2018-07-10

PIE-1, SUMOylation, and Epigenetic Regulation of Germline Specification in Caenorhabditis elegans

Heesun Kim
University of Massachusetts Medical School

PIE-1, SUMOYLATION, AND EPIGENETIC REGULATION OF GERMLINE SPECIFICATION IN *CAENORHABDITIS ELEGANS*

A Dissertation Presented

By

Heesun Kim

Submitted to the Faculty of the
University of Massachusetts Graduate School of Biomedical Sciences, Worcester
In partial fulfillment of the requirement for the degree of

DOCTOR OF PHILOSOPHY

July 10, 2018

Interdisciplinary graduate program

PIE-1, SUMOYLATION, AND EPIGENETIC REGULATION OF GERMLINE SPECIFICATION IN *CAENORHABDITIS ELEGANS*

A Dissertation Presented

By

Heesun Kim

This work was undertaken in the Graduate School of Biomedical Sciences
Interdisciplinary graduate program

Under the mentorship of

Craig C. Mello, Ph.D., Thesis Advisor

Erik J. Sontheimer, Ph.D., Member of Committee

Michael Francis, Ph.D., Member of Committee

Sean P. Ryder, Ph.D., Member of Committee

T. Keith Blackwell, M.D., Ph.D., External Member of Committee

Victor R. Ambros, Ph.D., Chair of Committee

Mary Ellen Lane, Ph.D.,
Dean of the Graduate School of Biomedical Sciences

July 10, 2018

DEDICATION

To my parents,

for your unconditional support

To my husband,

for always being with me whenever I need

To my son,

for just being with me

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my advisor Craig Mello. Thanks to his strong support, guidance, and passion for science, I was able to devote myself to my thesis project as well as many other interesting projects and grow my academic career as a dedicated scientist.

Victor Ambros was my QE chair, TRAC chair, and DEC chair. His advice and guidance have made it possible for me to enrich my thesis research. During my academic years with him, his knowledge and enthusiasm for science have always inspired me to be a better scientist.

I also want to thank my TRAC committee members, Sean Ryder, Marian Walhout, and Oliver Rando. They always asked thought-provoking questions and were willing to help me progress when I encountered a technical problem for SUMO purification assay.

I would like to thank Eric Sontheimer, Michael Francis, and Keith Blackwell for being in my dissertation committee and evaluate my thesis with critical and valuable comments.

Masaki is one of my great mentors whom I respect and appreciate. He is my bench mate I enjoyed working with. We shared every moment of exciting findings and frequently had scientific discussion.

I thank all of my colleagues I have spent time with. I had amazing years in the Mello lab as well as my unofficial second lab, Victor lab. It was always a pleasure coming to my lab every day and working with such kind and great

people. My interaction with each lab member has trained me to be a better scientist.

Lastly, I love and thank my parents, husband, and son for their unconditional support and love. Without them, I would have not been able to accomplish my PhD. and build my academic career.

ABSTRACT

In many organisms, the most fundamental event during embryogenesis is differentiating between germline cells and specialized somatic cells. In *C. elegans*, PIE-1 functions to protect the germline from somatic differentiation and appears to do so by blocking transcription and by preventing chromatin remodeling in the germline during early embryogenesis. Yet the molecular mechanisms by which PIE-1 specifies germline remain poorly understood. Our work shows that SUMOylation facilitates PIE-1-dependent germline maintenance and specification. *In vivo* SUMO purification in various CRISPR strains revealed that PIE-1 is SUMOylated at lysine 68 in the germline and that this SUMOylation is essential for forming NuRD complex and preserving HDA-1 activity. Moreover, HDA-1 SUMOylation is dependent on PIE-1 and enhanced by PIE-1 SUMOylation, which is required for protecting germline integrity. Our results suggest the importance of SUMOylation in the germline maintenance and exemplify simultaneous SUMOylation of proteins in the same functional pathway.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

PGCs: primordial germ cells

Bmp: bone morphogenetic protein

Blimp1: B-lymphocyte-induced maturation protein-1

Prdm14: PR domain-containing protein 14

Tcfap2c: transcription factor AP-2 gamma

HDAC: histone deacetylase

Gcl: germ cell-less

Pgc: polar granule component

Nos: Nanos

CTD: carboxy-terminal domain

P-TEFb: positive transcription elongation factor b

TAF-4: TATA-binding protein associated factor-4

NuRD: nucleosome remodeling and histone deacetylase

SUMO: small ubiquitin-like modifier

UBC: ubiquitin like conjugating

TDG: thymine DNA glycosylase

RNAi: RNA-mediated interference

TLS: translesion synthesis

CRISPR: Clustered Regularly Interspaced Short Palindromic Repeats

Cas9: CRISPR-associated nuclease

sgRNA: small guide RNA

Indels: small insertions and deletions

HR: homologous recombination

PCR: polymerase chain reaction

NGM: nematode growth medium

ORF: open reading frame

BSD: blasticidin S resistance gene

PAGE: polyacrylamide gel electrophoresis

NHEJ: non-homologous end joining

PAM: protospacer adjacent motif

PBS: predicted biological score

TBE: Tris/borate/EDTA

RFLP: restriction fragment length polymorphisms

IPTG: 1mM isopropyl β-d-thiogalactoside

PBS: phosphate-buffered saline

BSA: bovine serum albumin

DIC: differential interference contrast

GBP: GFP-binding protein

IP: immunoprecipitation

TCA: trichloroacetic acid

PVDF: polyvinylidene difluoride

ts: temperature-sensitive

TEV: a TEV protease cleavage site

him: high incidence of males

GCNA: germ cell nuclear antigen

CED-1: cell death abnormal 1

as: analog-sensitive

ZF: zinc finger

PIC: pol II initiation complex

MuDPIT: multidimensional protein identification technology

PREFACE

The CRISPR work presented in Chapter I has been previously published as *Kim, H., Ishidate, T., Ghanta, K.S., Seth, M., Conte, D., Jr., Shirayama, M., and Mello, C.C. A co-CRISPR strategy for efficient genome editing in Caenorhabditis elegans. Genetics 197, 1069-1080 (2014).* This paper was the result of the collaborative effort. I designed the experiment with Dr. Craig C. Mello and performed all the experimental works except a couple of things. Dr. Ishidate created the *unc-22* sgRNA, a Co-CRISPR marker. Dr. Shirayama provided method sources. I wrote the manuscript with contributions from all authors.

All of the works presented in Chapter II were performed by me. Dr. Craig C. Mello guided the experimental design and data analysis. I am preparing the manuscript for publication now.

The work presented in Chapter III is a preliminary body of work that explores another molecular mechanism of PIE-1 function. Dr. Craig C. Mello and I designed the experiments. I performed all of the works. Dr. Shirayama provided the idea for generation of the *pie-1(R109L)* mutant. Dr. Craig C. Mello and Dr. Shirayama guided data analysis.

CHAPTER I: INTRODUCTION

Germline Specification

As the precursor of gametes, germ cells contain all genetic and epigenetic information, and they function in transmitting this information from one generation to the next (Figure 1.1). Since this process is continuous, germ cells are considered to be immortal. In addition, germ cells are totipotent, having the potential to differentiate into any type of tissue. For example, the fertilized egg that is formed when a sperm fertilizes an oocyte produces all of the cells in an organism: these are the primordial germ cells that are the precursors of the germ cells and somatic cells that give rise to the rest of an organism (Figure 1.1). Protecting germ cell identity from somatic differentiation or mutation is, therefore, a crucial matter in all animals. The mechanisms that determine how the germline is distinguished from somatic tissues have been studied extensively in various model organisms.

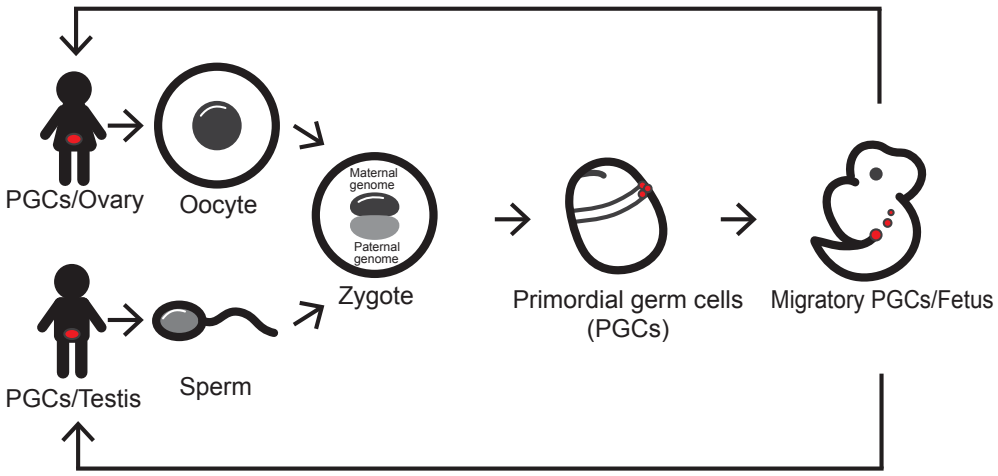


Figure 1.1 Life cycle of the germ cells

The life cycle of the germ cells is a continuous process, and the germ cells are immortal. A new life begins when a sperm fertilizes an oocyte and they together form a zygote. The zygote forms primordial germ cells that are the precursors of the gametes for the next generation and produces many differentiated somatic cells that give rise to the rest of the organism. The red dot represents germ cells.

Germline specification in mice, Drosophila, and C. elegans

In mice, primordial germ cells (PGCs) are induced in the epiblast by bone morphogenetic protein (Bmp) signaling (Figure 1.2A) (Lawson et al., 1999; Ying et al., 2000). WNT signaling also contributes to PGC generation because *wnt3* mutants fail to induce PGC although BMP signaling is activated (Figure 1.2A) (Ohinata et al., 2009). For the subsequent maintenance of the germline lineage, three other transcription factors play essential roles in somatic transcriptional regulation (Magnusdottir et al., 2013): B-lymphocyte-induced maturation protein-1 (Blimp1) directly represses somatic genes by binding to specific regulatory sequences (Ohinata et al., 2005; Vincent et al., 2005); PR domain-containing protein 14 (Prdm14) is required for the activation of pluripotency genes (Yamaji et al., 2008); and transcription factor AP-2 gamma (Tcfap2c) functions as a downstream effector of Blimp1 (Figure 1.2B) (Weber et al., 2010). PRDI-BF1, the human ortholog of mouse Blimp1, is known to repress transcription by binding and recruiting histone modification enzymes, such as histone deacetylase (HDAC) and histone H3 lysine methyltransferase G9a, to a transcription site (Gyory et al., 2004; Yu et al., 2000). Similarly, Blimp1 also interacts with Prmt5, an arginine-specific histone methyltransferase, which results in high levels of H2A/H4 R3 methylation to repress transcription (Ancelin et al., 2006).

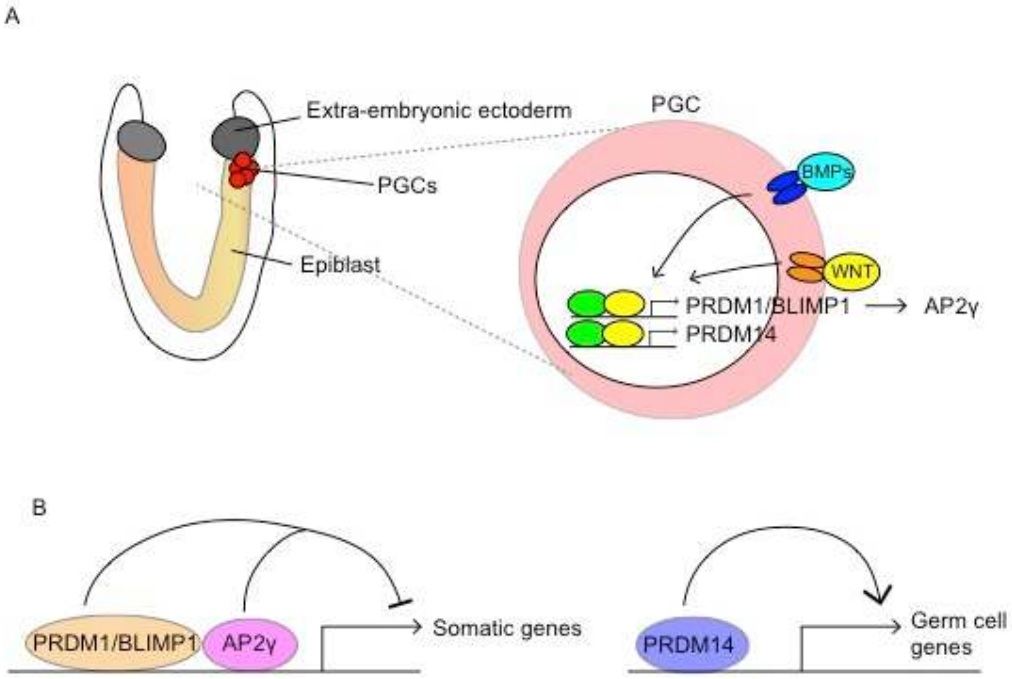


Figure 1.2 Specification of germline cells in mice

(A) PGCs are formed via BMP and WNT signaling in the epiblast. Both BMP and WNT signaling play a role in activating of PRDM1/BLMP1 and PRDM14. AP2γ is a downstream factor of PRDM1/BLMP1. (B) Roles of three key transcription factors in PGCs. Both PRDM1/BLMP1 and AP2γ together are required for repression of somatic genes and PRDM14 is responsible for activation of germ cell genes.

In *Drosophila*, PGCs are formed during oogenesis when Oskar, a germline determinant in germ granules, leads to germ/pole plasm assembly (Figure 1.3) (Ephrussi and Lehmann, 1992; Illmensee and Mahowald, 1974). The specification of the germline cells is facilitated by three maternally loaded factors, *germ cell-less (gcl), polar granule component (pgc),* and *nanos (nos)*, to inhibit transcription in PGCs (Figure 1.3). Pgc is required for inhibition of carboxy-terminal domain (CTD) Ser2 phosphorylation of RNA Pol II via the sequestration of positive transcription elongation factor b (P-TEFb) kinase in the germline (Hanyu-Nakamura et al., 2008). During the formation of PGCs, Gcl plays a role in transcriptional repression of a subset of somatic genes through unknown mechanisms (Leatherman et al., 2002). As a downstream effector of Pgc, Nos inhibits histone modification enzymes that are involved in transcriptional activation (Kobayashi et al., 1996; Sato et al., 2007).

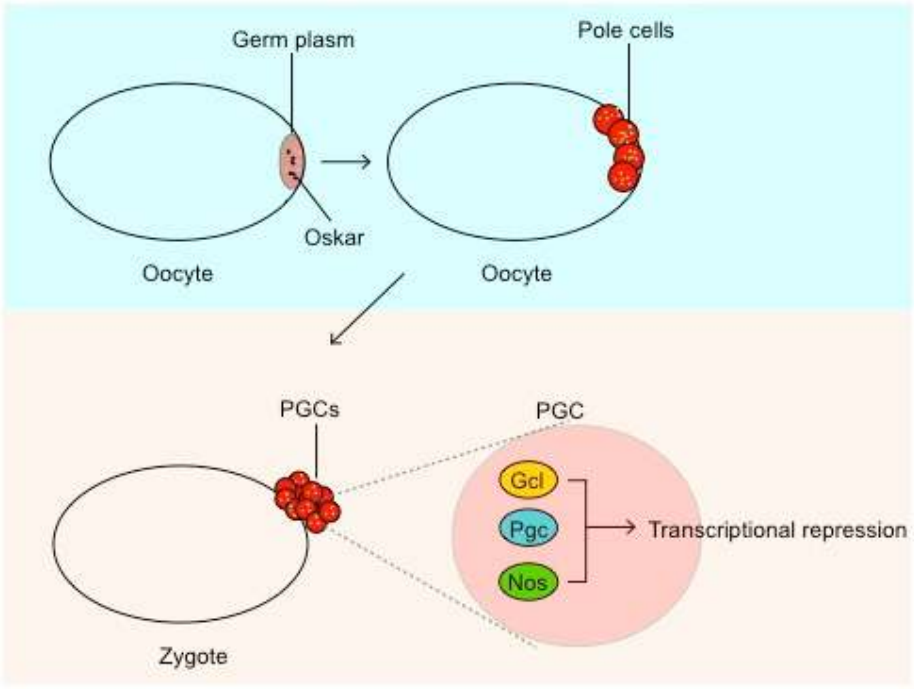


Figure 1.3 Specification of germline cells in *Drosophila*

In *Drosophila,* Oskar is responsible for germ plasm assembly in the oocyte. Three factors that are localized in germ cell nuclei, Gcl, Pgc, and Nos, are required for transcriptional repression.

In *Caenorhabditis elegans*, PGCs are established through a series of four asymmetric divisions of an embryo. The first asymmetric division of the zygote (P_0) gives rise to a larger somatic blastomere (AB) and a smaller germline blastomere (P_1). Asymmetric divisions continue until the formation of the P_4 germline blastomere, which is the precursor of the PGCs. Then, the P_4 germline blastomere undergoes equal cell division to generate two PGCs, Z2 and Z3, which resume divisions to generate the entire germline once the first stage (L1) larvae start feeding (Figure 1.4).

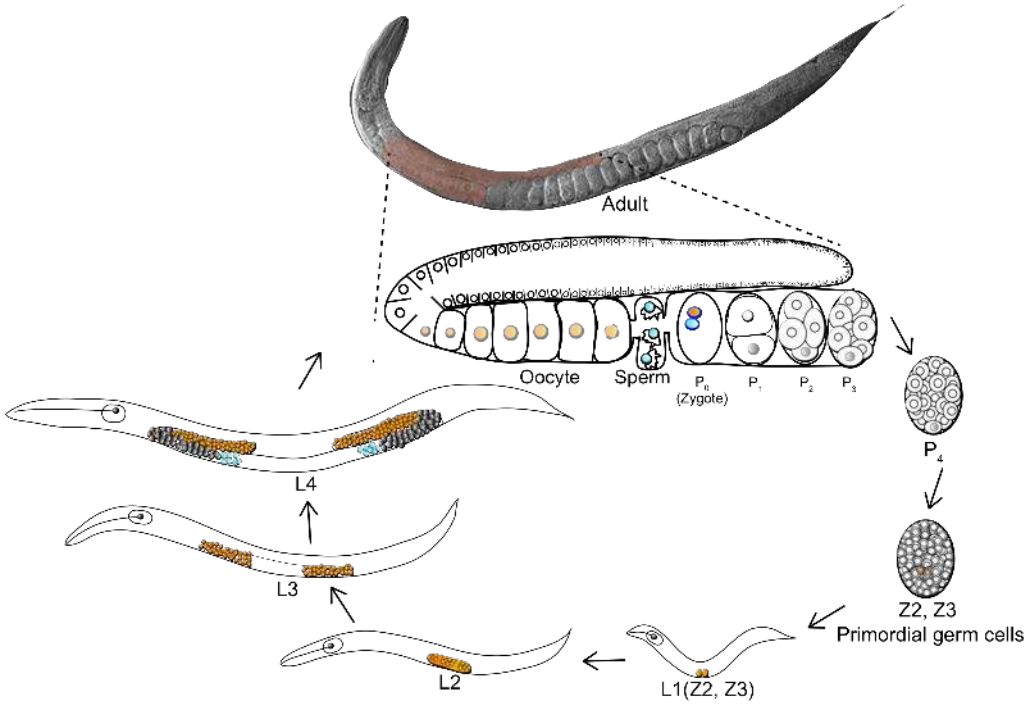


Figure 1.4 Specification of germline cells in *C. elegans*

In *C. elegans*, the adult gonad (indicated by red shading) contains many germ cells. These germ cells develop into gametes, oocytes and sperms, and they together form a fertilized egg denoted P_0. P_0 divides into one germ cell (P_1) and one somatic cell. P_1 gives rise to P_2 and its sister somatic cell. These divisions continue to occur until the P_4 blastomere generates two PGCs, Z2 and Z3. Once L1 is hatched, Z2 and Z3 resume divisions to proliferate germ cells, which lead to the entire germline in the adult.

During these asymmetric cleavages, cytoplasmic germ granules (called P-granules) are segregated to each germline blastomere. P-granules contain many maternally loaded RNAs and RNA-binding proteins (Strome, 2005). Among these are vasa-related RNA helicases GLH1–4 and the RGG domain proteins PGL1/3 that persist in P-granules at all developmental stages (Gruidl et al., 1996; Kawasaki et al., 2004; Kawasaki et al., 1998; Kuznicki et al., 2000). *pgl* and *glh* mutants show severe defects in larvae germ cell proliferation and gamete formation but not in germline specification (Kawasaki et al., 2004; Kuznicki et al., 2000; Spike et al., 2008). The contribution of P-granules to germline specification is, however, still inconclusive due to redundancy and a strong maternal effect (Wang and Seydoux, 2013).

In *C. elegans*, three maternally loaded factors OMA-1, OMA-2 and PIE-1 regulate transcriptional repression in the early germline lineage. OMA-1 and OMA-2 are cytoplasmic proteins expressed in oocytes and the very early embryo; thus, their functional studies of these proteins have focused on post-transcriptional regulation for oocyte maturation, instead of transcriptional regulation for germline specification (Detwiler et al., 2001; Kaymak and Ryder, 2013; Lin, 2003; Nishi and Lin, 2005). In addition to regulating oocyte mRNA expression, however, OMA-1 and OMA-2 also globally repress transcription in the early germline blastomeres, P_0 and P_1, where they bind and sequester to the cytoplasm TATA-binding protein associated factor-4 (TAF-4), an essential component for transcriptional initiation in the nucleus (Guven-Ozkan et al., 2008).

PIE-1 is a key regulator of germline specification, given that germline blastomeres in *pie-1* mutants adopt somatic fates (Mello et al., 1992). In addition, PIE-1 is segregated to each germline blastomere during the first four embryonic cleavages and its nuclear localization correlates with the role of PIE-1 in the transcriptional repression of somatic differentiation (Mello et al., 1996). The molecular mechanisms by which PIE-1 functions as a transcriptional repressor to protect the germline lineage from somatic differentiation have been extensively studied.

PIE-1 and transcriptional repression

The *pie-1* gene was first identified in a screen for maternal-effect lethal mutations, and *pie-1* mutant embryos produce excess numbers of pharyngeal and intestinal cells due to transformation of the germline blastomere P_2 into a somatic blastomere, like its sister EMS (Mello et al., 1992). PIE-1 protein contains two CCCH zinc finger motifs and segregates with each germline blastomere (Mello et al., 1996). Nuclear localization of PIE-1 in the germline blastomeres has raised the possibility that PIE-1 might regulate transcription to repress somatic fates in the germline lineage. For example, one known factor regulated by PIE-1 is the SKN-1 transcription factor that is required for somatic differentiation but is present in both the germline blastomere P_2 and the somatic blastomere EMS (Bowerman et al., 1993). In wild-type embryos, SKN-1 activity is thought to be down-regulated by PIE-1 to protect the germline lineage in P_2,

because SKN-1 drives germline blastomere differentiation into somatic tissues in *pie-1* mutants (Mello et al., 1992). Newly transcribed RNAs and CTD Ser2 phosphorylation of Pol II are detected in both germline and somatic blastomeres in *pie-1* mutants, whereas they are detected only in somatic blastomeres in wild-type embryos (Seydoux and Dunn, 1997; Seydoux et al., 1996). Studies in a vertebrate cell culture assay suggested that PIE-1 functions as a transcriptional repressor using a specific repression sequence YAPMAPT on the C-terminus of PIE-1 (Batchelder et al., 1999). This sequence resembles the Pol II CTD heptapeptide repeat (YSPTSPS) sequence where Ser2 is phosphorylated during transcriptional elongation and Ser5 is phosphorylated during initiation (Figure 1.5) (Batchelder et al., 1999; Zaborowska et al., 2016). Therefore, the YAPMAPT sequence competitively targets a Pol II CTD-binding complex to block transcription (Figure 1.5) (Batchelder et al., 1999). PIE-1 (DAQMEQT) transgenes with nonconservative mutations of the YAPMAPT motif eliminate transcriptional repression and significantly reduce the frequency of *pie-1* rescue compared to the wild-type transgene (Batchelder et al., 1999). The YAPMAPT motif binds to Cyclin T, a regulatory subunit of P-TEFb, to inhibit Pol II elongation (Figure 1.5), whereas PIE-1 (DAQMEQT) does not (Zhang et al., 2003). A subsequent study of the YAPMAPT motif was performed in *C. elegans* using low-copy integrated transgenic lines generated by ballistic transformation (Praitis et al., 2001). The results showed that the YAPMAPT is sufficient to inhibit Ser2 phosphorylation, but not Ser5 phosphorylation (Ghosh and Seydoux, 2008).

Thus, PIE-1 (DAQMEQT) still has the ability to inhibit transcription like the wild-type, which suggests PIE-1 regulates Ser5 phosphorylation by interacting with other components of the transcriptional machinery (Ghosh and Seydoux, 2008).

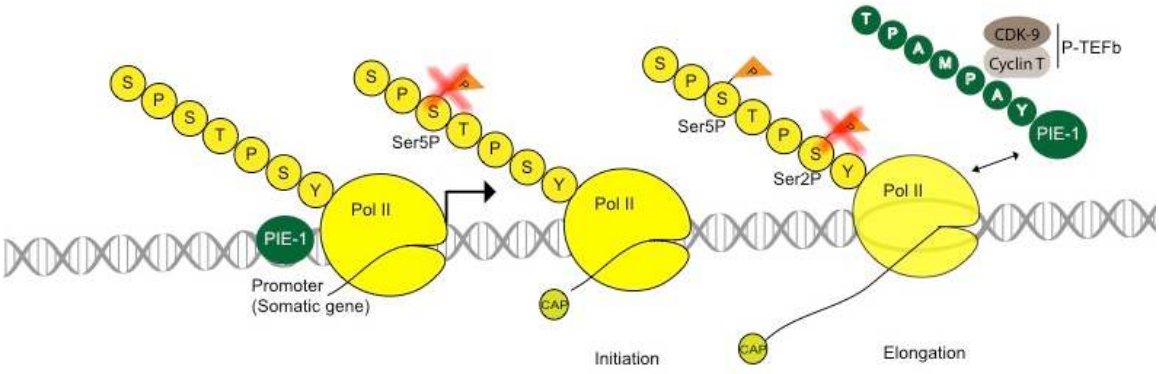


Figure 1.5 PIE-1 represses transcription in the germline blastomeres
PIE-1 contains the YAPMAPT sequence that resembles the sequence YSPTSPS of the pol II CTD where Ser2 is phosphorylated during transcriptional elongation and Ser5 is phosphorylated during initiation. The YAPMAPT sequence competitively binds to CyclinT, the subunit of P-TEFb that is required for Ser2 phosphorylation, and blocks P-TEFb recruitment to pol II CTD. PIE-1 also blocks Ser5 phosphorylation by an unknown mechanism.

