I

# THE ROLES OF SET-9 AND SET-26 IN LONGEVITY, GERMLINE FUNCTION AND RNAI PATHWAY

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Wenke Wang

May 2018

# THE ROLES OF SET-9 AND SET-26 IN LONGEVITY, GERMLINE FUNCTION AND RNAI PATHWAY

Wenke Wang, Ph. D.

Cornell University 2018

Aging, germline function and RNA interference are three distinct biological processes that associate with alterations in epigenetic state, including histone modifications. Our lab discovered two highly homologous SET domain containing genes, *set-9 and set-26*, in an RNAi screen of putative histone methyltransferases that regulate lifespan in *C. elegans*. Inactivating *set-9*/*set-26* consistently extends lifespan. However, both the molecular functions of SET-9 and SET-26 and whether they are also involved in other biological processes are largely unknown.

My work found that SET-26 but not SET-9 plays an important role in lifespan regulation. Endogenous SET-26 is broadly expressed while SET-9 is only detectable in the germline. Specifically, somatically expressed SET-26 is implicated in lifespan determination. Transcriptional profile revealed that the somatic SET-26 regulated DAF-16-dependent genes are enriched for lifespan determinant genes, which likely explain the long-lived lifespan phenotype of *set-26* mutant. Moreover, SET-9 and SET-26 act redundantly to regulate germline function. Molecular analyses showed that SET-9 and SET-26 bind to H3K4me3 both *in vitro* and *in vivo*. Germline-expressed SET-9 and SET-26 normally bind to and restrict H3K4me3. Loss of SET-9 and SET-26 causes spreading of this mark with a bias towards 3' end. The spreading

of H3K4me3 correlates with elevated expression of a group of germline specific genes, which could explain the fertility defect in the *set-9 set-26* double mutant. In addition to longevity and germline function, I found that SET-26 is also involved in RNAi-mediated gene silencing. SET-26 is required for dsRNA-induced H3K9me3 but not H3K27me3 deposition.

SET-9 and SET-26 are putative homologs of the human protein Mixed Lineage Leukemia 5 (hMLL5), which plays an important role in regulating hematopoietic homeostasis, cell cycle and survival. However, the molecular basis underlying hMLL5 activity is still unknown. Our findings on SET-9 and SET-26 will help us better understand the role of hMLL5 in human and also implicated hMLL5 in germline function and longevity.

# BIOGRAPHICAL SKETCH

Wenke Wang was born in Xiangtan, a small city of Hunan Province in China in 1990. He is the only child of his family and his parents put high expectation on him. He had a happy childhood in the peaceful small city and at age 12 he was sent to a boarding school where he starts to learn Science Olympiad. This was an important experience for him as it not only helped him develop effective study habits, but also give him the opportunity to gain admission to Tsinghua University, which is one of the top two in China, without needing to take the College Entrance Examination. His passion for science and biology starts there. During his time in Tsinghua University, he complete ohis undergraduate research at Guo lab studying stem cell biology. He received his BA degree in 2012 and decided to go to United States for pursuing a PhD. Degree right after that. Luckily, he entered the Graduate Field of Biochemistry, Molecular, and Cell Biology PhD. Program at Cornell University where he spent 5 years to investigate the genetic and molecular mechanisms of aging in Professor Siu Sylvia Lee's lab where he not only learned experimental skill and critical thinking ability but also developed interests in aging research. He decided to continue his postdoctoral training at the Buck Institute of aging in the Professor Garrison and Brem's lab.

ACKNOWLEDGMENTS

Firstly, I would like to express my sincere gratitude to my advisor Professor Sylvia Siu Lee for her continuous support of my Ph.D. study. She not only provided such a good environment for me to learn experimental skills, critical thinking ability, communication skills, and experimental designs and so on, but also helped me a lot with my personal growth. During my graduate study, there are countless difficulties that I can never overcome without her help. Her guidance also helped me a lot in all the time of research and writing this thesis. I could not have imagined to complete my Ph.D. study without her help.

Besides my advisor, I would also like to thank other members of my thesis committee: Professor Paul Soloway and Professor Charles Danko, for their insightful comments, suggestions and encouragement, but also for the criticisms and hard question in my research which incented me to widen my research from different perspectives.

I am grateful to the MBG department for providing such a welcoming and supportive environment for all the researchers. I thank Cornell Biotech Resource Center for providing a great high through put sequencing service. I also would like to thank all the members of Cornell worm group for their helpful suggestions and discussions.

I also would like to thank all the past and present lab members. I thank Dr. Zoey Ni for starting the SET-9/26 project and her work on SET-26 provided a good platform form me to start my projects. I thank Dr. Amaresh Chaturbedi for his contribution of germline characterization in my project. I thank Dr. Minghui Wang for his help in statistical analysis. I thank Dr. Satheeja Velayudhan for her validation of my RNA-seq results. I thank Serim An for her contribution to this project. I am also grateful for insightful suggestions and discussions by Dr. Veerle Rottiers, Dr. Mintie Pu, Dr. Cheng-Lin Li, Dr.

Ella Chang, Dr. Erin Jones and Seth Sagulo. Finally, I am grateful to Rada Omanovic who provides supports for the reagents and plates for my experiments.

Last but not the least, I would like to thank my family: my father Wang Shuhuai and my mother Xie Xiangmei for their endless support. Without their love and support, I could never finished my Ph.D.

# TABLE OF CONTENTS

# CHAPTER 1

# INTRODUCTION

*Caenorhabditis elegans* (*C. elegans*) is a free-living soil nematode. It has a small size, short life cycle (3-day) and can be easily handled in the laboratory. It feeds on bacteria and around 300 progenies are produced by a single self-fertilizing hermaphrodite. The male can be produced under certain conditions and fertilize the hermaphrodite, which offers convenience in genetic studies. Through different mutagenesis projects, genetic mutants for almost every gene were generated for *C. elegans*. *C. elegans* shares a common ancestor with humans. Many genes and molecular mechanisms, including those in humans, were present in this ancestor and therefore are conserved between *C. elegans* and humans. Understanding these genes and molecular mechanisms in *C. elegans* will help us better understand those in humans. All of these make *C. elegans* a great model for studying various biological processes including longevity, germline function and RNAi process.

SET-9 and SET-26 were discovered in our previous screen as two longevity determining factors in *C. elegans* (Ni et al. 2012). They are putative orthologs of the human protein Mixed Lineage Leukemia 5 (hMLL5) which plays important roles in regulating hematopoietic homeostasis, the cell cycle and survival (Zhang et al. 2017). Considering MLL5 is implicated in leukemia and glioblastoma (Zhang et al. 2017), understanding the functions and biological roles of SET-9 and SET-26 in *C. elegans* will also help us to better understand the functions and biological roles of MLL5, and may pave the way for potential future therapeutic development.

## 1.1 Longevity and germline in *C. elegans*

### Longevity

Aging is a highly conserved process and happens in nearly all organisms. Because of its short life cycle and ease of genetic manipulation (Figure 1.1), *C. elegans* is a widely used model for aging research. Several lifespan regulation pathways have been found in *C. elegans* such as Insulin/IGF-1 like signaling, the TOR pathway,

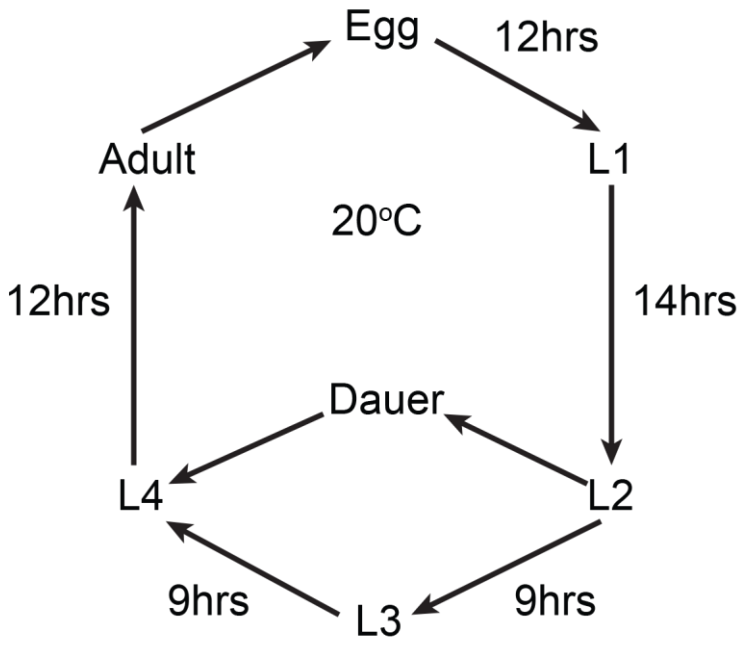



Figure 1.1.1 *C. elegans* life cycle

mitochondrial dysfunction and reproductive signals (Kenyon 2010). I will briefly discuss these pathways here.

## Insulin/IGF-1 like signaling pathway

Insulin/IGF-1 like signaling pathway is the first pathway that was found to influence aging in animals (Kenyon 2005). The insulin/IGF-1 like signaling pathway starts with the binding of insulin/IGF-1 with its receptor, the active form of which in turn initiates a PI(3)K/AKT/PDK kinase cascade. The activated kinase phosphorylates DAF-16/FOXO causing cytoplasmic retention and inhibition of DAF-16/FOXO, which leads to deregulation of various downstream genes (Figure 1.1.2). Perturbation in this pathway activates DAF-16/FOXO by increasing nuclear portion of DAF-16 and therefore affect expression of downstream genes (Kenyon 2005). In *C. elegans*, disfunctions in insulin/IGF-1 receptor DAF-2 and components of the PI(3)K)/AKT/PDK kinase cascade extend lifespan of the animals (Kenyon 2005). The long-lived phenotype of these mutants depends on the transcription factor DAF-16/FOXO (Kenyon et al. 1993), which regulates various genes that act cumulatively to influence the aging process of animals (Kenyon et al. 1993). The effect of insulin/IGF-1 like signaling pathway on lifespan is evolutionarily conserved (Kenyon 2005). In *Drosophila*, inhibiting insulin/IGF-1 signaling pathway or activating FOXO specifically in fat body extends lifespan of the animals (Hwangbo et al. 2004). In mice, fat-specific insulin receptor knockout mutant lives longer than normal mice (Bluher et al. 2003). In addition to insulin/IGF-1 receptor - PI(3)K)/AKT/PDK kinase cascade pathway, the transcriptional factor DAF-16 is regulated by diverse other factors and mediates longer lifespan in multiple mutants

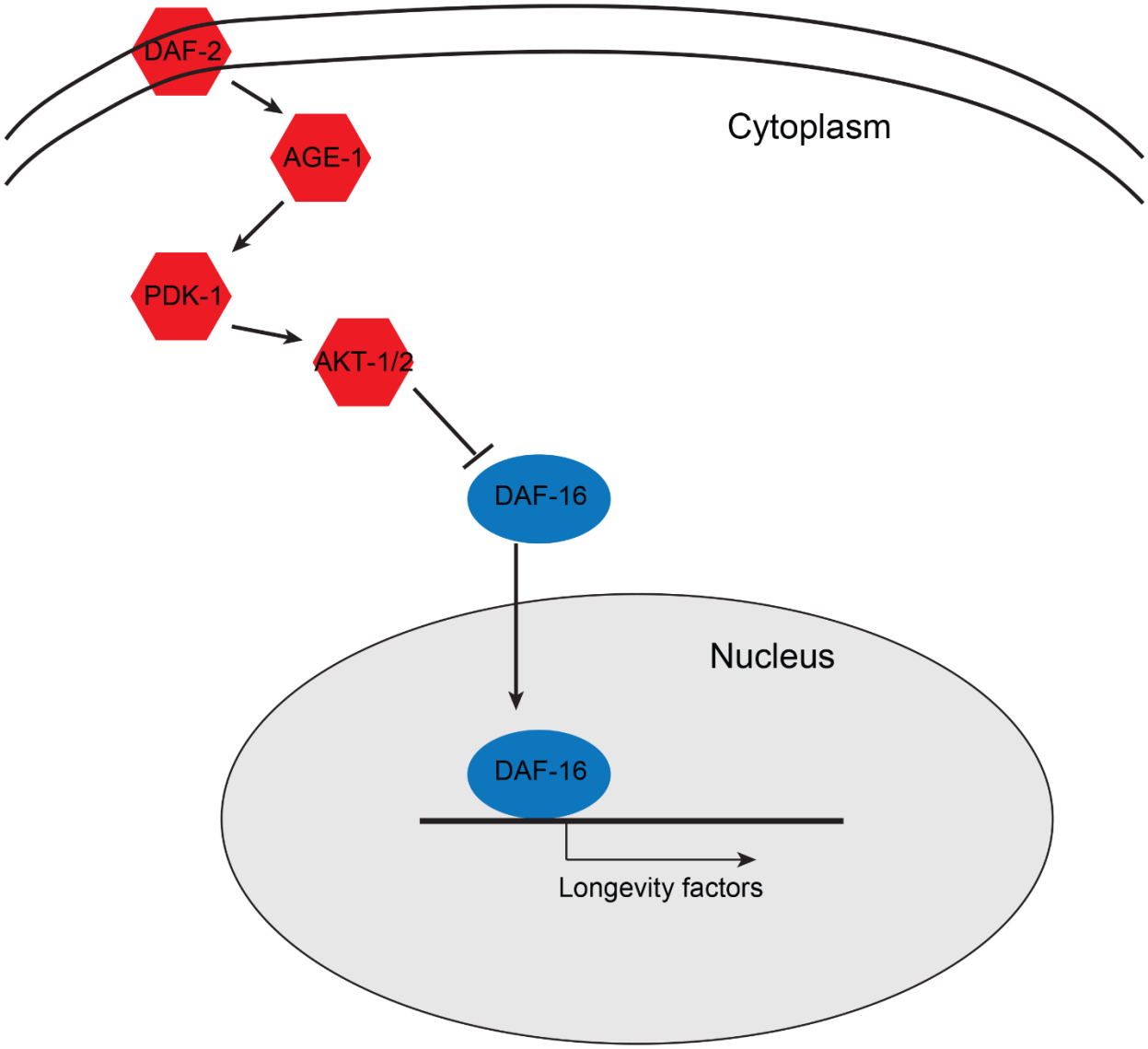



Figure 1.1.2 Insulin/IGF-1 like signaling pathway in *C. elegans*

(Hesp et al. 2015). For example, inactivating *hcf-1*, host cell factor 1 (HCF-1) homolog in *C. elegans*, extends lifespan of *C. elegans* in a daf-16-dependent manner (Li et al. 2008; Rizki et al. 2011). HCF-1 limits DAF-16 from accessing its target gene promoters by forming a complex with DAF-16 and thereby regulates gene expression (Li et al. 2008). Overall, the Insulin/IGF-1 like signaling pathway plays an important role in lifespan regulation.

## TOR pathway

The target of rapamycin (TOR) is a conserved kinase that involved in various biological processes (Wullschleger et al. 2006). It is inhibited by rapamycin and mainly acts as a mediator that senses IGF-1/insulin signal, energy status and nutrient (amino acid), and in turn to regulate downstream factors including S6K and 4E-BP to affect cell metabolism and growth (Wullschleger et al. 2006).

The TOR pathway is also important for lifespan regulation. Inhibiting *let-363*, the TOR homolog in *C. elegans*, extends lifespan in *C. elegans* (Vellai et al. 2003). The effect of TOR pathway in lifespan is evolutionarily conserved as repressing mTOR in mice by feeding rapamycin extends mice lifespan (Harrison et al. 2009). Unlike IGF-1/insulin signaling pathway, at least in *C. elegans*, the long-lived phenotype is independent of DAF-16/FOXO, suggesting that the TOR pathway acts as a novel pathway to affect the aging process (Vellai et al. 2003).

Activated IGF-1/insulin signaling pathway, high energy status and elevated levels of amino acid can activate TOR signaling pathway, which in turn regulates two well

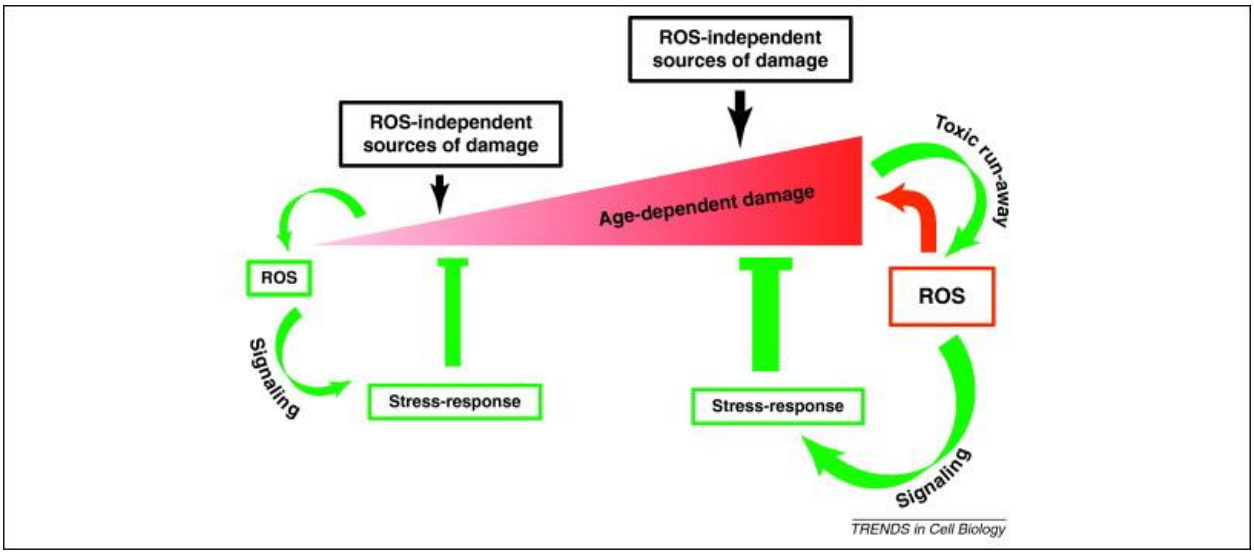
studied downstream factors S6K and 4E-BP by phosphorylation of the proteins (Evans et al. 2011). Phosphorylation of the ribosomal subunit S6K activates this protein (Magnuson et al. 2012), whereas phosphorylation of the translation inhibitor 4E-BP inactivates this protein (Musa et al. 2016), both promote protein translation in general. Repressing translation has been linked to longer lifespan as RNAi screens for lifespan factors in *C. elegans* have revealed that translational factors when knocked down can extend lifespan of the animals (Hamilton et al. 2005) and translation factor mutants show extended lifespan (Cattie et al. 2016). These results suggest that inhibiting TOR pathway extends lifespan at least partially through repressing translation.

TOR pathway is also a negative regulator of autophagy (Jung et al. 2010). Nutrient starvation, low growth factor and stress signals induce autophagy by inhibiting TOR pathway (Jung et al. 2010). In *C. elegans*, repressing genes required for autophagy suppresses lifespan extension in both dietary restriction mutant and TOR inhibiting worms (Hansen et al. 2008). These results suggest that autophagy is important for longer lifespan in those worms with TOR pathway repression.

## Mitochondrial dysfunction

Mitochondria is an intracellular organelle that is essential for many biological processes. The primary role of mitochondria is to produce ATP through oxidative phosphorylation. The electrons from NADH and FADH2 pass through electron transport chain to oxygen and produce proton gradient to drive the ATP production by ATP synthase (Friedman and Nunnari 2014). Defects in mitochondria will trigger

overproduction of reactive oxygen species (ROS) which is toxic for organisms (Murphy 2009). Interestingly, in *C. elegans*, mitochondria plays an essential role in aging as many mutants with defects in electron transport chain and ATP synthase show longer lifespan (Hansen et al. 2005; Feng et al. 2001; Lee et al. 2003). The effect of mitochondria in the aging process is evolutionarily conserved as inhibiting components of electron transfer chain in *Drosophila* and mice also extend lifespan of the animals (Copeland et al. 2009; Liu et al. 2005). Many models of aging have been proposed and the ROS theory is the most intriguing one among them. The mitochondrial free radical theory of aging proposes that ROS generated from mitochondria is the main cause of aging (HARMAN 1956). According to this theory, ROS is considered as an unwanted toxic byproducts of mitochondrial electron transfer that induce damage to cells due to its high chemical reactivity. Old animals with decline in mitochondrial function have increased ROS production which leads to aging of the animals (Bratic and Larsson 2013). However, ROS can be both beneficial and detrimental to lifespan at least in *C. elegans* (Lee et al. 2010; Yang and Hekimi 2010). The long-lived mitochondrial mutants contain higher levels of ROS and repressing ROS production by antioxidant treatment abolishes the longer lifespan suggesting the requirement of elevated ROS levels for the long lifespan (Lee et al. 2010; Yang and Hekimi 2010). They proposed a model that there is a threshold ROS level in lifespan determination. Animals with increased but below threshold ROS level exhibit longer lifespan. However, once the ROS level pass the threshold, the toxic ROS induce damage to the cell and cause shorter lifespan (Van Raamsdonk and Hekimi 2009; Lee et al. 2010; Yang and Hekimi 2010) (Figure 1.1.3). It will be interesting to know if the ROS threshold model also exists in other organisms.




*Hekimi S, Trends Cell Biol., 2011*

Figure 1.1.3 The ROS threshold model

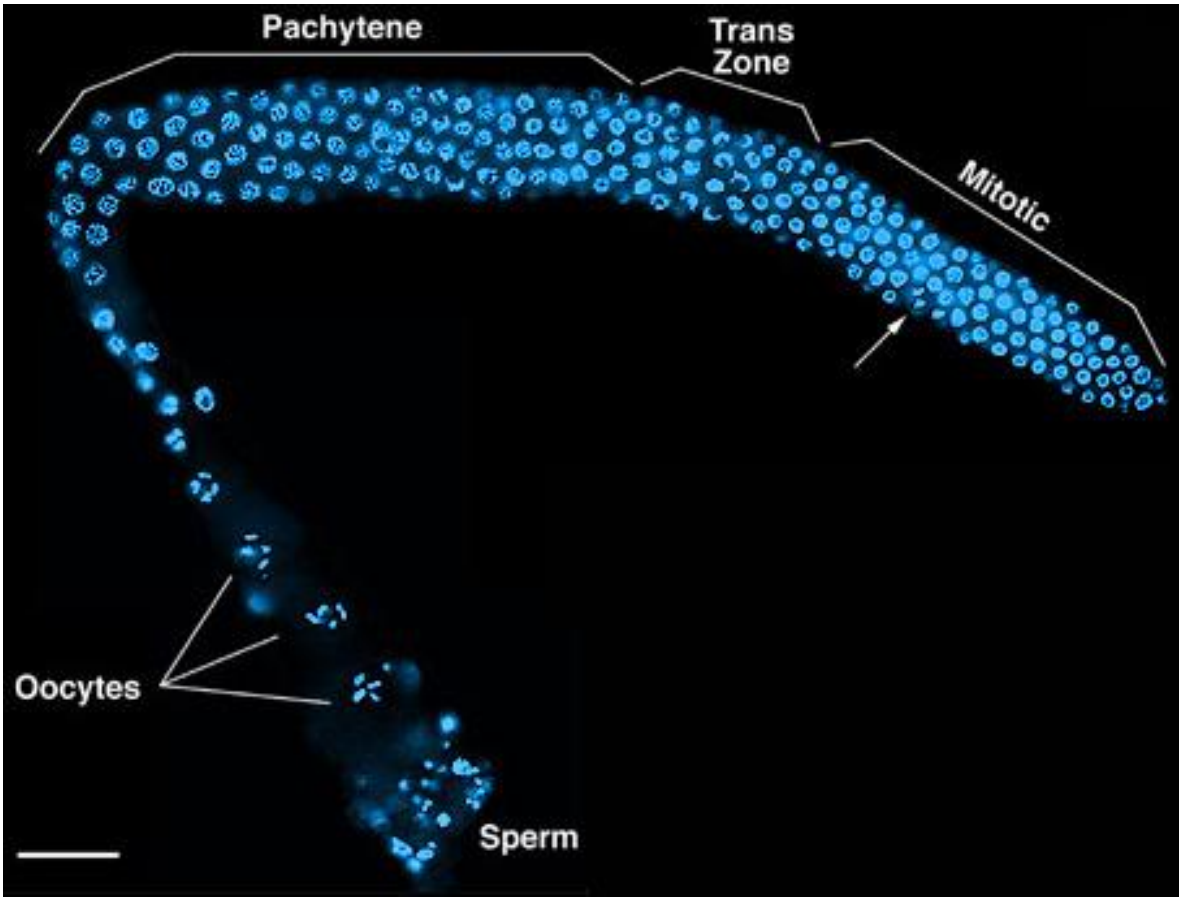
## Reproductive signals

*C. elegans* gonad, consisting of germ cells and somatic gonad, is not only important for reproduction but also for longevity. When it hatches, its gonad consists of four precursor cells: Z1, Z2, Z3 and Z4 (Kimble and Hirsh 1979). Z2 and Z3 cells give rise to germ cells whereas Z1 and Z4 cells give rise to somatic gonad (Kimble and Hirsh 1979). Removal of Z2 and Z3 cells by laser microsurgery results in longer lifespan and interestingly, this lifespan extension phenotype is abrogated by additional removal of Z1 and Z4 cells, suggesting that somatic gonad generates signals in response to germ cells removal to extend lifespan (Hsin and Kenyon 1999) (Figure 1.1.4). It is important to note that complete elimination of gonad, including both germ cells and somatic gonad, does not result in longer lifespan, arguing against the idea that longevity triggered by germline removal is a simple consequence of reproductive trade-off (Hsin and Kenyon 1999). Further study shows that germline stem cells are likely the source that secret signals that shorten lifespan (Arantes-Oliveira et al. 2002). In this study, the authors showed that sterile mutants with defects in spermatogenesis and oogenesis do not always have a longer lifespan (Arantes-Oliveira et al. 2002). However, mutations that reduce the number of germline stem cells, such as loss-of-function mutations in *glp-1* and *mes-1*, extend lifespan whereas mutants with overproliferated germline stem cells, such as *glp-1* and *gld-1*, have a shorter lifespan (Arantes-Oliveira et al. 2002). In response to removal of germline stem cells, KRI-1 and TCER-1 activate DAF-16 and trigger DAF-16 target gene expression and extend lifespan (Berman and Kenyon 2006; Ghazi et al. 2009). Additionally, loss of germline stem cells derepresses DAF-12 signaling pathway that may also contributes to longer lifespan in those germline less mutants (Arantes-

Oliveira et al. 2002). Similar longevity phenotype has also been observed in Drosophila with ablated germline suggesting that reproductive signals regulation in longevity is an evolutionarily conserved process (Flatt et al. 2008).

## Germline

*C. elegans* is a hermaphrodite that contains both female and male reproductive systems, which contains two U-shaped gonad arms joined at a common uterus (Kimble and Hirsh 1979) (Figure 1.1.5). As *C. elegans* hatches, the gonad contains four cells: Z1, Z2, Z3 and Z4. Z2 and Z3 cells give rise to germ cells whereas Z1 and Z4 cells give rise to somatic gonad (Kimble and Hirsh 1979). Each adult gonad arm exhibits distal-proximal polarity with mitotic cells localized at the distal end and meiotic cells at the proximal end (Kimble and Hirsh 1979). The distal end of each arm is capped by a distal tip cell (DTC) that covers the distal tip of mitotic cells, including germline stem cells, and acts as a niche cell for germline stem cells (Hubbard 2007). *C. elegans* germline generates both oocytes and sperms but at different developmental stages: oocytes are produced throughout adult; sperm are produced during L4 stage and then stored and used in adult spermatheca to fertilize oocytes (Kimble and Hirsh 1979). Defects in mitotic proliferation, meiotic proliferation, spermatogenesis and oogenesis, all result in sterility. I will briefly discuss mitotic and meiotic proliferation in *C. elegans* here.

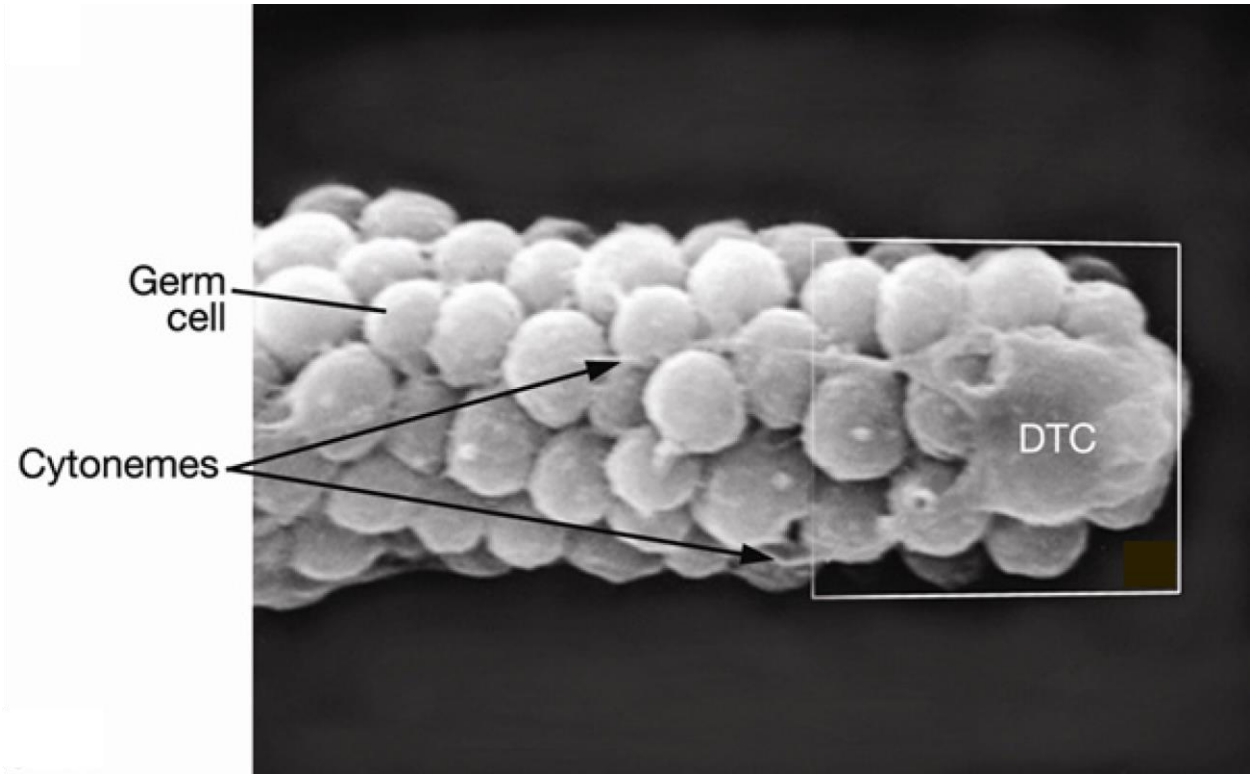



*Schedl Lab*

Figure 1.1.5 *C. elegans* Germline

## Mitotic proliferation

*C. elegans* germline mitotic proliferation zone contains around 230 cells with M-phase like nuclei (Maciejowski et al. 2006). Around two third of the cells that localize in the distal end of the mitotic proliferation zone are actively cycling with one third of the cells that localize in the proximal end are in meiotic S-phase (Fox et al. 2011). Out of a population of 230 cells, around 20 cells enter meiosis each hour (Fox et al. 2011). 90% of the 20 cells will act as nurse cells for supporting the formation of the oocytes and undergo apoptosis (Fox et al. 2011). The distal end of cells in mitotic proliferation zone are in close contact with DTC, which likely acts as niche for germline stem cells (Kimble and White 1981) (Figure 1.1.6). Although the factors that mediate this interaction are unknown, the close contact between the DTC and germ cells is likely important for self-renewal (Crittenden et al. 2006). Ablating the DTC with laser microsurgery results in all germ cells prematurely entering meiotic phase and loss of self-renewal ability (Kimble and White 1981). In addition, mis-positioning of the DTC to proximal position induce ectopic establishment of proliferative cell population near the new position of the DTC (Kimble and White 1981), suggesting that the DTC is sufficient and necessary for stem cells proliferation. The DTC secrets ligands that bind to GLP-1 receptor that localized on the surface of the mitotic cells and trigger gene expression required for self-renewal (Morgan et al. 2010; Austin and Kimble 1987). As the germ cells move proximately, less ligands are bound by GLP-1 receptors and the cells start to enter meiotic phase. Defects in the GLP-1/Notch signaling pathway result in a germline that lacking proliferation cells (Morgan et al. 2010; Austin and Kimble 1987), whereas *glp-1* gain-of-function results in over-proliferation of mitotic cells (Berry et al. 1997).




*WormAtlas*

Figure 1.1.6 *C. elegans* distal tip cell (DTC)

## Meiotic proliferation

Meiosis is a specialized cell division process that cell reduce its chromosome number by half and produce haploid gametes to be used for fertilization. In this process, DNA replication is followed by two rounds of chromosome segregation to generate four daughter cells. In *C. elegans* germline, the most distal end contains population of mitotic cells and meiotic S-phase cells (Fox et al. 2011). As cells move proximately, they enter transition zone, which is evidenced by a distinct crescent moon-shaped DNA morphology as the cells are in leptotene/zygotene of meiotic prophase (Dernburg et al. 1998). Cells that continue to move proximately enter into the pachytene region as they are in pachytene of meiotic prophase and followed by gamete formation.

## Spermatogenesis and Oogenesis

Spermatogenesis is the process in which an animal produces sperms. In *C. elegans*, hermaphrodites produce approximately 300 sperms during the L4 stage and produce exclusively oocytes throughout the adult stage. The sperms are stored in the spermatheca, where they fertilize oocytes during ovulation. Male continue to make sperm and transfer sperm to hermaphrodites through mating.

Oogenesis is the process in which an animal produces oocytes. In *C. elegans*, unlike spermatogenesis that sperm proceed through the meiotic process uninterrupted, oocytes arrest at diakinesis of meiotic prophase I (McCarter et al. 1999). When the oocyte matures, it enters the spermatheca and is fertilized by the sperm at ovulation.

## 1.2 Histone modification

DNA is a long molecule. It wraps around histone octamer, including two each of the histones H2A, H2B, H3, and H4, to form nucleosomes which can be then condensed to form chromatin. Nucleosome can regulate expression of some genes. For example, transcription start site for genes expressed in one tissue are depleted of nucleosome whereas those genes in other tissues are bound by nucleosome (Bargaje et al. 2012). In addition to nucleosome, the highly conserved histone tails can be modified by different enzymes to have different histone modifications which are implicated in regulation of various biological processes including chromatin structure, gene expression and DNA repair. In *C. elegans*, two types of histone modifications have been well characterized. Each of them will be discussed as follow.

### Histone acetylation

The acetylation of lysines group on histone tail can be dynamically regulated by the opposing actions of two classes enzyme: histone acetyltransferases (HATs) and histone deacetylases (HDACs) (Verdin and Ott 2015).

Using acetyl-CoA as donor, HATs transfer the acetyl group onto the lysine residues of histone tail. This reaction neutralizes the positive charges of lysine residues and therefore has the potential to interrupt the electronic interactions between positive charged histones and negative charged DNA (Bannister and Kouzarides 2011). As a result, acetylated histone inhibits the condensation of nucleosome into higher order structure and leads to a more open and active chromatin state, thereby allowing access

to RNA polymerase and transcriptional factors to initiate transcription (Bannister and Kouzarides 2011). On the other hand, HDACs catalyze the reverse reaction that restore the positive charge of histone by removing the acetyl group. This reaction has the potential to stabilize the interaction between histone and DNA, facilitating the folding of nucleosome into higher order chromatin structure and leads to a more compact and repressed chromatin state, thereby inhibiting transcription (Figure 1.2.1).

Almost every lysine residues on histone tail have the potential to be acetylated. These include but not limited to histone H3 lysine 4 (H3K4), H3K9, H3K14, H3K18, H3K23, H3K27, H3K36, H4K5, H4K8, H4K12 and H4K16. The high number of lysine residues that can be acetylated on one hand provides the potential for the positive charge in the hyper-acetylated genomic region to be effectively reduced by acetylation and therefore open chromatin will be generated, but on the other hand makes it more difficult for specificity of acetylation targeting sites. Therefore, limited evidence has suggested that HATs and HDACs can specifically modify acetylation targeting sites (Verdin and Ott 2015).

In addition to interrupting the interaction between histone and DNA, there are more and more evidence suggest that histone acetylation can recruit chromatin factors by specifically binding to their distinct domains (Bannister and Kouzarides 2011). The bromodomain is the first protein domain that has been found to specifically binds to acetylated histone (Owen et al. 2000). The bromodomain has been found in many proteins including but not limited to HATs and many transcriptional factors (Owen et al. 2000). For example, *Saccharomyces cerevisiae* Gcn5p is a transcription-associated HAT. It contains a bromodomain that target acetylated histone, which normally found at

promoter sites. The binding of Gcn5p to the promoters further open the chromatin structure and therefore facilitates the transcription of targeting genes (Owen et al. 2000). The other protein domain that binds to acetylated histone is the tandem PHD domain (Zeng et al. 2010). The human DPF3b has a tandem PHD domain which is

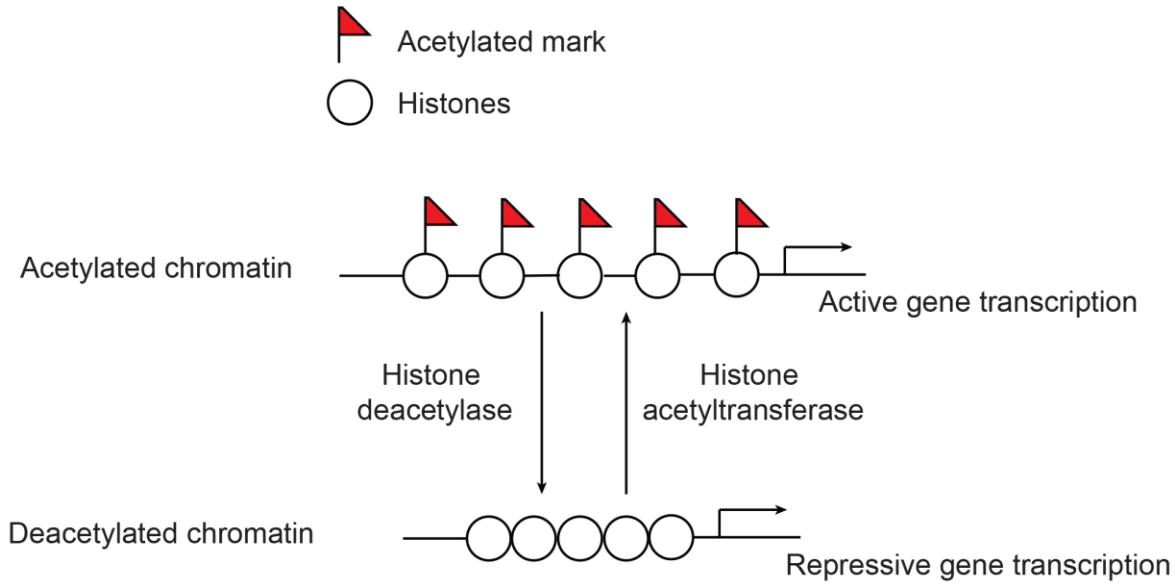



Figure 1.2.1 Histone acetylation and gene transcription

responsible for targeting DPF3b to acetylated histone. The binding of DPF3b to its target genes helps to recruit BAF complex and initiate gene transcription (Zeng et al. 2010).

## Histone methylation

In contrast to histone acetylation that is specifically for lysine residues, histone methylation occurs on all basic residues. However, the most extensively studied sites are the four lysine residues on histone H3: H3K4, H3K9, H3K27, H3K36. Similar to histone acetylation, there are two groups of enzymes that are responsible for the dynamic changes of histone methylation: histone methyltransferase and histone demethylase. Histone methyltransferase transfers methyl group from the donor S-adenosylmethionine to histone (MURRAY 1964) whereas histone demethylase remove methyl group (Shi et al. 2004). However, unlike histone acetylation, the addition of methyl group does not change the charge of the histone protein, therefore the regulation of chromatin state by histone methylation is more complicated. In addition, rather than acetylation, the lysine residue can be mono-, di- or tri-methylated, which add another level of complexity when considering histone methylation (Figure 1.2.2).

Due to the complexity of histone methylation, histone methyltransferase tends to be site specific. For example, in *C. elegans*, SET-domain containing protein SET-2, a histone methyltransferase, specifically methylates H3K4 (Greer et al. 2010). Loss of SET-2 leads to reduced H3K4me3 level in *C. elegans* (Greer et al. 2010). Conversely, the H3K4 demethylase RBR-2 specifically demethylate H3K4 (Greer et al. 2010). In the

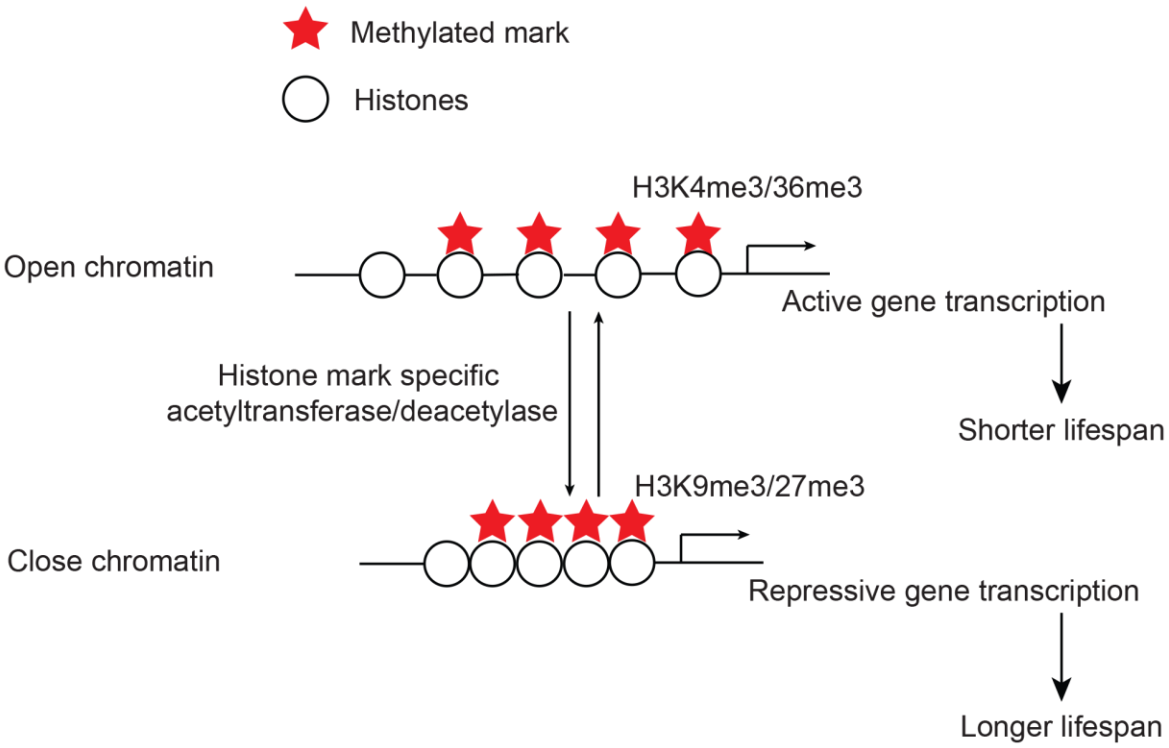



Figure 1.2.2 Histone acetylation, gene transcription and lifespan

absence of RBR-2, the worms will have an elevated H3K4me3 level (Greer et al. 2010). The ability of these enzymes to distinguish different lysine residues and different methylation states is an intrinsic property of the enzymes since purified proteins or even purified catalytic domains are able to catalyze corresponding reactions *in vitro* (Greer et al. 2014).

The different lysine residues to be methylated and different states of methylation provide another way for the organism to mark different chromatin state. For the most extensively studied marks, H3K4 and H3K36 methylation are generally associated with active chromatin state whereas H3K9 and H3K27 methylation are repressive histone marks (Greer and Shi 2012). Many protein domains have been found to distinguish different histone methylations. For example, H3K4me3, an active histone mark that is normally associated with promoter region, is recognized by PHD domain of many proteins including ING2 and PHF8 (Horton et al. 2010; Pena et al. 2006). The binding of PHD domain of ING2 to H3K4me3 bring ING2 to the promoter region to further recruit other transcriptional factors to initiate transcription (Pena et al. 2006), whereas the engagement of PHD domain of PHF8 to H3K4me3 helps the histone demethylase PHF8 to demethylates either H3K9me2 or H3K27me2, results in a more open chromatin structure to facilitate transcription (Horton et al. 2010). Another example is the repressive histone marks H3K9me3, which is normally associated with heterochromatin region. The H3K9me3 specific histone methyltransferase human SUV39H1 (Lachner et al. 2001) and *Schizosaccharomyces pombe* Clr4 (Bannister et al. 2001) place H3K9me3, which is then recognized by the chromo domain of HP1 to form heterochromatin structure.

## Cross-talk between histone marks

The existence of a large number of histone modification sites provide a way to tightly regulate chromatin state. In addition, since histone modifications do not act alone the cross-talk between different histone marks add another level of chromatin regulation.

Histone modification can be added onto different sites of the same histone, but some histone modifications are mutually exclusive. For instance, H3K4me2/3 and histone acetylation, active histone modification marks that generally localize at promoter regions of active transcribed genes, are normally not colocalized with repressive histone marks H3K9me2/3 (Greer et al. 2014). Gradual increase of H3K4me2 marks by loss of H3K4me1/2 demethylase SPR-5 accompanies by gradual loss of H3K9me3 marks in *C. elegans* (Greer et al. 2014). Similarly, H3K27me3 and H3K36me3 generally localize at mutually exclusive region in *C. elegans* (Gaydos et al. 2012). Loss of MES-4, a histone methyltransferase that responsible for H3K36 methylation, remove H3K36me3 marks and cause redistribution of H3K27me3, which causes mis-regulation of gene expression (Gaydos et al. 2012).

