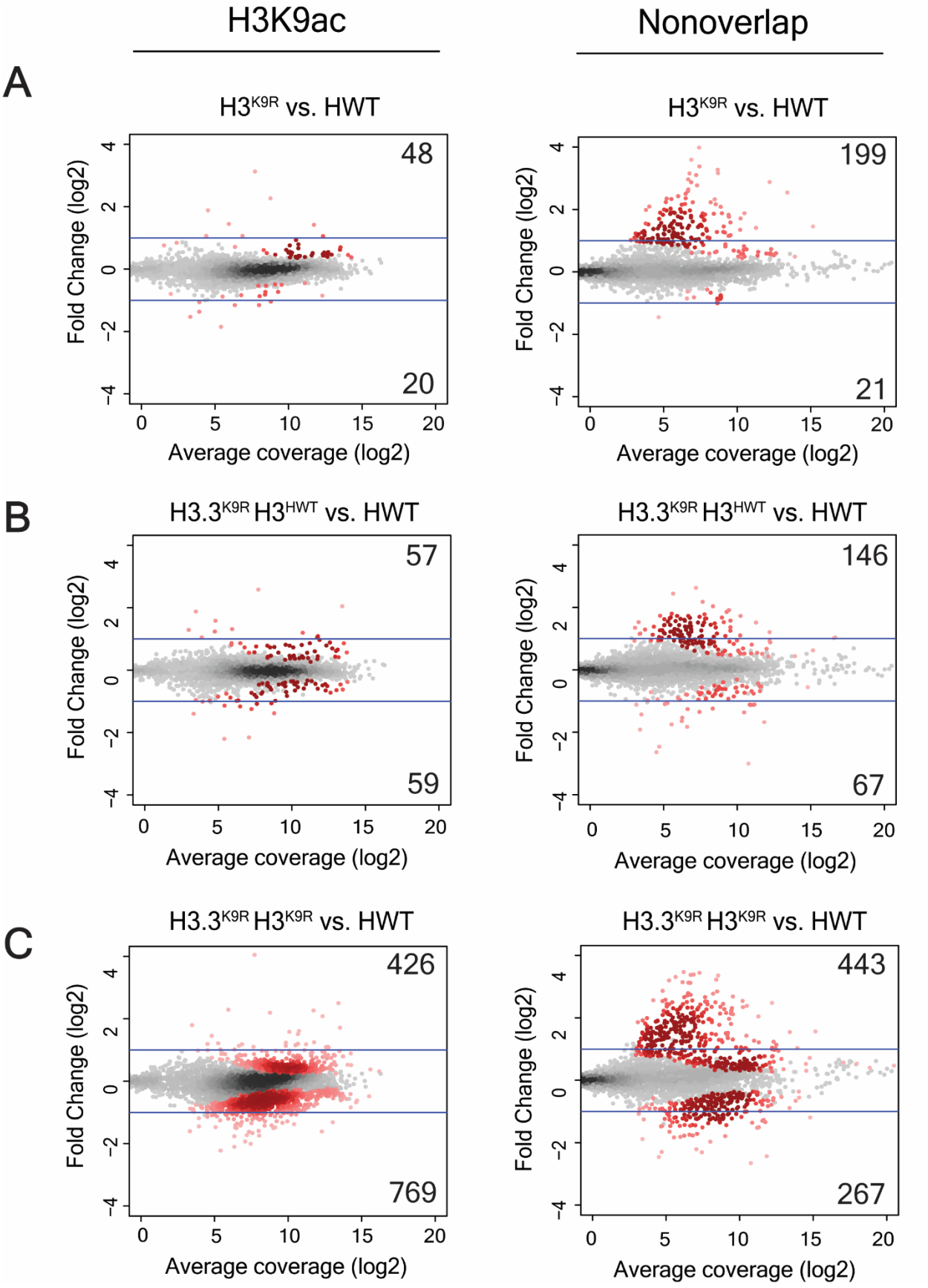

Significantly increased
% transcripts
0
10
20
30
40
H3K9R
H3.3K9R H3HWT
H3.3K9R H3K9R
1
2
3
4
5
6
7
8
9
chromatin state
Significantly decreased
% transcripts
0
5
10
15
20
H3K9R
H3.3K9R H3HWT
H3.3K9R H3K9R
1
2
3
4
5
6
7
8
9
chromatin state

Figure 3.16: K9ac associated transcripts are altered in *H3.3^{K9R} H3^{K9R}* double mutants. MA plot showing fold change of normalized RNA signal in *H3^{K9R}* (A), *H3.3^{K9R} H3^{HWT}* (B), and *H3.3^{K9R} H3^{K9R}* (C) mutants versus *HWT* at all transcripts from merged transcriptome. Average coverage on X-axis represents the mean expression level of a transcript. Transcripts that overlap an K9ac peak called from ChIP-seq data (GSM1363590 ; Pérez-Lluch et al. 2015) are shown in the left panel while those that do not are shown in the right panel. Significance (shown in red) was determined using DESeq2 (Love et al. 2014a) and an adjusted p value cutoff of 0.05.


H3K9ac
Nonoverlap
A
H3K9R vs. HWT
48
20
199
21
B
H3.3K9R H3HWT vs. HWT
57
59
146
67
C
H3.3K9R H3K9R vs. HWT
426
769
443
267
Fold Change (log2)
Average coverage (log2)

Discussion

Overlapping and distinct developmental functions of H3 and H3.3

In this study, we determined the distinct and overlapping roles that lysine 9 of variant and canonical histone H3 play in gene expression and heterochromatin function during *Drosophila* development. Our developmental genetic analyses demonstrate that H3.3K9 is necessary for fertility but not viability in *Drosophila*. In addition, we find that some euchromatic functions of H3K9 can be provided by either variant H3.3 or canonical H3, whereas H3.3K9 cannot completely compensate for H3K9 in some regions of heterochromatin as discussed below.

Several studies from multiple species have investigated the developmental functions of H3.3 and H3. In mice, single mutation of either *H3.3A* or *H3.3B* results in reduced viability and compromised fertility (Bush et al. 2013; Couldrey et al. 1999). Similarly, *Drosophila H3.3A* and *H3.3B* double mutants appear at lower than expected Mendelian ratios and are sterile (Sakai et al. 2009). H3.3 in *Tetrahymena thermophila* is also important for sexual reproduction, although it is not required for viability or maintenance of nucleosome density (Cui et al. 2006). Both *Tetrahymena* and *Drosophila* H3.3 and H3 can compensate for one another. In *Tetrahymena*, canonical H3 is dispensable if H3.3 is overexpressed (Cui et al. 2006). Similarly in *Drosophila*, transgenic expression of H3 can rescue both the semi-lethality (Sakai et al. 2009) and infertility (Hödl and Basler 2012) of *H3.3* mutants, indicating some functional redundancy between the two histones. Indeed, when expressed equivalently, *Drosophila* H3.3 can provide all of the developmental functions of H3 (Hödl and Basler

2012). Moreover, H3.3 is the sole H3 protein in *S. pombe* and *S. cerevisiae* yeast (Malik and Henikoff 2003).

H3.3K9 functions in heterochromatin

We find that under endogenous expression conditions, H3.3K9 functions in heterochromatin, including pericentromeric and telomeric regions of the genome. We detected H3.3K9 methylation in pericentromeric heterochromatin, congruous with previous data demonstrating that H3.3 in *Drosophila* is deposited at the chromocenter of polytene chromosomes in a replication-dependent manner (Schwartz and Ahmad 2005). We also observed that $H3.3^{K9R}$ mutants exhibited an abnormal chromocenter structure in polytene chromosomes. Moreover, we provide evidence that H3.3K9 is required for maintenance of telomeric chromatin architecture and repression of certain telomeric transcripts, indicating that replication-coupled expression of H3 cannot provide these particular H3K9 functions. These findings in *Drosophila* are consistent with studies in mouse embryonic stem cells showing that H3.3 is localized to telomeres, is methylated at K9, and functions in repression of telomeric repeat-containing RNAs (Goldberg et al. 2010; Udugama et al. 2015). Conversely, the genetic data we presented here and previously (Penke et al. 2016) indicate that H3K9 is essential for repression of transposon-derived transcripts in pericentric heterochromatin, and H3.3K9 cannot compensate for the lack of H3K9 at these regions of the genome. The role of H3.3K9 in telomere structure and function may be independent of HP1, as HP1 is recruited to telomeres via the terminin complex independently of H3K9me (Raffa et al. 2011; Vedelek et al. 2015; Badugu et al. 2003).

K9ac regulates euchromatic gene expression

Previous studies that mapped histone modifications across the genome identified K9ac as a characteristic of transcriptionally active regions (Kharchenko et al. 2011; Bernstein et al. 2005; Liang et al. 2004; Roh et al. 2005). Moreover, mutation of H3K9 acetyltransferases results in compromised transcriptional activity (Wang et al. 1998; Georgakopoulos and Thireos 1992; Kuo et al. 1998). However, H3K9 acetyltransferases have non-histone substrates in addition to H3K9, and decreased transcriptional output may be the result of pleiotropic effects (Glozak et al. 2005; Spange et al. 2009; Fillingham et al. 2008). Our study provides evidence that K9ac, rather than non-histone targets of H3K9 acetyltransferases, contributes to activating transcription, as *H3.3*^{K9R} and *H3*^{K9R} mutants exhibit reduced gene expression in regions normally enriched for K9ac. Importantly, these K9ac rich regions with reduced gene expression are not normally enriched in K9me2 or me3, indicating the observed phenotype is not due to changes in K9me2 or me3 and likely results from loss of K9ac. This change in gene expression was accompanied by a fully penetrant lethality early in larval development of *H3.3*^{K9R} *H3*^{K9R} combined mutant animals, raising the possibility that gene expression control via acetylation of H3K9 is critical for the completion of animal development. These data are also consistent with previous studies in *C. elegans* demonstrating that H3K9 methylation is not essential for viability (Towbin et al. 2012; Zeller et al. 2016).

Overlapping and distinct genomic functions of H3K9 and H3.3K9

Functional overlap of H3K9 and H3.3K9 appears to vary at different regions of the genome. Whereas H3.3K9 and H3K9 can perform similar functions in euchromatic regions of the genome and can fully compensate for one another, our RNA-seq data

demonstrate H3.3K9 can only partially compensate for H3K9 in regions of heterochromatin. Partial compensation by H3.3K9 in regions of K9me2/me3 is in line with previous studies showing H3.3 is found at heterochromatin (Goldberg et al. 2010; Lewis et al. 2010; Wong et al. 2010) and plays a role in transposon repression (Elsässer et al. 2015). In the genotypes we analyzed, mRNA encoding variant and canonical H3 are expressed from their native promoters. Thus, disparity in functional overlap might be due to differences in modes of expression and deposition and thus total amounts of variant and canonical H3 histones in particular regions of the genome. For instance, H3 is normally enriched in heterochromatin compared to H3.3 (Ahmad and Henikoff 2002), which may cause *H3*^{K9R} mutations to be more detrimental in these regions. However, H3.3 may be able to provide all H3 function when highly expressed in a replication-dependent manner, as a transgenic histone gene array in which the *H3.2* coding region was replaced by *H3.3* is nearly fully functional in larval imaginal discs (Hödl and Basler 2012). Thus, differences in expression and/or deposition into chromatin may be the only basis for functional differences between H3.3 and H3.2 that we observed.

Heterochromatin may be particularly sensitive to incorporation of non-modifiable K9 residues. H3K9 methylation serves as a binding site for the protein HP1, which can in turn recruit H3K9 methyltransferases (Elgin and Reuter 2013; Grewal and Jia 2007). Methylation of a neighboring nucleosome can restart the cycle and initiate propagation of a heterochromatic configuration along the chromosome. Introduction of even a small number of H3K9R containing nucleosomes may therefore disrupt this cycle and prevent proper heterochromatin formation and gene repression. Incorporation of H3.3B^{K9R} histones into regions of heterochromatin may dominantly affect chromatin structure,

resulting in the observed phenotypes at pericentromeres and telomeres in H3.3^{K9R} mutants. In contrast, incorporation of low amounts of H3K9R histones in euchromatin may not reduce K9ac levels sufficiently to disrupt gene expression. Finally, amino acid differences in variant and canonical H3 may direct distinct histone modification states on different histone types by influencing the binding of chromatin modifying enzymes (Jacob et al. 2014). Different histone modification states on H3.3 and H3 may underlie variation in compensation at different genomic regions.

In sum, our data investigating H3.3K9 and H3K9 function provide evidence that K9ac activates gene expression and advance our understanding of the overlapping and distinct functional roles of variant and canonical histones.

CHAPTER 4- DNA REPLICATION INITATES WITHIN A PERMISSIVE RATHER THAN DETERMINISTIC CHROMATIN LANDSCAPE

Introduction

Animal cells duplicate large complex genomes by regulating where in the genome, and when during S phase, DNA replication initiates. An evolutionarily conserved feature of this regulation is a temporal order of DNA replication initiation events resulting in characteristic early and late replicating regions of the genome (Rhind and Gilbert 2013). Such “replication timing” programs appear at early stages of animal development and ensure genome integrity during cell proliferation (Mantiero et al. 2011; Collart et al. 2013; Yuan and Farrell 2016; Hamperl and Cimprich 2016). Importantly, replication timing is associated with mutational burden and SNP density, as spontaneous mutations occur less frequently in early compared to late replicating regions of the genome(Stamatoyannopoulos et al. 2009; Donley and Thayer 2013). Furthermore, perturbed replication timing is thought to be an early epigenetic event that predisposes both cancer development and disease-associated genome rearrangement (Donley and Thayer 2013; Ryba et al. 2012). Notwithstanding their importance, mechanisms that regulate where and when DNA replication initiates within an animal genome remain poorly understood.

In contrast to replication initiation in single celled eukaryotes such as budding yeast, replication of animal genomes does not initiate at well-defined sequence motifs(Bell, Stephen P.; Stillman 1992; Miotto et al. 2016; MacAlpine et al. 2010). Rather,

two levels of genome organization have emerged as putative regulators of replication initiation: arrangement of DNA within the three dimensional space of the nucleus, and local chromatin structure, characterized in part by domains of differential DNA accessibility (i.e. differential nucleosome occupancy) (Pope et al. 2014; Hiratani et al. 2008). Current models posit that these features of genome organization regulate replication by influencing *trans*-acting factor recruitment to sites of replication initiation (i.e. origins) (Mantiero et al. 2011; Collart et al. 2013; Miotto et al. 2016; Pope et al. 2014; Das et al. 2015; Rivera-Mulia and Gilbert 2016). In all metazoan organisms examined to date, transcriptionally active, accessible euchromatin generally replicates early during S-phase, whereas transcriptionally repressive, inaccessible heterochromatin generally replicates late (Lubelsky et al. 2014; Eaton et al. 2011; Bell et al. 2010). Despite strong genome-wide correlations between replication and chromatin structure in animal cells, a causal relationship between the two has yet to be determined, largely due to imprecise methods for manipulating chromatin structure *in vivo*.

To ascertain how chromatin structure influences genome duplication, we asked if altering the distribution of accessible chromatin throughout the genome would affect the normal DNA replication program. Previous strategies to manipulate chromatin structure typically involved perturbation of factors that establish, interpret, or remove histone post-translational modifications (PTMs) (Beck et al. 2012). Although informative, these studies cannot precisely determine functional roles for histone PTMs in DNA replication because most histone-modifying enzymes also have non-histone substrates that may also participate in DNA replication (Glozak et al. 2005; Huang and Berger 2008; Sims and Reinberg 2008; Carlson and Gozani 2016). To reduce potential pleiotropic effects of

mutating histone-modifying enzymes, we employed a strategy in *Drosophila* to more precisely manipulate chromatin structure by mutating the histone genes themselves, an approach that is not currently feasible in other animal models. This strategy involves deleting the endogenous wild-type histone genes and replacing them with transgenic copies encoding a single amino acid substitution that prevents PTMs on a particular histone residue without concurrently affecting non-histone proteins (McKay et al. 2015). By determining how changes in chromatin structure in histone mutants affect DNA replication initiation throughout the genome, we can establish causal relationships between chromatin and genome duplication.

Materials and Methods

Sample preparation for FACS and sequencing

Overnight collections of HWT and H3K9R embryos were prepared as in (Penke et al. 2016) and dechorionated prior to embryo sorting. A Union Biometrica BioSorter for large particle flow cytometry equipped with a 488-nm solid state laser and accompanying FlowPilot software was used for identification and high throughput isolation of HWT and H3K9R animals.

Third instar wing imaginal discs were dissected over a period of four hours and stored in Grace's insect medium (supplemented with L-Glutamine, 3.33g/L Lactalbmin Hydrolysate, and 3.33g/L Yeastolate) on ice prior to nuclear isolation. Isolated nuclei were stained with 1.5µg/mL DAPI prior to FACS. Nuclei were sorted into G1, S, and G2 populations based on DNA content as measured by DAPI intensity on a FACSAria II or III (using NEB-0.1% Tween sheath). Isolated populations of nuclei were pelleted, flash

frozen, and stored at -80°C prior to DNA isolation and library preparation. Libraries were prepared with the Rubicon ThruPLEX DNA-seq kit and sequencing was performed on an Illumina HiSeq 2500 in the UNC-Chapel Hill High Throughput Sequencing Facility.

Sequence data analysis

Replication Timing Profiles

Reads from G1, S, and G2 samples were aligned to the dm6 reference genome (Release 6.04) using Bowtie2 (v2.3.2) default parameters. Two S phase replicates were generated for each genotype. Reads with a MAPQ score greater than 10 were retained using Samtools (v1.6). To generate a replication timing (RT) value for a particular window (100kb with a 10kb slide), the reads per million (RPM) value of each S phase replicate was divided by the RPM G1 value and averaged. RT profiles were generated by plotting the RT value at each window versus the genomic location. The limma statistical package was used to identify windows with significantly altered RT values between HWT and H3K9R animals (lmFit, adjusted p value, $p<0.01$). Coordinates of various chromatin states were obtained from Kharchenko et al. (Kharchenko et al. 2011) and converted to dm6 coordinates.

Wild-type Replication Timing Characterization

To calculate replication domain sizes, we identified the genomic coordinates halfway between each peak and valley of an RT profile and determined the distance from one halfway point to the next. We used modENCODE ChIP-seq data from whole 3rd instar larvae to calculate histone PTM enrichment at 100kb windows across the genome (ftp://data.modencode.org/D.melanogaster/Histone-Modification/ChIP-seq/raw-seqfile_fastq/). For each histone PTM, raw reads for two ChIP replicates and two input

replicates were aligned to the genome using Bowtie2 (v2.3.2) (Langmead and Salzberg 2012). Bedtools coverage (v2.25.0) was used to count the number of reads mapping to each 100kb window, and the resulting counts were normalized to read depth. Histone PTM enrichment for each replicate was calculated by dividing the ChIP normalized read counts by the input for each replicate; the resulting values were then averaged. All windows were ordered by RT value and split into five equally sized categories (early, early/mid, mid, mid/late, and late). Average PTM enrichment values of all windows in a category were calculated and represented as a heatmap. RNA-seq data from 3rd instar imaginal wing discs (McKay and Lieb 2013) was used to calculate transcript density or transcript activity at 100kb windows: the number of transcripts overlapping each 100kb window was determined and the normalized read per million of each transcript overlapping a window was summed, respectively.

FAIRE, HP1a, and RNA Analyses

FAIRE-seq and RNA-seq from 3rd instar imaginal wing discs and HP1a ChIP-seq from whole 3rd instar larvae were obtained from GSE85374 (Penke et al. 2016). FAIRE and HP1a reads from three H3K9R and two HWT replicates were aligned to the genome using Bowtie2 (v2.25.0) default parameters (Langmead and Salzberg 2012). The number of reads overlapping 10 kb windows were normalized to read depth (FAIRE) or the number of uniquely mapping *D. virilis* spike-in reads (HP1a). edgeR (v3.16.5) was used to calculate windows with significantly altered FAIRE or HP1a fold-change (p value <0.01) (Robinson et al. 2009).

RNA reads from three HWT and three H3K9R replicates were aligned using TopHat default parameters (v2.1.1) (Trapnell et al. 2014), and a transcriptome was

generated using Cufflinks (v2.2.1, see above for parameters). We combined the Cufflinks generated transcriptome with transposons annotated by RepeatMasker (Smit et al. 2013). Raw counts of RNA reads at each transcript were used as input for edgeR statistical analysis (p value <0.01) (Robinson et al. 2009). We then identified transcripts within or that overlapped each 10kb window and selected the transcript with the lowest p value.

To determine RT values at 10kb windows, we used the previously calculated log2 fold change and p values from 100kb windows. For each 10kb window, we calculated the median fold change and median p value of all overlapping 100kb windows. We used RT values from 100kb windows as this size closely matches average replication domain size (~100-200kb), but similar results were obtained using RT values determined from 10kb windows. 10kb windows were identified as having significantly altered RT between H3K9R and HWT if the p value was below 0.05 and the absolute log2 fold change was at least 0.1. To focus our analysis on more mappable regions of the genome, we analyzed 10kb windows on the major chromosome scaffolds (chr2L, chr2R, chr3L, chr3R, chr4, and chrX) that had an average FAIRE and HP1 counts per million (CPM) value of greater than zero. Comparisons of RT, FAIRE, HP1a, and RNA signal between H3K9R and HWT samples were performed with all reads or “uniquely” mapping reads (MAPQ>10) with similar results.

To calculate transposon families with significantly altered RNA levels, we summed raw counts of all individual transposons within a family and used edgeR as described above to determine significance (p value <0.05). In addition to transposon families, all transcripts identified in Cufflinks were included in this edgeR analysis to facilitate

modeling of variability. Data was visualized using the Integrative Genomics Viewer (Robinson et al. 2011).

Results

To probe the relationship between chromatin structure and replication in an intact animal, we adapted a genome-wide measure of replication called Repli-seq for use in *Drosophila* larval tissues (Sasaki et al. 2017; Siefert et al. 2017; Koren et al. 2014). We chose wing imaginal discs because they are composed of an epithelium of proliferating diploid precursor cells. Repli-seq is based on the premise that, in a population of S-phase cells, early replicating DNA sequences are over-represented relative to late replicating ones, due to a higher probability of replication initiation (Mantiero et al. 2011; Das et al. 2015; Rhind et al. 2010; Collart et al. 2013). Consequently, Repli-seq data are a proxy for the propensity of replication initiation in a particular region of the genome. We performed whole-genome sequencing on DNA isolated from populations of G1- and S-phase nuclei collected from wing discs by fluorescence-activated cell sorting (FACS) (Figure 4.1A). Replication profiles were generated by determining the S/G1 read count at 100kb intervals using a 10kb slide across the genome (Materials and Methods; Figure 4.2A). Larger S/G1 read count values indicate earlier replication and smaller values indicate later replication. Replication timing values generated from independent S-phase samples were highly reproducible (Figure 4.3).

Figure 4.1. Repli-seq measures genome-wide replication timing in vivo. A) Repli-seq experimental outline: (1) Nuclei were FACS sorted into G1 (yellow), S (red) and G2 (blue) populations based on DNA content. (2) DNA was sequenced and mapped back to the genome. More reads map to early than late replicating sequences. (3) S/G1 $\log_2$ ratio of mapped reads generates replication timing profiles. Normalizing to G1 and G2 phase controls gave similar results (Materials and Methods). B) LOESS regression line showing average *yw* S/G1 ($\log_2$) replication timing values in 100kb windows using a 10kb slide across chromosome 2 and 3 scaffolds. Approximate locations of constitutive heterochromatin (green) and largely euchromatic regions (blue) are indicated (Riddle et al. 2011; Hoskins et al. 2015). C) Heatscatter plot of *yw* S/G1 ($\log_2$) and gene density at all 10kb windows across the genome with LOESS regression line (black). D) Heatmap of relative modENCODE histone PTM enrichment in bins of equally sized replication timing quintiles (early, early/mid, mid, mid/late, and late) generated using S/G1 ($\log_2$) RT values within 100kb windows and normalized modENCODE data from whole third instar larvae. Color indicates average enrichment of all windows within a quintile. Scale of heatmap was capped at 1.4 to better represent distribution of values, as H3K9me2/me3 was greatly enriched in late replicating domains compared to other PTMs (see Figure S1E for non-capped H3K9me2/me3 heatmap). E) Plot of transposon number in 100kb windows across chromosome 3R with replication timing quintile (as determined in D) indicated by color.