PIE-1 and chromatin regulation

PIE-1 protein begins to disappear immediately after the birth of Z2 and Z3. At this time, Z2 and Z3 show a transiently active chromatin state, and then revert to and retain the repressive chromatin state until they resume cell divisions at the L1 larval stage (Furuhashi et al., 2010; Schaner et al., 2003). These specified-dynamic chromatin states in PGCs depend on MES chromatin regulators: MES-2, MES-3, MES-4 and MES-6. *mes* genes were identified from a screen for maternal effect sterility (Capowski et al., 1991), and even though they have a similar mutant phenotype, MES-4 functions differently from the others (Bender et al., 2004; Bender et al., 2006; Fong et al., 2002; Xu et al., 2001). For example, MES-2, MES-3, and MES-6 methylate Lys27 of histone H3, a repressive chromatin mark (Bender et al., 2004; Xu et al., 2001), whereas MES-4 methylates Lys 36 of histone H3, an active or potentially active chromatin mark (Bender et al., 2006; Fong et al., 2002; Furuhashi et al., 2010; Rechtsteiner et al., 2010). In embryos, MES-4 is present in all blastomeres until the 100-cell stage, but its expression level is diminished in somatic blastomeres except in Z2 and Z3 (Fong et al., 2002). Given that MES-4 plays a role in establishing an active or potentially active transcription mark, MES-4 activity seems to be regulated in somatic blastomeres to preserve the somatic fates and prevent germline potential. Genetic studies suggest that chromatin regulators in the synthetic multivulva B pathway antagonize MES-4 activity (Unhavaithaya et al., 2002; Wang et al., 2005). Among these is MEP-1, a Krüppel-class zinc-finger protein

(Belfiore et al., 2002) that interacts with two subunits of the nucleosome remodeling and histone deacetylase (NuRD) complex (Unhavaithaya et al., 2002): LET-418, a homolog of mammalian ATP-dependent nucleosome remodeling factor Mi-2 (von Zelewsky et al., 2000) and HDA-1, a homolog of mammalian histone deacetylase HDAC-1(Shi and Mello, 1998). PIE-1 interacts with MEP-1 and the NuRD complex in embryos to repress NuRD complex-mediated chromatin remodeling, thereby protecting a germline-specific chromatin state that is established by MES proteins (Unhavaithaya et al., 2002).

SUMOylation

SUMOylation is a post-translational modification that covalently and reversibly attaches small ubiquitin-like modifier (SUMO) to lysine residues of protein substrates. SUMOylation mainly occurs on a consensus SUMO acceptor site, consisting of the sequence ψKXE, where ψ represents a hydrophobic amino acid and X is any amino acid (Rodriguez et al., 2001), although a non-consensus motif also can be modified by SUMO (Tammsalu et al., 2014). In addition, not all proteins containing the consensus SUMO sequence are modified by SUMO. SUMOylation requires an enzymatic cascade involving E1-activating enzyme (Aos/Uba2p in yeasts), E2-conjugating enzyme (ubiquitin-like conjugating enzyme 9, Ubc9), and E3 ligase (Gareau and Lima, 2010). Ubc9 is a single SUMO E2 enzyme and has a unique ability to recognize and conjugate even without SUMO E3 ligase, although SUMO E3 ligases facilitate SUMO conjugation to a substrate (Hoeller et al., 2007; Kerscher et al., 2006). SUMOylation is a reversible modification because the conjugated SUMO is deconjugated from substrates by SUMO-specific proteases and free SUMO can be recycled in another round of SUMOylation (Gareau and Lima, 2010).

The reversible SUMOylation cycle and the enzymes required for the SUMOylation pathway are conserved in *C. elegans*. In *C. elegans*, SUMO is encoded by a single gene *smo-1*, whereas four SUMO genes have been identified in the vertebrates (Guo et al., 2004; Melchior, 2000). The *ubc-9* gene

encodes a SUMO E2 conjugating enzyme (Jones et al., 2002) and *gei-17* encodes a Siz/PIAS-type SUMO E3 ligase (Holway et al., 2005).

Unlike ubiquitination that leads to protein degradation, SUMOylation regulates protein functions for diverse biological processes including transcription, cell cycle progression, and DNA repair (Hay, 2005). For example, SUMO-modified transcription factors change their cellular localization, stability, or binding partners to affect transcriptional activity (Gill, 2005; Hay, 2005). Although many important biological functions require SUMOylation, identification of SUMO target proteins *in vivo* is challenging because the abundance of a particular protein that is SUMOylated at steady state is relatively very low compared to the total pool of the protein (usually less than 5%) and many SUMO-specific proteases actively remove SUMO from substrates. However, it is notable that even a small portion of SUMO-modified protein can cause a dramatic effect (Geiss-Friedlander and Melchior, 2007; Hay, 2005). In addition, a defect in SUMOylation of a single substrate shows a very mild phenotype in most cases, but results in strong defects in a few cases (Sarangi and Zhao, 2015).

SUMOylation and transcriptional repression

Although SUMOylation of a transcription factor can cause transcriptional activation, it usually leads to transcriptional repression by either recruiting repressors with chromatin remodeling activity or facilitating assembly of a transcriptional repression complex (Geiss-Friedlander and Melchior, 2007; Gill,

2005; Hay, 2005). HDAC recruitment is directly or indirectly involved in SUMOylation-mediated chromatin remodeling (Gill, 2005). For example, in *C. elegans*, SUMOylated LIN-1 represses transcription by recruiting the NuRD complex, which leads to inhibition of 1° vulval cell fate (Leight et al., 2005).

SUMOylation and DNA repair

Another essential biological function mediated by SUMOylation is DNA repair. Many DNA repair proteins are modified by SUMO, which affects their DNA binding affinity, promotes protein–protein interaction, and affects their protein stability to tightly control DNA damage repair and response (Sarangi and Zhao, 2015). Thymine DNA glycosylase (TDG), an essential repair enzyme for base excision repair, is a well-known SUMO substrate. TDG SUMOylation allows recycling of TDG: For example, TDG SUMOylation leads to release of TDG from its DNA substrate by reducing DNA affinity and then SUMO-specific proteases deconjugate the released TDG, which is free again to bind to mismatches in other DNA substrates (Hardeland et al., 2002). In *C. elegans*, depletion of *ubc-9* and *gei-17* using RNA-mediated interference (RNAi) results in high sensitivity to DNA damage and both *smo-1* and *ubc-9* null mutants fail to repair DNA breaks that occur in mitotically dividing nuclei in the premeiotic distal tip, suggesting SUMOylation plays an important role in the DNA damage repair and response (Boulton et al., 2004; Holway et al., 2006; Reichman et al., 2018). One subsequent study of *gei-17* suggested that GEI-17 leads to SUMOylation of

translesion synthesis (TLS) DNA polymerase POLH-1, and SUMOylated POLH-1 is protected from degradation before it can perform its function in response to DNA damage (Kim and Michael, 2008). Recently, ZTF-8 was identified as a SUMO target protein involved in DNA repair pathways in the germline (Kim and Colaiacovo, 2015). ZTF-8, a homolog of mammalian RHINO that interacts with the Rad9-Rad1-Hus1 complex, is required for DNA double strand break repair and DNA damage-checkpoint–mediated apoptosis (Kim and Colaiacovo, 2014).

A Co-CRISPR Strategy for Efficient Genome Editing in *C. elegans*

This work has been published previously

Abstract

Genome editing based on CRISPR (clustered regularly interspaced short palindromic repeats)-associated nuclease (Cas9) has been successfully applied in dozens of diverse plant and animal species including the nematode *Caenorhabditis elegans*. The rapid life-cycle and easy access to the ovary by microinjection make *C. elegans* an ideal organism both for applying CRISPR-Cas9 genome editing technology and for optimizing genome-editing protocols. Here we report efficient and straightforward CRISPR-Cas9 genome editing methods for *C. elegans*, including a Co-CRISPR strategy that facilitates detection of genome-editing events. We describe methods for detecting homologous recombination (HR) events, including direct screening methods as well as new selection/counter-selection strategies. Our findings reveal a surprisingly high frequency of HR-mediated gene conversion, making it possible to rapidly and precisely edit the *C. elegans* genome both with and without the use of co-inserted marker genes.

Introduction

Sequence-specific immunity mechanisms such as RNA interference (Grishok and Mello, 2002; Hannon, 2002; Voinnet, 2001; Zamore, 2001) and CRISPR-Cas9 (Bhaya et al., 2011; Horvath and Barrangou, 2010; Terns and Terns, 2011; Wiedenheft et al., 2012) provide sophisticated cellular defense against invasive nucleic acids. Understanding how these defense systems work has enabled researchers to re-direct them at cellular targets, providing powerful tools for manipulating both gene expression and the cellular genome itself. The CRISPR-Cas9 system is a bacterial anti-viral mechanism that captures fragments of viral DNA in specialized genomic regions for re-expression as small guide RNAs (sgRNAs) (Bhaya et al., 2011; Terns and Terns, 2011; Wiedenheft et al., 2012). In bacterial cells Cas9-sgRNA complexes provide acquired immunity against viral pathogens (Bhaya et al., 2011; Terns and Terns, 2011; Wiedenheft et al., 2012). When co-expressed along with an artificial sgRNA designed to target a cellular gene, the Cas9 nuclease has been shown to efficiently direct the formation of double-strand breaks at the corresponding target locus (Jinek et al., 2012). Remarkably, this mechanism works efficiently even within the context of eukaryotic chromatin (Gilbert et al., 2013). Genome editing using CRISPR-Cas9 has recently been demonstrated in numerous organisms, providing a powerful new tool with rapidly growing–if not infinite–potential for diverse biological applications (Bassett et al., 2013; Chang et al., 2013; Cho et al., 2013a; Cong et al., 2013; Dicarlo et al., 2013; Dickinson et al., 2013; Feng et al., 2014; Feng et

al., 2013; Friedland et al., 2013; Gratz et al., 2013; Jiang et al., 2013; Ma et al., 2014; Mali et al., 2013b; Wang et al., 2013; Yu et al., 2014; Zhou et al., 2014).

The CRISPR-Cas9 system has also been successfully applied to *C. elegans*. Methods that have been used to express Cas9 include mRNA injection and transgene-driven expression from a constitutive or an inducible promoter (Chen et al., 2013; Chiu et al., 2013; Cho et al., 2013b; Dickinson et al., 2013; Friedland et al., 2013; Katic and Grosshans, 2013; Lo et al., 2013; Tzur et al., 2013; Waaijers et al., 2013; Zhao et al., 2014). The U6 promoter has been used to drive sgRNA expression (Chiu et al., 2013; Dickinson et al., 2013; Friedland et al., 2013; Katic and Grosshans, 2013; Waaijers et al., 2013). The system has been used widely to produce small insertions and deletions (indels) that shift the reading frame of the target gene, often resulting in premature termination of translation and loss-of-function phenotypes (Chiu et al., 2013; Cho et al., 2013b; Friedland et al., 2013; Lo et al., 2013; Waaijers et al., 2013). In addition, single-stranded oligonucleotides have been used as donor molecules to precisely alter a target gene through homologous recombination (HR) (Zhao et al., 2014), and a selection scheme has been developed that allows the HR-mediated insertion of large sequence tags such as GFP (Chen et al., 2013; Dickinson et al., 2013; Tzur et al., 2013).

Despite these important advances, current CRISPR protocols for inducing indels and HR events in *C. elegans* could benefit from refinement. For example, different sgRNAs targeting the same gene can result in substantially variable

DNA cleavage efficiencies (Bassett et al., 2013; Chen et al., 2013; Wang et al., 2014); thus, identifying active sgRNAs can be time consuming and costly. In this study, we investigate several strategies to streamline CRISPR-Cas9-mediated genome editing in *C. elegans*. We describe a co-CRISPR strategy that can facilitate the identification of functional sgRNAs, and can enrich for transgenic animals carrying an HR event. We show that HR events can be identified without the need for selection at a rate of approximately 1% to as high as 10% of F1 transgenic animals scored. This high frequency allows HR events to be identified by directly scoring for GFP expression, or by polymerase chain reaction (PCR) screening to detect HR-induced DNA polymorphisms. Direct screening allows precise genome alterations that minimize the footprint of DNA alterations, such as inserted selectable markers, at the target locus. However in some cases, such as whole-gene deletion assays that may induce lethality, selection can be useful for both identifying and maintaining HR events. We therefore describe a straight-forward selection/counter-selection protocol that facilitates recovery of HR events where having a marker inserted at the target site might be tolerated or useful. Together the findings presented here take much of the guesswork out of using the CRISPR-Cas9 system in *C. elegans*, and the co-CRISPR strategy employed here may also prove useful in other organisms.

Materials and Methods

Genetics

All strains in this study were derived in the Bristol N2 background and maintained on nematode growth medium (NGM) plates seeded with OP50 (Brenner, 1974).

Selection of sgRNA target sequences

We manually searched for target sequences consisting of $G(N)_{19}NGG$ (Friedland et al., 2013; Wiedenheft et al., 2012) near the desired mutation site. For HR-mediated repair experiments such as *gfp* knock-in and introduction of point mutation, we selected the target sequences where it was possible to introduce a silent mutation in the PAM site. Target sequences are listed in Table 1.1.

Table 1.1 Summary of sgRNAs sequences and their efficiency

Name	Sequence	S/AS	% efficiency
avr-14 no.1	GAATATTGAAAGACTATGAT(TGG)	S	10
avr-14 no.2	GATTGGAGAGTTAGACCACG(TGG)	S	20
avr-15 no.9	GCAGAAAATGAATGTCATAC(AGG)	AS	HIGH
avr-15 no.10	GTTTGCAATATAAGTCACCC(AGG)	AS	HIGH
unc-22 no.2	GAACCCGTTGCCGAATACAC(AGG)	S	5
unc-22 no.9	GCCTTTGCTTCGATTTTCTT(TGG)	AS	0
unc-4 no.1	GTTATCGTCATCCGGTGACG(TGG)	AS	10
rde-3 no.3	GAATTTGAGCTTGAACGAGC(TGG)	AS	LOW
rde-3 no.4	GTCGATACTTCAAAATTAAT(TGG)	AS	LOW
lon-2 no.1	GGGAAACTATACCCTCACTG(TGG)	S	30
dpy-11 no.2	GCAAGGATCTTCAAAAAGCA(CGG)	S	0.4
dpy-11 no.4	GATGCTTGTAGTCTGGAACT(GGG)	AS	
unc-32 no.1	GATAGGAAGCATCAGATTGA(AGG)	AS	0
unc-32 no.2	GTTGCTGAACTGGGAGAGCT(CGG)	S	
bli-2 no.1	GGATTTGCTGCTACTGAATC(CGG)	AS	0
bli-2 no.2	GATGGACGGGATGGTAGAGA(TGG)	S	
dpy-5 no.2	GTCGGATTCGGCGCTGCATG(CGG)	S	0
dpy-5 no.3	GGTTTCCTGGAGCTCCGGCT(GGG)	AS	
ben-1 no.3	GGATATCACTTCCCAGAACT(TGG)	AS	0
ben-1 no.5	GGGAGAAAGTGATTTGCAGT(TGG)	S	
pie-1 a	GGCTCAGATTGACGAGGCGC(CGG)	S	24
pie-1 b	GCTGAGAGAAGAATCCATCG(GGG)	AS	15
pie-1 c	GGACAAAGAGAGGGGGTGAG(TGG)	AS	7.5
pie-1 d	GTTGAGTGCAGCCATTTGCT(CGG)	AS	5
smo-1 a	GCCGATGATGCAGCTCAAGC(AGG)	S	LOW
smo-1 b	GTGCACTTCCGTGTAAAGTA(TGG)	S	HIGH
smo-1 c	GTCTACCAAGAGCAGCTGGG(CGG)	S	HIGH
smo-1 d	GTATCTCAGTGGAAAAGGGA(TGG)	S	HIGH
vet-2	GTTGGATCATAGGATACCGG(TGG)	AS	38
C35E7.6	GGGCACCATACCGAGTGATG(GGG)	AS	100

Preparation of sgRNA constructs

We replaced the *unc-119* target sequence in pU6::unc-119 sgRNA vector (Friedland et al., 2013) with the desired target sequence using overlap extension PCR. The pU6::unc-119 sgRNA vector was diluted to 2 ng/µl and used as a template to generate two overlapping fragments. The first was amplified using the primers CMo16428 and sgRNA R, resulting in the U6 promoter fused to the GN_{19} target sequence (U6p::GN_{19}). The second was amplified using the primers CMo16429 and sgRNA F, resulting in the GN_{19} target sequence fused to the sgRNA scaffold and U6 3'UTR. These two PCR products were mixed together, diluted 1:50, and used as a template for a PCR reaction with primers CMo16428 and CMo16429. The resulting pU6::target sequence::sgRNA scaffold::U6 3'UTR fusion products were gel-purified and inserted into the pCR^{tm}-Blunt II-TOPO® vector (Invitrogen, cat. no. K2800-20). We used iProof^{TM} high-fidelity DNA polymerase (Bio-Rad, cat. no. 172-5300) in all PCR reactions above to minimize errors of PCR amplification, and all the constructs were confirmed by DNA sequencing. Primers sequences are listed in Table 1.2.

Table 1.2 Summary of primer sequences for sgRNA plasmid generation

Name	Sequence
CMO16428	TGAATTCCTCCAAGAACTCG
CMO16429	AAGCTTCACAGCCGACTATG
sgRNA_F	G(N)19GTTTTAGAGCTAGAAATAGC
avr-14 sgRNA_F	GATTGGAGAGTTAGACCACGGTTTTAGAGCTAGAAATAGC
avr-15 sgRNA_F	GTTTGCAATATAAGTCACCCGTTTTAGAGCTAGAAATAGC
unc-22 sgRNA_F	GAACCCGTTGCCGAATACACGTTTTAGAGCTAGAAATAGC
pie-1 a sgRNA_F	GGCTCAGATTGACGAGGCGCGTTTTAGAGCTAGAAATAGC
pie-1 b sgRNA_F	GCTGAGAGAAGAATCCATCGGTTTTAGAGCTAGAAATAGC
pie-1 c sgRNA_F	GGACAAAGAGAGGGGGTGAGGTTTTAGAGCTAGAAATAGC
smo-1 sgRNA_F	GTATCTCAGTGGAAAAGGGAGTTTTAGAGCTAGAAATAGC
vet-2 sgRNA_F	GTTGGATCATAGGATACCGGGTTTTAGAGCTAGAAATAGC
sgRNA_R	(N)19CAAACATTTAGATTTGCAATTC
avr-14 sgRNA_R	CGTGGTCTAACTCTCCAATCCAAACATTTAGATTTGCAATTC
avr-15 sgRNA_R	GGGTGACTTATATTGCAAACCAAACATTTAGATTTGCAATTC
unc-22 sgRNA_R	GTGTATTCGGCAACGGGTTCCAAACATTTAGATTTGCAATTC
pie-1 a sgRNA_R	GCGCCTCGTCAATCTGAGCCCAAACATTTAGATTTGCAATTC
pie-1 b sgRNA_R	CGATGGATTCTTCTCTCAGCCAAACATTTAGATTTGCAATTC
pie-1 c sgRNA_R	CTCACCCCCTCTCTTTGTCCCAAACATTTAGATTTGCAATTC
smo-1 sgRNA_R	TCCCTTTTCCACTGAGATACCAAACATTTAGATTTGCAATTC
vet-2 sgRNA_R	CCGGTATCCTATGATCCAACCAAACATTTAGATTTGCAATTC

Preparation of HR donor vectors

pie-1 donor plasmids (point mutations and gfp and flag-fusions): pie-1 genomic sequence (LGIII:12,425,767-12,428,049) was amplified using the primers C_PIE-1 PF and C-PIE-1 PR and the resulting PCR product was inserted into the pCR^{tm}-Blunt II-TOPO® vector (Invitrogen, cat. no. K2800-20).

The K68A and K68R mutations were introduced by PCR sewing (or overlap extension PCR). The *pie-1* plasmid described above was used as a template to generate overlapping PCR products with the corresponding site-specific mutations. The overlapping PCR products were mixed together (1:1), diluted 50-fold with water, and used as a template in the PCR-sewing step with an external primer pair. The fused PCR products were gel purified and cloned into the pCR^{tm}-Blunt II-TOPO® vector.

For building *gfp::pie-1* donor constructs, an *Nhe*I restriction site was inserted immediately after and in frame with the start codon of *pie-1* by PCR sewing. A plasmid containing wild-type or mutant *pie-1* sequence was used as a template to generate a left arm PCR product flanked by *Bsi*WI and *Nhe*I restriction sites and a right arm PCR product flanked by *Nhe*I and *Ngo*MIV restriction sites. The products were digested with *Nhe*I, purified using a PCR cleanup kit, and ligated together. The ligated products were cloned into the pCR^{tm}-Blunt II-TOPO® vector, and plasmids containing the appropriately ligated fragments were identified. A *Bsi*WI and *Ngo*MIV fragment, containing the in frame *Nhe*I site immediately after the start codon, was released and ligated to

similarly digested *pie-1* constructs. The GFP coding region amplified from pPD95.75 (Addgene) was inserted into the *Nhe*I site.

For *pie-1::gfp* or *pie-1::flag,* a 1.6 kb fragment (LGIII 12,428,172-12,429,798) was amplified from genomic DNA and inserted into the pCR^{tm}-Blunt II-TOPO® vector (Invitrogen, cat. no. K2800-20). Overlap extension PCR was used to introduce an *Nhe*I site immediately before the stop codon in this fragment of *pie-1*. A *3×flag* sequence (gattacaaagaccatgatggtgactataaggatcatgatattgactataaagacgatgacgataag) was inserted into the *Nhe*I site.

Finally, we used PCR sewing to introduce silent mutations that disrupt the PAM site (NGG to NTG) in each HR donor. The above plasmids were used as templates to generate the initial PCR products for PCR sewing. The final products were cloned into the pCR^{tm}-Blunt II-TOPO® vector. We used iProof^{TM} high-fidelity DNA polymerase (Bio-Rad, cat. no. 172-5300) in all PCR reactions above. Primers are listed in Table 1.3.

Table 1.3 Summary of primer sequences for repair template and PCR screening

Name	Sequence
C_PIE-1 PF	ATAGCCCGATTTTGGAGGTG
C_PIE-1 PR	CCTCGAATTTTGGCAATTTTTC
C_PIE-1 301L	ATGGATTTCTCGCCGTTTTTTC
C_PIE-1 318R	GTTGTATCCACGTCGTCTCG
C_PIE-1(K68A)_F	GGAAAATGGCTTCGTCCAGCACGTGAAGCG
C_PIE-1(K68A)_R	CTTGAGCGCTTCACGTGCTGGACGAAGCC
C_PIE-1(K68R)_F	GGAAAATGGCTTCGTCCTAGGCGTGAAGCG
C_PIE-1(K68R)_R	CTTGAGCGCTTCACGCCTAGGACGAAGCC
C_PIE-1 a MF	GCTATGTCTTTTAGTTGCAGGCGCCTC
C_PIE-1 a MR	CAGATTGACGAGGCGCCTGCAACTAA
SMO-1 PF	CGATTTTTCGGCTCATTTCG
SMO-1 PR	CCTCGTCAAATCCGAAATCG
SMO MF	CACCCATCAATCCCTTTTC
SMO MR	GAAAAGGGATTGATGGGTG
P1F	GTTTTTGCCCCCAAATTC
P1R	TGATGCTTCGATGCTGAAGA
P2F	GGCGTCAAAAGACATATGTAAAAG
P2R	CGCAATGGATGATTTTTGTC
P3F	GCCGAGCTATGTCTTTTAG
P3R	CTCAAGATCACTCCATTGGC
P4F	GGCGGTGCGTTTGAAGTGT
P4R	GGAAATAATAGTTGGTGGTGGC
P5F	CCATATTTTGTTTTGTATATTTATC
P5R	GGCACAAGTTCATTCACAGG
S1F	GAAGTGCACTTCCGTGTAAAGTATGGAACC
S1R	CCGGCTGCTATTTCATTGAT
mc.out.F1	GCTCAAGAAAGCCAATGGAG
mc.out R1	TTCTGAACCAGTCGATGCAG
mc.in F1	ATGGAGGGATCTGTCAATGG
mc.in R1	TGGCAGTCGAGACACTTCAG

mCherry::vet-2 and flag::vet-2 donor construct: A 2411 bp DNA fragment of the *vet-2* gene, including 1249 bp of sequence upstream and 1162 bp downstream of the *vet-2* start codon (corresponding to the genomic sequence LGI:10,845,543-10,847,953), was amplified from genomic DNA and inserted into pBluescript KS (+) vector (Addgene). An *Xma*I site was introduced by PCR immediately after the *vet-2* start codon. The *mCherry* coding sequence amplified from pCFJ90 (Addgene) or 3×*flag* sequence was inserted into the *Xma*I site.

smo-1::flag donor plasmid: smo-1 genomic sequence (LGI: 1,340,243-1,341,558) was amplified from N2 genomic DNA and inserted into the pCR^{tm}-Blunt II-TOPO® vector (Invitrogen, cat. no. K2800-20). Overlap extension PCR was used to introduce an *Nhe*I site immediately before the stop codon in this fragment of *smo-1*. The resulting PCR product was cloned into the pCR^{tm}-Blunt II-TOPO® vector. A 3×*flag* fragment with *Nhe*I overhangs was generated by annealing two overlapping oligonucleotides and ligated into the *smo-1* donor plasmid. We mutated the PAM site as described for the *pie-1* donors above.

gfp::pie-1 for MosSCI: A 3744 bp fragment (*Sca*I-*Not*I) containing the *pie-1* promoter was excised from pID3.01B (Addgene) and inserted into a modified MosSCI LGII vector (B1496) in which a *Not*I site was added to pCFJ151 (Frokjaer-Jensen et al., 2008; Shirayama et al., 2012). A 2631 bp PCR fragment containing the *pie-1* open reading frame (ORF) and 3' *UTR* was then inserted

into the resulting plasmid to make a *gfp::pie-1* plasmid for MosSCI. The plasmid was injected into the strain EG4322 at a concentration of 10 ng/ml by direct injection method to insert a single copy *gfp::pie-1* transgene on chromosome II at position 8,420,159.

BSD-fusion to pie-1: The nucleotide sequence of the Blasticidin S resistance gene (*BSD*) from *Aspergillus terreus* was codon optimized for *C. elegans* and an artificial intron (gtaagagatttttttaaaaatttatttttttacactgttttttctcag) was inserted into the middle of the *BSD* ORF: the entire gene was de novo synthesized by GenScript. The *BSD* fragment containing the *BSD* ORF (439 bp), *rpl-28* promoter (568 bp) and *rpl-28* 3'*utr* (568 bp) was inserted into pBluescript KS (+) vector (Addgene). The complete sequence of *BSD* marker is available upon request. A 1077 bp fragment of *pie-1* left homology was inserted into the *Xba*I site before the *rpl-28* promoter and a 1017 bp fragment of *pie-1* right homology was inserted into the *Sal*I-*Apa*I site after the *rpl-28* 3'*utr*. Blasticidin S (AG scientific, cat. no. B-1247) was used to select animals transformed with the *BSD* gene.

Preperation of heat shock-Cas9 plasmid

The *Mos1* transposase ORF in pJL44 (Addgene) was replaced with Cas9 from *Peft3::Cas9* vector (Friedland et al., 2013) to generate *hs::Cas9* (pWU34) construct.

Microinjection

DNA mixtures were microinjected into the gonads of young adult worms. Plasmids for injection were prepared using a midiprep plasmid purification kit (Qiagen, cat. no. 12143). For Co-CRISPR, we injected 50 ng/µl each vectors (Cas9 vector, *unc-22* sgRNA vector (Co-CRISPR), two untested-sgRNAs, and pRF4::*rol-6(su1006)*). Microinjection mixtures for HR contained 50 ng/ul each Cas9 vector, sgRNA vector, pRF4::*rol-6(su1006)*, and HR donor vector. The final concentration of DNA in the injection mix did not exceed 200 ng/µl. For injection mixes with 5 different plasmids, 40 ng/µl of each plasmid was added. For HR experiments, we injected about 40 to 60 worms, and for disruptions about 20 to 30 worms. After recovering from injection, each worm was placed onto an individual plate.

Screening for indels using Co-CRISPR

In order to validate untested sgRNAs we injected mixtures containing the *unc-22* sgRNA and up to several untested sgRNAs (as described above). Three days after injection, F1 rollers and F1 twitchers were picked individually to plates and allowed to produce F2 progeny for 2 to 3 days. F1 twitchers and F1 rollers with twitching F2 progeny were then transferred to 20 µl lysis buffer for PCR, polyacrylamide gel electrophoresis (PAGE) (see below) and/or DNA sequencing analysis. Total time from injection to indel detection was about 6 to 7 days.