In contrast, some histone modifications normally co-exist at certain regions. For example, H3K4me3 colocalizes with many histone acetylation marks at the promoter of active transcribed genes. There are many studies suggest that HATs, responsible for depositing histone acetylation, is recruited by H3K4me3 (Pena et al. 2006). *Saccharomyces cerevisiae* Sgf29, a component of SAGA complex, binds to H3K4me3 and recruit HATs (Bian et al. 2011). Deletion of Sgf29 leads to loss of SAGA at targeted sites and decreased level of histone acetylation (Bian et al. 2011). On the other hand,

HDACs, responsible for removing histone acetylation, have also been found to be recruited at H3K4me3 site (Pena et al. 2006). Therefore, the dynamic of histone acetylation rather than the mark itself plays an important role in chromatin regulation.

## 1.2.1 Histone modification and germline function in *C. elegans*

Dynamic regulation of histone modification has been observed during germline development in *C. elegans*. In general, consistent with underrepresented germline-specific gene on the X chromosome (Reinke et al. 2004), active histone marks, such as histone acetylation, H3K4me3 and H3K36me3, are more enriched on autosome whereas repressive histone marks, such as H3K9me3 and H3K27me3 are more enriched on X chromosome (Liu et al. 2011). The distribution and dynamic regulation of histone modification is crucial for the germline development and alteration of histone modification by depleting chromatin factors in many cases leads to defects in germline development.

### H3K27 methylation and H3K36 methylation

MES-2/-3/-6 complex is the histone methyltransferase that is responsible for H3K27me3 (Bender et al. 2004) in *C. elegans* whereas MES-4 is responsible for H3K36me3 (Bender et al. 2006). H3K27me3 and H3K36me3 generally mark different chromosome region: H3K27me3 is concentrated on the X chromosome and H3K36me3

is enriched on the autosome but depleted on the X chromosome (Bender et al. 2004, 2006). Consistent with its enzymatic activity, MES-4 colocalizes with H3K36me3 on the autosome and is absent on the X chromosome (Bender et al. 2006). Interestingly, this H3K36me3 pattern depends on H3K27me3 mark since removing H3K27me3 by depleting MES-2/-3/-6 complex leads to increase of H3K36me3 and MES-4 on the X chromosome suggesting that H3K27me3 protect the X chromosome to be marked by H3K36me3 (Bender et al. 2004, 2006). This regulation is crucial for germline development since *mes-2*, *mes-3*, *mes-4 and mes-6* mutants all results in a maternal sterile phenotype (Bender et al. 2004, 2006) (Figure 1.2.3).

## H3K4 methylation and H3K9 methylation

As discussed above, H3K4 methylation is generally associated with active transcription. MLL complex is the major enzyme that is responsible for depositing H3K4me3 mark in many organisms including *C. elegans* (Shilatifard 2012). The MLL complex in *C. elegans* contains four components: the histone methyltransferase *set-2*, *wdr-5.1*, *ash-2* and *rbbp-5* (Li and Kelly 2011). These members are important for H3K4me2/me3 in the germline and early embryo since depleting any of them significantly reduce H3K4me2/me3 levels. In addition, *set-2*, *wdr-5.1*, *ash-2* and *rbbp-5* mutants all exhibit germline defects suggesting the importance of the H3K4me2/me3 marks in the germline function (Li and Kelly 2011; Xiao et al. 2011) (Figure 1.2.3).

In contrast to H3K4 methylation, H3K9 methylation commonly correlates with transcriptional silencing. H3K9me2 and H3K9me3 have distinct pattern in *C. elegans*

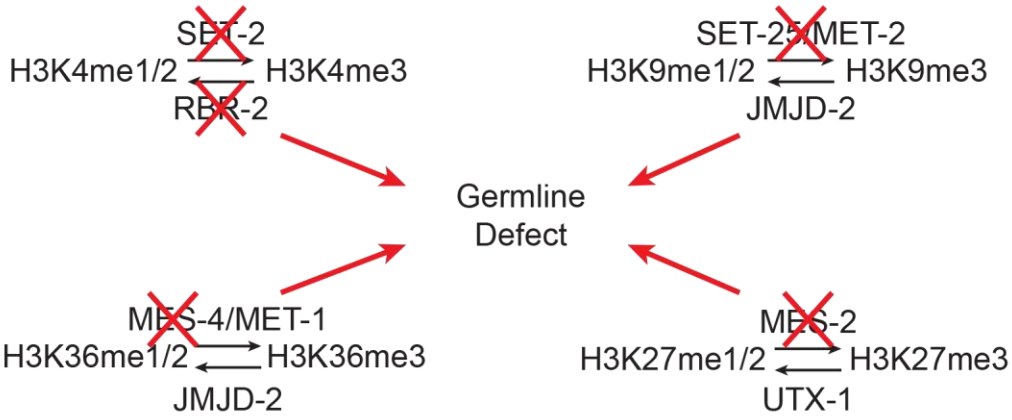



Figure 1.2.3 Histone methylation and germline defect

germline. H3K9me2 is enriched on unpaired sex chromosomes whereas H3K9me3 localizes to all chromosomes in all germ cells (Bessler et al. 2010). The SET domain containing protein MET-2 is responsible for H3K9me2 but not H3K9me3 in the germline (Bessler et al. 2010). Loss of MET-2 leads to almost complete depletion of H3K9me2 but H3K9me3 is unaffected (Bessler et al. 2010). On the other hand, high copied exchromosome array with somatic specific promoter is highly enriched for H3K9me3 correlating its role in transgene silencing (Bessler et al. 2010). The histone methyltransferase MES-2 is required for H3K9me3 but not H3K9me2 (Bessler et al. 2010). It is interesting to note that MES-2 is generally thought to be a H3K27me3 methyltransferase (Bender et al. 2004). It is possible that MES-2 is responsible for directly putting both H3K9me3 and H3K27me3 marks or the presence of H3K27me3/H3K9me3 is required for the deposition of H3K9me3/H3K27me3.

## Histone demethylation

In addition to histone methylation, histone demethylation is also important for germline function. Recently, SPR-5, the H3K4me1/me2 histone demethylase in *C. elegans*, has been found to play an important role in germline function (Katz et al. 2009; Kerr et al. 2014; Greer et al. 2014). Loss of *spr-5* leads to mortal germline phenotype, in which the brood size reduces progressively over generations (Katz et al. 2009). Accompanying reduced brood size, the global H3K4me2 level as well as H3K4me2 level in two primordial germ cells also increase over generations (Katz et al. 2009). Interestingly, the gradually increased global H3K4me2 level and reduced brood size are

rescued by removal of H3K4 methyltransferases (Greer et al. 2014) and are worsened by depleting H3K9 methyltransferases, MET-2 (Kerr et al. 2014) and SET-26 (Greer et al. 2014), suggesting the importance of cross-talk between different histone modifications. Consistent with the important role of H3K4me2 in germline function, the absence of RBR-2, another H3K4me1/me2 histone demethylase in *C. elegans*, also causes germline immortality (Alvares et al. 2014).

## Potential mechanisms

Many different potential mechanisms have been proposed to link the germline defects to the altered histone modification. These different possibilities might not be mutually exclusive.

First, the alteration in histone modification leads to changes in gene expression. The late generation *spr-5* mutant has increased H3K4me2 level, which generally correlates with increased gene expression (Katz et al. 2009). This elevated H3K4me2 in the late generation *spr-5* mutant associates with up-regulated gene expression in spermatogenesis-expressed genes (Katz et al. 2009; Kerr et al. 2014). It was proposed that SPR-5 acts to demethylate H3K4me2 to prevent the transmission of this mark to the successive generations and therefore inhibit mis-expression of different genes, especially spermatogenesis-expressed genes (Katz et al. 2009; Kerr et al. 2014), to maintain normal germline function.

Second, histone modification protects the germline fate of germ cells. Loss of both H3K4 demethylase SPR-5 and chromatin factor Mi2 LET-418 in *C. elegans* cause mis-

expression of somatic specific genes in the germline suggesting germ cells are transformed into somatic cells (Kaser-Pebernard et al. 2014). On the other hand, depleting MLL complex components: H3K4 methyltransferase SET-2 or WDR-5.1 also leads to ectopic expression of somatic markers in the germline (Robert et al. 2014). Soma-like cells can also be observed in the germline of *set-2* and *wdr-5.1* mutants confirming the transformation of germ cells into somatic cells (Robert et al. 2014). It is interesting to note that H3K4 methyltransferase SET-2 and H3K4 demethylase SPR-5 are two enzymes that catalyze the confronting reactions on H3K4, suggesting that the methylation on H3K4 must be appropriately regulated to maintain germline fate of germ cells.

Third, histone modification is important for DNA damage repair. SPR-5 is involved in meiotic DNA double-strand break repair in the germline of *C. elegans* (Nottke et al. 2011). Removal of SPR-5 causes progressive DNA damage repair defects including increased p53-dependent germ cell apoptosis, increased DNA damage repair markers RAD-51 and higher sensitivity under double strand break inducing treatments (Nottke et al. 2011). In addition, SPR-5 relocalizes rapidly in response to DNA double-strand breaks and genes with altered expression in the spr-5 mutant do not include known DNA damage repair gene, suggesting that SPR-5 acts directly in DNA double-strand break (Nottke et al. 2011).

All of these results suggest that histone modifications are important for germline and they must be carefully regulated to maintain normal germline function.

## 1.2.2 Histone modification and longevity regulation

Many studies have revealed that chromatin-based epigenetic changes accompany aging and that interfering chromatin factors affect lifespan in many model organisms. Global loss of core histone proteins have been observed during aging from yeast to mammalian cells (Liu et al. 2013b; Feser et al. 2010). Interestingly, overexpression of H3 and H4 extend the replicative lifespan of yeast by up to 50% (Feser et al. 2010), suggesting the importance of core histone level in lifespan regulation.

In addition to core histone level, various histone modifications have also been shown to play an important role in lifespan regulation.

### Histone methylation

H3K9me2 and H3K9me3 are the two well-known histone marks that associated with heterochromatin, the loss of which has been proposed as driving force to aging (Villeponteau 1997). Consistent with this model, global H3K9me3 level has been found reduced in the aged *C. elegans* (Ni et al. 2012). Progressive and stochastic age-dependent alterations have been observed in the nuclear architecture of old *C. elegans* and reduced lamin and lamin-associated proteins shorten lifespan (Haithcock et al. 2005), suggesting the importance of nuclear architecture in lifespan regulation. In addition to *C. elegans*, loss of heterochromatin has also been observed in Drosophila (Larson et al. 2012). Flies with increased heterochromatin have a longer lifespan whereas reducing heterochromatin shorten lifespan (Larson et al. 2012). Heterochromatin is important for suppressing ribosomal RNA synthesis, the

overexpression of which has been associated with shorter lifespan (Larson et al. 2012). In human, heterochromatin is also implicated in longevity regulation. A global loss of H3K9me3 has been found in human WS model in human embryonic stem cells (ESCs) and older individuals (Zhang et al. 2015). Loss of enzymatic activity of SUV39H1, the histone methyltransferase that is responsible for H3K9me3, accelerates cellular senescence in WS model (Zhang et al. 2015). However, the role of H3K9me3 in lifespan regulation is complicated. Depleting the methyltransferase SUV39H1 improves DNA repair and extends lifespan in a progeria mouse model (Liu et al. 2013a). Nevertheless these results suggest the importance of H3K9me3 in lifespan regulation.

In *C. elegans*, H3K27me3, another repressive histone mark, level is reduced during aging (Maures et al. 2011; Jin et al. 2011). Depleting histone methyltransferase and histone demethylase that targeting H3K27 both alter lifespan. Knockdown of H3K27 demethylase *utx-1* leads to increased H3K27me3 in old worms and extended lifespan (Maures et al. 2011; Jin et al. 2011) (Figure 1.2.4). On the other hand, knockdown of H3K27 methyltransferase *mes-2*, which leads to reduced H3K27me3 level, also cause longer lifespan in sterile worms (Ni et al. 2012), suggesting that H3K27me3 level positively correlates with lifespan in *C. elegans*. In contrast, reduced H3K27me3 level is associated with extended lifespan in Drosophila (Siebold et al. 2010). Depleting polycomb repressive complex-2 (PRC2), H3K27me3 histone methyltransferase, lead to reduced H3K27me3 level and extended lifespan (Siebold et al. 2010). Removal of polycomb's antagonist Trithorax group proteins restores the H3K27me3 level and rescues the increase in longevity, suggesting that H3K27me3 level negatively correlates with lifespan in Drosophila (Siebold et al. 2010). The

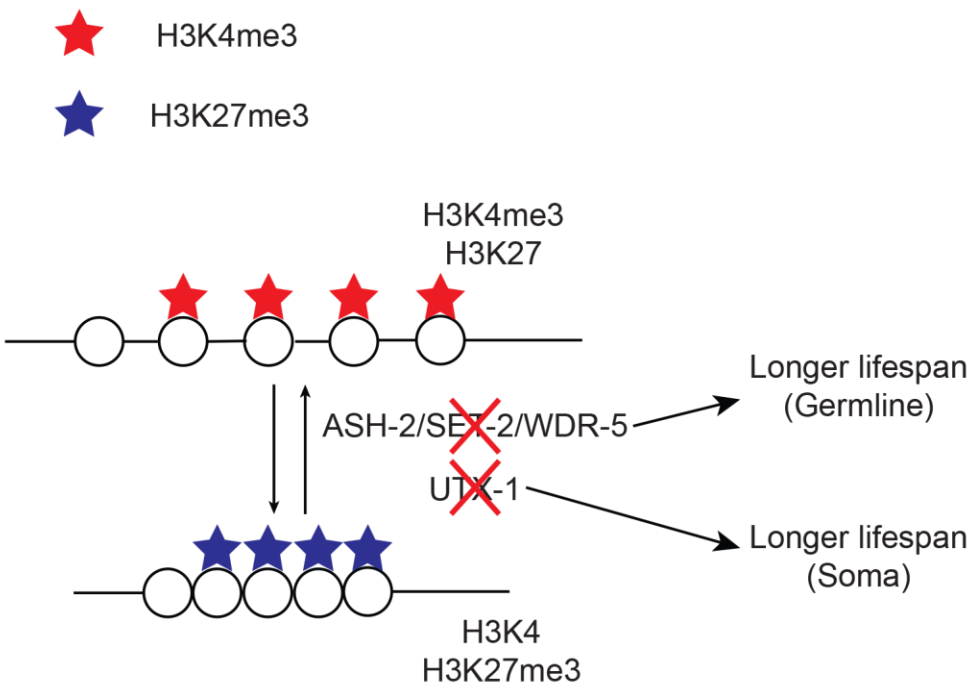



Figure 1.2.4 Loss of H3K4me3 in the germline and gain of H3K27me3 in soma extend lifespan

discrepancy between H3K27me3 level and lifespan in *C. elegans* and Drosophila might be the result of species differences.

H3K4me3, an active histone methylation mark, is also implicated in *C. elegans* lifespan regulation. Disrupting the histone methyltransferase complex for H3K4me3 causes decreased H3K4me3 and extends lifespan in a germline dependent manner (Greer et al. 2010) (Figure 1.2.4). Interestingly, this lifespan extension phenotype can be inherited epigenetically as the wild type progenies from their grandparent long-lived H3K4me3 methyltransferase mutant also has a comparable extended lifespan as their grandparent (Greer et al. 2011) (Figure 1.2.5). This transgenerational lifespan inheritance can be lasted for up to 5 generations (Greer et al. 2011). On the other hand, loss of H3K4me demethylase *rbr-2*, which leads to increased H3K4me3 level, also results in extended lifespan (Alvares et al. 2014). Interestingly, loss of H3K4me demethylase *spr-5*, cause a transgenerational increase in lifespan (Greer et al. 2016). Although early generation *spr-5* mutant has a normal lifespan, the late generation *spr-5* mutant has a significant extended lifespan compared to wild type (Greer et al. 2016). This lifespan extension phenotype depends on *daf-12* but is independent of germline as this lifespan phenotype can also be observed in the germless *glp-1* mutant background (Greer et al. 2016).

H3K36me3 is another active histone mark that is involved in the aging process. Rather than H3K4me3 that associates with promoter and involves in transcriptional initiation, H3K36me3 marks broadly within gene body and associates with transcriptional elongation (Bannister et al. 2005). A lifespan screen that aims to identify

amino acid of histones that regulate yeast replicative aging reveals that the H3K36 methylation is important for lifespan (Sen et al. 2015).

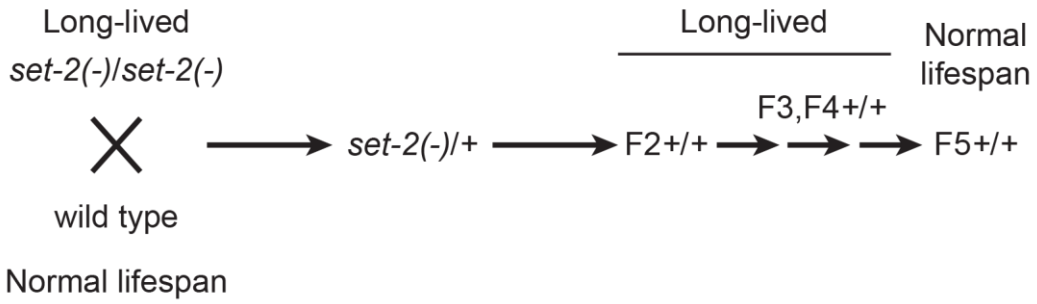


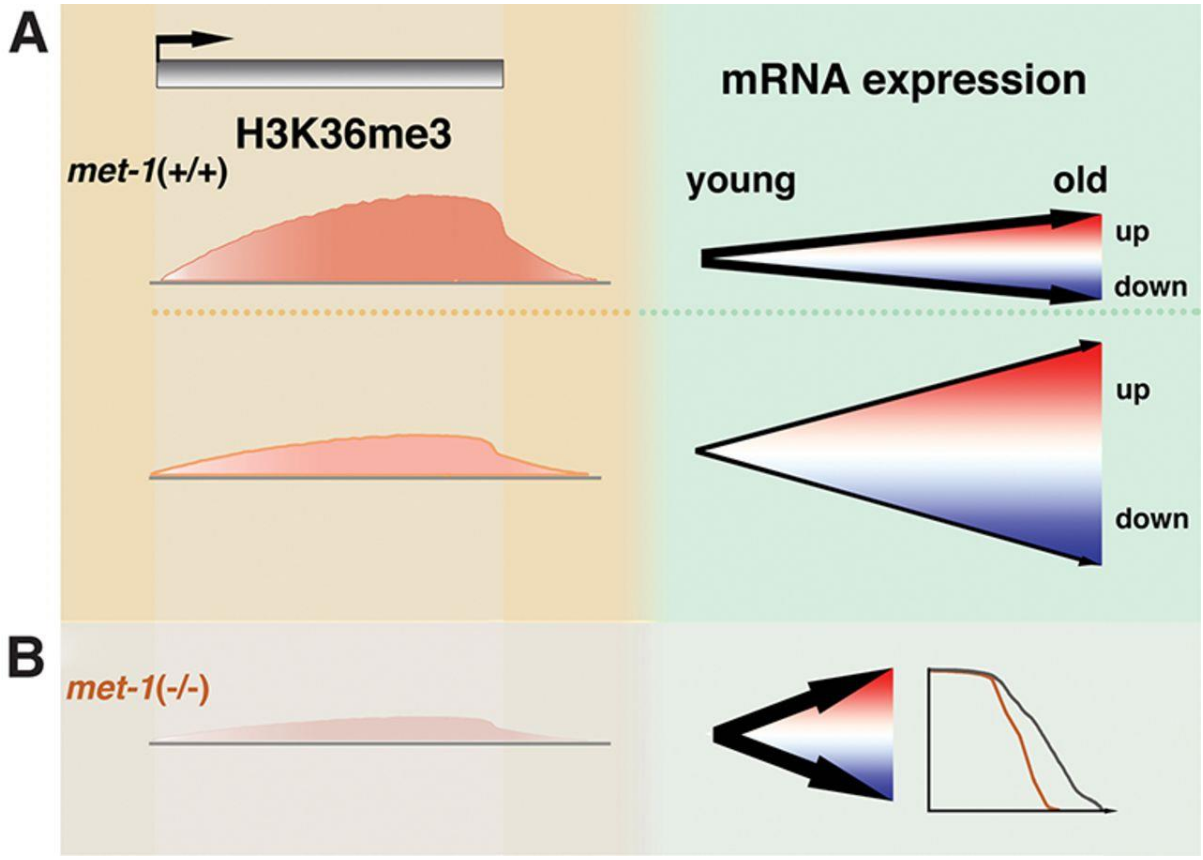

Figure 1.2.5 Transgenerational inheritance of lifespan in *set-2(-)* mutant

The lack of H3K36 methylation leads to shorter lifespan whereas removal of the K36me2/3 demethylase Rph1 causes increased H3K36me3 and extended lifespan suggesting that H3K36me3 positively correlates with lifespan (Sen et al. 2015). The marking of H3K36me3 over gene body is thought to prevent the cryptic transcription in the old animals and thus extends lifespan (Sen et al. 2015). Depletion of H3K36me3 causes increased cryptic transcription which is detrimental to the lifespan (Sen et al. 2015). Another study in *C. elegans* also supports the importance of H3K36me3 in longevity regulation (Pu et al. 2015). Consistent with yeast data, in the absence of MET-1, *C. elegans* has a shorter lifespan (Pu et al. 2015). There is no age-dependent change in the distribution of H3K36me3 (Pu et al. 2015). However, there is negative correlation between H3K36me3 level and differential gene expression upon aging (Pu et al. 2015), suggesting that H3K36me3 is important for maintaining transcriptional consistency during the aging process (Figure 1.2.6). This correlation has also been observed in the brain tissue of Drosophila suggesting the evolutional conservation of the role H3K36me3 (Pu et al. 2015).

## Histone acetylation

Histone acetylation is another histone modification that implicated in the regulation of aging and aging-related diseases.

In *C. elegans*, histone acetylation has been found involved in lifespan regulation (Edwards et al. 2014; Zhang et al. 2009a). The expression of CBP-1, a histone acetyltransferase, is induced by dietary restriction and in *daf-2* mutant, and results in increase histone acetylation (Zhang et al. 2009a). Histone acetylation level decreases




*Pu M, Genes Dev, 2015*

Figure 1.2.6 H3K36me3 restricts gene expression change with age and maintains life span

with age and cbp-1 RNAi knockdown reduces histone acetylation level and shorten lifespan whereas histone acetylation enhancer treatment extends lifespan, suggesting that histone acetylation positively correlates with lifespan (Zhang et al. 2009a). Consistent with this idea, beta-hydroxybutyrate (beta-HB), a histone deacetylase (HDAC) inhibitor, treatment increases lifespan in *C. elegans*. RNAi knockdown of HDACs hda-2 or hda-3 also extends lifespan (Edwards et al. 2014). These results together suggest the importance of histone acetylation in *C. elegans* longevity regulation.

In yeast, H3K16ac level increases with age (Dang et al. 2009). This increase is accompanied by the declined level of Sir2, a histone deacetylase that removes histone acetylation (Dang et al. 2009). Loss of Sir2 leads to elevated H4K16ac level and results in shorter lifespan whereas depleting Sas2, the H4K16 acetyltransferase that antagonizes Sir2, also antagonizes the effect of depleting Sir2 and causes reduced H4K16ac level and results in longer lifespan (Dang et al. 2009). The substitution of H4K16 to glutamine (H4K16Q), which mimics acetylated state, significantly reduces lifespan, further confirms the detrimental role of H4K16ac in lifespan determination (Dang et al. 2009). Sir2 may act at telomeres to prevent the spreading of H4K16ac and loss of Sir2 causes hyper-acetylation of H4K16 at telomeres and results in shorter lifespan (Dang et al. 2009). However, the role of H4K16ac in longevity regulation is complicated. In mammals, depleting Zmpste24, the lamin A-processing enzyme, causes the Hutchinson Gilford progeria syndrome (Varela et al. 2005). Hypo-acetylation of H4K16 is observed in the *Zmpste24*$^{-/-}$ mice (Krishnan et al. 2011). The reduction in the H4K16ac level is accompanied by the reduced nuclear matrix association of Mof, the

histone acetyltransferase that deposits histone acetylation (Krishnan et al. 2011). Feeding *Zmpste24*$^{-/-}$ mice with HDAC inhibitors restores H4K16ac level, rescues DNA repair defect and extends lifespan of *Zmpste24*$^{-/-}$ mice (Krishnan et al. 2011). These results suggest that the accelerated aging phenotype might be direct consequence of the reduced H4K16ac level.

All of these results suggest that histone modifications are important for longevity regulation.

## 1.2.3 Histone modification and RNAi pathway

RNA interference (RNAi) is a conserved cellular mechanism in which RNA inhibits transcription and/or translation of targeted genes. It involves post-transcriptional silencing and transcriptional silencing. Post-transcriptional silencing, or the canonical RNAi pathway, resides in the cytoplasm, where RNAs are processed into small interference RNA (siRNA) which paired with complementary mRNA strands by the RNA-Induced Silencing Complex (RISC). The mRNA bound by RISC are unable to be translated and are prompted for degradation (Holoch and Moazed 2015). In addition to post-transcriptional silencing, there exists a lesser-known pathway within the nucleus. This nuclear RNAi pathway is initiated when Argonaute (AGO) proteins shuttle cytoplasmic processed siRNA into the nucleus. Once inside, the siRNA guides the AGO complex to nascently-transcribed, complementary RNA, stalls transcription at these loci and modifies epigenetic marks (Figure 1.2.7). This process includes many processes to inhibit the synthesis of mRNA, so it is also called transcriptional silencing

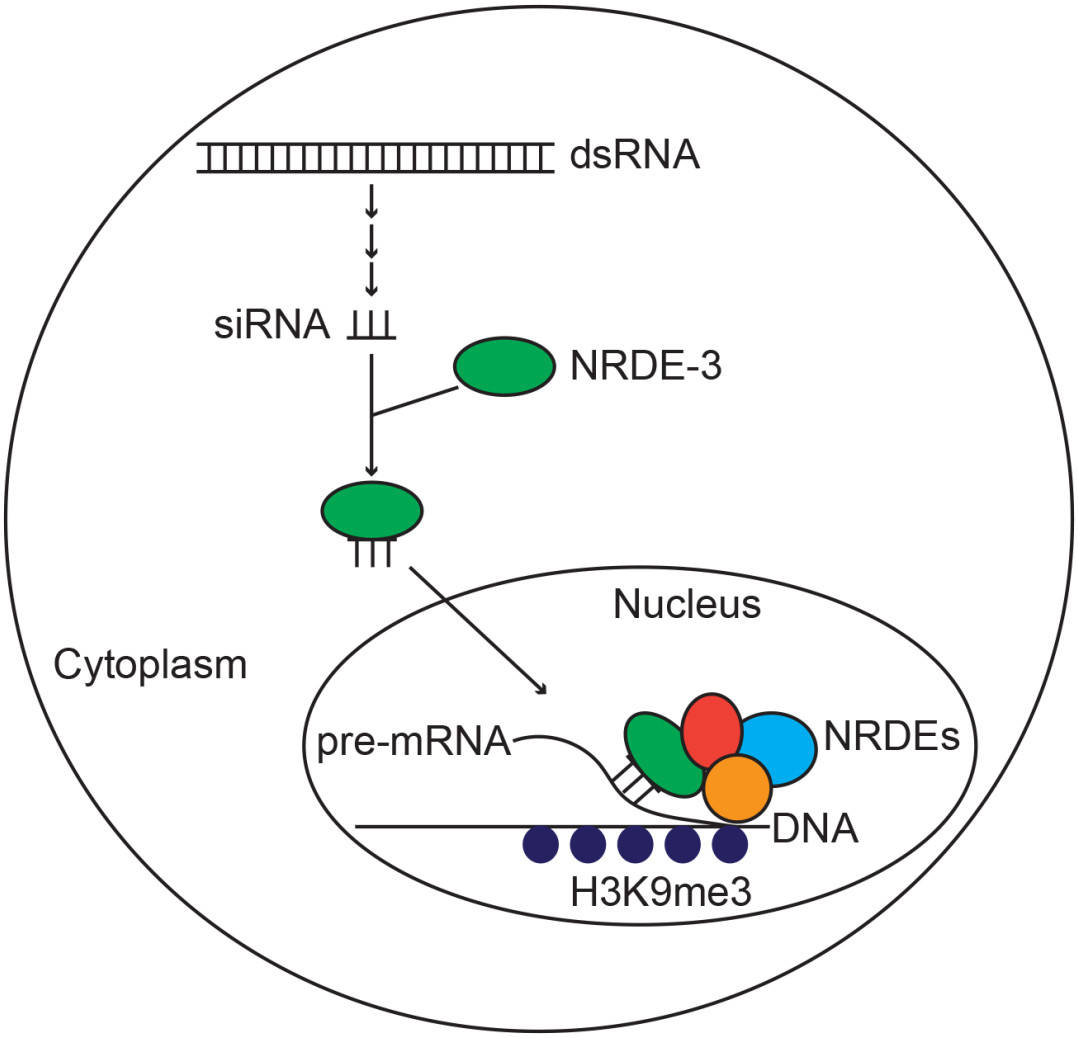
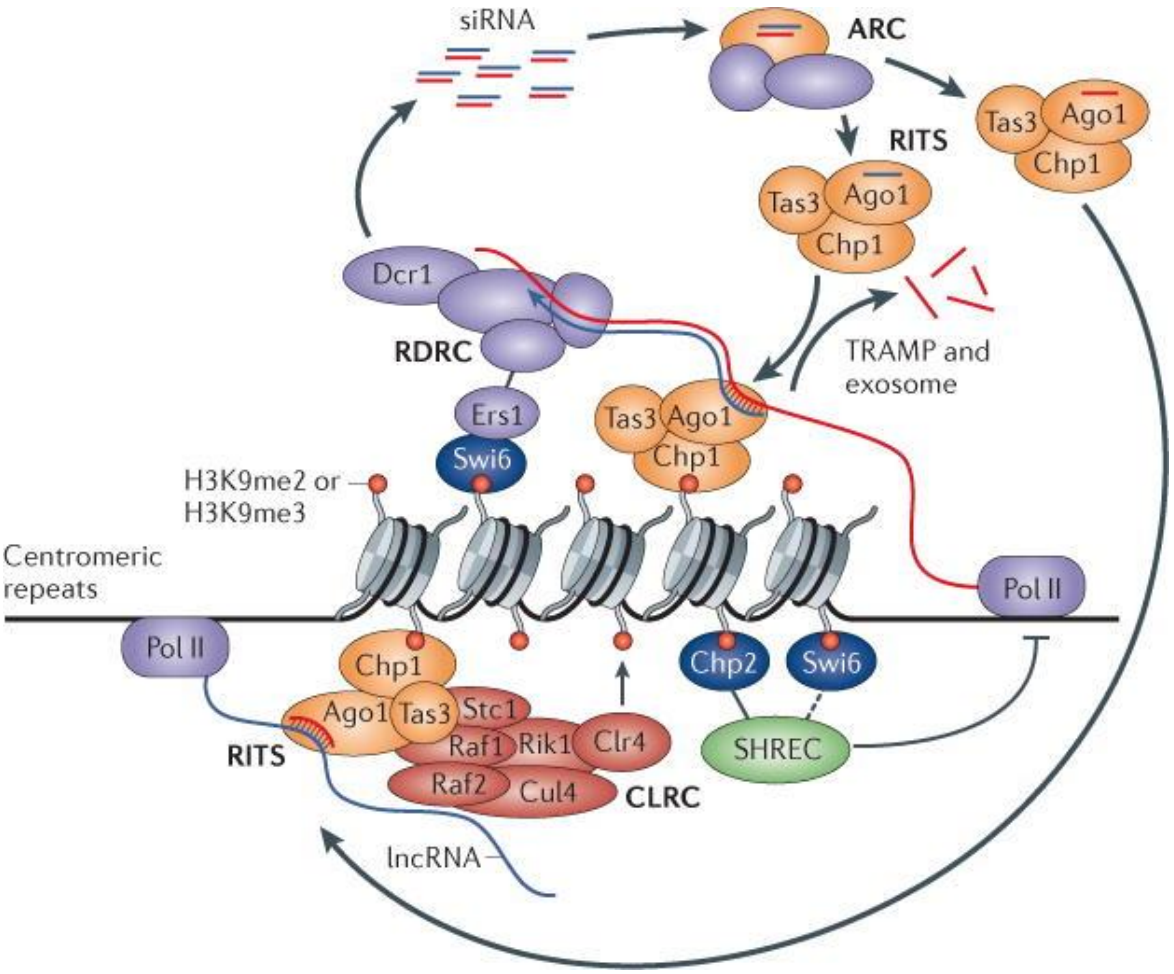



Figure 1.2.7 Nuclear RANi pathway

pathway (Holoch and Moazed 2015). In this thesis, I will mainly focus on the less known nuclear RNAi pathway.

## RNAi mediates heterochromatin assembly

Small RNAs are first found associated with heterochromatin assembly in fission yeast *Schizosaccharomyces pombe* (Buhler et al. 2006; Verdel et al. 2004; Motamedi et al. 2004). The RITS complex contains Ago1, the fission yeast Argonaute homolog, Chp1, a heterochromatin-associated chromodomain protein, and Tas3 (Verdel et al. 2004). These RITS complex components associate with small RNAs and bind to heterochromatin region in a RNAi dependent manner (Verdel et al. 2004). Depleting any of these components causes defects in heterochromatin formation (Verdel et al. 2004). The RNA-directed RNA polymerase complex (RDRC) includes RNA-directed RNA polymerase (Rdp1), Hrr1, an RNA helicase, and Cid12, a member of the polyA polymerase family (Motamedi et al. 2004). RDRC interacts with RITS in a Dicer ribonuclease (Dcr1), which involved in dsRNA processing, and the Clr4 histone methyltransferase, which involved in H3K9me3 deposition, dependent manner (Motamedi et al. 2004). These two complexes also associate with siRNA (Motamedi et al. 2004). Ectopic tethering of RITS complex to euchromatin region triggers H3K9me3 deposition (Buhler et al. 2006) (Figure 1.2.8). A recent study shows the unique roles for different H3K9me states in RNAi and heritable silencing of transcription (Jih et al. 2017). Using a Clr4 active-sites mutations that specifically lacks H3K9me3, but allow H3K9me2 activity, the authors show that H3K9me2 is sufficient for RNAi-dependent co-




*Moazed D, Nat Rev Genet, 2015*

Figure 1.2.8 RNAi mediated heterochromatin assembly in *S. pombe*

transcriptional gene silencing (Jih et al. 2017). The H3K9me3 regions are transcriptionally silent whereas H3K9me2 regions, which also mark by other active histone modifications, are transcriptionally active (Jih et al. 2017). These results suggest a model that siRNA guides RITS to its targeting genomic region to establish heterochromatin assembly by depositing different H3K9me states and the interaction between RITS and RDRC further promotes the siRNA synthesis to reinforce heterochromatin silencing.

## piRNA

piRNAs were first found associated with transposon silencing in mice and Drosophila (Malone and Hannon 2009). They silence the expression of transposable elements at both transcriptional and post transcriptional level (Siomi et al. 2011; Huang et al. 2013; Le Thomas et al. 2013; Rozhkov et al. 2013; Sienski et al. 2012). In the cytoplasm, piRNAs guide piwi complexes to its transposon target through complementary base pairing and endonucleolytic cleavage (Siomi et al. 2011). In the nucleus, piRNAs guide piRNA-complex, including Piwi protein, HP1a and Su(var)3-9 to the piRNA-complementary sequences in the genome (Le Thomas et al. 2013; Rozhkov et al. 2013). The recruitment of H3K9me3 methyltransferase Su(var)3-9 establishes heterochromatic H3K9me3 marks on its targeting region and represses transcription (Huang et al. 2013; Sienski et al. 2012; Le Thomas et al. 2013). Loss of components of piRNA-complex causes reduction of heterochromatic H3K9me3 marks on its targeting sites and derepression of transposon expression (Sienski et al. 2012; Rozhkov et al.

2013). Insertion of piRNA-complementary sequence to an ectopic site in the genome cause recruitment of piRNA-complex, including Piwi protein, HP1a and Su(var)3-9 and reduced RNA Pol II occupation as well as H3K9me2/3 enrichment (Le Thomas et al. 2013; Rozhkov et al. 2013) suggesting that piRNA is sufficient and necessary for piRNA-complex recruitment and transcriptional silencing.

## Nuclear RNAi

In *C. elegans*, the cytoplasm production of primary siRNA induced by exogenous dsRNA triggers secondary siRNA production (Yigit et al. 2006). In addition to facilitating mRNA degradation in the cytoplasm, secondary siRNA can be shuttled into the nucleus by Nuclear RNAi-Defective 3 (NRDE-3) protein (Guang et al. 2008). Once inside, RNA-bound NRDE-3 binds nascently-transcribed, complementary RNA and form complexes with other NRDE proteins around those loci (Guang et al. 2010). The NRDE complex then represses transcription by depositing heterochromatic marks H3K9me3 and H3K27me3 and stalling transcription at these loci (Mao et al. 2015; Guang et al. 2010). The RNAi-induced H3K9me3 and H3K27me3 can be inherited transgenerationally as these marks last for several generations in the absence of dsRNA trigger (Gu et al. 2012; Mao et al. 2015). The transgenerational inheritance of H3K9me3 and H3K27me3 involves transmission of germline siRNA and depends on the germline-expressed argonaute protein HRDE-1 (Gu et al. 2012; Mao et al. 2015). Lose of *hrde-1* causes mortal germline in worms suggesting that the inheritance of siRNA is important for germline immortality and fertility (Buckley et al. 2012).

## 1.3 SET-9 & SET-26

*C .elegans* SET-9 & SET-26 are two highly conserved proteins. They contain a SET (Su(var)3-9, Enhancer of Zeste, Trithorax) domain and a PHD (Plant HomeoDomain) domain. SET domain-containing proteins represent a major group of histone methyltransferases (Dillon et al. 2005), which is critical for methylating histones. The SET domain of SET-9 & SET-26 contains mutations in conserved residues thought to be key for methylating activities (Ni et al. 2012), making it unclear whether SET-9 & SET-26 could be active enzymes. Nevertheless, a recent study reported that the SET domain of SET-26 exhibits H3K9me3 activity *in vitro* (Greer et al. 2014). The PHD domain of some proteins have been reported to bind to histones in a sequence- and modification-dependent manner (Shi et al. 2007, 2006) raising the possibility that SET-9 and SET-26 could be recruited to chromatin via binding to specific histone marks.

MLL5 and UpSET represent likely homologs of SET-9 & SET-26 in mammals and in Drosophila. UpSET and MLL5 all harbor centrally localized SET and PHD domains, and they share ~30-40% sequence identity in their PHD and SET domains. The PHD domain of MLL5 was found to bind to H3K4me3 *in vitro* (Ali et al. 2013), and both MLL5 and UpSET have been shown to localize at promoter regions in cultured cells (Ali et al. 2013; Rincon-Arano et al. 2012). Like the SET domain of SET-9 and SET-26, the SET domains of MLL5 and UpSET lack the key catalytic amino acids and have been proposed to not be an active enzyme (Rincon-Arano et al. 2012; Mas-Y-Mas et al. 2016). MLL5 was discovered as a candidate myeloid leukemia tumor suppressor gene (Emerling et al. 2002) and has been implicated to play an important role in multiple

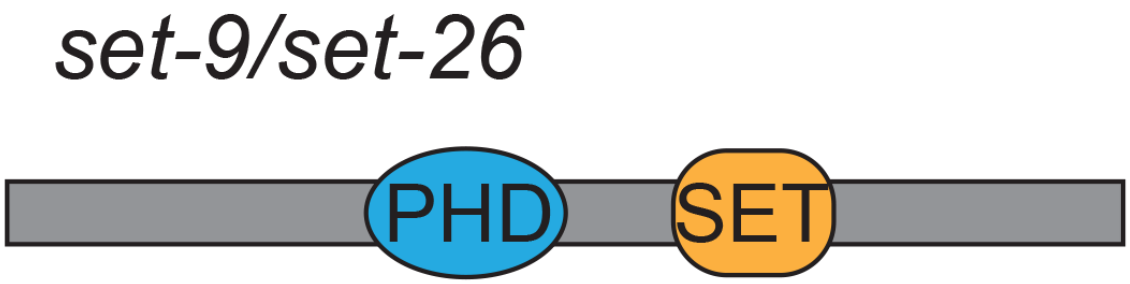



Figure 1.2.9 SET-9 and SET-26 contain a PHD and a SET domain

biological processes (Zhang et al. 2017). Knockdown of MLL5 causes cell cycle inhibition (Liu et al. 2012) and genomic instability (Liu et al. 2012) in cultured cells. MLL5-deficient mice exhibit hematopoietic stem cell defects and male infertility (Liu et al. 2012). However, the molecular mechanisms whereby MLL5 regulate these biological functions are largely unknown. Loss of UpSET, on the other hand, results in female sterility in Drosophila (Rincon-Arano et al. 2012). Recent studies have suggested that both MLL5 and UpSET have roles in regulating chromatin accessibility in cells (Rincon-Arano et al. 2012; Gallo et al. 2015).

*set-9* and *set-26* share 96-97% identity in both DNA sequences and protein sequences (Ni et al. 2012). In a previous targeted RNAi screen to identify the SET-domain containing proteins that play a role in longevity in C. elegans, we found that RNAi knockdown of *set-9* and *set-26* results in significant lifespan extension (Ni et al. 2012). We confirmed the extended lifespan by using the *set-26* mutant (Ni et al. 2012) and due to the similarity of the sequences between *set-9* and *set-26* we expected that the *set-9* mutant also has a longer lifespan although there was no *set-9* mutant available at the time.

In summary, SET-9 & SET-26 are two important proteins in *C. elegans*. However, the molecular roles of SET-9 & SET-26 are largely unknown. Elucidating the mechanisms by which SET-9 & SET-26 play in *C. elegans* will help us better understand the roles and molecular functions of MLL5 in humans.

# CHAPTER 2

# SET-9 and SET-26 are H3K4me3 readers and play critical roles in germline development and longevity[1]


## Abstract

SET-9 and SET-26 are highly homologous paralogs that share redundant function in germline development, but SET-26 alone plays a key role in longevity and heat stress response. SET-26 is broadly expressed, but SET-9 is only detectable in the germline, which likely account for their different biological roles. SET-9 and SET-26 bind to H3K4me3 with adjacent acetylation marks in vitro and in vivo. In the soma, SET-9 and SET-26 act through DAF-16 to modulate longevity. In the germline, SET-9 and SET-26 restrict H3K4me3 domains around SET-9 and SET-26 binding sites, and regulate the expression of specific target genes, with critical consequence on germline development. SET-9 and SET-26 are highly conserved and our findings provide new insights into the critical role of these H3K4me3 readers in germline function and longevity.


## Introduction

Dynamic regulations of histone methylation status have been linked to many biological processes (Santos and Dean 2004). Recent studies have revealed that specific histone

[1] This Chapter is modified from "Wenke Wang, Amaresh Chaturbedi, Minghui Wang, Serim An, Satheeja Santhi Velayudhan, Siu Sylvia Lee. SET-9 and SET-26 are H3K4me3 readers and play critical roles in germline development and longevity. eLife 2018;7:e34970 DOI: 10.7554/eLife.34970"

methyltransferases and demethylases can play key roles in regulating germline functions and/or modulate longevity (Han and Brunet 2012; Greer and Shi 2012). In *C. elegans*, loss of the COMPASS complex, which is critical for methylating histone H3 lysine 4, results in a global decrease in H3K4 trimethylation (H3K4me3) levels, and a reduced brood size (Li and Kelly 2011; Robert et al. 2014) and extended lifespan (Greer et al. 2010) phenotypes. Interestingly, deletions of *spr-5* or *rbr-2*, both of which encode demethylases that erase methyl marks on H3K4, also result in fertility defects (Alvares et al. 2014; Katz et al. 2009) and altered lifespan (Alvares et al. 2014; Greer et al. 2010). These findings suggest that histone methylations need to be precisely controlled to maintain longevity and germline function. It is important to note that the molecular mechanisms whereby these H3K4 modifying enzymes effect reproduction and longevity functions, including the genomic regions that they act on and how the altered H3K4 methylation levels contribute to the biological outcomes in these mutants are largely unknown.

SET (Su(var)3-9, Enhancer of Zeste, Trithorax) domain-containing proteins represent a major group of histone methyltransferases (Dillon et al. 2005). We previously carried out a targeted RNAi screen to identify the SET-domain containing proteins that play a role in longevity in *C. elegans*. We found that RNAi knockdown of *set-9* and *set-26*, two closely related genes, results in significant lifespan extension (Ni et al. 2012). SET-9 and SET-26 share 96% sequence identity and both proteins contain a highly conserved PHD and SET domain. PHD domains are known to bind to specific histone modifications (Shi et al. 2007, 2006), suggesting that SET-9 and SET-26 could be recruited to chromatin via binding to specific histone marks. Interestingly, the SET

domain of SET-9 & SET-26 contains mutations in conserved residues thought to be key for methylating activities (Ni et al. 2012), making it unclear whether SET-9 & SET-26 could be active enzymes. Nevertheless, a recent study reported that the SET domain of SET-26 exhibits H3K9me3 activity *in vitro* (Greer et al. 2014).