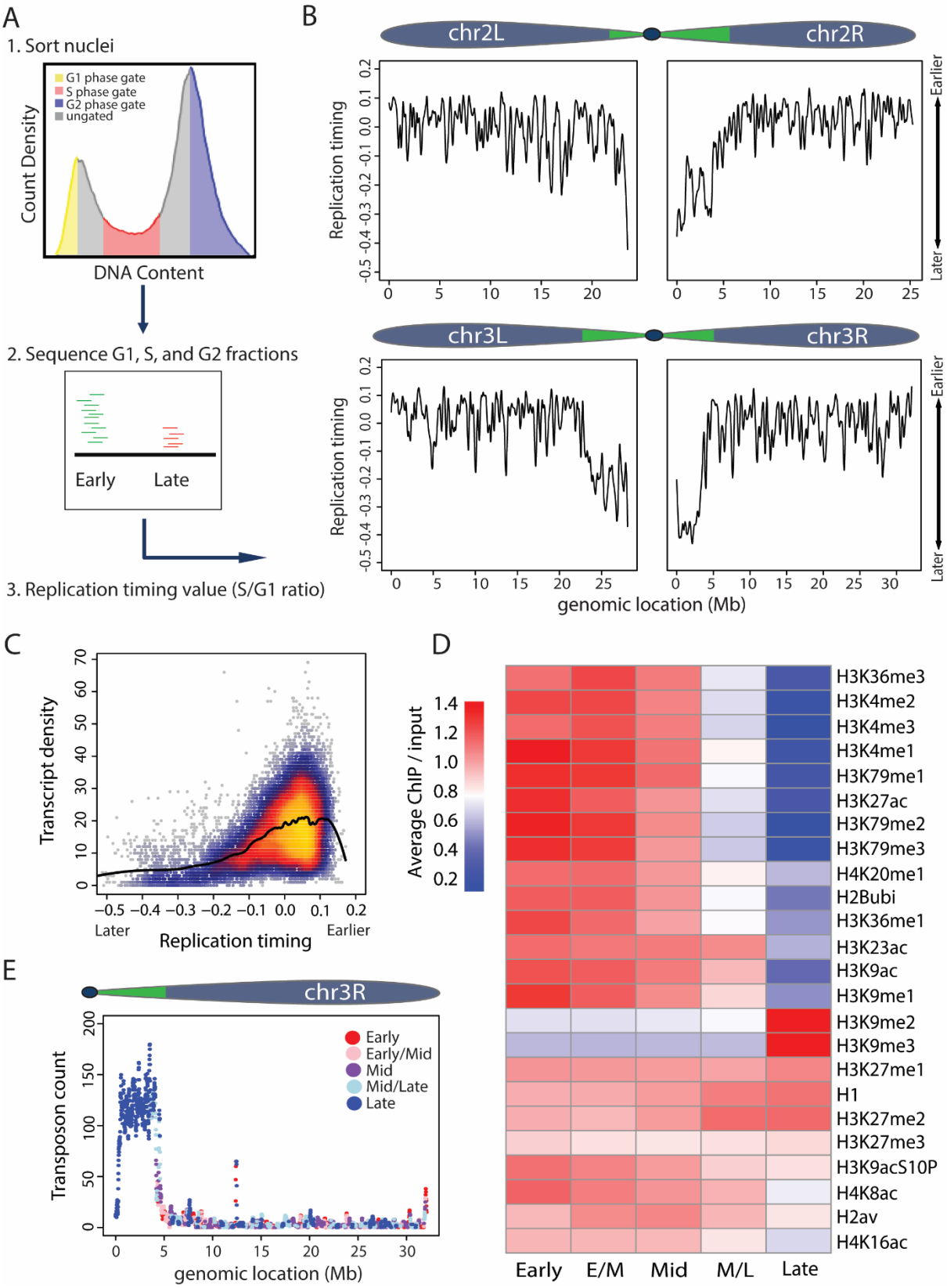

A
1. Sort nuclei
G1 phase gate
S phase gate
G2 phase gate
ungated
Count Density
DNA Content
2. Sequence G1, S, and G2 fractions
Early
Late
3. Replication timing value (S/G1 ratio)
B
chr2L
chr2R
chr3L
chr3R
Replication timing
Earlier
Later
genomic location (Mb)
C
Transcript density
Later
Replication timing
Earlier
D
Average ChIP / input
H3K36me3
H3K4me2
H3K4me3
H3K4me1
H3K79me1
H3K27ac
H3K79me2
H3K79me3
H4K20me1
H2Bubi
H3K36me1
H3K23ac
H3K9ac
H3K9me1
H3K9me2
H3K9me3
H3K27me1
H1
H3K27me2
H3K27me3
H3K9acS10P
H4K8ac
H2av
H4K16ac
Early
E/M
Mid
M/L
Late
E
chr3R
Transposon count
Early
Early/Mid
Mid
Mid/Late
Late
genomic location (Mb)

Figure 4.2. Replication timing in *Drosophila* wing discs correlates with features of active and repressive chromatin. A) Representative 5Mb region on chromosome 3R of S/G1 (log_2) replication timing values within 100kb windows with a 10kb slide. RT values are an average of replicate *yw* samples. LOESS regression line indicated in red. B) LOESS regression line showing average *yw* S/G1 (log_2) replication timing values in 100kb windows using a 10kb slide across chromosome X and 4 scaffolds. Approximate locations of constitutive heterochromatin (green) and largely euchromatic regions (blue) are indicated (Hoskins et al. 2015; Riddle et al. 2011). C) Histogram of *yw* replication domain sizes. D) Heatscatter plot of *yw* S/G1 (log_2) replication timing values and RNA expression levels (Materials and Methods) within all 10kb windows across the genome with LOESS regression line in black. E) Heatmap of relative H3K9me2 and H3K9me3 enrichment in bins of equally sized RT quintiles (early, early/mid, mid, mid/late, and late) generated using S/G1 (log_2) RT values within 100kb windows and normalized modENCODE H3K9me2/me3 data from whole third instar larvae. Color indicates average enrichment of all windows within each replication timing quintile. F) Average modENCODE histone PTM enrichment for all 100kb windows within each of the equally sized replication timing quintiles (E=early, E/M=early-mid, M=mid, M/L= mid-late, and L=late). G) Number of transposons within 100kb windows plotted versus genomic location. The color of each dot indicates the replication timing quintile of the window.

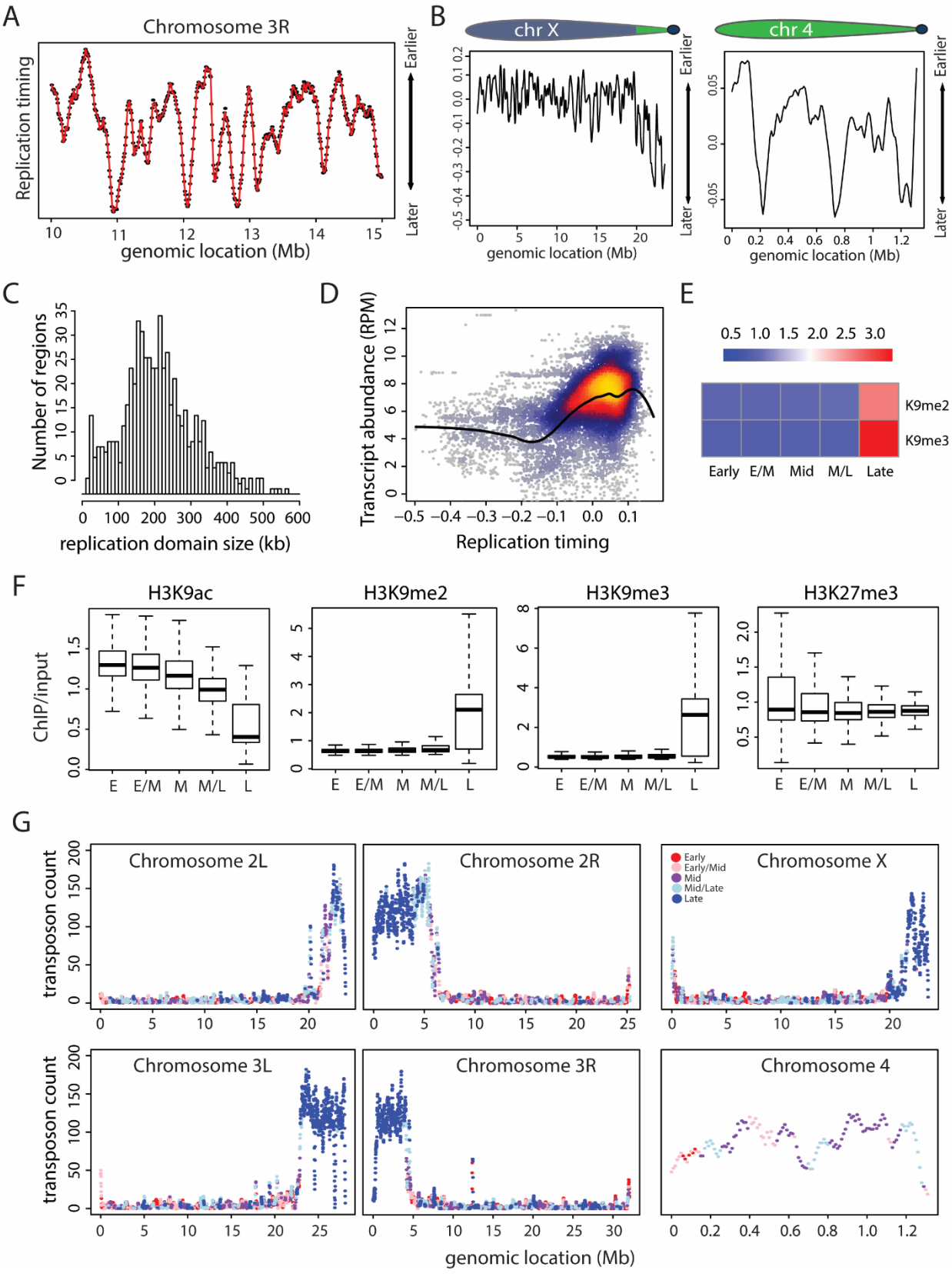
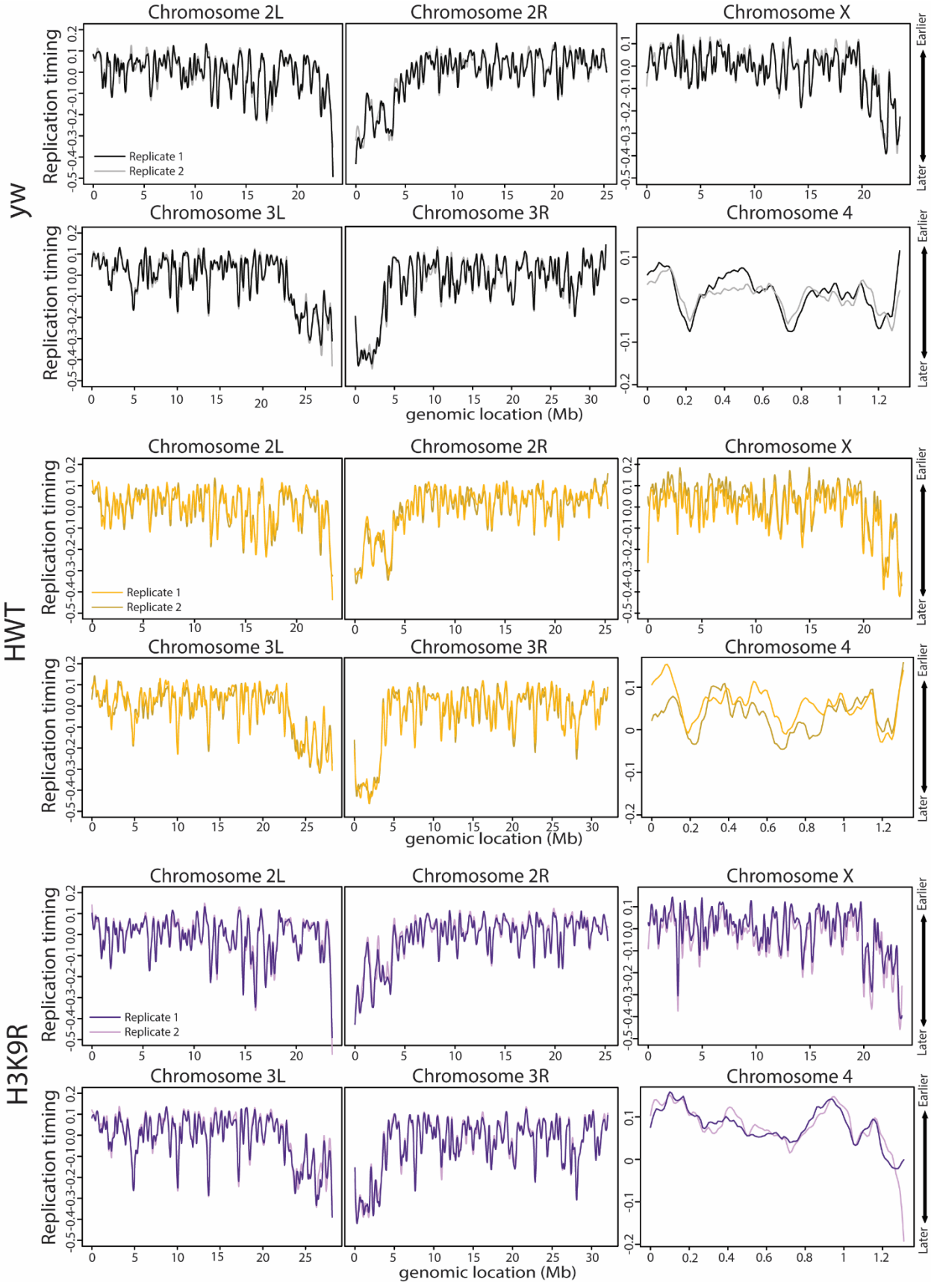
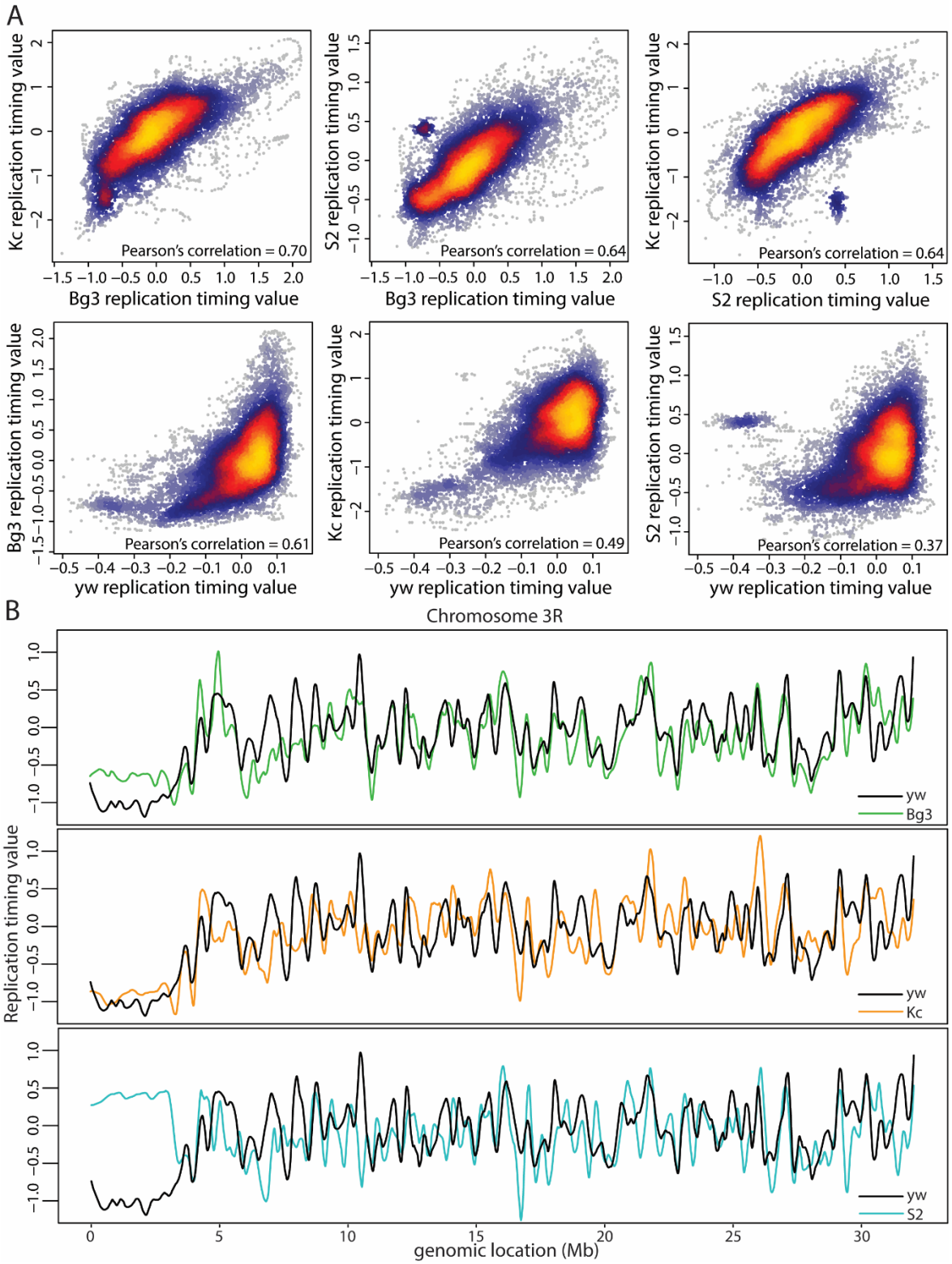

A
Chromosome 3R
Replication timing
genomic location (Mb)
Earlier
Later
B
chr X
chr 4
genomic location (Mb)
C
Number of regions
replication domain size (kb)
D
Transcript abundance (RPM)
Replication timing
E
K9me2
K9me3
Early E/M Mid M/L Late
F
H3K9ac
H3K9me2
H3K9me3
H3K27me3
ChIP/input
E E/M M M/L L
G
Chromosome 2L
Chromosome 2R
Chromosome X
Chromosome 3L
Chromosome 3R
Chromosome 4
transposon count
Early
Early/Mid
Mid
Mid/Late
Late
genomic location (Mb)

Figure 4.3. Replication timing profiling in *Drosophila* tissue is highly reproducible. Quantile normalized S/G1 (log_2) replication timing values for each replicate for the indicated genotypes were plotted versus genomic coordinate for all major chromosome scaffolds. Each replicate *yw*, *HWT*, and *H3K9R* profile is shown in a different shade of grey, yellow, and purple, respectively.


yw
HWT
H3K9R
Chromosome 2L
Chromosome 2R
Chromosome X
Chromosome 3L
Chromosome 3R
Chromosome 4
Replication timing
Replicate 1
Replicate 2
Earlier
Later
genomic location (Mb)

Repli-seq data revealed that in wild type wing discs the pericentromeric heterochromatin replicates later than the mostly euchromatic chromosome arms (Figure 4.1B; Figure 4.2B), consistent with prior cytological observations (Taylor 1960a). Despite replicating at largely different times on average, both pericentromeric and euchromatic regions contained earlier and later replicating domains within them, such that the earliest replicating domains in pericentromeric heterochromatin exhibited similar values to the latest replicating domains on chromosome arms. Replication domain sizes ranged from 20kb-570kb (Figure 4.2C), closely matching previous measurements (MacAlpine et al. 2004). Additionally, our wing disc replication profiles are similar to those previously generated from *Drosophila* cell lines and most closely correlate with replication timing data obtained from a cell line derived from the same developmental stage as wing discs (Figure 4.4) (Lubelsky et al. 2014). Consistent with previous studies in zebrafish embryos and in *Drosophila* and mammalian cultured cells (Lubelsky et al. 2014; Eaton et al. 2011; Bell et al. 2010; Siefert et al. 2017; Martin et al. 2011; Petryk et al. 2016), we found that earlier replication correlates with higher gene density (Figure 4.1), higher levels of transcription (Figure 4.2D), and activating histone PTMs such as H3K4me and H3K9ac (Figure 4.1D). In contrast, later replication occurred in gene-poor regions (Figure 4.1C) and was enriched in transposons (Figure 4.1E, Figure 4.2G) and repressive histone PTMs, such as H3K9me2/me3 (Figure 4.1D, Figure 4.2E,F). Thus, Repli-seq reveals highly reproducible replication profiles from *Drosophila* tissue that match general features of replication found in other systems.

Figure 4.4. Wild-type 3rd instar imaginal wing discs and cell culture replication timing profiles are highly correlated. A) Heatscatter plot of S/G1 (log_2) replication timing value at 100kb windows from *yw* imaginal wing discs and previously generated timing profiles from three *Drosophila* cell culture lines (Kc, S2, and Bg3) (Lubelsky et al. 2014). Top row shows correlations between each of the three cell culture lines, and bottom row shows the correlations between *yw* wing discs and the three cell culture lines. Windows with earlier timing values in S2 cells compared to other cell types are located in the pericentromeres and may be due to copy number differences in these regions. B) Comparison of replication timing profiles between *yw* wing discs and each of the three cell culture lines on chromosome 3R.


A
Kc replication timing value
Bg3 replication timing value
Pearson's correlation = 0.70
S2 replication timing value
Bg3 replication timing value
Pearson's correlation = 0.64
Kc replication timing value
S2 replication timing value
Pearson's correlation = 0.64
Bg3 replication timing value
yw replication timing value
Pearson's correlation = 0.61
Kc replication timing value
yw replication timing value
Pearson's correlation = 0.49
S2 replication timing value
yw replication timing value
Pearson's correlation = 0.37
B
Chromosome 3R
Replication timing value
yw
Bg3
yw
Kc
yw
S2
genomic location (Mb)

To determine how chromatin structure influences replication, we altered the relative proportion of accessible and inaccessible chromatin using an H3K9R mutation. Previously we showed that H3K9R mutants are significantly depleted of methylated H3K9 (H3K9me) and Heterochromatin Protein 1a (HP1a) within heterochromatin (Penke et al. 2016). HP1a binds H3K9me and facilitates heterochromatin formation through multimerization of HP1a molecules and recruitment of *trans*-acting factors (Canzio et al. 2011; Azzaz et al. 2014). In addition, we found that loci within H3K9R mutant pericentromeric heterochromatin are depleted of nucleosomes, as measured by increased FAIRE-seq signal (Formaldehyde-Assisted Isolation of Regulatory Elements), relative to controls (Penke et al. 2016). We interpret this elevated FAIRE-seq signal as an indication that heterochromatin has become more accessible in H3K9R mutants. Due to established correlations between DNA accessibility within chromatin and early replication, we expected large-scale advancement of replication timing at nucleosome-depleted H3K9R pericentromeres. Surprisingly, Repli-seq data revealed that ~96% of the H3K9R genome has similar replication timing compared to control (Figure 4.5A; Figure 4.6), consistent with FACS analysis indicating that cell cycle phasing is only slightly perturbed in H3K9R mutant cells (Figure 4.5C). These data suggest that replication proceeds normally across most of the genome in H3K9R mutants, including at much of the pericentromeric heterochromatin. However, in the remaining 4% of the genome, ~2% advanced and ~2% delayed replication timing (Materials and Methods; Figure 4.7A). Importantly, these changes are not likely to be caused by changes in the expression of genes encoding replication factors, as the H3K9R mutation does not significantly affect their expression (File S2) or that of other protein-coding genes (Penke et al. 2016).

Figure 4.5. Disruption of heterochromatin structure perturbs replication. A) S/G1 (log_2) averaged HWT (histone wild type) (yellow) and H3K9R (purple) replicates plotted across chromosome 3R at 100kb windows using a 10kb slide. See Figure 4.6 for other chromosomes. B) A 5 Mb region of pericentromeric heterochromatin of chromosome 3R with advanced windows in H3K9R compared to control highlighted in red (pvalue <0.01). C) Dean-Jett-Fox calculated cell cycle indices for HWT (yellow) and H3K9R (purple) acquired via FACS. In H3K9R mutants, we observed a slightly decreased and a slightly increased S and G2/M index, respectively. Error bars indicate standard deviation of three independent experiments (* = P<0.05, Student's T-test). D) All 10kb windows that advanced (red) or delayed (blue) replication in H3K9R mutants compared to HWT were assigned to the nine chromatin states previously defined in *Drosophila* (Kharchenko et al. 2011). Shown are the percentage of advanced or delayed windows that overlap each chromatin state. E) Average enrichment of modENCODE H3K9ac, H3K9me2, H3K9me3, and H3K27me3 signal from whole third instar larvae at 10kb windows of advanced (red), delayed (blue), or randomized set of windows.

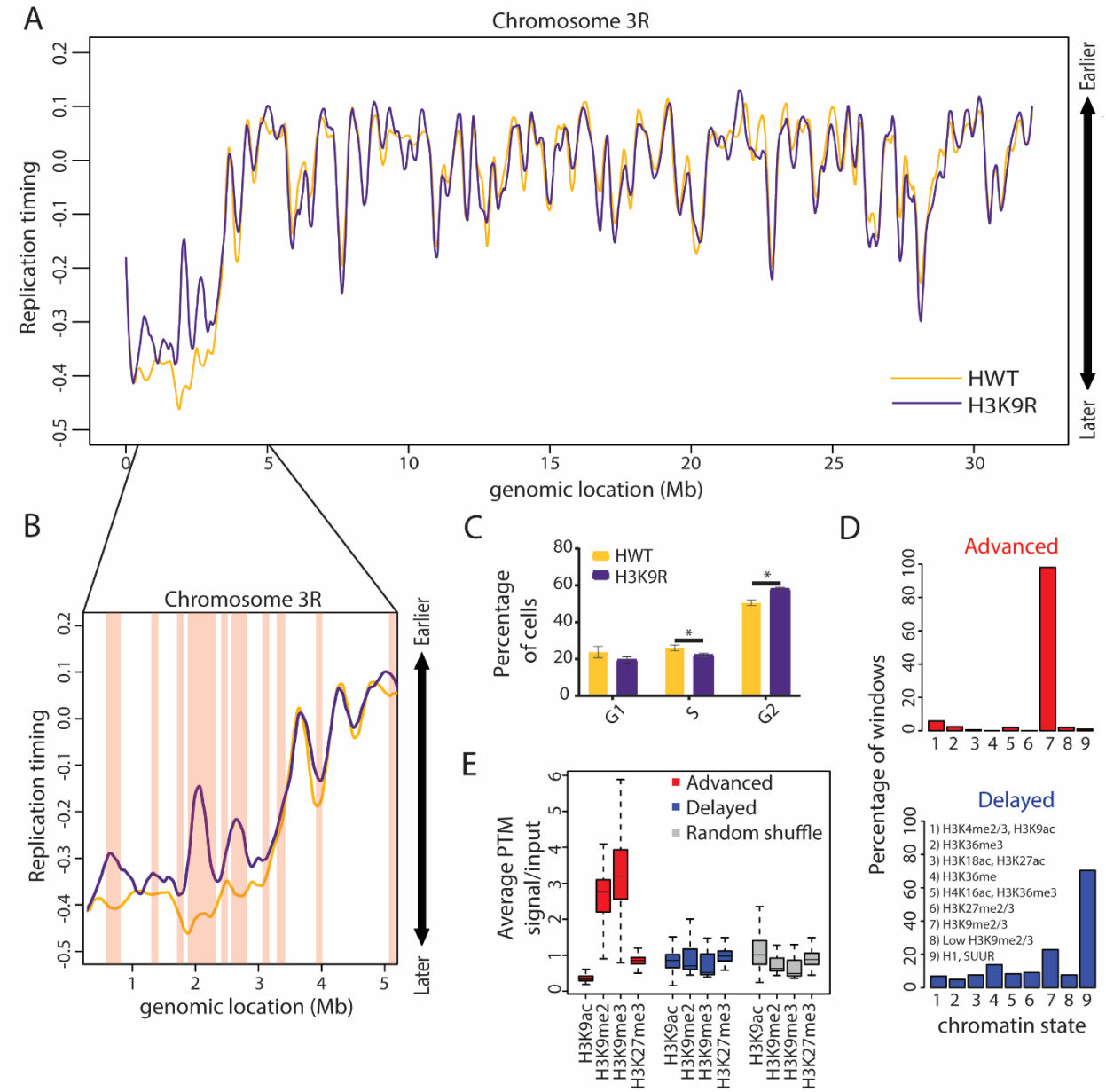

A
Chromosome 3R
Replication timing
genomic location (Mb)
Earlier
Later
HWT
H3K9R
B
Chromosome 3R
Replication timing
genomic location (Mb)
C
Percentage of cells
HWT
H3K9R
G1
S
G2
D
Advanced
Delayed
Percentage of windows
chromatin state
1) H3K4me2/3, H3K9ac
2) H3K36me3
3) H3K18ac, H3K27ac
4) H3K36me
5) H4K16ac, H3K36me3
6) H3K27me2/3
7) H3K9me2/3
8) Low H3K9me2/3
9) H1, SUUR
E
Average PTM signal/input
Advanced
Delayed
Random shuffle
H3K9ac
H3K9me2
H3K9me3
H3K27me3

Figure 4.6. Replication timing profile for H3K9R mutants and control. LOESS regression line applied to S/G1 (log_2) averaged replicates from HWT (yellow) and H3K9R (purple) plotted across all major chromosome scaffolds at 100kb windows with a 10kb slide.