Screening for HR events

Direct detection of GFP: This procedure works for donor vectors that cannot drive GFP without first integrating into the genomic target site. For GFP::PIE-1 it was necessary to mount gravid F1 rollers three at a time under cover slips on 2% agarose pads for screening at 40X magnification using a Zeiss Axioplan2 microscope. For bright GFP constructs, it should be possible to screen using a fluorescence dissecting scope. GFP-positive animals were recovered by carefully removing the cover slip and transferring to individual plates. After laying eggs for 1 day they were individually lysed in 20 µl lysis buffer for PCR confirmation of the GFP insertion. GFP-positive F2 homozygotes were then identified and correct insertion of GFP was confirmed by DNA sequencing. Total time from injection to recovery of heterozygotes was 3 to 4 days.

PCR detection: F1 rollers were picked individually to plates and allowed to lay eggs for 1 day. For the co-CRISPR assay, F1 rollers were allowed to produce F2 progeny for 2 to 3 days so that F2 twitcher progeny could be identified. (Note: F1 roller animals that segregate twitching progeny should be selected as these animals exhibit the highest HR frequency, while non-rolling F1 twitchers should be avoided [see Results and Discussion]). F1 animals were then transferred into lysis buffer in indexed PCR tubes and were screened using primers outside the homology arms followed by restriction enzyme digestion to detect the insertion. In some experiments, 1 µl of the initial PCR reaction was used as a template for

a second PCR reaction with primers within the donor sequences. Though useful, this latter procedure gave several false positives in our hands. Total time from injection to recovery of heterozygotes was 4 days. For the Co-CRISPR strategy, 3 more days were required to recover heterozygotes.

Selection/Counter-selection method: Four days after injection, gravid F1 rolling adults were placed in groups of 10 to 15 animals per plate onto media containing ivermectin and blasticidin. After 3 to 4 more days, the plates were scored for viable, fertile progeny. Insertion of *BSD* at the target locus was then confirmed by PCR and DNA sequencing (as described above). The total time from injection to recovery of HR events was 7 to 10 days. Though slightly longer in duration this procedure required approximately ten times less labor as only the relatively rare viable populations were subject to PCR and DNA sequence analysis. Primers for screening HR events are listed in Table 1.3.

Imaging

Images were captured with an ORCA-ER digital camera (Hamamatsu) and AxioVision (Zeiss) software.

Screening for small indels by PCR and PAGE

We designed primers to amplify (~30 cycles) PCR products of 60-65 bp encompassing the CRISPR-Cas9 target site. PCR products were resolved on

15% polyacrylamide gels to distinguish dsDNA molecules that differ by as little as 1 bp. We found that we could screen for indels even in HR experiments, but it required two PCR steps. In the first PCR reaction (~20 cycles), primers outside of the homology arms were used to avoid amplifying the donor sequence. In the second reaction (~15 cycles), 1 µl of the first PCR reaction was used as a template to generate the 60-65 bp PCR product encompassing the CRISPR-Cas9 target site. *TaKaRa Ex Taq*^{TM} (Takara, cat. no. RR001) was used for the PCR reactions above. Primer sequences are listed in Table 1.3.

Immunoblotting

One hundred adult worms were lysed in 80 µl of 1X sample buffer (25 µl of M9 containing 100 worms, 25 µl of 2X lysis buffer, 20 µl of 4X NuPAGE® LDS Sample buffer (Invitrogen, cat. no. NP0008), and 10 µl of β-mercaptoethanol by boiling for 20 min, freezing, and boiling again for 10 min. The worm lysate proteins were separated on 4-12% NuPAGE® Tris-Acetate Mini Gels (Invitrogen, cat. no. NP0335BOX). Proteins were transferred to Immun-Blot® PVDF Membrane (Bio-Rad, cat. no.162-0177) at 100 V for 1 hr at 4°C. Mouse monoclonal anti-PIE-1 antibody (P4G5) (Mello et al., 1996) and rabbit polyclonal anti-PGL-1 antibody was used at 1:50 and 1:500, respectively.

Results

Using a visible co-transformation marker enriches for genome-editing events

While conducting CRISPR-Cas9 experiments to induce mutations in the *pie-1* gene, we used the dominant co-injection marker *rol-6* to monitor injection quality. From 60 injected animals, we obtained 93 fertile F1 rollers. Remarkably, we noted that several of these F1 rollers (5/93) produced 100% dead embryos exhibiting the distinctive *pie-1* maternal-effect embryonic lethal phenotype (Mello et al., 1996) (Figure 1.6A). Genomic sequencing of these F1 adults identified lesions in the *pie-1* gene consistent with Cas9-directed cleavage (Figure 1.6B). In some cases the maternal and paternal alleles exhibited distinct lesions, while in others, the same lesion was found in both alleles (Figure 1.6B). Since the DNA was delivered into the ovary of an adult, after the switch from sperm to oocyte development, the paternal allele must have been targeted in the F1 zygote soon after fertilization. The fact that both alleles exhibit identical lesions in some animals suggests that a chromosome previously cut and repaired by NHEJ was used as a donor molecule to copy the lesion into the homolog.

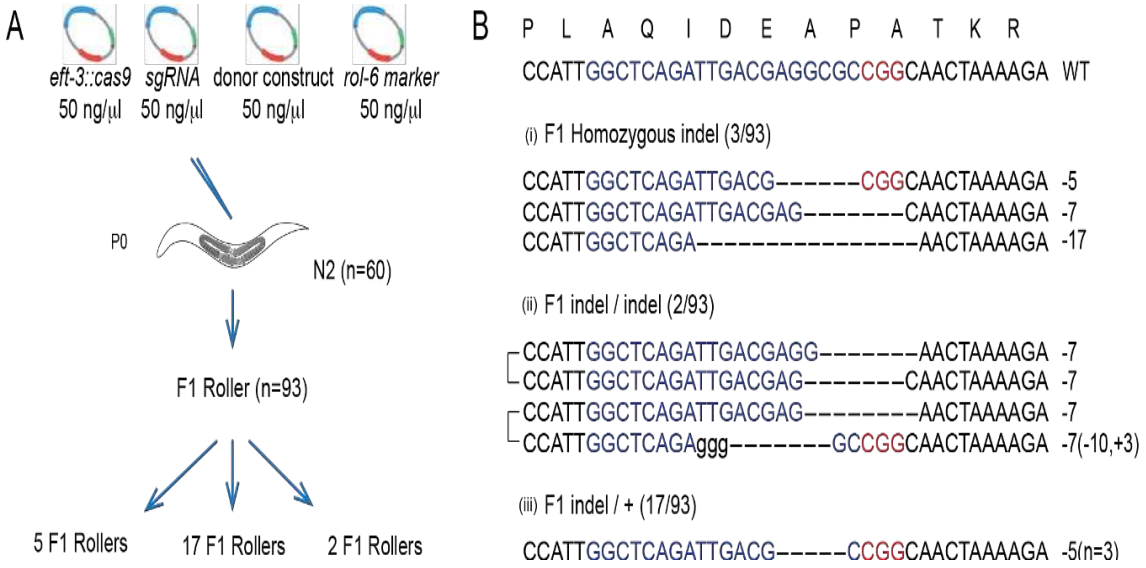


Figure 1.6 Efficient CRISPR-Cas9-mediated gene disruptions in transgenic animals (A) Schematic representation of screen for CRISPR-Cas9 genome editing events. The dominant transformation marker *rol-6* was co-injected with Cas9, *pie-1* a sgRNA, and donor plasmids. F1 rollers were screened for NHEJ-mediated indels by DNA sequencing. Among 93 F1 rollers, 22 indels were obtained. (B) Sequences of the wild-type *pie-1* target site (top) and CRISPR-Cas9-mediated indels among F1 animals: (i) *pie-1* homozygotes carrying the same indel on both alleles; (ii) *pie-1* homozygotes carrying a different indel on each allele; and (iii) *pie-1* heterozygotes. Lower case indicates insertion, and dash indicates deletion. The PAM is marked in red, and target sequences are marked in blue. The number of deleted (-) or inserted (+) bases is indicated to the right of each indel. The numbers in parentheses in (iii) represent the number of animals with the indels shown.

If the activation of Cas9 in the germline is broadly or non-specifically mutagenic, then some injected animals would be expected to segregate novel mutants, including mutants with non-*pie-1* dead-embryo phenotypes. To look for evidence of off-target mutagenesis, we screened among the progeny of F1 rollers for animals producing dead embryos, or other visible phenotypes. A careful examination of F2 and F3 populations revealed 17 populations from 93 F1 rollers that segregated numerous dead embryos (Figure 1.6A). However, examination of these dead embryos by Nomarski microscopy revealed the distinctive *pie-1* mutant phenotype, and no other phenotypes. Each of these 17 strains segregated *pie-1* homozygotes at the expected Mendelian frequency, indicating that the original F1 rollers were heterozygous for *pie-1* loss-of-function mutations. Sequencing of these strains revealed indels in the region of the *pie-1* gene targeted by CRISPR-Cas9 (Figure 1.6B). In some cases, to avoid the delay and added cost of DNA sequencing, genomic DNA prepared from lysates of each candidate was amplified as ~60 bp PCR products that were then analyzed on a 15% PAGE gel. This analysis easily detected lesions as small as 5 nt (Figure 1.6B and Figure 1.7A).

In addition to F1 rollers, we randomly selected F1 non-roller sibling progeny that were produced during the same time window as the F1 rollers. Among 100 non-roller siblings, we failed to find animals segregating dead embryos. Thus, using the dominant visible *rol-6* marker to identify F1 transgenic animals (rollers) also identified animals in which Cas9 was active. It is important

to note, however, that very few of the animals with *pie-1* mutations continued to exhibit the roller phenotype in subsequent generations, suggesting that the *rol-6* transgene expression was transient and present only in the F1 generation.

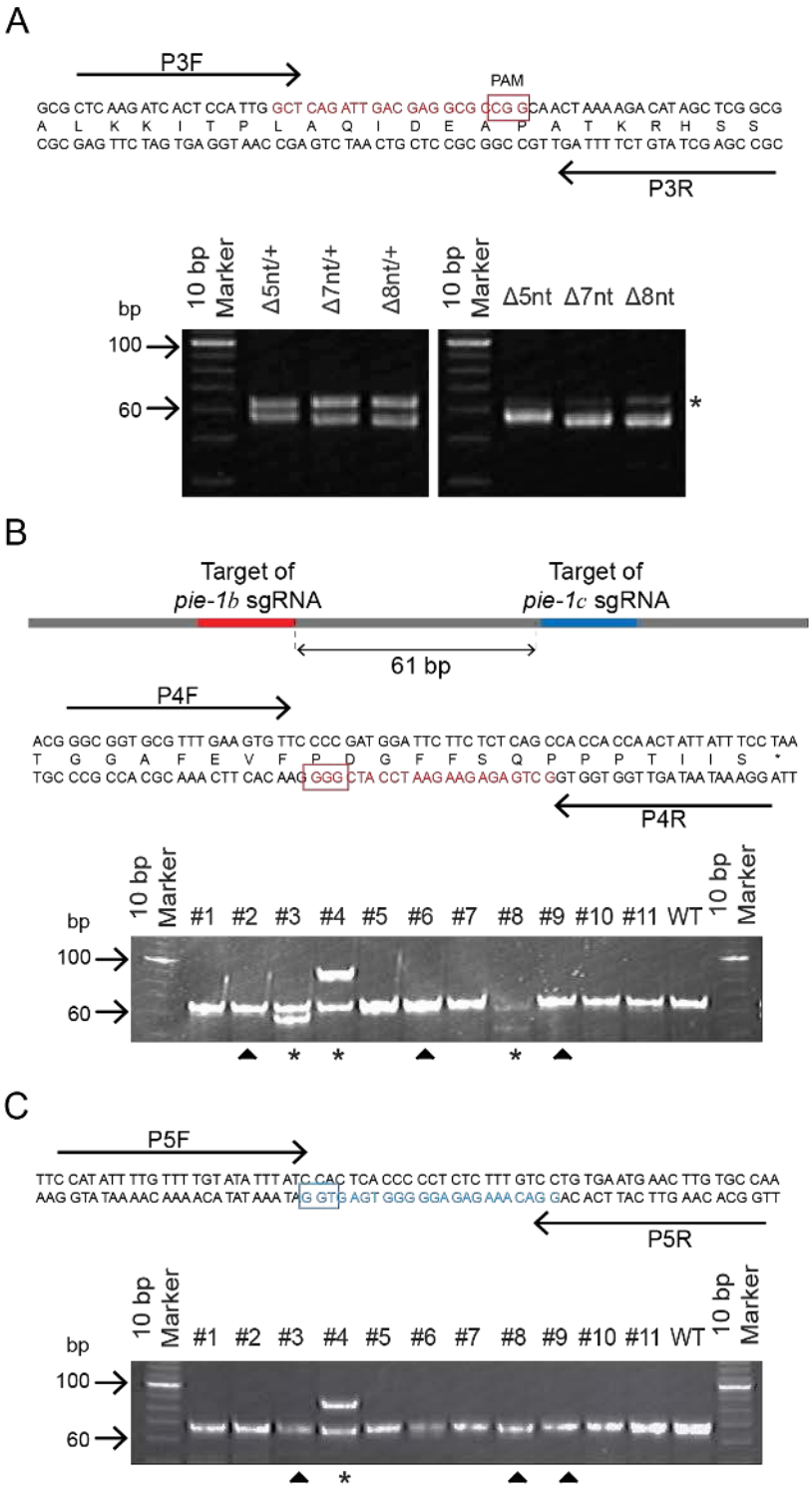

A
P3F
PAM
GCG CTC AAG ATC ACT CCA TTG GCT CAG ATT GAC GAG GCG CCG GCA ACT AAA AGA CAT AGC TCG GCG
A L K K I T P L A Q I D E A P A T K R H S S
CGC GAG TTC TAG TGA GGT AAC CGA GTC TAA CTG CTC CGC GGC CGT TGA TTT TCT GTA TCG AGC CGC
P3R
bp
100
60
10 bp Marker
Δ5nt/+
Δ7nt/+
Δ8nt/+
10 bp Marker
Δ5nt Δ7nt Δ8nt
*
B
Target of pie-1b sgRNA
Target of pie-1c sgRNA
61 bp
P4F
ACG GGC GGT GCG TTT GAA GTG TTC CCC GAT GGA TTC TTC TCT CAG CCA CCA CCA ACT ATT ATT TCC TAA
T G G A F E V F P D G F F S Q P P P P T I I S *
TGC CCG CCA CGC AAA CTT CAC AAG GGG CTA CCT AAG AAG AGA GTC GGT GGT GGT TGA TAA TAA AGG ATT
P4R
bp
100
60
10 bp Marker
#1 #2 #3 #4 #5 #6 #7 #8 #9 #10 #11 WT
10 bp Marker
C
P5F
TTC CAT ATT TTG TTT TGT ATA TTT ATC CAC TCA CCC CCT CTC TTT GTC CTG TGA ATG AAC TTG TGC CAA
AAG GTA TAA AAC AAA ACA TAT AAA TAG GTG AGT GGG GGA GAG AAA CAG GAC ACT TAC TTG AAC ACG GTT
P5R
bp
100
60
10 bp Marker
#1 #2 #3 #4 #5 #6 #7 #8 #9 #10 #11 WT
10 bp Marker

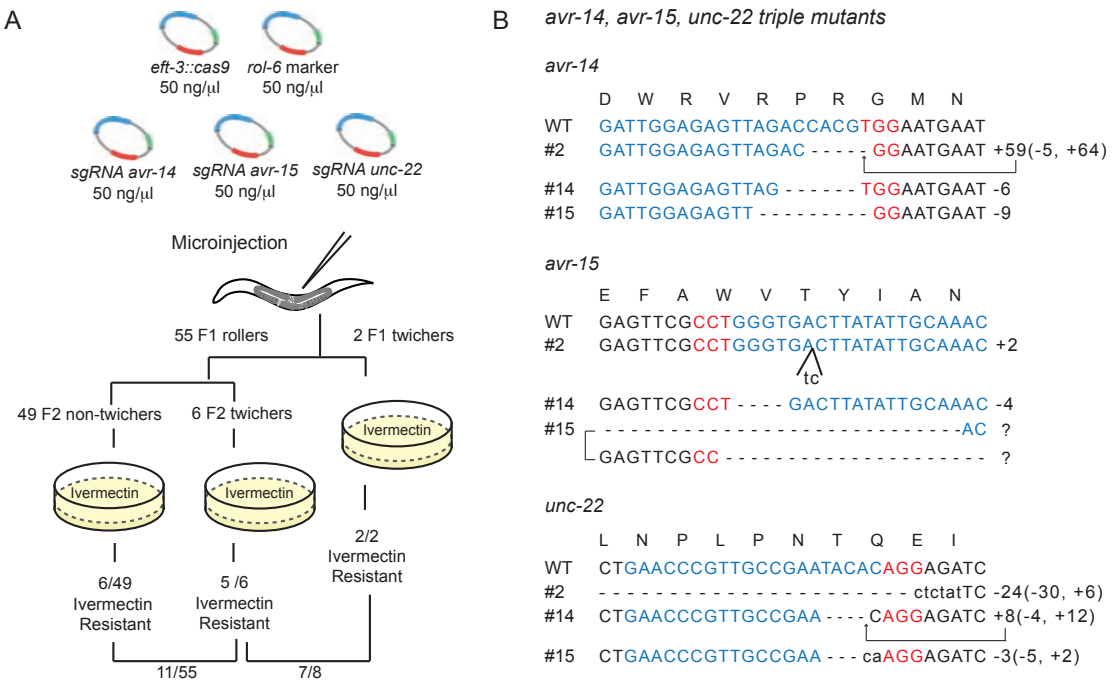
Figure 1.7 Detecting small indels on 15% polyacrylamide gels
(A) The indicated primers (arrows) were used to amplify sequences immediately surrounding the CRISPR-Cas9 target site (red). The indels in this experiment were from an HR experiment, so an initial PCR was performed using primers outside of the homology arms of the donor template (Figure 1.10A). The initial PCR was used as a template to amplify the target site using the indicated primers. PCR products from F1 heterozygotes (left) and F2 homozygotes (right) were separated on a non-denaturing 15% polyacrylamide gel and stained with ethidium bromide. The asterisk indicates the PCR product amplified from residual donor plasmids in the single worm lysate (B) and (C) Test of two uncharacterized *pie-1* sgRNAs using the Co-CRISPR strategy and PAGE analysis. The *pie-1* sgRNA vectors were combined and co-injected with the *unc-22* sgRNA, Cas9, and *rol-6* plasmids. The *pie-1* sgRNA target sites (shown in red and blue) are separated by 61 bp. As this experiment did not include an HR donor, only a single round of PCR was performed with the indicated primers (arrows). We lysed 11 F1 animals with the twitching phenotype (#3, #8, and #9-11) or that produced twitching progeny (#1-2 and #4-7). WT, wild-type N2 genomic DNA was used as a template. Asterisks indicate lanes in which small indels were detected. The filled triangles indicate lanes in which the primer pair could not detect the indels.

A co-CRISPR strategy to detect genome-editing events

In practice we have found that about half of sgRNAs tested are not effective. Thus, while the *rol-6* marker was clearly useful for finding animals with CRISPR-Cas9-induced lesions, we nevertheless frequently had to screen through dozens or even hundreds of F1 rollers by PCR or sequencing only to conclude that CRISPR-Cas9 was not active in the injection. We therefore reasoned that co-injecting a proven sgRNA (one that works well and results in an easily recognized visible phenotype) would allow us to more directly identify animals in which Cas9 is active. We tested this strategy using an sgRNA targeting the muscle structural gene *unc-22* (Benian et al., 1993; Moerman and Baillie, 1979). We chose this sgRNA both because *unc-22* loss of function causes a distinctive recessive paralyzed twitching phenotype that is easy to score and also because this sgRNA works moderately well compared to other sgRNAs (Table 1.1). Thus, F1 and F2 *unc-22* twitchers should arise from animals exposed to the greatest levels of Cas9 activity, and should therefore also have active Cas9 loaded with the co-injected sgRNAs.

To test the co-CRISPR strategy, we co-injected the *unc-22* sgRNA with two previously validated sgRNAs targeting *avr-14* and *avr-15* (Figure 1.8A), two genes whose wild-type activities redundantly confer sensitivity to the potent nematicide ivermectin (Dent et al., 2000). The *rol-6* marker was also included in these injections to facilitate the identification of twitchers that arise in the F2 among the progeny of F1 roller animals. We then measured, among 55 F1

rollers, the frequency of ivermectin resistant strains (20%, n=11) and twitcher strains (11%, n=6) (Figure 1.8A). Strikingly, selecting for the twitching phenotype dramatically enriched for animals exhibiting ivermectin resistance. For example, among 8 F1 animals that were either twitching themselves or produced twitcher progeny, 7 (88%) also produced progeny resistant to ivermectin (Figure 1.8A and 1.8C). We confirmed co-CRISPR activity by sequencing the lesions in several of these strains (Figure 1.8B).



C

Gene	Number Injected	F1 rollers producing F2 twitchers/Total F1 rollers	F1 having both rolling and twitching phynotype	F1 twitchers	Roller-based indel frequency	Twitcher-based indel frequency
avr-14 avr-15	20	6/55	0	2	11/55 (20%)	7/8 (88%)
pie-1 b *pie-1* c	30	2/31	1	8	2/32 (6%) 1/32 (3%)	6/11(55%) 4/11(36%)

Figure 1.8 "*unc-22*" Co-CRISPR as a marker to indicate actively expressed Cas9

(A) Schematic of Co-CRISPR strategy to identify functional sgRNAs targeting avr genes. sgRNAs targeting *avr-14* and *avr-15* were co-injected with a functional *unc-22* sgRNA, the Cas9 expression vector and the *rol-6* transformation marker. F1 rollers or twitchers were transferred to individual plates. The plates were allowed to starve, and then they were copied to plates containing 2 ng/ml ivermectin to identify CRISPR-Cas9-induced *avr-14; avr-15* double mutants. (B) Indel sequences in *avr-14; unc-22; avr-15* triple mutants. *avr-15* isolate #15 carried different indels on each allele. Sequences labeled with a question mark could not be precisely determined. (C) Comparison of twitcher-based indel frequency and roller-based indel frequency.

Similar results were obtained in several additional Co-CRISPR experiments (Figure 1.8C and data not shown). For example, we used this approach to test two uncharacterized sgRNAs targeting the 3' end of *pie-1* (Figure 1.8C and Figure 1.7B). Among 11 twitcher lines identified in the F1 or F2, we identified 3 indels by PCR and PAGE for one of the two sgRNAs (Figure 1.7B) and a single indel for the other sgRNA (Figure 1.7C). Sequence analysis confirmed these indels, which included a 6 nt deletion, a 24 nt insertion, and an 11 nt deletion for one sgRNA and a 16 nt insertion for the other. However, the PAGE detection method clearly underestimated the frequency of indels. Sequence analysis identified three heterozygous deletion mutations of 42 nt, 43 nt, and 603 nt that deleted primer binding sites and were thus too large to be detected by our PCR and PAGE analysis (Figure 1.7B). These unusually large deletions may reflect simultaneous cutting induced by the two adjacent sgRNAs whose targets are separated by 61 bp in this experiment (Figure 1.7B). In conclusion, these findings suggest that PAGE analysis of 10 to 20 Co-CRISPR lines should be sufficient to determine if an uncharacterized sgRNA is active. It should be noted that since many F1 rollers analyzed were heterozygous for *unc-22* lesions, it was usually possible to find non-uncs with the desired indel using the *unc-22* Co-CRISPR assay. However, if *unc-22* is inconvenient for a particular experiment our findings suggest that alternative Co-CRISPR sgRNAs targeting, for example, genes that when mutated confer resistance to ivermectin or benomyl or other genes with visible mutant phenotypes may be substituted.

The observation that using nearby sgRNAs can induce deletions that remove the intervening sequence is consistent with previous findings where large deletions were produced in this way (Horii et al., 2013; Ran et al., 2013; Ren et al., 2013; Wang et al., 2013; Zhou et al., 2014). Thus the co-CRISPR strategy should facilitate the identification of deletions that remove the interval between two sgRNAs. However, further experimentation will be required to determine how large an interval can readily be eliminated in this way. For the purpose of validating sgRNAs our findings suggest that large deletions produced by testing multiple nearby sgRNAs simultaneously may confound the analysis of which sgRNAs are active. On the other hand, pooling sgRNAs targeting a number of different genes that are distant from one another in the genome should, in principle, allow several sgRNAs to be validated in a single Co-CRISPR microinjection experiment.

Identification of HR events without a co-selectable marker

We next sought to use CRISPR-induced double-strand breaks to drive homologous recombination (HR). Several types of editing are possible, ranging from changing a single amino acid to inserting a protein tag such as GFP, or even deleting the entire target gene. In designing donor molecules to introduce point mutations or epitope fusions, we found it necessary to alter the sgRNA target sequence in the donor by mutating the protospacer adjacent motif (PAM) site, or by introducing mismatches within the seed region (Cong et al., 2013;

Jiang et al., 2013; Jinek et al., 2012; Sternberg et al., 2014). In our experience, failure to take this step frequently results in HR events containing CRISPR-Cas9-induced indels or a very low frequency of HR events: sometimes 0% (data not shown).

Previous studies successfully used single-stranded oligonucleotide donor molecules (Zhao et al., 2014) or double-stranded plasmid donor molecules (Dickinson et al., 2013) to induce HR events in *C. elegans.* However these studies relied on screening for a selectable phenotype introduced by the HR event. Given the high frequencies of NHEJ events detected in the studies above, we wondered if CRISPR-Cas9-mediated HR events could be recovered directly without the need for selection.

To test this idea, we decided to use CRISPR-Cas9-mediated HR to introduce the *gfp* coding sequence immediately downstream of the start codon in the endogenous *pie-1* locus (Figure 1.9A). The donor plasmid in this experiment contained *Nhe*I restriction sites flanking the *gfp* coding sequence, 1 kb homology arms, and a silent mutation that disrupts the PAM sequence at the sgRNA target site (Figure 1.9A). We generated three different donor constructs–*gfp::pie-1(WT)*, *gfp::pie-1(K68A)* and *gfp::pie-1(K68R)*. Each donor molecule was co-injected with vectors to express the sgRNA, Cas9, and *rol-6* marker. We then directly examined the resulting F1 rolling animals for GFP::PIE-1 expression in the germline and embryos using epifluorescence microscopy (Figure 1.9B, see materials and methods).

Using this approach, we obtained 9 independent lines out of 92 F1 rollers for *gfp::pie-1(K68A*), 1 out of 69 for *gfp::pie-1(K68R)*, and 1 out of 72 for *gfp::pie-1(WT)*. Subsequent analyses revealed that each of these F1 animals was heterozygous for *gfp::pie-1*, and each strain incorporated both the *gfp* coding sequence and the PAM site mutation, as well as the linked K68A and K68R missense mutations. For unknown reasons, we found that one of the nine *gfp::pie-1(K68A)* lines could not be maintained.