In this work, we demonstrated that, despite their high sequence identity, SET-26, but not SET-9, plays a key role in stress response and longevity. In addition, we revealed a novel redundant function of SET-9 and SET-26 in germline development. We also confirmed that SET-26 is broadly expressed, whereas SET-9 is only expressed in the germline, which likely account for their distinct and redundant functions in lifespan and reproduction. Indeed, genetic and transcriptomic analyses supported the notion that SET-26 acts through the FOXO transcription factor DAF-16 in the soma to modulate longevity. Furthermore, we showed that the PHD domain of SET-9 & SET-26 binds to H3K4me3 *in vitro* and that the genome-wide binding patterns of SET-9 and SET-26 are highly concordant with that of H3K4me3 marking in *C. elegans*, indicating that SET-9 and SET-26 are recruited to H3K4me3 marked regions *in vivo*. Although the SET domain of SET-26 was reported to methylate H3K9me3 *in vitro* (Greer et al. 2014), our results indicated that loss of *set-9* & *set-26* does not affect global H3K9me3 levels and the genome-wide binding patterns of SET-9 and SET-26 are highly divergent from that of H3K9me3. Instead, we found that loss of *set-9* & *set-26* results in expansion of H3K4me3 marking surrounding most if not all of the SET-9 and SET-26 binding sites, likely specifically in the germline, and significant RNA expression change of a subset of the SET-9 & SET-26 bound genes. We propose that SET-9 and SET-26 are recruited to the chromatin via binding to H3K4me3, where they function to restrict H3K4me3

spreading and to regulate the expression of specific genes, and together these activities contribute to the proper maintenance of germline development.

## Results

### *set-26*, but not *set-9*, single mutant exhibits prolonged lifespan and heightened resistance to heat stress

RNAi knockdown of the highly similar paralogs *set-26* and *set-9* were previously shown to significantly extend lifespan in *C. elegans* (Greer et al. 2010; Ni et al. 2012). However, due to their high sequence similarity, RNAi likely knocks down both *set-26* and *set-9* in those experiments. We confirmed the lifespan extension phenotype with multiple available *set-26* single mutants, but a *set-9* single mutant was not available at the time (Ni et al. 2012). To delineate whether *set-9*, like *set-26*, also plays a role in lifespan determination, we used CRISPR-cas9 to generate a *set-9* mutant (Figure 1A). The *set-9* mutant we obtained carries a mutation that causes a premature stop codon and is expected to produce a truncated SET-9 protein lacking the conserved PHD and SET domains (Figure 1A). We tested the lifespan phenotype of this *set-9* single mutant along with the *set-26* single and *set-9 set-26* double mutants. Consistent with previous results, the *set-26* single mutant lived longer than wild-type worms (Figure 1B). Surprisingly, although SET-9 and SET-26 proteins share 97% identity in protein sequence, the *set-9(rw5)* mutation did not alter lifespan in either wild-type or the *set-26* mutant background (Figure 1B). Similar to the lifespan phenotype, *set-26*, but not *set-9* single mutant, was more resistant to heat stress compared to wild-type worms (Figure

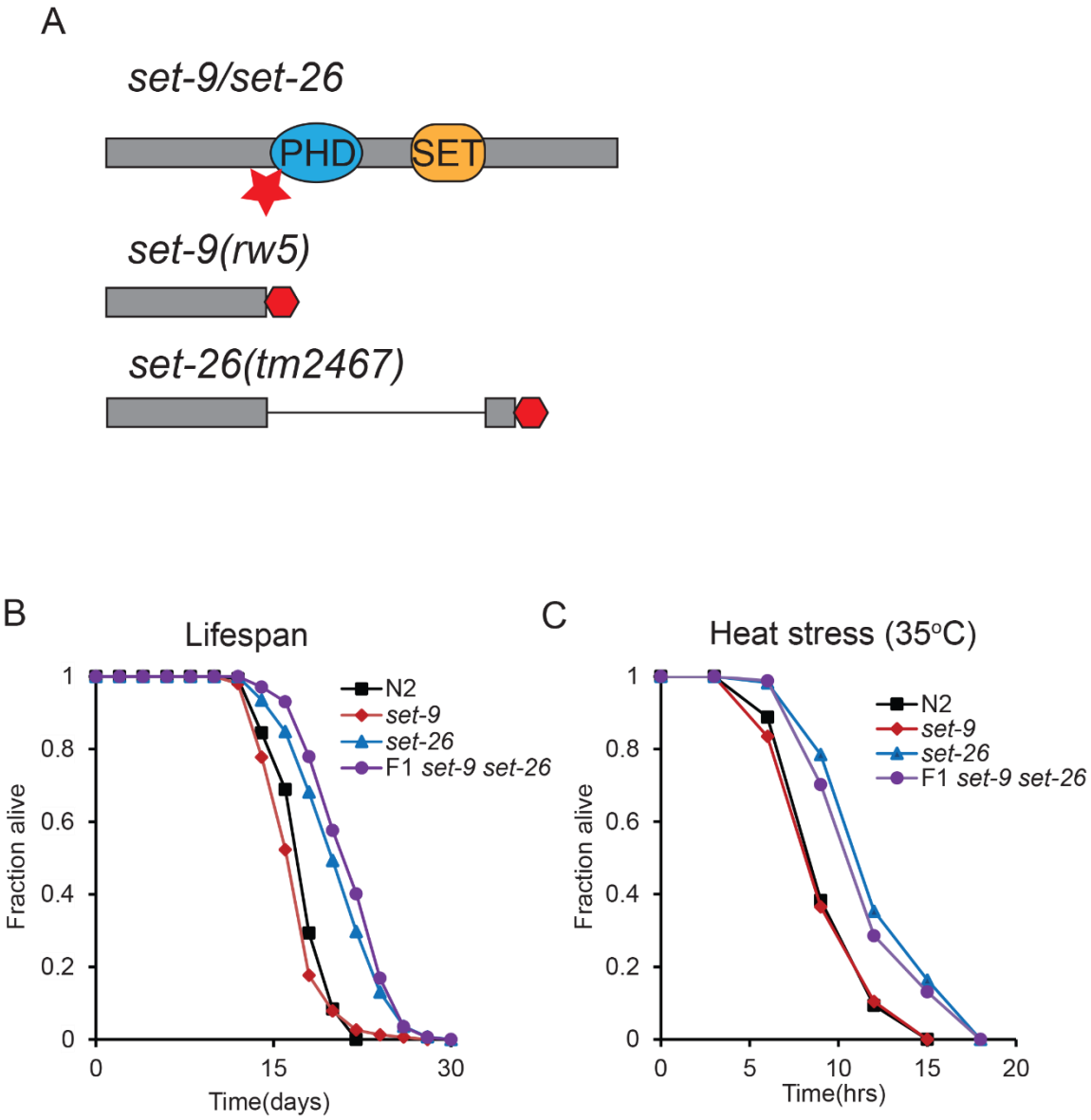

A
set-9/set-26
PHD
SET
set-9(rw5)
set-26(tm2467)
B
Lifespan
Fraction alive
Time(days)
N2
set-9
set-26
F1 set-9 set-26
C
Heat stress (35°C)
Fraction alive
Time(hrs)
N2
set-9
set-26
F1 set-9 set-26

Figure 1 *set-26* but not *set-9* is important for longevity.

(A) Schematic of the *set-9(rw5)* and *set-26(tm2467)* mutants. Red star indicates the position of the sgRNA (single guide RNA) targeting the *set-9* gene. Premature stop codons caused by deletions of 38 nucleotides in the *set-9* gene and 1,090 nucleotides in the *set-26* gene are depicted as red hexagons. Loss of *set-26* gene but not *set-9* gene extended lifespan (B), and increased resistance to heat stress (C). Survival curves for N2, *set-26(tm2467)*, *set-9(rw5)*, and *set-9(rw5) set-26(tm2467)* strains from representative experiments are shown. Quantitative data for all replicates are shown in Table S1.

A

Lifespan experiment 1

| Strains | # of worms | Mean +/- SD | p value vs N2 | p value vs set-26 |
|---|---|---|---|---|
| N2 | 46/57 | 16.96 +/- 0.35 | | |
| *set-9* | 48/60 | 17.27 +/- 0.37 | 1 | |
| *set-26* | 57/66 | 20.60 +/- 0.48 | 3.80E-08 | |
| *set-9 set-26* | 59/66 | 21.25 +/- 0.46 | 0.00E+00 | 1 |

Lifespan experiment 2

| Strains | # of worms | Mean +/- SD | p value vs N2 | p value vs *set-26* |
|---|---|---|---|---|
| N2 | 52/65 | 18.62 +/- 0.25 | | |
| *set-9* | 44/67 | 18.29 +/- 0.43 | 1 | |
| *set-26* | 43/72 | 21.35 +/- 0.53 | 5.00E-06 | |
| *set-9 set-26* | 42/67 | 23.14 +/- 0.42 | 0.00E+00 | 0.1797 |

Lifespan experiment 3

| Strains | # of worms | Mean +/- SD | p value vs N2 | p value vs *set-26* |
|---|---|---|---|---|
| N2 | 69/73 | 17.77 +/- 0.29 | | |
| *set-9* | 61/66 | 16.39 +/- 0.27 | 0.0016 | |
| *set-26* | 38/41 | 20.74 +/- 0.59 | 0.0000024 | |
| *set-9 set-26* | 71/74 | 22.64 +/- 0.38 | 0 | 0.0554 |

B

35°C heat resistance experiment 1

| Strains | # of worms | Mean +/- SD | p value vs N2 | p value vs *set-26* |
|---|---|---|---|---|
| N2 | 64 | 11.02 +/- 0.32 | | |
| *set-9* | 57 | 11.47 +/- 0.33 | 0.9591 | |
| *set-26* | 53 | 14.72 +/- 0.43 | 0 | |
| *set-9 set-26* | 27 | 15.22 +/- 0.59 | 1.2E-08 | 1 |

35°C heat resistance experiment 2

| Strains | # of worms | Mean +/- SD | p value vs N2 | p value vs *set-26* |
|---|---|---|---|---|
| N2 | 43 | 8.72 +/- 0.19 | | |
| *set-9* | 58 | 8.38 +/- 0.20 | 0.7592 | |
| *set-26* | 63 | 11.29 +/- 0.24 | 0 | |
| *set-9 set-26* | 57 | 10.95 +/- 0.23 | 6E-09 | 0.785 |

Table S1

1C). These results suggested that inactivation of *set-26,* but not *set-9,* extends lifespan and improves heat resistance.

### ***set-9*** **and** ***set-26*** **act redundantly to maintain germline function**

While propagating the *set-9 set-26* double mutant, we noticed a possible fertility defect. To more thoroughly assess the roles of SET-9 and SET-26 in reproduction, we assayed the brood size of *set-9, set-26* single and double mutants. We found that the progeny number produced by *set-9* and *set-26* single mutants was slightly smaller compared to that of wild-type worms (Figure 2A). Interestingly, the homozygous *set-9 set-26* double mutant derived from heterozygous parents (first generation, i.e. F1) also exhibited a mild brood size defect, and this defect became significantly more severe in the second and later generations (F2 to F6, Figure 2A). Deficiency of several histone modifiers has been previously reported to exhibit a “mortal germline” phenotype. We performed a classical mortal germline assay and found that the *set-9 set-26* double mutant indeed displayed a mortal germline phenotype (Figure 2C). Further detailed analyses indicated that the *set-9 set-26* double mutant exhibited a high sterile rate and a low brood size through the F2-F6 generations that we assayed (Figure 2A and Figure 2-figure supplement 1A).

Since we noted a large difference between the brood size of the *set-9 set-26* double mutant in the F1 and F2 generations and suspected a possible maternal influence, we therefore performed a series of crosses to test this possibility. We found that *set-9 set-26* double mutants derived from homozygous *set-9 set-26* hermaphrodites crossed with

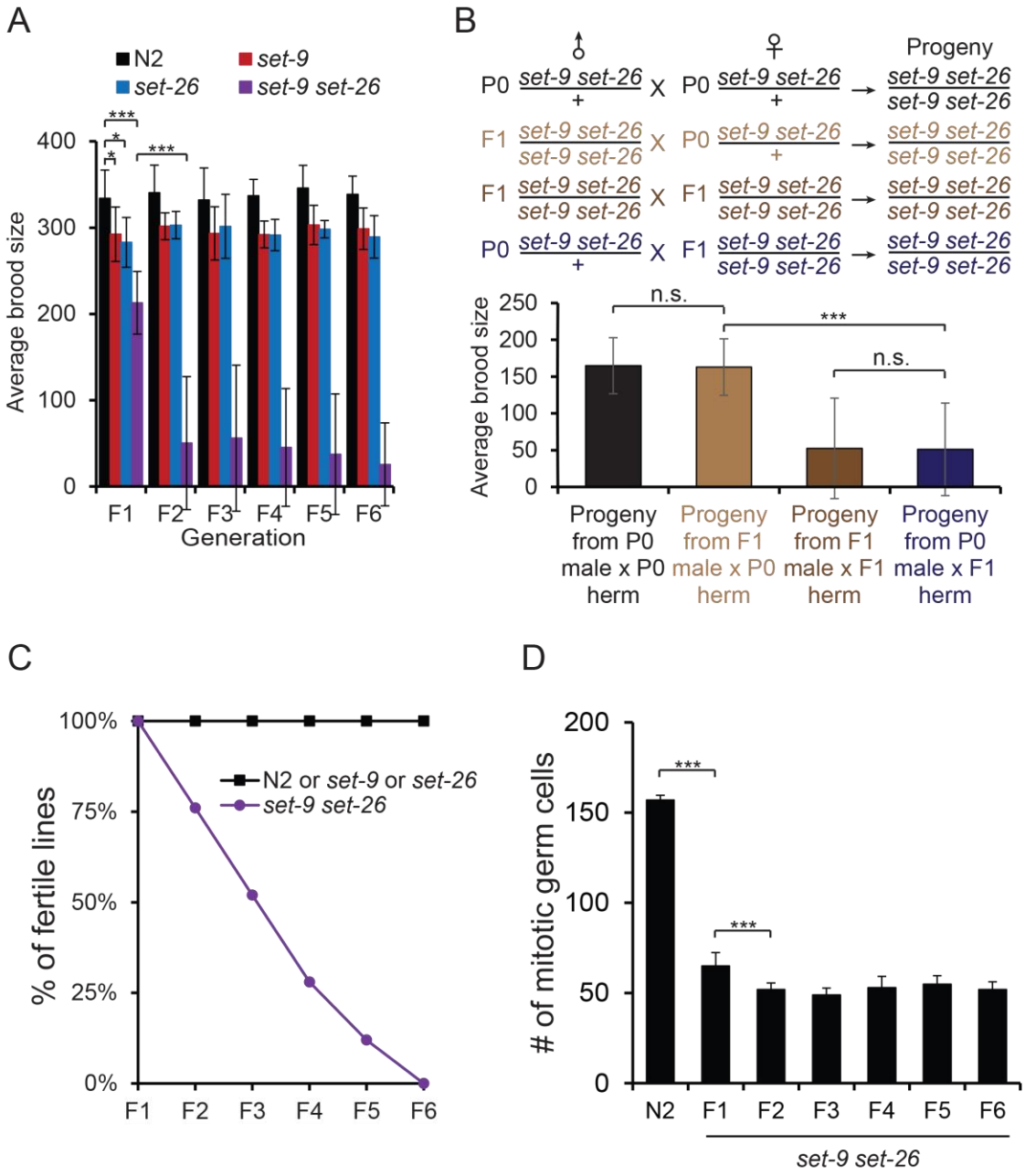
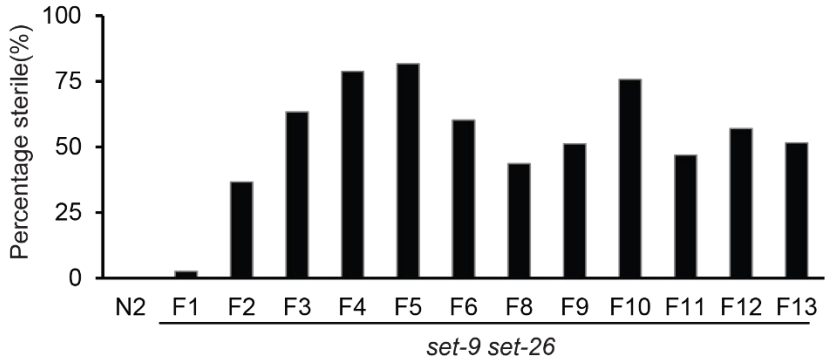

A
N2
set-9
set-26
set-9 set-26
Average brood size
Generation
F1
F2
F3
F4
F5
F6
B
Progeny
P0 set-9 set-26 / + X P0 set-9 set-26 / + → set-9 set-26 / set-9 set-26
F1 set-9 set-26 / set-9 set-26 X P0 set-9 set-26 / + → set-9 set-26 / set-9 set-26
F1 set-9 set-26 / set-9 set-26 X F1 set-9 set-26 / set-9 set-26 → set-9 set-26 / set-9 set-26
P0 set-9 set-26 / + X F1 set-9 set-26 / set-9 set-26 → set-9 set-26 / set-9 set-26
Average brood size
n.s.
***
n.s.
Progeny from P0 male x P0 herm
Progeny from F1 male x P0 herm
Progeny from F1 male x F1 herm
Progeny from P0 male x F1 herm
C
% of fertile lines
N2 or set-9 or set-26
set-9 set-26
D
# of mitotic germ cells
N2
set-9 set-26

Figure 2 *set-9* and *set-26* act redundantly to maintain fertility.

(A) The *set-9(rw5) set26(tm2467)* double mutant worms derived from heterozygous parents (F1) displayed a mild fertility defect. The double mutant worms displayed a much more severe fertility defect at later generations (F2-F6). Average brood size of N2, *set-26(tm2467), set-9(rw5),* and *set-9(rw5) set-26(tm2467)* strains at the indicated generation were shown (*p<0.05, ***p<0.001). The error bars represent standard errors. n=9~10 for N2, *set-9* mutant, *set-26* mutant, and F1 *set-9 set-26* double mutant worms; n=~50 for F2-F6 *set-9 set-26* double mutants. (B) Maternal contribution of *set-9* & *set-26* appeared important for alleviating the fertility defect in the double mutant. Average brood size of the *set-9(rw5) set26(tm2467)* double mutants derived from four different crosses were shown (***p<0.001, n.s. no significant). n=11~12 for assessing the brood size of the *set-9(rw5) set26(tm2467)* homozyous progeny from heterozygous male(P0) X heterozygous hermaphrodite(P0) and homozygous male(F1) X heterozygous hermaphrodite(P0); n=30~34 for progeny from homozygous male (F1) X heterozygous hermaphrodite(P0) and heterozygous male(P0) X homozygous hermaphrodite(F1).

(C) The *set-9(rw5) set26(tm2467)* double mutant exhibited a mortal germline phenotype. At each generation, 6 L1s for N2, *set-26(tm2467)*, *set-9(rw5)* and *set-9(rw5) set-26(tm2467)* strains were transferred to a new plate. Plates were scored as not fertile when no progeny were found. % of fertile lines indicated percentage of plates that were fertile. n=6 for N2, *set-9* and *set-26* mutants; n=25 for *set-9 set-26* double mutants. (D) The *set-9(rw5) set26(tm2467)* double mutant worms that remained fertile nevertheless exhibited reduced number of mitotic germ cells. Whole worms or dissected gonads of fertile *set-9(rw5) set26(tm2467)* mutants were stained by DAPI and the mitotic cells were counted. D2 adults were scored. n=18~27, ***p<0.001. Analyses of sterile *set-9(rw5) set26(tm2467)* double mutant worms are shown in Figure 2-figure supplement 1. Quantitative data are shown in Table S2.

A


Percentage sterile(%)
100
75
50
25
0
N2
F1
F2
F3
F4
F5
F6
F8
F9
F10
F11
F12
F13
set-9 set-26

B

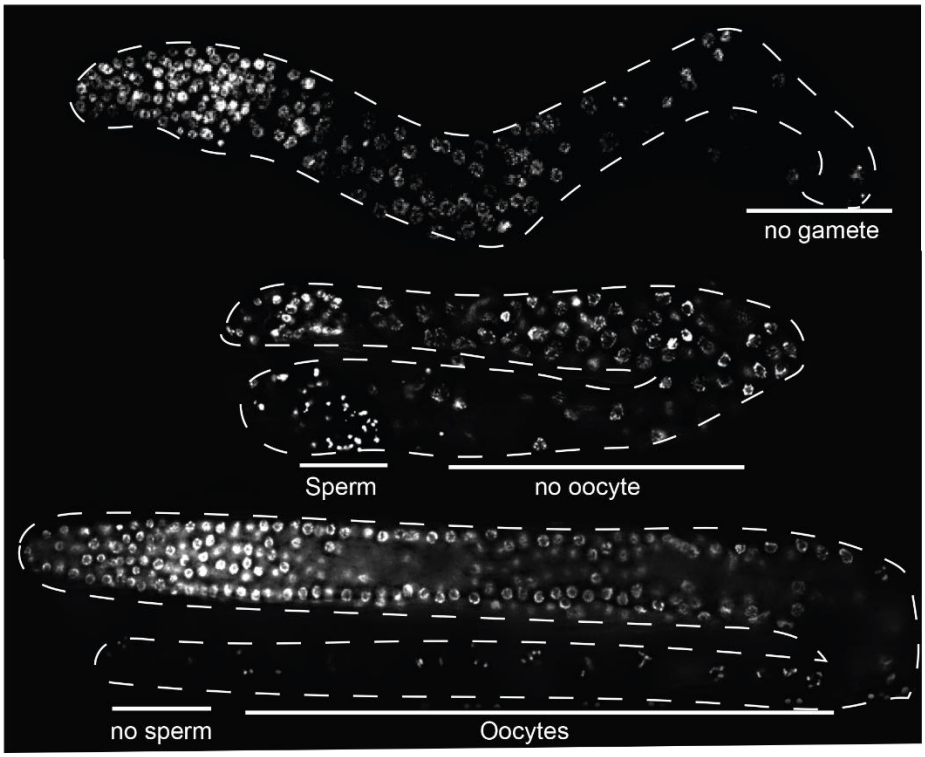

no gamete
Sperm
no oocyte
no sperm
Oocytes

Figure 2-figure supplement 1 Fertility defects of the *set-9 set-26* double mutant.

(A) The *set-9(rw5) set-26(tm2467)* double mutant exhibited variable percentage of sterility at the different generations. N=155-517. (B) Representative DAPI staining images showing the germline of three different sterile F3-F4 *set-9(rw5) set-26(tm2467)* double mutant worms. The germline of these sterile worms showed variable phenotypes, including no differentiated germ cells, no oocytes, or no sperms. D1-D2 adults were scored.

A

Brood size experiment

| | F1 | F2 | F3 | F4 | F5 | F6 |
|---|---|---|---|---|---|---|
| N2 | 334 +/- 33 | 340 +/- 32 | 332 +/- 37 | 337 +/- 19 | 346 +/- 26 | 339 +/- 21 |
| *set-9* | 292 +/- 32 | 302 +/- 16 | 293 +/- 31 | 292 +/- 16 | 303 +/- 23 | 299 +/- 24 |
| *set-26* | 283 +/- 29 | 303 +/- 16 | 302 +/- 37 | 291 +/- 18 | 298 +/- 10 | 289 +/- 25 |
| *set-9 set-26* | 213 +/- 36 | 50 +/- 77 | 56 +/- 85 | 45 +/- 68 | 37 +/- 70 | 26 +/- 48 |

B

Maternal effect brood size experiment

| Progeny | Mean +/- SD |
|---|---|
| Progeny from P0 male x P0 herm | 165 +/- 38 |
| Progeny from F1 male x P0 herm | 163 +/- 39 |
| Progeny from F1 male x F1 herm | 52 +/- 68 |
| Progeny from P0 male x F1 herm | 51 +/- 63 |

All crosses were tracked by *sur-5p::gfp*

C

Mortal germline assay

| | F1 | F2 | F3 | F4 | F5 | F6 |
|---|---|---|---|---|---|---|
| N2 | 100% | 100% | 100% | 100% | 100% | 100% |
| *set-9* | 100% | 100% | 100% | 100% | 100% | 100% |
| *set-26* | 100% | 100% | 100% | 100% | 100% | 100% |
| *set-9 set-26* | 100% | 76% | 52% | 28% | 12% | 0% |

D

Mitotic cells counting

| Strain | Mean +/- SD |
|---|---|
| N2 | 157 +/- 2.51 |
| F1 *set-9 set-26* | 65 +/- 7.4 |
| F2 *set-9 set-26* | 52 +/- 3.6 |
| F3 *set-9 set-26* | 49 +/- 3.7 |
| F4 *set-9 set-26* | 53 +/- 6.2 |
| F5 *set-9 set-26* | 55 +/- 4.6 |
| F6 *set-9 set-26* | 52 +/- 4.2 |

Table S2

heterozygous fathers exhibited a significantly more severe brood size defect compared to those from heterozygous hermaphrodites crossed with homozygous mutant fathers (Figure 2B). In other words, heterozygous mothers, but not heterozygous fathers, helped to maintain better germline function in the progeny. These data support the notion of a maternal contribution in germline maintenance in the *set-9 set-26* double mutant worms.

We next used DAPI staining to monitor the germ cells of the *set-9 set-26* double mutant at the F3 and F4 generations. For the double mutant worms that become sterile, we observed variable germline phenotypes, including a very small mitotic region with no differentiated cells, and a small mitotic region with sperms only or a largely normal mitotic region with oocytes only (Figure 2-figure supplement 1B), suggesting problems with both the germline stem cells and their subsequent differentiation. For the double mutant worms that remained fertile, we observed germlines with a smaller but stable number of mitotic cells (Figure 2D). The results together indicated that SET-9 and SET-26 act redundantly to maintain normal germline function and they may regulate both the proliferation and differentiation of the germline stem cells.

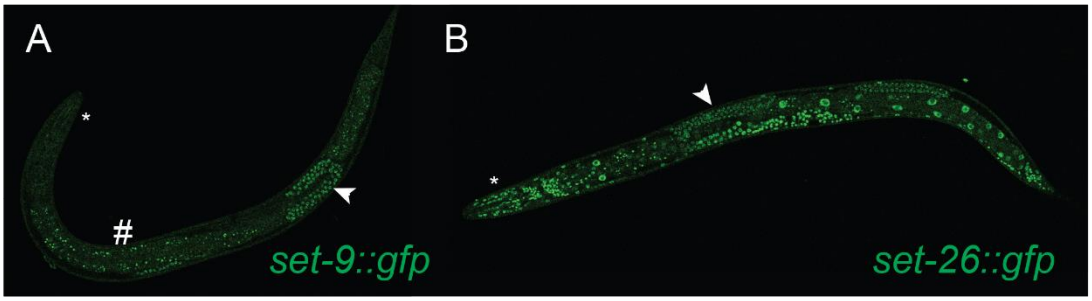
**SET-26 is broadly expressed but SET-9 is only detectable in the germline**

Given the high degree of sequence identity between SET-9 and SET-26, and given their differential roles in lifespan and heat resistance, we wondered whether these two proteins could be expressed in different tissues. In an attempt to resolve the expression patterns of SET-9 and SET-26, we previously used RT-PCR at precise temperatures, as

well as an antibody that recognized both SET-9 and SET-26, in wild-type, *set-26* single, and germlineless mutant worms, and deduced that SET-26 is likely broadly expressed and SET-9 is likely expressed in the germline (Ni et al. 2012). To unambiguously determine the expression patterns of the SET-9 and SET-26 proteins, we used CRISPR-cas9 to knock-in a GFP tag at the C-terminus of the endogenous *set-9* and *set-26* loci and monitored their expression patterns. Consistent with our previous report (Ni et al. 2012), we found that GFP-tagged SET-26 was broadly expressed in both the somatic and germline cells of *C. elegans* (Figure 3B). In contrast, the GFP-tagged SET-9 was only detected in germline cells (Figure 3A). As expected, expression of these two proteins was restricted to the nucleus, which is consistent with their possible roles in chromatin regulation. The ubiquitous expression of SET-26, but not SET-9, likely explains why SET-26 alone has a role in lifespan and heat resistance.

We note that the knock-in worms expressing GFP-tagged SET-26 lived slightly longer than wild-type (but significantly shorter than the *set-26* mutant) (Figure 3-figure supplement 1A) and had a slight heat resistance phenotype (Figure 3-figure supplement 1B), and the knock-in worms expressing both SET-9::GFP and SET-26::GFP had a slightly lower brood size compared to wild-type worms, but a significantly larger brood size than the *set-9 set-26* double mutant worms (Figure 3-figure supplement 1C). The data together suggested that the GFP-tag somewhat compromised the functions of SET-9 and SET-26, but the tagged proteins remain largely functional.

We next wondered whether the germline or somatic expression of SET-26 is important for lifespan modulation. We previously showed that RNAi knockdown of *set-9/-26* (RNAi targets the two genes due to high sequence identity) in *glp-1(e2141)* germlineless


A
B
*
#
*
set-9::gfp
set-26::gfp

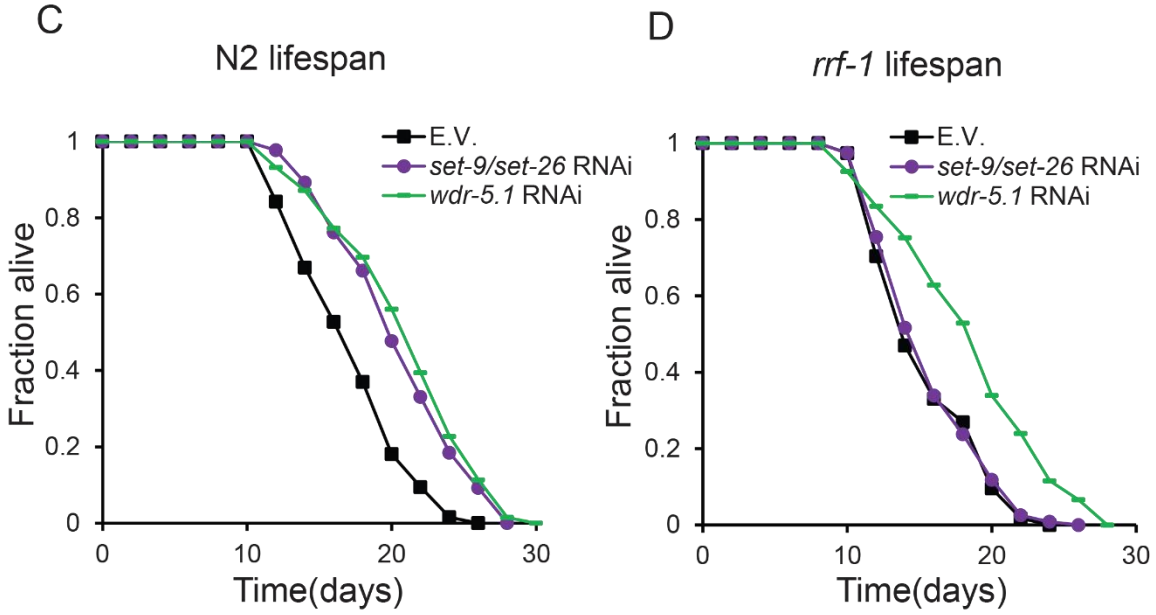
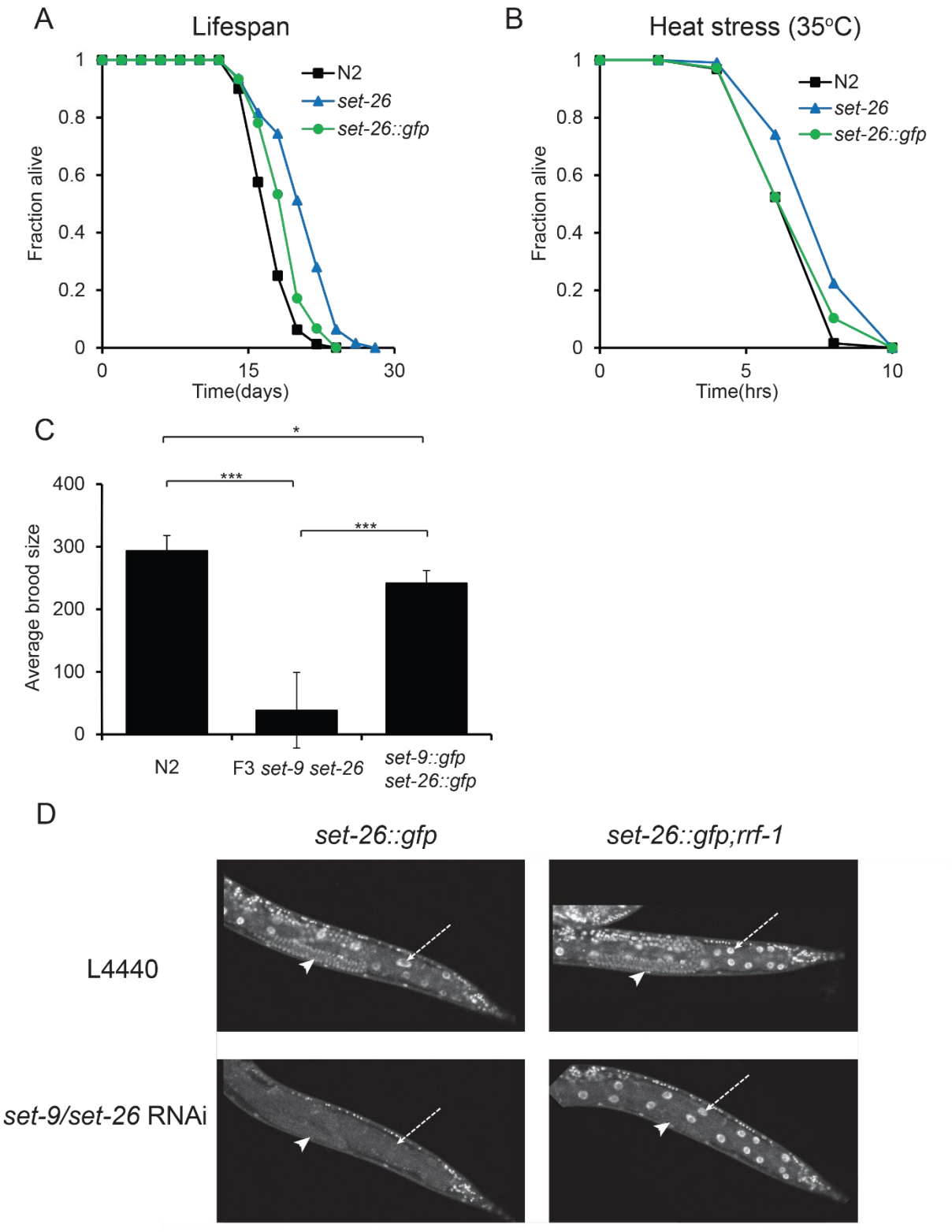

C
N2 lifespan
E.V.
set-9/set-26 RNAi
wdr-5.1 RNAi
Fraction alive
Time(days)
D
rrf-1 lifespan
E.V.
set-9/set-26 RNAi
wdr-5.1 RNAi
Fraction alive
Time(days)

Figure 3 SET-26 is broadly expressed and SET-9 is only detectable in the germline.

(A, B) Fluorescent micrographs of worms carrying gfp knock-in at the C-terminus of *set-9* or *set-26* gene (*set-9::gfp* and *set-26::gfp*). GFP-fused SET-26 was detected in all cells and GFP-fused SET-9 was only detected in the germline. Star indicates head, arrow indicates germline in the images. The signal outside of the germline detected in the *set-9::gfp* strain represented autofluroescence (marked by hashtag), which appeared yellow under the microscope. (C, D) Germline-specific knockdown of *set-9* & *set-26* was not sufficient to extend lifespan. RNAi Knockdown of *set-9* & *set-26* or *wdr-5.1* extended lifespan in N2 worms (C). RNAi knockdown of *wdr-5.1*, but not *set-9* & *set-26*, extended lifespan in the *rrf-1(pk1417)* mutant worms. Quantitative data are shown in Table S3.


A
Lifespan
Fraction alive
Time(days)
N2
set-26
set-26::gfp
B
Heat stress (35°C)
Fraction alive
Time(hrs)
N2
set-26
set-26::gfp
C
Average brood size
N2
F3 set-9 set-26
set-9::gfp
set-26::gfp
D
set-26::gfp
set-26::gfp;rrf-1
L4440
set-9/set-26 RNAi

## Figure 3-figure supplement 1 GFP-tagged SET-9 & SET-26 are largely functional.

(A-B) Worms carrying a *gfp* knockin at the C-terminus of *set-26* exhibited mild lifespan extension and heat resistance phenotypes, but the phenotypes were much weaker than those of the loss-of-function *set-26(tm2467)* mutant. (C) Worms carrying *gfp* knockins at the C-terminal of *set-9* and *set-26* genes exhibited a mild fertility defect, which was significantly different from the severe fertility defect observed in the *set-9(rw5) set-26(tm2467)* double mutant (*p<0.05, ***p<0.001). n=8-10 for *set-9::gfp set-26::gfp* and N2; n=28 for F3 *set-9(rw5) set-26(tm2467).* Quantitative data are shown in Table S4. (D) SET-26::GFP expression of RNAi Knockdown of *set-9 & set-26* and L4440 in the *set-26::gfp; rrf-1(pk1417)* and *set-26::gfp* strains are shown.

Table S3

A

N2 lifespan experiment 1

| Condition | # of worms | Mean +/- SD | p value vs N2+L4440 |
|---|---|---|---|
| N2+L4440 | 65/70 | 18.18 +/- 0.41 | |
| N2+*set-9/26* RNAi | 66/70 | 21.49 +/- 0.50 | 0.0000012 |
| N2+*wdr-5.1* RNAi | 64/67 | 21.63 +/- 0.61 | 0.00000066 |

N2 lifespan experiment 2

| Condition | # of worms | Mean +/- SD | p value vs N2+L4440 |
|---|---|---|---|
| N2+L4440 | 62/67 | 16.61 +/- 0.52 | |
| N2+*set-9/26* RNAi | 64/67 | 20.00 +/- 0.55 | 0.0001 |
| N2+*wdr-5.1* RNAi | 68/72 | 20.76 +/- 0.56 | 0.00000099 |

B

*rrf-1(pk1417)* lifespan experiment 1

| Condition | # of worms | Mean +/- SD | p value vs rrf-1+L4440 |
|---|---|---|---|
| *rrf-1* +L4440 | 56/70 | 16.84 +/- 0.47 | |
| *rrf-1+set-9/26* RNAi | 59/70 | 16.50 +/- 0.46 | 1 |
| *rrf-1+wdr-5.1* RNAi | 60/69 | 20.32 +/- 0.68 | 0.0000085 |

*rrf-1(pk1417)* lifespan experiment 2

| Condition | # of worms | Mean +/- SD | p value vs rrf-1+L4440 |
|---|---|---|---|
| *rrf-1* +L4440 | 59/67 | 14.79 +/- 0.44 | |
| *rrf-1+set-9/26* RNAi | 59/66 | 15.43 +/- 0.46 | 0.6413 |
| *rrf-1+wdr-5.1* RNAi | 61/67 | 17.58 +/- 0.56 | 0.0004 |

A

*set-26::gfp* lifespan experiment 1

| Strains | # of worms | Mean +/- SD | p value vs N2 | p value vs *set-26* |
|---|---|---|---|---|
| N2 | 45/58 | 17.51 +/- 0.33 | | |
| *set-26* | 55/61 | 20.62 +/- 0.46 | 0.00000075 | |
| *set-26::gfp* | 52/61 | 18.77 +/- 0.37 | 0.0195 | 0.0024 |

*set-26::gfp* lifespan experiment 2

| Strains | # of worms | Mean +/- SD | p value vs N2 | p value vs *set-26* |
|---|---|---|---|---|
| N2 | 35/49 | 17.71 +/- 0.35 | | |
| *set-26* | 70/86 | 20.88 +/- 0.42 | 0.000000054 | |
| *set-26::gfp* | 53/80 | 19.25 +/- 0.34 | 0.0042 | 0.0006 |

B

*set-26::gfp* 35°C heat stress experiment 1

| Strains | # of worms | Mean +/- SD | p value vs N2 | p value vs *set-26* |
|---|---|---|---|---|
| N2 | 61 | 7.02 +/- 0.15 | | |
| *set-26* | 60 | 7.97 +/- 0.19 | 0.0002 | |
| *set-26::gfp* | 52 | 7.15 +/- 0.19 | 1 | 0.0065 |

*set-26::gfp* 35°C heat stress experiment 2

| Strains | # of worms | Mean +/- SD | p value vs N2 | p value vs *set-26* |
|---|---|---|---|---|
| N2 | 67 | 7.01 +/- 0.14 | | |
| *set-26* | 56 | 7.86 +/- 0.18 | 0.0008 | |
| *set-26::gfp* | 55 | 7.24 +/- 0.20 | 0.5832 | 0.0691 |

C

*set-9::gfp set-26::gfp* brood size experiment

| | Mean +/- SD |
|---|---|
| N2 | 294 +/- 24 |
| F3 *set-9 set-26* | 39 +/- 60 |
| *set-9::gfp set-26::gfp* | 242 +/- 20 |

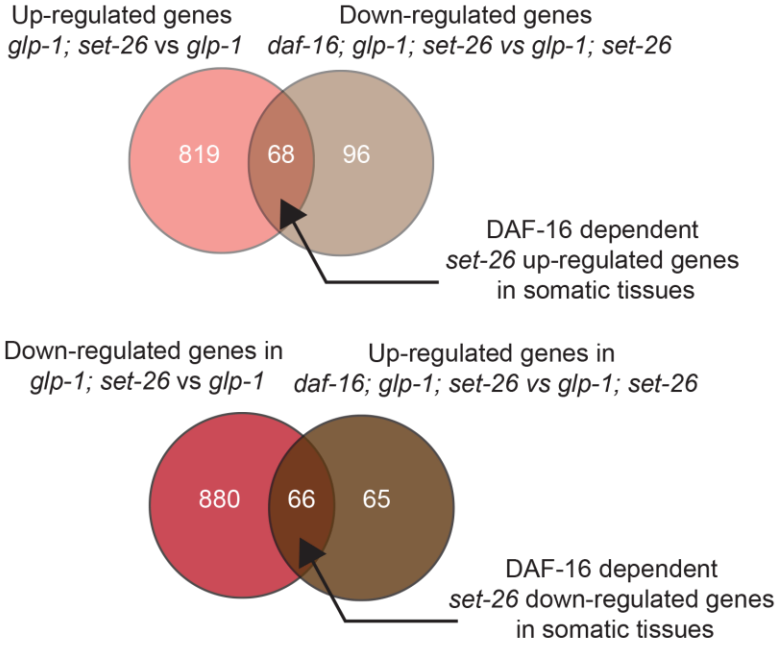
Table S4

mutant worms extended lifespan to a similar degree as in wild-type worms (Ni et al. 2012), suggesting that somatic *set-26* is important for lifespan modulation. To further test this possibility, we used the *rrf-1* mutant, in which RNAi is efficient in the germline but not somatic cells (Sijen et al. 2001), to assess whether knockdown of *set-26* (*& set-9*) in the germline alone can extend lifespan. As a control for tissue-specific RNAi, we monitored SET-26::GFP expression in wild-type or *rrf-1* mutant worms treated with *set-9/-26* RNAi. As expected, *set-9/-26* RNAi greatly reduced SET-26::GFP expression in most tissues except neurons in wild-type worms (Figure 3-figure supplement 1D), whereas *set-9/-26* RNAi treatment specifically knocked down SET-26::GFP expression in the germline in *rrf-1* mutant worms (Figure 3-figure supplement 1D). We next assessed the lifespan of wild-type or *rrf-1* mutant worms treated with *set-9/-26* RNAi. We included *wdr-5.1* RNAi as a positive control as *wdr-5.1* is known to act in the germline to modulate lifespan (Greer et al. 2010). As expected, RNAi knockdown of *wdr-5.1* extended lifespan in both wild-type and *rrf-1* mutant worms. In contrast, *set-9/-26* RNAi knockdown extended lifespan in wild-type but not in the *rrf-1* mutant background (Figure 3C and 3D), indicating that inactivation of *set-26* (*& set-9*) in the germline is not sufficient for lifespan modulation. These results corroborated with our previous findings, and indicated that SET-26 likely act in the somatic cells to modulate longevity and heat stress response, but SET-9 and SET-26 act redundantly in the germline to maintain reproductive function.