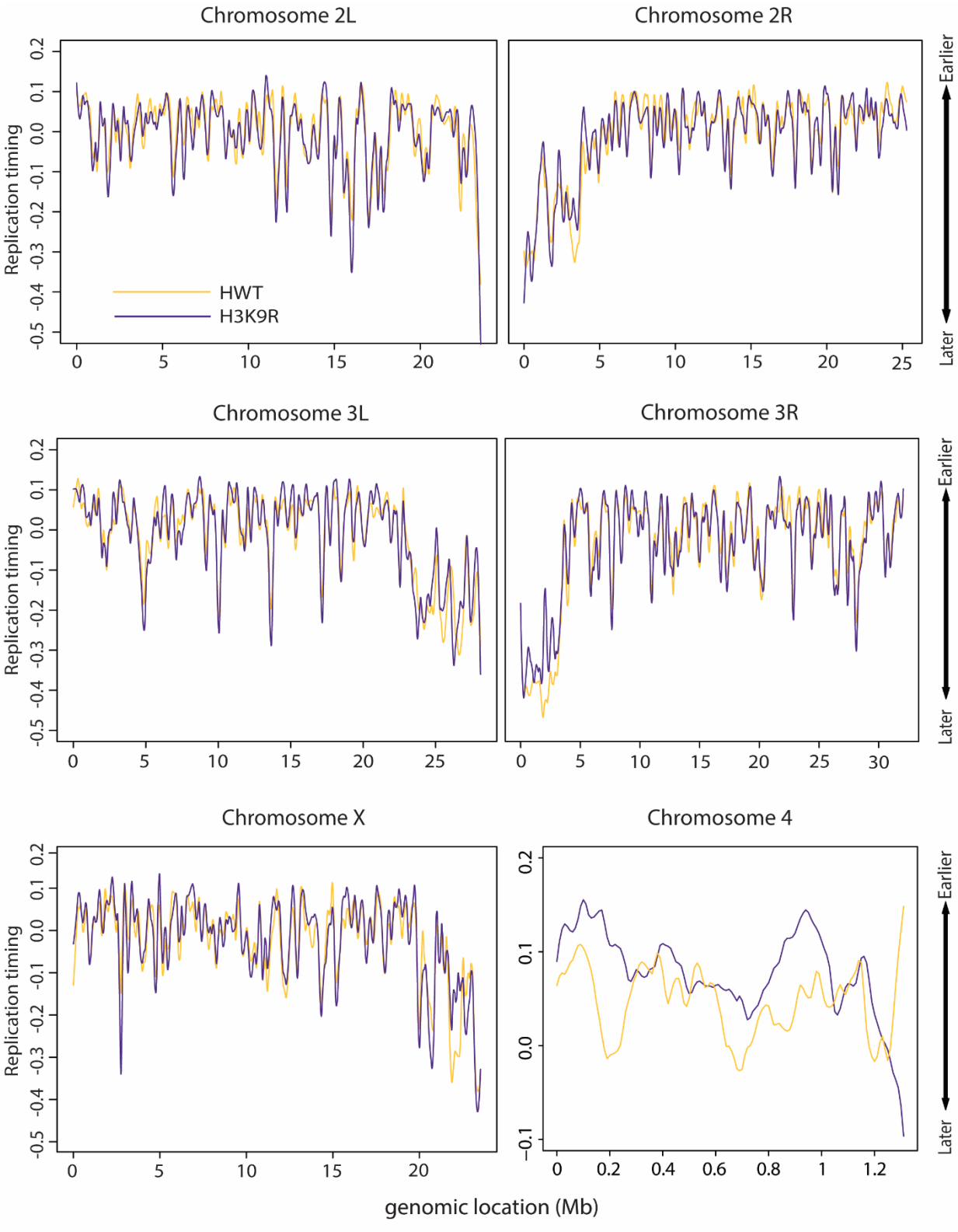
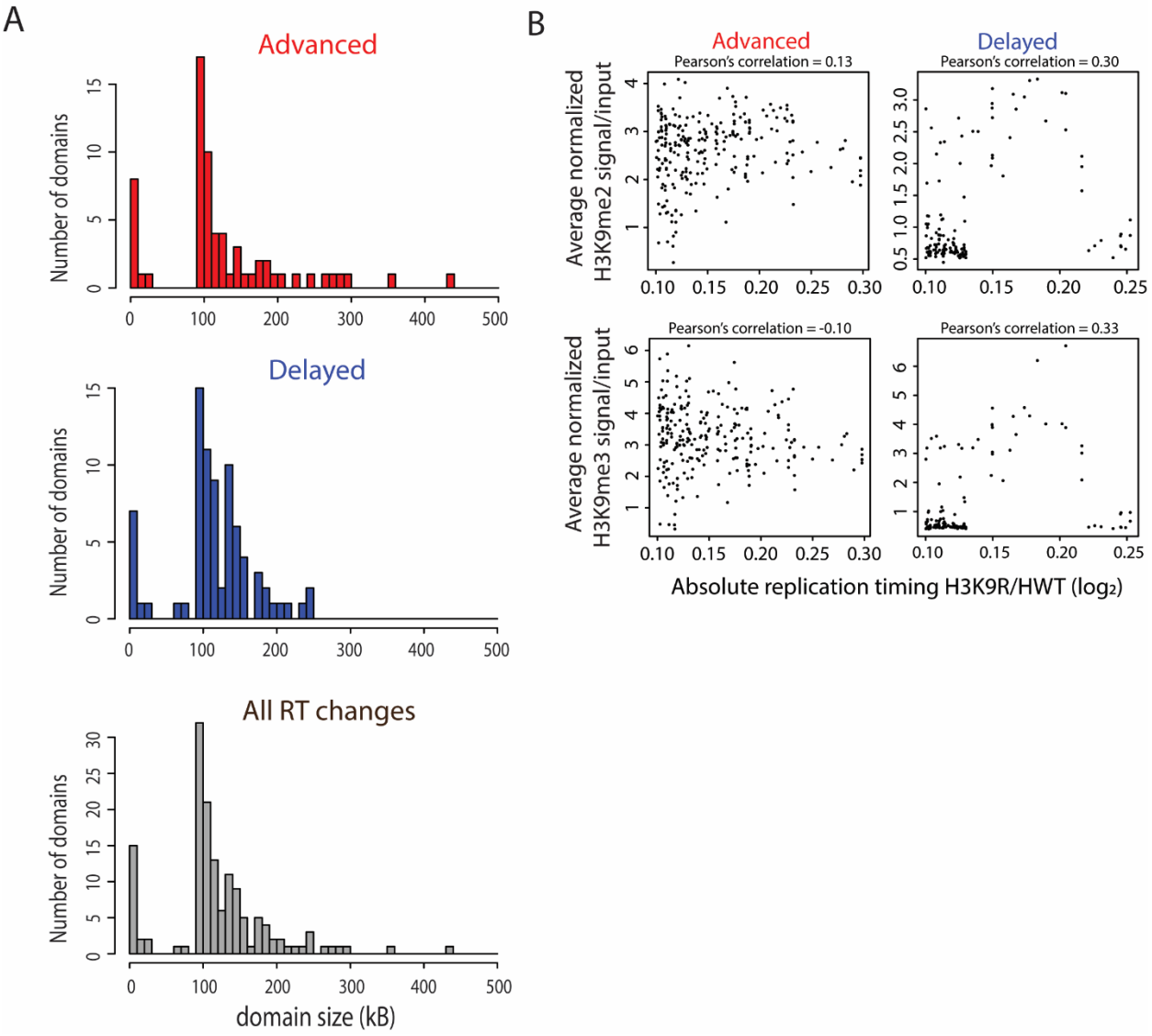

Chromosome 2L
Chromosome 2R
Chromosome 3L
Chromosome 3R
Chromosome X
Chromosome 4
Replication timing
HWT
H3K9R
Earlier
Later
genomic location (Mb)

Figure 4.7. Characterization of altered replication timing in H3K9R mutants. A) Histogram of the number of domain sizes with advanced (red), delayed (blue), or all replication timing change (grey). B) Correlation analysis of the absolute H3K9R/HWT log_2 RT fold change versus the average enrichment of H3K9me2 (top) or H3K9me3 (bottom) signal at 10kb windows with significantly advanced (left) or delayed (right) replication. ChIP-seq enrichment was determined from modENCODE datasets from wild-type whole 3rd instar larvae.


A
Advanced
Number of domains
Delayed
Number of domains
All RT changes
Number of domains
domain size (kB)
B
Advanced
Delayed
Pearson's correlation = 0.13
Pearson's correlation = 0.30
Average normalized H3K9me2 signal/input
Pearson's correlation = -0.10
Pearson's correlation = 0.33
Average normalized H3K9me3 signal/input
Absolute replication timing H3K9R/HWT (log2)

The majority (82.1%) of windows that replicate earlier in H3K9R mutants are located in pericentromeric heterochromatin (Figure 4.5B) or on the small 4th chromosome (Figure 4.6), which is primarily heterochromatic (Haynes et al. 2007). In contrast, the majority (76.2%) of windows that replicate later are located along chromosome arms (Figure 4.5A; Figure 4.6). Notably, windows of advanced replication in H3K9R mutants are enriched in a wild type genome for H3K9me2/me3 and not for other histone PTMs such as H3K27me3, a marker of facultative heterochromatin (Figure 4.5D,E; Figure 4.7B). This observation suggests that advanced replication is a direct effect of the H3K9R mutation. In contrast, delayed replication was not correlated with H3K9me2/me3, and instead occurred preferentially in chromatin environments relatively devoid of histone PTMs, referred to as “Black” chromatin (Figure 4.5D,E; Figure 4.7B) (Filion et al. 2010).

We hypothesized that if chromatin structure directly influences replication, then replication timing changes should occur at newly accessible chromatin in H3K9R mutants. We therefore assigned FAIRE-seq (Penke et al. 2016) and replication timing values to 10kb windows across the genome (Materials and Methods) to compare chromatin accessibility and replication timing in H3K9R mutants and control. Nearly all windows (230/243) that exhibit significantly advanced replication in H3K9R mutants also have increased FAIRE signal, suggesting that a more accessible chromatin environment is necessary for early replication (Figure 4.8A,B; Figure 4.9A). In contrast, most delayed windows exhibit no change in FAIRE signal (Figure 4.8A,B; Figure 4.9A). While most regions of the pericentromeres included in the current genome assembly are more accessible in H3K9R mutants compared to control (Penke et al. 2016), strikingly, the vast majority (92.9%) of windows with increased FAIRE signal do not display altered

replication timing (Figure 4.8A), suggesting that a more accessible chromatin environment is not sufficient to alter replication.

Similar results were obtained when considering HP1a chromatin binding (Figure 4.9B), which we previously showed by ChIP-seq is depleted from regions of the H3K9R genome that largely overlap regions of increased chromatin accessibility (Figure 4.9C) (Penke et al. 2016). HP1a is depleted at 217 of the 243 advanced windows in H3K9R mutants (Figure 4.8C,D; Figure 4.9B,D). However, the majority (94.7%) of windows that lose HP1a in H3K9R mutants do not have altered replication timing. We further observed that the extent of altered replication does not correlate with the magnitude of change in either FAIRE or HP1a signal (Figure 4.9E,F), suggesting that once a threshold of accessibility is reached there is no further effect on replication timing. We note that domains of altered replication timing in H3K9R mutants do not match those previously identified after HP1a knockdown in *Drosophila* cultured cells (Schwaiger et al. 2010), potentially due to H3K9-independent functions of HP1a or to the exclusion of repetitive DNA from the microarray based assay previously used (Figure 4.10).

Our comparison of Repli-seq data with FAIRE-seq and HP1a ChIP-seq data revealed two important features of the relationship between chromatin and replication. First, altered chromatin accessibility and HP1a loss occurs independently of, and therefore likely precedes, changes in replication initiation. Second, some factor must actively function within accessible chromatin to advance replication, as only a subset of domains that increase accessibility in H3K9R mutants change replication timing.

Figure 4.8. Open chromatin is permissive to advancement but not delay of replication timing. A) Heatscatter plot of the H3K9R/HWT ratio of normalized replication timing values (S/G1 (log_2)) plotted versus the H3K9R/HWT ratio of normalized FAIRE signal at all 10kb windows across the major chromosome scaffolds. Advanced (red) and delayed (blue) 10kb windows are indicated ($p<0.05$ and log_2 fold change>0.1; limma). Darker color indicates higher density of windows. B) Cumulative count of significantly advanced (red) or delayed (blue) 10kb windows ordered by increasing FAIRE signal in H3K9R compared to HWT. C) Heatscatter plot of the H3K9R/HWT ratio of normalized HP1a ChIP signal versus the H3K9R/HWT ratio of normalized FAIRE signal at all 10kb windows across the major chromosome scaffolds. Significance in C determined as in A. D) Venn-diagram of all 10kb windows with significantly altered FAIRE or HP1a signal in H3K9R compared to HWT ($p<0.01$; edgeR) and significantly altered RT determined as in A and C.

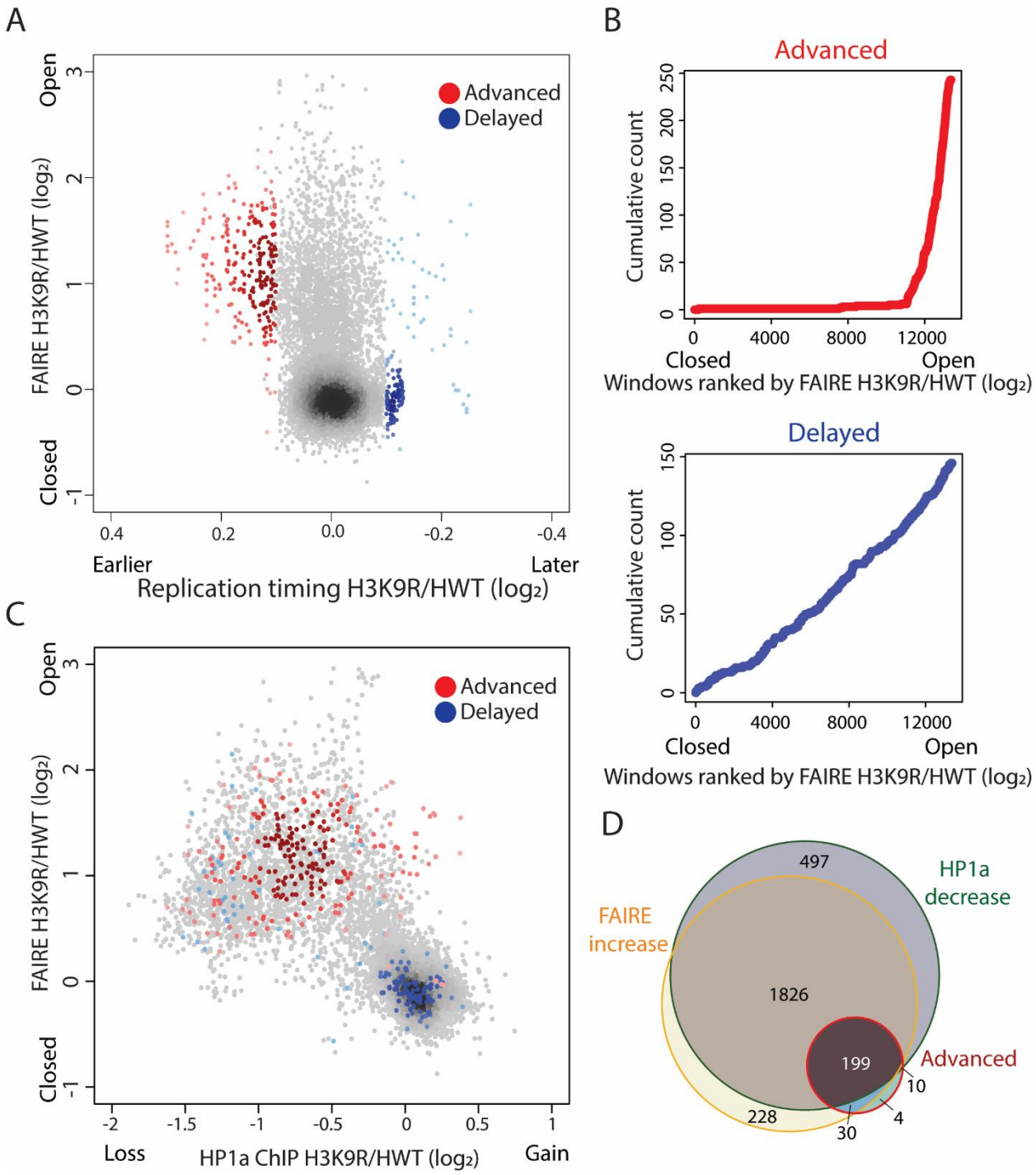

A
Open
Closed
FAIRE H3K9R/HWT (log2)
Advanced
Delayed
Earlier
Later
Replication timing H3K9R/HWT (log2)
B
Advanced
Cumulative count
Closed
Open
Windows ranked by FAIRE H3K9R/HWT (log2)
Delayed
Cumulative count
Closed
Open
Windows ranked by FAIRE H3K9R/HWT (log2)
C
Open
Closed
FAIRE H3K9R/HWT (log2)
Advanced
Delayed
Loss
HP1a ChIP H3K9R/HWT (log2)
Gain
D
497
HP1a decrease
FAIRE increase
1826
199
Advanced
10
228
30
4

Figure 4.9. Disrupting heterochromatin is not sufficient for inducing altered replication. A) Venn-diagram of 10kb windows with significantly altered FAIRE signal and significantly advanced or delayed replication in H3K9R mutants compared to control. B) Heatscatter plot of the H3K9R/HWT ratio of normalized replication timing values (S/G1 (log_2)) plotted versus the H3K9R/HWT ratio of normalized HP1a ChIP signal at all 10kb windows across the major chromosome scaffolds. C) Venn-diagram of 10kb windows with significantly increased FAIRE signal and decreased HP1a ChIP signal. D) Venn-diagram of 10kb windows with significantly altered HP1a ChIP signal and significantly advanced or delayed RT. E-F) Absolute change in FAIRE signal (E) or HP1a ChIP signal (F) between H3K9R mutants and controls plotted versus the absolute change in replication timing between the two genotypes. The magnitude of altered chromatin accessibility or HP1a localization is not correlated with the magnitude of replication timing change.

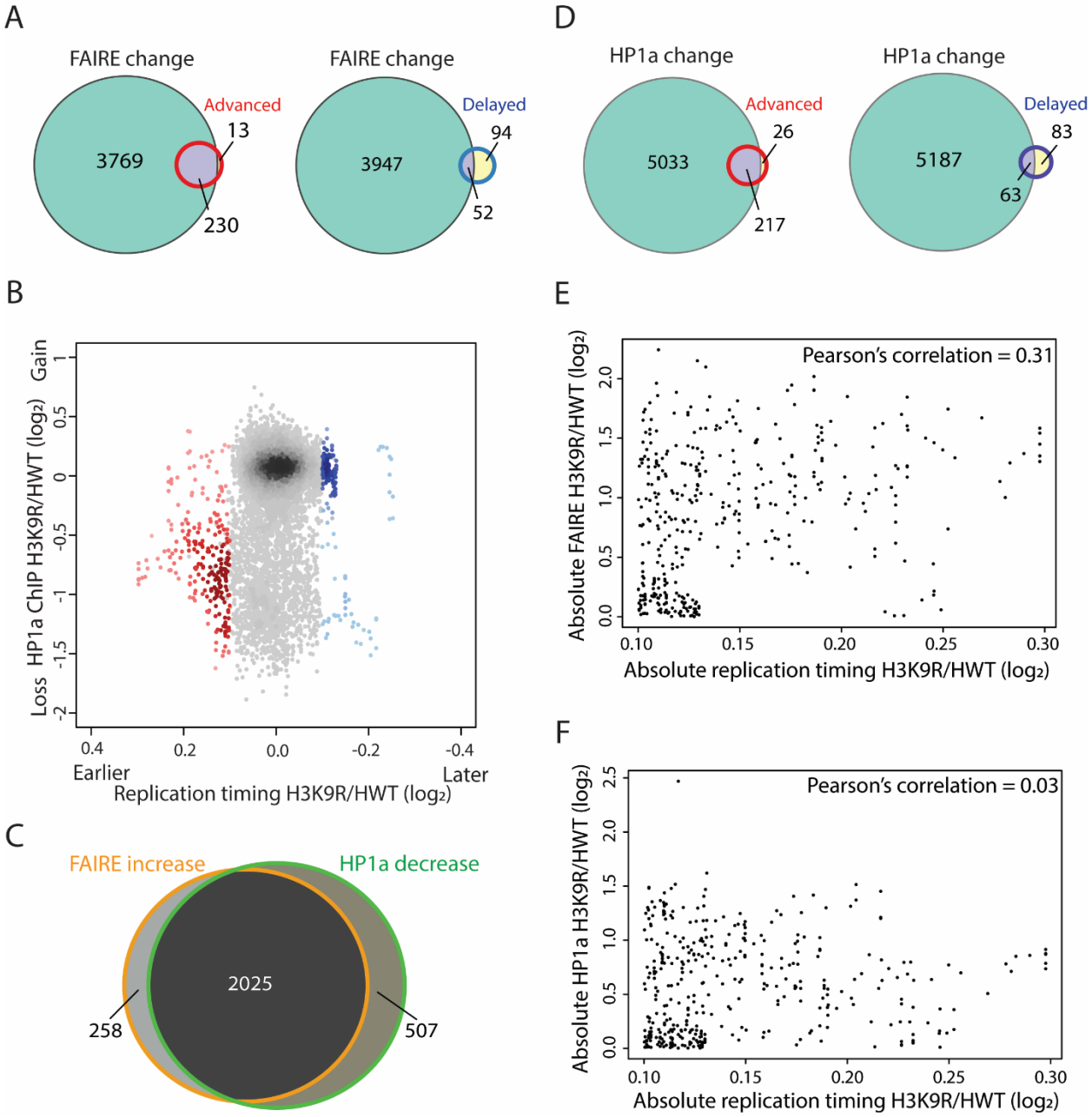
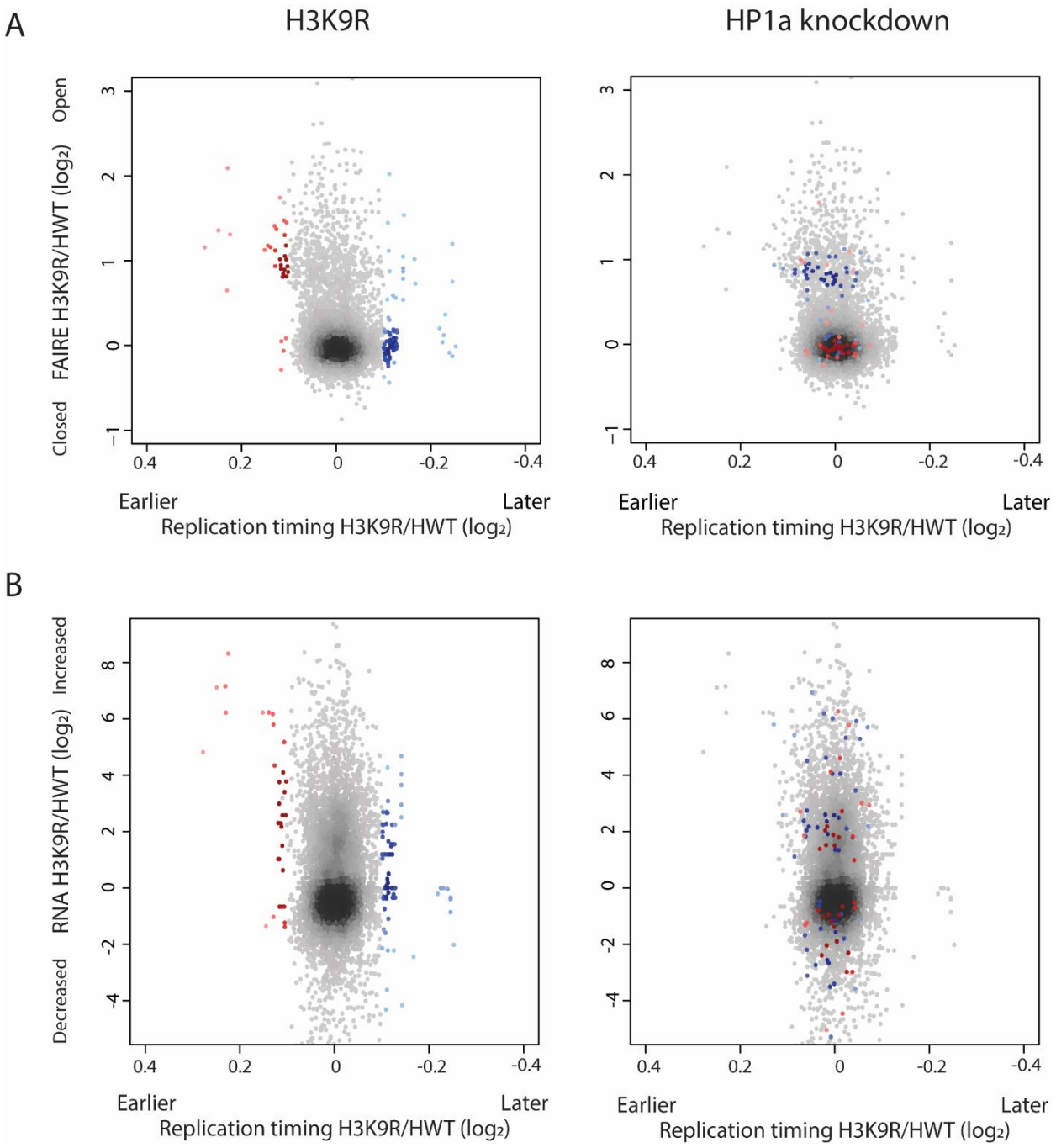

A
FAIRE change
Advanced
3769
13
230
FAIRE change
Delayed
3947
94
52
D
HP1a change
Advanced
5033
26
217
HP1a change
Delayed
5187
83
63
B
Gain
Loss
HP1a ChIP H3K9R/HWT (log2)
1
0.5
0
-0.5
-1
-1.5
-2
0.4
0.2
0.0
-0.2
-0.4
Earlier
Later
Replication timing H3K9R/HWT (log2)
E
Pearson's correlation = 0.31
Absolute FAIRE H3K9R/HWT (log2)
2.0
1.5
1.0
0.5
0.0
0.10
0.15
0.20
0.25
0.30
Absolute replication timing H3K9R/HWT (log2)
C
FAIRE increase
HP1a decrease
2025
258
507
F
Pearson's correlation = 0.03
Absolute HP1a H3K9R/HWT (log2)
2.5
2.0
1.5
1.0
0.5
0.0
0.10
0.15
0.20
0.25
0.30
Absolute replication timing H3K9R/HWT (log2)

Figure 4.10. Domains of altered replication in H3K9R mutants do not overlap those identified after HP1a knockdown. A-B) Heatscatter plot of the H3K9R/HWT ratio of normalized replication timing values (S/G1 (log_2)) plotted versus the H3K9R/HWT ratio of normalized FAIRE (A) or RNA-seq (B) signal at Hidden Markov Model determined replication domains identified by Schwaiger et al. (2010). Significantly advanced windows are indicated in red and significantly delayed windows are indicated in blue. Domains identified as significantly altered using Repli-seq in H3K9R mutants are shown in the left panels and domains identified by Schwaiger et al. using BrdU ChIP coupled with microarrays are shown in the right panels. Differences between the two datasets could be due to H3K9-independent roles for HP1 or tissue specific differences (3^{rd} instar imaginal wing disc vs. embryo derived Kc cells). We speculate that regions of advanced replication timing identified in our study that were not identified in Schwaiger et al. were due to the removal of repetitive DNA sequences in the microarray designed used previously.