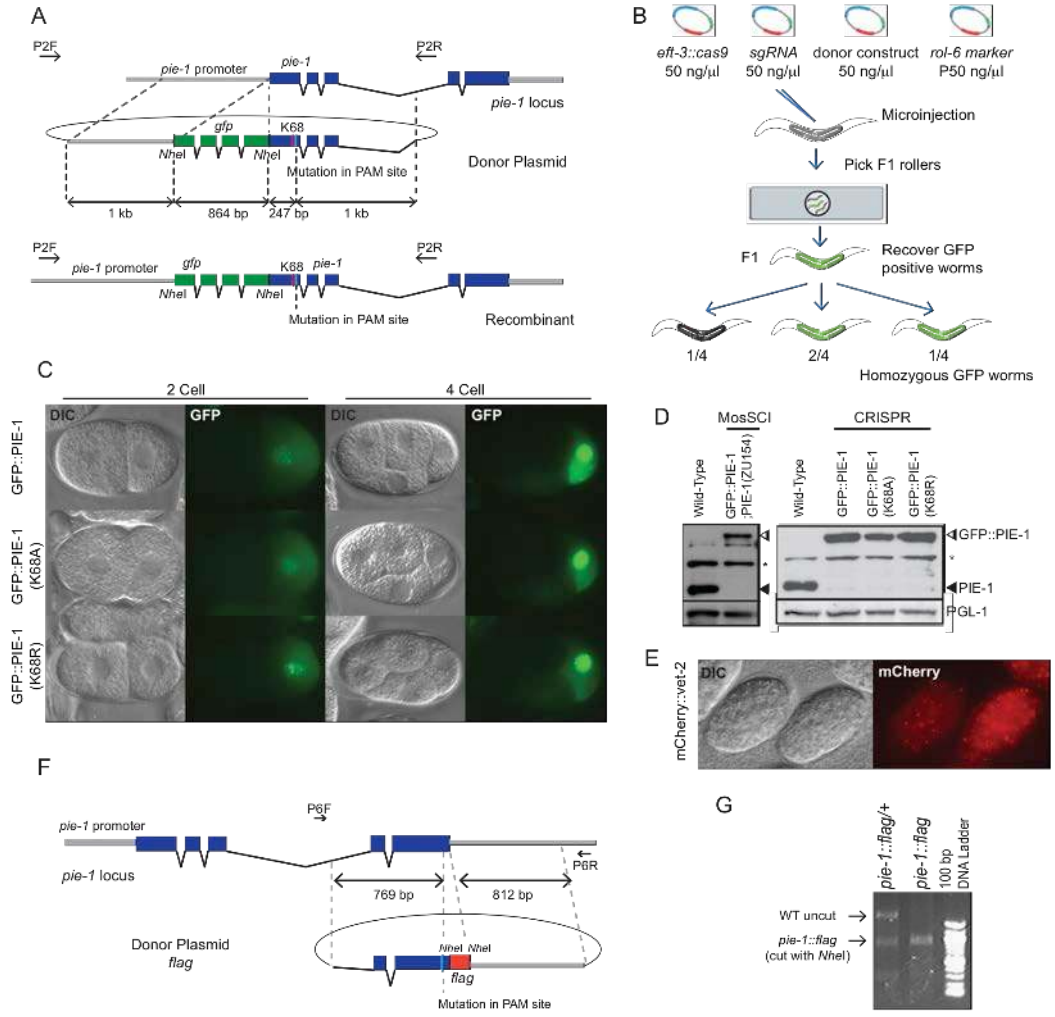

A
P2F
P2R
pie-1 promoter
pie-1
pie-1 locus
gfp
K68
NheI
NheI
Mutation in PAM site
Donor Plasmid
1 kb
864 bp
247 bp
1 kb
P2F
pie-1 promoter
gfp
K68
pie-1
P2R
NheI
NheI
Mutation in PAM site
Recombinant
B
eft-3::cas9
50 ng/µl
sgRNA
50 ng/µl
donor construct
50 ng/µl
rol-6 marker
P50 ng/µl
Microinjection
Pick F1 rollers
F1
Recover GFP
positive worms
1/4
2/4
1/4
Homozygous GFP worms
C
2 Cell
4 Cell
DIC
GFP
DIC
GFP
GFP::PIE-1
GFP::PIE-1 (K68A)
GFP::PIE-1 (K68R)
D
MosSCI
CRISPR
Wild-Type
GFP::PIE-1 PIE-1(ZU154)
Wild-Type
GFP::PIE-1
GFP::PIE-1 (K68A)
GFP::PIE-1 (K68R)
GFP::PIE-1
PIE-1
PGL-1
E
mCherry::vet-2
DIC
mCherry
F
pie-1 promoter
P6F
pie-1 locus
P6R
769 bp
812 bp
Donor Plasmid
flag
NheI NheI
flag
Mutation in PAM site
G
pie-1::flag/+
pie-1::flag
100 bp DNA Ladder
WT uncut
pie-1::flag
(cut with NheI)

Figure 1.9 HR-mediated knock-in to generate fusion genes at endogenous loci
(A) Schematic of the Cas9/sgRNA target site and the donor plasmid for *gfp::pie-1* knock-ins. The donor plasmid contains the gfp coding sequence inserted immediately after the start codon of *pie-1*, 1 kb of homology flanking the CRISPR-Cas9 cleavage site, and a silent mutation in the PAM site. (B) Strategy to screen for *gfp* knock-in lines. We placed 3 F1 rollers at a time on a 2% agar pad and screened for GFP expression using epifluorescence microscopy. GFP-expressing worms were individually recovered and allowed to make F2 progeny for one day before being lysed for PCR and DNA sequence analysis. We confirmed Mendelian inheritance of *gfp* knock-in alleles among F2 progeny. (C) GFP::PIE-1 expression in the germline of 2- to 4-cell embryos of *gfp::pie-1* knock-in strains. (D) Immunoblot analysis showing PIE-1 expression levels in wild-type animals, MosSCI-mediated *gfp::pie-1* knock-in animals, and CRISPR-Cas9-mediated *gfp* knock-in animals. A MosSCI strain of *gfp::pie-1; pie-1(zu154)* was obtained by crossing *gfp::pie-1* (LGII) with the *pie-1(zu154)* (LGIII) null mutant. (E) mCherry expression in late embryos of the *mCherry::vet-2* knock-in strain. (F) Schematic of Cas9/sgRNA target site, PAM site, and donor plasmid for *pie-1::flag* knock-in. The PAM is located in the last exon of *pie-1*. The donor plasmid includes flag coding sequence immediately before the *pie-1* stop codon and ~800 bp homology arms flanking the target site. (G) PCR and restriction analysis of an HR event. PCR products were generated using the primers indicated in (F), and the products were digested with *Nhe*I. The *pie-1::flag* gene conversion introduces an *Nhe*I RFLP that is observed in F1 heterozygous and F2 homozygous *pie-1::flag* animals.

The high rates of HR observed in the above study, suggested that it should also be possible to recover HR events by screening DNA isolated from F1 rollers using PCR. To test this idea, we designed donor molecules to insert the *pie-1* lysine 68 lesions described above with either no tag sequences or with sequences encoding the FLAG epitope immediately before the stop codon of the *pie-1* gene (Figure 1.9F and Figure 1.10A). For these HR experiments, we used 300 bp (no tag) and 800 bp (*flag* tag) flanking homology arms along with previously tested sgRNAs (Figure 1.9F, Figure 1.7B, and Figure 1.10A). We then used PCR to amplify the genomic DNA sequence from F1 rollers and restriction analysis to identify F1 heterozygotes carrying the desired insertion (Figure 1.9G, Figure 1.10B, and Figure 1.10C). These studies identified two K68A HR events among 93 F1 rollers (60 injected worms), and 3 *flag* HR events among 84 F1 rollers (40 injected worms) (data not shown). A similar PCR-detection strategy was used to introduce *mcherry* into the gene *vet-2.* In this experiment, mCherry expression was not visible in adult F1 rollers, but was easily detected among the F2 embryos produced by PCR-positive animals (Figure 1.9E). Taken together these findings show that CRISPR-Cas9-mediated HR occurs at a remarkably high frequency in *C. elegans*.

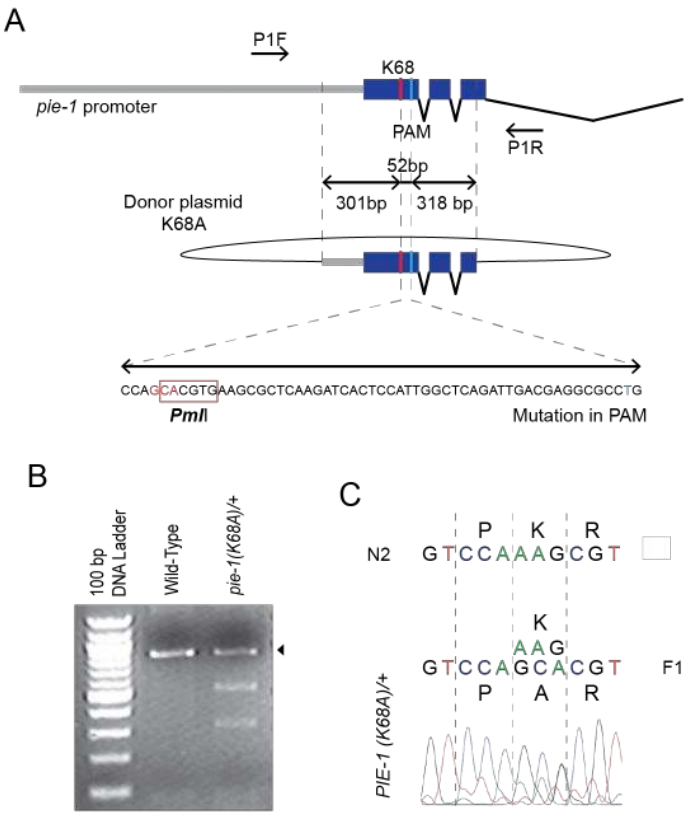


Figure 1.10 Site-specific mutagenesis of *pie-1* by HR

(A) Schematic of the Cas9/sgRNA target sites in *pie-1* locus and donor plasmids. The K68A donor plasmid contains ~300 bp of homology flanking the 52 bp target region between the K68 codon and PAM site and introduces a *Pml*I restriction site (red box). The PAM site of each donor was disrupted by silent mutations so that it will not be targeted by CRISPR-Cas9. The blue bar indicates the PAM site, and the red bar indicates the position of K68. (B) PCR and restriction enzyme analysis of wild-type control worms and F1 rollers from K68A CRISPR-Cas9-mediated HR experiments. PCR primers outside of the donor homology arms (P1F and P1R for K68A) are indicated in (A). Restriction analysis following PCR shows the RFLP in *pie-1(K68A)/+*. The wild-type product is indicated by the filled triangle. (C) DNA sequence analyses to confirm the desired point mutations. Note that the PCR products for sequencing were amplified using the primers outside of donor plasmid, as indicated in (A).

We compared the expression and localization of GFP::PIE-1 protein in our newly generated *gfp::pie-1* knock-in strains to strains in which *gfp::pie-1* was inserted at a heterologous site in the genome by MosSCI (Frokjaer-Jensen et al., 2008; Shirayama et al., 2012). The knock-in strains showed the expected localization of PIE-1 in 2- to 4-cell embryos (Figure 1.9C). Strikingly, immunoblot analysis using the PIE-1 monoclonal antibody (P4G5) revealed that GFP::PIE-1 protein was expressed at a much higher level in the CRISPR-Cas9-induced knock-in strains, similar to the expression level of endogenous PIE-1 protein (Figure 1.9D). These results are consistent with a previous study (Dickinson et al., 2013) and indicate, perhaps not surprisingly, that insertion of GFP into the endogenous locus can achieve near optimal expression levels of the tagged protein.

Co-CRISPR for identifying HR events

Most of the HR work described above was done before we realized the utility of co-CRISPR markers for validating sgRNAs. To determine if the co-CRISPR strategy could facilitate recovery of HR events, we co-injected *unc-22* sgRNAs along with CRISPR HR injection mixes targeting *vet-2*, *pie-1* and *smo-1* genes (Figure 1.11). The findings from these studies suggest that using a co-CRISPR marker can increase the frequency of HR events in the range of approximately 2- to 4-fold over those observed by first selecting F1 roller animals (Table 1.4).

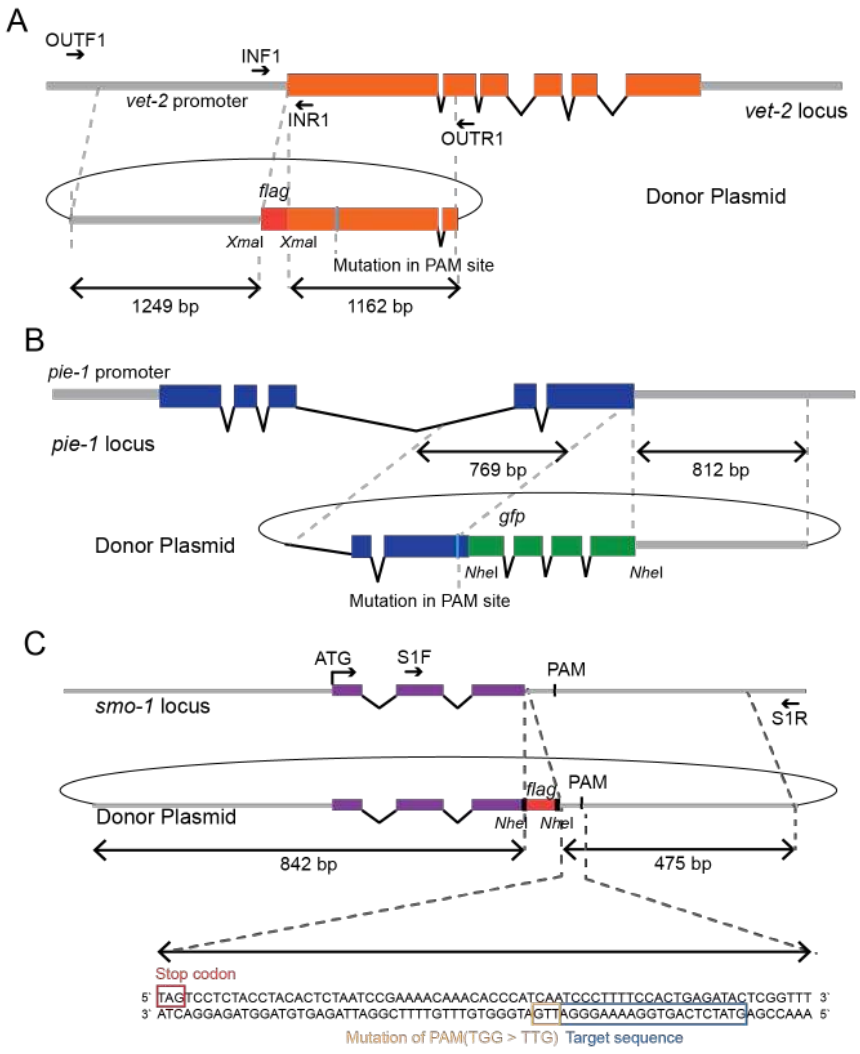


Figure 1.11 HR donor plasmids used in Co-CRISPR experiments

(A) Schematic of the *flag::vet-2* donor plasmid. The *flag* coding sequence was inserted immediately after the *vet-2* start codon and flanked by ~1200 bp homology arms. (B) Schematic of the *pie-1::gfp* donor plasmid. (C) Schematic of the *smo-1::flag* donor plasmid. The donor plasmid includes *flag* coding sequence immediately before the *smo-1* stop codon and asymmetrical homology arms (~800 bp and ~500 bp) flanking the target site, and the Cas9/sgRNA target sequence is located in the 3'UTR of *smo-1*. The PAM sites mutated in each donor indicate the locations of the Cas9/sgRNA target sites.

Interestingly, however, these studies required a modification to the Co-CRISPR screening strategy. For testing sgRNAs using Co-CRISPR, we found that F1 and F2 twitchers were equally likely to exhibit co-induction of indels with the second sgRNA. However, our data suggest that HR events were not enriched and might be depleted among non-rolling, F1 twitcher animals. One possible explanation for this surprising finding is that Cas9-sgRNA complexes may assemble in the germline cytoplasm and then segregate into maturing oocytes independently of the co-injected DNA (including both the roller DNA and of course the donor DNA plasmids). Zygotes inheriting programmed Cas9 could undergo Non-homologous end joining (NHEJ) but HR-directed repair would not be possible without the donor vector. Consistent with this reasoning, we found that in most cases HR events were only enriched among F1 animals that were both rolling, and thus had inherited the injected DNA, and also segregated twitching progeny, indicating that Cas9 was active (Table 1.4). For example among 145 F1 rollers, we found 7 animals heterozygous for a 3' insertion of *gfp* into the *pie-1* locus. Among the F1 twitchers that were non-rolling, zero were GFP-positive, while among the 4 F1 rollers that segregated twitching progeny, two (50%) were GFP positive. One convenient aspect of searching for HR events among F1 rollers heterozygous for *unc-22* twitchers was that the unlinked twitcher phenotype could easily be segregated away in subsequent generations. These findings suggest that Co-CRISPR screening can enhance the detection of HR events. Indeed, we always found at least one HR event among the F1 rollers with

twitcher progeny (3/29, 2/4 and 1/12). However, in most cases additional HR events were also recovered by scoring all the F1 rollers (Table 1.4).

Table 1.4 Co-CRISPR strategy for HR events

HR donor/ Targeted gene	Number Injected	F1 rollers producing F2 twitchers/ Total F1 rollers	F1 having both rolling and twitching phenotype	F1 twitchers	F1 Twitcher-based HR frequency	Roller-based HR frequency	Roller producing F2 twitchers-based HR frequency
flag::vet-2/ vet-2	40	29/65	0	62	2/62 (3%)	4/65 (6%)	3/29 (10%)
pie-1::gfp/ pie-1	40	4/145	0	7	0/7 (0%)	7/145 (5%)	2/4 (50%)
smo::flag / smo-1	40	12/55	10	22	1/22 (5%)	1/55 (2%)	1/12 (8%)

A selection/counter-selection strategy for recovering HR events

The above findings demonstrate that selections are not necessary for identifying and recovering HR events using the CRISPR-Cas9 system. However, for some experiments a dominant selection could save considerable time and expense, especially where insertion of heterologous DNA is likely to be tolerated, for example when generating a null allele of a gene or when one wishes to precisely delete non-coding genes or regulatory elements. The inserted marker also has the potential benefit of providing a selection for maintaining strains that may not be homozygous viable. Previous studies have described several selection strategies, including *unc-119*, NeomycinR, PuromycinR and HygromycinR (Frokjaer-Jensen et al., 2012; Frokjaer-Jensen et al., 2008; Giordano-Santini et al., 2010; Radman et al., 2013; Semple et al., 2010). In order to test a selection counter-selection scheme for CRISPR-induced HR, we decided to employ a new worm antibiotic-resistance marker expressing the bacterial *BSD* as selectable marker, and the *avr-15* gene as a counter-selectable marker. We have previously shown that introducing an *avr-15(+)* plasmid into extrachromosomal arrays and balancer chromosomes can be used to facilitate their counter selection (Duchaine et al., 2006; Shirayama et al., 2012). The *avr-15(+)* vector expresses a gene that confers sensitivity to the potent nematicide drug, ivermectin. This counter-selection can be used to remove *BSD(+)* extrachromosomal transgenes, thus facilitating the recovery of animals bearing an HR-induced insertion of the selectable marker. This counter-selection approach requires a starting strain

resistant to ivermectin, which is conferred by lesions in both the *avr-14* and *avr-15* genes. Ivermectin resistant double mutant strains are essentially wild-type in appearance, and as noted above, new strains can readily be rendered ivermectin resistant by simply co-injecting sgRNAs targeting *avr-14* and *avr-15* (Figure 1.8).

To test this selection/counter-selection strategy we first designed a donor plasmid containing the *BSD* gene flanked with 1 kb *pie-1* homology arms (Figure 1.12A). We injected 58 ivermectin-resistant animals with a mix containing this donor plasmid along with a validated *pie-1* sgRNA vector, the *Peft-3::Cas9* vector, the *rol-6* vector and the *avr-15(+)* vector. Gravid F1 rollers were then placed approximately 11 per plate directly onto plates containing both blasticidin and ivermectin (Figure 1.12B, see materials and methods). After 3 to 4 days we found that two of fourteen plates produced blasticidin-resistant, fertile animals (Figure 1.12B). In a second experiment, we injected the same injection mixture into 40 animals and obtained 103 F1 rollers, from which we identified four blasticidin-resistant strains.

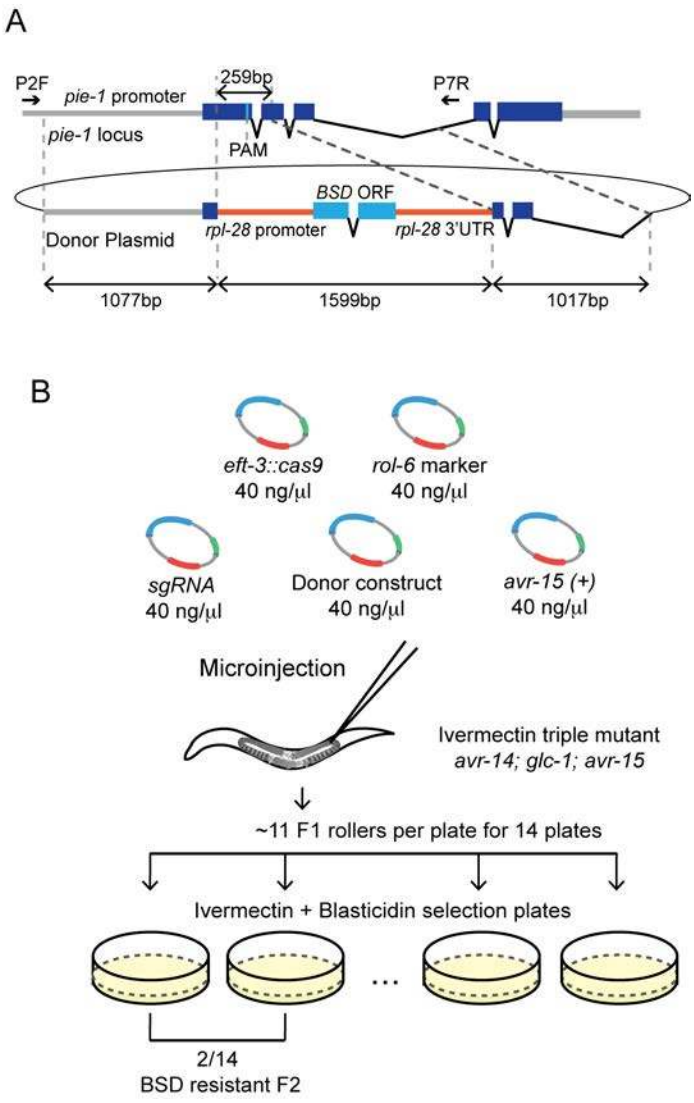


Figure 1.12 A blasticidin-resistance marker to select *pie-1* knockout mutants
(A) Schematic of the Cas9/sgRNA target site and an HR donor plasmid in which a heterologous blasticidin-resistance (BSD) gene replaces a region of *pie-1* and is flanked by 1 kb homology arms. The BSD gene is under the control of the *rpl-28* promoter (568 bps) and 3' utr (568 bps). (B) Schematic representation of the blasticidin selection strategy to precisely delete the *pie-1* gene. *pie-1* a sgRNA was co-injected with the Cas9 expression vector, the *rol-6* transformation marker, the *pie-1*Δ*::BSD* donor construct, and the *pCCM416::Pmyo-2::avr-15(+)* counter-selection vector. The indicated number of F1 rollers was transferred to the plates containing 2 ng/ml ivermectin to select against the extra-chromosomal array, and 100 µg/ml blasticidin to identify BSD knock-in lines. We

identified two plates with resistant, fertile adults among 14 plates, 3 to 4 days after transferring animals.

Discussion

Optimizing sgRNA and donor molecule selection

There is much work still to do in order to optimize CRISPR methodology for *C. elegans*. For example, it remains unclear at this point why upwards of half of the sgRNAs tested fail to induce events. The sgRNAs that we have tested and the activities observed are summarized in Table 1.1, and all of the active sgRNA vectors will be made available through (Addgene). Another area requiring more study is how best to optimize HR donors. All of the HR donor molecules used in the experiments described here were circular plasmids with at least 300 bp homology arms (Figure 1.10A). For GFP insertion we used either 800 bp or 1 kb homology arms, and observed roughly equal frequencies of HR in both cases (Figure 1.9A and Figure 1.11B). Future studies should explore shorter homology arms and other types of donor molecules including linear dsDNA donor molecules produced, for example, by PCR, as well as chemically synthesized ssDNA. It will also be important to explore the optimal distance between the cut site and the homology arm. Increasing this distance requires longer gene-conversion tracts, and in other organisms is correlated with reduced frequency of the desired homologous event (Paques and Haber, 1999). Our findings suggest that gene-conversion tracts of approximately 250 bp are common in *C. elegans*. Optimizing HR conditions for each type of donor molecule will likely require extensive experimentation in order to generate statistically significant findings on relative efficiencies. Although there is still much work to do, the efficiencies

reported here are already remarkably high. For example, indels were frequently identified in greater than 10% of F1 rollers, and the ratio of HR events to the total number of CRISPR-Cas9-induced-repair events was consistently in the range of 10% over the course of all of our experiments where both HR and indels were monitored.

Conclusions

Our findings provide a versatile framework for using CRISPR-Cas9 genome editing in *C. elegans,* and the Co-CRISPR strategy we employ is likely to be of value for CRISPR-Cas9 studies in other organisms. The tools described here, however, are likely to be just the beginning of what will be possible in the near future. For example, the use of catalytically inactive Cas9 fusion proteins to tether regulators to DNA targets has not been described yet in *C. elegans,* but it is already finding many exciting applications in other systems (Cheng et al., 2013; Kearns et al., 2014; Larson et al., 2013; Mali et al., 2013a; Qi et al., 2013). CRISPR-Cas9 technology should also dramatically facilitate the use of other nematode models, including species distantly related to *Caenorhaditis elegans* and perhaps parasitic nematodes. The ability to efficiently engineer genomes will only enhance the utility of model organisms where gene variants can now be generated and analyzed rapidly and cost-effectively, facilitating the production of new animals models for human disease-associated alleles. Moreover, CRISPR-Cas9-engineered strains with special alleles of

important genes can be used as starting strains in forward genetic screens, including suppressor and enhancer screens, which are extremely powerful in *C. elegans.* It is now easier than ever for researchers to use *C. elegans* to explore the function of conserved genes of interest. Indeed, the CRISPR-Cas9 technology lowers the barrier to move from one system to another, effectively making all organisms one, when exploring conserved cellular mechanisms.

CHAPTER II: PIE-1, SUMOYLATION, AND EPIGENETIC REGULATION OF GERMLINE SPECIFICATION IN *C. ELEGANS*

Introduction

During embryogenesis, the most fundamental event in many organisms is the differentiation of germline cells and specialized somatic cells (Ikenishi, 1998). In C. *elegans*, the separation of the germline from the soma occurs during the first four embryo cleavages (Sulston et al., 1983). PIE-1 is a tandem CCCH-type zinc-finger protein essential for the germline fate (Mello et al., 1996). In the germline blastomere, PIE-1 inhibits embryonic transcription in the nucleus (Mello et al., 1992; Seydoux et al., 1996) and promotes maternal gene expression in the cytoplasm (Tenenhaus et al., 2001). PIE-1 interacts with the NuRD complex, which is required to repress the germline fate in somatic cells (Unhavaithaya et al., 2002), suggesting that PIE-1 and NuRD have antagonistic activities. Consistent with this idea, somatic expression of PIE-1 mimics MEP-1 loss of function and results in ectopic expression of germline genes and stem cell fates in the soma (Unhavaithaya et al., 2002). However, the molecular mechanisms by which PIE-1 functions in germline blastomeres remain to be determined.

Here, we investigated the functional consequence of the interaction between PIE-1 and SUMOylation, a post-translational modification with important biological functions. We aimed to address whether the SUMOylation of PIE-1 and associated chromatin-remodeling factors is required for transcriptional repression and epigenetic regulation in the *C. elegans* germline. Understanding the epigenetic mechanisms involved in germline specification will provide information

about the molecular mechanisms that control the maintenance and differentiation of germline stem cells.