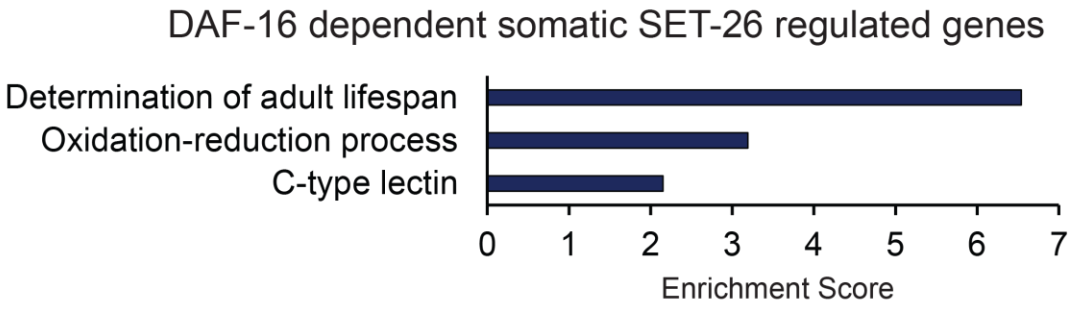
**Transcriptional profiling revealed candidate longevity and germline function genes regulated by SET-9 and SET-26**

To gain insights into the molecular changes that may contribute to the somatic SET-26 effect on lifespan, we investigated the transcriptional profiles of the long-lived germlineless *glp-1; set-26* double mutant. We isolated total RNA from *glp-1; set-26* double and *glp-1* single mutant worms, and performed RNA sequencing after removing ribosomal RNAs (ribo-minus RNA-seq). We next used edgeR, an RNA-seq analysis tool in the R package (Robinson et al. 2010), to identify the genes that showed statistically significant expression change in the *glp-1; set-26* double mutant compared to *glp-1* mutant (Figure 4A). We identified 887 up-regulated and 946 down-regulated genes in response to *set-26* loss in the soma (Figure 4A), and gene ontology (GO) analyses indicated that these genes were over-represented in multiple functional groups (Figure 4-figure supplement 1A). Since we previously showed that somatic *set-26* largely acts through *daf-16*, which encodes the Forkhead box O (FOXO) transcription factor, to modulate lifespan (Ni et al. 2012), we sought to further identify the transcriptional changes in response to somatic *set-26* loss that are also dependent on *daf-16*. Using similar RNA-seq experiments, we investigated the transcriptional profiles of *daf-16; glp-1; set-26* triple and *glp-1; set-26* double mutants (Figure 4A). We identified 164 genes that were up-regulated, and 131 genes that were down-regulated in the *daf-16; glp-1; set-26* triple mutant (Figure 4A). By comparing these gene lists with the gene lists discussed above for the *glp-1; set-26* double mutant vs. *glp-1*, we deduced the somatic genes whose expression become significantly up-regulated or down-regulated when *set-26* is deleted, but those expression changes were reverted when *daf-16* is simultaneously lost (down-regulated or up-regulated in the *daf-16; glp-1; set-26* triple mutant, respectively) (Figure 4A). We termed these DAF-16-depednent somatic SET-26

A


Up-regulated genes
glp-1; set-26 vs glp-1
Down-regulated genes
daf-16; glp-1; set-26 vs glp-1; set-26
819
68
96
DAF-16 dependent
set-26 up-regulated genes
in somatic tissues
Down-regulated genes in
glp-1; set-26 vs glp-1
Up-regulated genes in
daf-16; glp-1; set-26 vs glp-1; set-26
880
66
65
DAF-16 dependent
set-26 down-regulated genes
in somatic tissues

B


DAF-16 dependent somatic SET-26 regulated genes
Determination of adult lifespan
Oxidation-reduction process
C-type lectin
0
1
2
3
4
5
6
7
Enrichment Score

## Figure 4 DAF-16 dependent somatic SET-26 regulated genes are enriched for lifespan determinant genes

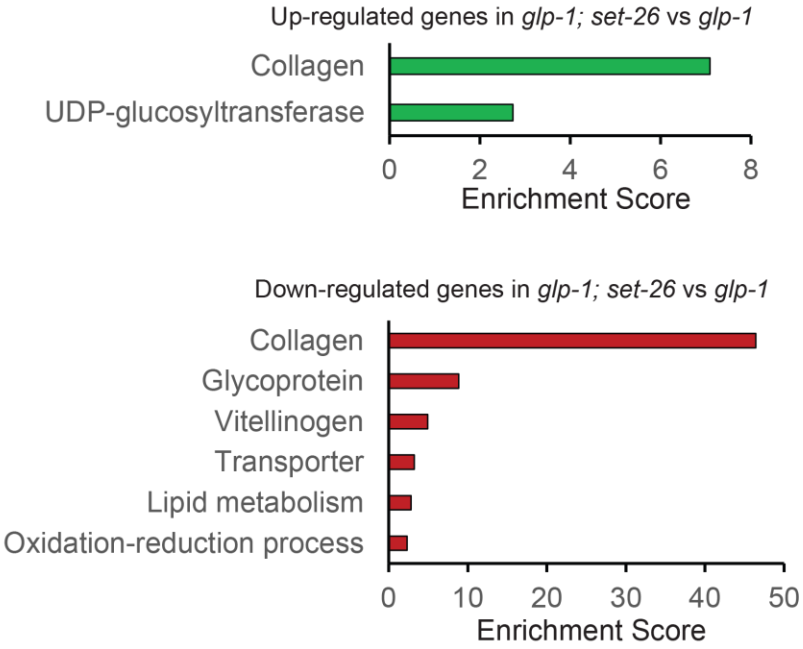
(A) Venn diagrams show the overlap between up-regulated genes in *glp-1(e2141); set-26(tm2467)* (comparing with *glp-1(e2141)*) and down-regulated genes in *daf-16(mgDf47); glp-1(e2141); set-26(tm2467)* (comparing with *glp-1(e2141); set-26(tm2467)*); and the overlap between down-regulated genes in *glp-1(e2141); set-26(tm2467)* (comparing with *glp-1(e2141)*) and up-regulated genes in *daf-16(mgDf47); glp-1(e2141); set-26(tm2467)* (comparing with *glp-1(e2141); set-26(tm2467)*). (B) GO term analysis of DAF-16 dependent somatic SET-26 regulated genes.

regulated genes. Interestingly, GO term analyses revealed that the functional group “determination of adult lifespan” was highly enriched in these DAF-16-dependent somatic SET-26 regulated genes (Figure 4B). Therefore, the transcriptomic analysis corroborated the genetic analysis, and supports a model that DAF-16-mediated gene regulation likely contributes to the lifespan phenotype of the *set-26* mutant.

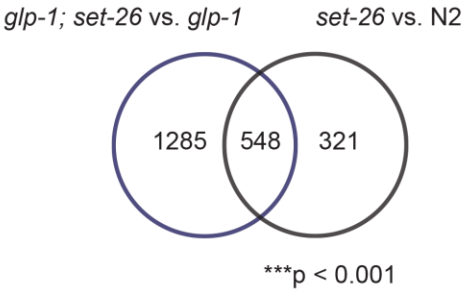
We additionally investigated the transcriptional profiles of the long-lived fertile *set-26* single mutant and revealed that 869 genes showed significant expression change in the *set-26* mutant compared with wild-type worms. As expected, there was a significant and substantial overlap between the genes that exhibited expression change in response to whole-body loss of *set-26* and the somatic SET-26 regulated genes discussed above (Figure 4-figure supplement 1B). Interestingly, the analysis using germlineless worms revealed far greater number of genes with expression change compared to that using whole worms. This could be due to technical variations between experiments, but might also suggest that some genes exhibit selective expression changes only in somatic cells, and those expression changes could be masked when germ cells were included in the analyses.

To gain insights into the molecular changes that may underlie the germline phenotypes, we next compared the transcriptional profiles of the *set-9* single, *set-26* single, and F1 *set-9 set-26* double mutants. We identified 162, 334, 1888 genes that were up-regulated, and 545, 534, 1644 genes that were down-regulated in the *set-9*, *set-26*, and F1 *set-9 set-26* mutants respectively (Figure 5A). Interestingly, although there were significant overlap among the three gene sets, a substantial number of genes appeared to only show expression changes in the F1 *set-9 set-26* double mutant (Figure 5A,

A


Up-regulated genes in glp-1; set-26 vs glp-1
Collagen
UDP-glucosyltransferase
0
2
4
6
8
Enrichment Score
Down-regulated genes in glp-1; set-26 vs glp-1
Collagen
Glycoprotein
Vitellinogen
Transporter
Lipid metabolism
Oxidation-reduction process
0
10
20
30
40
50
Enrichment Score

B


glp-1; set-26 vs. glp-1
set-26 vs. N2
1285
548
321
***p < 0.001

## Figure 4-figure supplement 1 Up-regulated germline function candidate genes are enriched for sperm-specific genes.

(A) GO term analysis of up-regulated genes in *glp-1(e2141); set-26(tm2467)* (comparing with *glp-1(e2141)*) and GO term analysis of down-regulated genes in *glp-1(e2141); set-26(tm2467)* (comparing with *glp-1(e2141)*). (B) Venn diagram shows the overlap between *set-26* regulated genes in *glp-1(e2141); set-26(tm2467)* (comparing with *glp-1(e2141)*) and *set-26* regulated genes in *set-26(tm2467)* (comparing with N2). (C) Venn diagram shows the overlap between DAF-16 dependent somatic SET-26 regulated genes and *set-26* regulated genes (comparing with N2). (D) GO term analysis of DAF-16 dependent somatic SET-26 regulated genes overlaps with *set-26* regulated genes (comparing with N2).

1,430 down-regulated, 1,781 up-regulated), suggesting a redundant role of SET-9 and SET-26 in regulating gene expression. Since the F1 *set-9 set-26* mutant had a mild brood-size phenotype, and gave rise to progeny that exhibited severe defects in germline development (Figure 2), we speculated that many of these SET-9 and SET-26 co-regulated genes could be important for germline function. Indeed, GO analyses revealed that the genes that showed expression change in response to the simultaneous loss of *set-9* and *set-26* were over-represented for a number of different functional groups, including genes with annotated functions in sperm development and function (Figure 5B and 5C). We further compared these SET-9 and SET-26 co-regulated with genes previously determined to be germline-, oocyte-, and sperm-specific (Reinke et al. 2004). Interestingly, we found a significant over-representation of germline-specific genes among the genes that exhibited up-regulated expression in the F1 *set-9 set-26* double mutant (Figure 5D), but not the genes that exhibited down-regulated expression. Both sperm- and oocyte-specific genes were among these germline-specific genes that were up-regulated in the F1 *set-9 set-26* double mutant. It is possible that up-regulated expression of these germline-specific genes contribute to the reproductive defects of the *set-9 set-26* double mutant (Greer et al. 2014; Katz et al. 2009; Kerr et al. 2014).

Considering the maternal effect of SET-9 and SET-26 on fertility (Figure 2), we also profiled the transcriptome of the F3 *set-9 set-26* double mutant and compared that with the transcriptional profile of the F1 *set-9 set*-26 double mutant discussed above. We note that the germline of the F3 *set-9 set-26* is morphologically quite different from wild-type and the F1 *set-9 set-26* double mutant (Figure 2). Interestingly, we found that the

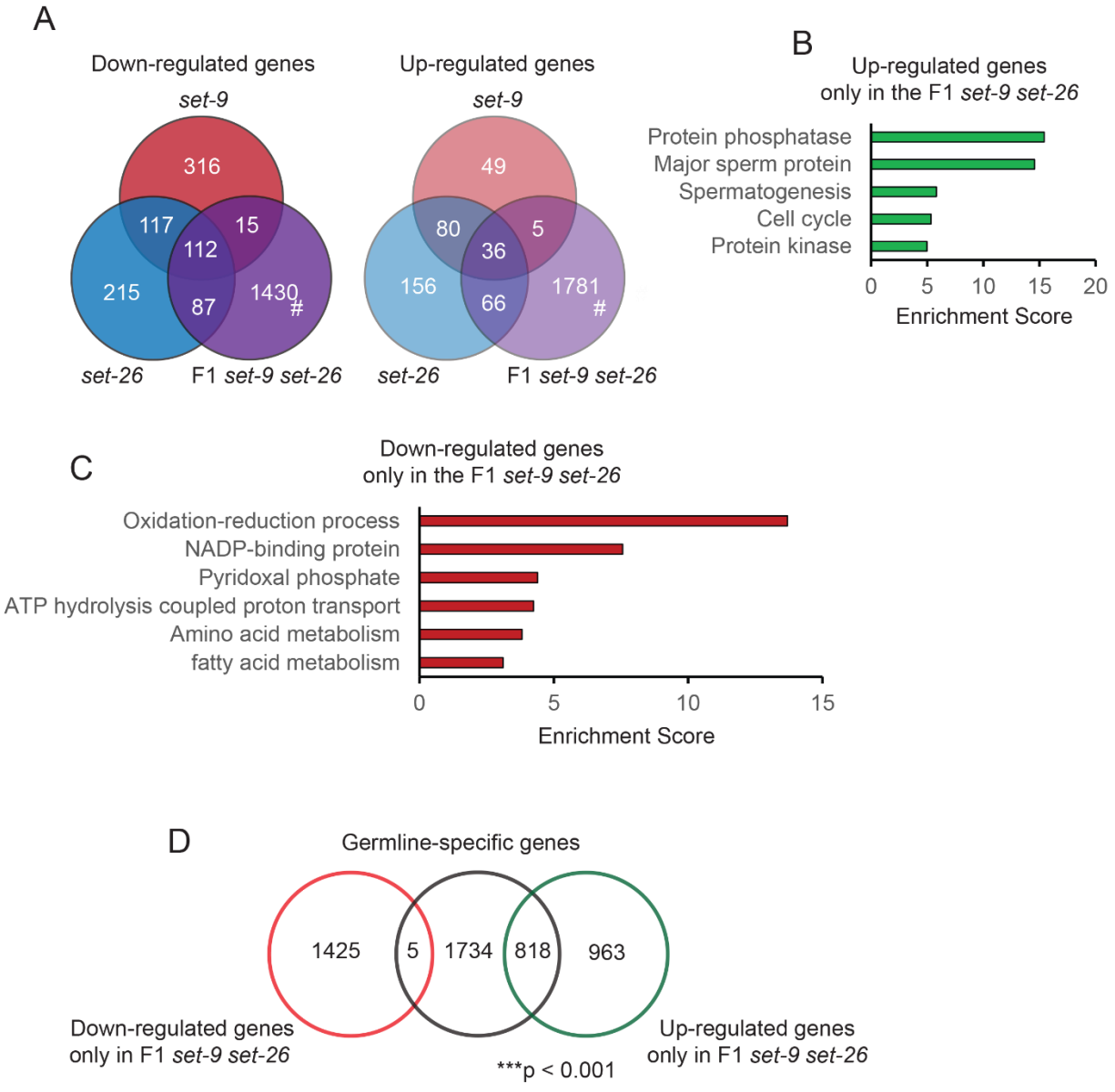

A
Down-regulated genes
set-9
316
117
15
112
215
87
1430
#
set-26
F1 set-9 set-26
Up-regulated genes
set-9
49
80
5
36
156
66
1781
#
set-26
F1 set-9 set-26
B
Up-regulated genes
only in the F1 set-9 set-26
Protein phosphatase
Major sperm protein
Spermatogenesis
Cell cycle
Protein kinase
0
5
10
15
20
Enrichment Score
C
Down-regulated genes
only in the F1 set-9 set-26
Oxidation-reduction process
NADP-binding protein
Pyridoxal phosphate
ATP hydrolysis coupled proton transport
Amino acid metabolism
fatty acid metabolism
0
5
10
15
Enrichment Score
D
Germline-specific genes
1425
5
1734
818
963
Down-regulated genes
only in F1 set-9 set-26
Up-regulated genes
only in F1 set-9 set-26
***p < 0.001

**Figure 5 Transcriptional profiles of *set-9(rw5)*, *set-26(tm2467),* and F1 *set-9(rw5) set-26(tm2467)* mutants.**

(A) Venn diagrams show the overlap among *set-9(rw5)*, *set-26(tm2467)*, and *set-9(rw5) set-26(tm2467)* down-regulated (left) and up-regulated (right) gene sets. hashtag indicates genes that only show expression change in the F1 *set-9(rw5) set-26(tm2467)* double mutant. (B) GO term analysis of up-regulated genes that only show expression change in the F1 *set-9(rw5) set-26(tm2467)* double mutant. (C) GO term analysis of down-regulated genes that only show expression change in the F1 *set-9(rw5) set-26(tm2467)* double mutant.

genes with significant expression change in the F1 *set-9 set-26* and F3 *set-9 set-26* double mutants not only substantially overlapped, but they were also enriched for similar functional groups based on GO term analyses (Figure 5-figure supplement 1A, 1B and 1C). The fold change of gene expression, compared to wild-type, in the F1 *set-9 set-26* and the F3 *set-9 set-26* double mutants also positively correlated (Figure 5-figure supplement 1D and 1E). These results together suggested that the transcriptional profiles of the F1 and the F3 *set-9 set-26* double mutants are highly correlative despite that the F3 *set-9 set-26* double mutant has a more severely defective germline. We note that the GO term “development and reproduction” was unique for the F3 *set-9 set-26* (Figure 5-figure supplement 1B), which may reflect the more severe germline defects in these mutant worms.

**Loss of SET-9 and SET-26 does not affect the global levels of H3K9me3**

We next investigated the possible normal functions of SET-9 and SET-26, which could inform how their inactivations lead to the gene expression changes and biological phenotypes discussed above. The SET domain of SET-26 was recently reported to show H3K9me3 methylation activity *in vitro* (Greer et al. 2014). This was a somewhat surprising result, as the SET domain of SET-26 (and SET-9) contains multiple mutations in the critical residues generally thought to be key for the methylating enzymatic activity of SET domain proteins (Figure 6-figure supplement 1A), and we have previously speculated that SET-9 and SET-26 are likely not active enzymes (Dillon et al. 2005; Ni et al. 2012). Furthermore, the likely homologs of SET-9 and SET-26 in flies (UpSET) and mammals (MLL5) have been reported to lack methylating activity (Sebastian et al.

2009; Rincon-Arano et al. 2012). Nevertheless, because of the reported *in vitro* results, we sought to monitor the global levels of H3K9me3 in the *set-9* and *set-26* mutants. We reasoned that if SET-9 & SET-26 are major enzymes for depositing H3K9me3 in *C. elegans*, then we would detect reduced H3K9me3 levels in the *set-9 set-26* double mutant strain. Using Western blotting, we showed that the global levels of H3K9me3 were not detectably altered in the *set-9, set-26,* or the *set-9 set-26* double mutants compared to wild-type worms at the L4 stage (Figure 6-figure supplement 1B and Figure 7-figure supplement 1A). We further investigated the genomic distribution of H3K9me3 in wild-type N2 and the *set-9 set-26* double mutant worms using ChIP-seq (chromatin immunoprecipitation coupled with next generation sequencing). Inspection of the genome-wide H3K9me3 distribution between N2 and the F3 *set-9 set-26* double mutant revealed highly similar patterns (Figure 6-figure supplement 1C). Furthermore, the Diffbind data analysis pipeline (Bardet et al. 2011) identified little significant differences between the two data sets (data not shown). Taken together the Western and ChIP-seq results, we concluded that SET-9 & SET-26 are likely not the major enzymes required for H3K9me3 deposition in *C. elegans*. However, we could not rule out the possibility that the SET-9 & SET-26 could deposit H3K9me3 in specific cells, or during a specific time. Consistent with our data, SET-25 is known to deposit the majority of H3K9me3 in *C. elegans* (Towbin et al. 2012).

**The PHD domains of SET-9 and SET-26 bind to H3K4me3 with adjacent acetylation marks *in vitro***

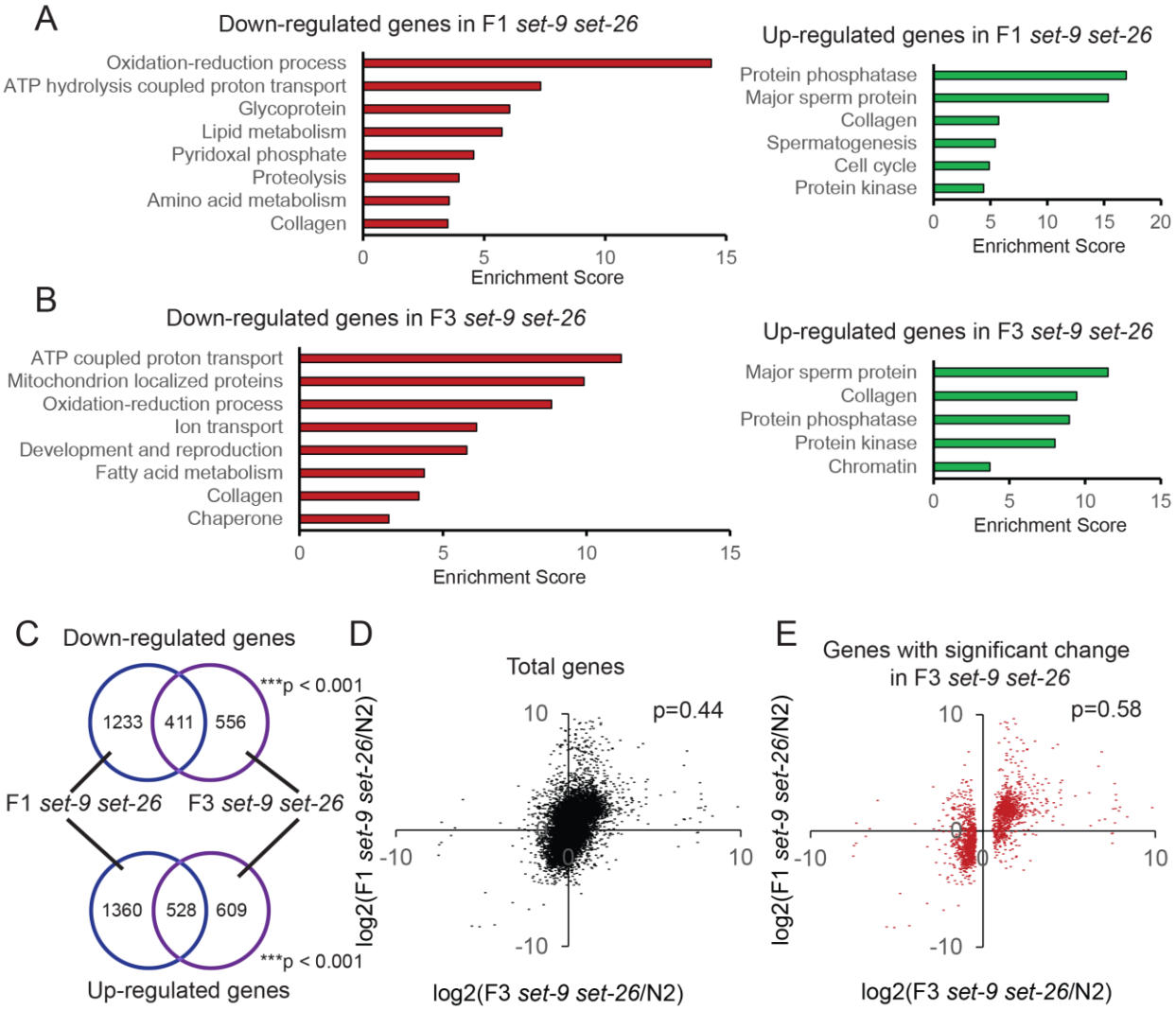
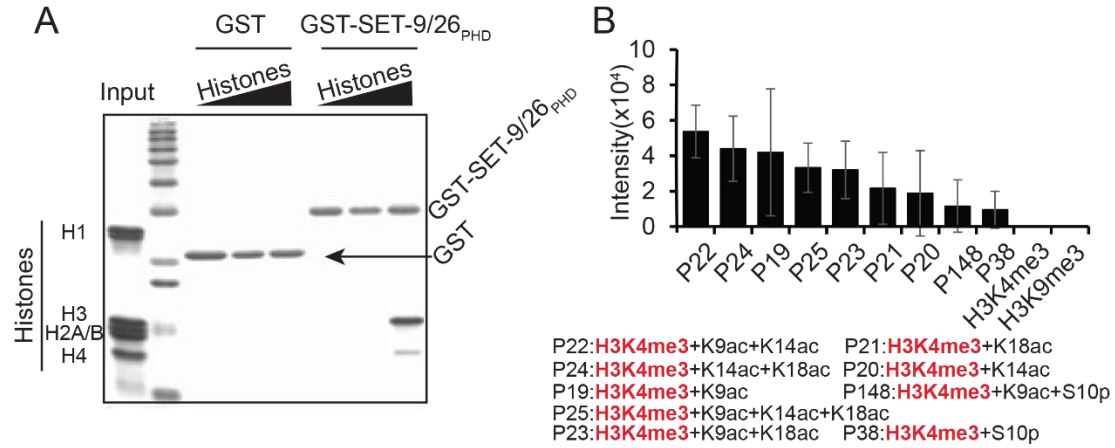
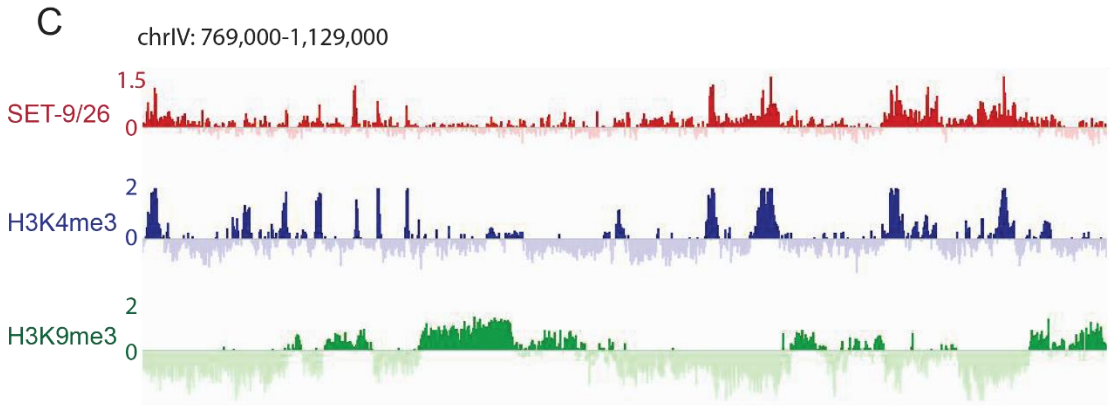
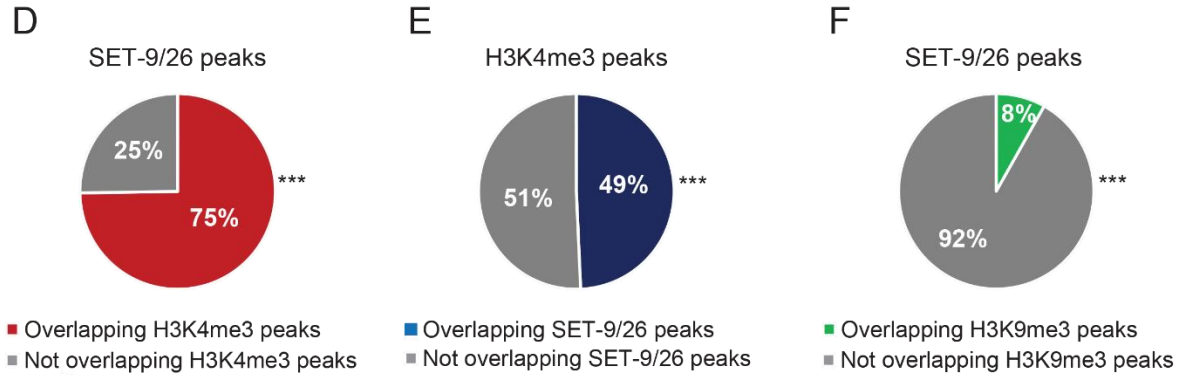

A
Down-regulated genes in F1 set-9 set-26
Oxidation-reduction process
ATP hydrolysis coupled proton transport
Glycoprotein
Lipid metabolism
Pyridoxal phosphate
Proteolysis
Amino acid metabolism
Collagen
0
5
10
15
Enrichment Score
Up-regulated genes in F1 set-9 set-26
Protein phosphatase
Major sperm protein
Collagen
Spermatogenesis
Cell cycle
Protein kinase
0
5
10
15
20
Enrichment Score
B
Down-regulated genes in F3 set-9 set-26
ATP coupled proton transport
Mitochondrion localized proteins
Oxidation-reduction process
Ion transport
Development and reproduction
Fatty acid metabolism
Collagen
Chaperone
0
5
10
15
Enrichment Score
Up-regulated genes in F3 set-9 set-26
Major sperm protein
Collagen
Protein phosphatase
Protein kinase
Chromatin
0
5
10
15
Enrichment Score
C
Down-regulated genes
***p < 0.001
1233
411
556
F1 set-9 set-26
F3 set-9 set-26
1360
528
609
***p < 0.001
Up-regulated genes
D
Total genes
p=0.44
10
-10
0
10
-10
log2(F1 set-9 set-26/N2)
log2(F3 set-9 set-26/N2)
E
Genes with significant change
in F3 set-9 set-26
p=0.58
10
-10
0
10
-10
log2(F1 set-9 set-26/N2)
log2(F3 set-9 set-26/N2)

**Figure 5-figure supplement 1 Transcriptional profiles comparison between F1 and F3 *set-9(rw5) set-26(tm2467) double mutant*.**

(A) GO term analysis of down-regulated and up-regulated genes in F1 *set-9(rw5) set-26(tm2467)* double mutant. (B) GO term analysis of down-regulated and up-regulated genes in F3 *set-9(rw5) set-26(tm2467)* double mutant. (C) Venn diagrams show the overlaps between down-regulated and up-regulated genes in the F1 and F3 *set-9(rw5) set-26(tm2467)* double mutant. (D) Scatter plot of log2 fold change of all mRNA in F3 and F1 *set-9(rw5) set-26(tm2467)* double mutant. (E) Scatter plot of log2 fold change of genes, expression level of which are significantly changed in *set-9(rw5) set-26(tm2467)* double mutant, in F3 and F1 *set-9(rw5) set-26(tm2467)* double mutant.

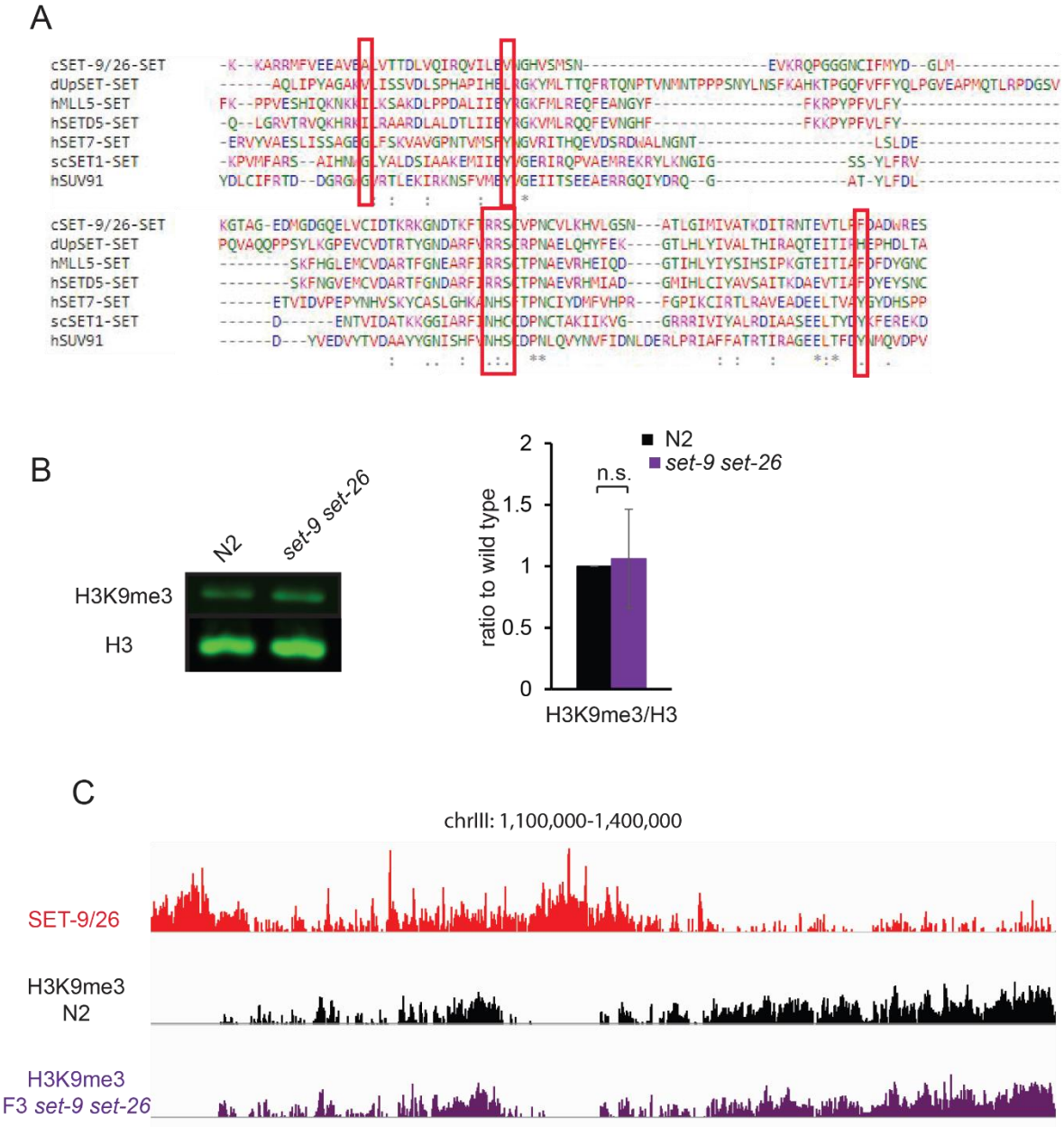

A
GST
GST-SET-9/26PHD
Input
Histones
Histones
GST-SET-9/26PHD
GST
Histones
H1
H3
H2A/B
H4
B
Intensity(x10⁴)
10
8
6
4
2
0
P22
P24
P19
P25
P23
P21
P20
P148
P38
H3K4me3
H3K9me3
P22:H3K4me3+K9ac+K14ac
P21:H3K4me3+K18ac
P24:H3K4me3+K14ac+K18ac
P20:H3K4me3+K14ac
P19:H3K4me3+K9ac
P148:H3K4me3+K9ac+S10p
P25:H3K4me3+K9ac+K14ac+K18ac
P23:H3K4me3+K9ac+K18ac
P38:H3K4me3+S10p


C
chrIV: 769,000-1,129,000
SET-9/26
1.5
0
H3K4me3
2
0
H3K9me3
2
0


D
SET-9/26 peaks
25%
75%
***
Overlapping H3K4me3 peaks
Not overlapping H3K4me3 peaks
E
H3K4me3 peaks
51%
49%
***
Overlapping SET-9/26 peaks
Not overlapping SET-9/26 peaks
F
SET-9/26 peaks
8%
92%
***
Overlapping H3K9me3 peaks
Not overlapping H3K9me3 peaks

## Figure 6 SET-9 & SET-26 bind to H3K4me3.

(A) PHD domain of SET-9 & SET-26 specifically pulled down H3 *in vitro*. Coomassie-blue stained gel showing pulldown results using GST–SET-9/26$_{PHD}$ and GST control. (B) Binding intensity of GST-SET-9/26$_{PHD}$ to the significant hits from histone peptide arrays. Quantitative data are shown in Table S5. (C) Genome browser view showing ChIP z-scores (standardized log2 ratios of ChIP/Input or ChIP/H3ChIP signals) for SET-9/26, H3K4me3 and H3K9me3 at a representative region. (D) 75% of the SET-9 & SET-26 peaks overlapped with H3K4me3 peaks. (***p<0.001 indicates overlapping more than expected) (E) 49% of the H3K4me3 peaks overlapped with SET-9 & SET-26 peaks. (***p<0.001 indicates overlapping more than expected) (F) 8% of the SET-9/26 peaks overlapped with H3K9me3 peaks. (***p<0.001 indicates overlapping less than expected)


A
cSET-9/26-SET
dUpSET-SET
hMLL5-SET
hSETD5-SET
hSET7-SET
scSET1-SET
hSUV91
B
N2
set-9 set-26
H3K9me3
H3
N2
set-9 set-26
n.s.
ratio to wild type
2
1.5
1
0.5
0
H3K9me3/H3
C
chrIII: 1,100,000-1,400,000
SET-9/26
H3K9me3
N2
H3K9me3
F3 set-9 set-26

## Figure 6-figure supplement 1 SET-9 & SET-26 are not the major enzyme for H3K9me3 *in vivo.*

(A) Clustal Omega alignment of the SET domains of SET-9 and SET-26 from *C. elegans*, MLL5, SET7 and SUV91 from human, UpSET from *Drosophila melanogaster* and SET1 from *Saccharomyces*. Red boxes indicate the proposed key residues. The online tool used is: http://www.ebi.ac.uk/Tools/msa/clustalo/ . (B) Western blotting showing H3K9me3 levels in N2 and the *set-9(rw5) set-26(tm2467)* double mutant (left). Synchronized L4 worms for N2 and the *set-9(rw5) set-26(tm2467)* double mutant were used. Quantification of three independent Western experiments (right). (C) Representative genome browser views of H3K9me3 profiles in N2 (black) and the F3 *set-9(rw5) set-26(tm2467)* double mutant (purple). z-scores for normalized H3K4me3 and SET-9 & SET-26 ChIP signals in N2 and the *set-9(rw5) set-26(tm2467)* are shown.

A

| Peptide # | Mean +/- SD |
|---|---|
| P22 | 53716 +/- 14910 |
| P24 | 43989 +/- 18399 |
| P19 | 41971 +/- 35922 |
| P25 | 33272 +/- 13841 |
| P23 | 32016 +/- 16214 |
| P21 | 21673 +/- 20341 |
| P20 | 18864 +/- 24133 |
| P148 | 11574 +/- 14806 |
| P38 | 9542 +/- 10432 |
| H3K4me3 | 0 +/- 0 |
| H3K9me3 | 0 +/- 0 |

Table S5

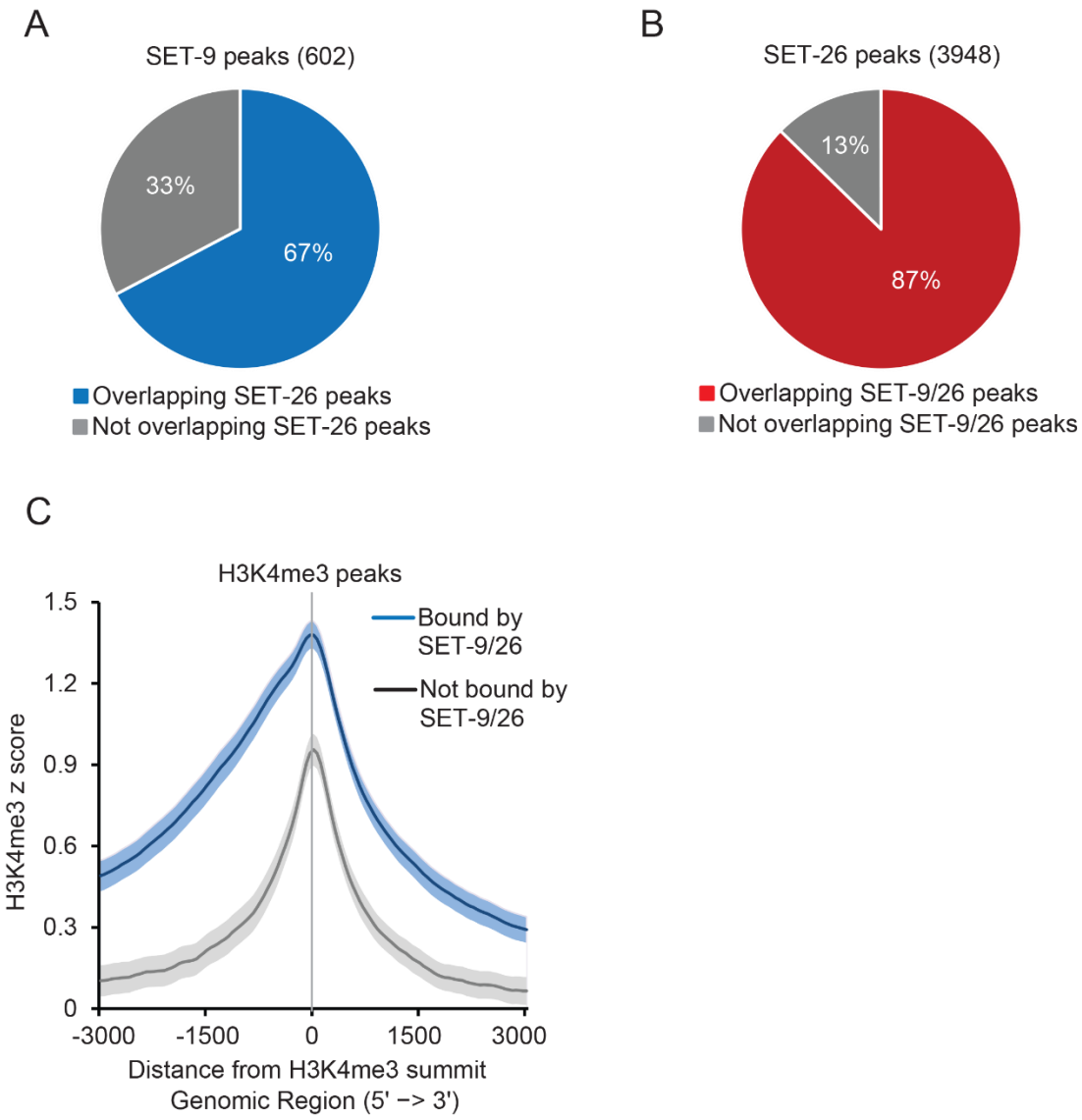
We next turned our attention to the PHD domains of SET-9 & SET-26, which are 100% identical. PHD domains are known to recognize specific histone modifications, we therefore tested whether the PHD domain of SET-9 & SET-26 also binds to specific histone modifications *in vitro*. We first used GST-tagged PHD domain of SET-9/SET-26 to perform an *in vitro* pull-down experiment using histones from calf thymus. We found that the PHD domain of SET-9 & SET-26 pulled down histones, in particular histone H3 (Figure 6A). A small amount of histone H4 was also recovered, which may be due to H3 and H4 associating as histone octamers in cells. We next screened for the specific histone modifications that are recognized by the PHD domain of SET-9 & SET-26 using a histone peptide array containing 95 unique modifications and 265 synthetic histone peptides. This assay revealed that the PHD domain of SET-9 & SET-26 specifically interacted with H3 peptides containing the K4me3 modification in combination with nearby acetylation (K9ac, K14ac and/or K18ac), but not H3K4me3 alone (Figure 6B).

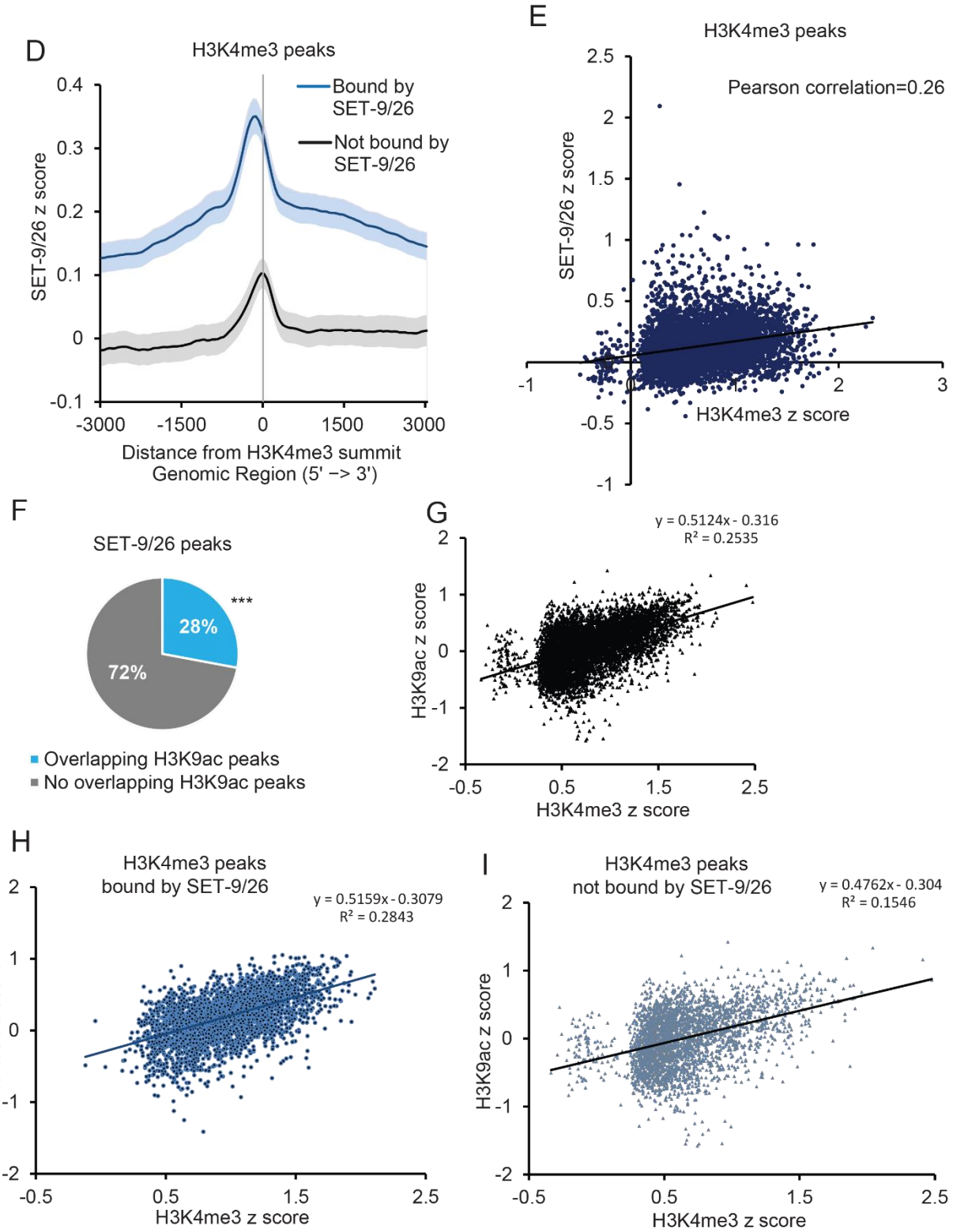
**Genome-wide profiles of SET-9 and SET-26 are highly concordant with that of H3K4me3**

We further examined the genome-wide binding profiles of SET-9 and SET-26 in *C. elegans* using ChIP-seq. Because antibodies capable of immunoprecipitating endogenous SET-9 and SET-26 were not available, we utilized the GFP knock-in strains discussed above. We performed anti-GFP ChIP-seq using the *set-9::gfp*, *set-26::gfp*, and *set-9::gfp set-26::gfp* strains. The double GFP strain was used in the hope that higher levels of GFP expression would provide more robust ChIP-seq data. MACS2 was used to identify the genomic regions significantly enriched for SET-9 and SET-26

binding. From the analyses, 602, 3948, and 5903 peaks were identified as bound by SET-9, SET-26, and SET-9 & SET-26 together, respectively. Interestingly, 67% of the SET-9 peaks (Figure 6-figure supplement 2A) overlapped with the SET-26 peaks, and 87% of the SET-26 peaks overlapped with the SET-9 & SET-26 peaks (Figure 6-figure supplement 2B). That SET-9 was found to bind many fewer regions than SET-26, and that the SET-9-bound peaks largely overlapped with those bound by SET-26, are consistent with the earlier data indicating SET-9 has a much more restricted expression pattern and a more limited function in the germline that is redundant with SET-26. Interestingly, ChIP-seq analysis from the double GFP strain nevertheless revealed many more peaks compared to SET-26::GFP alone. We interpreted these results to suggest that the double GFP strain simply represented a better reagent for capturing the SET-9 and SET-26 binding profiles than either of the single GFP strain, because the ChIP assay worked more effectively with higher levels of GFP expression. For further analyses, we used the peak regions identified in the double GFP strain as representation of the binding sites of SET-9 & SET-26.