A
H3K9R
HP1a knockdown
Open
Closed
FAIRE H3K9R/HWT (log2)
3
2
1
0
−1
0.4
0.2
0
-0.2
-0.4
Earlier
Later
Replication timing H3K9R/HWT (log2)
B
Increased
Decreased
RNA H3K9R/HWT (log2)
8
6
4
2
0
-2
-4

One such factor could be the transcriptional activity within a chromatin domain. We therefore compared replication timing with transcriptome profiles of wing discs from H3K9R and control animals (Penke et al. 2016), including in our analysis previously annotated genes and transposons (Materials and Methods). We focused on transcripts most likely to drive replication timing changes by identifying a high confidence transcript within each 10kb window that differed between H3K9R and control samples (i.e. the transcript with the lowest p-value). We then compared the fold-change of this transcript to the replication timing value of the same 10kb window. Similar to our chromatin accessibility analysis, we found that only a small fraction (6.8%) of the 3,371 10kb windows containing a transcript with a significant expression change also exhibited a replication timing change (Figure 4.11A,B; Figure 4.12A). This observation indicates that altered transcription is not sufficient to induce replication changes. We further observed that the majority (76.5 %) of windows with advanced replication in H3K9R mutants contain a change in gene expression (Figure 4.11A,B,D). Moreover, because most (97.3%) of these changes were increases in expression, we speculate that transcription might promote early replication initiation.

Figure 4.11. Altered transposon expression occurs at domains of advanced replication timing in H3K9R mutants. A) Heatscatter plot of the H3K9R/HWT ratio of normalized replication timing values (S/G1 (log_2)) plotted versus the H3K9R/HWT ratio of normalized RNA-seq signal at all 10kb windows across the major chromosome scaffolds. The transcript with the lowest p-value across the 10kb window was plotted, and significantly advanced (red) and delayed (blue) 10kb windows are indicated ($P<0.05$ and log_2 fold change>0.1; limma). B) Histogram of the number of differentially expressed transcripts in 10kb windows of advanced replication (red) (left). Venn-diagram comparing the number of windows with differentially expressed transcripts and number of windows with advanced replication (right). C) Histogram of the number of transposons belonging to a differentially expressed transposon family in 10kb windows of advanced replication (red) (left). Venn-diagram comparing the number of windows with a transposon belonging to a differentially expressed transposon family to the number of windows with advanced replication (right). D) Browser shot of 10kb window (chr3R-2130000-2140000) with advanced replication. HWT (yellow) and H3K9R (purple) normalized FAIRE-seq, HP1a ChIP-seq, and RNA-seq data plotted in the context of mappability, genes, and transposons.

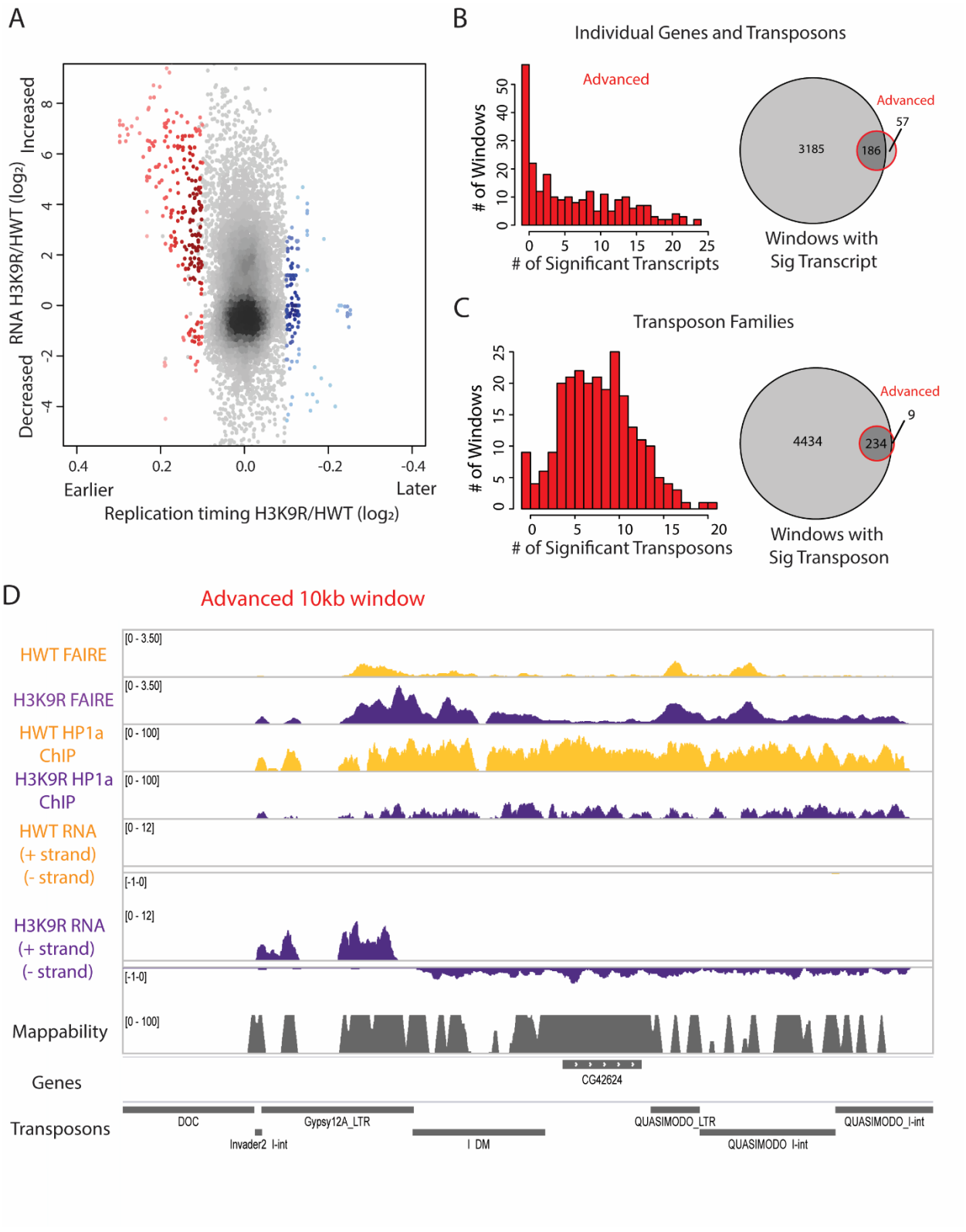
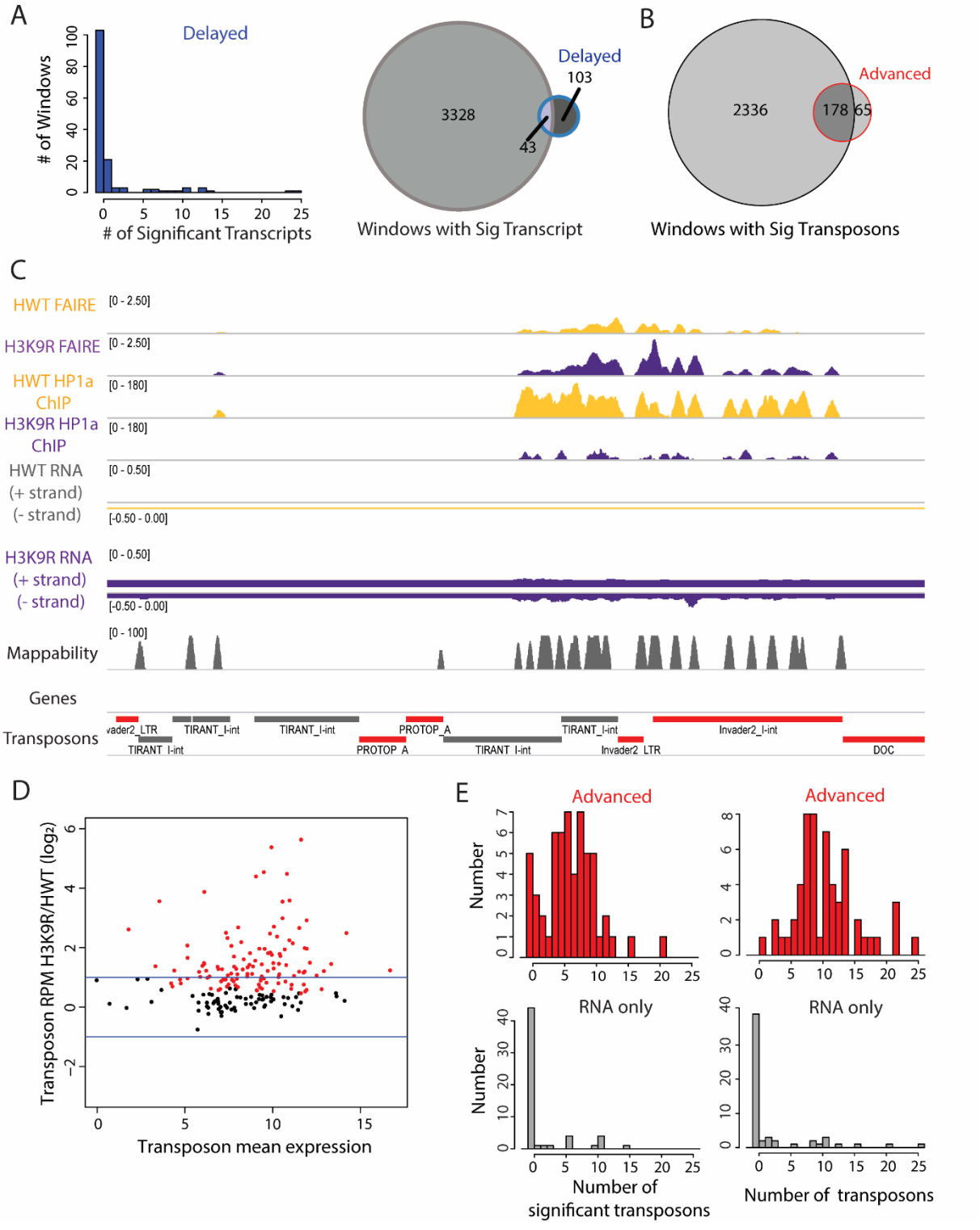

A
Increased
Decreased
RNA H3K9R/HWT (log2)
8
6
4
2
0
-2
-4
0.4
0.2
0.0
-0.2
-0.4
Earlier
Later
Replication timing H3K9R/HWT (log2)
B
Individual Genes and Transposons
Advanced
# of Windows
0
10
20
30
40
50
0
5
10
15
20
25
# of Significant Transcripts
Advanced
57
3185
186
Windows with
Sig Transcript
C
Transposon Families
# of Windows
0
5
10
15
20
25
0
5
10
15
20
# of Significant Transposons
Advanced
9
4434
234
Windows with
Sig Transposon
D
Advanced 10kb window
HWT FAIRE
[0 - 3.50]
H3K9R FAIRE
[0 - 3.50]
HWT HP1a
ChIP
[0 - 100]
H3K9R HP1a
ChIP
[0 - 100]
HWT RNA
(+ strand)
(- strand)
[0 - 12]
[-1-0]
H3K9R RNA
(+ strand)
(- strand)
[0 - 12]
[-1-0]
Mappability
[0 - 100]
Genes
CG42624
Transposons
DOC
Gypsy12A_LTR
QUASIMODO_LTR
QUASIMODO_I-int
Invader2 I-int
I DM
QUASIMODO I-int

Figure 4.12. Regions of advanced replication in H3K9R mutants exhibit altered transposon expression. A) Histogram of the number of differentially expressed transcripts in 10kb windows of delayed replication (blue) (left). Venn-diagram comparing the number of windows with differentially expressed transcripts and number of windows with delayed replication (right). B) Venn-diagram comparing the number of windows with a differentially expressed transposon to the number of windows with advanced replication (see also Figure 4.11C). Because high transposon density (Figure 4.13A,B) and low sequence mappability of these regions likely masked our ability to detect transcriptional changes, we examined expression levels of transposon families rather than individual transposons (Materials and Methods). Counts from individual transposons were summed based on RepeatMasker categorization of transposon families. C) Genome browser shot of a 10kb window with significantly advanced replication in H3K9R mutants but no detectable accompanying change in RNA expression via edgeR analysis. FAIRE-seq, HP1a ChIP-seq, or RNA-seq signal are shown for H3K9R (purple) and HWT (yellow) samples. Note the low mappability of this region due to high transposon density. Red transposons indicate individual transposons belonging to a family that is differentially expressed in H3K9R mutants. Browser shot provides a representative example of transcriptional changes that are likely occurring but cannot be directly examined due to low mappability. D) MA plot showing differential expression of transposon families between HWT control and H3K9R samples. Each dot represents a transposon family with red indicating statistical significance as determined by edgeR ($p<0.01$; see Materials and Methods). Blue lines indicate two-fold change. E) Histograms in the top left panel show the number of transposons belonging to a family that is differentially expressed in H3K9R

mutants compared to control at 57 10kb windows that exhibited advanced replication in H3K9R mutants but no initially detected transcriptional change (see also Venn diagram in Figure 4B). Bottom left panel shows number of transposons bellowing to a differentially expressed family at all 10kb windows that exhibit a transcriptional change but no replication timing change (RNA only). Histograms in right panel show the number of transposons at the 57 advanced windows (top) and RNA only windows (bottom). 52 of the 57 windows we did not initially score as having changed expression (Figure 4.11B) contained at least one transposon belonging to a family with significantly altered expression in H3K9R mutants. The remaining 5 windows were surrounded by 10kb windows containing several transposons with significantly altered expression. These data suggest that altered transcription is necessary for advanced replication in H3K9R mutants.


A
Delayed
# of Windows
# of Significant Transcripts
Delayed
103
3328
43
Windows with Sig Transcript
B
Advanced
2336
178
65
Windows with Sig Transposons
C
HWT FAIRE
H3K9R FAIRE
HWT HP1a ChIP
H3K9R HP1a ChIP
HWT RNA (+ strand) (- strand)
H3K9R RNA (+ strand) (- strand)
Mappability
Genes
Transposons
TIRANT_I-int
PROTOP_A
Invader2_LTR
Invader2_I-int
DOC
D
Transposon RPM H3K9R/HWT (log2)
Transposon mean expression
E
Advanced
RNA only
Number
Number of significant transposons
Number of transposons

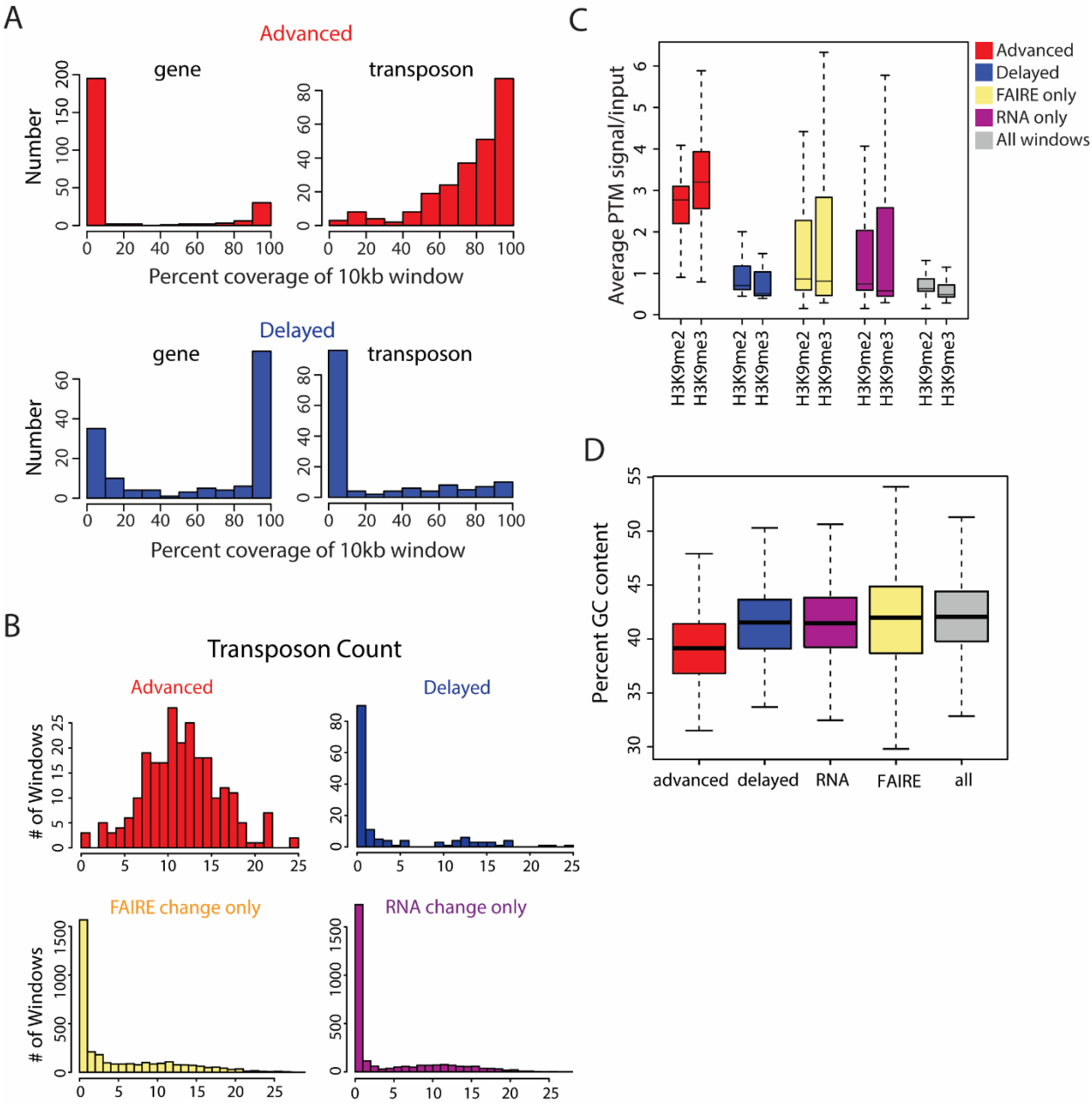
Windows with advanced replication have a high transposon density, unlike delayed windows which are gene-rich (Figure 4.13A). The low sequence mappability of most transposons likely inhibited our ability to detect all transcriptional changes within advanced replication domains (Figure 4.12C). Therefore, in addition to analyzing individual transposons, we also identified transposons belonging to families that were differentially expressed between H3K9R and control (Materials and Methods; Figure 4.11C; Figure 4.12B-E). All 243 windows of advanced replication in H3K9R mutants contain either a transposon belonging to a family that was differentially expressed in H3K9R mutants compared to control (96.4%) or neighbored a window containing multiple differentially expressed transcripts (Figure 4.11C). Although we cannot determine whether individual transposons within all 243 advanced windows are undergoing transcriptional changes, these data suggest that altered transcription is required for advancement of replication.

We have shown that an increase in chromatin accessibility and a change in gene expression accompany advanced replication in H3K9R mutants (Figure 4.11D), but our data indicate that these factors are not sufficient for promoting earlier replication initiation. What then distinguishes domains of advanced replication within accessible chromatin from those that don't change replication timing? Along with transposon enrichment (Figure 4.13B), advanced replication domains are enriched for H3K9me2/me3 (Figure 4.13C) and exhibited a lower GC content (Figure 4.13D) compared to domains of increased chromatin accessibility or increased RNA expression with unaltered replication (FAIRE only and RNA only, respectively). Although transposon density distinguished advanced domains, the majority of domains with altered transposon expression do not

alter replication (Figure 4.11C; Figure 4.12B). Therefore, we surmise that altered transposon expression is necessary, but additional events must occur within accessible chromatin to advance replication.

These events could include origin specification or activation. Origins of replication are licensed during G1 phase by origin specification factors, and a subset of licensed origins is then subsequently activated during S phase. Certain models for a temporal program of replication initiation posit a stochastic process in which a higher density of licensed origins in accessible, euchromatin increases the probability of replication initiation compared to inaccessible heterochromatin (Sasaki et al. 2017; Rhind et al. 2010; Löb et al. 2016; Hawkins et al. 2013). Additionally, *trans*-acting factors act upon the licensed origin landscape to either promote or inhibit origin activation (Yamazaki et al. 2012). Regulation of either molecular event could demarcate domains that advance replication within the permissive open chromatin environment created by the H3K9R mutation. In addition to advanced domains, we also showed that delayed replication domains in H3K9R mutants are largely independent of altered chromatin accessibility or transcriptional changes. We therefore hypothesize that elevated accessibility of pericentromeric heterochromatin functions as a “sink” for limiting replication factors, resulting in delayed replication of domains along chromatin arms as proposed for other replication factors (Yoshida et al. 2014; Foti et al. 2016).

Figure 4.13. Transposon density and H3K9me2/me3 status are distinguishing features of regions with advanced replication. A) Histogram of the number of significantly advanced (red) or delayed (blue) 10kb windows within 10 bins representing two categories: the percentage of each window covered by genes (left panels) or transposons (right panels). B) Histogram of the number of transposons in 10kb windows of advanced replication (red), delayed replication (blue), FAIRE change without replication change (FAIRE only), RNA change without replication change (RNA only), and all 10kb windows. C) Average enrichment of modENCODE H3K9me2 and H3K9me3 signal from wild-type whole third instar larvae at 10kb windows within the categories described in panel B. D) Boxplot of the percent GC content of 10kb windows within the categories described in panel B.

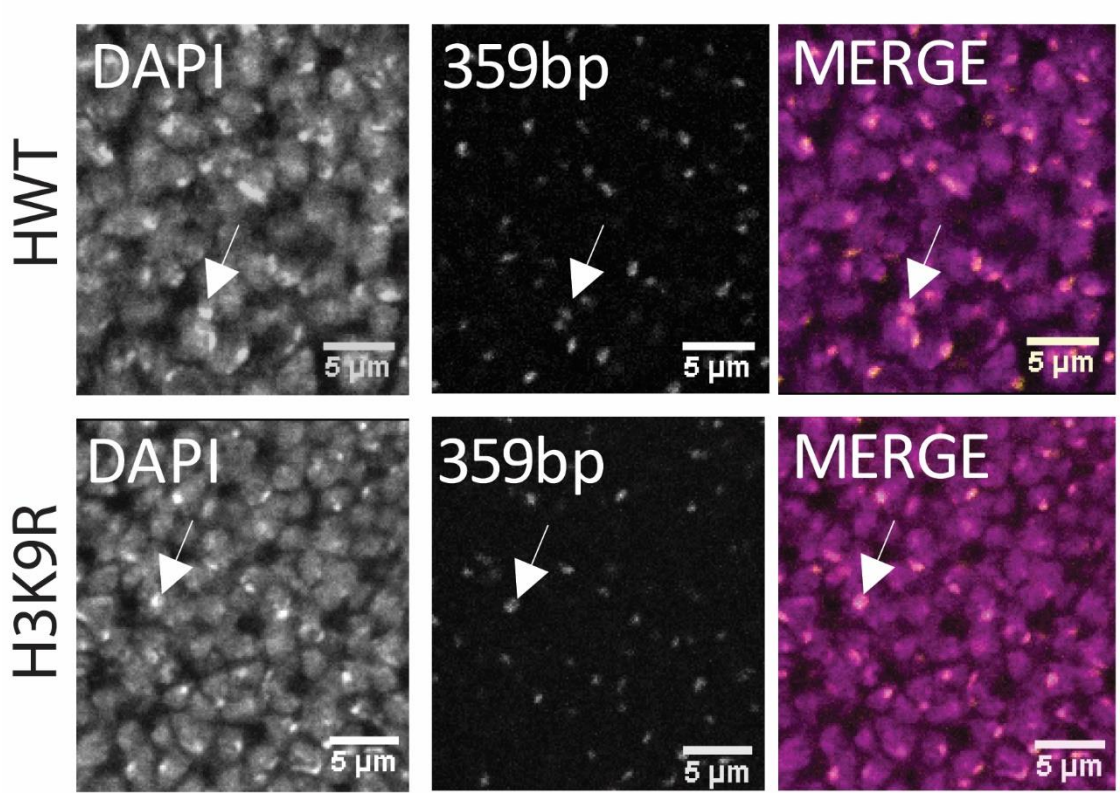

A
Advanced
gene
transposon
Number
Percent coverage of 10kb window
Delayed
gene
transposon
Number
Percent coverage of 10kb window
B
Transposon Count
Advanced
Delayed
# of Windows
FAIRE change only
RNA change only
# of Windows
C
Average PTM signal/input
Advanced
Delayed
FAIRE only
RNA only
All windows
H3K9me2
H3K9me3
D
Percent GC content
advanced
delayed
RNA
FAIRE
all

In H3K9R mutants, pericentromeric heterochromatin replicated later than the largely euchromatic chromosome arms, with both earlier and later replication domains occurring within each region. These data support a model in which the two regions are organized in different nuclear compartments wherein *trans*-acting factors act to locally determine replication timing domains. Previous work in metazoans has identified a largely euchromatic compartment “A” and a largely heterochromatic compartment “B” that may correspond to the arms and pericentromeric regions of *Drosophila* chromosomes, respectively (Lieberman-Aiden et al. 2009). Interestingly, H3K9R domains that advanced replication within pericentromeric heterochromatin replicated with a timing similar to later replicating domains along chromosome arms (Figure 4.5A,B; Figure 4.6), suggesting that overall compartmentalization of euchromatin and heterochromatin is not disrupted by H3K9R mutation. In support of this idea, H3K9R diploid nuclei have a DAPI-bright chromocenter that colocalizes with fluorescence in-situ hybridization (FISH) probes for the pericentromeric 359-bp repeat, similar to wild type nuclei (Figure 4.14). This observation indicates that the H3K9R mutant retains some features of heterochromatin. We propose that H3K9R mutation alters replication timing by disrupting local chromatin accessibility without affecting overall compartmentalization of heterochromatin (Strom et al. 2017).