Materials and Methods

Yeast two-hybrid analysis

The yeast two-hybrid screen was performed by Hybrigenics services (Paris, France, http://www/hybrigenics-services.com). The coding sequence for amino acids 2–335 of *C. elegans pie-1* (NM_001268237.1) was amplified by PCR from N2 cDNA and cloned into pB66 via C-terminal fusion with the Gal4 DNA-binding domain (Gal-4-PIE-1). The construct was checked by sequencing and used as a bait to screen a random-primed *C. elegans* Mixed Stages cDNA library constructed into pP6. 5 million clones (5-fold the complexity of the library). The library was screened using a mating approach with YHGX13 (Y187 ade2-101::loxp-kanMX-loxP, matα) and CG1945 (mata) yeast strains as previously described (Fromont-Racine et al., 1997). 153 His+ colonies were selected on a medium lacking tryptophan, leucine, and histidine, and supplemented with 0.5 mM 3-aminotriazole to prevent bait autoactivation. The prey fragments of the positive clones were amplified by PCR and sequenced at their 5' and 3' junctions. The resulting sequences were used to identify the corresponding interacting proteins in the GenBank database (NCBI) using a fully automated procedure. A confidence score (predicted biological score [PBS]) was attributed to each interaction as previously described (Formstecher et al., 2005).

C. elegans strains and genetics

The strains and alleles used in this study were N2 Bristol (wild-type), *pie-1(zu154) unc-25 (e156)/qC1, rde-3(ne3370),* and lines newly generated using CRISPR/Cas9-mediated genome editing as described in Table 2.1. Worms were, unless otherwise stated, maintained at 20°C on NGM plates seeded with OP50 *Escherichia coli*, and genetic analyses were performed essentially as described (Brenner, 1974).

Table 2.1 Summary of experimental models: CRISPR strains

ubc-9(G56R)	ubc-9(ne4446[UBC-9(G56R)])	This study
3xflag::smo-1	smo-1(ne4309[3xFLAG::SMO-1])	This study
his$_{10}$::smo-1	smo-1(ne4346[10XHIS::SMO-1])	This study
pie-1(K68R)	pie-1(ne4303[PIE-1(K68R)])	(Kim et al., 2014)
pie-1(K68R)::3xflag	pie-1(ne4304[PIE-1(K68R)::3xFLAG)	This study
pie-1::3xflag	pie-1(ne4302[PIE-1::3xFLAG])	(Kim et al., 2014)
pie-1::gfp	pie-1(ne4301[PIE-1::GFP])	(Kim et al., 2014)
pie-1(K68R)::gfp	pie-1(ne4377[PIE-1(K68R)::GFP]	This study
mep-1::gfp::tev::3xflag	mep-1(ne4380[MEP-1::GFP::TEV::3xFLAG])	This study
pgl-1::mCherry	pgl-1(ne4394[PGL-1::mCherry])	This study
pie-1(DAQMEQT)::3xflag	pie-1(ne4389[PIE-1(DAQMEQT)::3xFLAG])	This study
pie-1(DAQMEQT)::gfp	pie-1(ne4378[PIE-1(DAQMEQT)::GFP])	This study
cdk-12-as	cdk-12([CDK-12(F383G)])	Gift from Hermand Damien, University of Namur, Belgium
gcna-1(Δ2bp)	gcna-1(ne4334[GCNA-1(Δ2bp)])	(Carmell et al., 2016)
gcna-1(Δ1748bp)	gcna-1(ne4356[GCNA-1(Δ1748bp)]	(Carmell et al., 2016)

CRISPR/Cas9 genome editing

The Co-CRISPR strategy (Kim et al., 2014) using *unc-22* sgRNA as a co-CRISPR marker was used to enrich HR events for tagging a gene of interest with the non-visualizing epitope (*10xhis* and *3xflag*) or introduction of a point mutation (G56R). To screen for insertions of *10xhis* and *3xflag*, we used two-round PCR: the First PCR was performed with primers (F: cctcaaaaaccaagcgaaaacc R: ccggctgctatttcattgat), and 1 µl of the 1st PCR product was used as a template for the 2nd PCR with primers (F:gagactcccgctataaacga R:ctcaagcaggcgacaacgcc). To detect *10xhis* knock-ins, the final products were run either on 2% Tris/borate/EDTA (TBE) gel or 10% PAGE gel. *The ubc-9(G56R)* mutation introduced an *Hae*III restriction fragment length polymorphisms (RFLP) that was used to screen for G56R conversion in PCR products (F: cattacatggcaagtcggg, R: cgttgccgcatacagaaatag). For visualization of either GFP or mCherry tag, F1 rollers were mounted under coverslips on 2% agarose pads to directly screen for GFP or mCherry expressing animals as described previously (Kim et al., 2014).

sgRNA construct

*P*reviously generated *pie-1* sgRNA plasmid (Kim et al., 2014) was used for *pie-1 (DAQMEQT)::3xflag*. *ubc-9* sgRNA (ggctcgaacttgcactttgg), *smo-1* sgRNA (gagactcccgctataaacga), *pgl-1* sgRNA (cgtggacgtggtggttacgg), and *mep-1* sgRNA (gcgcaaaagaaggaagacgg) were constructed by ligating annealed sgRNA oligonucleotides to *Bsa*I-cut pRB1017 (Arribere et al., 2014).

Donor template

Unless otherwise stated, a silent mutation to disrupt the PAM site in each HR donor was introduced by PCR sewing.

ubc-9(G56R): For the *ubc-9(G56R)* donor construct, a *ubc-9* fragment was amplified with primers (F: cattacatggcaagtcggg, R: gacgactaccacgaagcaagc) and this fragment was cloned into the Blunt II-TOPO vector (Invitrogen, K2800-20). To introduce the point mutation (G56R) and mutate the seeding region, overlap extension PCR was used (Figure 2.1C).

his$_{10}$::smo-1 and *3xflag::smo-1*: Using PCR sewing, either a *10xhis* fragment (caccatcaccatcaccatcaccaccatcac) or *Nhe*I site was introduced immediately after the start codon in the previously generated *smo-1* fragment (Kim et al., 2014). The resulting PCR product was cloned into the Blunt II-TOPO vector. A *3xflag* fragment with *Nhe*I overhangs was generated by annealing two oligonucleotides and ligated into the *smo-1* donor plasmid by the *Nhe*I site. Tagging with either *10xhis* or *3xflag* on the N-terminus of *smo-1* disrupted the PAM site.

pgl-1::mCherry: The *pgl-1* genomic sequence was amplified from N2 genomic DNA with primers (F: gtagctctgccaccggtatc, R: gcgggaagaccatcgaaaaatag) and inserted into pCR-Blunt II-TOPO vector. Overlap extension PCR was used to

introduce a *Nhe*I site immediately before the stop codon in this fragment of *pgl-1*. The mCherry-coding sequence amplified from pCFJ90 (Addgene) was inserted into the *Nhe*I site.

pie-1(DAQMEQT)::3xflag and *pie-1(DAQMEQT)::gfp* : To introduce the DAQMEQT mutations into either the *pie-1::flag* or *pie-1::gfp* donor plasmid previously built in pCR-Blunt II TOPO (Kim et al., 2014), PCR sewing was performed with primers: 1st left arm product (F: gaaaacggcttctcagacca R: gtagtatgtctgctccatctgcgcatccattggctggctattc), 1st right arm product (F: gccagccaatggatgcgcagatggagcagacatactactatc, R: tacaaagtccgcaactgtgc), and 2nd sewing PCR product (F: gaaaacggcttctcagacca R: tacaaagtccgcaactgtgc). The 2nd PCR products were gel-purified and cloned into the pCR-Blunt II Topo vector.

mep-1::gfp::tev::3xflag: For the *mep-1::gfp::tev::3xflag* donor construct, a *mep-1* fragment was amplified with primers (F1:gaaattcgctggcagtttct R1: ctgcaacttcgatcaatcga) from N2 genomic DNA and inserted into the pCR-Blunt II-TOPO vector. Overlap extension PCR was used to introduce an *Nhe*I site immediately before the stop codon in this *mep-1* fragment. The *Nhe*I site was used to insert the *gfp::tev::3xflag* coding sequence.

RNAi

RNAi was performed by feeding worms *E. coli* strain HT115 (DE3) transformed with the control vector or a gene-targeting construct from the *C. elegans* RNAi Collection (Kamath and Ahringer, 2003). L4 larval stage animals were placed on RNAi plates (NGM plates containing 1mM isopropyl β-d-thiogalactoside (IPTG) and 100 μg/ml ampicillin seeded with dsRNA-containing bacteria) and allowed to develop into adults. After 24h, adult animals were transferred to fresh RNAi plates and allowed to lay eggs over night. On the following day, the unhatched eggs were analyzed 12 h later after adults were removed from the test plates.

Immunofluorescence

Gonads were dissected and embryos were prepared by bleaching synchronous adult animals. Samples were placed on glass slides coated with 0.1% (w/v) Poly-L-lysine solution (Sigma-Aldrich, P8920) and fixed with 4% paraformaldehyde in phosphate-buffered saline (PBS) for 20min, freeze-cracking on dry ice, and –20°C cold methanol for 5min. After fixation, slides containing samples were washed three times with PBST (PBS containing 0.1% Triton X-100), blocked with 0.1% bovine serum albumin (BSA) in PBST for 30 min, and then incubated with primary antibodies overnight at 4°C or for 1h at room temperature. After three washes with PBST, secondary antibodies were applied for 1h at room temperature. Sample slides were washed three times with PBS and mounted with 10 μl of Vectashield with DAPI (VWR, 101098-44). The primary antibodies

(1:100) used were: anti-acetyl-histone H4 (Thermo Fisher Scientific, 06855MI), anti-acetyl-histone H3Lys9 (Millipore Sigma, 07-352), anti-trimethyl-histone H3Lys9 (Millipore Sigma, 051250), and anti-NPP-9 (Novus Biologicals, 48610002). The secondary antibodies (1:1000) used were: goat anti-rabbit IgG(H+L) Alexa Fluor 647(Thermo Fisher Scientific, A32733) and goat anti-mouse IgG(H+L) Alexa Fluor 488 (Thermo Fisher Scientific, A11001).

Microscopy

For live imaging, worms and embryos were mounted in M9 on a 2% agarose pad with or without 1mM levamisole (Sigma-Aldrich, 16595-80-5). Epi-fluorescence and differential interference contrast (DIC) microscopy were performed using an Axio Imager M2 Microscope (Zeiss). Images were captured with an ORCA-ER digital camera (Hamamatsu) and processed using Axiovision software (Zeiss). Confocal images were acquired using a Zeiss Axiover 200M microscope equipped with a Yokogawa CSU21 spinning disk confocal scan head and custom laser launch and relay optics (Solamere Technology Group). Stacks of images were taken and analyzed using ImageJ.

Immunoprecipitation

Either synchronous adult worms (~200,000 animals) or early embryos (bleached from 10×100,000 adult animals) were collected and washed three times with M9 buffer before being homogenized in lysis buffer [20mM HEPES pH 7.5, 125mM

$Na_3C_6H_5O_7$ (sodium citrate), 0.1%(v/v) Tween 20, 0.5%(v/v) Triton X-100, 2mM $MgCl_2$, 1mM DTT, and a Mini Protease Inhibitor Cocktail Tablet (Roche)] using a FastPrep-24 benchtop homogenizer (MP Biomedicals). Worm or embryo lysates were centrifuged twice at 14,000×g for 30 min at 4°C, and supernatants were incubated with GFP-binding protein (GBP) beads for 1.5 h at 4°C on a rotating shaker. The beads were washed three times with immunoprecipitation (IP) buffer containing protease inhibitor for 5 min each wash and then washed twice with high-salt wash buffer (50mM HEPES pH 7.5, 500mM KCl, 0.05% NP40, 0.5mM DTT, and protease inhibitor). Immune complexes were eluted with elution buffer (50mM Tris-Cl pH 8.0, 1xSDS) for 10 min at 70°C.

In vivo SUMO purification

Synchronous adult worms (~200,000) or 500 µl of packed embryos (bleached from (~1,000,000 synchronous adult worms) were homogenized in lysis buffer at pH 8.0 (6M guanidine-HCl, 100mM Na_2HPO_4/NaH_2PO_4 pH 8.0, and 10mM Tris-HCl pH8.0) using a FastPrep-24 benchtop homogenizer (MP Biomedicals). Lysates were cleared by centrifugation at 14,000×g for 30 min at 4°C and equalized using quick start Bradford 1x dye reagent (BioRad, 5000205). Ni-NTA resin was washed three times with lysis buffer containing 20mM imidazole pH 8.0 and 5mM β-mercaptoethanol while samples were prepared. To equalized samples, we added imidazole pH 8.0 to 20mM and β-mercaptoethanol to 5mM, and then the samples were incubated with 100 µl of pre-cleared 50% slurry of Ni-

NTA resin (Qiagen, 30210) for 2–3 h at room temperature on a rotating shaker. Ni-NTA resin was washed in 1-ml aliquots of the following series of buffers: Buffer 1 pH 8.0 (6M guanidine-HCl, 100mM Na_2HPO_4/NaH_2PO_4 pH 8.0, 10mM Tris-HCl pH 8.0, 10mM imidazole pH 8.0, 5mM β-mercaptoethanol, and 0.1% Triton X-100), Buffer 2 pH 8.0 (8M urea, 100mM Na_2HPO_4/NaH_2PO_4 pH 8.0, 10mM Tris-HCl pH 8.0, 10mM imidazole pH 8.0, 5mM β-mercaptoethanol, and 0.1% Triton X-100), Buffer 3 pH 7.0 (8M urea, 100mM Na_2HPO_4/NaH_2PO_4 pH 7.0, 10mM Tris-HCl pH 7.0, 10mM imidazole pH 7.0, 5mM β-mercaptoethanol, and 0.1% Triton X-100), Buffer 4 pH 6.3 (8M urea, 100mM Na_2HPO_4/NaH_2PO_4 pH6.3, 10mM Tris-HCl pH 6.3, 10mM imidazole pH 6.3, 5mM β-mercaptoethanol, and 0.1% Triton X-100), Buffer 5 pH 6.3 (8M urea, 100mM Na_2HPO_4/NaH_2PO_4 pH 6.3, 10mM Tris-HCl pH 6.3, and 5mM β-mercaptoethanol). To purify the polyhistidine-tagged proteins under denaturing conditions, 8M urea was added to all wash buffers. Triton-X-100, the non-ionic detergent was used to reduce nonspecific hydrophobic interactions. Imidazole (10mM) was used to increase the stringency of the wash by reducing nonspecific protein binding to the resin. The use of wash buffers with gradually decreasing pH (pH8 to pH6.3) also reduced nonspecific binding of proteins by protonating the neutral histidine and thereby removing the weakly bound proteins that may contain tandem repeats of the histidine. The SUMOylated proteins were eluted with elution buffer pH 7.0 (7M urea, 100mM Na_2HPO_4/NaH_2PO_4 pH 7.0, 10mM Tris-HCl pH 7.0, and 500mM imidazole pH 7.0) For western blotting, input samples containing

guanidine-HCl were diluted with H_2O (1:6) and then purified by trichloroacetic acid (TCA) precipitation: an equal volume of 10% TCA was added to diluted samples, which were then incubated on ice for 20 min and centrifuged for 20 min at 4°C; the obtained pellet was washed with 100 µl of ice-cold ethanol and then resuspended in Tris-HCl buffer pH 8.0.

Western blot analysis

NuPage LDS sample buffer (4x) (Thermo Fisher Scientific, NP0008) was added to samples, which were then loaded on precast NuPAGE Novex 4-12% Bis-Tris protein gel (Thermo Fisher Scientific) and transferred onto a polyvinylidene difluoride (PVDF) membrane (Bio-Rad) using Mini Trans-Blot cells (Bio-Rad) at 80 V for 2.2 h at 4°C. Membranes were blocked with 5% skim milk and probed with primary antibodies: anti-FLAG (1:1000) (Sigma-Aldrich, F1804), anti-GFP (1:500) (Roche, 11814460001), anti-HDA-1 (1:1000) (Novus Biologicals, 38660002), anti-Let-418 (1:1000) (Novus Biologicals, 48960002), anti-SMO-1 (1:1000) (purified from Hybridoma cell cultures), and anti-PIE-1(P4G5) (1:100). Antibody binding was detected with secondary antibodies (1:2500): goat anti-mouse (Abcam, ab6789) and goat anti-rabbit (Thermo Fisher Scientific, 656120).

Results

Interaction of PIE-1 with SMO-1, UBC-9, and GEI-17

To explore the molecular mechanisms by which PIE-1 functions in transcriptional regulation for germline specification, we sought to identify putative interacting partners of PIE-1 by yeast two-hybrid screening. In this screen, we used the full-length *pie-1* cDNA as bait and a mixed-stage *C. elegans* cDNA library as prey. Interestingly, SMO-1 (SUMO) and two enzymes involved in the SUMOylation pathway, UBC-9 (E2 SUMO enzyme) and GEI-17 (E3 SUMO ligase), were identified as PIE-1 protein interactors in this screen (Table 2.2). The failure to identify E1 SUMO activation enzyme is consistent with its known function in binding SUMO proteins but not substrates (Gareau and Lima, 2010). This result implies that PIE-1 function may be dependent on SUMOylation, because all SUMO components that can bind to a protein substrate were identified in this screen.

Table 2.2 PIE-1 interactors identified in yeast two-hybrid screen

The score that was computed to assess interaction reliability represents the probability of non-specific interaction. A is the lowest probability and E is the highest probability.

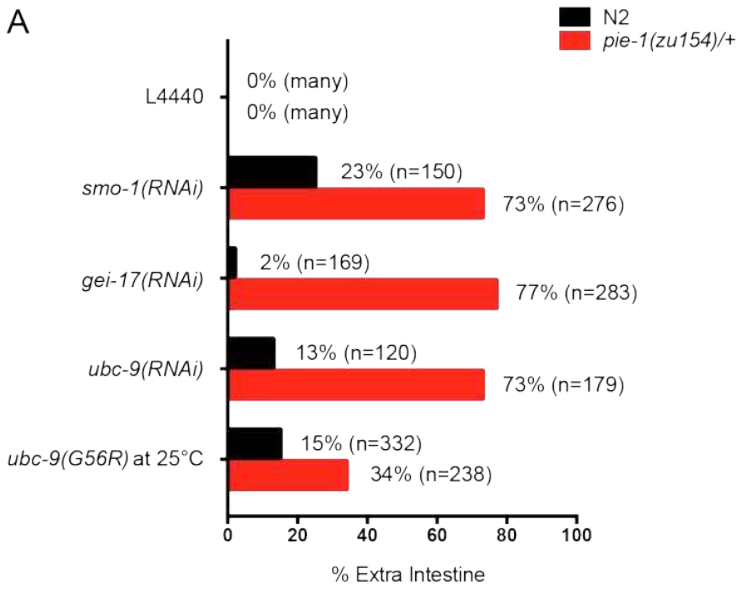
Gene	Domain/function	Score
ZIF-1	ZF interacting protein/ubiquitination	A
GEI-17	E3 ligase for sumo	A
GLD-3	KH domain/Germ line development	B
MBL-1	MUSCLEBLIND-mRNA splicing regulator	B
PIE-1	Probable homodimer	B
POP-1	TCF family member/posterial pharynx defect	B
LON-1	PR-protein superfamily/ a target of TGF beta family	B
SMO-1	SUMO	C
UBC-9	E2 ligase for sumo	E
MEP-1	C2H2/function with NuRD complex	E
RAD-26	An ortholog of human RAD54L2 A transcription-coupled repair factor	E

We next analyzed genetic interactions between *pie-1* and *smo-1, ubc-9, and gei-17* to investigate whether SUMOylation facilitates PIE-1 function. Depletion of *smo-1*, *ubc-9*, or *gei-17* by RNAi resulted in the death of all the progeny of wild-type animals during embryogenesis, but the dead embryos showed low penetrance extra intestinal cells derived from the germline blastomere, a phenotype similar to the *pie-1* null mutant (Figure 2.1A). RNAi of these genes caused *pie-1* heterozygotes to make high penetrance dead embryos with the *pie-1* mutant phenotype (Figure 2.1A), suggesting that SUMOylation is required for PIE-1 function.

Generation of temperature-sensitive ubc-9 alleles and analysis of genetic interaction between pie-1 and ts ubc-9

For a detailed genetic study of SUMOylation with *pie-1*, we aimed to generate *temperature-sensitive* (*ts*) *ubc-9* alleles by CRISPR/Cas9-mediated genome editing, because the homozygous *ubc-9* deletion allele (tm2610) showed strong zygotic effect-sterility. Previously, a *ts* Ubc9 allele (P69S) was generated in the yeast *Saccharomyces cerevisiae* by mutating a conserved proline residue across other members of the Ubc protein family (Betting and Seufert, 1996). We were also able to find the conserved proline residue in the *C. elegans* UBC-9 protein as the UBC-9 protein sequence is highly homologous to yeast UBC9 (94%). The newly generated *ubc-9 (P69S)*, however, showed a strong sterile phenotype like the null allele at 20°C although the strain was fertile, but produced dead embryos

at 15°C. The G56 residue corresponds to another identified ts allele, G58 in yeast Cdc34 (Prendergast et al., 1995), which is also conserved in all UBC proteins across other organisms (Figure 2.1B). Using CRISPR/Cas9-mediated genome editing, we successfully generated the *ts* mutant *ubc-9 (G56R)* (Figure 2.1C) that has a wild-type appearance at 15°C but produces 100% dead embryos at 25°C (Figure 2.1A). Employing the *ts* mutants of *C. elegans*, we asked whether the *ubc-9* homozygous strain while *pie-1* heterozygous, produces dead embryos with a *pie-1* phenotype (extra intestine). The occurrence of a *pie-1* phenotype was significantly increased in *ubc-9 (G56R); pie-1/+* mutants compared to *ubc-9 (G56R)* mutants at 25°C (Figure 2.1A). However, we noticed that *ts ubc-9* mutants in *pie-1* heterozygotes (25°C) showed a lower penetrance *pie-1* phenotype than after *ubc-9* RNAi (20°C) (Figure 2.1A). One possibility is that among the target proteins of UBC-9 may be *ts* genes that cause a different phenotype from *ubc-9* RNAi. Considering that RNAi-mediated depletion usually does not completely shut off the target gene, another possible explanation is that *ubc-9 (G56R)* at 25°C can completely remove functional UBC-9 protein, whereas using RNAi depletes UBC-9 partially, thereby causing *ubc-9(G56R)* (25°C) mutation to gives rise to a phenotype more biased towards the *ubc-9* mutant than *ubc-9* RNAi.


A
N2
pie-1(zu154)/+
L4440
0% (many)
0% (many)
smo-1(RNAi)
23% (n=150)
73% (n=276)
gei-17(RNAi)
2% (n=169)
77% (n=283)
ubc-9(RNAi)
13% (n=120)
73% (n=179)
ubc-9(G56R) at 25°C
15% (n=332)
34% (n=238)
0
20
40
60
80
100
% Extra Intestine


B
Yeast CDC34 (UBC3) MVLNEDSIYHGGFFKAQMRFPEDFPFSPPQFRFTPAIYHPNVY-RDGRLCISILHQSGDP
Yeast UBC4 MG-PADSPYAGGVFFLSIHFPTDYPFKPPKISFTTKIYHPNIN-ANGNICLDILKDQ---
Yeast UBC5 MG-PSDSPYAGGVFFLSIHFPTDYPFKPPKVNFTTKIYHPNIN-SSGNICLDILKDQ---
Yeast UBC1 LG-PPGTPYEGGKFVVDIEVPMEYPFKPPKMQFDTKVYHPNISSVTGAICLDILKNA---
Yeast RAD6 (UBC2) IG-PADTPYEDGTFRLLLEFDEEYPNKPPHVKFLSEMFHPNVY-ANGEICLDILQNR---
Human UBC9 PG-KKGTPWEGGLFKLRMLFKDDYPSSPPKCKFEPPLFHPNVY-PSGTVCLSILEEDKD-
Mouse UBC9 PG-KKGTPWEGGLFKLRMLFKDDYPSSPPKCKFEPPLFHPNVY-PSGTVCLSILEEDKD-
Drosophila UBC9 PG-KKSTPWEGGLYKLRMIFKDDYPTSPPKCKFEPPLFHPNVY-PSGTVCLSLLDEEKD-
Yeast UBC9 PG-KEGTNWAGGVYPITVEYPNEYPSKPPKVKFPAGFYHPNVY-PSGTICLSILNEDQD-
C. elegans UBC-9 PG-RKDTIWEGGLYRIRMLFKDDFPSTPPKCKFEPPLFHPNVY-PSGTVCLSLLDENKD-
Gly56
Pro69
ts alleles in yeast

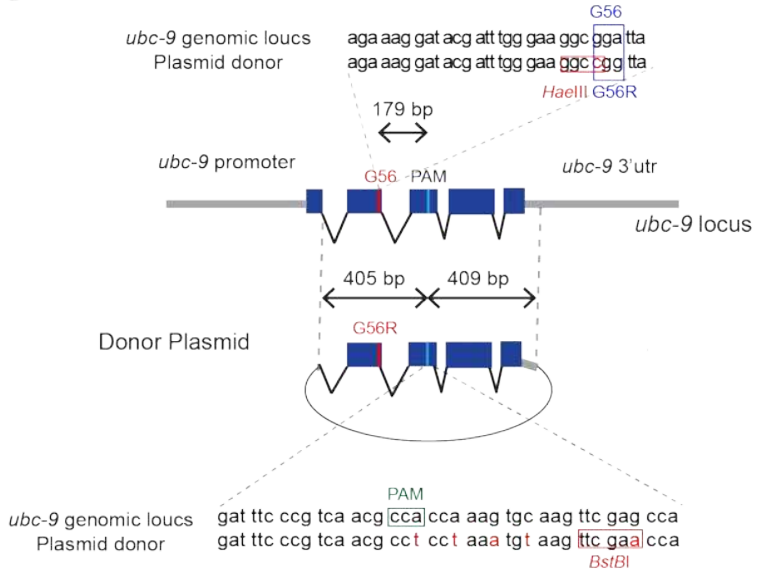

C
G56
ubc-9 genomic loucs
aga aag gat acg att tgg gaa ggc gga tta
Plasmid donor
aga aag gat acg att tgg gaa ggc cgg tta
HaeIII G56R
179 bp
ubc-9 promoter
G56
PAM
ubc-9 3'utr
ubc-9 locus
405 bp
409 bp
Donor Plasmid
G56R
PAM
ubc-9 genomic loucs
gat ttc ccg tca acg cca cca aag tgc aag ttc gag cca
Plasmid donor
gat ttc ccg tca acg cct cct aaa tgt aag ttc gaa cca
BstBI

Figure 2.1 Genetic interactions between *pie-1* and *smo-1*, *ubc-9*, and *gei-17*
(A) Genetic interaction of *pie-1* with SUMOylation pathway. Defective SUMOylation pathway by RNAi feeding-mediated each indicated gene knockdown causes high frequency of *pie-1* dead embryos at 20°C. The *ubc-9(G56R) ts* allele makes 2-fold increased *pie-1* dead embryos at 25°C, but shows a lower penetrance extra gut derived from the germline blastomere compared to RNAi feeding experiments. (B) Partial sequence alignment of UBC enzymes including UBC9. Residues conserved in all UBC proteins are shown in red. The indicated G56 and P69 residues are isolated *ts* alleles in yeast Cdc34 and both are conserved in all UBC proteins over other organisms (blue box). Among many conserved residues, the G56 residue is identified as a *ts* allele in C. elegans. (C) Schematic of the Cas9/sgRNA target sites in *ubc-9* locus and donor plasmid. The G56R donor contains ~400 bp of homology flanking the PAM site (green box) and introduces a *Hae*III (by codon substitution) and *BstB*I (by mutation in the seed region) restriction sites (red box). The blue bar indicates the PAM site, and the red bar indicates the position of G56.