Given the *in vitro* binding of the PHD domain of SET-9 & SET-26 to H3K4me3, we wondered whether SET-9 and SET-26 also bind to H3K4me3 *in vivo*. To test this, we examined the genome-wide pattern of H3K4me3 in wild-type worms using ChIP-seq. A representative genome browser view revealed that the SET-9 & SET-26 binding profile generally correlated well with the H3K4me3 profile (Figure 6C). In contrast, the SET-9 & SET-26 binding profile was largely different from that of H3K9me3 (Figure 6C), again consistent with our earlier conclusion that SET-9 & SET-26 likely do not modify H3K9me3.


A
SET-9 peaks (602)
33%
67%
Overlapping SET-26 peaks
Not overlapping SET-26 peaks
B
SET-26 peaks (3948)
13%
87%
Overlapping SET-9/26 peaks
Not overlapping SET-9/26 peaks
C
H3K4me3 peaks
Bound by SET-9/26
Not bound by SET-9/26
H3K4me3 z score
1.5
1.2
0.9
0.6
0.3
0
-3000
-1500
0
1500
3000
Distance from H3K4me3 summit
Genomic Region (5' –> 3')


D
H3K4me3 peaks
Bound by SET-9/26
Not bound by SET-9/26
SET-9/26 z score
0.4
0.3
0.2
0.1
0
-0.1
-3000
-1500
0
1500
3000
Distance from H3K4me3 summit
Genomic Region (5' -> 3')
E
H3K4me3 peaks
Pearson correlation=0.26
SET-9/26 z score
2.5
2
1.5
1
0.5
-0.5
-1
-1
0
2
3
H3K4me3 z score
F
SET-9/26 peaks
***
28%
72%
Overlapping H3K9ac peaks
No overlapping H3K9ac peaks
G
y = 0.5124x - 0.316
R² = 0.2535
H3K9ac z score
2
1
0
-1
-2
-0.5
0.5
1.5
2.5
H3K4me3 z score
H
H3K4me3 peaks
bound by SET-9/26
y = 0.5159x - 0.3079
R² = 0.2843
H3K9ac z score
2
1
0
-1
-2
-0.5
0.5
1.5
2.5
H3K4me3 z score
I
H3K4me3 peaks
not bound by SET-9/26
y = 0.4762x - 0.304
R² = 0.1546
H3K9ac z score
2
1
0
-1
-2
-0.5
0.5
1.5
2.5
H3K4me3 z score

## Figure 6-figure supplement 2 SET-9 & SET-26 bind to H3K4me3 *in vivo*.

(A) A large proportion of SET-9 peaks overlapped with SET-26 peak. (B) A large proportion of SET-26 peaks overlapped with SET-9 & SET-26 peaks. (C) Metagene plots showing the average H3K4me3 z-score (standardized log2 ratios of H3K4me3ChIP/H3ChIP signals) for the H3K4me3 peaks bound (shown in blue) and not bound by SET-9 & SET-26 (shown in grey) in N2 worms. Regions 3000bp upstream and downstream of the peak summits are shown. The light blue and light grey indicate 95% confidence intervals. (D) Metagene plots showed the average SET-9 & SET-26 z-score (standardized log2 ratios of SET-9 & SET-26 ChIP/input signals) for the H3K4me3 peaks bound (shown in blue) and not bound by SET-9 & SET-26 peaks (shown in grey). The light blue and light grey indicate 95% confidence intervals. (E) Scatter plot of H3K4me3 z-score and SET-9 & SET-26 z-score for H3K4me3 peak regions. (F) 28% of the SET-9 & SET-26 peaks overlapped with H3K9ac peaks. (***p<0.001 indicates overlapping more than expected) H3K9ac ChIP-seq data is from modENCODE(G) Scatter plot of H3K4me3 z-score and H3K9ac z-score for all H3K4me3 peak regions. (H) Scatter plot of H3K4me3 z-score and H3K9ac z-score for H3K4me3 peak regions that bound by SET-9 & SET-26. (I) Scatter plot of H3K4me3 z-score and H3K9ac z-score for H3K4me3 peak regions not bound by SET-9 & SET-26.

To more rigorously assess whether the SET-9 & SET-26 profile is concordant with the H3K4me3 profile, we tested whether SET-9 & SET-26, and H3K4me3 were enriched in overlapping genomic regions. To this end, we used MACS2 to identify 5996 genomic regions enriched for H3K4me3 marking (peaks), and then compared the degree of overlap between the SET-9 & SET-26 peaks with those of H3K4me3. In this comparison, we found that 75% of the SET-9 & SET-26 binding regions overlapped with the H3K4me3 peaks (Figure 6D). The reciprocal comparison revealed that 49% of the H3K4me3 peaks overlapped with the SET-9 & SET-26 binding regions (Figure 6E). We sought to further characterize the overlapping profiles between SET-9 & SET-26 and H3K4me3 using meta-analysis. We separated the H3K4me3 peaks into the group that bound by SET-9 & SET-26 and the group that did not, and we plotted the average normalized H3K4me3 levels centered around the summits of the H3K4me3 peaks and oriented at a 5' to 3' direction according the nearest genes associated with the peaks. We found that, on average, the H3K4me3 peaks bound by SET-9 & SET-26 had higher levels of H3K4me3 marking and their H3K4me3 marking was somewhat higher 5' to the summit (Figure 6-figure supplement 2C). This asymmetrical marking of H3K4me3 might relate to that these peaks generally localized around annotated transcriptional/translational start sites (TSSs) and H3K4me3 levels tend to be higher 5' to the start sites of the genes. In contrast, the average plot of the H3K4me3 peaks that were not bound by SET-9 & SET-26 was more symmetrical, and this correlated with these H3K4me3 peaks generally localized to gene body regions (data not shown).

We next used a similar meta-analysis approach but plotted the normalized SET-9 & SET-26 ChIP signal for the two groups of H3K4me3 peaks, still oriented at the summits

of the H3K4me3 peaks (Figure 6-figure supplement 2D). This analysis allowed us to determine how far the average summit of the SET-9 & SET-26 peaks was relative to that of the H3K4me3 peaks. The results indicated that the average summit of the SET-9 & SET-26 peaks was ~100-200bp upstream of the H3K4me3 summit (Figure 6-figure supplement 2D). Interestingly, for the group of H3K4me3 peaks that were not identified to share SET-9 & SET-26 enrichment based on MACS2, we nevertheless detected a small amount of SET-9 & SET-26 binding exactly at the summits of the H3K4me3 peaks (Figure 6-figure supplement 2D). We interpreted these results to suggest that SET-9 & SET-26 likely bind to most if not all of the H3K4me3 enriched regions, but some of the binding was too weak (or strong binding only in a subset of the *C. elegans* cells) to be called by a statistical program like MACS2. Lastly, we used scatter plot analysis to compare the H3K4me3 enriched signal vs. the SET-9 & SET-26 binding signal for each of the H3K4me3 peak region (Figure 6-figure supplement 2E). The results showed that the H3K4me3 signal intensity is positively correlated with that of SET-9 & SET-26 (Figure 6-figure supplement 2E). The data thus far support a model that SET-9 & SET-26 bind to regions marked by H3K4me3 in *C. elegans*, and that the detectable SET-9 & SET-26 binding sites are generally marked by higher levels of H3K4me3.

Since the *in vitro* histone peptide array results suggested that histone acetylations are also important for the binding, we next compared the degree of overlap between the SET-9 & SET-26 peaks with those of H3K9ac using data from modENCODE. In this comparison, we found that 28% of the SET-9 & SET-26 binding regions overlapped with the H3K9ac peaks (***$p<0.001$) (Figure 6-figure supplement 2F), and 98% of these shared peak regions were also marked by H3K4me3 (data not shown). The significant

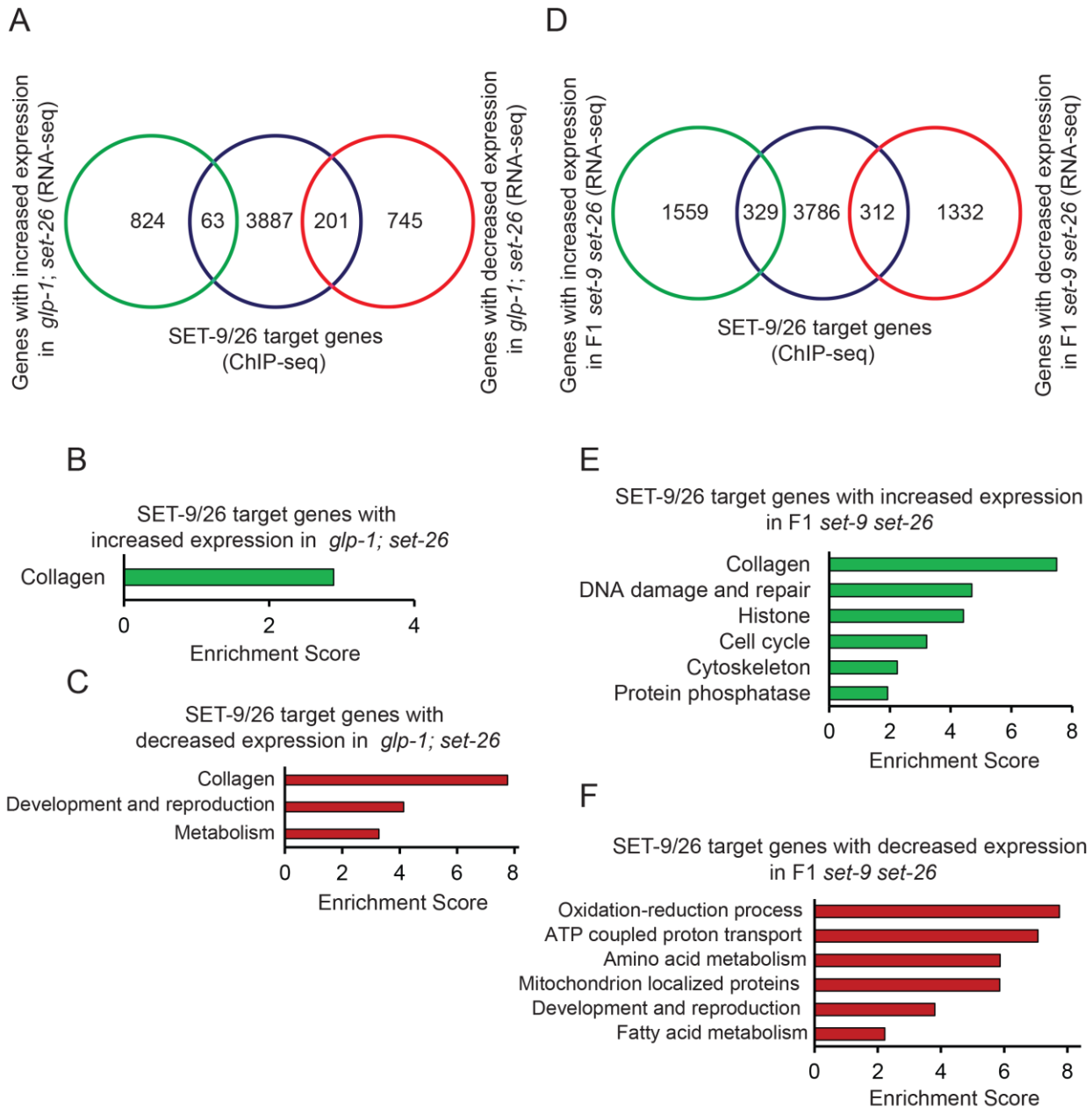
but lower percentage of overlap between SET-9 & SET-26 binding with the H3K9ac enriched regions (relative to the overlap with H3K4me3 enriched regions) is consistent with the *in vitro* observation that SET-9 & SET-26 bind to H3K4me3 with adjacent acetylation, but the exact acetylated residues could vary (Figure 6B), but is also likely partly due to technical differences, e.g. the different ChIP-seq data were generated using worms at different stages and acetyl specific antibodies generally have lower specificity. We additionally compared the overall correlation between H3K4me3 and H3K9ac enriched regions and whether SET-9 & SET-26 may affect their co-occurance. As expected, we found that the genome-wide distribution of H3K4me3 correlated well with that of H3K9ac (Figure 6-figure supplement 2G). Interestingly, the correlation between H3K4me3 and H3K9ac was higher for the peaks that were bound by SET-9 & SET-26 compared to those that were not (Figure 6-figure supplement 2H and Figure 6-figure supplement 2I). These genomic data are consistent with the model that SET-9 & SET-26 bind to genomic regions marked by H3K4me3 with nearby acetylations in *C. elegans*.

**SET-9 & SET-26 regulate the RNA expression of some of their target genes**

We next asked whether SET-9 & SET-26 binding could influence gene expression. To test this, we assigned the 5903 regions bound by SET-9 & SET-26 (based on ChIP-seq results, Figure 6) with their closest genes, and then filtered them for genes that were detectably expressed in our RNA-seq data sets, which yielded 4427 potential targets of SET-9 & SET-26 that were also actively expressed in our experimental conditions. We then compared the list of SET-9 & SET-26 target genes with the lists of genes that

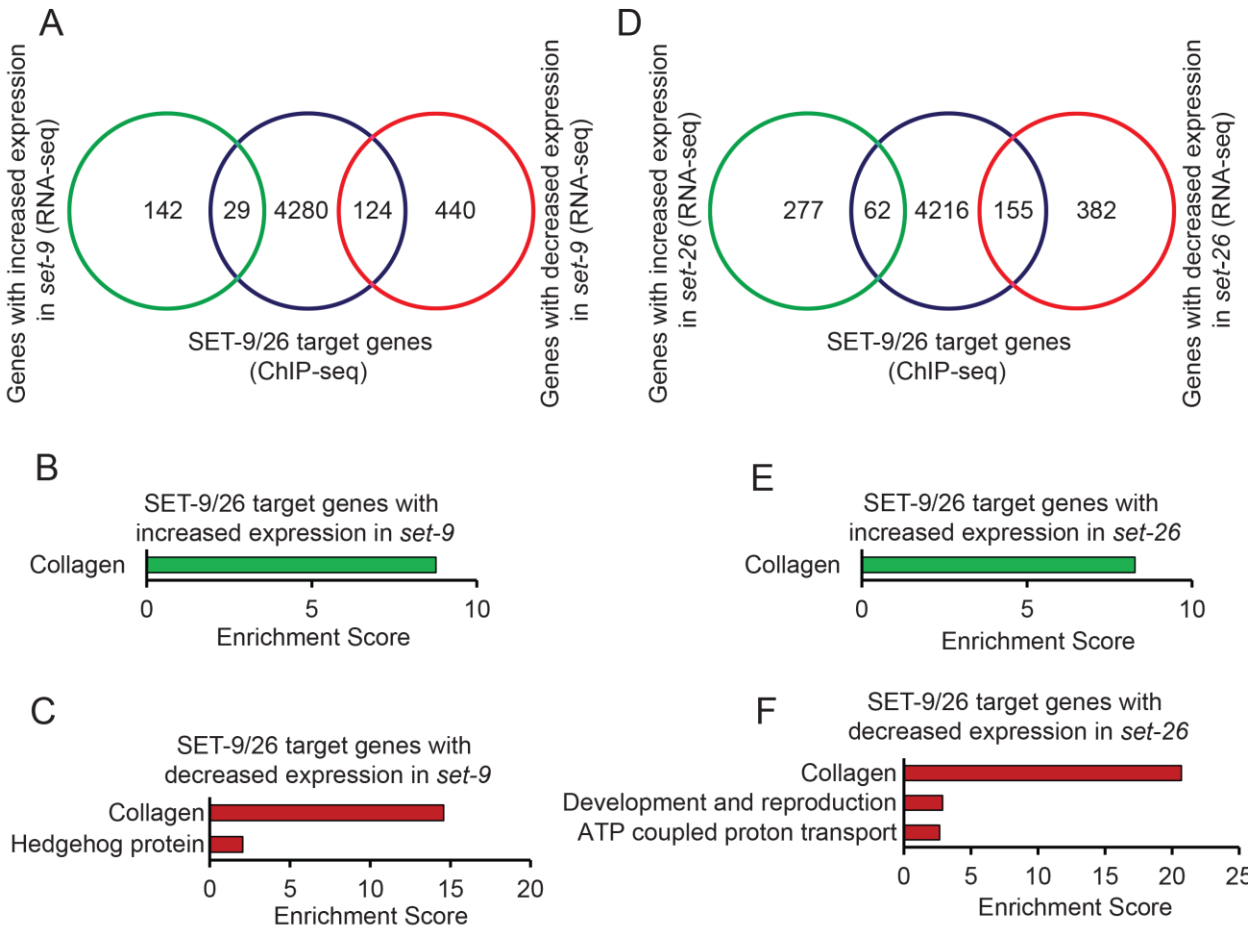
exhibited significant gene expression change in the *set-9* single, *set-26* single, and the F1 *set-9 set-26* double mutants compared to wild-type worms based on our RNA-seq data (Figure 7D and Figure 7-figure supplement 1A and 1D). Through this comparison, we identified the putative SET-9 & SET-26 target genes that changed expression when *set-9* and/or *set-26* were inactivated (Figure 7D and Figure 7-figure supplement 1A and 1D). Out of the 735, 876, 3532 differentially expressed genes in the *set-9*, *set-26*, and F1 *set-9 set-26* double mutant strains, SET-9 & SET-26 bound to 153, 217, 641 of them respectively (Figure 7D and Figure 7-figure supplement 1A and 1D). GO analyses revealed some interesting over-represented functional groups among these, especially for the SET-9 & SET-26 targets that exhibited expression change in the F1 double mutant (Figures 7E-F and Figure 7-figure supplement 1).

In an attempt to identify the somatic SET-26 target genes with a role in longevity, we compared the genes bound by SET-9 & SET-26 to the genes that showed expression change in the germlineless *glp-1; set-26* double mutant (Figure 7A). GO analysis did not reveal functional groups that are directly linked to longevity (Figure 7B and 7C). Moreover, for the DAF-16-dependent somatic SET-26 regulated genes that were highly enriched for the functional group “determination of lifespan” (Figure 4), only 3 out of the 134 genes were bound by SET-9 & SET-26, a representation that is lower than expected based on random chance. Together, these data suggested that SET-26 indirectly impact DAF-16 activity and DAF-16-mediated longevity change.


A
Genes with increased expression in glp-1; set-26 (RNA-seq)
824
63
3887
201
745
SET-9/26 target genes (ChIP-seq)
Genes with decreased expression in glp-1; set-26 (RNA-seq)
D
Genes with increased expression in F1 set-9 set-26 (RNA-seq)
1559
329
3786
312
1332
SET-9/26 target genes (ChIP-seq)
Genes with decreased expression in F1 set-9 set-26 (RNA-seq)
B
SET-9/26 target genes with increased expression in glp-1; set-26
Collagen
0
2
4
Enrichment Score
C
SET-9/26 target genes with decreased expression in glp-1; set-26
Collagen
Development and reproduction
Metabolism
0
2
4
6
8
Enrichment Score
E
SET-9/26 target genes with increased expression in F1 set-9 set-26
Collagen
DNA damage and repair
Histone
Cell cycle
Cytoskeleton
Protein phosphatase
0
2
4
6
8
Enrichment Score
F
SET-9/26 target genes with decreased expression in F1 set-9 set-26
Oxidation-reduction process
ATP coupled proton transport
Amino acid metabolism
Mitochondrion localized proteins
Development and reproduction
Fatty acid metabolism
0
2
4
6
8
Enrichment Score

## Figure 7 Binding of SET-9 & SET-26 regulate the RNA expression of specific target genes.

(A) Venn diagram showing comparisons of the up-regulated and down-regulated somatic SET-26 regulated genes and the SET-9 & SET-26 target genes. (B) GO term analyses of the up-regulated somatic SET-26 regulated genes bound by SET-9 & SET-26. (C) GO term analyses of the down-regulated somatic SET-26 regulated genes bound by SET-9 & SET-26. (D) Venn diagram showing comparisons of the genes with increased and decreased expression changes in the F1 *set-9(rw5) set-26(tm2467)* double mutant and the SET-9 & SET-26 target genes. (E) GO term analyses of the up-regulated genes in the F1 *set-9(rw5) set-26(tm2467)* double mutant that bound by SET-9 & SET-26. (F) GO term analyses of the down-regulated genes in the F1 *set-9(rw5) set-26(tm2467)* double mutant that bound by SET-9 & SET-26.


A
Genes with increased expression in set-9 (RNA-seq)
142
29
4280
124
440
Genes with decreased expression in set-9 (RNA-seq)
SET-9/26 target genes (ChIP-seq)
B
SET-9/26 target genes with increased expression in set-9
Collagen
0
5
10
Enrichment Score
C
SET-9/26 target genes with decreased expression in set-9
Collagen
Hedgehog protein
0
5
10
15
20
Enrichment Score
D
Genes with increased expression in set-26 (RNA-seq)
277
62
4216
155
382
Genes with decreased expression in set-26 (RNA-seq)
SET-9/26 target genes (ChIP-seq)
E
SET-9/26 target genes with increased expression in set-26
Collagen
0
5
10
Enrichment Score
F
SET-9/26 target genes with decreased expression in set-26
Collagen
Development and reproduction
ATP coupled proton transport
0
5
10
15
20
25
Enrichment Score

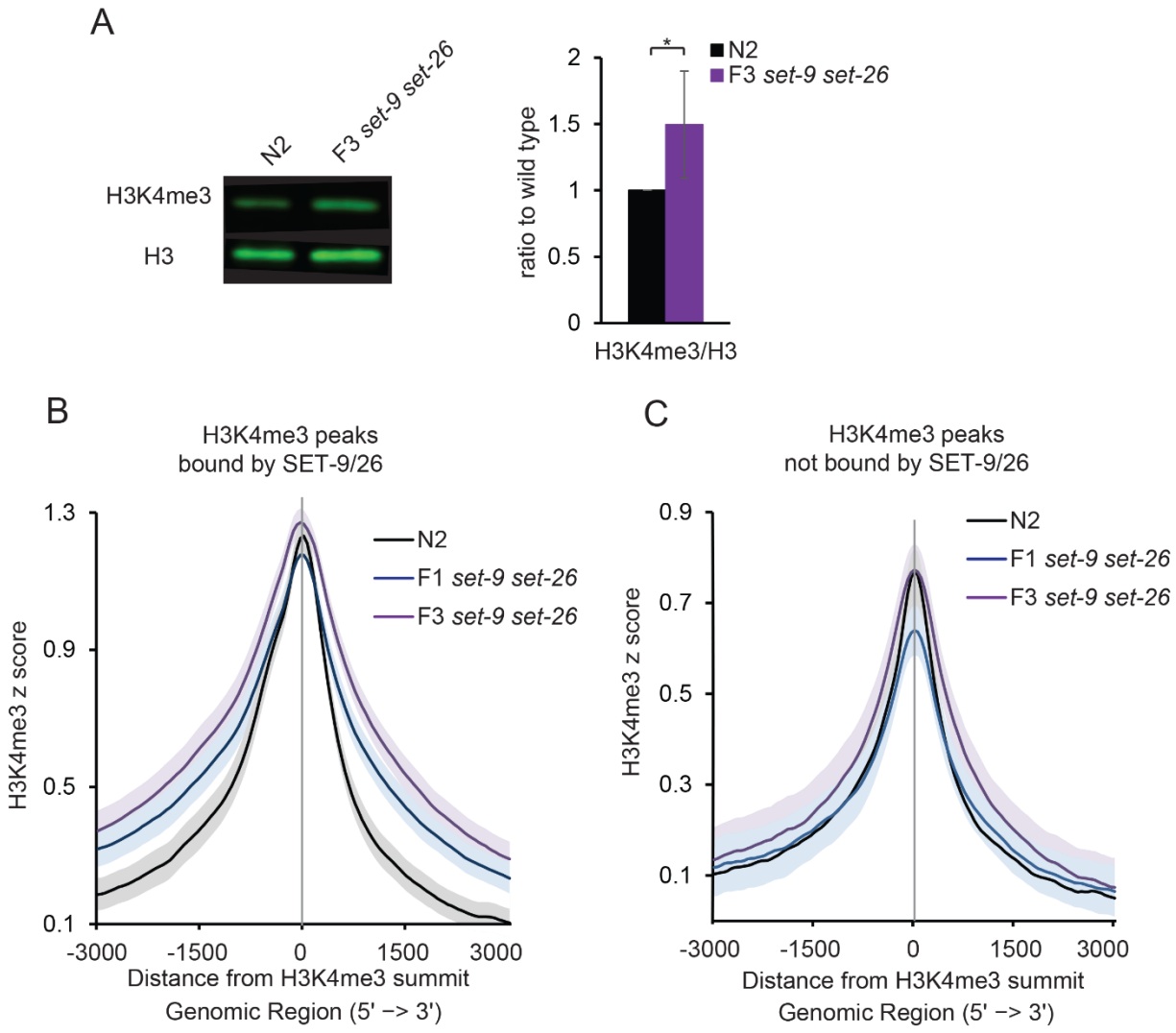
## Figure 7-figure supplement 1 SET-9 & SET-26 target genes.

(A) Venn diagrams showing comparisons of the genes with increased and decreased expression changes in the *set-9(rw5)*. (B) GO term analyses of the up-regulated genes in *set-9(rw5)* bound by SET-9 & SET-26. (C) GO term analyses of the down-regulated genes in *set-9(rw5)* bound by SET-9 & SET-26. (D) Venn diagrams showing comparisons of the genes with increased and decreased expression changes in the *set-26(tm2467)*. (E) GO term analyses of the up-regulated genes in *set-26(tm2467)* bound by SET-9 & SET-26. (F) GO term analyses of the down-regulated genes in *set-26(tm2467)* bound by SET-9 & SET-26.

**Loss of SET-9 and SET-26 results in elevated levels of H3K4me3**

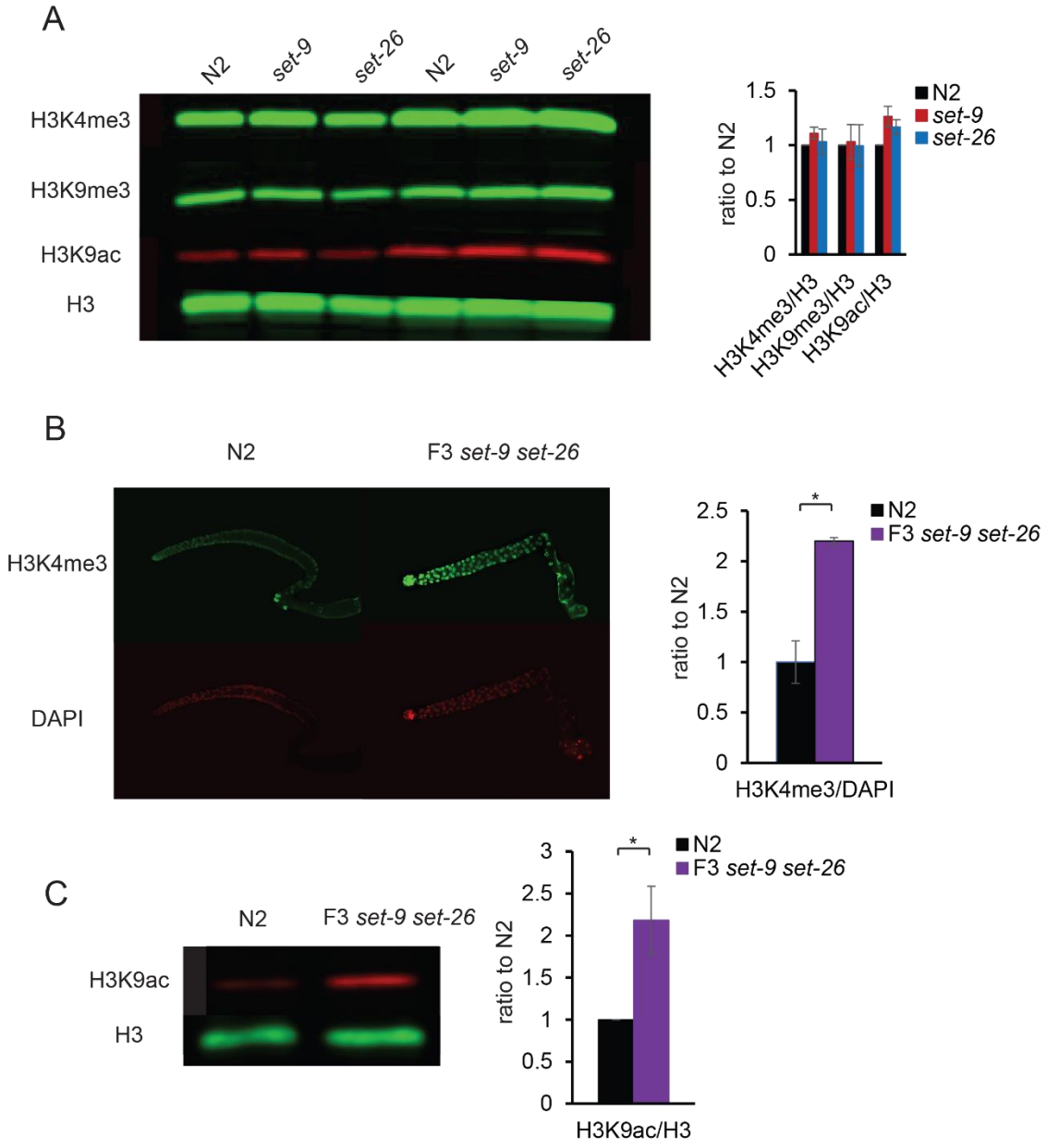
We next tested whether SET-9 & SET-26, in addition to binding to H3K4me3, can influence the marking of H3K4me3 in anyway. We first examined the global H3K4me3 levels in the *set-9* single, *set-26* single and F3 *set-9 set-26* double mutants. We observed a moderate but significant increase in H3K4me3 levels in the F3 *set-9 set-26* double mutant but not in the *set-9* or *set-26* single mutants using Western blotting (Figure 8A and Figure 8-figure supplement 1A). This increase in H3K4me3 levels was also readily detectable in dissected gonads of the *set-9 set-26* double mutant (Figure 8-figure supplement 1B).

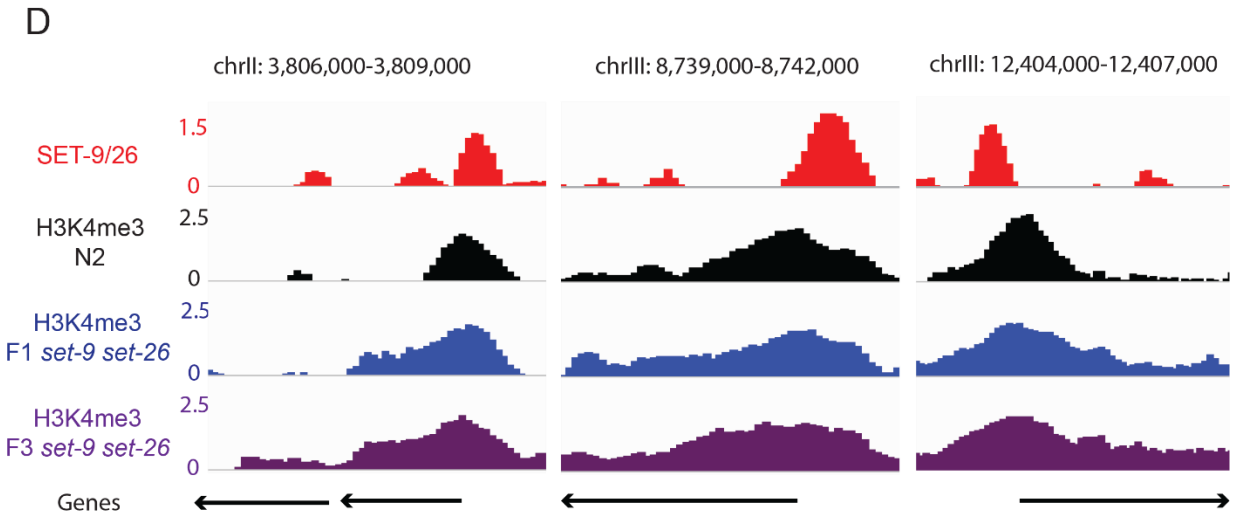
Given our earlier data indicating that SET-9 & SET-26 bind to H3K4me3, the elevated H3K4me3 levels in the *set-9 set-26* double mutant strain could be a result of the lost recruitment of SET-9 and SET-26 at H3K4me3 sites. To test this hypothesis, we examined the genome-wide patterns of H3K4me3 in wild-type, the F1 and F3 *set-9 set-26* double mutant worms. Inspection of the genome-wide H3K4me3 distribution in the three genotypes revealed a “spreading” of H3K4me3 enriched regions around SET-9 & SET-26 binding sites in worms lacking *set-9* & *set-26* (Figure 8-figure supplement 1D). To quantify this “spreading” globally, we separated the H3K4me3 enriched regions into those bound by SET-9 & SET-26 *vs.* those that were not, and compared their average profiles using meta analysis (Figures 8B and 8C). The meta analysis plots were centered around the summits of the H3K4me3 peaks and orientated in the 5’ to 3’ direction according to the closest gene associated with each of the peak (Figure 8B and 8C). For the H3K4me3 peaks bound by SET-9 & SET-26, we again observed a “spreading” of H3K4me3 marking, especially towards the 3’ direction, where significantly

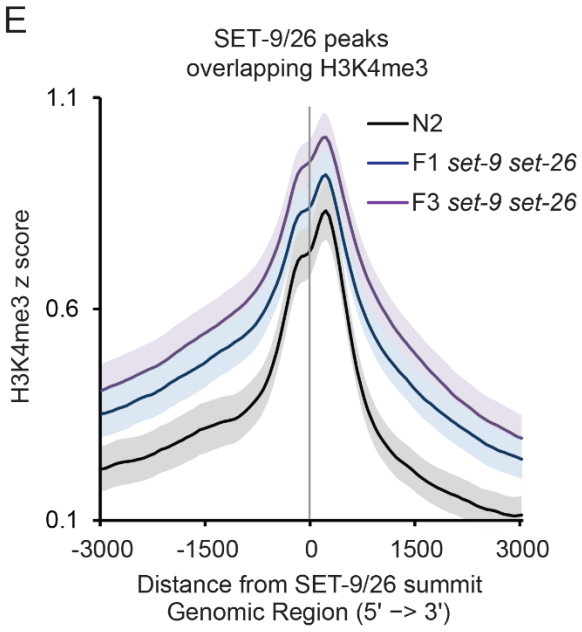

A
N2
F3 set-9 set-26
H3K4me3
H3
ratio to wild type
2
1.5
1
0.5
0
*
N2
F3 set-9 set-26
H3K4me3/H3
B
H3K4me3 peaks
bound by SET-9/26
H3K4me3 z score
1.3
0.9
0.5
0.1
N2
F1 set-9 set-26
F3 set-9 set-26
-3000
-1500
0
1500
3000
Distance from H3K4me3 summit
Genomic Region (5' -> 3')
C
H3K4me3 peaks
not bound by SET-9/26
H3K4me3 z score
0.9
0.7
0.5
0.3
0.1
N2
F1 set-9 set-26
F3 set-9 set-26
-3000
-1500
0
1500
3000
Distance from H3K4me3 summit
Genomic Region (5' -> 3')

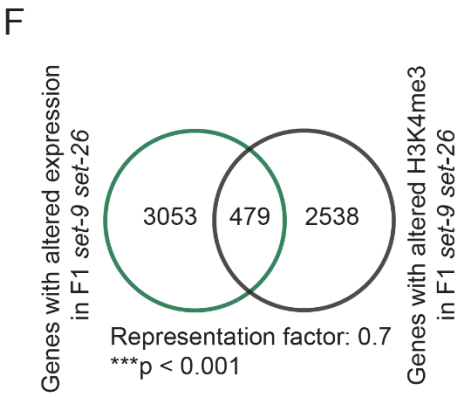
## Figure 8 SET-9 & SET-26 restrict the spreading of H3K4me3.

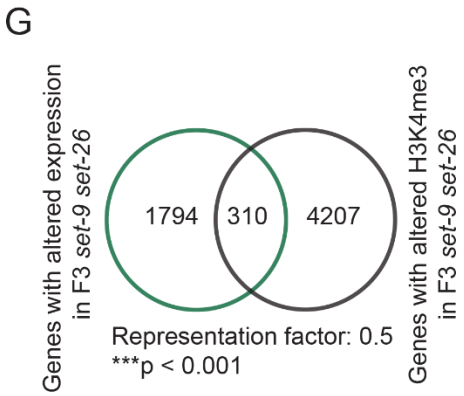
(A) Western blotting indicated that H3K4me3 levels increased in the F3 *set-9(rw5) set-26(tm2467)* double mutant (left). Synchronized L4 worms for N2 and the F3 *set-9(rw5) set-26(tm2467)* double mutant were used. Quantification of normalized H3K4me3 levels from three independent Western experiments (right). (*p<0.05) (B) Metagene plots showing the average H3K4me3 z-score (standardized log2 ratios of ChIP/H3ChIP signals) for the peaks bound by SET-9 & SET-26 in N2 (black) and the F1 (blue) and F3 (purple) *set-9(rw5) set-26(tm2467)* double mutants. Regions 3000bp upstream and downstream of the peak summits are shown. The grey areas indicate 95% confidence intervals. (C) Metagene plots showing the average H3K4me3 z-score for the peaks not bound by SET-9 & SET-26 in N2 (black) and the F1 (blue) and F3 (purple) *set-9(rw5) set-26(tm2467)* double mutant. The grey areas indicate 95% confidence intervals.

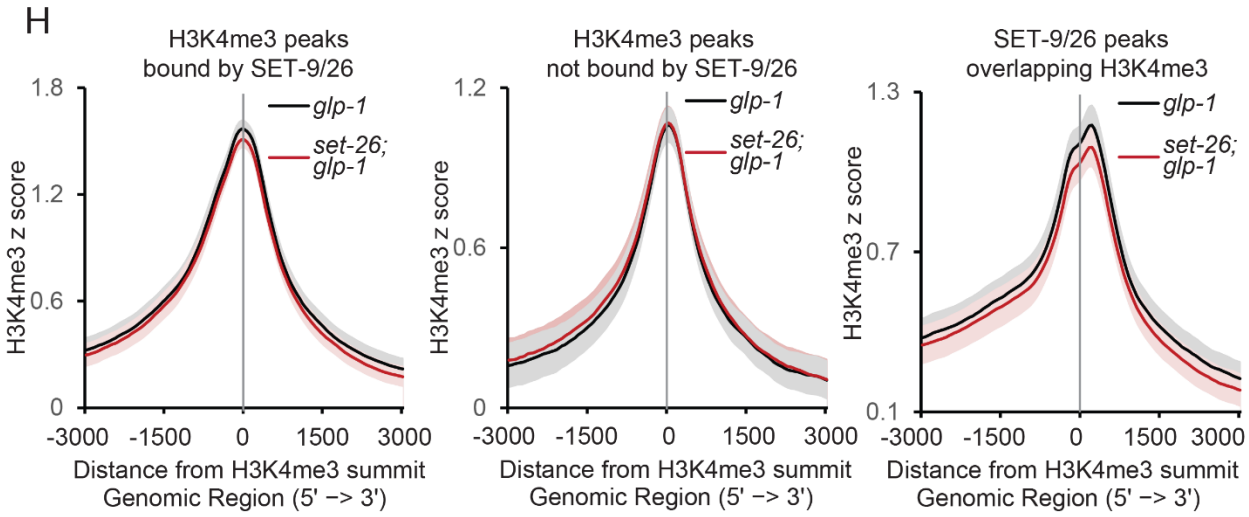

A
N2
set-9
set-26
N2
set-9
set-26
H3K4me3
H3K9me3
H3K9ac
H3
ratio to N2
1.5
1
0.5
0
N2
set-9
set-26
H3K4me3/H3
H3K9me3/H3
H3K9ac/H3
B
N2
F3 set-9 set-26
H3K4me3
DAPI
ratio to N2
2.5
2
1.5
1
0.5
0
*
N2
F3 set-9 set-26
H3K4me3/DAPI
C
N2
F3 set-9 set-26
H3K9ac
H3
ratio to N2
3
2.5
2
1.5
1
0.5
0
*
N2
F3 set-9 set-26
H3K9ac/H3


D
chrII: 3,806,000-3,809,000
chrIII: 8,739,000-8,742,000
chrIII: 12,404,000-12,407,000
SET-9/26
1.5
0
H3K4me3
N2
2.5
0
H3K4me3
F1 set-9 set-26
2.5
0
H3K4me3
F3 set-9 set-26
2.5
0
Genes


E
SET-9/26 peaks
overlapping H3K4me3
1.1
0.6
0.1
H3K4me3 z score
N2
F1 set-9 set-26
F3 set-9 set-26
-3000
-1500
0
1500
3000
Distance from SET-9/26 summit
Genomic Region (5' -> 3')


F
Genes with altered expression in F1 set-9 set-26
3053
479
2538
Genes with altered H3K4me3 in F1 set-9 set-26
Representation factor: 0.7
***p < 0.001


G
Genes with altered expression in F3 set-9 set-26
1794
310
4207
Genes with altered H3K4me3 in F3 set-9 set-26
Representation factor: 0.5
***p < 0.001


H
H3K4me3 peaks
bound by SET-9/26
H3K4me3 peaks
not bound by SET-9/26
SET-9/26 peaks
overlapping H3K4me3
glp-1
set-26;
glp-1
H3K4me3 z score
1.8
1.2
0.6
0
1.2
0.6
0
1.3
0.7
0.1
-3000
-1500
0
1500
3000
Distance from H3K4me3 summit
Genomic Region (5' −> 3')

## Figure 8-figure supplement 1 SET-9 & SET-26 restrict the spreading of H3K4me3.

(A) H3K4me3 and H3K9me3 levels in N2, *set-9(rw5)* and *set-26(tm2467)* mutants. Synchronized L4 worms were used. Quantification showed results from three independent experiments. (B) Elevated H3K4me3 levels were observed in the dissected gonads of the F3 *set-9(rw5) set-26(tm2467)* double mutant. Synchronized D1-D2 adults for N2 and the F3 *set-9(rw5) set-26(tm2467)* double mutant were used. Representative immunostaining images are shown. Right panel shows quantification of images shown on left. (*$p<0.05$) (C) Western blotting showing elevated H3K9ac levels in the *set-9(rw5) set-26(tm2467)* double mutant (left). Synchronized L4 worms for N2 and the *set-9(rw5) set-26(tm2467)* double mutant were used. Quantification of normalized H3K9ac levels from three independent Western experiments (right). (*$p<0.05$) (D) Representative genome browser views of H3K4me3 profiles in N2 (black) and the F1 (blue) and F3 (purple) *set-9(rw5) set-26(tm2467)* double mutant surrounding SET-9 & SET-26 binding sites (red). z-scores for normalized H3K4me3 and SET-9 & SET-26 ChIP signals in N2 and the F1 and F3 *set-9(rw5) set-26(tm2467)* are shown.

(E) Metagene plots showing the average H3K4me3 z-score (standardized log2 ratios of ChIP/H3ChIP signals) in N2 (black) and the F1 (blue) and F3 (purple) *set-9(rw5) set-26(tm2467)* double mutant (purple). Regions 3000bp upstream and downstream of the peak summits of SET-9 & SET-26 binding are shown. The light grey and light purple areas indicate 95% confidence intervals. (F), (G) Venn diagrams showing genes with altered expression in the F1 and F3 *set-9(rw5) set-26(tm2467)* double mutant identified by RNA-seq and genes with altered H3K4me3 in the F1 and F3 *set-9(rw5) set-26(tm2467)* double mutant identified by ChIP-seq. (***p<0.001 indicate less than expected). H) Metagene plots showing the average H3K4me3 z-score (standardized log2 ratios of ChIP/H3ChIP signals) in *glp-1(e2141)* (black) and *glp-1(e2141); set-26(tm2467)* (orange) for H3K4me3 peaks bound by SET-9 & SET-26 (left), not bound by SET-9 & SET-26 (middle) and SET-9 & SET-26 peaks overlapping H3K4me3 (right). The light grey and light purple areas indicate 95% confidence intervals.

elevated levels of H3K4me3 were detected starting at around +500bp and -1000bp beyond the summit in the F1 and F3 *set-9 set-26* double mutant compared to wild-type worms (Figure 8B). Interestingly, for the H3K4me3 peaks that showed no detectable SET-9 & SET-26 binding, the average plot showed no detectable “spreading” of the H3K4me3 marking (Figure 8C). We further examined the observed “spreading” of H3K4me3 marking by producing meta plot of the H3K4me3 peaks that overlapped with SET-9 & SET-26 binding that centered around the summits of SET-9 & SET-26 binding peaks (Figure 8-figure supplement 1E). We again observed an insignificant increase at the center of the plot, where SET-9 & SET-26 binding peaked, but a detectable increase of H3K4me3 levels in regions flanking the summit of SET-9 & SET-26 binding sites in the F1 and F3 *set-9 set-26* double mutant compared to wild-type worms. It is interesting to note that the F3 *set-9 set-26* double mutant showed a more obvious elevation of H3K4me3 compared to the F1 *set-9 set-26* double mutant, even though this difference was not statistically significant (Figure 8B, 8C and Figure 8-figure supplement 1E). We previously reported that the levels of H3K4me3 as detected by Western blotting was not impacted by the loss of *set-26* in the germlineless *glp-1* mutant (Ni et al. 2012), suggesting that inactivation of SET-26 does not impact the global levels of H3K4me3 in the soma. To further investigate this possibility, we performed similar H3K4me3 ChIP-seq analyses in the germlineless *glp-1; set-26* and *glp-1* mutants. Consistent with our previous observation, we observed no significant difference in the genome-wide patterns of H3K4me3 in germlineless worms with or without SET-26 (Figure 8-figure supplement 1H). Together, these results suggested that SET-9 & SET-26 binding normally help to restrict H3K4me3 marking, likely specifically in the germline.