Using a histone mutation to specifically manipulate chromatin structure, we demonstrate that establishment of early and late replication domains is a downstream consequence of the open chromatin landscape across the genome, and that transcription is a strong predictor of early replication. Accordingly, maintenance of inaccessible chromatin and repression of transcription are critical for late replication of pericentromeric

heterochromatin. Importantly, neither accessible chromatin nor transcription is sufficient to trigger earlier replication, indicating modes of control for replication initiation that are independent of these two features of animal genome structure and activity.

Figure 4.14. DAPI-bright chromocenter in H3K9R mutants colocalizes with 359bp FISH probe (X chromosome pericentromeric heterochromatin). HWT and H3K9R female eye imaginal disc stained with 359-bp FISH probe (yellow) and counterstained with DAPI (magenta). White arrow designates representative DAPI bright chromocenter and 359-bp probe colocalization in H3K9R and control. Shown is a single representative slice of a Z projection. If a nucleus does not contain a chromocenter in this representative image, the chromocenter was captured in a Z plane not shown.


HWT
H3K9R
DAPI
359bp
MERGE
5 µm

CHAPTER 5- ROLE OF H3K9 IN POLYPLOID REPLICATION

Introduction

The *Drosophila* larval salivary gland consists of polyploid cells that can achieve a ploidy of greater than 1000C (Fox and Duronio 2013; Lilly and Duronio 2005). Polyploidy is established through a modified cell cycle called the endocycle, in which cells alternate between G1 and S phases without cell division. Interestingly, the pericentromeric regions along with several intercalary heterochromatic regions on the chromosome arms often do not complete replication before the following phase of replication (Yarosh and Spradling 2014). This slow replication results in their under-replication relative to the rest of the genome. In contrast to polyploid salivary gland cells, replication in diploid cells occurs uniformly across the genome. Through better understanding of under-replication, we may elucidate functions of regulatory mechanisms normally present in diploid cells to ensure this uniformity. For example, stalled replication forks occur in regions of under-replication but do not lead to cell death due to repression of p53-mediated apoptosis in polyploid cells (Andreyeva et al. 2008; Mehrotra et al. 2008). We can therefore examine relationships between replication progression and DNA damage in the absence of this pathway. Furthermore, under-replicated regions share characteristics of mammalian fragile states and may help us understand general features of genome instability (Hua and Orr-Weaver 2017).

Mechanisms that control under-replication have come to light through the discovery and characterization of the Suppressor of Under-Replication (SuUR) protein

(Belyaeva et al. 1998; Zhimulev et al. 2003a). The SuUR proteins binds regions of late replication (Makunin et al. 2002), and SuUR mutants result in complete loss of under-replication in intercalary heterochromatin and partial loss in pericentromeric heterochromatin (Zhimulev et al. 2003a; Belyaeva et al. 1998). Conversely, overexpression of SuUR enhances under-replication (Zhimulev et al. 2003b). More recently, direct evidence of SuUR's role in establishing late replication was demonstrated. SuUR was discovered to modulate replication independently of origin specification by slowing replication fork progression (Sher et al. 2012; Nordman et al. 2014). The function of slowed replication is unclear as SuUR mutants are viable, but consequences of under-replication include increased susceptibility to somatic DNA alterations (Yarosh and Spradling 2014).

An important outstanding question is how SuUR is localized to chromatin, indicating which genomic regions are selected for under-replication. Several findings have pointed to a role for chromatin in this selection. For instance, 60% of intercalary heterochromatin regions that are under-replicated overlap with Polycomb group binding sites (Zhimulev et al. 2003b). Furthermore, SuUR interacts with HP1, and the chromatin localization of the two components are interdependent (Pindyurin et al. 2008). Mutation or overexpression of either component alters chromatin localization patterns of the other. The connection between HP1 and SuUR is further supported by the fact that tethering either HP1 or SuUR is sufficient to induce under-replication (Pokholkova et al. 2015). Finally, SuUR influences methylation status of H3K9 and H3K27 (Koryakov et al. 2011). Although the interdependence of HP1 and SuUR is well-established, the role of histone modifications in SuUR localization has remained unclear. To understand the

role of H3K9 is under-replication we investigated endocycle progression and relative genome copy number in H3K9R mutant salivary glands.

Materials and Methods

Polytene chromosome preparation

Salivary gland polytene chromosome spreads were prepared as in Cai et al. (2010). *HWT* and *K9R* salivary glands were squashed on the same slide and genotypes were differentiated using anti-HP1a (1:1250). Anti-PCNA antibody was used at a 1:500 dilution. Anti-HP1a was obtained from the Developmental Hybridoma Bank (C1A9), and anti-PCNA was obtained from Abcam (ab29). PCNA patterns were categorized blindly based on Kolesnikova et al. (2013).

Salivary gland sequencing

25 salivary glands were dissected in PBS-T (PBS with 0.1% Triton-X100) from 3rd instar wandering larvae and frozen at -80C. Salivary glands were subsequently resuspended in 500 µl of PBS and added to a 7ml Dounce homogenizer pretreated with 5% BSA in PBS-T. Glands were homogenized with a loose pestle for 10 strokes and the tight pestle for 10 strokes. Homogenate was transferred to an Eppendorf tube and pelleted at 1500g for 5 minutes at 4C. Pelleted nuclei were resuspended in 1.5 mL of FAIRE lysis buffer (McKay and Lieb 2013) and sonicated to 500-1000bp using a Branson probe sonifier. Fragmented DNA was treated sequentially with RNase A and Proteinase K before phenol chloroform extraction and ethanol precipitation. Next-generation sequencing library preparation was performed usng the ThruPlex DNA-seq kit from Rubicon Genomics.

Results

H3K9 is essential for normal endocycle progression

To determine if H3K9 played a role in cell cycle progression of salivary gland nuclei, we stained polytene chromosome spreads from salivary glands with DAPI to visualize DNA and with Proliferating Cell Nuclear Antigen (PCNA), a marker of cells undergoing replication. PCNA is a DNA clamp that functions in polymerase processivity and associates with active replication forks (Moldovan et al. 2007). We first examined overall cell cycle progression by counting the percentage of cells in G1 (PCNA negative) and S phase (PCNA positive) in *HWT* and *K9R* polytene chromosomes. Whereas ~65% of cells in *HWT* were PCNA positive, only ~35% of *K9R* cells were positive, indicating cell cycle progression is misreguated in *K9R* mutants (Figure 5.1).

The decreased percentage of nuclei in S phase could be a consequence of altered progression through G1 and/or S phase. We next examined S phase progression in more detail by comparing PCNA localization patterns in *HWT* and *K9R* polytene chromosomes. Previously, PCNA staining patterns have been used to subdivide and characterize S phase progression (Kolesnikova et al. 2013). We used these guidelines to bin all PCNA positive chromosomes into a particular pattern and determined the percentage of chromosomes in each pattern (Figure 5.2). Interestingly, *K9R* mutants displayed an increased percentage of chromosomes in the earliest and latest S phase patterns, suggesting the timing of S phase is disrupted.

Figure 5.1: H3K9 is essential for endocycle progression. Polytene chromosomes were stained with DAPI and anti-PCNA. PCNA positive staining marked chromosomes undergoing replication and are in S phase (blue). PCNA negative staining indicates chromosomes in G1 phase (orange). Experiment performed in collaboration with Robin Armstrong.

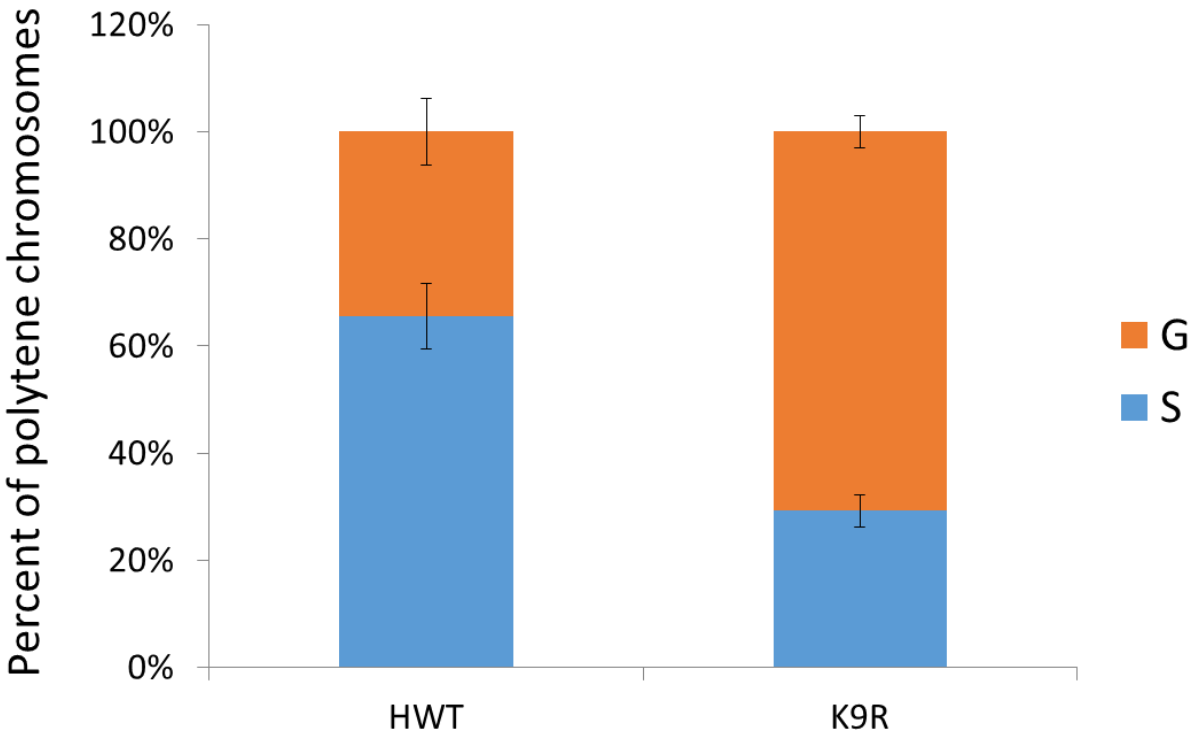
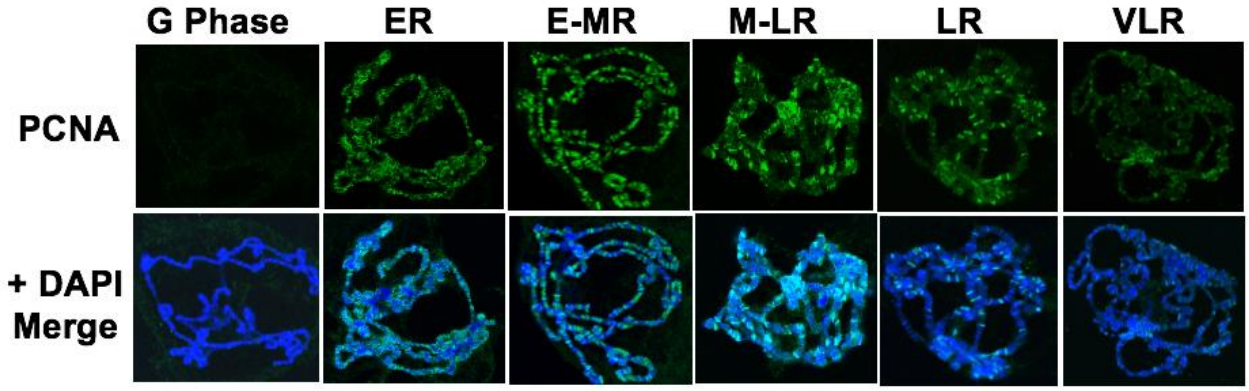
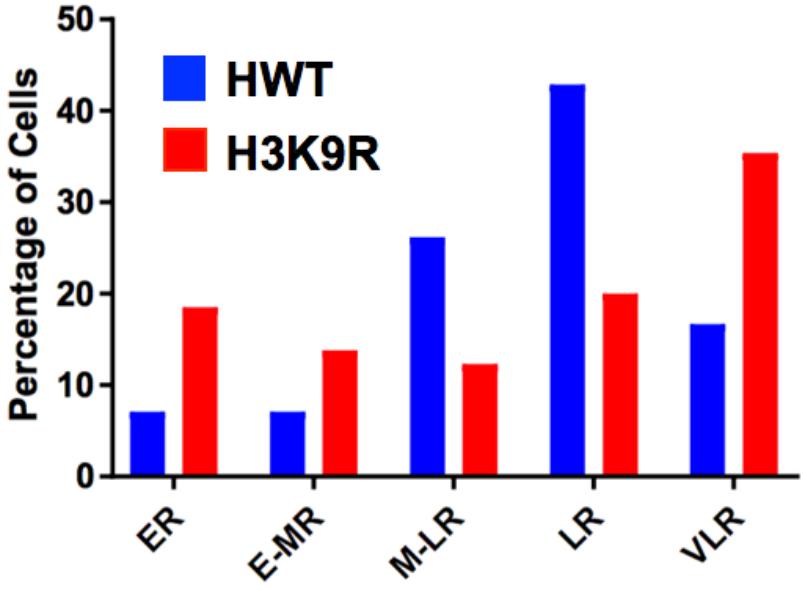

Percent of polytene chromosomes
120%
100%
80%
60%
40%
20%
0%
HWT
K9R
G
S

Figure 5.2: S phase progression is disrupted in K9R salivary gland nuclei. PCNA staining patterns on endoreplicating salivary gland polytene chromosomes identifies various stages of S phase from early (ER) to late (VLR). Bottom bar graph indicates percentage of total PCNA positive polytene chromosomes categorized into 5 different stages of S phase. Number of polytenes categorized was greater than 100 for each genotype. Experiment performed in collaboration with Robin Armstrong and Samuel Chao.


G Phase
ER
E-MR
M-LR
LR
VLR
PCNA
+ DAPI
Merge

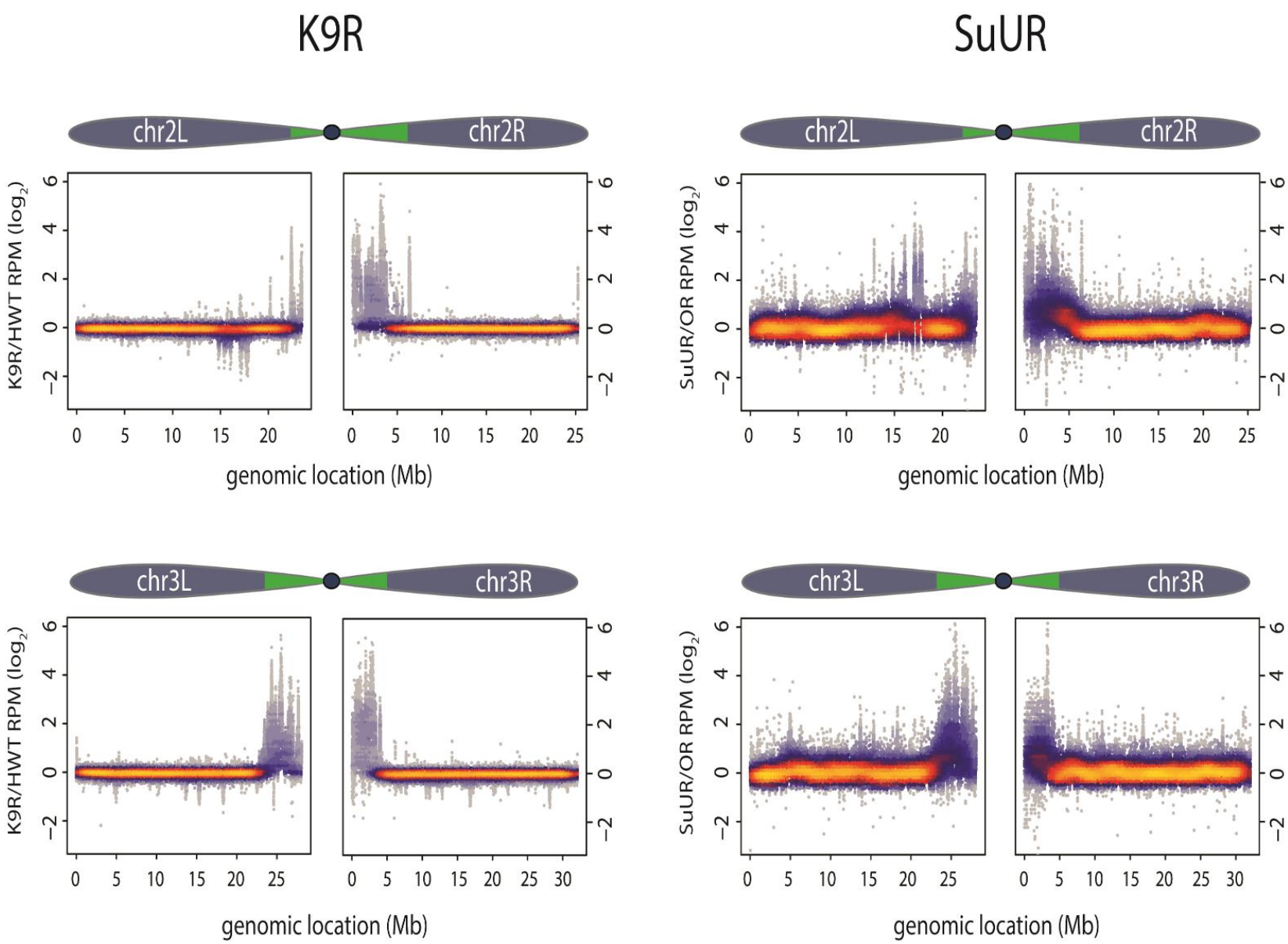

Percentage of Cells
50
40
30
20
10
0
HWT
H3K9R
ER
E-MR
M-LR
LR
VLR

Role of H3K9 is regulating under-replication

One possible explanation for this phenotype is altered under-replication. If regions of intercalary or pericentromeric heterochromatin are being replicated more frequently in *K9R* mutants, we might expect an increase in the percentage of chromosomes with a late replicating PCNA pattern. We examined relative copy number of *HWT* and *K9R* salivary gland genomes through whole-genome sequencing. Briefly, nuclear extracts from whole salivary glands were sonicated, and DNA was precipitated for next-generation sequencing library preparation. Sequencing reads were mapped to the genome using Bowtie2 (Langmead and Salzberg 2012), and bedtools was used to count the number of sequencing reads overlapping 1 kb windows tiled across the genome (Quinlan and Hall 2010). We then normalized the resulting counts by read depth and plotted the K9R/HWT $\log_2$ ratio of normalized read counts at each 1 kb window versus the genomic coordinate of the window. Windows that have a value greater than 0 indicate regions where under-replication is suppressed, whereas regions where under-replication is enhanced exhibit a value less than 0. Importantly, pericentromeric regions of *H3K9R* mutants have suppressed the under-replication normally present in *HWT* salivary glands (Figure 5.3). These data indicate that regions of pericentromeric heterochromatin have a higher copy number in *K9R* mutants compared to *HWT*. In contrast, in regions of intercalary heterochromatin along chromosome arms, under-replication is maintained in *K9R* salivary glands. Using previously generated data from salivary glands of SuUR mutants and a wild-type control (OR), we examined copy number differences between the two genotypes with the above pipeline (Nordman et al. 2014). As reported, under-replication is suppressed in SuUR mutants at both pericentromeric and intercalary heterochromatin (Figure 5.3),

suggesting that SuUR operates through an H3K9-dependent mechanism in pericentromeric heterochromatin and an H3K9-independent mechanism in intercalary heterochromatin.

Figure 5.3: H3K9R mutants suppress under-replication at pericentromeric heterochromatin. Mutant/control ratio of normalized whole genome sequencing reads from H3K9R or SuUR larval salivary glands within 1-kb windows (dots) tiled across chromosomes 2 and 3. SuUR data taken from Nordman et al. 2014. Values above 0 indicate higher copy number, and thus less under-replication. Approximate locations of euchromatin and pericentromeric heterochromatin in blue and green, respectively (Riddle et al. 2011; Hoskins et al. 2015). Experiment performed in collaboration with Robin Armstrong.


K9R
SuUR
chr2L
chr2R
chr3L
chr3R
K9R/HWT RPM (log2)
SuUR/OR RPM (log2)
genomic location (Mb)

Discussion

S phase progression in *K9R* mutants

Through PCNA staining of salivary gland polytene chromosomes we determined that *K9R* mutants exhibit a smaller percentage of replicating cells compared to *HWT*, as well as variation in the percentage of chromosomes at different stages of S phase. In diploid cells we have also observed a change in cell cycle progression. Mitotic recombination to generate twin spots was used to demonstrate that the size of K9R clones was two thirds the size of *HWT* clones (Figure 2.1); however, the number of nuclei per area was equivalent between *HWT* and *K9R* clones (data not shown). These data suggest that diploid *K9R* cell proliferation is slightly slower than *HWT*. Moreover, through cell cycle analysis by flow cytometry, we observed *K9R* imaginal wing discs have a smaller percentage of cells in S phase and a slightly larger percentage of cells in G2 phase compared to *HWT* (Figure 4.1).

The cause of the disruption to the salivary gland endocycle is unclear. The small percentage of replicating cells in *K9R* samples could be the result of developmental variation between larvae of the two genotypes, as the number of actively replicating nuclei decreases as the 3rd instar larval stage finishes. This possibility is unlikely as K9R mutants are developmentally delayed by 1-2 days and should therefore exhibit an increase in the percentage of actively replicating cells if developmental differences underlay this phenotype. An alternative possibility is that G1 phase is longer in *K9R* mutants or S phase is shorter. Additionally, the entire endocycle could progress faster in *K9R* mutants and could finish earlier in development. If replication timing was disrupted, and late replicating regions replicated along with early replicating regions, duplication of

the entire genome may occur faster. This hypothesis is supported by an increase in the percentage of S phase polytene chromosomes with an early PCNA pattern (Figure 5.2). Although many chromosomal regions may replicate earlier in *K9R* mutants, some pericentromeric regions of the salivary gland genome may remain late replicating, similar to what we observe in *K9R* diploid wing discs cells via Repli-seq (Figure 4.5). The advancement of some pericentromeric regions but not others would result in an increase in the percentage of chromosomes with the latest PCNA pattern (VLR), which is the phenotype we observe (Figure 5.2). In other words, if only a portion of normally late regions begin to replicate earlier, chromosomes that would normally have mid S phase PCNA patterns (M-LR and LR) would look more like early or very late S phase patterns (ER or VLR); however, this interpretation is speculative.

To distinguish between these possibilities, future experiments should determine the absolute length of S phase. We have crossed a PCNA-RFP transgene (gift from the Di Talia lab at Duke University) into the *HWT* and *K9R* backgrounds to permit live imaging of salivary glands. Fluorescent signal of this transgene is intense and easily distinguishable from background signal in live salivary glands, but future work must optimize imaging culture conditions.

Role of H3K9 in under-replication of salivary gland nuclei

We also observed that in *K9R* mutants under-replication was suppressed in pericentromeric regions but not intercalary heterochromatin on chromosome arms (Figure 5.3). Under-replication was unaltered in other histone mutants including, *K20A*, *K16R*, and a *K27R* mutant that is heterozygous for the histone deletion and contains 40 copies of the K27R histone repeat (*K27R* mutants with no endogenous histones are

embryonic lethal). In SuUR mutants under-replication is suppressed in both types of heterochromatin. These data suggest that SuUR mediated repression is dependent on H3K9 in pericentromeric heterochromatin but not intercalary heterochromatin. SuUR recruitment at sites of under-replication on chromosome arms is thought to be dependent on H3K27me, as 60% of these regions overlap Polycomb group binding sites (Zhimulev et al. 2003b). We did not observe suppression of under-replication in a K27R mutant with ~100 copies of wild-type histone repeats and 40 copies of K27R histone repeats; however, H3K27me on endogenous histones present may have been sufficient to mediate under-replication (data not shown).

The suppression of under-replication at pericentromeres in both K9R and SuUR mutants is consistent with the ability of H3K9me to recruit HP1 in these regions (Penke et al. 2016) and the interdependence of HP1 and SuUR (Pindyurin et al. 2008). However, a direct test of this model has yet to be completed. Importantly, future experiments should determine if H3K9 is necessary for SuUR recruitment by examining the localization of SuUR in an *K9R* mutant. This experiment can be accomplished through the use of a SuUR-GFP transgene or a SuUR antibody. Overall, our data demonstrate an essential role for H3K9 in mediating under-replication likely through recruitment of SuUR.