PIE-1 mislocalization in SMO-1-depleted oocytes

As implied by the analysis of genetic interactions, if SUMOylation regulates PIE-1 function in germline specification, PIE-1 protein stability or localization in the germline lineage might be affected by SMO-1 depletion. Alternatively, SUMOylation might regulate PIE-1 at the molecular levels, such as inducing PIE-1 conformational changes to facilitate or inhibit interactions with its binding partners. To assess these possibilities, we first examined PIE-1 localization in previously generated *pie-1::gfp* CRISPR lines (Kim et al., 2014). The dynamic distribution of PIE-1 in the early embryos has already been well studied. As expected, embryonic localization of PIE-1::GFP in the CRISPR lines was consistent with previous findings that PIE-1 is present in the nuclear and cytoplasmic P-granule of the germline blastomeres (Figure 2.2A) (Mello et al., 1996; Reese et al., 2000; Tenenhaus et al., 1998). In the adult germline, however, we observed that PIE-1 localization was different from the previous reports of PIE-1 detection in the cytoplasm of oocytes both using a monoclonal antibody (Tenenhaus et al., 1998) and by generating transgenic lines using the complex array method (Reese et al., 2000). PIE-1 protein was first detected in pachytene nuclei at low levels but gradually increased throughout the germline development until fertilization (Figure 2.2A). We also observed PIE-1 in the nuclei of oocytes where the PIE-1 expression pattern appears to be associated with DNA (Figures 2.2A and 2.2B).

We then examined if *smo-1* RNAi causes PIE-1 mislocalization. In oocytes depleted of SMO-1 by RNAi, we observed that PIE-1 was obviously sequestered in the nucleolar structure (Figures 2.2B and 2.2C), suggesting that proper PIE-1 localization requires SUMOylation. The oocytes with the sequestered PIE-1 in the nucleolus were fertilized normally, but embryonic lethality followed. Unlike in the oocytes, we did not observe an obvious difference in the nuclear localization of PIE-1 between the control and SMO-1-depleted embryos, except that the punctate cytoplasmic expression of PIE-1 in P-granules appeared mildly diminished at the P2 blastomere in SMO-1-depleted embryos (data not shown).

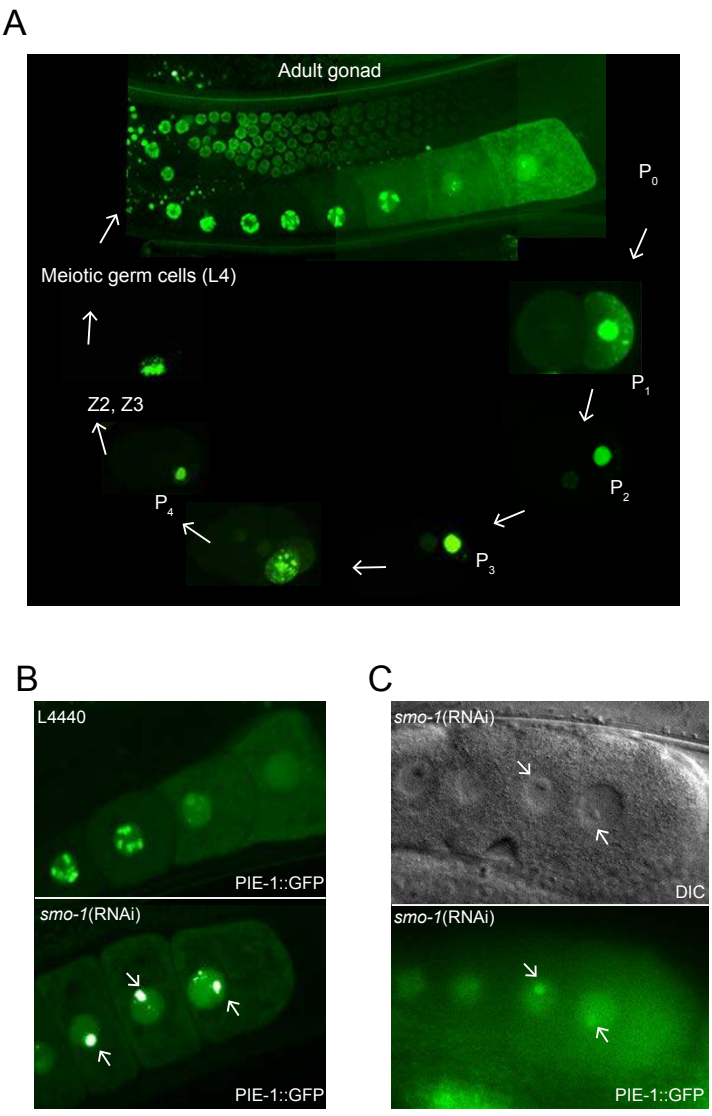


Figure 2.2 PIE-1 expression in the adult germline

(A) Confocal images of adult germline and embryos in a strain expressing PIE-1::GFP. (B) Confocal images of adult germline expressing PIE-1::GFP either without (L4440) or with *smo-1*(RNAi). (C) Nomarski DIC and GFP fluorescence micrographs of oocytes in SMO-1-depleted adults by RNAi. The white arrow indicates the PIE-1 expression in the nucleoli.

PIE-1 is SUMOylated on lysine 68 residue in the C. elegans germline

We then asked whether PIE-1 is indeed SUMOylated in *C. elegans*. PIE-1 contains only one consensus SUMO acceptor site, consisting of the sequence ψKXE (Figure 2.3A) where SUMOylation usually occurs (Rodriguez et al., 2001). These residues are perfectly conserved in *C.briggsae* and *C. elegans* (data not shown). Thus, we investigated whether the consensus SUMO acceptor site is required for PIE-1 SUMOylation if PIE-1 is SUMOylated. To perform an *in vivo* SUMO purification assay in *C. elegans*, we generated an endogenously tagged *his$_{10}$::smo-1* CRISPR strain (Figure 2.3B) that allows SUMO-modified proteins to be enriched by nickel affinity chromatography under denaturing conditions. Purification under denaturing conditions improves detection of SUMOylated proteins at levels similar to that of the endogenous protein by removing non-covalent interactions and inhibiting SUMO-specific protease activity (Tatham et al., 2009). Previously generated wild-type *pie-1::flag* and the consensus SUMO acceptor site mutant *pie-1(K68R)::flag* CRISPR alleles using each plasmid donor (Figure 2.3C) (Kim et al., 2014) were introduced into *his$_{10}$::smo-1* CRISPR lines by genetic mating, because the PIE-1 antibody (P4G5) is not effective for detecting PIE-1 SUMOylation. This is because the peptide used to make P4G5 antibody contains amino acids 54–73 which include a potential SUMO-binding site, K68 (Mello et al., 1996; Tenenhaus et al., 1998) and we observed that the P4G5 antibody no longer recognized PIE-1 protein when we engineered a strain expressing a single copy of the genomic fusion

gfp::pie-1(K68R) transgene (data not shown) (Frokjaer-Jensen et al., 2008). We also created a multiple epitope, GFP::TEV (a TEV protease cleavage site, ENLYFQG)::3xFlag, tagged *mep-1* CRISPR line using a plasmid donor (Figure 2.3D) for further functional studies of PIE-1, because PIE-1 directly interacts with the MEP-1 and NuRD complex (Unhavaithaya et al., 2002). SUMO purification followed by western blotting revealed a slowly migrating PIE-1 band that represents the SUMOylated PIE-1 in *his10::smo-1; pie-1::flag,* but not in *his10::smo-1; pie-1(K68R)::flag* (Figure 2.3F). The SUMOylated PIE-1 band, however, was detected only in adult lysates and not in embryo lysates, whereas the SUMOylated MEP-1 band was present in both lysates (the SUMOylated MEP-1 was used as a positive control of SUMO purification after confirming SUMO is conjugated to MEP-1 in both lysates) (Figures 2.3E and 2.3F). Taken together, these data suggest that PIE-1 is indeed a SUMO substrate and the lysine 68 residue is a crucial site for PIE-1 SUMOylation. In addition, PIE-1 SUMOylation appears to occur in the germline before fertilization.

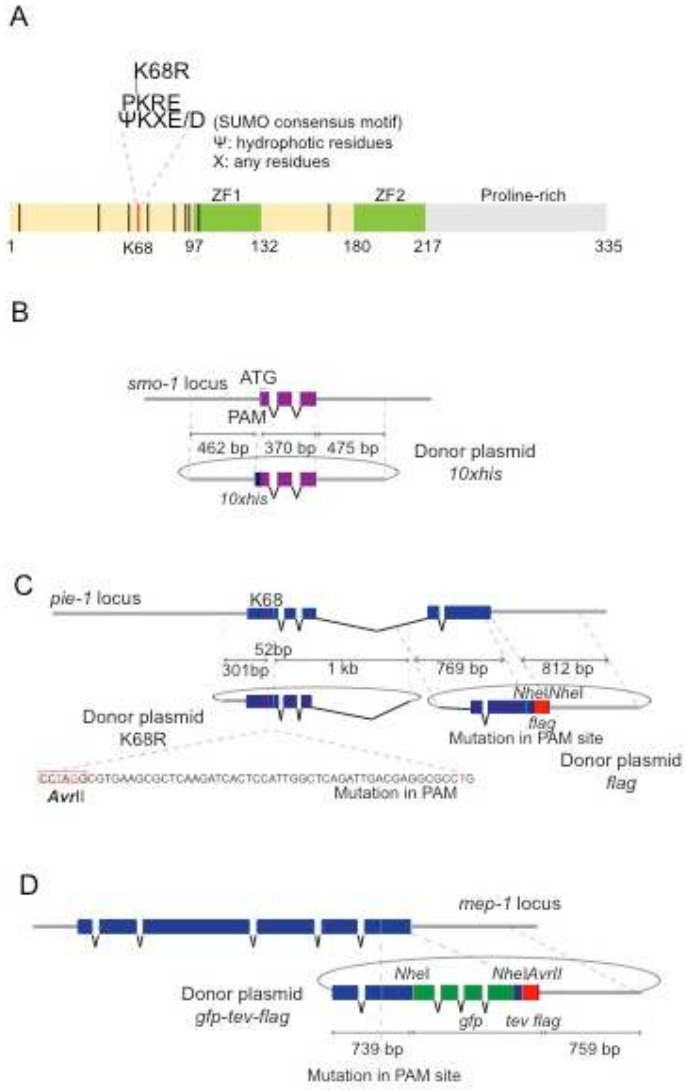
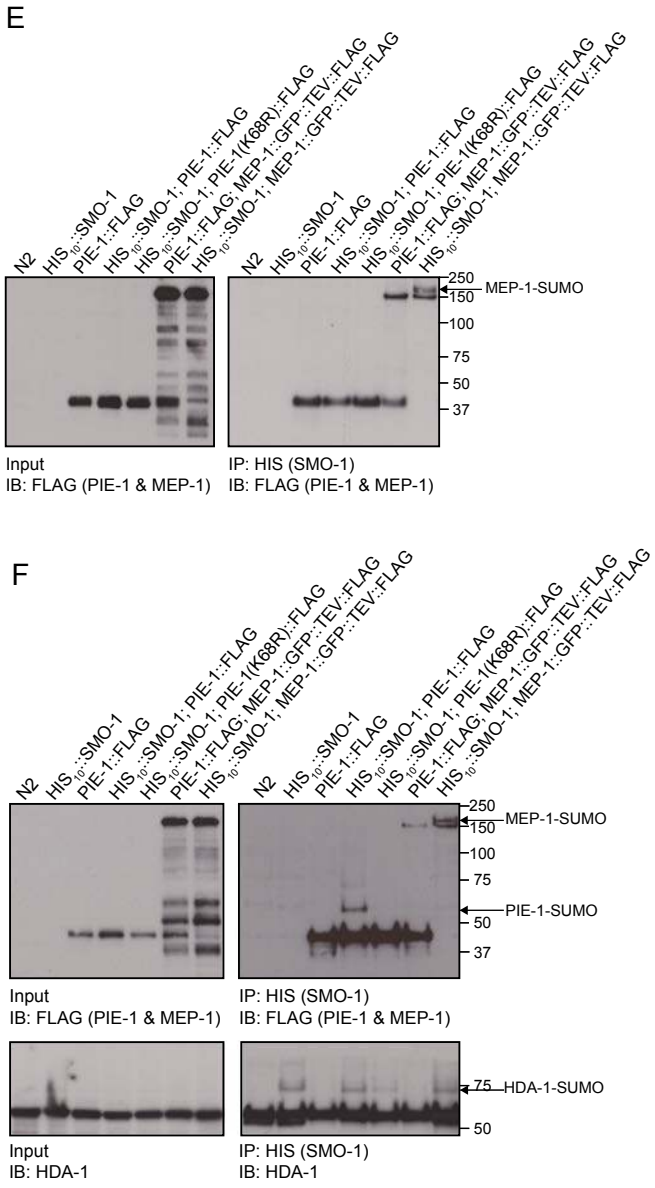


Figure 2.3 PIE-1 is SUMOylated on K68 residue in the *C. elegans* germline
(A) Schematic of PIE-1 containing the consensus SUMO acceptor site (ψKXE), K68. The black bars represent lysine residues on PIE-1 and the K68 is indicated with red bar. (B) Schematic of the *his$_{10}$::smo-1* donor plasmid. The 1*0xhis* coding sequence is inserted immediately after the *smo-1* start codon. (C). Schematic of the *pie-1* donors used to generate either *pie-1(K68)*, *pie-1::flag*, or *pie-1(K68R)::flag*. (D) Schematic of the *mep-1::gfp::tev::flag* donor plasmid



(Figure 2.3 continued)

(E) *In vivo* SUMO purification assay followed by western blotting in embryo lysates. PIE-1 SUMOylation is not detected while MEP-1 SUMOylation is detected (arrow). (F) *In vivo* SUMO purification assay followed by western blotting in adult lysates revealed endogenously SUMOylated PIE-1 and the K68R abolished PIE-1 SUMOylation (arrow). HDA-1 SUMOylation was detected, but dependent on the K68 residue (arrow).

SUMO acceptor site mutant pie-1(K68R) has a weak hypomorphic phenotype

Although *pie-1(K68R)* mutant animals were completely devoid of PIE-1 SUMOylation (Figure 2.3F), the functional consequence of PIE-1 SUMOylation seemed to be different from that suggested by our genetic interactions data because *pie-1(K68R)* mutants did not produce dead embryos like the *pie-1* null mutants (Figures 2.1A and 2.4B). One possible explanation is that the functional effect of PIE-1 SUMOylation on germline specification can be more critical when PIE-1 levels are low, like in the case of heterozygous *pie-1/+*. Consistent with this hypothesis, when we generated single-copy *pie-1* transgenes (Frokjaer-Jensen et al., 2008), *gfp::pie-1* and *gfp::pie-1(K68R)*, where PIE-1 expression level is lower than in endogenously tagged CRISPR line (Kim et al., 2014), the *gfp::pie-1(K68R)* poorly rescued a putative *pie-1* null mutant (zu154) at 5%(n=1250), while the wild-type *gfp::pie*-1 transgene rescued at 93% (n=1222) (data not shown). Alternatively, when PIE-1 levels were normal, the defective phenotype of PIE-1 SUMOylation appeared weak as shown in the *pie-1(K68R)* CRISPR line. In addition, our finding that PIE-1 is present in the nucleus of adult germline cells suggests that PIE-1 functions not only in the germline specification for embryogenesis but also in genomic stability. Thus, we sought to extensively characterize the *pie-1(K68R)* mutant phenotype. Interestingly, we detected a notable frequency of diverse abnormal phenotypes while propagating the *pie-1(K68R)*, such as high incidences of males (him), burst worms, tail

morphogenesis, or sterility (data not shown). To exclude a possibility of mutations caused by off-target effects of CRISPR/Cas9-mediated genome editing, we outcrossed the *pie-1(K68R)* strain several times and compared the resultant lies to other independently generated *pie-1(K68R)* CRISPR lines. We confirmed that 5% of adult progeny displayed different spontaneous mutations in the outcrossed *pie-1(K68R)* (n=6880, data not shown), while this percentage is normally 0.1–0.2% for wild-type (Ahmed and Hodgkin, 2000; Harris et al., 2006; Hodgkin et al., 1979), indicating that the effect we observed in *pie-1(K68R)* was not due to off-target effects. Moreover, the brood sizes of individual *pie-1(K68R)* mutants were notably variable, although the average number of progeny did not show a significant difference compared to wild-type (Figure 2.4A). Interestingly, the size of the *pie-1(K68R)* gonad was obviously smaller than that of wild-type and the number of germ cells was also lower than in wild-type (Figure 2.4C).

The SUMOylation pathway has been suggested to activate DNA damage responses (Boulton et al., 2004; Holway et al., 2006; Kim and Colaiacovo, 2015; Kim and Michael, 2008; Reichman et al., 2018). Therefore, loss of functional SUMO in the *pie-1(K68R)* mutant should result in a more severe defect in genome stability if PIE-1 SUMOylation is indeed crucial for protecting the genomic integrity of the germline. *3xflag::smo-1* CRISPR lines, unexpectedly, displayed a hypomorphic mutation phenotype including a high rate of male progeny, embryonic lethality, and L1 lethality. We observed a significantly reduced brood size when we introduced a *pie-1(K68R)* hypomorphic allele into

the *3xflag::smo-1* strain by crossing, compared to each single mutant, whereas there was no detectable genetic interaction in embryonic lethality (Figures 2.4A and 2.4B).

We further tested the genetic interactions of *pie-1(K68R)* with another crucial factor for genome stability. A beta-nucleotidyl transferase RDE-3/MUT-2, one of the downstream factors in the RNAi pathway, is required for genome stability based on its silencing of transposons and high-copy number transgenes (Chen et al., 2005; Collins and Anderson, 1994). Interestingly, *pie-1(K68R)* showed strong genetic interactions with *rde-3*; for example, the *rde-3(ne3370); pie-1(K68R)* showed a 100% sterile phenotype (data not shown). A germ cell nuclear antigen (GCNA) protein that is homologous to IDR-containing proteins implicated in DNA damage repair is required for genome integrity (G, Dokshin, personal communication) (Carmell et al., 2016). In *C. elegans*, two independent deletion mutants, *gcna-1(ne4334)* and *gcna-1(ne4356)*, were generated using CRISPR/Cas9-mediated genome editing (Carmell et al., 2016). Consistent with the idea that PIE-1 SUMOylation may play a role in protecting germline integrity, we observed a dramatically increased sterile phenotype both in *pie-1(K68R), gcna-1(ne4334)* and in *pie-1(K68R), gcna-1(ne4356)*, while each *gcna-1* mutant alone did not show an apparent phenotype (data not shown). Taken together, these data suggest that PIE-1 SUMOylation is involved in genome stability. The state of genome stability in the adult germline may affect gamete production,

which results in the failure to protect the germline lineage after fertilization in a *pie-1* dosage-dependent manner.

To further investigate the effect of PIE-1(K68R) on germline specification in embryos, we compared PIE-1 expression patterns in the wild-type and K68R mutant. We did not detect any apparent differences between them. However, PIE-1 expression in the P-granule was gradually diminished during embryogenesis (from P_2 to P_4) in the K68R mutant, but not in the wild-type (Figures 2.4D and 2.4E).

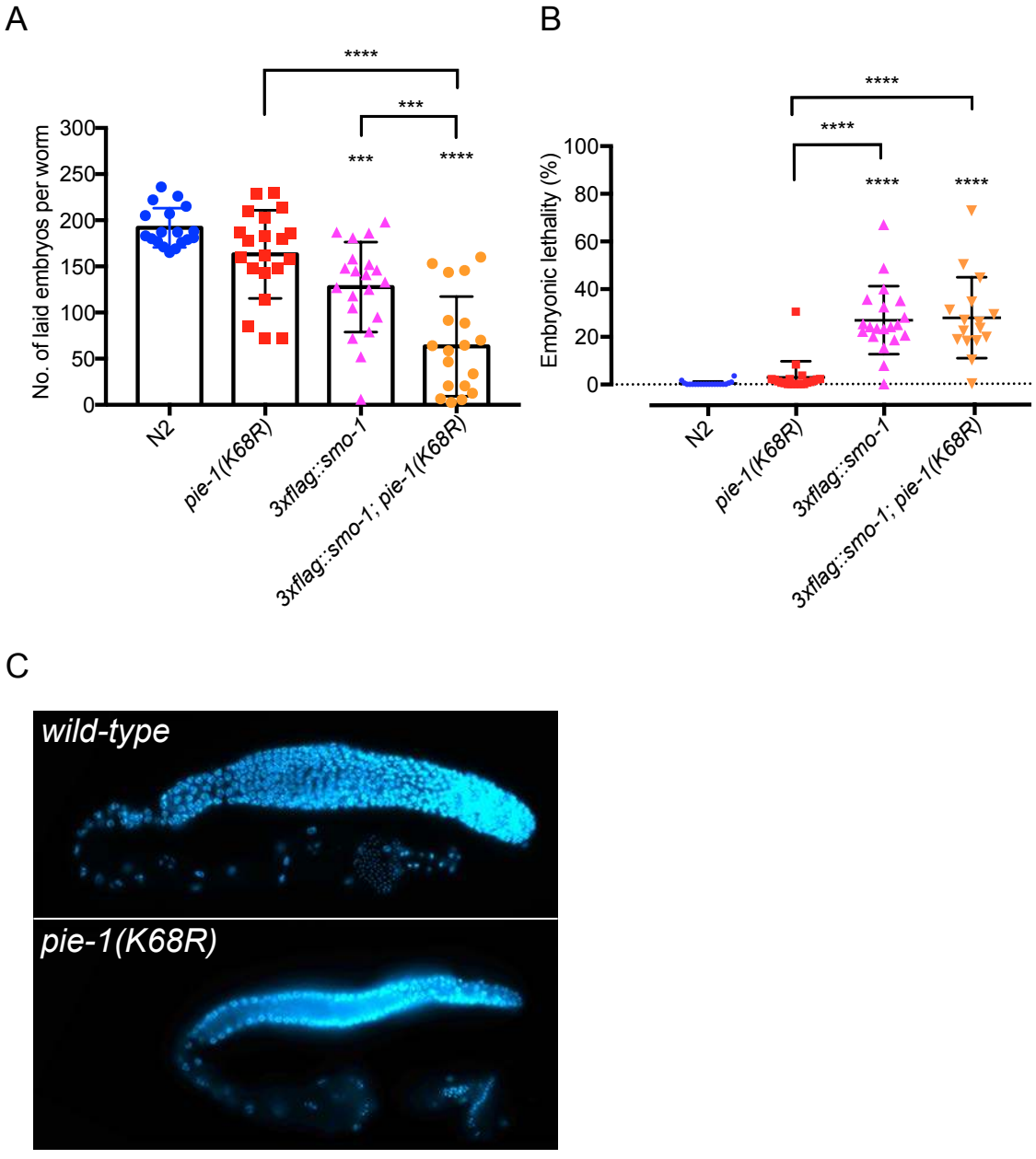
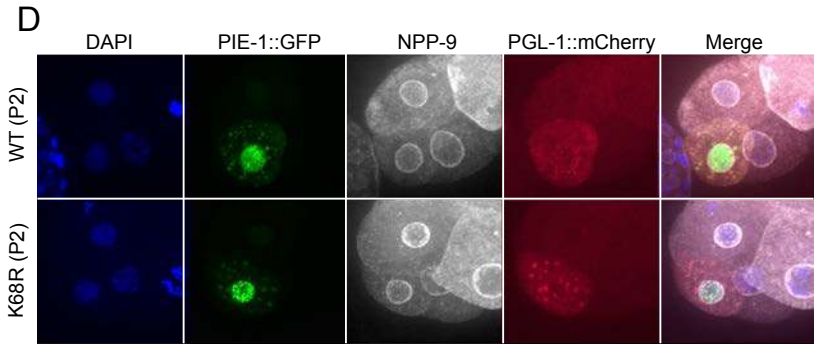
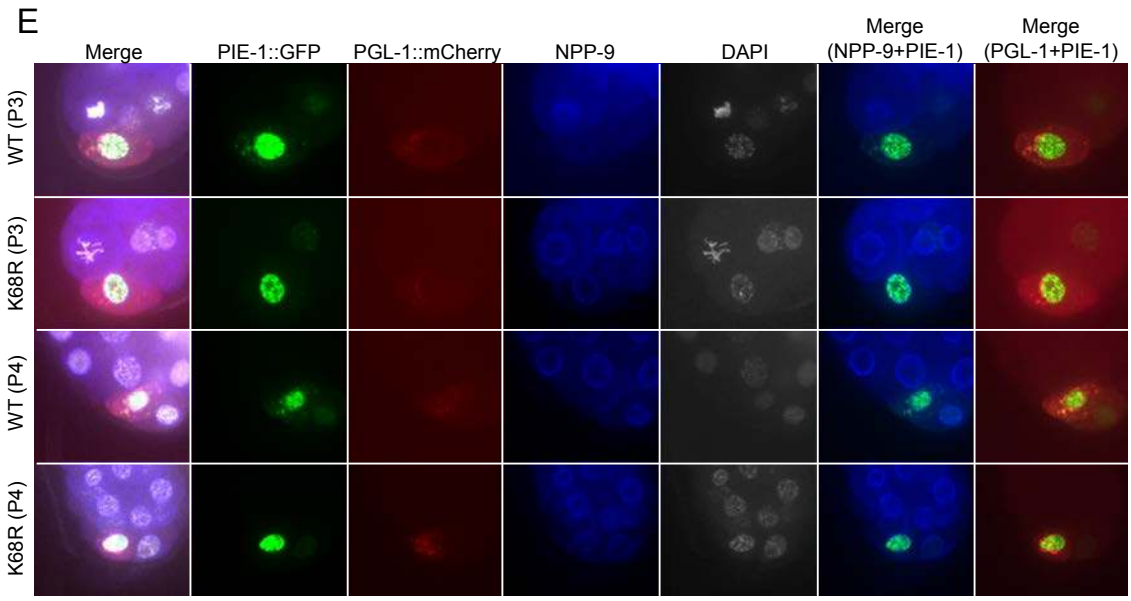


Figure 2.4 Characterization of *pie-1(K68R)*

(A) Brood sizes and (B) Embryonic lethality of N2, *pie-1(K68R)*, *3xflag::smo-1*, and *3xflag::smo-1; pie-1(K68R)*. Two-tailed t-test: ***P<0.0005, ****P<0.00001. (C) Immunofluorescence micrographs of DAPI in adult gonad of WT and K68R.