**A subset of germline-specific genes showed correlated changes in H3K4me3 and RNA expression levels in the F1 and F3 *set-9 set-26* double mutants**

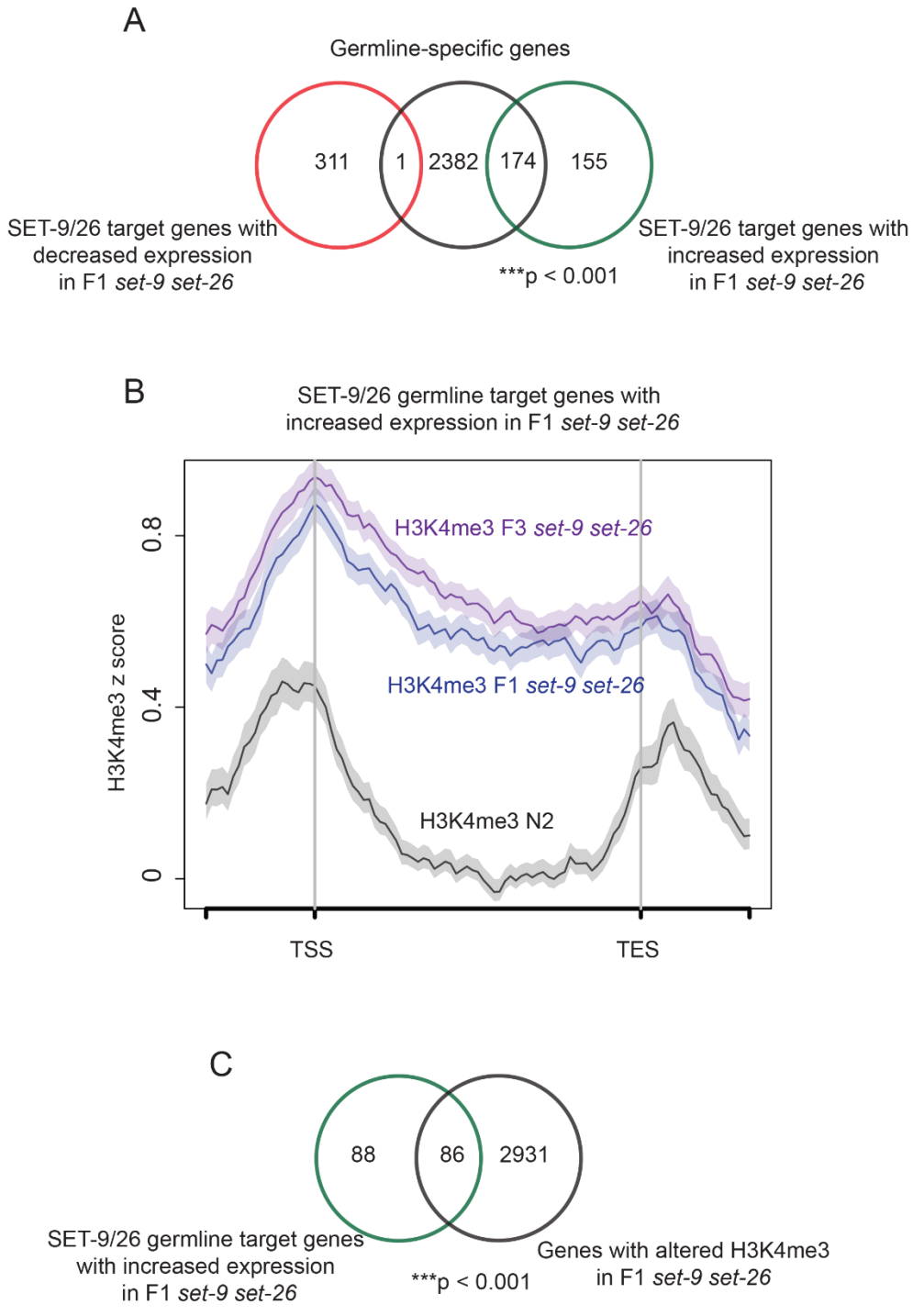
H3K4me3 has been generally associated with active gene expression (Sims et al. 2007). We wondered whether the regions with expanded H3K4me3 marking could be associated with gene expression changes in the F1 and F3 *set-9 set-26* double mutant. To assess this, we used csaw, an R package for differential binding analysis of ChIP-seq data using sliding windows, to identify 3438 and 5456 regions that exhibited statistically significant different H3K4me3 marking in F1 and F3 *set-9 set-26* double mutant compared with wild-type worms (data not shown). Consistent with a global increase in H3K4me3 levels in the F1 and F3 *set-9 set-26* double mutant, 92% and 99% of the differential H3K4me3 regions showed elevated H3K4me3 levels in the F1 and F3 *set-9 set-26* double mutant. We next assigned these differential H3K4me3 peaks to their closest genes and identified 3017 and 4517genes that are associated with altered H3K4me3 marking (Table S6). We then compared these gene lists with the list of genes that exhibited significant expression change between the F1 and F3 *set-9 set-26* double mutant compared to wild-type worms based on our earlier RNA-seq data. In this comparison, only 479 and 310 genes with elevated H3K4me3 marking showed expression change in F1 and F3 *set-9 set-26* double mutant (Figure 8-figure supplement 1F and 1G), an overlap that was significantly lower than what would be expected based on random chance. These results suggested that the expanded H3K4me3 regions, which likely contribute to the overall elevated H3K4me3 levels detected in the F1 and F3 *set-9 set-26* double mutant, were not generally accompanied

with detectable gene expression changes based on comparison with RNA-seq analyses from whole worms.

Since our data suggested that the “spreading” of H3K4me3 surrounding SET-9 & SET-26 binding sites likely occur specifically in the germline, we wondered whether the SET-9 & SET-26 target genes in the germline would show a correlation between H3K4me3 “spreading” and gene expression increase. To investigate this possibility, we compared the H3K4me3 profiles for the genes that were bound by SET-9 & SET-26 and showed expression increase in F1 *set-9 set-26* double mutant compared with wild-type worms (Figure 7), and were determined to be “germline-specific” based on previous reports (Figure 5 & Figure 8- figure supplement 2A). We found that their average H3K4me3 markings were highly upregulated in the F1 and F3 *set-9 set-26* double mutant compared with wild-type worms, and the elevation was noticeable at both the TSS and the TES and throughout the gene body (Figure 8- figure supplement 2B). A comparison of this list of germline-specific SET-9 & SET-26 target genes also indicated that their H3K4me3 elevation (based on csaw) was significantly correlated with gene expression increase (based on RNA-seq) (Figure 8- figure supplement 2C).

**RNAi screens revealed that SET-2, the H3K4me3 methyltransferase, functions cooperatively with SET-9 & SET-26 to regulate germline function**

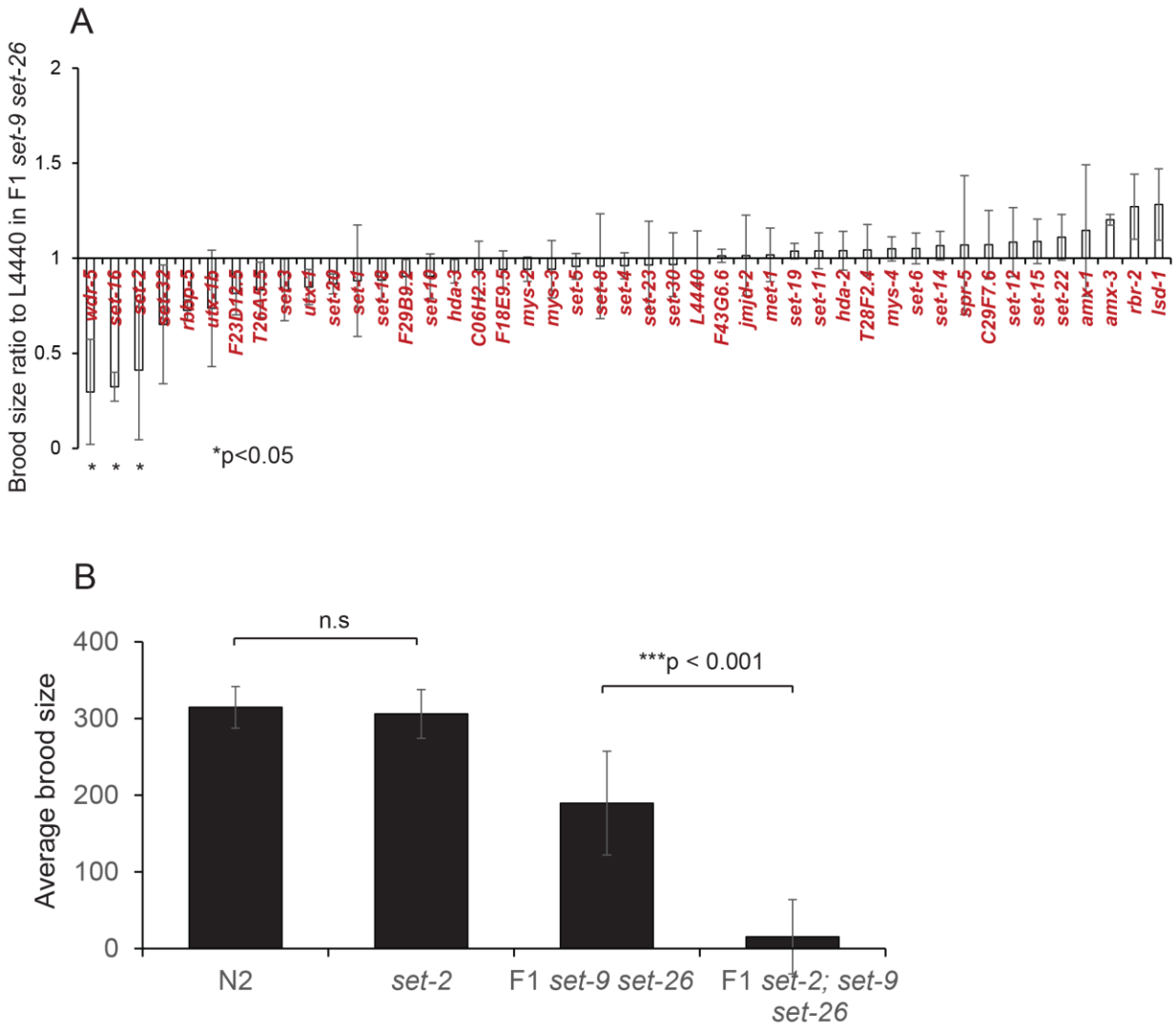
Since many histone methyltransferases and demethylases are known to play important roles in germline function in *C. elegans* (Katz et al. 2009; Li and Kelly 2011; Nottke et al. 2011; Xiao et al. 2011; Robert et al. 2014; Kerr et al. 2014; Greer et al. 2014), we


A
Germline-specific genes
311
1
2382
174
155
SET-9/26 target genes with decreased expression in F1 set-9 set-26
***p < 0.001
SET-9/26 target genes with increased expression in F1 set-9 set-26
B
SET-9/26 germline target genes with increased expression in F1 set-9 set-26
H3K4me3 z score
0.8
0.4
0
H3K4me3 F3 set-9 set-26
H3K4me3 F1 set-9 set-26
H3K4me3 N2
TSS
TES
C
88
86
2931
SET-9/26 germline target genes with increased expression in F1 set-9 set-26
***p < 0.001
Genes with altered H3K4me3 in F1 set-9 set-26

## Figure 8-figure supplement 2 Germline-specific genes show correlated changes in H3K4me3 and RNA expression.

(A) Venn diagrams show the overlaps between SET-9/26 target genes with expression change in F1 *set-9(rw5) set-26(tm2467)* and germline-specific genes. (B) H3K4me3 z-score of SET-9/26 germline target genes with increased expression in F1 *set-9(rw5) set-26(tm2467)* in N2 (black) and the F1 (blue) and F3 (purple) *set-9(rw5) set-26(tm2467)* double mutant. (C) Venn diagram shows the overlap between SET-9/26 germline target genes with increased expression in F1 *set-9(rw5) set-26(tm2467)* double mutant and genes with altered H3K4me3 in F1 *set-9(rw5) set-26(tm2467)* double mutant. Gene lists can be found in table S6.

performed an RNAi screen targeting putative histone methyltransferases and demethylases to uncover genes that potentially work with SET-9 & SET-26 to regulate germline function. We treated F1 *set-9 set-26* double mutant worms, which showed a mild reproductive defect (Figure 2), with each of the RNAi and assayed their consequent brood size. Whereas most of the RNAi treatment did not substantially affect the brood size of the F1 *set-9 set-26* double mutant, we found that three components of the MLL complex, including *set-2*, *set-16* and *wdr-5.1*, significantly reduce the brood size of the F1 *set-9 set-26* double mutant when knocked down (Figure 9A). To rule out off-target effects, we crossed the partial loss-of-function *set-2(ok952)* mutant with the *set-9 set-26* double mutant and assayed their brood size at the F1 generation. We again observed a synergistic effect, where the *set-2(ok952)* single mutant had a normal brood size as previously reported, but the F1 *set-2; set-9 set-26* triple mutant had a drastically reduced brood size compared with the F1 *set-9 set-26* double mutant (Figure 9B). This result is consistent with the model that SET-2 and SET-9 & SET-26 regulate H3K4me3 marking in different ways. Our data above indicated that SET-9 & SET-26 bind to H3K4me3 and restrict H3K4me3 domain, especially in the germline. The results with the weak loss-of-function *set-2* mutation suggested that suboptimal deposition of H3K4me3 together with broadening of H3K4me3 marking could cause detrimental effect to germline function.


A
Brood size ratio to L4440 in F1 set-9 set-26
2
1.5
1
0.5
0
*p<0.05
wdr-5
set-16
set-2
set-32
rbbp-5
utx-1b
F23D12.5
T26A5.5
set-3
utx-1
set-20
set-1
set-18
F29B9.2
set-10
hda-3
C06H2.3
F18E9.5
mys-2
mys-3
set-5
set-9
set-4
set-23
set-30
L4440
F43G6.6
jmjd-2
met-1
set-19
set-11
hda-2
T28F2.4
mys-4
set-6
set-14
spr-5
C29F7.6
set-12
set-15
set-22
amx-1
amx-3
rbr-2
lsd-1
B
Average brood size
400
300
200
100
0
n.s
***p < 0.001
N2
set-2
F1 set-9 set-26
F1 set-2; set-9 set-26

## Figure 9 SET-9 & SET-26 and SET-2 act synergistically to regulate fertility.

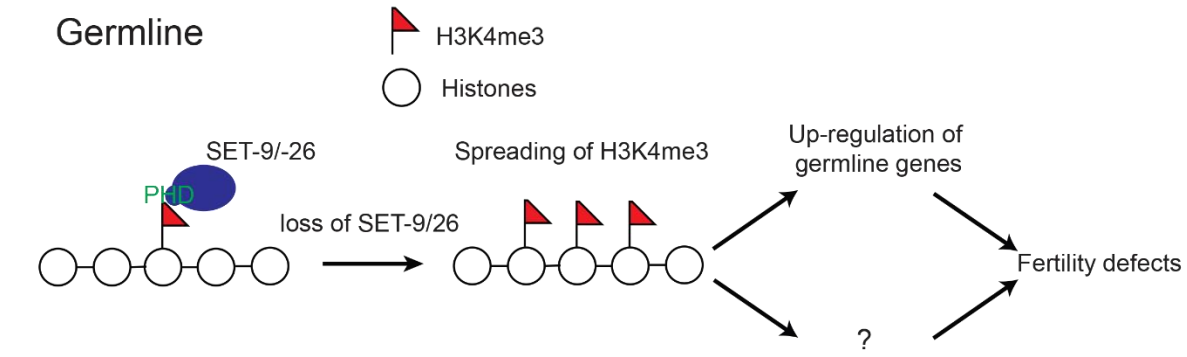
(A) Ratio of eggs laid by F1 *set-9(rw5) set-26(tm2467)* double mutant worms fed dsRNA of *C. elegans* potential histone modifiers (mainly methyltransferases and demethylase) or empty vector (L4440) for one generation. (B) Average brood size in N2, *set-2(ok952)*, F1 *set-9(rw5) set-26(tm2467)* double mutant and F1 *set-2(ok952)*; *set-9(rw5) set-26(tm2467)* triple mutant.

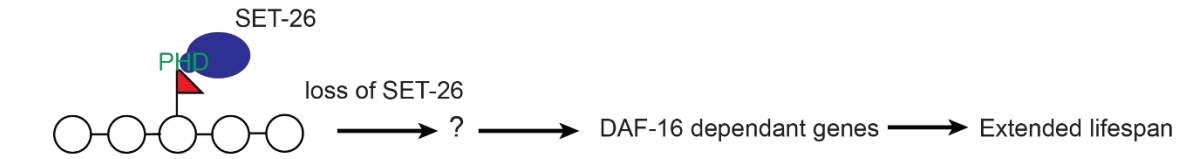
# Discussion

In this study, we revealed that the highly homologous paralogs SET-9 and SET-26 have unique as well as redundant functions. We found that SET-9 and SET-26 share redundant function in germline development, but only SET-26, not SET-9, plays a role in modulating lifespan and resistance to heat stress. We confirmed that SET-9 is only detectable in the germline whereas SET-26 is broadly expressed, and we concluded that their differential expression patterns likely account for the different phenotypes associated with the loss of each gene. Our transcriptomic analyses corroborated previous genetic studies and implicated SET-26 to act through DAF-16 to modulate longevity. We further demonstrated that SET-9 and SET-26 bind to H3K4me3 *in vitro* and *in vivo*, and that the loss of *set-9* & *set-26* results in the broadening of H3K4me3 domains surrounding most SET-9 & SET-26 binding regions, likely specifically in the germline. The expanded H3K4me3 domains were associated with increased expression of germline-specific genes bound by SET-9 & SET-26. We propose that in the soma, SET-9 and SET-26 somehow impact DAF-16 activity to modulate longevity, whereas in the germline, they are critical for restricting H3K4me3 domains and regulating the expression of specific target genes with important consequence on germline development (Figure 10).

### *set-9* and *set-26* are duplicated genes

*set-9* and *set-26* share high sequence identity both within the coding regions and in the non-coding sequences flanking the coding regions (near 90% identity in the +/− 500bp


Germline
H3K4me3
Histones
SET-9/-26
PHD
loss of SET-9/26
Spreading of H3K4me3
Up-regulation of germline genes
Fertility defects
?


Somatic tissues
SET-26
PHD
loss of SET-26
?
DAF-16 dependant genes
Extended lifespan

## Figure 10 SET-9 & SET-26 regulate H3K4me3 and target gene expression.

In germline, the loss of *set-9* & *set-26* results in the broadening of H3K4me3 domains surrounding most SET-9 & SET-26 binding regions and results in up-regulation of germline genes. In somatic tissues, loss of *set-26* modulates lifespan by indirectly regulating DAF-16 dependent genes. We propose that SET-9 and SET-26 are critical for organizing local chromatin environment and regulating the expression of specific target genes, and these activities together contribute to their roles in germline development and longevity.

regions), and even in the genes 3' of *set-9* and *set-26* (Y24D9B.1 and Y51H4A.13 respectively). Although *set-9 & set-26* are highly conserved from worms to yeast to mammals, only one homolog in each of the other diverse species has been detected based on sequence and domain structure alignment (Rincon-Arano et al. 2012; Zhang et al. 2017). Even the closely related *Caenorhabditis* species, such as *C. briggsae*, *C. remanei*, *C. brenneri* and *C. japonica*, only harbor one gene that is highly homologous to *set-9 & set-26*. These results suggested that *set-9* and *set-26* have arisen from gene duplication specifically in *C. elegans*. In the future, it will be interesting to explore how this duplication event allowed *set-26* and *set-9* to adopt unique functions.

**SET-9 and SET-26 in chromatin regulation**

Based on the similarity in sequence and domain structure, *Drosophila* UpSET and mammalian SETD5 and MLL5 represent the likely homologs of SET-9 & SET-26. SET-9 & SET-26, UpSET, SETD5 and MLL5 all harbor centrally localized SET and PHD domains, and they share ~30-40% sequence identity in their PHD and SET domains.

The SET domains of SET-9, SET-26, UpSET, SETD5 and MLL5 proteins are highly conserved and all share similar mutations that suggest they should lack enzymatic activities (Figure 6-figure supplement 1A). Indeed, UpSET, SETD5 and MLL5 have not been found to harbor methylating activities (Osipovich et al. 2016; Rincon-Arano et al. 2012; Madan et al. 2009). Whereas the SET domain of SET-26 has been suggested to methylate H3K9me3 *in vitro* (Greer et al. 2014), our data indicated that SET-9 and SET-

26 do not have major roles in H3K9me3 deposition in *C. elegans*. The exact function(s) of the SET domain in these proteins remain to be elucidated.

The PHD domains of SET-9, SET-26, UpSET and MLL5 proteins are also highly conserved. The PHD domain of MLL5 was found to bind to H3K4me3 *in vitro* (Ali et al. 2013), and both MLL5 and UpSET have been shown to localize at promoter regions in cultured cells (Rincon-Arano et al. 2012; Ali et al. 2013). Our results demonstrated that SET-9 & SET-26 also bind to H3K4me3 *in vitro* and *in vivo*. We further revealed that the PHD domains of SET-9 & SET-26 bind to H3K4me3 with adjacent acetylation (e.g. K9, K14, K18) with much greater affinity (Figure 6B). This is an important finding and potentially links SET-9 & SET-26 to the regulation of both H3K4me3 and histone acetylation. Indeed, we showed that loss of *set-9* & *set-26* results in a global elevation of H3K4me3 and H3K9Ac levels. We additionally uncovered that the elevated global levels of H3K4me3 represent expansion of H3K4me3 domains surrounding SET-9 & SET-26 bound regions, and we predict that similar expansion of H3K9Ac likely occurs in the *set-9 set-26* double mutant. Interestingly, loss of *set-26* in somatic tissues does not result in similar H3K4me3 elevation (Figure 8-figure supplement 1H). These data are reminiscent to those reported for MLL5 and UpSET (Rincon-Arano et al. 2012; Gallo et al. 2015). In human glioblastoma cells with self-renewing potential, it was found that knockdown of MLL5 leads to increased global levels of H3K4me3 and a more open chromatin environment (Gallo et al. 2015). Importantly, this anti-correlation between MLL5 and H3K4me3 was only detected in primary glioblastoma cells with self-renewing potential, but not in bulk glioblastoma samples, nor in non-neoplastic brain samples or colon cancer cells (Gallo et al. 2015), indicating a regulatory process that is highly cell

type specific. A possible parallel in *C. elegans* is that the anti-correlation between SET-9 & SET-26 status and H3K4me3 levels is particularly obvious in the gonads of adult worms (Figure 6-figure supplement 1B), which harbor proliferative germline stem cells. In *Drosophila* Kc cells, UpSET knockdown has been shown to increase the levels and spreading of H3K9Ac, H3K16Ac, and H3K4me2/3 around some TSSs, which results in a generally more accessible chromatin environment (Rincon-Arano et al. 2012). UpSET achieves this partly through recruitment of specific histone deacetylases. Although the role of SETD5 PHD domain has not been studied, SETD5 also has been found to interact with histone deacetylase complex (Osipovich et al. 2016). Loss of SETD5 causes elevation in histone acetylation at transcriptional start sites and near downstream regions (Osipovich et al. 2016), suggesting that it acts in a similar manner to *Drosophila* UpSET. Taking together the *C. elegans*, *Drosophila*, and mammalian data, it appears that the SET-9, SET-26, UpSET, SETD5, MLL5 family of factors play an important role in binding to H3K4me3, possibly with flanking acetylation sites, and regulating local chromatin accessibility, partly through restricting the spread of histone modifications such as H3K4 methylation and H3K9 acetylation. Extending from the findings with UpSET in *Drosophila* and SETD5 in mammals, SET-9 and SET-26 likely also recruit demethylating and deacetylating enzymes to confine local methylation and acetylation domains.

**Biological functions of SET-9 and SET-26**

In this study, we uncovered a novel redundant function of SET-9 and SET-26 in germline function. It is interesting to note that an earlier report suggested that loss of

*set-26*, but not *set-9*, accelerated the progressive sterility of the *spr-5* mutant (Greer et al. 2014). However, our analyses indicated that the *set-9(n4949)* mutant used in the study (Greer et al. 2014) represents a deletion/duplication allele, as we can PCR amplify the *set-9* gene from the mutant (data not shown). Our data using the newly generated *set-9(rw5)* deletion/frame-shift mutant supported a role of SET-9 in collaboration with SET-26 to regulate germline development.

In considering the possible homologs of SET-9 and SET-26 in other species, MLL5 has been implicated in male fertility (Heuser et al. 2009; Madan et al. 2009), whereas UpSET is important for female fertility (Rincon-Arano et al. 2012). Our data indicated that SET-9 & SET-26 have a more pleiotropic role in the germline, affecting both germline stem cell proliferation, and the subsequent differentiation into oocytes and sperms. This difference could be due to that *C. elegans* are hermaphrodites and SET-9 & SET-26 have adopted broader functions in the germline. In addition, MLL5 has been implicated in regulating cell cycle progression (Deng et al. 2004) and stem cell pluripotency (Zhang et al. 2009b), which may parallel the redundant roles of SET-9 & SET-26 in germline stem cells proliferation. Considering the highly conserved functions of SET-9, SET-26, UpSET, SETD5 and MLL5 at the molecular and phenotypic levels, an intriguing possibility is that UpSET, SETD5 and MLL5 have a yet-to-be characterized role in stress response and longevity.

A key important question is how SET-9 & SET-26 can mediate their effects on stress response, longevity and germline development. Our data clearly supported that SET-26 acts in the soma and through DAF-16 to modulate longevity. The exact functional

relationship between SET-26 and DAF-16 remains unclear. It was striking that the comparison of the SET-9 & SET-26 bound targets from whole worms and the DAF-16-mediated transcriptomic changes in germlineless mutant not only did not show a significant overlap, it in fact showed an overlap that was much lower than expected by random chance. It seems clear that SET-9 and SET-26 do not bind to candidate DAF-16 regulated genes. Therefore, SET-9 and SET-26 likely mediate their effect on DAF-16 indirectly.

We also note that while we were able to identify SET-9 & SET-26 targets whose expression show significant change in response to the loss of *set-9* & *set-26*, the SET-9 & SET-26 binding targets were not enriched for genes that showed expression change in the *set-9 set-26* double mutant. In order words, genes bound by SET-9 & SET-26 were not more likely to show expression change when *set-9* & *set-26* were deleted. Our data cannot resolve whether SET-9- & SET-26-binding has a direct consequence on gene expression regulation.

We detected significant broadening of H3K4me3 marking surrounding the SET-9 & SET-26 bound regions. Although this change was observed in whole worms, it was not detectable in worms lacking germline (Figure 8-figure supplement 1H). We interpreted this to mean that the expansion of H3K4me3 domains likely happen either specifically in the germline, or most noticeable in the germline. Somewhat surprisingly, such broadening of H3K4me3 domains did not appear to correlate with gene expression using whole worm analyses. This finding could partly be due to technical caveats, as the ChIP-seq and RNA-seq data were generated using whole worms, which could mask

correlated changes in specific cells. Consistent with this speculation, we detected a significant and positive correlation between H3K4me3 expansion with RNA expression change of germline-specific genes. In the future, ChIP-seq and RNA-seq analyses using dissected gonads *vs* somatic tissues will help to further establish this prediction.

In addition to gene regulation, the expanded H3K4me3 regions likely represent perturbed chromatin environment that could interfere with processes other than gene regulation, such as genome maintenance. In fact, in *C. elegans*, altered H3K4 methylations have been shown to predispose mutant worms to DNA damage and genome instabilities, which result in germline defects (Nottke et al. 2011). Interestingly, inactivating UpSET in *Drosophila* and MLL5 in mammals also lead to increased genome instability and DNA damage (Tasdogan et al. 2016; Rincon-Arano et al. 2012). In *C. elegans*, we observed an increased number of germ cell apoptosis in the *set-9 set-26* double mutant (data not shown). Moreover, RNAi knockdown of *mre-11*, a double-strand break repair protein, synergistically aggravated the fertility defects of the *set-9 set-26* double mutant (data not shown). These data together suggested that DNA damage and genome instability could be increased in the *set-9 set-26* double mutant, which may contribute to their germline defects.

In summary, our findings provided new mechanistic insights into the functions of SET-9 & SET-26 with important implications for their roles in longevity and germline development. We revealed that SET-9 & SET-26 are recruited to H3K4me3 marked regions and participate in confining H3K4me3 domains, particularly in the germline. Loss of *set-9* & *set-26* in these regions likely leads to a more open and accessible

chromatin environment. In addition, SET-26 acts in the soma to modulate longevity and it appears achieve this by regulating DAF-16-mediated transcription indirectly. We propose that SET-9 & SET-26 functions are important for their roles in germline development, stress response, and longevity. Our molecular findings are consistent with the possibility that human MLL5 and SETD5 represent functional homologs of SET-9 & SET-26. Our findings provide new insights into how MLL5 and SETD5 may act and also implicated them in stress response and longevity.

## Experimental Procedures

### Worm strains

The N2 strain was used as the wild-type (WT). The mutants used in the study are: *set-2(ok952), set-9(rw5)*, *set-26(tm2467)*, *rrf-1(pk1417)*, the *set-9(rw5) set-26(tm2467)* double mutant was maintain as balanced heterozygote *set-9(rw5) set-26(tm2467)* /nT1. The GFP knock-in strains constructed in this study are: *set-9::gfp(rw24)*, *set-26::gfp(rw25)*.

### Antibodies

The antibodies used were anti-H3K4me3(Millipore, Billerica, MA 17-614), anti-H3K9me3(abcam, Cambridge, United Kingdom ab8898), anti-GFP(abcam, Cambridge, United Kingdom ab290), anti-H3(abcam, Cambridge, United Kingdom ab1791), anti-H3K9ac(Wako, Richmond, VA 309-32379).

**Brood size assay**

Brood size assay was performed as described (Li et al. 2008). Each single L4 worm was picked onto individual plate and was transferred to a new plate every 24 hours until the end of its reproductive phase. Dead eggs and alive progenies were counted as its total brood size. All experiments were repeated two to three times. Student's t-test was used to calculate the p-values.

**Lifespan assay**

Lifespan assay was performed as described (Li et al. 2008). All experiments were performed at 20°C. For RNAi plates, the *set-9/26* RNAi construct was taken from the Ahringer RNAi library. RNAi bacteria were grown in LB with 100ug/ml Carbenicillin (Carb) and 15ug/ml Tetracycline (Tet) at 37°C to OD600 around 0.8. The culture was concentrated 5-fold, and seeded onto plates with Carb and Tet. Sufficient IPTG stock was added to plates so that final IPTG concentration is 4mM. Let plates dry and induce for ~4 hrs before use. For lifespan assays using NGM plates seeded with OP50 bacteria, the OP50 bacteria were grown in LB overnight at 37°C and the culture was concentrated 3-fold and seeded onto plates. For all lifespan assays, Worms were picked onto RNAi plates to lay ~40 eggs per plate and the progeny were grown on the plate until they reached adulthood. The worms were transferred to a new plate every day during their reproductive phase and then transferred to a new plate every 4 days. Worms were scored every other day, and those that failed to respond to a gentle

prodding with a platinum wire were scored as dead. Animals that bagged, exploded, or crawled off the plate were considered as censored. We defined the day when we transferred the young adult worms as day 0 of adult lifespan. All the lifespan experiments were repeated at least two independent times. Standard survival analysis was performed using SPSS and OASIS software (Yang et al. 2011). The survival function was estimated using the Kaplan-Meier method, and the p-values were calculated using a log-rank test.

**Heat stress assay**

Heat stress assay was performed as described (Li et al. 2008). Synchronous D0 adult worms grown on OP50-NGM plates at 20°C were shifted to 35°C. Worms were scored every 3-4 hours. All the heat stress experiments were repeated at least two independent times. All experiments were repeated two to three times. Standard analysis was performed using OASIS software. The survival function was estimated using the Kaplan-Meier method, and the p-values were calculated using a log-rank test.

**Analysis of germline mortality**

Mortal Germline assays were performed by transferring 6 L1 larvae per plate to fresh NGM plates every generation, as previously described (Ahmed and Hodgkin 2000). Percentage fertile lines were calculated as the number of fertile plates divided by the number of total plates.

**Immunofluorescence Staining**

The gonads from 100-150 worms were dissected out using syringe needle on a poly-lysin coated slide. Gonads were permeabilized by the standard freeze-crack method and were fixed in 4% formaldehyde fixative (PBS / 4 % formaldehyde) for 30 minutes followed by 5-minute incubation in chilled methanol at -20°C in a Coplin jar. Slide was washed twice with PBST for 5 minutes each and finally incubated with blocking buffer (PBS + 5% BSA + 0.1% tween-20 + 0.1% triton-100) for 1 hour. Blocking buffer was removed and gonads were incubated with 25 µl of 1:50 dilution of H3K4me3 antibodies for overnight at 4°C in a humidity chamber. Next day, gonads were washed twice with PBST for 10 minutes each and incubated with 25 µl of 1:50 dilution of secondary antibody for 1-2 hours at RT. DAPI (final concentration 25 ng/ µl) was also added with secondary antibody. Gonads were further washed three times with PBST for 10 minutes each and mounted in 10 µl vectashield mounting medium (Vecta Laboratories). Experiments were repeated twice. 4 worms and 5-7 germline nuclei for each worm were used for quantification of H3K4me3 and DAPI. Student's t-test was used to calculate the p-values.

**DAPI staining and measuring the number of germ cells within the mitotic region:**

The gonads from 100-150 worms were dissected out using syringe needle. Gonads were fixed in 4% formaldehyde fixative (PBS / 4 % formaldehyde) for 1 hour. After the removal of fixative, gonads were washed twice with 1 ml PBST (PBS buffer + 0.1 %

Tween 20) each. Fixed worms were incubated with DAPI (25 ng/ µl in PBS) for 30 minutes. Gonads were further washed three times with 1 ml PBST for 5 minutes each and mounted on agarose pads in 10 µl vectashield mounting medium (Vecta Laboratories). DAPI-stained gonad images were taken using the Z-stacking function of the microscope. To count mitotic germ cell numbers, we marked the boundary between mitotic zone and transition zone by observing the transition zone-specific nuclear morphology (crescent shape) and counted the number of nuclei in each focal plane within the mitotic region. Experiments were repeated three times. 6-9 worms were used for quantification. Student's t-test was used to calculate the p-values.

**GST protein pull-down assay**

pGEX-2TK was used for generating bacteria expression construct. SET-9/26$_{PHD}$ cDNA was amplified from wildtype cDNA by PCR using the primers: SET-9/26$_{PHD}$-F: CTCAGGATCCGATTCCGAATCCGAGGGAA; SET-9/26$_{PHD}$-R: GCGTGAATTCCCGCTCGAAGTCGATTCAAAA; and subcloned into pGEX-2TK.

The expression construct was transformed into BL21 bacteria. Bacteria containing cDNA of SET-9/26$_{PHD}$ was culture to OD600 equal to around 0.6 and the SET-9/26$_{PHD}$ expression was induced by adding IPTG to final concentration equals to 0.5mM. Bacteria were lysed by sonication and protein was purified by Glutathione Sepharose (Sigma).

GST protein pull-down assay was performed as described (Tsai et al. 2010). 25 µg GST-PHD protein were incubated with 10, 100, 1000 µg of calf thymus total histones

(Sigma) in 500 µl NTP overnight at 4 °C. 90 µl of a 50% slurry of GST-beads were added and incubated for 2 h at 4 °C, recovered by centrifugation and washed 6 times (10 min at 4 °C) with NTP buffer (50 mM Tris-HCl 7.4, 300mM NaCl, 0.1% NP-40). The protein bound beads were analyzed by SDS-PAGE and detected by Coomassie stain.

**Array binding assay**

The array binding assay was performed by EpiCypher. Briefly, GST-PHD were applied on an EpiTitan™ array that was separated by a gasket such that two chambers were delineated. After the protein incubation, a series of the anti-GST (primary) and the fluorescent AlexaFluor 647 (secondary) antibody incubation steps were carried out to detect the bound protein. Two independent experiments were performed.

**Immunoblotting and Quantification**

Immunoblotting was performed as described (Ni et al. 2012). Synchronized embryos were put onto NGM-OP50 plates and grown at 20°C until reaching mid-L4. Worms were harvested and washed 3 times using ice-cold M9 buffer. Worm pellets were lysed with boiling SDS sample and equal amount of lysates were used for SDS-PAGE(18%) and transferred to nitrocellulose membranes. The membranes were incubated with primary antibodies overnight (H3K4me3/H3K9me3/H3K9ac, 1:1,000; H3, 1:2,000). IRDye secondary antibody was used and the result was quantified using Odyssey imaging system. All experiments were repeated 3-4 times. Modified histone levels were normalized to H3 levels. Student’s t-test was used to calculate the p-values.

**Plasmid and homologous DNA repair template construction**

pU6::set-9 sgRNA was generated using pU6::unc-119 sgRNA as described (Friedland et al. 2013). The pU6::unc-119 sgRNA was used as template to amplify two overlapping PCR fragments using the primers U6prom EcoRI F and set-9 gRNA R or set-9 gRNA F and U6prom HindIII R. These PCR products were gel-purified and then mixed together in a second PCR with primers U6prom EcoRI F and U6prom HindIII R. This final PCR product was digested with EcoRI and HindIII and ligated into a pU6::unc-119 sgRNA plasmid that had been digested with EcoRI and HindIII, creating the vector pU6::set-9 sgRNA. The pU6::set-9/set26 sgRNAs (1 and 2) that used for generating *set-9::gfp* and *set-26::gfp* strains were constructed using the same strategy. The homologous DNA repair templates that used for generating *set-9::gfp* and *set-26::gfp* strains were designed and synthesized as described (Paix et al. 2014). Homologous arms flanking gfp DNA sequence were generated using the pPD90 plasmid that contains gfp sequence as template and the primers: armF12-GFP: CGAGACGAAGCCGaTCtACgCGaTGGAAcagtaaaggagaagaacttttcactggagttg; armR12-GFP: caagtttttcgcagattccttgCTAtttgtatagttcAtccatgccatgtgtaatccc; The DNA repair template with ~30bp homologous arms was generated using the primer: armF1-complete: ccctcaatttttttcagCTGAAACAAACTCGAGACGAAGCCG; and armR1-complete: gggacaattttattcttcaagtttttcgcagattcc. The DNA repair template with ~60bp homologous arms was generated using the primer: armF2-complete: ccaaaaaatctccttaaaaaccctcaatttttttcagCTGAAACAAACTCGAGACGAAGC; and armR2-complete: cgagatagaaagagatgatatgggacaattttattcttcaagtttttcgcagattcc.

**CRISPR-mediated genome editing**

CRISPR-mediated genome editing was performed as described (Paix et al. 2014; Friedland et al. 2013). For generating *set-9* mutant, day1 adult animals were injected with pDD162 (Peft-3::Cas9::tbb-2 3'UTR), pCFJ90 (pmyo-2::mCherry) and pU6::set-9 sgRNA and grown overnight at 16°C. The survived worms were separated, transferred to 20°C and their F2 mCherry-positive animals were genotyped for *set-9* mutation. For generating *set-9::gfp* and *set-26::gfp* strains, day1 adult animals were injected with pDD162 (Peft-3::Cas9::tbb-2 3'UTR), pCFJ90 (pmyo-2::mCherry), pU6::set-9/set26 sgRNAs (1 and 2) and homologous DNA repair templates and grown overnight at 16°C. The survived worms were separated, transferred to 20°C and their F2 mCherry-positive animals were genotyped for GFP knockin strains. sgRNAs that target two loci and two DNA repair templates with different length of homologous arms ( one ~30bp and the other ~60bp) were co-injected to increase efficiency. The *set-9::gfp* and *set-26::gfp* strains were mounted on a microscope slide and visualized using a Zeiss 710 confocal system.

**Chromatin immunoprecipation following sequencing (ChIP-seq)**

ChIP experiments for SET-9::GFP and SET-26::GFP were performed as described (Zhong et al. 2010) with the following modifications. Approximately 70,000-100,000 L4s were harvested and crosslinked in 2% formaldehyde-M9 solution for 25 minutes at room temperature with rotation. The worms were then washed with 100 mM Tris pH 7.5 to

quench formaldehyde solution, washed two times with M9, and once with FA buffer (50 mM HEPES/KOH pH 7.5, 1 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate; 150 mM NaCl) supplemented with 2X protease inhibitors (Roche Cat#11697498001, cOmplete Protease Inhibitor Cocktail Tablets). Worms were then collected in a 15 ml conical tube, snap-freezed in liquid N2 and stored at -80°C. The pellet was resuspended in 1 ml FA buffer plus protease inhibitors. Using a Bioruptor sonicator, the sample was sonicated on ice/salt water 30 times with the following settings: 30 sec on, 60 sec off. The chromatin was then further sheared by the covaris s2 40 times with the following settings: 20% duty factor, intensity 8, 200 cycles per burst, 60 sec on, 45 sec off. The tube spun containing worm extract was then spun at 13,000g for 30 minutes at 4°C. The supernatant was then transferred to a new tube and the protein concentration of the supernatant was then determined by Bradford assay. Extract containing approximately 2 mg of protein was incubated in a microfuge tube with 6-12ul anti-GFP antibodies overnight at 4°C with gentle rotation. 10% of the material was removed and used as input DNA. Then 30 ul of protein A conjugated to sepharose beads (EMD Millipore) were added to each ChIP sample and rotated at 4°C for 4 hours. The beads were then spun at 2000rpm for 1 min to collect and washed twice for 5 mins each at 4°C in 1 ml of FA buffer, once in FA with 500 mM NaCl and once in FA with 1M NaCl with gentle rotation. The beads were then washed in TEL buffer (0.25 M LiCl, 1% NP-40, 1% sodium deoxycholate, 1 mM EDTA, 10 mM Tris-HCl, pH 8.0) for 5 min and twice in TE for 5 min. To elute the immunocomplexes, 50 ul Elution Buffer (1% SDS in TE with 250 mM NaCl) was added and the tube incubated at 65°C for 15 min, with brief vortexing every 5 min. The beads were spun down at 2000rpm for 1 min and the supernatant

transferred to a new tube. The elution was repeated and supernatants combined. To each sample, RNAseA was added and incubated at 37°C for 15 min, and proteinase K was added and incubated for 1 hour at 55°C, then 65°C overnight to reverse crosslinks. The DNA was then purified with the Qiaquick PCR purification kit (Qiagen), and eluted with 40 ul H2O. The immunoprecipitated DNA was either checked by qPCR or subjected to high-throughput sequencing library preparation. The protocol for library preparation for SET-9/26 ChIP–Seq is NEBNext® Ultra™ II DNA Library Prep Kit for Illumina® (NEB).

ChIP experiments for H3K4me3 and H3K9me3 were performed as described (Pu et al. 2015). L4 worm pellets were ground with a mortar and pestle and cross-linked with 1% formaldehyde in PBS for 10 min at room temperature. Worm fragments were collected by spinning at 3000g for 5 min and resuspended in FA buffer followed by sonication with Bioruptor. Chromatin extract was incubated with antibody overnight at 4°C. Precipitated DNA (10–15 ng) from each sample was used for Illumina sequencing library preparation. DNA from ChIP was first end-repaired to generate a blunt end followed by adding single adenine base for adaptor ligation. The ligation product with adaptor was size-selected and amplified by PCR with primers targeting the adaptor. Up to 12 samples were multiplexed in one lane for single-end 50-nt Illumina HiSeq sequencing.

All ChIP-seq experiments were repeated at least two times.

**ChIP-seq data analysis**

The data analysis pipeline was performed as described (Pu et al. 2015). Low quality reads were removed using the FASTX Toolkit. The sequencing reads from two independent experiments were combined and then aligned to the WS250 *C. elegans* genome using bowtie2. PCR duplicates were removed and bam files were generated using SAMtools. The bam files were then used for calling peaks by MACS2 (combined broad and narrow peaks). A GLM model was then applied to compute the counts of all the peaks to identify significant ones using two independent replicates. H3 ChIP data were used as a control for H3K4me3 ChIP and genomic DNA input was used as a control for SET-9 & SET-26 ChIP. ChIP signals (z-score) were normalized and calculated using bamCoverage software.