CHAPTER 6- FUTURE DIRECTIONS FOR ENTERPRISING POST-DOCS OR GRADUATE STUDENTS

Histone Compensation

Through characterization of the histone replacement platform we determined that although there are ~200 copies of the histone repeat in wild-type flies, 12 copies are sufficient to rescue the viability and fertility of histone deletion mutants. We previously demonstrated that mRNA and protein levels of histones in wild-type and HWT (12x rescue) were equivalent, indicating the existence of a mechanism for histone dosage compensation (McKay et al. 2015). Several possibilities can explain this phenomenon. First, in wild-type flies that contain all the endogenous histone genes, many of these histone clusters may be inactive throughout the majority of development or in most cells. In a specific context, such as oogenesis where high levels of histone mRNA or protein must be packed into the developing egg, the usually inactive repeats may become activated. Indeed, in HWT females we observe defects in oogenesis and meiotic recombination, suggesting that 12 copies of the histone repeat may be insufficient for this developmental stage. Interestingly, increasing the copy number of ectopic histone repeats to 40 ameliorates the meiotic recombination defect (Talia Hatkevich, personal communication). This model raises questions of which histone repeats are active in wild-type flies and how activation and inactivation of repeats is regulated.

Alternatively, histone mRNA or protein levels from the ectopic 12 histone repeats may be increased in the absence of endogenous histones, analogous to the upregulation of the male X chromosome. The potential ability of the ectopic histone locus to upregulate histone expression may be due to the chromatin environment at VK33. This genomic location may be more amenable to histone expression than the heterochromatic location of the endogenous locus on chromosome 2L near the pericentromere. Lower levels of histone expression may also result in a feedback mechanism that regulates trans-acting factors to increase histone expression. For example, the addition of transcriptional activators or the removal of transcriptional repressors could result in increased histone expression. One possible trans-acting factor that may regulate histone levels is Mute, a factor that localizes to the histone locus body. The loss of mute results in increased histone H3 and H4 transcript levels, indicating Mute functions as a transcriptional repressor of the histone locus (Bulchand et al. 2010). Loss of Mute from the ectopic histone repeats could allow increased histone expression in HWT animals.

Finally, decreased histone expression from replication-dependent histone genes could be compensated for by increased expression of histone variants. Sakai et al. demonstrated that in variant H3.3 null flies canonical H3 is upregulated (2009). A similar mechanism could exist to compensate for low levels of canonical H3 in HWT animals. A simple experiment to test this model is to remove H3.3 from HWT animals and measure viability. Both H3.3 null mutants and HWT animals are viable, but their combination would result in a synthetic lethal phenotype if HWT animals rely on H3.3 to maintain proper histone levels. One problem with this model is that no variant for H2B has been

identified in *Drosophila*. Canonical histone transcripts upregulated in H3.3 null mutants were polyadenylated (Sakai et al. 2009), suggesting the endogenous, canonical histone genes can be transcribed distinctly from their normal stem loop dependent regulatory mechanism. Transcription through this alternative mechanism could increase histone levels by producing transcripts throughout the cell cycle as opposed to just S phase. The H2A-H2B shared promoter region may allow upregulation of these two transcripts from the endogenous locus to complement potential H3.3 and His4r (histone H4 replacement) upregulation in HWT animals.

Lethality of $H3^{K9R}$ and $H3.3^{K9R}$ $H3^{K9R}$ Mutants

One of the most challenging unanswered questions from my work is what causes the lethality of K9R mutants. Several different processes are disrupted in canonical K9R mutants including, transposon activation and mobilization, chromosome segregation, cell proliferation, increased DNA damage, and altered replication timing. Furthermore, although few protein coding genes are altered in canonical K9R mutants, we cannot rule out the possibility that those few changes are affecting viability. In contrast, variant and canonical K9R mutants have numerous protein coding genes that are significantly altered compared to controls (Figure 3.13). Many of these genes occur in regions enriched in K9ac (Figure 3.16). The abundance of misregulated processes in K9R mutants, in addition to those we have not discovered yet, makes teasing apart the cause of lethality a complicated problem.

Analysis of additional amino acid substitutions at K9 may shed some light on this issue. Specifically, examination of a canonical K9Q mutant could separate functions of

K9me and K9ac. Glutamine is often used as an acetyl mimic, and if K9Q functions analogously to an acetylated K9, we could establish conditions where K9me is prevented but K9ac is present. Analyzing these phenotypes will be complicated, however, as K9me and K9ac are almost certainly required at different regions of the genome, and amino acid substitutions using the histone replacement platform result in global changes. Alternatively, if one could rescue the viability of K9R mutants by suppressing one of the misregulated processes, we could make conclusions regarding the cause of lethality. For example, if transposon mobilization could be suppressed by some mechanism, we could determine how detrimental mobilization is for the animal. These specific rescue experiments would be difficult to carry out, but would be the most informative.

Perhaps, a combination of misregulated processes leads to the early death of H3K9R mutants, and it is less important to determine the cause of death than to identify the host of regulatory processes that H3K9 contributes to. At least, that's what I tell myself to help me sleep at night.

Chromatin Organization of $H3^{K9R}$ Mutants

We observed that canonical K9R mutants had increased chromatin accessibility at pericentromeric heterochromatin as measured by FAIRE-seq (Figure 2.3), which measures regions of nucleosome depletion. The defect in chromatin organization that leads to increased FAIRE-seq signal in K9R mutants is unclear. One possibility is that increased transcription of transposons and other repetitive transcripts displaces nucleosomes as RNA polymerases traverse DNA. Alternatively, higher order chromatin

organization or folding could be disrupted, allowing factors access to genomic regions that normally would be kept in an inaccessible chromatin environment. The activity of these factors, which could include polymerases or chromatin remodelers, may result in altered nucleosome occupancy. In this model the mechanism of establishing heterochromatic inaccessibility could take several forms. 3D architecture may play a role in this inaccessibility, and it would be of great interest to measure changes in 3D chromatin organization in K9R mutants via Hi-C or an analogous approach (Rao et al. 2017). Additionally, heterochromatin was recently demonstrated to exist in a phase transition (Strom et al. 2017). The boundaries of a liquid-liquid separation may prevent various trans-acting factors from accessing heterochromatin and altering nucleosome occupancy. Determining if phase transitions are disrupted in K9R mutants deserves attention.

Histone PTM Landscape in K9R Mutants

A plethora of different histone PTMs exist on all types of histones (Rothbart and Strahl 2014; Hake et al. 2006). An interesting question is how various histone PTMs influence the deposition or removal of other histone PTMs. For example, H3K9me has been suggested to influence H4K20me levels (Schotta et al. 2004). A largely unexplored question in my work is how the absence of H3K9 methylation or acetylation in H3K9R mutants influences the histone PTM landscape. Preliminary experiments suggest that H3K27me3 levels are unchanged in H3K9R mutants as evidenced by polytene chromosome stains (data not shown). However, conclusions regarding the status of H3K27me levels or other histone PTMs will require more quantitative and

sensitive measures such as western blot and ChIP-seq experiments. Of particular interest, mass spectrometry analysis of histone proteins represents an exciting potential avenue to explore and characterize the PTM landscape in H3K9R mutants.

Local Regulation of Replication in $H3^{K9R}$ Mutants

In Chapter 4 we demonstrated that increased chromatin accessibility and altered RNA expression were permissive but not instructive for altered replication timing in K9R mutants. Interestingly, many regions of altered chromatin and transcription did not change replication timing, suggesting another factor differentiates these regions from those that did change timing. One possibility is that permissive regions that alter replication license more origins than permissive regions that do not change replication. To test this hypothesis, it will be important to measure Orc or MCM localization across the genome in HWT and K9R samples. These experiments may be accomplished through a newly established protocol called CUT&RUN, which provides genome-wide protein localization information from lower inputs with higher signal-to-noise ratios (Skene and Henikoff 2017). CUT&RUN relies on targeting MNase to a protein using an antibody. MNase digests the DNA surrounding the protein and releases a fragment that can be collected and sequenced. In collaboration with the McKay lab, we have used this procedure to generate a genome-wide binding profile for the transcription factor E93 that matches the profile generated from a traditional ChIP-seq approach. Applying this protocol to Orc2 would permit a genome-wide measure of origin licensing in imaginal wing discs, which would be difficult to acquire using ChIP-seq due to a high input requirement.

In addition to origin licensing, differences in origin activation may influence whether a permissive region is capable of altering replication timing in K9R mutants. Two trans-acting factors HP1 and Rif1 have previously been shown to modulate replication timing (Schwaiger et al. 2010; Yamazaki et al. 2012). Both of these factors are known to associate with heterochromatin and could influence altered replication timing in K9R mutants. Altered HP1 localization is closely correlated with changes in chromatin accessibility in K9R mutants (Figure 4.9); therefore, similar to chromatin accessibility, HP1 does not appear to instruct replication timing changes. Using CUT&RUN to measure Rif1 localization in HWT and K9R mutants would determine if Rif1 contributes to observed replication timing changes.

Consequences of Altered Replication Timing in K9R Mutants

Several groups have observed altered replication timing in a number of cancers, including renal, ovarian, leukemia, and lymphoma (Ryba et al. 2012; Korenstein-Ilan et al. 2002; Smith et al. 2001; Sun et al. 2001). For example, bone marrow samples from children with leukemia exhibited regions of the genome that replicated at different times than non-cancerous cells (Ryba et al. 2012). These regions with altered replication timing were also more likely to have mutations or chromosomal rearrangements. Both mutations and chromosomal rearrangements are linked to disease progression, suggesting misregulation of origin licensing or activation predisposes cells to advanced disease states. Importantly, the link between DNA replication and advancement of disease is a correlation; we do not know if changes in replication timing *cause* mutations or rearrangements.

Mutations and chromosomal rearrangements can both result from improper repair of DNA damage. Changes in replication timing could result in DNA damage, possibly due to replication fork stalling or fork collision. We have observed that K9R mutants have elevated levels of DNA damage compared to HWT samples as measured by γ-H2Av signal (data not shown). Similarly, ~35% of mitotic cells from 3rd instar brains exhibit anaphase bridging. The cause of this DNA damage is as of yet unclear; however, we hypothesize that at least some of this damage is the result of misregulated replication. Future experiments could examine polytene chromosomes to determine if regions of active replication (marked by PCNA staining) overlap regions of DNA damage (marked by γ-H2Av staining). Results of this experiment may motivate a genome-wide approach to examining DNA damage patterns in HWT and K9R animals using γ-H2Av CUT&RUN. These correlative approaches might suggest a link between altered replication timing and DNA damage.

Concluding Remarks

The fortuitous organization of histone genes in *Drosophila* combined with the power and versatility of its genetic tools has afforded a unique and potent opportunity to explore histone post-translational modifications in an animal model. Using the histone replacement platform, I have investigated the roles of H3K9 and H3.3K9 in various biological processes and have used H3K9R mutants as a tool to investigate general principals of genome regulation. As described above, my interrogation into the complexities of heterochromatin biology represents only a fraction of the knowledge we can obtain through study of the H3K9R mutant, which in and of itself, represents only a

single application of the histone replacement approach. I hope the work laid out in this dissertation has laid a solid foundation for study of numerous histone residues in many other scientific journeys.

APPENDIX I

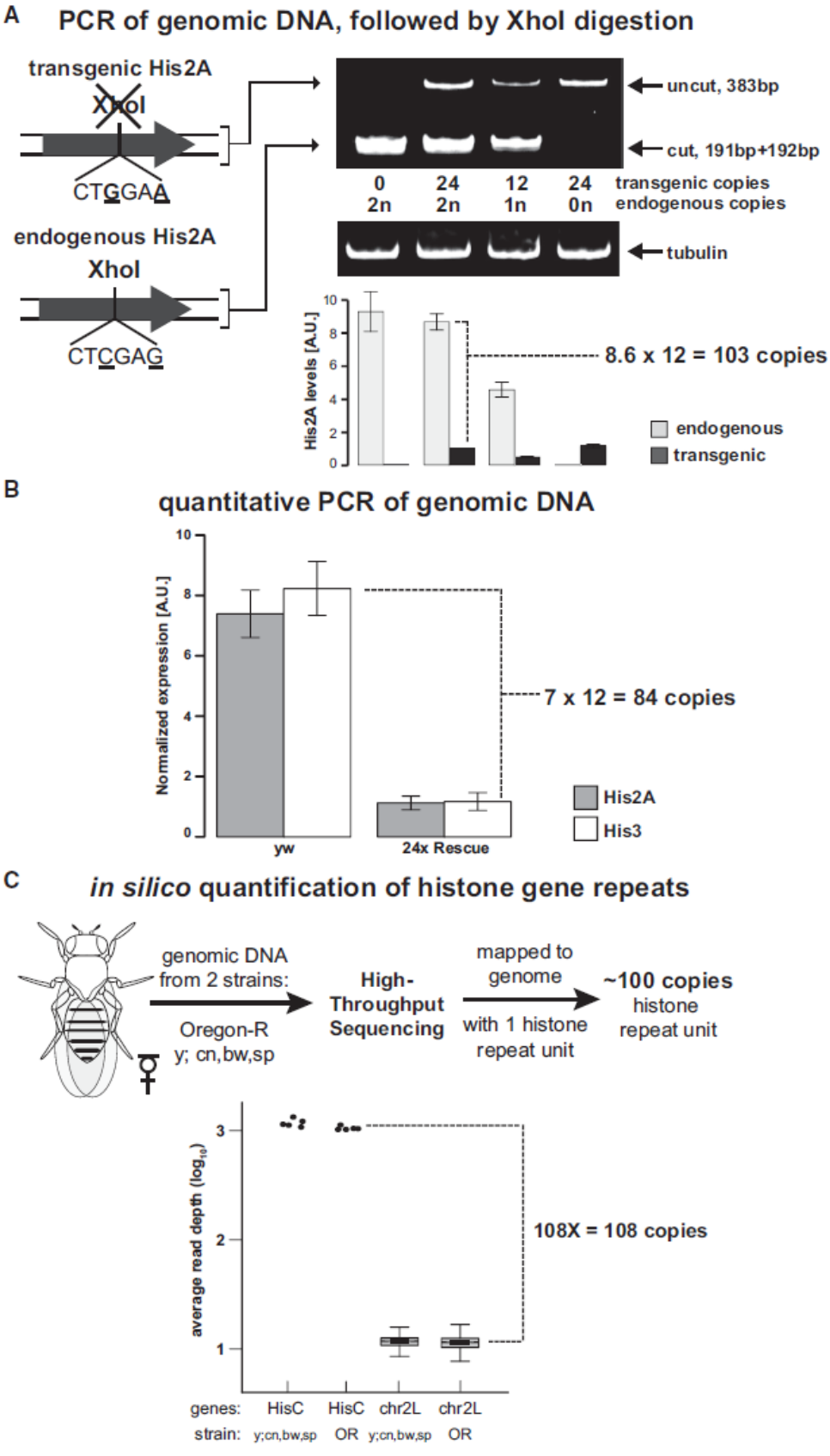
Histone replacement platform reveals histone dosage compensation mechanism

Surprisingly, despite the presence of ~100 histone repeats in wild-type animals, both the viability and fertility of endogenous histone deletion mutants can be rescued with an ectopic array of only 12 histone repeats (12x-Rescue) (McKay et al. 2015). Early studies of histone gene organization in *Drosophila melanogaster* used re-association kinetics to estimate that ~100 copies of the histone repeat were tandemly repeated in wild-type animals (Lifton et al. 1978); however, recent assemblies of the *Drosophila* genome based on whole-genome sequencing included only 23 copies of the histone repeat. We therefore took two complementary approaches to determine histone copy number. To discriminate between endogenous and transgenic His2A DNA, we engineered a silent mutation in an XhoI site within the transgenic His2A gene (Figure 2A). Using PCR primers that recognize both endogenous and transgenic templates, we amplified His2A genomic DNA from four genotypes and digested the PCR products with XhoI, cutting the endogenous His2A fragment in two equal halves while leaving the transgenic His2A product intact. Following electrophoresis, quantification of band intensities revealed that the endogenous His2A template is 8-fold more abundant than the transgenic His2A template (Figure A1.1A). Importantly, semiquantitative PCR reactions from both the endogenous and transgenic His2A templates are within the linear range, as shown by XhoI digestion assays using genomic DNA from four genotypes with different histone gene copy numbers (Figure A1.1A). Consistent with measurements from the XhoI digestion assay, real-time PCR indicates that the His2A

and His3 genes are 7-fold more abundant in wild-type flies than 24x-Rescue flies (homozygous for 12x-Rescue transgene) (Figure A1.1B). These experiments indicate that the haploid Drosophila genome contains approximately 100 histone repeats.

Second, we calculated the histone gene copy number using high-throughput sequencing analysis. We reasoned that the abundance of histone sequences relative to those of other genes on chromosome 2L would reflect the number of copies of histone genes in the genome. To accurately measure their abundance, we sequenced genomic DNA from two different strains and mapped reads to a custom Drosophila genome containing a single histone gene repeat unit. Comparison of the average read density across the coding sequence of each histone gene to the average read density across coding sequences of the remaining annotated genes on chromosome 2L revealed that the histone genes are ~100-fold more abundant (Figure A1.1C), consistent with our PCR assays and the original estimates (Lifton et al. 1978)

Figure A1.1 The Haploid Drosophila Genome Contains 100 Copies of the Histone Repeat Unit. (A) Ethidium bromide-stained gel of XhoI-digested PCR products of endogenous and transgenic H2A genes. For each of the four genotypes barplots of normalized band intensity are shown below each lane. Error bars represent SEM. (B) Barplots of normalized real-time PCR results for wild-type (yw) and 24x Rescue (homozygous 12x-Rescue transgene) genotypes using primers to H2A and H3. Error bars represent SEM. (C) Flow chart and plots of in silico quantification of histone gene repeats for two wild-type strains (Oregon R [OR], and y;cn,bw,sp). HisC: total read depth for each of the five replication-dependent genes; chr2L: box plots of average read depth for the remaining genes on chromosome 2L. The box represents the inner quartile range (IQR), and whiskers represent 1.5-times IQR. For clarity, outliers were not plotted. Experiments in panels A and B of this figure were completed by Taylor Penke and panel C by Daniel McKay.

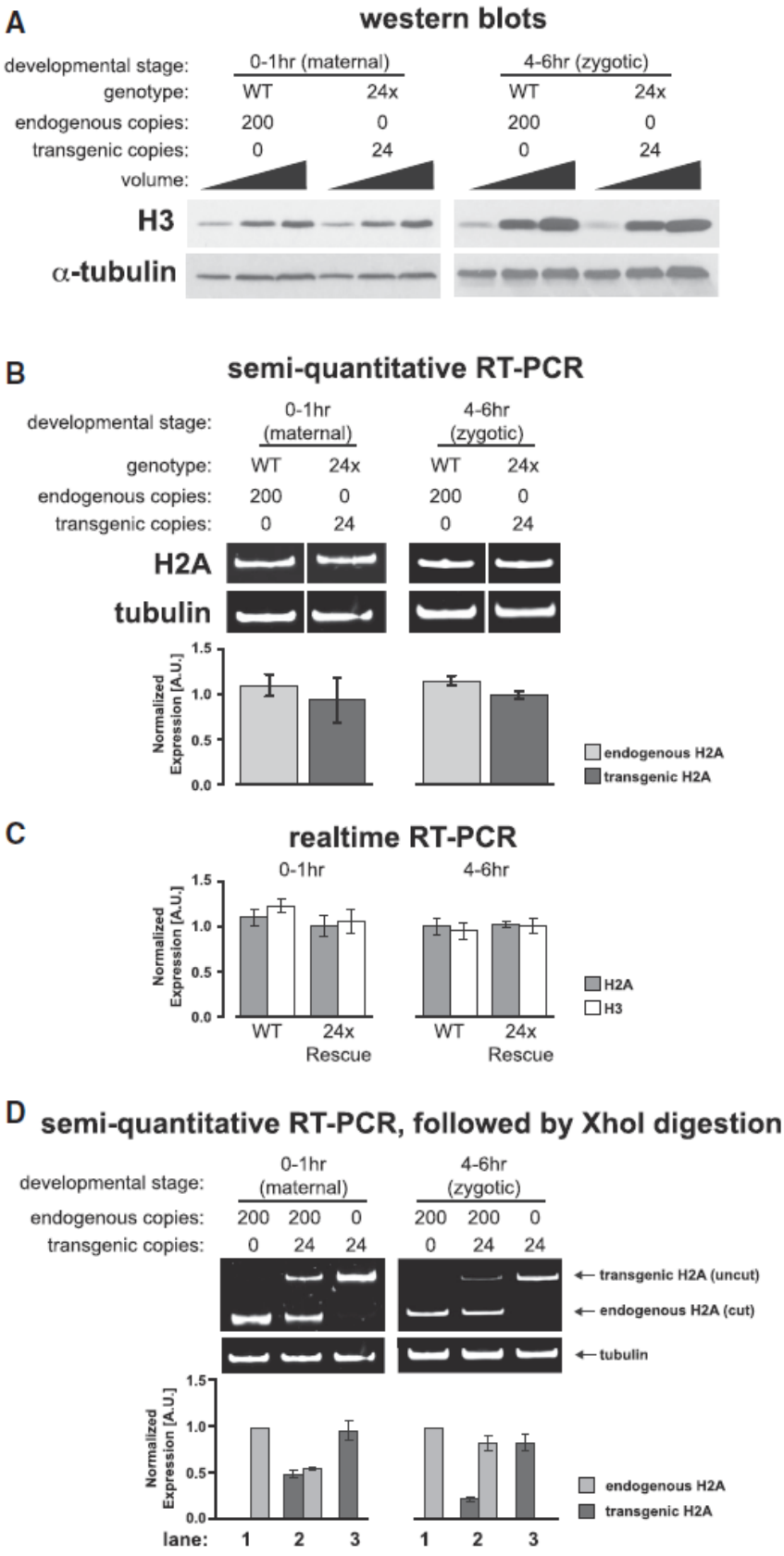

A
PCR of genomic DNA, followed by XhoI digestion
transgenic His2A
XhoI
CTCGAG
endogenous His2A
XhoI
CTCGAG
uncut, 383bp
cut, 191bp+192bp
0 24 12 24 transgenic copies
2n 2n 1n 0n endogenous copies
tubulin
His2A levels [A.U.]
8.6 x 12 = 103 copies
endogenous
transgenic
B
quantitative PCR of genomic DNA
Normalized expression [A.U.]
7 x 12 = 84 copies
His2A
His3
yw
24x Rescue
C
in silico quantification of histone gene repeats
genomic DNA from 2 strains:
Oregon-R
y; cn,bw,sp
High-Throughput Sequencing
mapped to genome
with 1 histone repeat unit
~100 copies
histone repeat unit
average read depth (log10)
108X = 108 copies
genes: HisC HisC chr2L chr2L
strain: y;cn,bw,sp OR y;cn,bw,sp OR

The preceding experiments show that wild-type diploid flies contain ~200 copies of the histone repeat unit, and yet a single 12x histone transgene is sufficient to support development of flies lacking all endogenous histone genes. We therefore compared expression levels between the endogenous and transgenic histone genes. Western blot and RT-PCR analysis at two stages of embryogenesis (0–1 hr and 4–6 hr) showed no significant differences in histone protein or mRNA levels between wild-type and 24x-Rescue flies (Figures A1.2A-C). Because the zygotic histone genes are not active in 0–1 hour embryos, histone levels at this time point reflect maternal protein and mRNA derived from the activity of the histone genes during oogenesis. The 4–6 hour time point includes cell cycle 15, when zygotic histone gene activity is first required due to destruction of the maternal histone supply. Despite different demands on histone gene activity between these two stages, the 24x transgenic histone genes produce the same amount of protein and mRNA as 200 copies of the endogenous histone genes (Figures A1.2A-C). Thus, histone replacement flies express equivalent steady-state levels of histones as wild-type flies, despite a 10-fold difference in gene copy number.

Because both the protein levels and the amino acid sequences of the endogenous and transgenic histones are identical, we infer that the nucleosome and higher-order chromatin organization is similar across the genome in wild-type and 24x-Rescue flies. In addition, 12x- and 24x-Rescue flies show no increase in sensitivity to the DNA-replication inhibiting agent hydroxyurea (data not shown), as we hypothesize would occur if histone production during S phase was limiting in these animals. The similar amount of mRNA produced in the 24x-Rescue and wild-type flies suggests the existence of a histone gene dosage compensation mechanism. To test whether such a

mechanism exists, we compared the levels of mRNA in wild-type and 24x-Rescue flies to those in flies containing both endogenous and transgenic histone genes (‘‘endogenous + 24x’’), discriminating between them using the XhoI digestion assay described above (Figure A1.1A). Similar to the results from undigested samples (Figure A1.2B), His2A mRNA levels are the same in wild-type and 24x-Rescue embryos (Figure 4D, lane 1 and lane 3). In contrast, His2A mRNA levels originating from both endogenous and ectopic histone genes are reduced in ‘‘endogenous + 24x’’ embryos compared with wild-type and 24x-Rescue embryos (Figure A1.2D, lane 2). Importantly, the sum of endogenous plus ectopic His2A mRNA in ‘‘endogenous + 24x’’ flies equals the levels observed in wild-type or 24x-Rescue embryos. Thus, the total amount of histone mRNA at a given stage of embryogenesis is the same for each genotype, suggesting that the steady state level of RNA expressed from individual histone genes is scaled to the total number of histone genes present in the genome.

Figure A1.2 Transgenic Histone Arrays Are Expressed at Levels Similar to the Endogenous Genes (A) Western blot of wild-type (WT) and 24x Rescue genotypes at 0–1 hr and 4–6 hr after egg laying. (B) Ethidium bromide stained gel of RT-PCR products from 0–1 hr and 4–6 hr wild-type (WT) and 24x Rescue embryos. Barplots of normalized band intensity are shown below each lane. Error bars represent SEM. (C) Barplots of normalized real-time RT-PCR results for H2A and H3 in 0–1 hr and 4–6 hr wild-type (WT) and 24x Rescue embryos. Error bars represent SEM. (D) Ethidium bromide stained gel of XhoI-digested RT-PCR products from 0–1 hr and 4–6 hr embryos for three genotypes: wild-type (lane 1); +/+;12xWT/ 12xWT (lane 2); 24x Rescue (lane 3). Barplots of normalized band intensity are shown. Error bars represent SEM. Experiments performed by Taylor Penke.