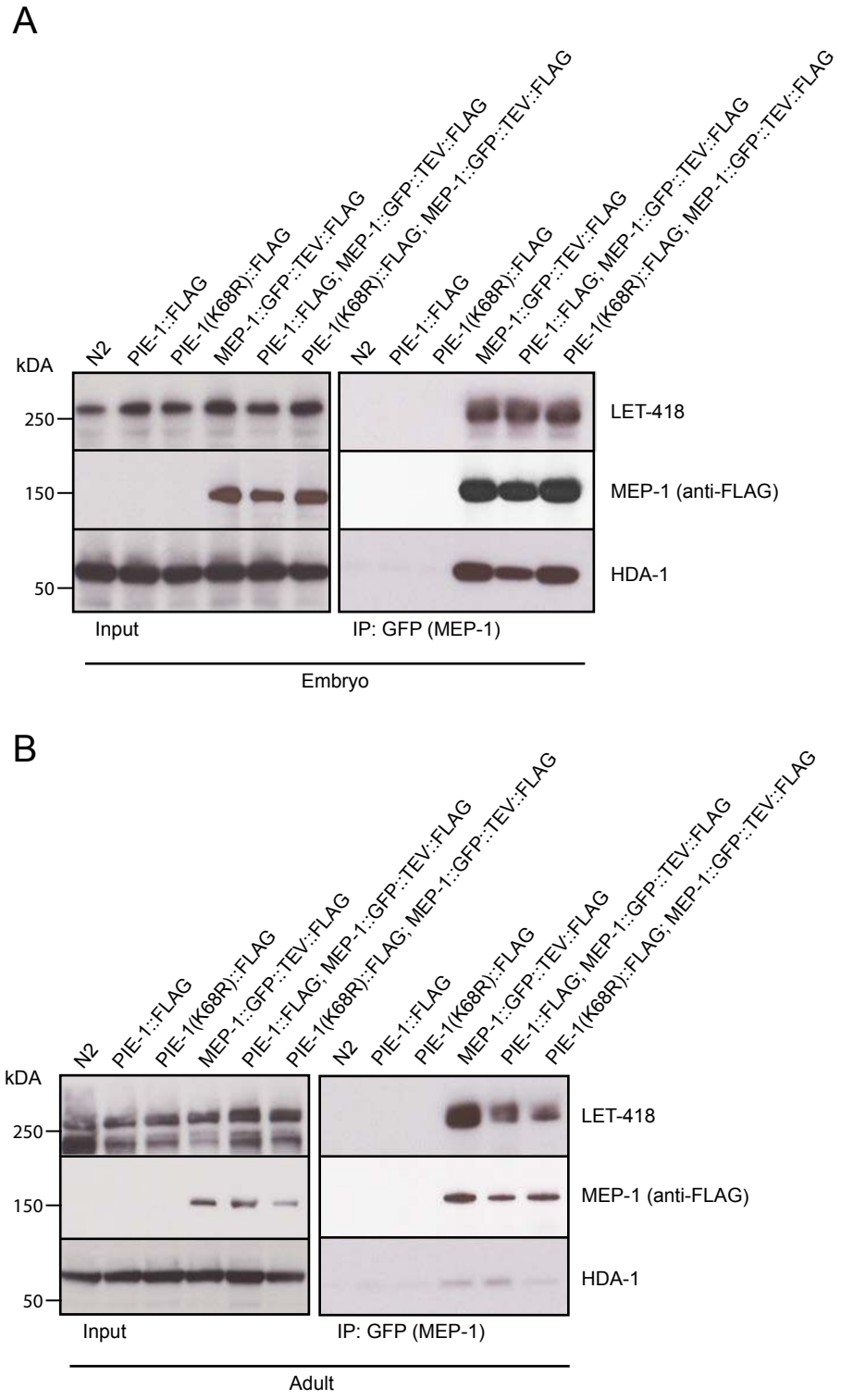
(Figure 2.4 continued)

(D) and (E) Confocal images of embryos expressing PIE-1::GFP (in WT and K68R as indicated) and PGL-1::mCherry, and confocal immunofluorescence micrographs of anti-NPP-9 and DAPI.

PIE-1 SUMOylation facilitates HDA-1 SUMOylation

Given our genetic data and the functional consequence of PIE-1 SUMOylation, we asked whether the genome instability of *pie-1(K68R)* involves regulation of a PIE-1–dependent NuRD complex. Our previous studies suggested that PIE-1 inhibits the activity of the MEP-1 and NuRD complex, including LET-418/Mi-2 and HDA-1/HDAC-1, to protect the MES-dependent germline chromatin state during the early embryogenesis (Unhavaithaya et al., 2002). SUMOylation promotes binding of multiple functionally related proteins by acting as a 'SUMO glue' (Matunis et al., 2006; Shen et al., 2006), and functionally related proteins are SUMOylated together (Golebiowski et al., 2009; Psakhye and Jentsch, 2012; Tammsalu et al., 2014; Tatham et al., 2011). Therefore, PIE-1 SUMOylation also may facilitate SUMOylation of MEP-1, LET-418, or HDA-1. Using a well-validated antibody for HDA-1 detection, we easily detected a slowly migrating HDA-1 band in animals having the *his$_{10}$::smo-1* allele, which indicates the presence of a SUMOylated HDA-1(Figure 2.3F). Surprisingly, the level of SUMOylated HDA-1 was significantly reduced only in the SUMO acceptor site mutant *pie-1(K68R)* (Figure 2.3F), supporting our hypothesis that PIE-1 SUMOylation may enhance SUMOylation of a PIE-1 binding partner. However, MEP-1 SUMOylation was not affected by defective PIE-1 SUMOylation (data not shown) and it was technically impossible to determine whether LET-418 is SUMOylated because the size of unmodified LET-418 is already too big (~250 kDa) for the protein to be detected and distinguished from SUMOylated LET-418.

Next, we addressed whether PIE-1 SUMOylation requires the formation of the functional NuRD complex. However, we encountered technical issues in PIE-1 IP, including the insolubility of nuclear PIE-1 proteins in lysis buffer, the lack of an available PIE-1 antibody that recognizes peptides outside of the SUMO acceptor site, and an epitope tag-cleavage issue during the IP process. In MEP-1 IP, we confirmed previous findings that MEP-1 interacts with the NuRD complex in embryo lysates (Figure 2.5A) (Unhavaithaya et al., 2002). However, as reported previously (Unhavaithaya et al., 2002), PIE-1 was not detected in MEP-1 IP (data not shown). Perhaps, detection of PIE-1 protein is challenging due to the insolubility of PIE-1. The interaction between MEP-1 and the NuRD complex was still intact in embryos of *pie-1(K68R*) mutants (Figure 2.5A). As our SUMO purification data suggested that PIE-1 SUMOylation occurs in the adult germline before fertilization, we further investigated their interactions in adult lysates. In contrast to embryos, HDA-1 was not bound to MEP-1 in adults of *pie-1(K68R)* mutants, while the levels of LET-418 bound to MEP-1 were not changed (Figure 2.5B). Therefore, these data suggest that PIE-1 SUMOylation is required for HDA-1 SUMOylation, which is an essential step in the formation of a functional NuRD complex in the adult germline.



(Figure 2.5 continued)

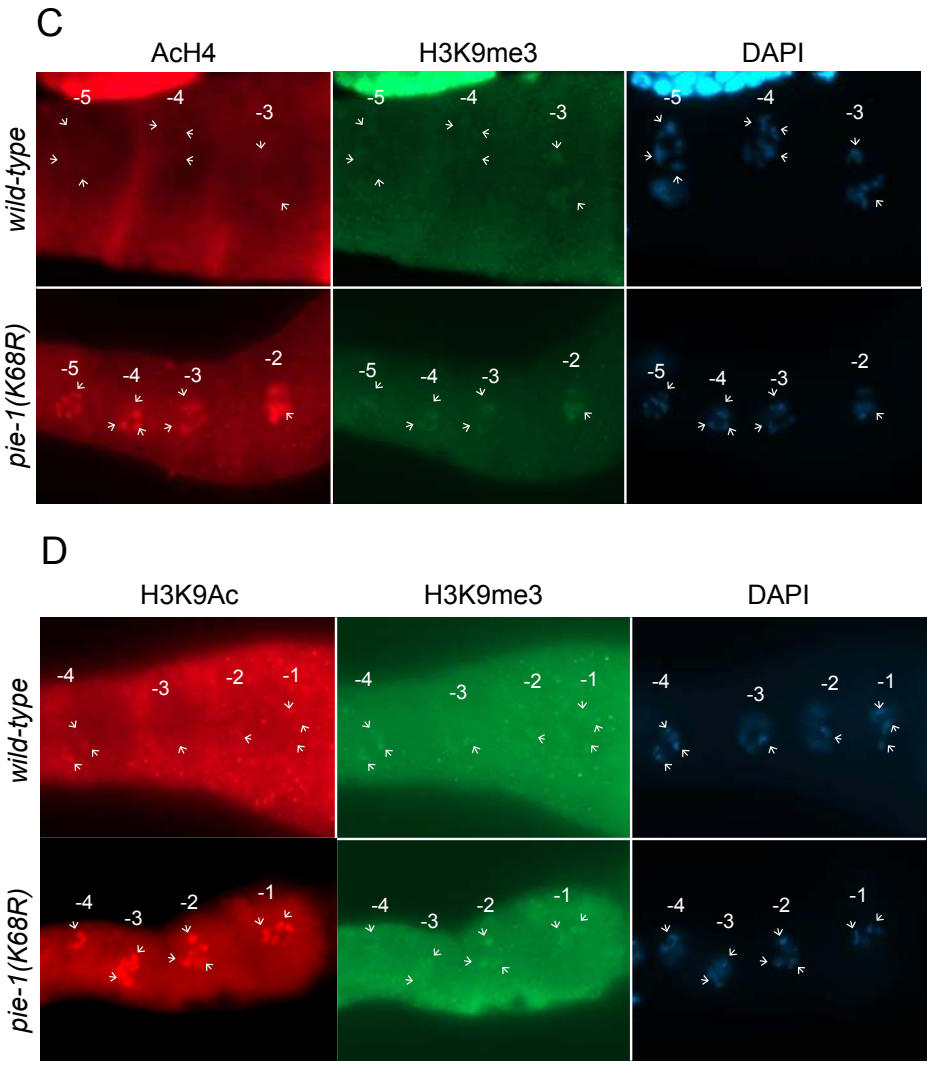


Figure 2.5 PIE-1 SUMOylation is required for activity of histone deacetylase

(A) and (B) Co-immunoprecipitation experiment showing physical interaction between MEP-1 and NuRD complex (LET-418 and HDA-1). Immunoprecipitation was performed in embryo and adult lysates using GBP beads, and the indicated proteins (MEP-1, LET-418, and HDA-1) were detected with anti-FLAG antibody, anti-LET-418 antibody, and anti-HDA-1 antibody, respectively. (C) Immunofluorescence micrographs of anti-H4ac and anti-H3K9me3 double staining in adult oocytes of WT and the K68R mutant. (D) Immunofluorescence micrographs of anti H3K9Ac and anti-H3K9me3 double staining in adult oocytes of WT and the K68R mutant. The arrows indicate stained chromosomes

Increased levels of histone acetylation in pie-1(K68R)

SUMOylation of HDAC1 has been proposed to enhance both its activity and its role in transcriptional repression (Cheng et al., 2004; David et al., 2002; Gill, 2005). On the other hand, HDAC1 SUMOylation enhances the expression of an anti-apoptotic gene by releasing a transcription factor (Tao et al., 2017). Interestingly, intermediate functions of HDAC1 SUMOylation were also reported (Citro et al., 2013; Joung et al., 2018). For example, basal level of SUMOylated HDAC1 represses a transcription factor MyoD, but enhanced HDAC1 SUMOylation changes its binding partners for myogenesis from MyoD to E2F/RB (Joung et al., 2018). These discrepant previous results may imply that the function of HDAC SUMOylation is highly dependent on the physiological context. Given our findings that unSUMOylated HDA-1 did not associate with MEP-1 or LET-418, the NuRD complex may lose its role in transcriptional repression due to increased acetylation level. To test this possibility, we examined histone acetylation levels in the adult germline. In the oocytes, we observed visibly increased acetylation levels of both H4 and H3K9 in the *pie-1(K68R)* mutant compared to wild-type (Figures 2.5C and 2.5D), suggesting that decreased HDA-1 SUMOylation and thus loss of NuRD complex formation represses the histone deacetylase activity of HDA-1.

Many studies have suggested that HDACs play an important role in DNA damage responses by maintaining chromatin structure during DNA repair, inhibiting transcription in the repair region, or directly regulating the important

repair factors (Dobbin et al., 2013; Gong and Miller, 2013; Hsiao and Mizzen, 2013; Miller et al., 2010; Nikolova et al., 2017; Stengel and Hiebert, 2015). As a consequence of HDAC inhibition, accumulation of double-strand breakes, enhanced apoptosis, and genomic instability have been reported (Nikolova et al., 2017; Robert and Rassool, 2012; Thurn et al., 2013). The phagocytic receptor cell death abnormal 1 (CED-1) is a component of the apoptotic pathway and initiates a signaling pathway for engulfment of apoptotic cells in *C. elegans (Yu et al., 2008; Zhou et al., 2001)*. Therefore, we examined whether *pie-1(K68R)* increases the number of apoptotic cells in the adult germline. Interestingly, *pie-1(K68R)* mutant gonads displayed a 2-fold increase in CED-1::GFP-expressing cells that represent engulfing cells undergoing apoptosis (Lu et al., 2009; Zhou et al., 2001), compared to wild-type (Figures 2.6A and 2.6B). In addition, we detected that 3% of scored gonads (n=155) in the distal tip of *pie(K68R)* contained an abnormally high number of apoptotic cells, whereas the wild-type did not show any apoptosis in the distal tip (0%, n=109, Figures 2.6A and 2.6B). These findings strongly support our idea that the increased number of apoptotic cells and genomic instability in the *pie-1(K68R)* mutant may be due to HDA-1 inhibition by loss of SUMOylation.

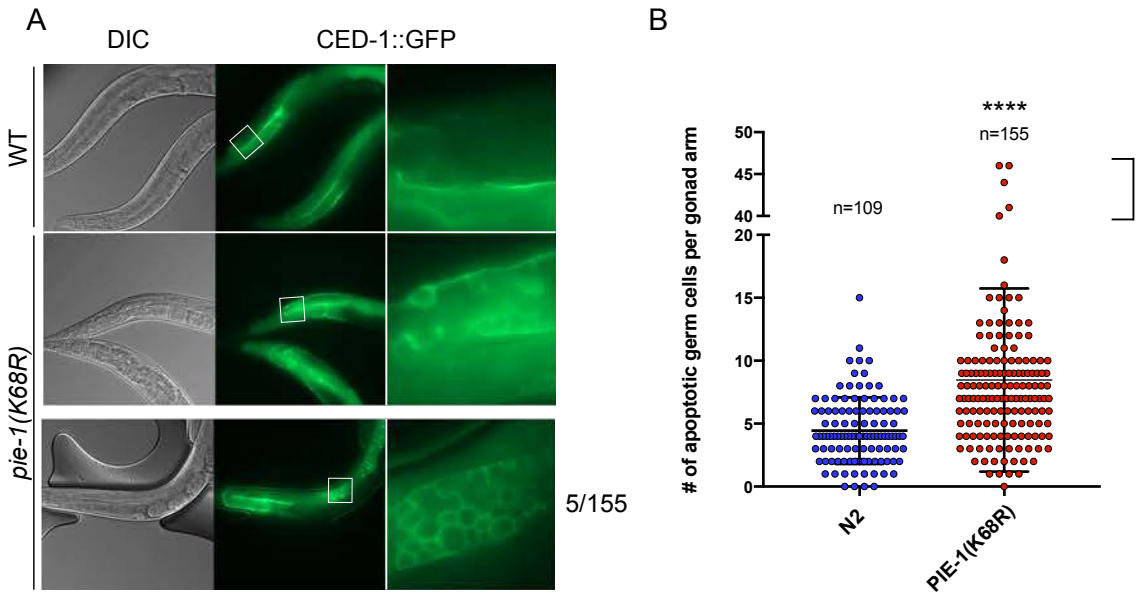


Figure 2.6 Increased number of apoptotic cells in the adult gremlin of *pie-1(K68R)*. (A) GFP fluorescence micrographs of adult germline expressing CED-1::GFP, an apoptosis marker, to show an example image in WT and *pie-1(K68R)*. In the K68R mutant, most germ cells in the distal tip express CED-1 ::GFP in 3% of scored gonads (5/155). (B) The quantification of the number of apoptotic cells using CED-1 GFP per gonad arm. The square bracket indicates animals showing a higher number of apoptotic germ cells in the distal tip. Two-tailed t-test: ****P<0.0001.

Discussion

The results of the present study demonstrate that PIE-1 protein is a SUMO substrate, and their interaction is required for NuRD complex formation and the activity of histone deacetylase in the *C. elegans* germline. We found that PIE-1 protein was present in the nuclei of pachytene-stage germ cells and oocytes in the adult germline, implying that PIE-1 may function in meiotic progression of germ cells. Our *in vivo* SUMO purification assay revealed that PIE-1 was SUMOylated in the adult germline, but not in embryos, and the K68 residue of PIE-1 was a SUMO acceptor site. The SUMO acceptor site mutant *pie-1(K68R)* decreased HDA-1 SUMOylation and abolished MEP-1 interaction with the NuRD complex. In addition, depletion of PIE-1 SUMOylation caused an increase in the number of apoptotic germ cells in the pachytene region, and occasionally, most mitotic germ cells underwent apoptosis in the *pie-1(K68R)* mutant. In conclusion, our study suggests that a novel function of PIE-1 in the adult germline is to protect a NuRD complex-mediated epigenetic state required for genome stability, which may affect germline specification after fertilization.

Our genetic experiments suggest that SUMOylation is required for PIE-1 functions in germline specification during embryogenesis. In contrast to this strong genetic interaction data, the SUMO acceptor mutant *pie-1(K68R)*, however, resulted in a very mild phenotype at a glance. Like with the *pie-1(K68R)* mutation, mutation of most SUMO substrates has resulted in no notable phenotype, and thus, the functional consequences of their SUMOylation remain

enigmatic. Recently, a reasonable explanation for the paradoxical discrepancy between the strong phenotype of SUMO pathway mutants and the mild phenotype of a single SUMO substrate mutant has been proposed: SUMOylation targets a functional protein group rather than one specific protein, and thus, the SUMOylated proteins act together to trigger synergistic effect on an important biological function (Golebiowski et al., 2009; Psakhye and Jentsch, 2012; Tammsalu et al., 2014; Tatham et al., 2011). Therefore, if we could identify and mutate all SUMO substrates in the same functional group with PIE-1, we might detect a strong phenotype like that of the *pie-1* null mutant. Our current results at least indicate that HDA-1 is a factor in the same functional group as PIE-1 in the *C. elegans* germline.

In contrast to our finding that histone acetylation levels are increased in the *pie-1(K68R)* mutant compared to the wild-type*,* PIE-1 has been suggested previously to inhibit HDA-1 activity (Unhavaithaya et al., 2002). In addition, ectopic overexpression of PIE-1 (via the heat-shock promoter, *hsp16-1*, (Seydoux et al., 1996)) in somatic tissues induces a *synMuv* phenotype like the *mep-1* mutant phenotype in a *lin-15A* background, suggesting that PIE-1 antagonizes the MEP-1 and NuRD complex (Unhavaithaya et al., 2002). However, an engineered strain expressing a single-copy *hsp::pie-1* (Frokjaer-Jensen et al., 2008) did not show any *synMuv* phenotype, although overexpression of PIE-1 was confirmed by western blotting (data not shown). Thus, we were concerned that the result previously reported by Unhavaithaya et

al. might have been obtained due to random integration of extrachromosomal array *hsp::pie-1* into an unknown genomic region that disrupts a functional gene. Nevertheless, our genetic data still appear to support the previous model that PIE-1 protects the germline fate from MEP-1 and NuRD complex–induced chromatin remodeling by antagonizing the MEP-1 and NuRD complex. For example, the *pie-1(K68R)* mutation suppressed ~26% of the L1 lethality caused by mep-1 loss of function (data not shown). One possible model is that the functions of PIE-1 differ between the adult germline and the germline blastomeres of zygotic embryos, and these different functions are regulated by the SUMOylation pathway. Alternatively, the integrity of the adult germline may affect germline specification in the embryos. Therefore, we propose that in the adult germline, PIE-1 SUMOylation is required for the formation of a functional NuRD complex to create proper epigenetic information before fertilization, which is required for inhibiting somatic gene transcription after fertilization, and to protect genomic stability from DNA damage by facilitating HDA-1 SUMOylation (and SUMOylation of possibly other proteins), which is required for a DNA repair pathway.

Considering our finding that the *pie-1(K68R)* mutant shows an increased level of histone acetylation, we will prioritize studies to determine whether loss of PIE-1 SUMOylation affects transcription in the germline. In addition, if we determine and mutate a SUMO acceptor site of MEP-1, HDA-1, or another possible component in the same functional group, we will be able to address

whether they act together with PIE-1 to trigger a synergistic effect on germline specification in the embryos.

CHAPTER III: GENERATION OF *pie-1(DAQMEQT)::gfp* CRISPR LINES AND *IN VIVO* MODEL STUDY FOR TRANSCRIPTIONAL REGULATION

Introduction

In *C. elegans*, maternally loaded PIE-1 functions as a crucial germline determinant during early embryogenesis (P_1–P_4). PIE-1 depletion causes P_2 germline blastomere to adopt the fate of its sister blastomere, EMS (Mello et al., 1992). This dramatic transformation of cell fate has been an intriguing subject of investigations to understand how animals protect their germline lineages from inappropriate somatic differentiation. Seydoux et al first found that mRNA transcription is silenced in the early germline blastomere (Seydoux et al., 1996). In addition, Ser2 phosphorylation of a pol II elongation marker is not detected in the early germline blastomeres but appears when *pie-1* is absent (Seydoux and Dunn, 1997). Studies using a human cell culture system suggested that PIE-1 contains a mimic sequence (YAPMAPT) that resembles the heptapeptide repeat sequence (YSPTSPS) of pol II CTD, where phosphorylation occurs for elongation (Ser2) and initiation (Ser5), and thus, PIE-1 competitively targets and sequesters CTD kinase away from pol II (Batchelder et al., 1999). In addition, the YAPMAPT mutation (DAQMEQT) rescues the *pie-1* null allele at a reduced frequency compared to wild-type (Batchelder et al., 1999), indicating that the YAPAMPT sequence is important for PIE-1 functions.

Previously, a strain engineered to carry a single copy or low copy number of DAQMEQT transgenes via ballistic transformation was used to test the importance of the YAPMAPT sequence in transcriptional repression (Ghosh and Seydoux, 2008). However, studies using these transgenes suggested that the

YAPMAPT motif is essential for inhibition of Ser2 phosphorylation but not for transcriptional repression (Ghosh and Seydoux, 2008). We can now introduce the transgene at an endogenous locus using CIRSPR/Cas9-mediated genome editing, which allows us to monitor the mutant phenotype in the biological context.

Results

We generated the DAQMEQT CRISPR lines to explore the mechanism by which PIE-1 functions in germline specification as demonstrated in the previous studies. We also introduced either GFP or FLAG epitope tag to the C terminus of the PIE-1 genomic locus, thereby creating *pie-1(DAQMEQT)::gfp* and *pie-1(DAQMEQT)::flag*. The DAQMEQT mutant showed a maternal-effect sterile phenotype at a high frequency (72%) but not embryonic lethality like a pie-1 null phenotype (Figure 3.1A). For the most part, the P2 blastomere was properly specified with little sign of cell fate change as expected from a pie-1 null allele. Because the YAPMAPT sequence was proposed to be a competitive non-phosphorylatable inhibitor of CTD kinase (Batchelder et al., 1999), we reasoned that lowering CDK12 activity might suppress the DAQMEQT mutant. While the p-TEFb complex, CDK-9/cyclin T, has been thought to be required for CTD Ser2 phosphorylation (Ghosh and Seydoux, 2008; Wood and Shilatifard, 2006; Zhang et al., 2003), another Ser2 kinase complex, CDK-12/cyclin K was identified in *C. elegans*, *Drosophila*, and human (Bartkowiak et al., 2010; Bowman et al., 2013). In the *C. elegans* germline, Ser2 phosphorylation requires CDK-12, but not CDK-9 (Bowman et al., 2013). To test this possibility we used a *cdk-12-analog-sensitive* (*-as*) allele that has a single amino acid mutation (F383G) in the ATP-binding pocket created by CRISPR/Cas9-mediated genome editing (gift from Hermand Damien, University of Namur, Belgium). The *as* allele can bind to a bulky ATP analog, thereby blocking kinase activity when an ATP analog drug is

added (Lera and Burkard, 2012). Strikingly, we found that even without the drug, the homozygous *cdk-12-as pie-1(DAQMEQT)::gfp* showed a significantly suppressed sterile phenotype to 2% (instead of 72% in the DAQMEQT) (Figure 3.1A), strongly supporting the previous model that the YAPMAPT motif is responsible for the inhibition of CTD kinase. With drug treatment, *cdk-12-as* showed a fully penetrant L1 arrest phenotype like *cdk-12* (RNAi). The DAQMEQT mutant also caused significant suppression of larval arrest as with the *cdk-12-as* allele. Instead of only 2% of *cdk-12-as* mutants escaping and reaching adulthood after treatment with the 1μM ATP analog dose, we found that 23% of the *cdk-12-as, pie-1(DAQMEQT)::gfp* animals reached adulthood. These findings clearly suggest a relationship between PIE-1 and CDK-12.

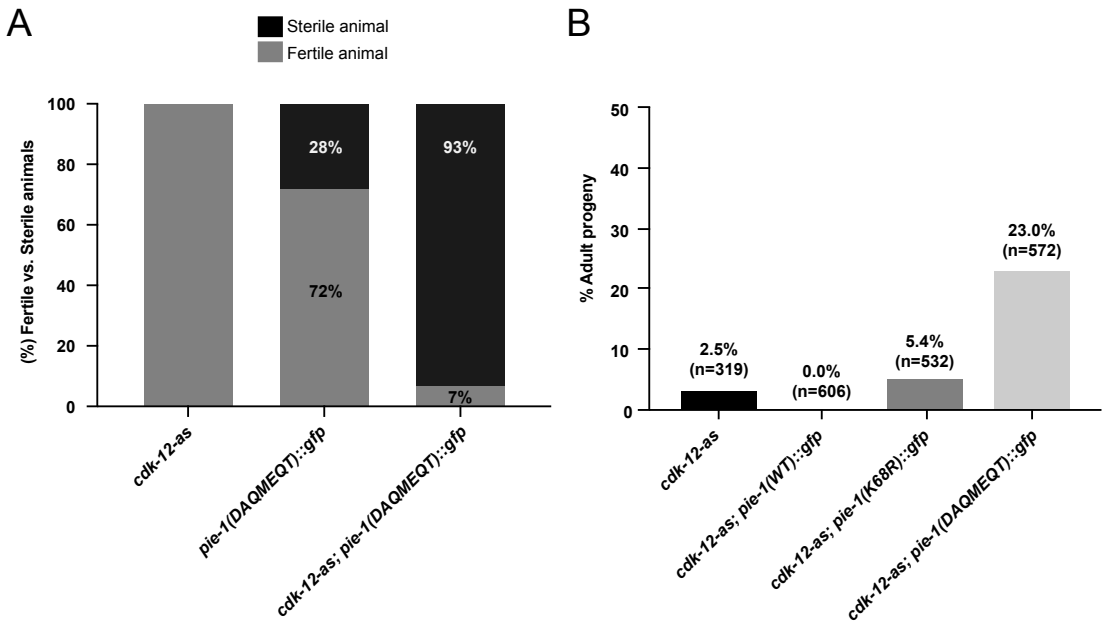
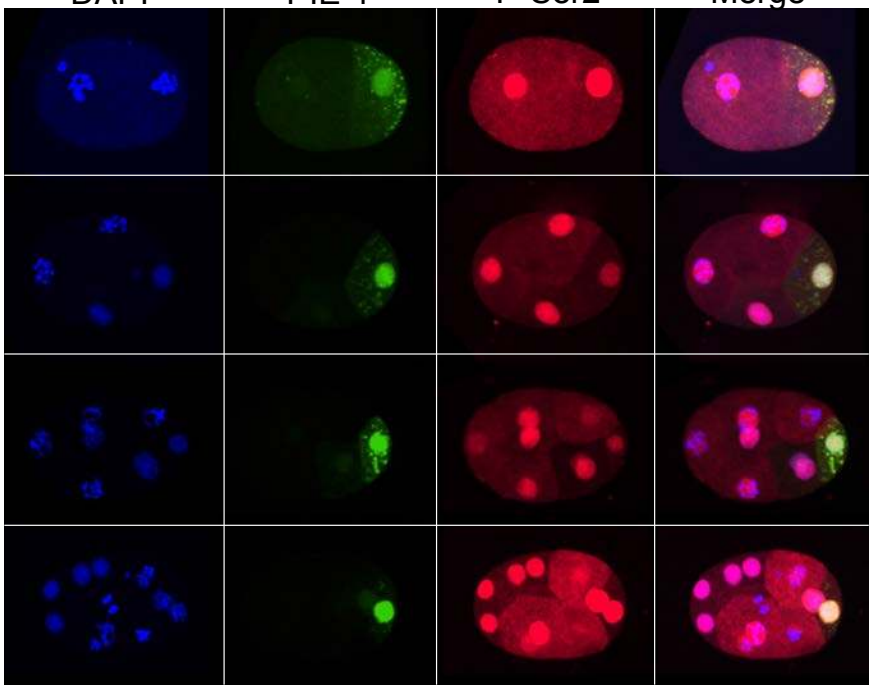


Figure 3.1 Genetic interaction between *pie-1(DAQMEQT)::gfp* and *cdk-12-as*
(A) *cdk-12-as* allele suppresses a sterile phenotype of the DAQMEQT mutant. (B) The DAQMEQT mutant rescues L1 arrest phenotype of *cdk-12-as* in 1 μM of ATP-analog drug, 3MB-PP-1 (Toronto Research Chemicals, A602960). 100 μM of stock solution was prepared in DMSO and added to the NGM agar before pouring plate and OP50 containing 1 μM of ATP-analog drug was seed to the plates. Synchronous L3-L4 animals were placed to the NGM plate containing 1μM of ATP-analog drug.