For analyses using genes, each peak was associated with its closest gene. Overlapping peaks were determined by 1bp overlap between two peaks using bedtools. For oriented meta plots, 3000bp upstream and downstream of the summit were included and 50bp windows were used for normalized counts in each window. The summits determined by MACS2 were assigned to their closest genes. If a gene is on the "-" strain, the normalized counts for -3000bp to 3000bp of that summit were counted in a reverse direction. For example, if a peak, which was assigned to a gene in the "-" strain, has a summit of X, the extended region for this summit in 5' to 3' direction is X+3000bp to X-3000bp; if a peak, which has assigned to a gene in the "+" strain, has a summit of Y, the extended region for this summit in 5' to 3' direction is Y-3000bp to Y+3000bp. The 95% confidence interval was calculated using bootstrap method. Briefly, 10% of the total summits were randomly selected and the mean was calculated. This random selection was repeated1000 times and the 95% confidence interval was calculated based on the

estimate that these 1000 mean number follow normal distribution (Hesterberg et al. 2005). Fishers' exact test was used when comparing two lists of genes or peaks.

H3K9ac ChIP-seq data was downloaded from: http://data.modencode.org/cgi-bin/findFiles.cgi?download=3578

**RNA isolation and library preparation for RNA-seq.**

RNA isolation was performed as described (Li et al. 2008). Synchronized mid/late L4 staged worms that grown at 20°C were homogenized in 1 ml Tri-reagent for 30 min at room temperature. 0.1mL of BCP was added to the sample and mixed well. The sample was then spun at 12,000 g 15 min at 4°C and the aqueous phase was transferred to a new tube. 0.5ml Isopropanol was added to the sample, incubated at room temperature for 10 min and spun at 12,000g 10 min. The RNA pellet was washed twice with 75% EtOH and dissolved in water. The RNA sample was then was then purified to remove DNA using RNeasy Mini Kit (Qiagen). The protocol for library preparation was using Ovation Human FFPE RNA-Seq Library Systems (NuGEN).

**RNA-seq data analysis**

The data analysis pipeline was performed as described (Pu et al. 2015). Low quality reads were removed using the FASTX Toolkit. Illumina primers (adaptors) were then removed using cutadapt. And tRNA and rRNA reads were removed using Bowtie and the remaining reads were aligned to WS250 *C. elegans* genome using TopHat2 with no

novel junctions. Mapped reads were then input into Cufflinks to calculate raw counts for each gene, which were then used for differential expression analysis by edgeR. Genes with less than 20 reads mapped to them in all samples were removed and the remaining genes were used as input to test for differential expression. PCA analysis was performed using the built-in function in edgeR. 5% false discovery rate (FDR) was used to determine differential expression.

**Heatmap**

Heatmaps were generated using ngs.plot software (Shen et al. 2014). H3K4me3 ChIP-seq results from wild-type and the *set-9 set-26* double mutant were used to generate the ranked heatmap, and the RNA-seq and SET-9 & SET-26 ChIP-seq data were plotted according to their order in the H3K4me3 heatmaps.

**RNAi screen**

RNA constructs were obtained from the Ahringher libarary. HT115 bacteria containing RNAi constructs were grown at 37°C and seeded on nematode growth medium (NGM) plates containing carbenicillin and tetracycline and dry overnight. dsRNA expression was induced by adding IPTG to a final concentration 0.4 mM. Heterozygous adult *set-9 set-26*/+ worms were put on plates for 1~2 hours to lay eggs and F1 homozygous *set-9 set-26* worm was picked onto new plates with RNAi bacteria. Brood size of 3-4 RNAi treated worms were scored.

**Gene ontology analysis**

Gene ontology (GO) analysis was carried out using the DAVID 6.8 Bioinformatics Database (http://david.abcc.ncifcrf.gov) (Huang et al. 2009).

# CHAPTER 3

# The role of SET-26 in nuclear RNAi


## Abstract

In *C. elegans*, SET-26 is a ubiquitously-expressed nucleus-localized putative histone methyltransferase. In recent years, our lab has shown an association of SET-26 with aging, as well as its role in regulating global H3K4me3 patterns. Moreover, we have observed an RNAi-resistant phenotype in *set-26* mutants. Because SET-26 localizes to the nucleus of cells, this latter finding implicates its involvement in the nuclear RNAi pathway. Further investigations show that SET-26 is specifically required for depositing H3K9me3 but not H3K27me3 marks at region targeted by RNSi trigger. Our findings reveal a novel role of SET-26 in RNAi-mediated gene silencing and more work needs to be done to uncover how SET-26 participants in the RNAi pathway.


## Introduction

RNA interference (RNAi) is a conserved cellular mechanism in which RNA inhibits transcription and/or translation of targeted genes. The canonical RNAi pathway resides in the cytoplasm, where double-stranded RNAs (dsRNAs) are processed into primary small interference RNAs (siRNAs) which paired with complementary mRNA strands by the RNA-Induced Silencing Complex (RISC) (Pratt and MacRae 2009). The double-stranded RNA formed by RISC are unable to be translated and are prompted for

degradation. Since the silencing happens after transcription, this is also called post-transcriptional gene silencing (PTGS) (Chekulaeva and Filipowicz 2009).

In addition to the canonical RNAi pathway, there exists an alternative, less-known pathway within the nucleus in *C. elegans* (Guang et al. 2008, 2010). This nuclear RNAi pathway is initiated when Nuclear RNAi-Defective 3 (NRDE-3) protein shuttles cytoplasmic siRNA into the nucleus. Once inside, siRNA-bound NRDE-3 bind nascently-transcribed, complementary RNA and form complexes with other NRDE proteins around those loci (Guang et al. 2008, 2010). The NRDE complex then stalls transcription at these loci and deposits repressive histone marks H3K9me3 and H3K27me3 (Guang et al. 2010; Mao et al. 2015). Loss of components of NRDE complex results in reduced levels of H3K9me3 and H3K27me3 (Guang et al. 2010; Mao et al. 2015). SET-25 and MET-2, two major H3K9me3 methyltransferases in *C. elegans*, are required for H3K9me3 methylation induction, but loss of both *set-25* and *met-2* did not exhibit a nuclear RNAi defective phenotype (Mao et al. 2015), whereas *mes-2* is required for both H3K27me3 deposition and nuclear RNAi defective phenotype (Mao et al. 2015), suggesting that H3K9me3 and H3K27me3 play distinct roles in nuclear RNAi. These results indicate that dsRNA-triggered chromatin modifications play an important role in nuclear RNAi.

In this work, we demonstrated that, despite their high sequence identity, SET-26, but not SET-9, plays a key role in nuclear RNAi. In addition, we showed that SET-26 is specifically required for H3K9me3 but not H3K27me3 deposition at dsRNA targeting sites. Together, we revealed a novel role of SET-26, a putative H3K9me3 methyltransferase, in nuclear RNAi pathway.

## Results

### *set-26* mutant exhibits RNAi resistance phenotype

Many nucleus localized proteins involved in histone modifications also have a potential role in RNAi. We previously found that the highly similar paralogs SET-9 & SET-26 localize exclusively in the nucleus, raising the question of whether they also play a role in RNAi. To this end, we tested RNAi response of *set-9*, *set-26* single and double mutants by feeding *E. coli* that express dsRNA targeting *lin-29* and *nhr-23*, knock down of which cause bursting and larva arrest phenotype respectively (Rougvie and Ambros 1995; Kostrouchova et al. 2001). Interestingly, although SET-9 and SET-26 proteins share 97% identity in protein sequence, *set-26* single and *set-9 set-26* double mutants exhibited a dosage dependent RNAi resistance phenotype for both *lin-29* and *nhr-23* RNAi treatment but *set-9* mutant had a similar response as wild type (Figure 1A and 1B). These results suggested that inactivation of *set-26*, but not *set-9*, causes a dosage dependent RNAi resistance phenotype in *C. elegans*. We previously found that SET-26 is broadly expressed but SET-9 is only detectable in the germline (Ni et al. 2012). The ubiquitous expression of SET-26, but not SET-9, may explain why SET-26 alone has a role in RNAi.

### *set-26* transgene rescues RNAi resistance phenotype in *set-26* single mutant

To further validate the role of *set-26* in RNAi response, we used previously constructed transgenic strains overexpressing *set-26* and assayed their response to RNAi. In order to be more effectively quantify the RNAi effect, we treated worms with *E. coli* expressed dsRNA targeting *dpy-13*, knock

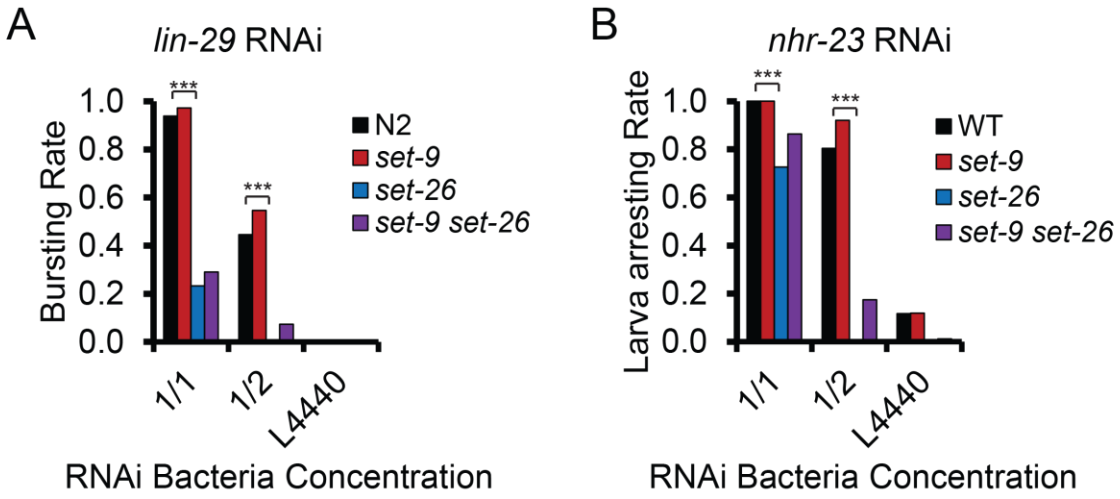

A
lin-29 RNAi
Bursting Rate
1.0
0.8
0.6
0.4
0.2
0.0
***
***
N2
set-9
set-26
set-9 set-26
1/1
1/2
L4440
RNAi Bacteria Concentration
B
nhr-23 RNAi
Larva arresting Rate
1.0
0.8
0.6
0.4
0.2
0.0
***
***
WT
set-9
set-26
set-9 set-26
1/1
1/2
L4440
RNAi Bacteria Concentration

Figure 1. *set-26* but not *set-9* is important for RNAi.

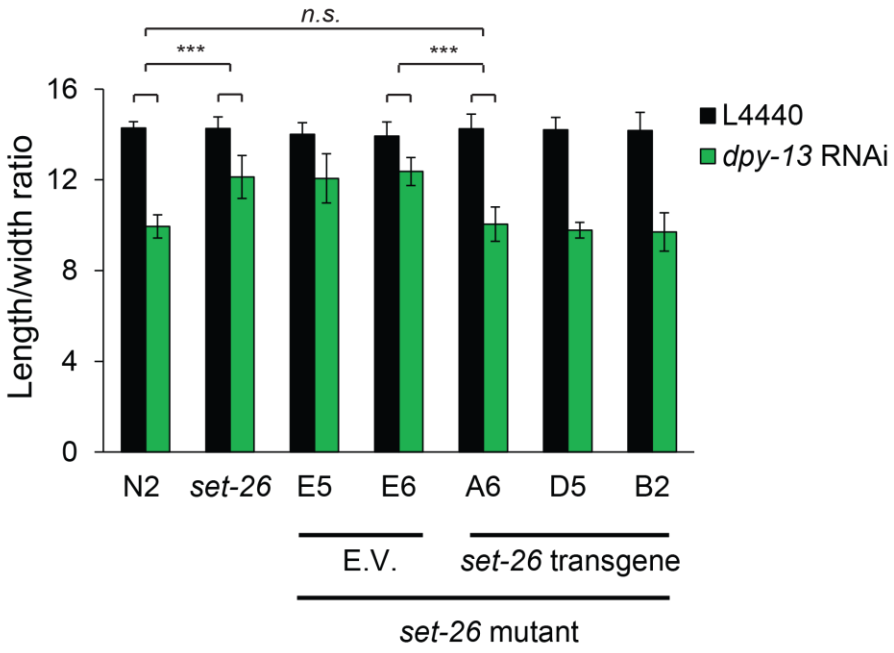
(A) Bursting rate for indicated strains treated with different *lin-29* RNAi concentration are shown. 1/2 represents *lin-29* RNAi bacteria diluted with L4440 RNAi bacteria at one to one ratio. (B) Larva arresting rate for indicated strains treated with different *nhr-23* RNAi concentration are shown. 1/2 represents *nhr-23* RNAi bacteria diluted with L4440 RNAi bacteria at one to one ratio. (***$p<0.001$)

down of which results in a short and chunky body shape (von Mende et al. 1988), and measured the length over width ratio. Knocking down *dpy-13* reduced the length over width ratio in wild type (Figure 2). Consistent with a dosage dependent RNAi resistant phenotype in *set-26* mutant, although *dpy-13* RNAi treatment also reduced the length over width ratio in *set-26* mutant, the degree of reduction is significantly lower than that in wild type (Figure 2). In addition, this RNAi resistance phenotype can be rescued by expression of wild-type *set-26* transgenes in the *set-26* mutant (Figure 2). Taken together, our results strongly suggested that *set-26* plays a role in RNAi in *C. elegans*.

**_set-26_ is important for nuclear RNAi**

Considering SET-26 localizes exclusively in the nuclear, we wonder whether *set-26* plays a role in nuclear RNAi. The *lin-15b* and *lin-15a* are transcribed as one pre-mRNA that is spliced into *lin-15b* and *lin-15a* mRNAs in the nucleus (Huang et al. 1994; Clark et al. 1994). Worms lacking both *lin-15b* and *lin-15a* genes, but not either gene alone, exhibit a multivulva (Muv) phenotype (Huang et al. 1994; Clark et al. 1994). Wild type worms treated with lin-15b RNAi does not have any obvious phenotype. However, *eri-1* mutant with enhanced sensitivity to dsRNAs (Kennedy et al. 2004) treated with *lin-15b* RNAi will deplete both *lin-15b* and *lin-15a* gene activity and results in a multivulva (Muv) phenotype (Guang et al. 2008). Therefore *lin-15b* RNAi treatment in *eri-1* mutant could be a tool to test whether silencing occurs at the transcriptional level (Guang et al. 2008). To this end, we assayed the Muv rate in *eri-1*, *set-9 eri-1* and *set-26 eri-1* mutants. NRDE-3 is known as a factor that shuttles siRNA into nucleus and is required for nuclear RNAi and was included as a control. We found that *set-26 eri-1* mutant has a

reduced Muv rate compared to *eri-1* mutant but *set-9 eri-1* mutant has a similar response as *eri-1* mutant (Figure 3). The *set-9 set-26 eri-1* triple mutant arrested at L1 stage due to unknown reason and


n.s.
***
***
16
12
8
4
0
Length/width ratio
L4440
dpy-13 RNAi
N2
set-26
E5
E6
A6
D5
B2
E.V.
set-26 transgene
set-26 mutant

*

Figure 2. *set-26* transgene rescues *dpy-13* RNAi resistant phenotype of *set-26* mutant

Length over width ratio for indicated strains treated with *dpy-13* RNAi and L4440 are shown. A6, D5 and B2 are *set-26* mutants with *set-26* transgene. E5 and E6 are *set-26* mutants with empty vectors. *set-26* mutants are less sensitive to *dpy-13* RNAi. And *set-26* transgenes restore the sensitivity to *dpy-13* RNAi. (***p<0.001, n.s. no significant)

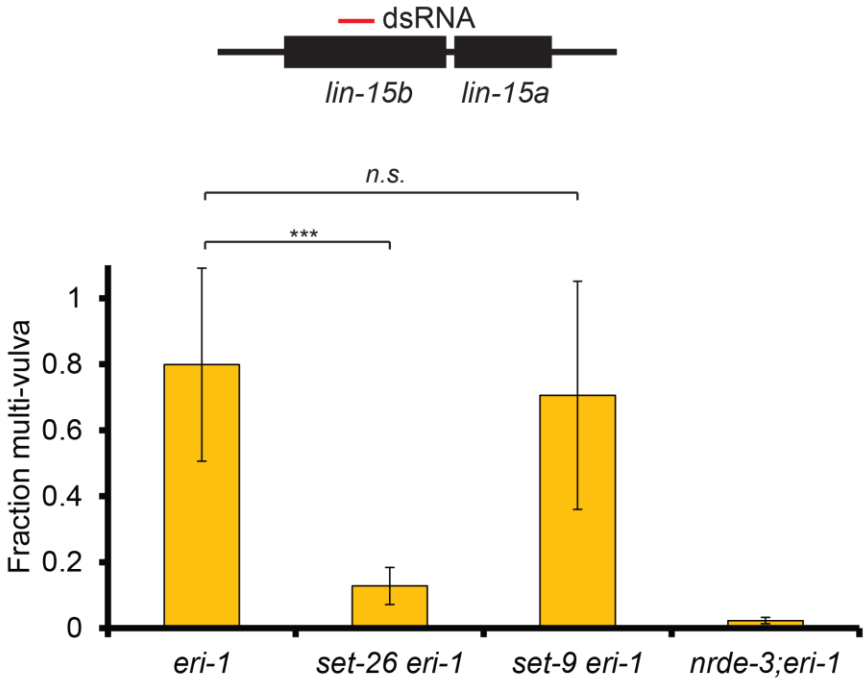

dsRNA
lin-15b
lin-15a
n.s.
***
Fraction multi-vulva
1
0.8
0.6
0.4
0.2
0
eri-1
set-26 eri-1
set-9 eri-1
nrde-3;eri-1

Figure 3. *set-26* is important for nuclear RNAi pathway

Multi-vulva phenotype requires knock down of both *lin-15b* and *lin-15a* gene in *eri-1* mutant. Fraction of multi-vulva phenotype are shown for indicated strains treated with *lin-15b* RNAi. (***p<0.001, n.s. not significant)

was not included in this assay. This result suggested that *set-26,* but not *set-9,* is required for nuclear RNAi in *C. elegans*.

### ***set-26*** **is only required for depositing H3K9me3, but not H3K27me3, at dsRNA targeting sites**

H3K9me3 and H3K27me3 are known to mark dsRNA targeting region in the genome to silence transcription (Guang et al. 2010; Gu et al. 2012; Mao et al. 2015). Defects in nuclear RNAi result in reduced levels of both H3K9me3 and H3K27me3 at dsRNA targeting sites, which likely contribute to reduced silencing and RNAi resistance phenotype to *lin-15b* RNAi treatment (Guang et al. 2010; Gu et al. 2012; Mao et al. 2015). Since the SET domain of SET-26 has the potential to methylate H3K9me3 *in vitro* (Greer et al. 2014), although its effect could be indirect, we next wondered whether SET-26 acts as an H3K9me2/3 methyltransferase at dsRNA targeting region in response to RNAi treatment. To this end, we assayed H3K9me3 and H3K27me3 levels around dsRNA targeting region. Consistent with their phenotypic responses, exogenous *lin-15b* RNAi treatment triggers H3K9me3 and H3K27me3 deposition around dsRNA targeting region in *eri-1* mutant but not in *eri-1; nrde-3* double mutant, which is a nuclear RNAi defective mutant (Figure 4 and Figure 5). Interestingly, although the elevated H3K27me3 levels at *lin-15b* targeting sites remained unchanged in *set-26 eri-1* mutant with *lin-15b* RNAi treatment (Figure 5), the H3K9me3 levels weresignificantly reduced (Figure 4). These results suggest that SET-26 is specifically required for H3K9me3 but not H3K27me3 deposition in response to dsRNA treatment.

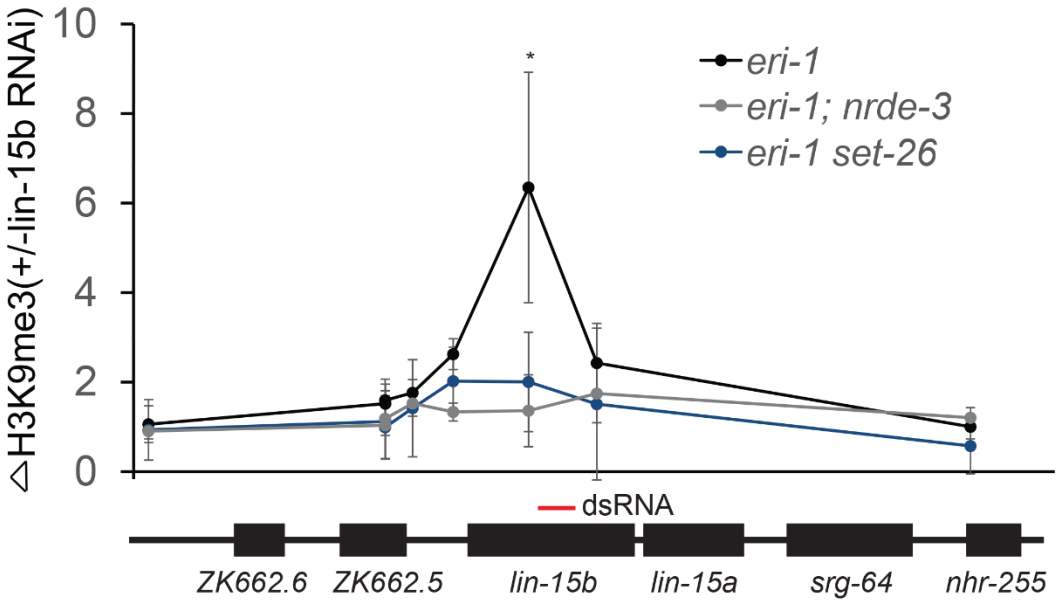

△H3K9me3(+/-lin-15b RNAi)
10
8
6
4
2
0
*
eri-1
eri-1; nrde-3
eri-1 set-26
dsRNA
ZK662.6
ZK662.5
lin-15b
lin-15a
srg-64
nhr-255

Figure 4. *set-26* is required for siRNA-directed H3K9me3 deposition

Chromatin Immunoprecipitation (ChIP) with anti-H3K9me3 was performed on adult worms exposed to *lin-15b* RNAi or L4440 bacteria. Co-precipitating H3K9me3 DNA was quantified with qRT-PCR and data are expressed as ratios of samples exposed to *lin-15b* RNAi or L4440 bacteria. (*$p<0.05$, n=3)

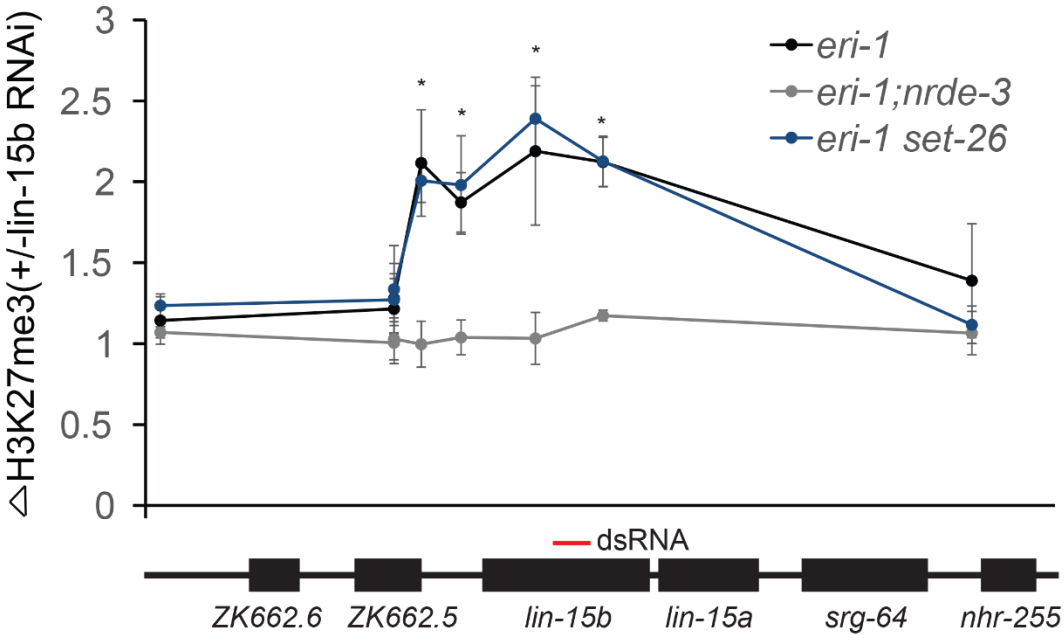

△H3K27me3(+/-lin-15b RNAi)
3
2.5
2
1.5
1
0.5
0
eri-1
eri-1;nrde-3
eri-1 set-26
dsRNA
ZK662.6
ZK662.5
lin-15b
lin-15a
srg-64
nhr-255

Figure 5. *set-26* is dispensable for siRNA-directed H3K27me3 deposition

Chromatin Immunoprecipitation (ChIP) with anti-H3K27me3 was performed on adult worms exposed to *lin-15b* RNAi or L4440 bacteria. Co-precipitating H3K27me3 DNA was quantified with qRT-PCR and data are expressed as ratios of samples exposed to *lin-15b* RNAi or L4440 bacteria. (*$p<0.05$, $n=3$)

## Discussion

In this study, we revealed a novel role of *set-26* in RNAi pathway. We found that SET-26 but not SET-9 plays a role in RNAi. We demonstrated that loss of *set-26* causes reduced H3K9me3 but not H3K27me3 levels in dsRNA induced heterochromatic regions. Further work needs to be done to show that whether SET-26 regulates RNAi directly.

H3K9me3 and H3K27me3 have been linked to heterochromatin formation in response to dsRNA treatment (Holoch and Moazed 2015). In *C. elegans*, dsRNA triggers nuclear RNAi pathway dependent small-RNA-directed H3K27me3 at dsRNA targeting sites (Mao et al. 2015). The H3K27me3 marks can be inherited to progeny for multiple generations (Mao et al. 2015). MES-2, the major H3K27me3 methyltransferase, is required for the H3K27me3 induction and the loss of MES-2 leads to a nuclear RNAi defective phenotype (Mao et al. 2015), suggesting that H3K27me3 plays an important role in nuclear RNAi.

However, the importance of H3K9me3 in nuclear RNAi in *C. elegans* is debatable as loss of major H3K9me3 methyltransferases abolish H3K9me3 induction but does not result in a nuclear RNAi defective phenotype (Kalinava et al. 2017; Mao et al. 2015). Loss of *met-2*, *set-25* and *set-32* causes reduced H3K9me3 at dsRNA targeting sites in response to RNAi treatment (Mao et al. 2015). The H3K9me3 induction and the requirement of these H3K9me3 methyltransferases are further confirmed by whole genome analysis of H3K9me3 in RNAi treated worms (Kalinava et al. 2017; Gu et al. 2012). However, these H3K9me3 methyltransferases are not required for dsRNA-

induced transcriptional silencing and heritable RNAi (Kalinava et al. 2017; Mao et al. 2015). *met-2*, *set-25* and *set-32* single, double and triple mutants, although show no H3K9me3 induction in response to RNAi treatment, fail to exhibit pronounced nuclear RNAi defects at both phenotypic and transcriptional level (Kalinava et al. 2017; Mao et al. 2015).

In this study, we revealed a novel role of SET-26 in nuclear RNAi. The loss of *set-26* results in a nuclear RNAi defective phenotype (Figure 3). In addition, we found that SET-26 is required for H3K9me3 but not H3K27me3 deposition at dsRNA targeting sites. *In vitro* study shows that the SET domain of SET-26 has the potential to methylate H3K9me3 (Greer et al. 2014). In contrast to other H3K9me3 enzymes that are only required for H3K9me3 deposition at targeting sites but not transcriptional silencing, SET-26 is required for both H3K9me3 deposition and transcriptional silencing at the phenotypic level (Figure 3 and Figure 5). One possibility is although MET-2, SET-25 and SET-32 are the major methyltransferases that trigger H3K9me3, SET-26 mediates both the H3K9me3 induction and transcriptional silencing.

# Experimental Procedures

## Worm strains

The N2 strain was used as the wild-type (WT). The mutants used in the study are: *set-9(rw5)*, *set-26(tm2467), eri-1(mg366), nrde-3(gg66)*.

## Antibodies

The antibodies used were anti-H3K9me3(abcam, Cambridge, United Kingdom ab8898), anti-GFP(abcam, Cambridge, United Kingdom ab290), anti-H3(abcam, Cambridge, United Kingdom ab1791), anti-H3K27me3 (Millipore, United States 07-449).

## RNAi

For *lin-29, nhr-23* and *dpy-13* RNAi, RNA constructs were obtained from the Ahringher library. HT115 bacteria containing RNAi constructs were grown at 37°C and seeded on nematode growth medium (NGM) plates containing carbenicillin and tetracycline and dry overnight. dsRNA expression was induced by adding IPTG to a final concentration 0.4 mM. *lin-15b* RNAi experiments were performed as described (Guang et al. 2008). *lin-15b* RNAi construct was a gift from Scott Kennedy (Harvard University). HT115 bacteria containing RNAi constructs were grown at 37°C for 16hours and 1M IPTG was added to the culture to a final concentration 4mM at room temperature for 2 hours. The culture was then used for seeding RNAi plates and used in the same day. The *eri-1* mutant, which is hypersensitive to RNAi, was used in this experiment to maximize our ability to score the phenotype. We scored the multivulva phenotype in adult stage at the second generation (Guang et al. 2008).

## ChIP

ChIP experiments for H3K4me3 and H3K27me3 were performed as described (Pu et al. 2015). Adult worm pellets were ground with a mortar and pestle and cross-linked with 1% formaldehyde in PBS for 10 min at room temperature. Worm fragments were collected by spinning at 3000g for 5 min and resuspended in FA buffer followed by sonication with Bioruptor. Chromatin extract was incubated with antibody overnight at 4°C. Precipitated DNA from each sample was used for quantitative PCR. All experiments were repeated three times and t-test was used to calculate statistical significance.

# CHAPTER 4

# CONCLUSIONS AND FUTURE DIRECTION

Alterations in epigenetic state have been observed during aging, germline development, and RNAi. *Caenorhabditis elegans* is a widely used animal model for aging, germline and RNAi research. Several epigenetic regulators modulating longevity, germline function and RNAi pathway have been previously identified in *C. elegans*. However, studies on the influence of genome-wide alteration of histone marks in longevity, germline function and RNAi pathway were limited. My work has focused on studying the functions of two highly homologous putative histone modifiers SET-9 and SET-26 on longevity, germline function and RNAi pathway.

My work in Chapter 2 described and characterized the distinct functions of SET-9 and SET-26 on longevity and germline function. It suggested that SET-9 & SET-26 bind to H3K4me3 and restrict the spreading of H3K4me3, especially in the germline. In addition, I showed that SET-9 & SET-26 are not the major enzymes for H3K9me3 under normal condition. Interestingly, in Chapter 3, I showed that *set-26* mutant exhibits a RNAi resistance phenotype and has a much lower H3K9me3 level at siRNA targeting site upon RNAi treatment. This suggested that in response to RNAi trigger, SET-26 is required for H3K9me3 deposition around siRNA targeting site. There are various possibilities that account for its different role in H3K9me3. First, considering the *in vitro* H3K9me3 methyltransferase activity, it is possible that SET-26 itself is an H3K9me3

methyltransferase that specifically for dsRNA induced H3K9me3. Second, as a histone binding protein, it is also possible that although SET-26 is not a major H3K9me3 enzyme, SET-26 mediated dsRNA trigger-dependent recruitment of H3K9me3 methyltransferase to its targeting sites in an indirect manner, such as acting as a scaffolding protein at the dsRNA targeting sites. In the future, it will be interesting to figure this out. One possible way is to perform Immunoprecipitation–Mass Spectrometry (IP-MS) assay, in which I can label SET-26 protein with GFP and immunoprecipitate SET-26 protein with GFP antibody with or without dsRNA treatment. The proteins co-precipitate with SET-26 after dsRNA treatment will be candidates for H3K9me3 deposition upon dsRNA treatment. The results from IP-MS assay might also support the model that SET-26 acts as a mediator protein in RNAi pathway. An alternative way is to perform a genetic screen for candidate genes that required for RNAi pathway. I can use the worm mutants that have mutations in candidate genes and treat those mutants with RNAi to see any of them has a RNAi resistance phenotype. In addition, my work in Chapter 2 demonstrated that SET-26 plays an important role in lifespan regulation and I showed in Chapter 3 that SET-26 is required for RNAi pathway. It would be interesting to investigate in the future that whether the lifespan phenotype in *set-26* mutant connects to its RNAi resistance phenotype. Furthermore, in Chapter 2 I found that the *set-9 set-26* double mutant has a mortal germline phenotype and in Chapter 3 I observed that the *set-26* mutant is resistant to RNAi treatment. These findings raise a question of whether the role that SET-26 plays in the RNAi pathway relates to its function in germline. It would be interesting to study the siRNA profile changes in the *set-9 set-26* double mutant. In addition, it would be also interesting to investigate

whether the *set-9 set-26* double mutant is resistant to RNAi in a transgenerational inheritance manner.

# APPENDIX 1

# INVESTIGATING THE ROLE OF HISTONE ACETYLATION IN LONGEVITY

Histone acetylation has been found to be very important for chromatin structure. Under gene repression stage, DNA tightly binds to histone and form this condensed structure. But, there is a group of enzymes that called histone acetyltransferase (HAT) that transfer acetyl group to the lysine residue on the histone, which open the chromatin and activate the gene transcription. The reverse reaction is catalyzed by another group of enzymes called histone deacetylase (HDAC) that remove the acetyl group from the histone and close the chromatin structure (Grunstein 1997).

Histone acetylation is important for lifespan. Histone H4 lysine16 acetylation (H4K16ac) regulates cellular lifespan in yeast (Dang et al. 2009). H4K16ac increases with aging (Dang et al. 2009). Loss of Sas2, a histone acetyltransferase, or overexpression of Sir2, a histone deacetylase, decrease the level of H4K16ac and extends cellular lifespan in yeast (Dang et al. 2009). Histone deacetylase inhibitor treatment that suppresses HDAC function extends lifespan in *C. elegans* (Zhang et al. 2009a), suggesting that the influence of histone acetylation on longevity is different between yeast and *C. elegans*. In order to investigate the influence of histone acetylation on lifespan in germline-less *C. elegans*, we first asked whether global histone acetylation mark will change during aging. We found that while histone

acetylation on histone H4 increase during aging, H3K14ac decreases during aging. Interestingly, loss of SET-26 function rescued the decreased H3K14ac level in old worm suggesting that SET-26 might be important for age-dependent change in histone acetylation. In addition, to examine the influence of histone deacetylase in lifespan regulation, we found that the lifespan was extended in germline-less worms treated with HDAC inhibitor. However, the long-lived lifespan with HDAC inhibitor treatment is independent with SET-26, suggesting that they act in parallel to modulate lifespan. We also screened for histone deacetylase that affect lifespan in *C. elegans* and found that *hda-2* when knocked down extends lifespan. Collectively, our data suggests that histone acetylation plays an important role in modulating lifespan in *C. elegans*.

**Aging affects histone acetylation marks differentially**

To investigate how aging affects histone acetylation marks in somatic cells, I assayed global histone acetylation mark level using young (day 2) and old (day 12) *glp-1* mutant, which lacks germline. I first assayed histone acetylation levels on H4K5, H4K8, H4K12 and H4K16. We observed a global increase in H4K5ac, H4K8ac and H4K16ac levels in old worms compared with young worms (Figure AI.1) (Table AI.1). This increase in histone acetylation levels was also observed in long-lived *glp-1; age-1* and *glp-1; set-26* double mutant, suggesting that *age-1* and *set-26* act independently with global H4K5ac, H4K8ac and H4K16ac level to affect lifespan. I

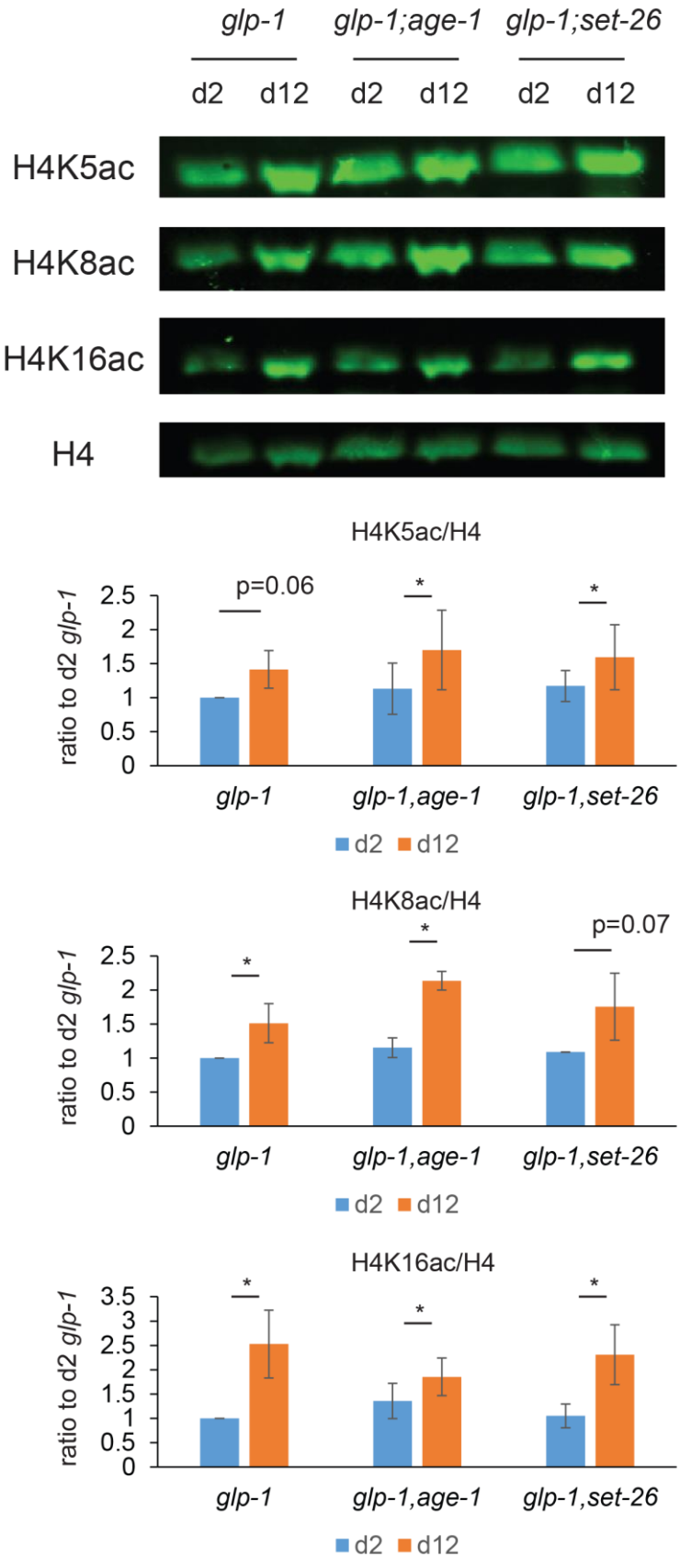

glp-1
glp-1;age-1
glp-1;set-26
d2
d12
d2
d12
d2
d12
H4K5ac
H4K8ac
H4K16ac
H4
H4K5ac/H4
p=0.06
*
*
ratio to d2 glp-1
2.5
2
1.5
1
0.5
0
glp-1
glp-1,age-1
glp-1,set-26
d2
d12
H4K8ac/H4
*
*
p=0.07
ratio to d2 glp-1
2.5
2
1.5
1
0.5
0
glp-1
glp-1,age-1
glp-1,set-26
d2
d12
H4K16ac/H4
*
*
*
ratio to d2 glp-1
3.5
3
2.5
2
1.5
1
0.5
0
glp-1
glp-1,age-1
glp-1,set-26
d2
d12

Figure AI.1 Increased H4K5ac, H4K8ac and H4K16ac levels in old worms.

H4K5ac, H4K8ac and H4K16ac levels of *glp-1(e2141)*, *glp-1(e2141); age-1(hx546)* and *glp-1(e2141);set-26(tm2467)* at day2 and day12 stage are shown (*p<0.05, n=3)

| | Change in D12 *glp-1* |
|---|---|
| H4K5ac | Increase |
| H4K8ac | Increase |
| H4K12ac | No change |
| H4K16ac | Increase |
| H3K9ac | No change |
| H3K14ac | Decrease |
| H3K18ac | No change |
| H3K23ac | No change |
| H3K27ac | No change |

Table AI.1 Histone acetylation marks examined in young and old worms.

Global levels of histone acetylation marks were examined in day2 (young) and day12 (old) glp-1(e2141) worms.

next assayed histone acetylation levels on H3K9, H3K14, H3K18, H3K23 and H3K27. No significant difference was observed for H3K9ac, H3K18ac, H3K23ac and H3K27ac (Table AI.1), but we observed a decreased level of H3K14ac in old worms compared with young worms (Figure AI.2) (Table AI.1). Surprisingly, although this decrease in H3K14ac level was also observed in long-lived *glp-1; age-1*, loss of *set-26* partially rescued the decreased H3K14ac level in old worms (Figure AI.2) suggesting that *set-26* is required for reducing H3K14ac level in old worm. These results suggested that aging affects specific histone acetylation marks.

**HDAC inhibitor sodium butyrate (NaB) increase lifespan in *glp-1* mutant**

Next, I wondered whether histone acetylation plays a role in longevity. To this end, I examined the effect of HDAC inhibitor NaB on lifespan. NaB treatment reduces HDAC function and causes increased histone acetylation (Zhang et al. 2009a). I observed extended lifespan in *glp-1* worms treated with NaB (Figure AI.3). Since the above results showed that *set-26* is required for reducing H3K14ac level in old worms, I wondered whether *set-26* mutation and NaB treatment act in a same pathway to affect lifespan. If both NaB treatment and loss of *set-26* modulate lifespan in the same manner, I would expect that NaB treatment would not extend lifespan in the *glp-1; set-26* double mutant. However, I observed that NaB treatment further extends lifespan in the *glp-1; set-26* double mutant suggesting that NaB treatment and SET-26 act in parallel to modulate lifespan.

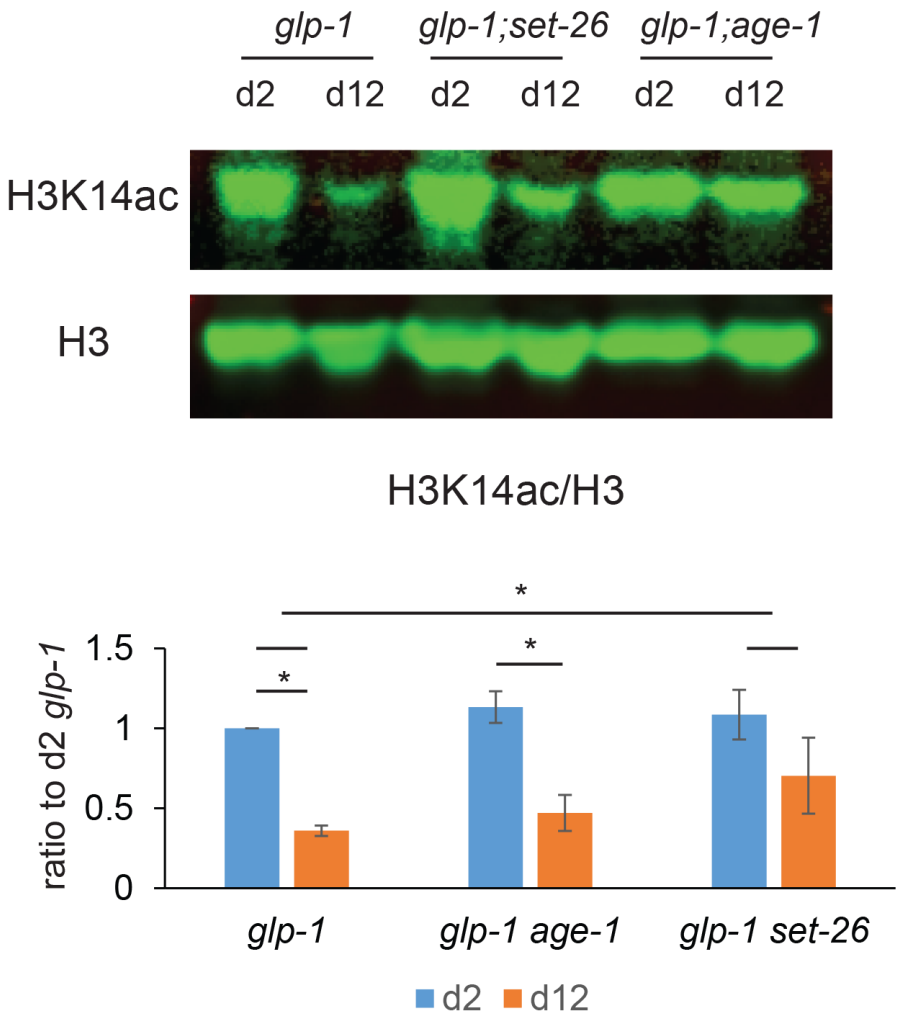

glp-1
glp-1;set-26
glp-1;age-1
d2
d12
d2
d12
d2
d12
H3K14ac
H3
H3K14ac/H3
ratio to d2 glp-1
1.5
1
0.5
0
*
*
*
glp-1
glp-1 age-1
glp-1 set-26
d2
d12

Figure AI.2 Loss of *set-26* partially rescues the decreased H3K14ac level in old worms.