A
western blots
developmental stage:
0-1hr (maternal)
4-6hr (zygotic)
genotype:
WT
24x
WT
24x
endogenous copies:
200
0
200
0
transgenic copies:
0
24
0
24
volume:
H3
α-tubulin
B
semi-quantitative RT-PCR
developmental stage:
0-1hr (maternal)
4-6hr (zygotic)
genotype:
WT
24x
WT
24x
endogenous copies:
200
0
200
0
transgenic copies:
0
24
0
24
H2A
tubulin
Normalized Expression [A.U.]
1.5
1.0
0.5
0.0
endogenous H2A
transgenic H2A
C
realtime RT-PCR
0-1hr
4-6hr
Normalized Expression [A.U.]
1.5
1.0
0.5
0.0
H2A
H3
WT
24x Rescue
WT
24x Rescue
D
semi-quantitative RT-PCR, followed by XhoI digestion
developmental stage:
0-1hr (maternal)
4-6hr (zygotic)
endogenous copies:
200
200
0
200
200
0
transgenic copies:
0
24
24
0
24
24
transgenic H2A (uncut)
endogenous H2A (cut)
tubulin
Normalized Expression [A.U.]
1.5
1.0
0.5
0.0
endogenous H2A
transgenic H2A
lane:
1
2
3
1
2
3

REFERENCES

Ahmad K, Henikoff S. 2002. The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. *Mol Cell* **9**: 1191–1200.

Alberts B, Johnson A, Lewis J. 2002. *Biology of the Cell*. 4th ed. New York: Garland Science.

Allfrey VG, Faulkner R, Mirsky AE. 1964. Acetylation and Methylation of Histones and Their Possible Role in the Regulation of Rna Synthesis. *Proc Natl Acad Sci U S A* **51**: 786–94.

Allis CD, Wiggins JC. 1984. Histone rearrangements accompany nuclear differentiation and dedifferentiation in Tetrahymena. *Dev Biol* **101**: 282–294.

Allshire RC, Nimmo ER, Ekwall K, Javerzat JP, Cranston G. 1995. Mutations derepressing silent centromeric domains in fission yeast disrupt chromosome segregation. *Genes Dev* **9**: 218–233.

Andreyeva EN, Kolesnikova TD, Belyaeva ES, Glaser RL, Zhimulev IF. 2008. Local DNA underreplication correlates with accumulation of phosphorylated H2Av in the Drosophila melanogaster polytene chromosomes. *Chromosom Res* **16**: 851–862.

Azzaz AM, Vitalini MW, Thomas AS, Price JP, Blacketer MJ, Cryderman DE, Zirbel LN, Woodcock CL, Elcock AH, Wallrath LL, et al. 2014. Human heterochromatin protein 1α promotes nucleosome associations that drive chromatin condensation. *J Biol Chem* **289**: 6850–6861.

Badugu RK, Shareef MM, Kellum R. 2003. Novel Drosophila Heterochromatin Protein I (HP1)/Origin Recognition Complex-associated Protein (HOAP) Repeat Motif in HP1/HOAP Interactions and Chromocenter Associations. *J Biol Chem* **278**: 34491–34498.

Bannister a J, Zegerman P, Partridge JF, Miska E a, Thomas JO, Allshire RC, Kouzarides T. 2001. Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* **410**: 120–4.

Bannister AJ, Schneider R, Myers FA, Thorne AW, Crane-Robinson C, Kouzarides T. 2005. Spatial distribution of di- and tri-methyl lysine 36 of histone H3 at active genes. *J Biol Chem* **280**: 17732–17736.

Beck DB, Oda H, Shen SS, Reinberg D. 2012. PR-Set7 and H4K20me1: at the crossroads of genome integrity, cell cycle, chromosome condensation, and transcription. *Genes Dev* **26**: 325–37.

Beermann W. 1963. Cytological Aspects of Information Transfer in Cellular

Differentiation. *Am Zool* **3**: 23–32.

Bell, Stephen P.; Stillman B. 1992. ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. *Nature* **357**.

Bell O, Schwaiger M, Oakeley EJ, Lienert F, Beisel C, Stadler MB, Schübeler D. 2010. Accessibility of the Drosophila genome discriminates PcG repression, H4K16 acetylation and replication timing. *Nat Struct Mol Biol* **17**: 894–900.

Belyaeva ES, Zhimulev IF, Volkova EI, Alekseyenko a a, Moshkin YM, Koryakov DE. 1998. Su(UR)ES: a gene suppressing DNA underreplication in intercalary and pericentric heterochromatin of Drosophila melanogaster polytene chromosomes. *Proc Natl Acad Sci U S A* **95**: 7532–7.

Bernard P, Maure JF, Partridge JF, Genier S, Javerzat JP, Allshire RC. 2001. Requirement of heterochromatin for cohesion at centromeres. *Science* **294**: 2539–42.

Bernstein BE, Kamal M, Lindblad-Toh K, Bekiranov S, Bailey DK, Huebert DJ, McMahon S, Karlsson EK, Kulbokas EJ, Gingeras TR, et al. 2005. Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell* **120**: 169–181.

Biggar KK, Li SS-C. 2014. Non-histone protein methylation as a regulator of cellular signalling and function. *Nat Rev Mol Cell Biol* **16**: 5–17.

Blower MD, Karpen GH. 2001. The role of Drosophila CID in kinetochore formation, cell-cycle progression and heterochromatin interactions. *Nat Cell Biol* **3**: 730–9.

Brennecke J, Aravin A a, Stark A, Dus M, Kellis M, Sachidanandam R, Hannon GJ. 2007. Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. *Cell* **128**: 1089–103.

Brink RA. 1964. *The Role of Chromosomes in Development*. Academic Press.

Brower-Toland B, Riddle NC, Jiang H, Huisinga KL, Elgin SCR. 2009. Multiple SET methyltransferases are required to maintain normal heterochromatin domains in the genome of Drosophila melanogaster. *Genetics* **181**: 1303–19.

Brown SW. 1966. Heterochromatin. *Science (80-)* **151**: 417–425.

Brown SW, Nur U. 1964. Heterochromatic Chromosomes in the Coccids. *Science* **145**: 130–136.

Bulchand S, Menon SD, George SE, Chia W. 2010. Muscle wasted: a novel component of the Drosophila histone locus body required for muscle integrity. *J Cell Sci* **123**: 2697–707.

Bush KM, Yuen BT, Barrilleaux BL, Riggs JW, O'Geen H, Cotterman RF, Knoepfler PS. 2013. Endogenous mammalian histone H3.3 exhibits chromatin-related functions during development. *Epigenetics Chromatin* **6**: 7.

Cai W, Jin Y, Girton J, Johansen J, Johansen KM. 2010. Preparation of Drosophila Polytene Chromosome Squashes for Antibody Labeling. *Jove* 2–5.

Canzio D, Chang EY, Shankar S, Kuchenbecker KM, Simon MD, Madhani HD, Narlikar GJ, Al-Sady B. 2011. Chromodomain-mediated oligomerization of HP1 suggests a nucleosome-bridging mechanism for heterochromatin assembly. *Mol Cell* **41**: 67–81.

Carlson SM, Gozani O. 2016. Nonhistone Lysine Methylation in the Regulation of Cancer Pathways. *Cold Spring Harb Perspect Med* **6**: a026435.

Celniker SE, Dillon LAL, Gerstein MB, Gunsalus KC, Henikoff S, Karpen GH, Kellis M, Lai EC, Lieb JD, MacAlpine DM, et al. 2009. FEATURE Unlocking the secrets of the genome. *Nature* **459**: 927–930.

Cenci G, Ciapponi L, Gatti M. 2005. The mechanism of telomere protection: A comparison between Drosophila and humans. *Chromosoma* **114**: 135–145.

Collart C, Allen GE, Bradshaw CR, Smith JC, Zegerman P. 2013. Titration of Four Replication Factors Is Essential for the Xenopus laevis Midblastula Transition. *Science (80-)* **341**: 893–896.

Corradini N, Rossi F, Giordano E, Caizzi R, Verní F, Dimitri P. 2007. Drosophila melanogaster as a model for studying protein-encoding genes that are resident in constitutive heterochromatin. *Heredity (Edinb)* **98**: 3–12.

Couldrey C, Carlton MB, Nolan PM, Colledge WH, Evans MJ. 1999. A retroviral gene trap insertion into the histone 3.3A gene causes partial neonatal lethality, stunted growth, neuromuscular deficits and male sub-fertility in transgenic mice. *Hum Mol Genet* **8**: 2489–2495.

Cryderman DE, Grade SK, Li Y, Fanti L, Pimpinelli S, Wallrath LL. 2005. Role of Drosophila HP1 in euchromatic gene expression. *Dev Dyn* **232**: 767–774.

Csink A, Henikoff S. 1996. Genetic modification of heterochromatic association and nuclear organization in Drosophila. *Nature* **381**: 529–531.

Cui B, Liu Y, Gorovsky MA. 2006. Deposition and function of histone H3 variants in Tetrahymena thermophila. *Mol Cell Biol* **26**: 7719–30.

Czech B, Malone CD, Zhou R, Stark A, Schlingeheyde C, Dus M, Perrimon N, Kellis M, Wohlschlegel JA, Sachidanandam R, et al. 2008. An endogenous small interfering RNA pathway in Drosophila. *Nature* **453**: 798–802.

Danzer JR, Wallrath LL. 2004. Mechanisms of HP1-mediated gene silencing in Drosophila. *Development* **131**: 3571–80.

Das SP, Borrman T, Liu VWT, Yang SCH, Bechhoefer J, Rhind N. 2015. Replication timing is regulated by the number of MCMs loaded at origins. *Genome Res* **25**: 1886–1892.

De Wit E, Greil F, Van Steensel B. 2007. High-resolution mapping reveals links of HP1 with active and inactive chromatin components. *PLoS Genet* **3**: 0346–0357.

Demerec M, Slizynska H. 1937. Mottled White 258-18 of Drosophila Melanogaster. *Genetics* **22**: 641–649.

Dernburg AF, Broman KW, Fung JC, Marshall WF, Philips J, Agard DA, Sedat JW. 1996. Perturbation of Nuclear Architecture by Long-Distance Chromosome Interactions. *Cell* **85**: 745–759.

Donley N, Thayer MJ. 2013. DNA replication timing, genome stability and cancer. Late and/or delayed DNA replication timing is associated with increased genomic instability. *Semin Cancer Biol* **23**: 80–89.

Dorer DR, Henikoff S. 1994. Expansions of transgene repeats cause heterochromatin formation and gene silencing in Drosophila. *Cell* **77**: 993–1002.

Durand-Dubief M, Ekwall K. 2008. Heterochromatin tells CENP-A where to go. *Bioessays* **30**: 526–9.

Eaton ML, Prinz JA, MacAlpine HK, Tretyakov G, Kharchenko P V., MacAlpine DM. 2011. Chromatin signatures of the Drosophila replication program. *Genome Res* **21**: 164–174.

Ebert A, Lein S, Schotta G, Reuter G. 2006. Histone modification and the control of heterochromatic gene silencing in Drosophila. *Chromosome Res* **14**: 377–92.

Ebert A, Schotta G, Lein S, Kubicek S, Krauss V, Jenuwein T, Reuter G. 2004. Su(var) genes regulate the balance between euchromatin and heterochromatin in Drosophila. *Genes Dev* **18**: 2973–83.

Eissenberg JC, Elgin SCR. 2014. HP1a: a structural chromosomal protein regulating transcription. *Trends Genet* **30**: 103–110.

Eissenberg JC, James TC, Foster-Hartnett DM, Hartnett T, Ngan V, Elgin SC. 1990. Mutation in a heterochromatin-specific chromosomal protein is associated with suppression of position-effect variegation in Drosophila melanogaster. *Proc Natl Acad Sci U S A* **87**: 9923–7.

Elgin SCR, Reuter G. 2013. Position-effect variegation, heterochromatin formation, and gene silencing in Drosophila. *Cold Spring Harb Perspect Biol* **5**.

Elsässer SJ, Noh K-M, Diaz N, Allis CD, Banaszynski L a. 2015. Histone H3.3 is required for endogenous retroviral element silencing in embryonic stem cells. *Nature* **522**: 240–246.

ENCODE Project Consortium TEP, Birney E, Stamatoyannopoulos JA, Dutta A, Guigó R, Gingeras TR, Margulies EH, Weng Z, Snyder M, Dermitzakis ET, et al. 2007. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* **447**: 799–816.

Eskeland R, Eberharter A, Imhof A. 2007. HP1 Binding to Chromatin Methylated at H3K9 Is Enhanced by Auxiliary Factors □. *Mol Cell Biol* **27**: 453–465.

Estella C, McKay DJ, Mann RS. 2008. Molecular Integration of Wingless, Decapentaplegic, and Autoregulatory Inputs into Distalless during Drosophila Leg Development. *Dev Cell* **14**: 86–96.

Figueiredo ML a, Philip P, Stenberg P, Larsson J. 2012. HP1a recruitment to promoters is independent of H3K9 methylation in Drosophila melanogaster. *PLoS Genet* **8**.

Filion GJ, van Bemmel JG, Braunschweig U, Talhout W, Kind J, Ward LD, Brugman W, de Castro IJ, Kerkhoven RM, Bussemaker HJ, et al. 2010. Systematic protein location mapping reveals five principal chromatin types in Drosophila cells. *Cell* **143**: 212–24.

Fillingham J, Recht J, Silva AC, Suter B, Emili A, Stagljar I, Krogan NJ, Allis CD, Keogh M-C, Greenblatt JF. 2008. Chaperone control of the activity and specificity of the histone H3 acetyltransferase Rtt109. *Mol Cell Biol* **28**: 4342–53.

Foti R, Gnan S, Cornacchia D, Dileep V, Bulut-Karslioglu A, Diehl S, Buness A, Klein FA, Huber W, Johnstone E, et al. 2016. Nuclear Architecture Organized by Rif1 Underpins the Replication-Timing Program. *Mol Cell* **61**: 260–273.

Fox DT, Duronio RJ. 2013. Endoreplication and polyploidy: insights into development and disease. *Development* **140**: 3–12.

Georgakopoulos T, Thireos G. 1992. Two distinct yeast transcriptional activators require the function of the GCN5 protein to promote normal levels of transcription. *EMBO J* **11**: 4145–52.

Gerace EL, Halic M, Moazed D. 2010. The methyltransferase activity of Clr4Suv39h triggers RNAi independently of histone H3K9 methylation. *Mol Cell* **39**: 360–72.

Ghildiyal M, Zamore PD. 2008. Endogenous siRNAs Derived from Transposons and mRNAs in Drosophila Somatic Cells. *Science* **320**: 1077–1081.

Gibson DG, Young L, Chuang R-Y, Venter JC, Hutchison C a, Smith HO, Iii CAH, America N. 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* **6**: 343–5.

Glozak MA, Sengupta N, Zhang X, Seto E. 2005. Acetylation and deacetylation of non-histone proteins. *Gene* **363**: 15–23.

Goldberg AD, Banaszynski LA, Noh K, Lewis PW, Elsaesser SJ, Stadler S, Dewell S, Law M, Guo X, Li X, et al. 2010. Distinct Factors Control Histone at Specific Genomic Regions. *Cell* **140**: 678–691.

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

Grewal SIS, Jia S. 2007. Heterochromatin revisited. *Nat Rev Genet* **8**: 35–46.

Günesdogan U, Jäckle H, Herzig A. 2010. A genetic system to assess in vivo the functions of histones and histone modifications in higher eukaryotes. *EMBO Rep* **11**: 772–6.

Hake SB, Garcia BA, Duncan EM, Kauer M, Dellaire G, Shabanowitz J, Bazett-jones DP, Allis CD, Hunt DF. 2006. Expression Patterns and Post-translational Modifications Associated with Mammalian Histone H3 Variants. *J Biol Chem* **281**: 559–568.

Hamperl S, Cimprich KA. 2016. Conflict Resolution in the Genome: How Transcription and Replication Make It Work. *Cell* **167**: 1455–1467.

Han BW, Wang W, Zamore PD, Weng Z. 2015. PiPipes: A set of pipelines for piRNA and transposon analysis via small RNA-seq, RNA-seq, degradome-and CAGE-seq, ChIP-seq and genomic DNA sequencing. *Bioinformatics* **31**: 593–595.

Hannah A. 1951. Localization and Function of Heterochromatin in Drosophila Melanogaster. *Adv Genet* **4**: 87–125.

Hartmann-Goldstein IJ. 1967. On the relationship between heterochromatization and variegation in Drosophila, with special reference to temperature sensitive periods. *Genet Res* **10**: 143.

Hawkins M, Retkute R, Müller CA, Saner N, Tanaka TU, deMoura APS, Nieduszynski CA. 2013. High-Resolution Replication Profiles Define the Stochastic Nature of Genome Replication Initiation and Termination. *Cell Rep* **5**: 1132–1141.

Haynes KA, Gracheva E, Elgin SCR. 2007. A distinct type of heterochromatin within Drosophila melanogaster chromosome 4. *Genetics* **175**: 1539–1542.

Heitz E. 1928. Das Heterochromatin der Moose. *Jahrb Wiss Bot* **69**: 762–818.

Heitz E. 1929. Heterochromatin, Chromocentren, Chromomeren. *Ber Bot Ges* **47**: 274–284.

Heitz E, Bauer H. 1933. Beweise fur die Chromosomennatur der Kernschleifen in Knauelkernen von bibio hortulanus L. *Z Zellforsch Mikrosk Anat* **17**: 67–82.

Henikoff S. 1990. Position-effect variegation after 60 years. *Trends Genet* **6**: 422–426.

Henikoff S. 1979. Position effects and variegation enhancers in an autosomal region of Drosophila melanogaster. *Genetics* **93**: 105–115.

Henikoff S, Ahmad K. 2005. Assembly of Variant Histones into Chromatin. *Annu Rev Cell Dev Biol*.

Herz H-M, Morgan M, Gao X, Jackson J, Rickels R, Swanson SK, Florens L, Washburn MP, Eissenberg JC, Shilatifard a. 2014. Histone H3 lysine-to-methionine mutants as a paradigm to study chromatin signaling. *Science (80-)* **345**: 1065–1070.

Hiratani I, Ryba T, Itoh M, Yokochi T, Schwaiger M, Chang CW, Lyou Y, Townes TM, Schübeler D, Gilbert DM. 2008. Global reorganization of replication domains during embryonic stem cell differentiation. *PLoS Biol* **6**: 2220–2236.

Ho JWK, Jung YL, Liu T, Alver BH, Lee S, Ikegami K, Sohn K-A, Minoda A, Tolstorukov MY, Appert A, et al. 2014. Comparative analysis of metazoan chromatin organization. *Nature* **512**: 449–452.

Hödl M, Basler K. 2012. Transcription in the Absence of Histone H3.2 and H3K4 Methylation. *Curr Biol* 2253–2257.

Hoskins RA, Carlson JW, Wan KH, Park S, Mendez I, Galle SE, Booth BW, Pfeiffer BD, George RA, Svirskas R, et al. 2015. The Release 6 reference sequence of the Drosophila melanogaster genome. *Genome Res* **25**: 445–458.

Hsu TC, Schmid W, Stubblefield E. 1964. *DNA Replication Sequences in Higher Animals*. ed. M. Locke. Academic Press, Inc.

Hua BL, Orr-Weaver TL. 2017. DNA Replication Control During Drosophila Development: Insights into the Onset of S Phase, Replication Initiation, and Fork Progression Insights into Regulation of DNA Replication from Localized Changes in DNA Copy Number 35. *Genetics* **207**: 29–47.

Huang J, Berger SL. 2008. The emerging field of dynamic lysine methylation of non-histone proteins. *Curr Opin Genet Dev* **18**: 152–8.

Huang RCC, Bonner J, Murray K. 1964. Physical and biological properties of soluble nucleohistones. *J Mol Biol* **8**: 54–64.

Huisinga KL, Brower-toland B, Elgin SCR. 2006. The contradictory definitions of heterochromatin : transcription and silencing. *Chromosoma* 110–122.

Jacob Y, Bergamin E, Donoghue MT a, Mongeon V, LeBlanc C, Voigt P, Underwood

CJ, Brunzelle JS, Michaels SD, Reinberg D, et al. 2014. Selective methylation of histone H3 variant H3.1 regulates heterochromatin replication. *Science* **343**: 1249–53.

James TC, Elgin SC. 1986. Identification of a nonhistone chromosomal protein associated with heterochromatin in Drosophila melanogaster and its gene. *Mol Cell Biol* **6**: 3862–3872.

Jia S, Yamada T, Grewal SIS. 2004. Heterochromatin Regulates Cell Type-Specific Long-Range Chromatin Interactions Essential for Directed Recombination. *Cell* **119**: 469–480.

Jin C, Felsenfeld G. 2007. Nucleosome stability mediated by histone variants H3.3 and H2A.Z. *Genes Dev* 1519–1529.

Jin C, Zang C, Wei G, Cui K, Peng W, Zhao K, Felsenfeld G, Heart TN. 2011. H3.3/H2A.Z double variant-containing nucleosomes mark “nucleosome-free regions” of active promoters and other regulatory regions in the human genome. *Nat Genet* **41**: 941–945.

Kaminker J, Bergman C, Kronmiller B, Carlson J, Svirskas R, Patel S, Frise E, Wheeler D, Lewis S, Rubin G, et al. 2002. The transposable elements of the Drosophila melanogaster euchromatin: a genomics perspective. *Genome Biol* **3**: 1–20.

Kawamura Y, Saito K, Kin T, Ono Y, Asai K, Sunohara T, Okada TN, Siomi MC, Siomi H. 2008. Drosophila endogenous small RNAs bind to Argonaute 2 in somatic cells. *Nature* **453**: 793–797.

Kellum R, Alberts BM. 1995. Heterochromatin protein 1 is required for correct chromosome segregation in Drosophila embryos. *J Cell Sci* **108**: 1419–31.

Kharchenko P, Alekseyenko A, Schwartz YB, Minoda A, Riddle NC, Sabo PJ, Larschan EN, Gorchakov AA, Karpen GH, Park PJ. 2011. Comprehensive analysis of the chromatin landscape in Drosophila melanogaster. *Nature* **471**: 480–485.

Khesin RB, Leibovitch BA. 1978. Influence of deficiency of the histone gene-containing 38B-40 region on X-chromosome template activity and the White gene position effect variegation in Drosophila melanogaster. *MGG Mol Gen Genet* **162**: 323–328.

Khush GS, Rick CM, Robinson RW. 1964. Genetic Activity in a Heterochromatic Chromosome Segment of the Tomato. *Science (80-)* **145**: 1432–1434.

Klattenhoff C, Xi H, Li C, Lee S, Xu J, Khurana JS, Zhang F, Schultz N, Koppetsch BS, Nowosielska A, et al. 2009. The Drosophila HP1 homolog Rhino is required for transposon silencing and piRNA production by dual-strand clusters. *Cell* **138**: 1137–49.

Kolesnikova TD, Posukh O V., Andreyeva EN, Bebyakina DS, Ivankin A V., Zhimulev

IF. 2013. Drosophila SUUR protein associates with PCNA and binds chromatin in a cell cycle-dependent manner. *Chromosoma* **122**: 55–66.

Kondo S, Ueda R. 2013. Highly Improved gene targeting by germline-specific Cas9 expression in Drosophila. *Genetics* **195**: 715–721.

Konrad. 1963. Protein Synthesis and Rna Synthesis During Mitosis in Animal Cells. 267–277.

Koren A, Handsaker RE, Kamitaki N, Karlić R, Ghosh S, Polak P, Eggan K, McCarroll SA. 2014. Genetic variation in human DNA replication timing. *Cell* **159**: 1015–1026.

Korenstein-Ilan A, Amiel A, Lalezari S, Lishner M, Avivi L. 2002. Allele-specific replication associated with aneuploidy in blood cells of patients with hematologic malignancies. *Cancer Genet Cytogenet* **139**: 97–103.

Koryakov DE, Walther M, Ebert A, Lein S, Zhimulev IF, Reuter G. 2011. The SUUR protein is involved in binding of SU(VAR)3-9 and methylation of H3K9 and H3K27 in chromosomes of Drosophila melanogaster. *Chromosom Res* **19**: 235–249.

Kraushaar DC, Jin W, Maunakea A, Abraham B, Ha M, Zhao K. 2013. Genome-wide incorporation dynamics reveal distinct categories of turnover for the histone variant H3 . 3. *Genome Biol* **14**: 1.

Kuo M-H, Zhou J, Jambeck P, Churchill MEA, Allis CD. 1998. Histone acetyltransferase activity of yeast Gcn5p is required for the activation of target genes in vivo. *Genes Dev* **12**: 627–639.

Kwon SH, Workman JL. 2011. HP1c casts light on dark matter. *Cell Cycle* **10**: 625–630.

Lachner M, O'Carroll D, Rea S, Mechtler K, Jenuwein T. 2001. Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* **410**: 116–120.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**: 357–360.

Levin HL, Moran J V. 2011. Dynamic interactions between transposable elements and their hosts. *Nat Rev Genet* **12**: 615–627.

Levis RW, Ganesan R, Houtchens K, Tolar LA, Sheen F miin. 1993. Transposons in place of telomeric repeats at a Drosophila telomere. *Cell* **75**: 1083–1093.

Lewis EB. 1950. The phenomenon of position effect. *Adv Genet* **3**: 73–115.

Lewis PW, Elsaesser SJ, Noh K, Stadler SC, Allis CD. 2010. Daxx is an H3.3-specific histone chaperone and cooperates with ATRX in replication-independent chromatin assembly at telomeres. *PNAS* **107**: 1–6.

Li W, Prazak L, Chatterjee N, Grüninger S, Krug L, Theodorou D, Dubnau J. 2013. Activation of transposable elements during aging and neuronal decline in Drosophila. *Nat Neurosci* **16**: 529–531.

Liang G, Lin JCY, Wei V, Yoo C, Cheng JC, Nguyen CT, Weisenberger DJ, Egger G, Takai D, Gonzales FA, et al. 2004. Distinct localization of histone H3 acetylation and H3-K4 methylation to the transcription start sites in the human genome. *Proc Natl Acad Sci* **101**: 7357–7362.

Lieberman-Aiden E, Berkum NL Van, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, et al. 2009. Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome. *Science (80-)* **33292**: 289–294.

Lifton RP, Goldberg ML, Karp RW, Hogness DS. 1978. The Organization of the Histone Genes in Drosophila melanogaster: Functional and Evolutionary Implications. *Cold Spring Harb Symp Quant Biol* **42**: 1047–1051.

Lilly MA, Duronio RJ. 2005. New insights into cell cycle control from the Drosophila endocycle. *Oncogene* **24**: 2765–2775.

Löb D, Lengert N, Chagin VO, Reinhart M, Casas-Delucchi CS, Cardoso MC, Drossel B. 2016. 3D replicon distributions arise from stochastic initiation and domino-like DNA replication progression. *Nat Commun* **7**: 11207.

Locke J, Kotarski MA, Tartoff KD. 1988. Dosage dependent modifiers of position effect variegation in Drosophila melanogaster and a mass action model that explains their effect. *Genetics* **120**: 181–198.

Love MI, Anders S, Huber W. 2014a. *Differential analysis of count data - the DESeq2 package*.

Love MI, Huber W, Anders S. 2014b. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 1–21.

Lu BY, Emtage PC, Duyf BJ, Hilliker a J, Eissenberg JC. 2000. Heterochromatin protein 1 is required for the normal expression of two heterochromatin genes in Drosophila. *Genetics* **155**: 699–708.

Lubelsky Y, Prinz J a, Denapoli L, Li Y, Belsky J a, Macalpine DM. 2014. DNA replication and transcription programs respond to the same chromatin cues DNA. 1102–1114.

Luger K, Mäder a W, Richmond RK, Sargent DF, Richmond TJ. 1997. Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* **389**: 251–260.

Lundberg LE, Stenberg P, Larsson J. 2013. HP1a, Su(var)3-9, SETDB1 and POF stimulate or repress gene expression depending on genomic position, gene length

and expression pattern in Drosophila melanogaster. *Nucleic Acids Res* **41**: 4481–94.

MacAlpine DM, Macalpine DM, Rodriguez HK, Bell SP, Bell SP. 2004. Coordination of replication and transcription along a. *Genes Dev* 3094–3105.

MacAlpine HK, Gordân R, Powell SK, Hartemink AJ, MacAlpine DM. 2010. Drosophila ORC localizes to open chromatin and marks sites of cohesin complex loading. *Genome Res* **20**: 201–211.

Makunin I V., Volkova EI, Belyaeva ES, Nabirochkina EN, Pirrotta V, Zhimulev IF. 2002. The Drosophila suppressor of underreplication protein binds to late-replicating regions of polytene chromosomes. *Genetics* **160**: 1023–1034.

Malik HS, Henikoff S. 2003. Phylogenomics of the nucleosome. *Nat Struct Biol* **10**: 882–91.

Mani SR, Juliano CE. 2013. Untangling the web: the diverse functions of the PIWI/piRNA pathway. *Mol Reprod Dev* **80**: 632–64.

Mantiero D, Mackenzie A, Donaldson A, Zegerman P. 2011. Limiting replication initiation factors execute the temporal programme of origin firing in budding yeast. *EMBO J* **30**: 4805–4814.

Martin MM, Ryan M, Kim R, Zakas AL, Fu H, Lin M, Reinhold WC, Davis SR, Bilke S, Liu H, et al. 2011. Genome-wide depletion of replication inititation events in highly transcribed regions. *Genome Res* **21**: 1822–1832.

Marzluff WF, Gongidi P, Woods KR, Jin J, Maltais LJ. 2002. The Human and Mouse Replication-Dependent Histone Genes. *Genomics* **80**: 487–498.

McClinktock B. 1951. Chromosome organization and genic expression. *Cold Spring Harb Symp Quant Biol* **16**: 13–47.

McKay DJ, Klusza S, Penke TJR, Meers MP, Curry KP, McDaniel SL, Malek PY, Cooper SW, Tatomer DC, Lieb JD, et al. 2015. Interrogating the Function of Metazoan Histones using Engineered Gene Clusters. *Dev Cell* **32**: 373–386.

McKay DJ, Lieb JD. 2013. A Common Set of DNA Regulatory Elements Shapes Drosophila Appendages. *Dev Cell* **27**: 306–318.

McKittrick E, Gafken PR, Ahmad K, Henikoff S. 2004. Histone H3.3 is enriched in covalent modifications associated with active chromatin. *Proc Natl Acad Sci U S A* **101**: 1525–30.

Mehrotra S, Maqbool SB, Kolpakas A, Murnen K, Calvi BR. 2008. Endocycling cells do not apoptose in response to DNA rereplication genotoxic stress. *Genes Dev* **22**: 3158–3171.

Mellone BG, Allshire RC. 2003. Stretching it: putting the CEN(P-A) in centromere. *Curr Opin Genet Dev* **13**: 191–198.

Mellone BG, Ball L, Suka N, Grunstein MR, Partridge JF, Allshire RC. 2003. Centromere Silencing and Function in Fission Yeast Is Governed by the Amino Terminus of Histone H3. *Curr Biol* **13**: 1748–1757.

Miotto B, Ji Z, Struhl K. 2016. Selectivity of ORC binding sites and the relation to replication timing, fragile sites, and deletions in cancers. *Proc Natl Acad Sci* **113**: E4810–E4819.

Mis J, Ner SS, Grigliatti T a. 2006. Identification of three histone methyltransferases in Drosophila: dG9a is a suppressor of PEV and is required for gene silencing. *Mol Genet Genomics* **275**: 513–26.

Mito Y, Henikoff JG, Henikoff S. 2005. Genome-scale profiling of histone H3.3 replacement patterns. *Nat Genet* **37**.

Mohn F, Sienski G, Handler D, Brennecke J. 2014. The rhino-deadlock-cutoff complex licenses noncanonical transcription of dual-strand piRNA clusters in Drosophila. *Cell* **157**: 1364–79.

Moldovan GL, Pfander B, Jentsch S. 2007. PCNA, the Maestro of the Replication Fork. *Cell* **129**: 665–679.

Moore G, Procunier J, Cross D, Grigliatti T. 1979. Histone gene deficiencies and position–effect variegation in Drosophila. *Nature* **282**: 312–4.

Mottus R, Reeves R, Grigliatti TA. 1980. Butyrate suppression of position-effect variegation in Drosophila Melanogaster. *Mol Gen Genet* **469**: 465–469.

Muller HJ. 1930. Types of visible variations induced by X-rays in Drosophila. *J Genet* **22**: 299–334.

Nakayama J, Rice JC, Strahl BD, Allis CD, Grewal SIS. 2001. Role of Histone H3 Lysine 9 Methylation in Epigenetic Control of Heterochromatin Assembly. *Science (80-)* **292**: 110–113.

Nordman JT, Kozhevnikova EN, Verrijzer CP, Pindyurin A V, Andreyeva EN, Shloma V V., Zhimulev IF, Orr-Weaver TL. 2014. DNA copy-number control through inhibition of replication fork progression. *Cell Rep* **9**: 841–849.

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.

Pal-Bhadra M, Leibovitch B a, Gandhi SG, Chikka MR, Rao M, Bhadra U, Birchler J a, Elgin SCR. 2004. Heterochromatic silencing and HP1 localization in Drosophila are

dependent on the RNAi machinery. *Science* **303**: 669–72.

Passarge E. 1979. Emil Heitz and the concept of heterochromatin: longitudinal chromosome differentiation was recognized fifty years ago. *Am J Hum Genet* **31**: 106–15.

Peng JC, Karpen GH. 2008. Epigenetic regulation of heterochromatic DNA stability. *Curr Opin Genet Dev* **18**: 204–11.

Peng JC, Karpen GH. 2009. Heterochromatic genome stability requires regulators of histone H3 K9 methylation. *PLoS Genet* **5**: e1000435.

Pengelly AR, Copur Ö, Jäckle H, Herzig A, Müller J. 2013. A histone mutant reproduces the phenotype caused by loss of histone-modifying factor Polycomb. *Science* **339**: 698–9.

Penke TJR, McKay DJ, Strahl BD, Gregory Matera A, Duronio RJ. 2016. Direct interrogation of the role of H3K9 in metazoan heterochromatin function. *Genes Dev* **30**: 1866–1880.

Pérez-Lluch S, Blanco E, Tilgner H, Curado J, Ruiz-Romero M, Corominas M, Guigó R. 2015. Absence of canonical marks of active chromatin in developmentally regulated genes. *Nat Genet* **47**: 1158–67.

Petryk N, Kahli M, d'Aubenton-Carafa Y, Jaszczyszyn Y, Shen Y, Silvain M, Thermes C, Chen C-L, Hyrien O. 2016. Replication landscape of the human genome. *Nat Commun* **7**: 10208.

Piacentini L, Fanti L, Berloco M, Perrini B, Pimpinelli S. 2003. Heterochromatin protein 1 (HP1) is associated with induced gene expression in Drosophila euchromatin. *J Cell Biol* **161**: 707–714.

Piacentini L, Fanti L, Negri R, Del Vescovo V, Fatica A, Altieri F, Pimpinelli S. 2009. Heterochromatin protein 1 (HP1a) positively regulates euchromatic gene expression through RNA transcript association and interaction with hnRNPs in Drosophila. *PLoS Genet* **5**.

Pindyurin A V, Boldyreva L V, Shloma V V, Kolesnikova TD, Pokholkova G V, Andreyeva EN, Kozhevnikova EN, Ivanoschuk IG, Zarutskaya E a, Demakov S a, et al. 2008. Interaction between the Drosophila heterochromatin proteins SUUR and HP1. *J Cell Sci* **121**: 1693–1703.

Pokholkova G V., Koryakov DE, Pindyurin A V., Kozhevnikova EN, Belyakin SN, Andreyenkov O V., Belyaeva ES, Zhimulev IF. 2015. Tethering of SUUR and HP1 proteins results in delayed replication of euchromatic regions in Drosophila melanogaster polytene chromosomes. *Chromosoma* **124**: 209–220.

Pope BD, Ryba T, Dileep V, Yue F, Wu W, Denas O, Vera DL, Wang Y, Hansen RS,

Canfield TK, et al. 2014. Topologically associating domains are stable units of replication-timing regulation. *Nature* **515**: 402–405.

Prescott DM, Bender MA. 1962. Synthesis of RNA and protein during mitosis in mammalian tissue culture cells. *Exp Cell Res* **26**: 260–268.

Price BD, Andrea ADD. 2014. Chromatin Remodeling at DNA Double Strand Breaks. *Cell* **152**: 1344–1354.

Quinlan AR, Hall IM. 2010. BEDTools : a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**: 841–842.

Quivy J-P, Gérard A, Cook AJL, Roche D, Almouzni G. 2008. The HP1–p150/CAF-1 interaction is required for pericentric heterochromatin replication and S-phase progression in mouse cells. *Nat Struct Mol Biol* **15**: 972–979.

Raffa GD, Ciapponi L, Cenci G, Gatti M. 2011. Terminin: A protein complex that mediates epigenetic maintenance of Drosophila telomeres. *Nucleus* **2**: 383–391.

Rahman R, Chirn G-W, Kanodia A, Sytnikova YA, Brembs B, Bergman CM, Lau NC. 2015. Unique transposon landscapes are pervasive across Drosophila melanogaster genomes. *Nucleic Acids Res* **43**: 10655–72.

Rangan P, Malone CD, Navarro C, Newbold SP, Hayes PS, Sachidanandam R, Hannon GJ, Lehmann R. 2011. piRNA production requires heterochromatin formation in Drosophila. *Curr Biol* **21**: 1373–9.

Rao SSP, Huang SC, Glenn St Hilaire B, Engreitz JM, Perez EM, Kieffer-Kwon KR, Sanborn AL, Johnstone SE, Bascom GD, Bochkov ID, et al. 2017. Cohesin Loss Eliminates All Loop Domains. *Cell* **171**: 305–320.e24.

Rea S, Eisenhaber F, O'Carroll D, Strahl BD, Sun ZW, Schmid M, Opravil S, Mechtler K, Ponting CP, Allis CD, et al. 2000. Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* **406**: 593–9.

Reuter G, Wolff I. 1981. Isolation of dominant suppressor mutations for position-effect variegation in Drosophila melanogaster. *MGG Mol Gen Genet* **182**: 516–519.

Rhind N, Gilbert DM. 2013. DNA Replication Timing. *Cold Spring Harb Perspect Med* **3**: 1–26.

Rhind N, Yang SCH, Bechhoefer J. 2010. Reconciling stochastic origin firing with defined replication timing. *Chromosom Res* **18**: 35–43.

Riddle NC, Minoda a., Kharchenko PV, Alekseyenko a. a., Schwartz YB, Tolstorukov MY, Gorchakov a. a., Jaffe JD, Kennedy C, Linder-Basso D, et al. 2011. Plasticity in patterns of histone modifications and chromosomal proteins in *Drosophila* heterochromatin. *Genome Res* **21**: 147–163.

Rivera-Mulia JC, Gilbert DM. 2016. Replication timing and transcriptional control: beyond cause and effect- part III. *Curr Opin Cell Biol* **40**: 168–178.

Roberts PA. 1965. Difference in the behavior of eu- and hetero-chromatin: crossing over. *Nature* **205**: 725–726.

Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander E, Getz G, Mesirov J. 2011. Integrative genomics viewer. *Nat Biotechnol* **29**: 24–26.

Robinson MD, McCarthy DJ, Smyth GK. 2009. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**: 139–140.

Roh T, Cuddapah S, Zhao K. 2005. Active chromatin domains are defined by acetylation islands revealed by genome-wide mapping -- Roh et al. 19 (5): 542 -- Genes and Development. 542–552.

Rothbart SB, Strahl BD. 2014. Interpreting the language of histone and DNA modifications. *Biochim Biophys Acta* 1–17.

Roy S, Ernst J, Kharchenko P V., Kheradpour P, Negre N, Eaton ML, Landolin JM, Bristow CA, Ma L, Lin MF, et al. 2010. Identification of Functional Elements and Regulatory Circuits by Drosophila modENCODE. *Science (80-)* **330**: 1787–1797.

Ryba T, Battaglia D, Chang BH, Shirley JW, Buckley Q, Pope BD, Devidas M, Druker BJ, Gilbert DM. 2012. Abnormal developmental control of replication-timing domains in pediatric acute lymphoblastic leukemia. 1833–1844.

Ryder E, Blows F, Ashburner M, Bautista-Llacer R, Coulson D, Drummond J, Webster J, Gubb D, Gunton N, Johnson G, et al. 2004. The DrosDel collection: A set of P-element insertions for generating custom chromosomal aberrations in Drosophila melanogaster. *Genetics* **167**: 797–813.

Saito K, Siomi MC. 2010. Small RNA-Mediated Quiescence of Transposable Elements in Animals. *Dev Cell* **19**: 687–697.

Sakai A, Schwartz BE, Goldstein S, Ahmad K. 2009. Transcriptional and developmental functions of the H3.3 histone variant in Drosophila. *Curr Biol* **19**: 1816–20.

Sasaki T, Rivera-Mulia JC, Vera D, Zimmerman J, Das S, Padget M, Nakamichi N, Chang BH, Tyner J, Druker BJ, et al. 2017. Stability of patient-specific features of altered DNA replication timing in xenografts of primary human acute lymphoblastic leukemia. *Exp Hematol* **51**: 71–82.

Schotta G, Ebert A, Dorn R, Reuter G. 2003a. Position-effect variegation and the genetic dissection of chromatin regulation in Drosophila. *Semin Cell Dev Biol* **14**: 67–75.

Schotta G, Ebert A, Krauss V, Fischer A, Hoffmann J, Rea S, Jenuwein T, Dorn R, Reuter G. 2002. Central role of Drosophila SU(VAR)3-9 in histone H3-K9 methylation and heterochromatic gene silencing. *EMBO J* **21**: 1121–31.

Schotta G, Ebert A, Reuter G. 2003b. SU(VAR)3-9 is a conserved key function in heterochromatic gene silencing. *Genetica* **117**: 149–158.

Schotta G, Lachner M, Sarma K, Ebert A, Sengupta R, Reuter G, Reinberg D, Jenuwein T. 2004. A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. *Genes Dev* **18**: 1251–62.

Schwaiger M, Kohler H, Oakeley EJ, Stadler MB, Schübeler D. 2010. Heterochromatin protein 1 (HP1) modulates replication timing of the Drosophila genome. *Genome Res* **20**: 771–80.

Schwartz BE, Ahmad K. 2005. Transcriptional activation triggers deposition and removal of the histone variant H3 . 3. *Genes Dev* 804–814.

Scully R, Xie A. 2013. Double strand break repair functions of histone H2AX. *Mutat Res* **750**: 5–14.

Senti K-A, Brennecke J. 2010. The piRNA pathway: a fly's perspective on the guardian of the genome. *Trends Genet* **26**: 499–509.

Seum C, Reo E, Peng H, Rauscher FJ, Spierer P, Bontron S. 2007. Drosophila SETDB1 is required for chromosome 4 silencing. *PLoS Genet* **3**.

Sher N, Bell GW, Li S, Nordman J, Eng T, Eaton ML, Macalpine DM, Orr-weaver TL. 2012. Developmental control of gene copy number by repression of replication initiation and fork progression. *Genome Res* 64–75.

Siefert JC, Georgescu C, Wren JD, Koren A, Sansam CL. 2017. DNA replication timing during development anticipates transcriptional programs and parallels enhancer activation. *Genome Res* **27**: 1406–1416.

Simon JM, Giresi PG, Davis IJ, Lieb JD. 2013. A detailed protocol for formaldehyde-assisted isolation of regulatory elements (FAIRE). *Curr Protoc Mol Biol* **Chapter 21**: Unit21.26.

Sims RJ, Reinberg D. 2008. Is there a code embedded in proteins that is based on post-translational modifications? *Nat Rev Mol Cell Biol* **9**: 815–20.

Sinclair DAR, Mottus RC, Grigliatti TA. 1983. Genes which suppress position-effect variegation in Drosophila melanogaster are clustered. *MGG Mol Gen Genet* **191**: 326–333.

Skene PJ, Henikoff S. 2017. An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. *Elife* **6**: 1–35.

Slotkin RK, Martienssen R. 2007. Transposable elements and the epigenetic regulation of the genome. *Nat Rev Genet* **8**: 272–85.

Smit A, Hubley R, Green P. 2013. RepeatMasker Open-3.0.

Smith L, Plug A, Thayer M. 2001. Delayed replication timing leads to delayed mitotic chromosome condensation and chromosomal instability of chromosome translocations. *Proc Natl Acad Sci U S A* **98**: 13300–5.

Soruco MML, Chery J, Bishop EP, Siggers T, Tolstorukov MY, Leydon AR, Sugden AU, Goebel K, Feng J, Xia P, et al. 2013. The CLAMP protein links the MSL complex to the X chromosome during Drosophila dosage compensation. *Genes Dev* **27**: 1551–1556.

Spange S, Wagner T, Heinzel T, Krämer OH. 2009. Acetylation of non-histone proteins modulates cellular signalling at multiple levels. *Int J Biochem Cell Biol* **41**: 185–198.

Stamatoyannopoulos JA, Adzhubei I, Thurman RE, Kryukov G V, Mirkin SM, Sunyaev SR. 2009. Human mutation rate associated with DNA replication timing. *Nat Genet* **41**: 393–395.

Strahl B, Allis C. 2000. The language of covalent histone modifications. *Nature* **403**: 41–45.

Strom AR, Emelyanov A V., Mir M, Fyodorov D V., Darzacq X, Karpen GH. 2017. Phase separation drives heterochromatin domain formation. *Nature* **547**: 241–245.

Sun Y, Wyatt RT, Bigley A, Krontiris TG. 2001. Expression and replication timing patterns of wildtype and translocated BCL2 genes. *Genomics* **73**: 161–170.

Swanson C. 1957. *Cytology and cytogenetics*.

Szenker E, Ray-Gallet D, Almouzni G. 2011. The double face of the histone variant H3.3. *Cell Res* **21**: 421–434.

Tagami H, Ray-Gallet D, Almouzni G, Nakatani Y. 2004. Histone H3.1 and H3.3 Complexes Mediate Nucleosome Assembly Pathways Dependent or Independent of DNA Synthesis. *Cell* **116**: 51–61.

Talbert PB, Henikoff S. 2010. Histone variants--ancient wrap artists of the epigenome. *Nat Rev Mol Cell Biol* **11**: 264–275.

Talbert PB, Henikoff S. 2017. Histone variants on the move: substrates for chromatin dynamics. *Nat Rev Mol Cell Biol* **18**: 115–126.

Tamura T, Smith M, Kanno T, Dasenbrock H, Nishiyama A, Ozato K. 2009. Inducible deposition of the histone variant H3.3 in interferon-stimulated genes. *J Biol Chem* **284**: 12217–12225.

Taylor JH. 1960a. Asynchronous Duplication of Chromosomes in Cultured Cells of Chinese Hamster. *J Biophys Biochem Cytol* **7**: 455–463.

Taylor JH. 1960b. Nucleic Acid Synthesis in Relation To the Cell Division Cycle. *Ann N Y Acad Sci* **90**: 409–421.

Taylor JH. 1958. The mode of chromosome duplication in Crepis capillaris. *Exp Cell Res* **15**: 350–357.

Towbin BD, González-Aguilera C, Sack R, Gaidatzis D, Kalck V, Meister P, Askjaer P, Gasser SM. 2012. Step-wise methylation of histone H3K9 positions heterochromatin at the nuclear periphery. *Cell* **150**: 934–47.

Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L. 2014. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc*.

Tschiersch B, Hofmann A, Krauss V, Dorn R, Korgel G, Reuter G. 1994. The protein encoded by the Drosophila position- combines domains of antagonistic regulators of homeotic gene complexes. **13**: 3822–3831.

Udugama M, Chang FTM, Chan FL, Tang MC, Pickett H a, McGhie JDR, Mayne L, Collas P, Mann JR, Wong LH. 2015. Histone variant H3.3 provides the heterochromatic H3 lysine 9 tri-methylation mark at telomeres. *Nucleic Acids Res* gkv847-.

Vakoc CR, Mandat S a, Olenchock B a, Blobel G a. 2005. Histone H3 lysine 9 methylation and HP1gamma are associated with transcription elongation through mammalian chromatin. *Mol Cell* **19**: 381–91.

Vedelek B, Blastyák A, Boros IM. 2015. Cross-species interaction between rapidly evolving telomere-specific drosophila proteins. *PLoS One* **10**: 1–16.

Verreault A, Kaufman PD, Kobayashi R, Stillman B. 1996. Nucleosome assembly by a complex of CAF-1 and acetylated histones H3/H4. *Cell* **87**: 95–104.

Walen KH. 1964. Somatic Crossing Over in Relationship To Heterochromatin in Drosophila. *Genetics* **49**: 905–923.

Wallrath LL, Elgin SC. 1995. Position effect variegation in Drosophila is associated with an altered chromatin structure. *Genes Dev* **9**: 1263–1277.

Wallrath LL, Vitalini MW, Elgin SCR. 2014. Fundamentals of Chromatin eds. J.L. Workman and S.M. Abmayr. 529–552.

Wang L, Liu L, Berger SL. 1998. Critical residues for histone acetylation by Gcn5, functioning in Ada and SAGA complexes, are also required for transcriptional function in vivo. *Genes Dev* 640–653.

Wang SH, Nan R, Accardo MC, Sentmanat M, Dimitri P, Elgin SCR. 2014. A Distinct Type of Heterochromatin at the Telomeric Region of the Drosophila melanogaster Y Chromosome. *PLoS One* **9**.

Wang Z, Zang C, Rosenfeld JA, Schones DE, Barski A, Cuddapah S, Cui K, Roh T-Y, Peng W, Zhang MQ, et al. 2008. Combinatorial patterns of histone acetylations and methylations in the human genome. *Nat Genet* **40**: 897–903.

Wit E De, Laat W De. 2012. A decade of 3C technologies : insights into nuclear organization. *Genes Dev* **26**: 11–24.

Wong LH, Mcghie JD, Sim M, Anderson MA, Ahn S, Hannan RD, George AJ, Morgan KA, Mann JR, Choo KHA. 2010. ATRX interacts with H3.3 in maintaining telomere structural integrity in pluripotent embryonic stem cells. *Genome Res* 351–360.

Xu M, Long C, Chen X, Huang C, Chen S, Zhu B. 2010. Partitioning of Histone H3-H4 Tetramers During DNA Replication-Dependent Chromatin Assembly. *Science (80-)* **328**: 94–98.

Xu T, Rubin GM. 1993. Analysis of genetic mosaics in developing and adult Drosophila tissues. *Development* **117**: 1223–1237.

Yamazaki S, Ishii A, Kanoh Y, Oda M, Nishito Y, Masai H. 2012. Rif1 regulates the replication timing domains on the human genome. *EMBO J* **31**: 3667–3677.

Yarosh W, Spradling AC. 2014. Incomplete replication generates somatic DNA alterations within Drosophila polytene salivary gland cells. *Genes Dev* 1840–1855.

Yasuhara JC, Wakimoto BT. 2006. Oxymoron no more: the expanding world of heterochromatic genes. *Trends Genet* **22**: 330–338.

Yoshida K, Bacal J, Desmarais D, Padioleau I, Tsaponina O, Chabes A, Pantesco V, Dubois E, Parrinello H, Skrzypczak M, et al. 2014. The Histone Deacetylases Sir2 and Rpd3 Act on Ribosomal DNA to Control the Replication Program in Budding Yeast. *Mol Cell* **54**: 691–697.

Yuan K, Farrell PHO. 2016. TALE-light imaging reveals maternally emergence of functional heterochromatin in Drosophila embryos. *Genes Dev* **30**: 1–15.

Zeller P, Padeken J, Schendel R Van, Kalck V, Tijsterman M, Gasser SM. 2016. Histone H3K9 methylation is dispensable for Caenorhabditis elegans development but suppresses RNA : DNA hybrid-associated repeat instability. *Nat Genet* 1–13.

Zhang K, Fischer T, Porter RL, Dhakshnamoorthy J, Zofall M, Zhou M, Veenstra T, Grewal SIS. 2011. Clr4/Suv39 and RNA quality control factors cooperate to trigger RNAi and suppress antisense RNA. *Science* **331**: 1624–7.

Zhang X, Huang Y, Shi X. 2015. Emerging roles of lysine methylation on non-histone

proteins. *Cell Mol Life Sci*.

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, et al. 2008. Open Access Model-based Analysis of ChIP-Seq (MACS). *Genome Biol* **9**.

Zhimulev IF, Belyaeva ES, Makunin I V, Pirrotta V, Volkova EI, Alekseyenko a a, Andreyeva EN, Makarevich GF, Boldyreva L V, Nanayev R a, et al. 2003a. Influence of the SuUR gene on intercalary heterochromatin in Drosophila melanogaster polytene chromosomes. *Chromosoma* **111**: 377–398.

Zhimulev IF, Belyaeva ES, Semeshin VF, Shloma V V, Makunin I V, Volkova EI. 2003b. Overexpression of the SuUR gene induces reversible modifications at pericentric, telomeric and intercalary heterochromatin of Drosophila melanogaster polytene chromosomes. *J Cell Sci* **116**: 169–76.