H3K14ac levels of *glp-1(e2141)*, *glp-1(e2141); age-1(hx546)* and *glp-1(e2141);set-26(tm2467)* at day2 and day12 stage are shown (*p<0.05, n=3)

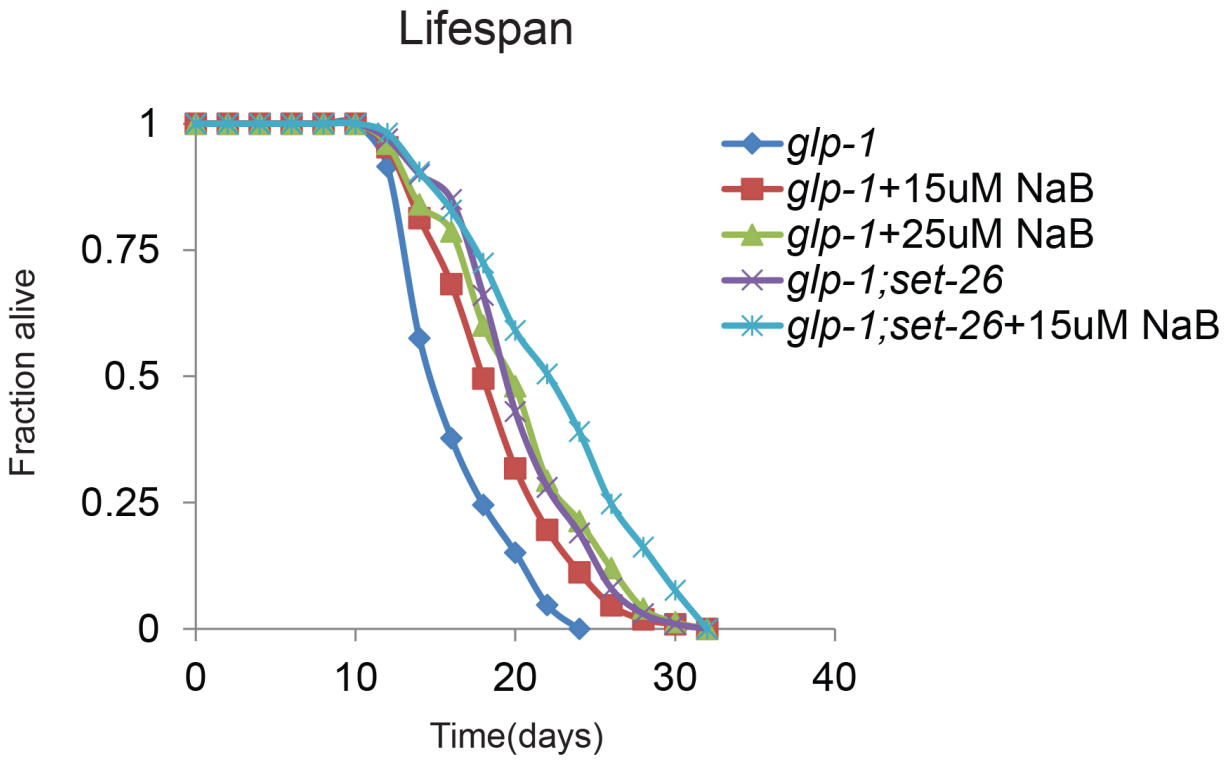

Lifespan
Fraction alive
1
0.75
0.5
0.25
0
0
10
20
30
40
Time(days)
glp-1
glp-1+15uM NaB
glp-1+25uM NaB
glp-1;set-26
glp-1;set-26+15uM NaB

Figure AI.3 NaB extends lifespan in *glp-1* and *glp-1;set-26* mutants

Survival curves for *glp-1(e2141)* and *glp-1(e2141);set-26(tm2467)* strains treated with indicated concentration of NaB are shown.

**Screen for histone acetylation modifiers that play a role in longevity**

The results above suggested that there might be histone acetyltransferase or/and histone deacetylase that are involved in longevity modulation, I then asked which specific histone acetylation modifier is important for lifespan. To this end, I performed an RNAi screen for histone acetyltransferase and histone deacetylase that when knocked down extend lifespan in *C. elegans*. From the RNAi screen of HDACs in *C. elegans*, I found that although most of the HDACs when knocked down do not affect or even shorten lifespan, knocking down *hda-2* extended lifespan in *C. elegans* (Figure AI.4). It will be interesting to further investigate whether HDA-2 is responsible for the lifespan extension phenotype of worm treated with HDAC inhibitor. In addition, I also screened other histone acetylation modifiers but none of them showed a lifespan extension phenotype (Figure AI.5). However, it is still possible that these factors could be important for lifespan regulation under other conditions.

In conclusion, in this study I found that global levels of H4K5ac, H4K8ac and H4K16ac increase whereas H3K14ac decrease during aging. Loss of SET-26 rescues the reduced H3K14ac but not other increased histone acetylation marks in old worms. In addition, I found that NaB treatment extends lifespan in a manner that is independent of SET-26, suggesting that NaB treatment acts in parallel with SET-26 to modulate lifespan. Furthermore, I revealed that *hda-2* when knocked down extends lifespan. My results suggested that histone acetylation is important for longevity. However, the specific histone acetylation modifiers that responsible for the long-lived phenotype with

worms treated with NaB are unknown. This could be an interesting future direction to study.

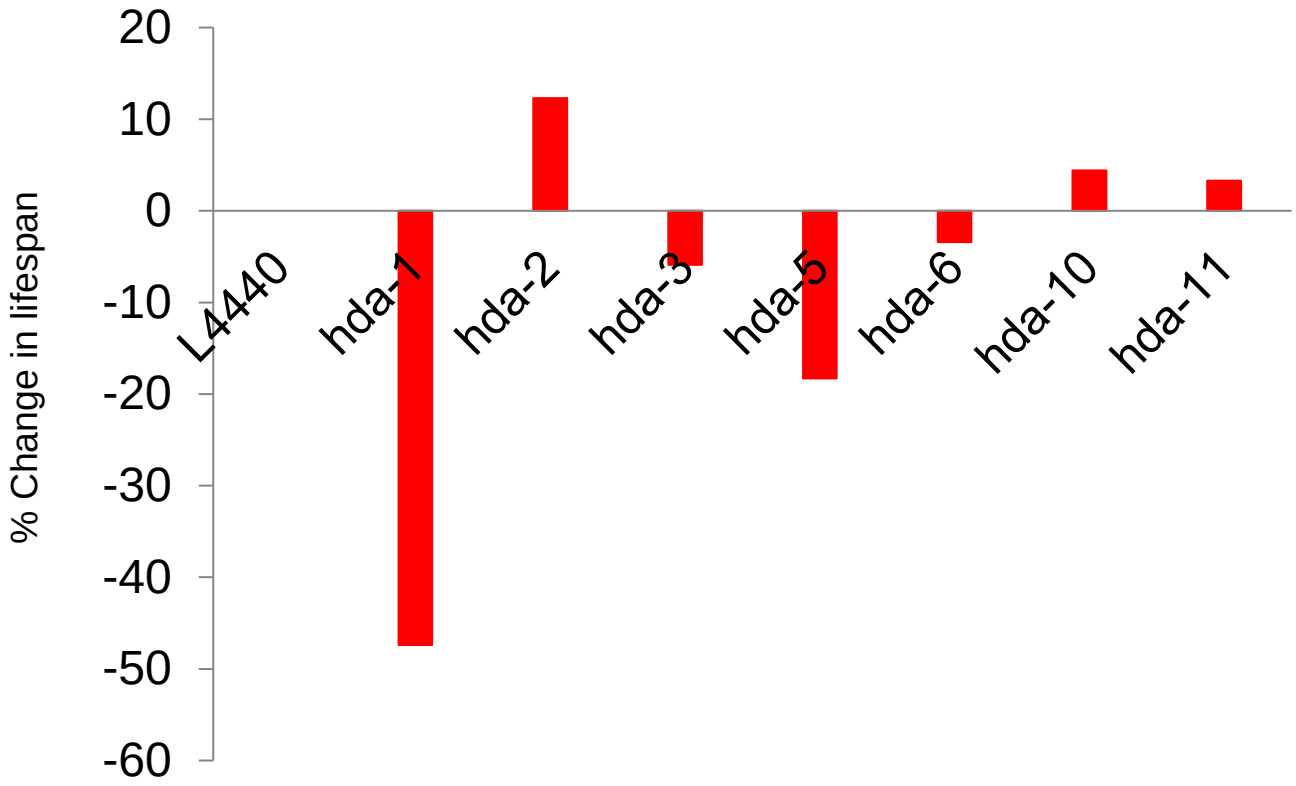

% Change in lifespan
20
10
0
-10
-20
-30
-40
-50
-60
L4440
hda-1
hda-2
hda-3
hda-5
hda-6
hda-10
hda-11

Figure AI.4 RNAi screen revealed *hda-2* when knocked down extends lifespan in *glp-1* mutant.

Percentage of lifespan changes in *glp-1* worms treated with indicated RNAi are shown.

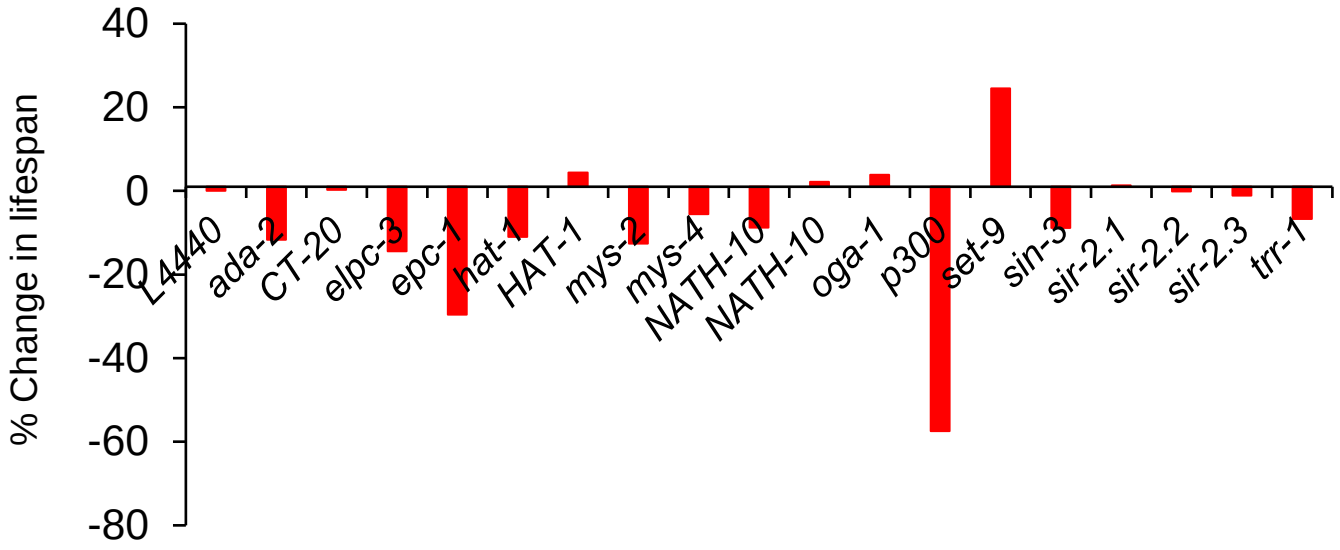

% Change in lifespan
40
20
0
-20
-40
-60
-80
L4440
ada-2
CT-20
elpc-3
epc-1
hat-1
HAT-1
mys-2
mys-4
NATH-10
NATH-10
oga-1
p300
set-9
sin-3
sir-2.1
sir-2.2
sir-2.3
trr-1

Figure AI.5 RNAi screen of histone acetyltransferase and deacetylase in lifespan.

Percentage of lifespan changes in *glp-1* worms treated with indicated RNAi are shown. *set-9* RNAi is used as a positive control.

# Reference

Ahmed S, Hodgkin J. 2000. MRT-2 checkpoint protein is required for germline immortality and telomere replication in C. elegans. *Nature* **403**: 159–164.

Ali M, Rincon-Arano H, Zhao W, Rothbart SB, Tong Q, Parkhurst SM, Strahl BD, Deng L-W, Groudine M, Kutateladze TG. 2013. Molecular basis for chromatin binding and regulation of MLL5. *Proc Natl Acad Sci U S A* **110**: 11296–11301.

Alvares SM, Mayberry GA, Joyner EY, Lakowski B, Ahmed S. 2014. H3K4 demethylase activities repress proliferative and postmitotic aging. *Aging Cell* **13**: 245–253.

Arantes-Oliveira N, Apfeld J, Dillin A, Kenyon C. 2002. Regulation of life-span by germ-line stem cells in Caenorhabditis elegans. *Science* **295**: 502–505.

Austin J, Kimble J. 1987. glp-1 is required in the germ line for regulation of the decision between mitosis and meiosis in C. elegans. *Cell* **51**: 589–599.

Bannister AJ, Kouzarides T. 2011. Regulation of chromatin by histone modifications. *Cell Res* **21**: 381–395.

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.

Bannister AJ, Zegerman P, Partridge JF, Miska EA, Thomas JO, Allshire RC, Kouzarides T. 2001. Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* **410**: 120–124.

Bardet AF, He Q, Zeitlinger J, Stark A. 2011. A computational pipeline for comparative

ChIP-seq analyses. *Nat Protoc* **7**: 45–61.

Bargaje R, Alam MP, Patowary A, Sarkar M, Ali T, Gupta S, Garg M, Singh M, Purkanti R, Scaria V, et al. 2012. Proximity of H2A.Z containing nucleosome to the transcription start site influences gene expression levels in the mammalian liver and brain. *Nucleic Acids Res* **40**: 8965–8978.

Bender LB, Cao R, Zhang Y, Strome S. 2004. The MES-2/MES-3/MES-6 complex and regulation of histone H3 methylation in C. elegans. *Curr Biol* **14**: 1639–1643.

Bender LB, Suh J, Carroll CR, Fong Y, Fingerman IM, Briggs SD, Cao R, Zhang Y, Reinke V, Strome S. 2006. MES-4: an autosome-associated histone methyltransferase that participates in silencing the X chromosomes in the C. elegans germ line. *Development* **133**: 3907–3917.

Berman JR, Kenyon C. 2006. Germ-cell loss extends C. elegans life span through regulation of DAF-16 by kri-1 and lipophilic-hormone signaling. *Cell* **124**: 1055–1068.

Berry LW, Westlund B, Schedl T. 1997. Germ-line tumor formation caused by activation of glp-1, a Caenorhabditis elegans member of the Notch family of receptors. *Development* **124**: 925–936.

Bessler JB, Andersen EC, Villeneuve AM. 2010. Differential localization and independent acquisition of the H3K9me2 and H3K9me3 chromatin modifications in the Caenorhabditis elegans adult germ line. *PLoS Genet* **6**: e1000830.

Bian C, Xu C, Ruan J, Lee KK, Burke TL, Tempel W, Barsyte D, Li J, Wu M, Zhou BO,

et al. 2011. Sgf29 binds histone H3K4me2/3 and is required for SAGA complex recruitment and histone H3 acetylation. *EMBO J* **30**: 2829–2842.

Bluher M, Kahn BB, Kahn CR. 2003. Extended longevity in mice lacking the insulin receptor in adipose tissue. *Science* **299**: 572–574.

Bratic A, Larsson N-G. 2013. The role of mitochondria in aging. *J Clin Invest* **123**: 951–957.

Buckley BA, Burkhart KB, Gu SG, Spracklin G, Kershner A, Fritz H, Kimble J, Fire A, Kennedy S. 2012. A nuclear Argonaute promotes multigenerational epigenetic inheritance and germline immortality. *Nature* **489**: 447–451.

Buhler M, Verdel A, Moazed D. 2006. Tethering RITS to a nascent transcript initiates RNAi- and heterochromatin-dependent gene silencing. *Cell* **125**: 873–886.

Cattie DJ, Richardson CE, Reddy KC, Ness-Cohn EM, Droste R, Thompson MK, Gilbert W V, Kim DH. 2016. Mutations in Nonessential eIF3k and eIF3l Genes Confer Lifespan Extension and Enhanced Resistance to ER Stress in Caenorhabditis elegans. *PLoS Genet* **12**: e1006326.

Chekulaeva M, Filipowicz W. 2009. Mechanisms of miRNA-mediated post-transcriptional regulation in animal cells. *Curr Opin Cell Biol* **21**: 452–460.

Clark SG, Lu X, Horvitz HR. 1994. The Caenorhabditis elegans locus lin-15, a negative regulator of a tyrosine kinase signaling pathway, encodes two different proteins. *Genetics* **137**: 987–997.

Copeland JM, Cho J, Lo TJ, Hur JH, Bahadorani S, Arabyan T, Rabie J, Soh J, Walker

DW. 2009. Extension of Drosophila life span by RNAi of the mitochondrial respiratory chain. *Curr Biol* **19**: 1591–1598.

Crittenden SL, Leonhard KA, Byrd DT, Kimble J. 2006. Cellular analyses of the mitotic region in the Caenorhabditis elegans adult germ line. *Mol Biol Cell* **17**: 3051–3061.

Dang W, Steffen KK, Perry R, Dorsey JA, Johnson FB, Shilatifard A, Kaeberlein M, Kennedy BK, Berger SL. 2009. Histone H4 lysine 16 acetylation regulates cellular lifespan. *Nature* **459**: 802–807.

Deng L-W, Chiu I, Strominger JL. 2004. MLL 5 protein forms intranuclear foci, and overexpression inhibits cell cycle progression. *Proc Natl Acad Sci U S A* **101**: 757–762.

Dernburg AF, McDonald K, Moulder G, Barstead R, Dresser M, Villeneuve AM. 1998. Meiotic recombination in C. elegans initiates by a conserved mechanism and is dispensable for homologous chromosome synapsis. *Cell* **94**: 387–398.

Dillon SC, Zhang X, Trievel RC, Cheng X. 2005. The SET-domain protein superfamily: protein lysine methyltransferases. *Genome Biol* **6**: 227.

Edwards C, Canfield J, Copes N, Rehan M, Lipps D, Bradshaw PC. 2014. D-beta-hydroxybutyrate extends lifespan in C. elegans. *Aging (Albany NY)* **6**: 621–644.

Emerling BM, Bonifas J, Kratz CP, Donovan S, Taylor BR, Green ED, Le Beau MM, Shannon KM. 2002. MLL5, a homolog of Drosophila trithorax located within a segment of chromosome band 7q22 implicated in myeloid leukemia. *Oncogene* **21**: 4849–4854.

Evans DS, Kapahi P, Hsueh W-C, Kockel L. 2011. TOR signaling never gets old: aging, longevity and TORC1 activity. *Ageing Res Rev* **10**: 225–237.

Feng J, Bussiere F, Hekimi S. 2001. Mitochondrial electron transport is a key determinant of life span in Caenorhabditis elegans. *Dev Cell* **1**: 633–644.

Feser J, Truong D, Das C, Carson JJ, Kieft J, Harkness T, Tyler JK. 2010. Elevated histone expression promotes life span extension. *Mol Cell* **39**: 724–735.

Flatt T, Min K-J, D'Alterio C, Villa-Cuesta E, Cumbers J, Lehmann R, Jones DL, Tatar M. 2008. Drosophila germ-line modulation of insulin signaling and lifespan. *Proc Natl Acad Sci U S A* **105**: 6368–6373.

Fox PM, Vought VE, Hanazawa M, Lee M-H, Maine EM, Schedl T. 2011. Cyclin E and CDK-2 regulate proliferative cell fate and cell cycle progression in the C. elegans germline. *Development* **138**: 2223–2234.

Friedland AE, Tzur YB, Esvelt KM, Colaiacovo MP, Church GM, Calarco JA. 2013. Heritable genome editing in C. elegans via a CRISPR-Cas9 system. *Nat Methods* **10**: 741–743.

Friedman JR, Nunnari J. 2014. Mitochondrial form and function. *Nature* **505**: 335–343.

Gallo M, Coutinho FJ, Vanner RJ, Gayden T, Mack SC, Murison A, Remke M, Li R, Takayama N, Desai K, et al. 2015. MLL5 Orchestrates a Cancer Self-Renewal State by Repressing the Histone Variant H3.3 and Globally Reorganizing Chromatin. *Cancer Cell* **28**: 715–729.

Gaydos LJ, Rechtsteiner A, Egelhofer TA, Carroll CR, Strome S. 2012. Antagonism

between MES-4 and Polycomb repressive complex 2 promotes appropriate gene expression in C. elegans germ cells. *Cell Rep* **2**: 1169–1177.

Ghazi A, Henis-Korenblit S, Kenyon C. 2009. A transcription elongation factor that links signals from the reproductive system to lifespan extension in Caenorhabditis elegans. *PLoS Genet* **5**: e1000639.

Greer EL, Becker B, Latza C, Antebi A, Shi Y. 2016. Mutation of C. elegans demethylase spr-5 extends transgenerational longevity. *Cell Res* **26**: 229–238.

Greer EL, Beese-Sims SE, Brookes E, Spadafora R, Zhu Y, Rothbart SB, Aristizabal-Corrales D, Chen S, Badeaux AI, Jin Q, et al. 2014. A histone methylation network regulates transgenerational epigenetic memory in C. elegans. *Cell Rep* **7**: 113–126.

Greer EL, Maures TJ, Hauswirth AG, Green EM, Leeman DS, Maro GS, Han S, Banko MR, Gozani O, Brunet A. 2010. Members of the H3K4 trimethylation complex regulate lifespan in a germline-dependent manner in C. elegans. *Nature* **466**: 383–387.

Greer EL, Maures TJ, Ucar D, Hauswirth AG, Mancini E, Lim JP, Benayoun BA, Shi Y, Brunet A. 2011. Transgenerational Epigenetic Inheritance of Longevity in C. elegans. *Nature* **479**: 365–371.

Greer EL, Shi Y. 2012. Histone methylation: a dynamic mark in health, disease and inheritance. *Nat Rev Genet* **13**: 343–357.

Grunstein M. 1997. Histone acetylation in chromatin structure and transcription. *Nature* **389**: 349–352.

Gu SG, Pak J, Guang S, Maniar JM, Kennedy S, Fire A. 2012. Amplification of siRNA in Caenorhabditis elegans generates a transgenerational sequence-targeted histone H3 lysine 9 methylation footprint. *Nat Genet* **44**: 157–164.

Guang S, Bochner AF, Burkhart KB, Burton N, Pavelec DM, Kennedy S. 2010. Small regulatory RNAs inhibit RNA polymerase II during the elongation phase of transcription. *Nature* **465**: 1097–1101.

Guang S, Bochner AF, Pavelec DM, Burkhart KB, Harding S, Lachowiec J, Kennedy S. 2008. An Argonaute transports siRNAs from the cytoplasm to the nucleus. *Science* **321**: 537–541.

Haithcock E, Dayani Y, Neufeld E, Zahand AJ, Feinstein N, Mattout A, Gruenbaum Y, Liu J. 2005. Age-related changes of nuclear architecture in Caenorhabditis elegans. *Proc Natl Acad Sci U S A* **102**: 16690–16695.

Hamilton B, Dong Y, Shindo M, Liu W, Odell I, Ruvkun G, Lee SS. 2005. A systematic RNAi screen for longevity genes in C. elegans. *Genes Dev* **19**: 1544–1555.

Han S, Brunet A. 2012. Histone methylation makes its mark on longevity. *Trends Cell Biol* **22**: 42–49.

Hansen M, Chandra A, Mitic LL, Onken B, Driscoll M, Kenyon C. 2008. A role for autophagy in the extension of lifespan by dietary restriction in C. elegans. *PLoS Genet* **4**: e24.

Hansen M, Hsu A-L, Dillin A, Kenyon C. 2005. New genes tied to endocrine, metabolic, and dietary regulation of lifespan from a Caenorhabditis elegans genomic RNAi

screen. *PLoS Genet* **1**: 119–128.

HARMAN D. 1956. Aging: a theory based on free radical and radiation chemistry. *J Gerontol* **11**: 298–300.

Harrison DE, Strong R, Sharp ZD, Nelson JF, Astle CM, Flurkey K, Nadon NL, Wilkinson JE, Frenkel K, Carter CS, et al. 2009. Rapamycin fed late in life extends lifespan in genetically heterogeneous mice. *Nature* **460**: 392–395.

Hesp K, Smant G, Kammenga JE. 2015. Caenorhabditis elegans DAF-16/FOXO transcription factor and its mammalian homologs associate with age-related disease. *Exp Gerontol* **72**: 1–7.

Hesterberg T, Moore DS, Monaghan S, Clipson A, Epstein R. 2005. Bootstrap methods and permutation tests. *Introd to Pract Stat* **5**: 1–70.

Heuser M, Yap DB, Leung M, de Algara TR, Tafech A, McKinney S, Dixon J, Thresher R, Colledge B, Carlton M, et al. 2009. Loss of MLL5 results in pleiotropic hematopoietic defects, reduced neutrophil immune function, and extreme sensitivity to DNA demethylation. *Blood* **113**: 1432–1443.

Holoch D, Moazed D. 2015. RNA-mediated epigenetic regulation of gene expression. *Nat Rev Genet* **16**: 71–84.

Horton JR, Upadhyay AK, Qi HH, Zhang X, Shi Y, Cheng X. 2010. Enzymatic and structural insights for substrate specificity of a family of jumonji histone lysine demethylases. *Nat Struct Mol Biol* **17**: 38–43.

Hsin H, Kenyon C. 1999. Signals from the reproductive system regulate the lifespan of

C. elegans. *Nature* **399**: 362–366.

Huang DW, Sherman BT, Lempicki RA. 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4**: 44–57.

Huang LS, Tzou P, Sternberg PW. 1994. The lin-15 locus encodes two negative regulators of Caenorhabditis elegans vulval development. *Mol Biol Cell* **5**: 395–411.

Huang XA, Yin H, Sweeney S, Raha D, Snyder M, Lin H. 2013. A major epigenetic programming mechanism guided by piRNAs. *Dev Cell* **24**: 502–516.

Hubbard EJA. 2007. Caenorhabditis elegans germ line: a model for stem cell biology. *Dev Dyn* **236**: 3343–3357.

Hwangbo DS, Gershman B, Tu M-P, Palmer M, Tatar M. 2004. Drosophila dFOXO controls lifespan and regulates insulin signalling in brain and fat body. *Nature* **429**: 562–566.

Jih G, Iglesias N, Currie MA, Bhanu N V, Paulo JA, Gygi SP, Garcia BA, Moazed D. 2017. Unique roles for histone H3K9me states in RNAi and heritable silencing of transcription. *Nature* **547**: 463–467.

Jin C, Li J, Green CD, Yu X, Tang X, Han D, Xian B, Wang D, Huang X, Cao X, et al. 2011. Histone demethylase UTX-1 regulates C. elegans life span by targeting the insulin/IGF-1 signaling pathway. *Cell Metab* **14**: 161–172.

Jung CH, Ro S-H, Cao J, Otto NM, Kim D-H. 2010. mTOR regulation of autophagy. *FEBS Lett* **584**: 1287–1295.

Kalinava N, Ni JZ, Peterman K, Chen E, Gu SG. 2017. Decoupling the downstream

effects of germline nuclear RNAi reveals that H3K9me3 is dispensable for heritable RNAi and the maintenance of endogenous siRNA-mediated transcriptional silencing in Caenorhabditis elegans. *Epigenetics Chromatin* **10**: 6. http://epigeneticsandchromatin.biomedcentral.com/articles/10.1186/s13072-017-0114-8.

Kaser-Pebernard S, Muller F, Wicky C. 2014. LET-418/Mi2 and SPR-5/LSD1 cooperatively prevent somatic reprogramming of C. elegans germline stem cells. *Stem cell reports* **2**: 547–559.

Katz DJ, Edwards TM, Reinke V, Kelly WG. 2009. A C. elegans LSD1 demethylase contributes to germline immortality by reprogramming epigenetic memory. *Cell* **137**: 308–320.

Kennedy S, Wang D, Ruvkun G. 2004. A conserved siRNA-degrading RNase negatively regulates RNA interference in C. elegans. *Nature* **427**: 645–649.

Kenyon C. 2005. The plasticity of aging: insights from long-lived mutants. *Cell* **120**: 449–460.

Kenyon C, Chang J, Gensch E, Rudner A, Tabtiang R. 1993. A C. elegans mutant that lives twice as long as wild type. *Nature* **366**: 461–464.

Kenyon CJ. 2010. The genetics of ageing. *Nature* **464**: 504–512.

Kerr SC, Ruppersburg CC, Francis JW, Katz DJ. 2014. SPR-5 and MET-2 function cooperatively to reestablish an epigenetic ground state during passage through the germ line. *Proc Natl Acad Sci U S A* **111**: 9509–9514.

Kimble J, Hirsh D. 1979. The postembryonic cell lineages of the hermaphrodite and male gonads in Caenorhabditis elegans. *Dev Biol* **70**: 396–417.

Kimble JE, White JG. 1981. On the control of germ cell development in Caenorhabditis elegans. *Dev Biol* **81**: 208–219.

Kostrouchova M, Krause M, Kostrouch Z, Rall JE. 2001. Nuclear hormone receptor CHR3 is a critical regulator of all four larval molts of the nematode Caenorhabditis elegans. *Proc Natl Acad Sci U S A* **98**: 7360–7365.

Krishnan V, Chow MZY, Wang Z, Zhang L, Liu B, Liu X, Zhou Z. 2011. Histone H4 lysine 16 hypoacetylation is associated with defective DNA repair and premature senescence in Zmpste24-deficient mice. *Proc Natl Acad Sci U S A* **108**: 12325–12330.

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.

Larson K, Yan S-J, Tsurumi A, Liu J, Zhou J, Gaur K, Guo D, Eickbush TH, Li WX. 2012. Heterochromatin formation promotes longevity and represses ribosomal RNA synthesis. *PLoS Genet* **8**: e1002473.

Le Thomas A, Rogers AK, Webster A, Marinov GK, Liao SE, Perkins EM, Hur JK, Aravin AA, Toth KF. 2013. Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state. *Genes Dev* **27**: 390–399.

Lee S-J, Hwang AB, Kenyon C. 2010. Inhibition of respiration extends C. elegans life span via reactive oxygen species that increase HIF-1 activity. *Curr Biol* **20**: 2131–

2136.

Lee SS, Lee RYN, Fraser AG, Kamath RS, Ahringer J, Ruvkun G. 2003. A systematic RNAi screen identifies a critical role for mitochondria in C. elegans longevity. *Nat Genet* **33**: 40–48.

Li J, Ebata A, Dong Y, Rizki G, Iwata T, Lee SS. 2008. Caenorhabditis elegans HCF-1 functions in longevity maintenance as a DAF-16 regulator. *PLoS Biol* **6**: e233.

Li T, Kelly WG. 2011. A role for Set1/MLL-related components in epigenetic regulation of the Caenorhabditis elegans germ line. *PLoS Genet* **7**: e1001349.

Liu B, Wang Z, Zhang L, Ghosh S, Zheng H, Zhou Z. 2013a. Depleting the methyltransferase Suv39h1 improves DNA repair and extends lifespan in a progeria mouse model. *Nat Commun* **4**: 1868.

Liu J, Cheng F, Deng L-W. 2012. MLL5 maintains genomic integrity by regulating the stability of the chromosomal passenger complex through a functional interaction with Borealin. *J Cell Sci* **125**: 4676–4685.

Liu L, Cheung TH, Charville GW, Hurgo BMC, Leavitt T, Shih J, Brunet A, Rando TA. 2013b. Chromatin modifications as determinants of muscle stem cell quiescence and chronological aging. *Cell Rep* **4**: 189–204.

Liu T, Rechtsteiner A, Egelhofer TA, Vielle A, Latorre I, Cheung M-S, Ercan S, Ikegami K, Jensen M, Kolasinska-Zwierz P, et al. 2011. Broad chromosomal domains of histone modification patterns in C. elegans. *Genome Res* **21**: 227–236.

Liu X, Jiang N, Hughes B, Bigras E, Shoubridge E, Hekimi S. 2005. Evolutionary

conservation of the clk-1-dependent mechanism of longevity: loss of mclk1 increases cellular fitness and lifespan in mice. *Genes Dev* **19**: 2424–2434.

Maciejowski J, Ugel N, Mishra B, Isopi M, Hubbard EJA. 2006. Quantitative analysis of germline mitosis in adult C. elegans. *Dev Biol* **292**: 142–151.

Madan V, Madan B, Brykczynska U, Zilbermann F, Hogeveen K, Dohner K, Dohner H, Weber O, Blum C, Rodewald H-R, et al. 2009. Impaired function of primitive hematopoietic cells in mice lacking the Mixed-Lineage-Leukemia homolog MLL5. *Blood* **113**: 1444–1454.

Magnuson B, Ekim B, Fingar DC. 2012. Regulation and function of ribosomal protein S6 kinase (S6K) within mTOR signalling networks. *Biochem J* **441**: 1–21.

Malone CD, Hannon GJ. 2009. Small RNAs as guardians of the genome. *Cell* **136**: 656–668.

Mao H, Zhu C, Zong D, Weng C, Yang X, Huang H, Liu D, Feng X, Guang S. 2015. The Nrde Pathway Mediates Small-RNA-Directed Histone H3 Lysine 27 Trimethylation in Caenorhabditis elegans. *Curr Biol* **25**: 2398–2403.

Mas-Y-Mas S, Barbon M, Teyssier C, Demene H, Carvalho JE, Bird LE, Lebedev A, Fattori J, Schubert M, Dumas C, et al. 2016. The Human Mixed Lineage Leukemia 5 (MLL5), a Sequentially and Structurally Divergent SET Domain-Containing Protein with No Intrinsic Catalytic Activity. *PLoS One* **11**: e0165139.

Maures TJ, Greer EL, Hauswirth AG, Brunet A. 2011. The H3K27 demethylase UTX-1 regulates C. elegans lifespan in a germline-independent, insulin-dependent

manner. *Aging Cell* **10**: 980–990.

McCarter J, Bartlett B, Dang T, Schedl T. 1999. On the control of oocyte meiotic maturation and ovulation in Caenorhabditis elegans. *Dev Biol* **205**: 111–128.

Morgan DE, Crittenden SL, Kimble J. 2010. The C. elegans adult male germline: stem cells and sexual dimorphism. *Dev Biol* **346**: 204–214.

Motamedi MR, Verdel A, Colmenares SU, Gerber SA, Gygi SP, Moazed D. 2004. Two RNAi complexes, RITS and RDRC, physically interact and localize to noncoding centromeric RNAs. *Cell* **119**: 789–802.

Murphy MP. 2009. How mitochondria produce reactive oxygen species. *Biochem J* **417**: 1–13.

MURRAY K. 1964. THE OCCURRENCE OF EPSILON-N-METHYL LYSINE IN HISTONES. *Biochemistry* **3**: 10–15.

Musa J, Orth MF, Dallmayer M, Baldauf M, Pardo C, Rotblat B, Kirchner T, Leprivier G, Grunewald TGP. 2016. Eukaryotic initiation factor 4E-binding protein 1 (4E-BP1): a master regulator of mRNA translation involved in tumorigenesis. *Oncogene* **35**: 4675–4688.

Ni Z, Ebata A, Alipanahiramandi E, Lee SS. 2012. Two SET domain containing genes link epigenetic changes and aging in Caenorhabditis elegans. *Aging Cell* **11**: 315–325.

Nottke AC, Beese-Sims SE, Pantalena LF, Reinke V, Shi Y, Colaiacovo MP. 2011. SPR-5 is a histone H3K4 demethylase with a role in meiotic double-strand break

repair. *Proc Natl Acad Sci U S A* **108**: 12805–12810.

Osipovich AB, Gangula R, Vianna PG, Magnuson MA. 2016. Setd5 is essential for mammalian development and the co-transcriptional regulation of histone acetylation. *Development* **143**: 4595–4607.

Owen DJ, Ornaghi P, Yang JC, Lowe N, Evans PR, Ballario P, Neuhaus D, Filetici P, Travers AA. 2000. The structural basis for the recognition of acetylated histone H4 by the bromodomain of histone acetyltransferase gcn5p. *EMBO J* **19**: 6141–6149.

Paix A, Wang Y, Smith HE, Lee C-YS, Calidas D, Lu T, Smith J, Schmidt H, Krause MW, Seydoux G. 2014. Scalable and versatile genome editing using linear DNAs with microhomology to Cas9 Sites in Caenorhabditis elegans. *Genetics* **198**: 1347–1356.

Pena P V, Davrazou F, Shi X, Walter KL, Verkhusha V V, Gozani O, Zhao R, Kutateladze TG. 2006. Molecular mechanism of histone H3K4me3 recognition by plant homeodomain of ING2. *Nature* **442**: 100–103.

Pratt AJ, MacRae IJ. 2009. The RNA-induced silencing complex: a versatile gene-silencing machine. *J Biol Chem* **284**: 17897–17901.

Pu M, Ni Z, Wang M, Wang X, Wood JG, Helfand SL, Yu H, Lee SS. 2015. Trimethylation of Lys36 on H3 restricts gene expression change during aging and impacts life span. *Genes Dev* **29**: 718–731.

Reinke V, Gil IS, Ward S, Kazmer K. 2004. Genome-wide germline-enriched and sex-biased expression profiles in Caenorhabditis elegans. *Development* **131**: 311–323.

Rincon-Arano H, Halow J, Delrow JJ, Parkhurst SM, Groudine M. 2012. UpSET recruits HDAC complexes and restricts chromatin accessibility and acetylation at promoter regions. *Cell* **151**: 1214–1228.

Rizki G, Iwata TN, Li J, Riedel CG, Picard CL, Jan M, Murphy CT, Lee SS. 2011. The evolutionarily conserved longevity determinants HCF-1 and SIR-2.1/SIRT1 collaborate to regulate DAF-16/FOXO. *PLoS Genet* **7**: e1002235.

Robert VJ, Mercier MG, Bedet C, Janczarski S, Merlet J, Garvis S, Ciosk R, Palladino F. 2014. The SET-2/SET1 histone H3K4 methyltransferase maintains pluripotency in the Caenorhabditis elegans germline. *Cell Rep* **9**: 443–450.

Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**: 139–140.

Rougvie AE, Ambros V. 1995. The heterochronic gene lin-29 encodes a zinc finger protein that controls a terminal differentiation event in Caenorhabditis elegans. *Development* **121**: 2491–2500.

Rozhkov N V, Hammell M, Hannon GJ. 2013. Multiple roles for Piwi in silencing Drosophila transposons. *Genes Dev* **27**: 400–412.

Santos F, Dean W. 2004. Epigenetic reprogramming during early development in mammals. *Reproduction* **127**: 643–651.

Sebastian S, Sreenivas P, Sambasivan R, Cheedipudi S, Kandalla P, Pavlath GK, Dhawan J. 2009. MLL5, a trithorax homolog, indirectly regulates H3K4 methylation,

represses cyclin A2 expression, and promotes myogenic differentiation. *Proc Natl Acad Sci U S A* **106**: 4719–4724.

Sen P, Dang W, Donahue G, Dai J, Dorsey J, Cao X, Liu W, Cao K, Perry R, Lee JY, et al. 2015. H3K36 methylation promotes longevity by enhancing transcriptional fidelity. *Genes Dev* **29**: 1362–1376.

Shen L, Shao N, Liu X, Nestler E. 2014. ngs.plot: Quick mining and visualization of next-generation sequencing data by integrating genomic databases. *BMC Genomics* **15**: 284.

Shi X, Hong T, Walter KL, Ewalt M, Michishita E, Hung T, Carney D, Pena P, Lan F, Kaadige MR, et al. 2006. ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression. *Nature* **442**: 96–99.

Shi X, Kachirskaia I, Walter KL, Kuo J-HA, Lake A, Davrazou F, Chan SM, Martin DGE, Fingerman IM, Briggs SD, et al. 2007. Proteome-wide analysis in Saccharomyces cerevisiae identifies several PHD fingers as novel direct and selective binding modules of histone H3 methylated at either lysine 4 or lysine 36. *J Biol Chem* **282**: 2450–2455.

Shi Y, Lan F, Matson C, Mulligan P, Whetstine JR, Cole PA, Casero RA, Shi Y. 2004. Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* **119**: 941–953.

Shilatifard A. 2012. The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis. *Annu Rev Biochem* **81**: 65–95.

Siebold AP, Banerjee R, Tie F, Kiss DL, Moskowitz J, Harte PJ. 2010. Polycomb Repressive Complex 2 and Trithorax modulate Drosophila longevity and stress resistance. *Proc Natl Acad Sci U S A* **107**: 169–174.

Sienski G, Donertas D, Brennecke J. 2012. Transcriptional silencing of transposons by Piwi and maelstrom and its impact on chromatin state and gene expression. *Cell* **151**: 964–980.

Sijen T, Fleenor J, Simmer F, Thijssen KL, Parrish S, Timmons L, Plasterk RH, Fire A. 2001. On the role of RNA amplification in dsRNA-triggered gene silencing. *Cell* **107**: 465–476.

Sims RJ 3rd, Millhouse S, Chen C-F, Lewis BA, Erdjument-Bromage H, Tempst P, Manley JL, Reinberg D. 2007. Recognition of trimethylated histone H3 lysine 4 facilitates the recruitment of transcription postinitiation factors and pre-mRNA splicing. *Mol Cell* **28**: 665–676.

Siomi MC, Sato K, Pezic D, Aravin AA. 2011. PIWI-interacting small RNAs: the vanguard of genome defence. *Nat Rev Mol Cell Biol* **12**: 246–258.

Tasdogan A, Kumar S, Allies G, Bausinger J, Beckel F, Hofemeister H, Mulaw M, Madan V, Scharfetter-Kochanek K, Feuring-Buske M, et al. 2016. DNA Damage-Induced HSPC Malfunction Depends on ROS Accumulation Downstream of IFN-1 Signaling and Bid Mobilization. *Cell Stem Cell* **19**: 752–767.

Towbin BD, Gonzalez-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**: 934–947.

Tsai W-W, Wang Z, Yiu TT, Akdemir KC, Xia W, Winter S, Tsai C-Y, Shi X, Schwarzer D, Plunkett W, et al. 2010. TRIM24 links a non-canonical histone signature to breast cancer. *Nature* **468**: 927–932.

Van Raamsdonk JM, Hekimi S. 2009. Deletion of the mitochondrial superoxide dismutase sod-2 extends lifespan in Caenorhabditis elegans. *PLoS Genet* **5**: e1000361.

Varela I, Cadinanos J, Pendas AM, Gutierrez-Fernandez A, Folgueras AR, Sanchez LM, Zhou Z, Rodriguez FJ, Stewart CL, Vega JA, et al. 2005. Accelerated ageing in mice deficient in Zmpste24 protease is linked to p53 signalling activation. *Nature* **437**: 564–568.

Vellai T, Takacs-Vellai K, Zhang Y, Kovacs AL, Orosz L, Muller F. 2003. Genetics: influence of TOR kinase on lifespan in C. elegans. *Nature* **426**: 620.

Verdel A, Jia S, Gerber S, Sugiyama T, Gygi S, Grewal SIS, Moazed D. 2004. RNAi-mediated targeting of heterochromatin by the RITS complex. *Science* **303**: 672–676.

Verdin E, Ott M. 2015. 50 years of protein acetylation: from gene regulation to epigenetics, metabolism and beyond. *Nat Rev Mol Cell Biol* **16**: 258–264.

Villeponteau B. 1997. The heterochromatin loss model of aging. *Exp Gerontol* **32**: 383–394.

von Mende N, Bird DM, Albert PS, Riddle DL. 1988. dpy-13: a nematode collagen gene that affects body shape. *Cell* **55**: 567–576.

Wullschleger S, Loewith R, Hall MN. 2006. TOR signaling in growth and metabolism. *Cell* **124**: 471–484.

Xiao Y, Bedet C, Robert VJP, Simonet T, Dunkelbarger S, Rakotomalala C, Soete G, Korswagen HC, Strome S, Palladino F. 2011. Caenorhabditis elegans chromatin-associated proteins SET-2 and ASH-2 are differentially required for histone H3 Lys 4 methylation in embryos and adult germ cells. *Proc Natl Acad Sci U S A* **108**: 8305–8310.

Yang J-S, Nam H-J, Seo M, Han SK, Choi Y, Nam HG, Lee S-J, Kim S. 2011. OASIS: online application for the survival analysis of lifespan assays performed in aging research. *PLoS One* **6**: e23525.

Yang W, Hekimi S. 2010. A mitochondrial superoxide signal triggers increased longevity in Caenorhabditis elegans. *PLoS Biol* **8**: e1000556.

Yigit E, Batista PJ, Bei Y, Pang KM, Chen C-CG, Tolia NH, Joshua-Tor L, Mitani S, Simard MJ, Mello CC. 2006. Analysis of the C. elegans Argonaute family reveals that distinct Argonautes act sequentially during RNAi. *Cell* **127**: 747–757.

Zeng L, Zhang Q, Li S, Plotnikov AN, Walsh MJ, Zhou M-M. 2010. Mechanism and regulation of acetylated histone binding by the tandem PHD finger of DPF3b. *Nature* **466**: 258–262.

Zhang M, Poplawski M, Yen K, Cheng H, Bloss E, Zhu X, Patel H, Mobbs C V. 2009a. Role of CBP and SATB-1 in aging, dietary restriction, and insulin-like signaling. *PLoS Biol* **7**: e1000245.

Zhang W, Li J, Suzuki K, Qu J, Wang P, Zhou J, Liu X, Ren R, Xu X, Ocampo A, et al. 2015. Aging stem cells. A Werner syndrome stem cell model unveils heterochromatin alterations as a driver of human aging. *Science* **348**: 1160–1163.

Zhang X, Novera W, Zhang Y, Deng L-W. 2017. MLL5 (KMT2E): structure, function, and clinical relevance. *Cell Mol Life Sci*.

Zhang Y, Wong J, Klinger M, Tran MT, Shannon KM, Killeen N. 2009b. MLL5 contributes to hematopoietic stem cell fitness and homeostasis. *Blood* **113**: 1455–1463.

Zhong M, Niu W, Lu ZJ, Sarov M, Murray JI, Janette J, Raha D, Sheaffer KL, Lam HYK, Preston E, et al. 2010. Genome-wide identification of binding sites defines distinct functions for Caenorhabditis elegans PHA-4/FOXA in development and environmental response. *PLoS Genet* **6**: e1000848.