Accordingly, we expected that Ser2 phosphorylation would be altered in the DAQMEQT mutant as previously published (Ghosh and Seydoux, 2008). The antibody we used (3E10) is well known to be specific for pol II in *C. elegans* (Bowman et al., 2013; Furuhashi et al., 2010). As expected, we observed that significantly reduced Ser2 phosphorylation in *cdk-12* (RNAi) embryos using this antibody (Figure 3.2C) (Bowman et al., 2013). However, unexpectedly we could not find any significant difference between germline blastomeres and somatic blastomeres on immunostaining using this Ser2 phosphorylation-specific antibody in the wild-type embryos (Figure 3.2A).

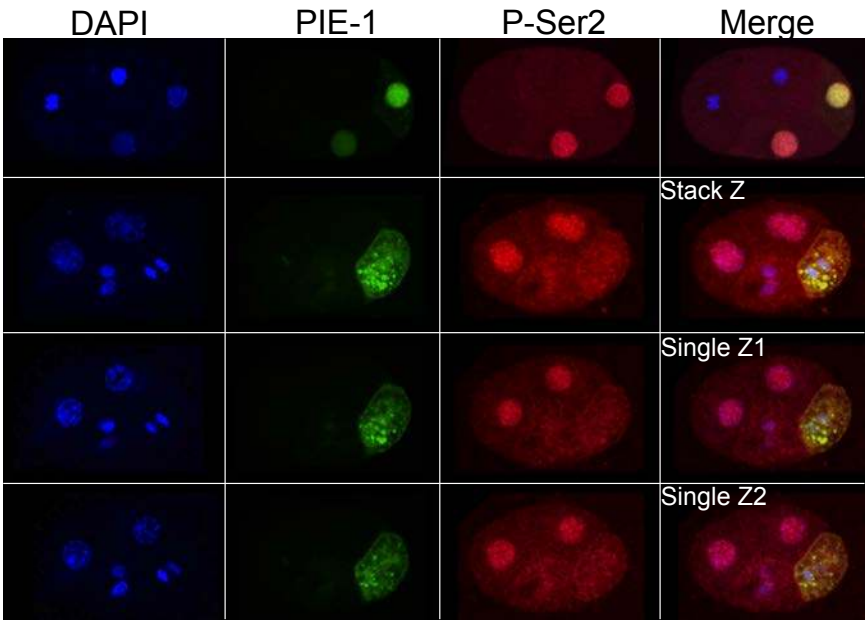
A

pie-1(WT)::gfp



B

pie-1(DAQMEQT)::gfp



C

pie-1(WT)::gfp; cdk-12(RNAi)

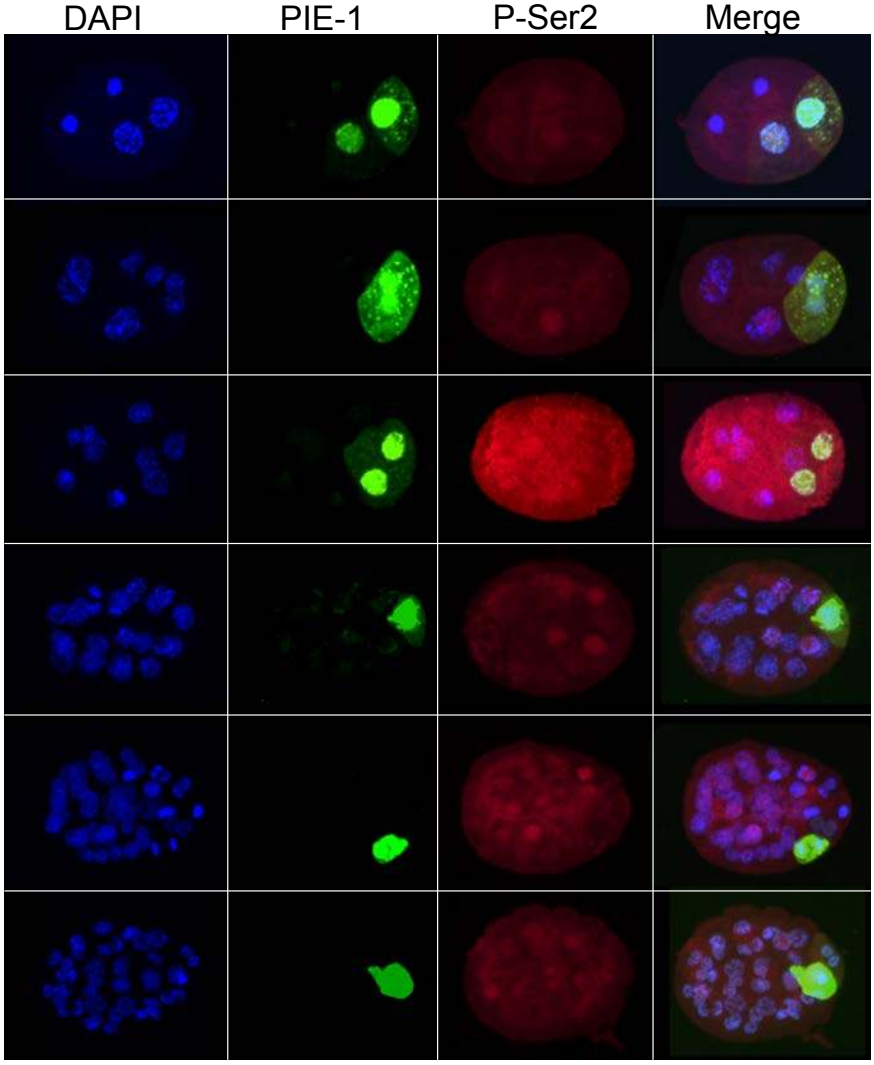
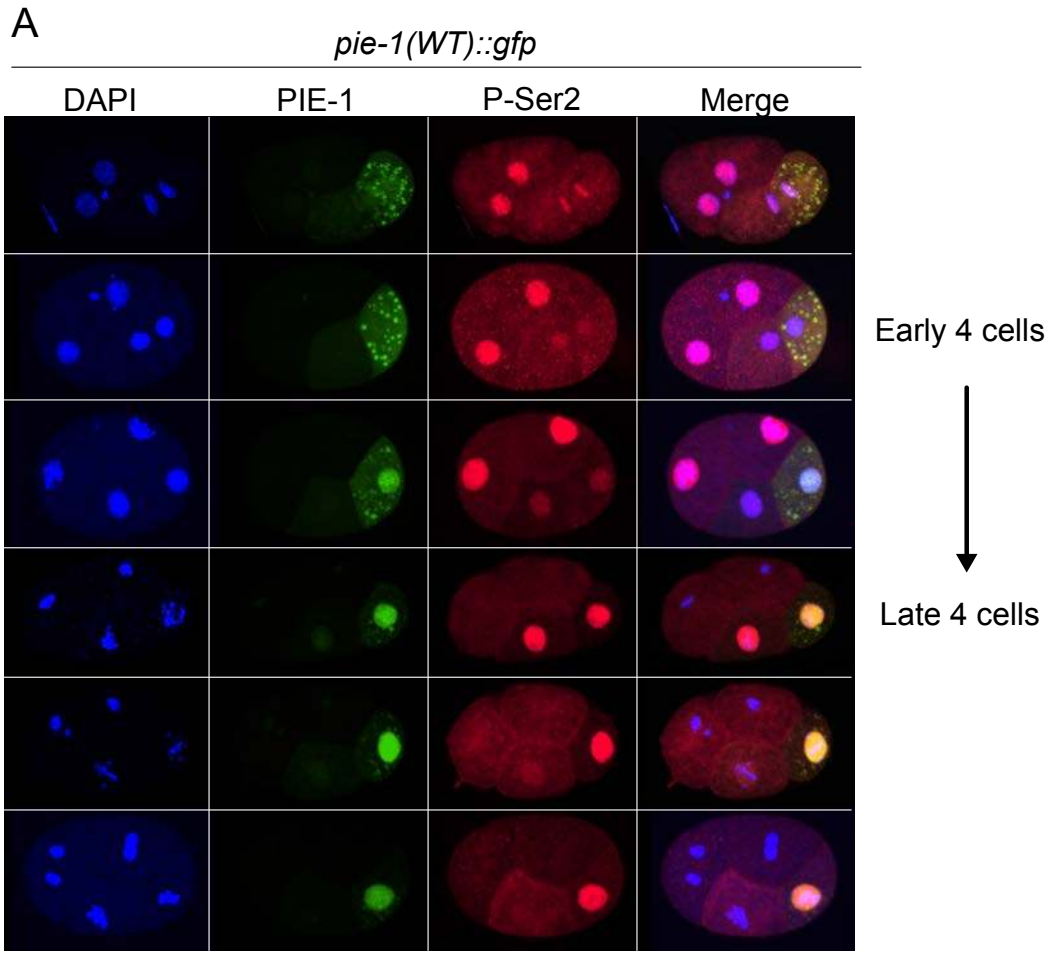


Figure 3.2 Ser2 phosphorylation staining in germline blastomeres

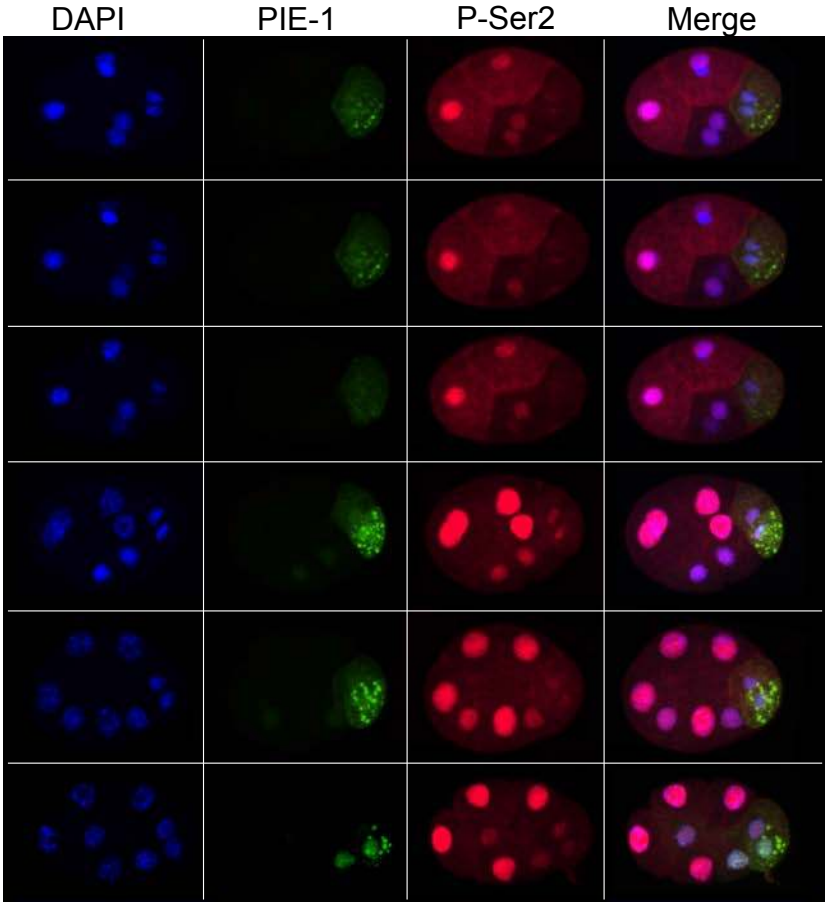
Confocal fluorescence micrographs of embryos expressing (A) PIE-1::GFP, (B) PIE-1(DAQMEQT)::GFP, and (C) PIE-1::GFP in *cdk-12*(RNAi) and confocal immunofluorescence micrographs of anti-Ser2P in (A) *pie-1::gfp*, (B) *pie-1(DAQMEQT)::gfp*, and (C) *pie-1::gfp; cdk-12*(RNAi).

In the germline blastomere, as in all cells, we observed pronounced cell cycle–dependent regulation of the nuclear epitope detected by this antibody. The immunostaining signals from antibodies was dispersed in the cytoplasm during nuclear breakdown and then became brighter gradually in the newly formed daughter nuclei over time (Figure 3.3) These findings raise concern that the previous claims could have been mistaken due to a timing issue, as the germline blastomere divides slower. For example, if we observed enough embryos, we found that the germline blastomeres were dimmer than their sisters due to this timing difference (Figure 3.3).


A
pie-1(WT)::gfp
DAPI
PIE-1
P-Ser2
Merge
Early 4 cells
Late 4 cells

B

pie-1(WT)::gfp



C

pie-1(WT)::gfp

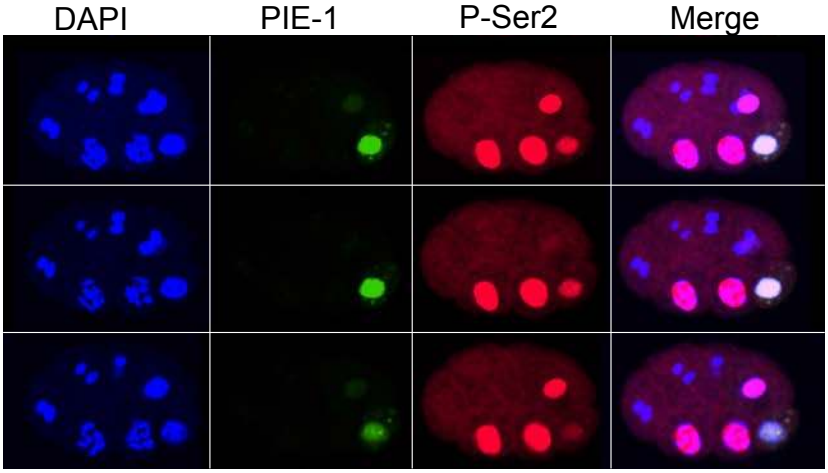


Figure 3.3 Ser2 phosphorylation staining is dependent on cell cycle
Confocal fluorescence micrographs of PIE-1::GFP and confocal immunofluorescence micrographs of anti-Ser2P from 4cell embryos to 12 cell embryos.

Discussion

While our immunofluorescence results detecting Ser2 phosphorylation in the wild-type germline blastomeres is controversial due to a discrepancy with a large body of evidence accumulated over the last 20 years, another CRIPSR strain expressing PIE-1(R109L)::GFP, an escape from ZIF-1 target, appears to support our observations. ZIF-1 interacts with CCCH finger proteins including PIE-1 to degrade them in somatic cells (DeRenzo et al., 2003). While the first CCCH zinc fingers (ZF1s) from PIE-1, MEX-1, and POS-1 and ZF2 from MEX5 have been identified as ZIP target proteins (DeRenzo et al., 2003), a critical residue in the ZF domain that requires this interaction remains to be determined. In an attempt to find the important residue, we aligned all ZF domains and identified the R109 residue as conserved among the ZIF-1–interacting domains above, but not in non-ZIF-1 target regions (data not shown). *pie-1(R109L)::gfp* mutants showed PIE-1 expression in the nuclei of somatic blastomeres (data not shown), indicating the R109 is an important residue for ZIF-1 targeting. However, the somatic expression of PIE-1 did not impair embryogenesis at all, and we did not detect any defect throughout animal development, which is not consistent with the embryonic lethal phenotype of *zif-1*(RNAi) (DeRenzo et al., 2003) and an idea that PIE-1 represses transcription globally.

Nevertheless, based on previous *in situ* hybridization studies (Seydoux et al., 1996), germline blastomeres definitely exhibit limited transcription compared to somatic blastomeres and pie-1–depleted germline blastomeres. Moreover, in

our genetic interaction investigation between *cdk-12* and the DAQMEQT mutant, a mutually rescuing result strongly suggested that PIE-1 is involved in the regulation of germline transcription via interaction with pol II CTD kinase, but not essential for transcriptional repression. One possible explanation is that maternally loaded epigenetic information to which PIE-1 SUMOylation contributes in the adult germline gives rise to a selectively limited transcription for germline specification and the YAPMAPT motif of PIE-1 interacts with CDK-12 and/or other pol II binding partners to ensure transcriptional inhibition in the germline blastomere. Consistent with this possibility, although we tried to generate a *pie-1* CRISPR strain containing both K68R and DAQMEQT by introducing the DAQMEQT sequence to the K68R mutant animals, a heterozygous DAQMEQT allele already caused embryonic lethality in the K68R homozygotes.

CHAPTER IV: DISCUSSION

The goal of this research was to understand how PIE-1 functions in germline specification. The Co-CRISPR strategy described in chapter I allowed us to generate all animal models with the desired alleles for this research. Because the genes of interest were modified by the CRISPR-Cas9 system at their endogenous loci, this research using the CRISPR lines could more accurately reflect the real biological situation, compared to using conventional transgenic approaches. For example, we demonstrated the nuclear localization of PIE-1 in the oocytes and meiotic germ cells of the *pie-1::gfp* CRISPR lines, while the previous transgenic lines did not show PIE-1 expression in the nuclei of adult germ cells (Reese et al., 2000). The first detection of PIE-1 localization in the adult germline was the crucial step in gaining new insight into how PIE-1 functions in protecting germline integrity for gamete production. In chapter II, we determined the molecular mechanism by which PIE-1 plays a role in germline integrity. We showed that PIE-1 is SUMOylated in the adult germline, and its SUMOylation facilitates both HDA-1 SUMOylation and NuRD complex formation. Furthermore, the loss of PIE-1 SUMOylation results in increased levels of histone acetylation. Taken together, our results presented in chapter II suggest that PIE-1 SUMOylation is important for preserving a NuRD complex-mediated proper epigenetic state in the adult germ line, which may affect germline integrity. However, we still do not have a clear answer as to how PIE-1 regulates transcription to maintain germline fate in the early embryos.

A Role for the YAPMAPT sequence in Transcriptional Regulation

PIE-1 inhibits both Ser2 phosphorylation and Ser5 phosphorylation of pol II CTD (Seydoux and Dunn, 1997), which results in repression of pol II elongation and pol II initiation, respectively. Previously, a model was suggested in which PIE-1 inhibits CTD Ser2 phosphorylation of Pol II by sequestering p-TEFb using the CTD-like motif YAPMAPT (Batchelder et al., 1999; Ghosh and Seydoux, 2008). In an effort to confirm this previous model, we generated and characterized the YAPMAPT mutant, *pie-1(DAQMEQT)* CRISPR strains, in the experiments presented in chapter III. Although the *pie-1(DAQMEQT)* homozygotes produced many sterile progenies and no dead embryos were found, the *pie-1(DAQMEQT)/pie-1(zu154)* strain, which has one copy of *pie-1(DAQMEQT)* and one copy of *pie-1* null allele, gave rise to all dead embryos with the *pie-1* mutant phenotype. Thus, one functional copy of the *pie-1* gene is enough for viability similar to that for wild-type animals. In addition, as discussed in chapter III, one copy of the *pie-1(DAQMEQT)* allele also caused embryonic lethality in the K68R homozygotes when we introduced the DAQMEQT sequence to the K68R mutant animals via the CRISPR-Cas9 system. These results imply that the YAPMAPT sequence may be a partially redundant with the K68R and/or other functional *pie-1* alleles.

PIE-1 Cleavage Event

The molecular mechanisms that determine how PIE-1 regulates CTD Ser5 phosphorylation of pol II still remain to be studied. The state of epigenetic marks that are maternally loaded during gamete production can be a crucial setting for the dynamic interplay among many components that inhibits or activates pol II. Consistent with this idea, PIE-1 is expressed and SUMOylated in the adult germline cells before fertilization. Perhaps, the PIE-1 SUMOylation may be required for inhibiting pol II initiation complex (PIC) formation by recruiting cofactors involved in epigenetic regulation. In an attempt to test these possibilities, we performed IP-multidimensional protein identification technology (MuDPIT, a mass spectrometry-based approach) proteomics on PIE-1. A C-terminal GFP fusion PIE-1 was purified using GBP beads that had been previously used to successfully purify other GFP fusions proteins including MEP-1 (in chapter II), UBC-9, and CDK-12. We found that PIE-1 solubility was markedly lower in IP buffer (see the Materials and Methods section in chapter II) than that of other proteins. Most purified PIE-1 protein from the soluble PIE-1 in the IP buffer, unexpectedly, was cleaved on the C-terminal region of PIE-1. We further investigated the exact cleavage site by mass spectrometry-based proteomics and found that the serine 327 residue is cleaved by AC3.5 (in collaboration with Shan Lu, National Institute of Biological Sciences, Beijing, China), which is predicted to have metallopeptidase activity based on protein domain information (http://www.wormbase.org, release WS264, date 08 march

2018). We next substituted the serine residue with alanine using CRISPR/Cas9-mediated genome editing and examined whether the S327A mutation abolished PIE-1 cleavage. Indeed, the cleavage event was completely absent in the S327A mutant. Interestingly, we found that both K68R mutation and depletion of SMO-1 by RNAi reduced the PIE-1 cleavage event somehow. Perhaps, the proteolytic cleavage event also can be regulated by PIE-1 SUMOylation. It will be interesting in the future to determine the functional consequence of the PIE cleavage event.

MEP-1/NuRD Function in Maintaining Germline–Soma Distinction

MEP-1 was previously identified as a PIE-1 interactor in a yeast two-hybrid screen (Unhavaithaya et al., 2002). Consistent with this result, we also found that MEP-1 interacts with PIE-1 in a yeast two-hybrid screen described in chapter II. Depletion of MEP-1 by RNAi induces a L1 larval arrest phenotype in which PGL-1 protein is ectopically expressed in the somatic cells (Unhavaithaya et al., 2002). Somatic expression of PIE-1 using the heat-shock promoter causes ectopic expression of PGL-1 in the intestinal cells, which mimics *mep-1* loss of function (Unhavaithaya et al., 2002). Based on these results, PIE-1 was proposed to antagonize MEP-1 function (Unhavaithaya et al., 2002). However, the *hsp::pie-1* transgenic strain does not induce significant transcriptional repression under conditions that inhibit MEP-1 (Unhavaithaya et al., 2002). It is not clear if the conditions are below the threshold for transcriptional repression (Unhavaithaya et al., 2002), because the overexpressed PIE-1 in the engineered

strain expressing a single-*copy hsp::pie-1* (Frokjaer-Jensen et al., 2008) did not repress both MEP-1 and transcription. As discussed in chapter II, the previous finding that *hsp::pie-1* mimics *mep-1* loss of function (Unhavaithaya et al., 2002) may have been obtained due to an unexpected disruption of a functional gene by randomly integrated extrachromosomal array *hsp::pie-1* in the genome. Alternatively, PIE-1 inhibits transcription, perhaps at specific loci, but not globally, given that somatic expression of PIE-1 in the *pie-1(R109L)::gfp* strain did not impair embryogenesis at all as discussed in chapter III.

In chapter II, we showed that SUMOylation is required for PIE-1 function in germline specification. Consistent with a previous study (Wu et al., 2012), we detected ectopic expression of PGL-1 protein in the intestinal cells of *pgl-1::mCherry* CRISPR lines when we depleted SMO-1 by RNAi. The genetic double strain *pgl-1::mCherry, ubc-9(G56R)* also showed ectopic PGL-1 expression in somatic cells at 25°C. Like PGL-1::mCherry, both GFP::PRG-1 and GFP::CSR-1 were also ectopically expressed in somatic cells when either SMO-1 or UBC-9 was depleted. These data suggest that SUMOylation is required to inhibit the expression of germline genes in somatic cells. Therefore, consistent with our results in chapter II, SUMOylation may facilitate NuRD complex-mediated chromatin remodeling for a proper epigenetic state that is important for germline–soma distinctions. It will be interesting to investigate whether the MEP-1–NuRD complex functions as a general repressor in both somatic and germline cells to protect their cell fate. Perhaps, PIE-1 and a key determinant of somatic

cell fate may recruit the MEP-1–NuRD complex to specific loci to define patterns of gene expression for germline specification and somatic specification, respectively.

Functional Domains of PIE-1

Two separate zinc-finger domains of PIE-1 contribute to asymmetric segregation of PIE-1 with the germline lineage (Figure 4.1) (Reese et al., 2000). ZF1 is targeted by ZIF-1, which interacts with the E3 ubiquitin ligase and promotes PIE-1 degradation in somatic blastomeres (DeRenzo et al., 2003), while ZF2 targets PIE-1 to P-granules that are actively transported into the presumptive germ cell after cell division (Reese et al., 2000). The YAPMAPT sequence of the C-terminal proline-rich region of PIE-1 is involved in transcriptional repression (Figure 4.1) (Batchelder et al., 1999). Interestingly, the proline-rich region is both sufficient and necessary for PIE-1 to interact with MEP-1 (Figure 4.1) (Unhavaithaya et al., 2002). In this research, we determined additional functional residues: K68 is required for PIE-1 SUMOylation, R109 on ZF1 is the target site of ZIF-1, and S327 is responsible for a PIE-1 cleavage event (Figure 4.1).

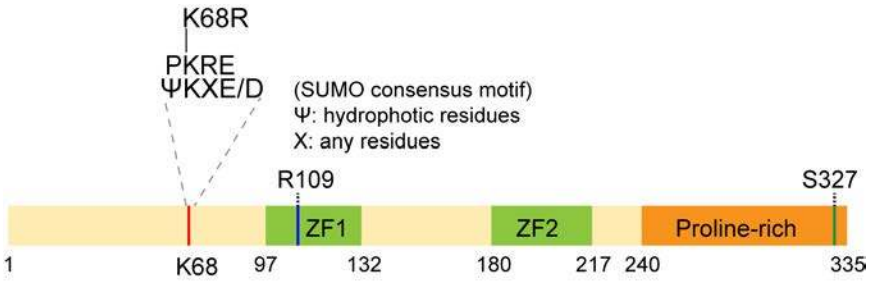


Figure 4.1 Functional domains of PIE-1

ZF1 (amino acids 97–132) is required for degradation by ZIF-1 in somatic blastomere, and ZF2 (amino acids 180–217) associates with P granules. The proline-rich region (amino acids 240–335) is important for transcriptional repression and interaction with MEP-1. The red bar represents the consensus SUMO acceptor site of PIE-1. The blue line indicates R109 residue that is the target site of ZIF-1. The green line indicates the cleavage site, S327.

Conclusions

Both the functional domains summarized in Figure 4.1 and our findings from this research (Figure 4.2) imply that PIE-1 is likely a multi-functional protein that regulates epigenetic modification, transcription, and post-transcriptional modification. However, the absence of a PIE-1 homolog in another organism makes the function of PIE-1 unpredictable. In addition, a certain threshold requirement for PIE-1 function (Batchelder et al., 1999; Tenenhaus et al., 1998; Unhavaithaya et al., 2002) (discussed in chapters II and IV) makes PIE-1 studies more complicated and challenging. Fortunately, efficient genome editing using CRIPSR-Cas9 technology allows us to explore the underlying molecular mechanisms more accurately and rapidly. *In vivo* SUMO purification from the various CRISPR strains that we employed here provides a powerful tool for exploring the function of genes of interest that are predicted to be SUMOylated. Using these major approaches, we have shown here that SUMOylation facilitates PIE-1–dependent germline maintenance and specification. PIE-1 SUMOylation is required for not only the formation of the NuRD complex but also for the promotion of other unsolved biological functions that are discussed in this chapter (Figure 4.2). Although the functional consequences and molecular mechanisms remain to be further elucidated, it is likely that SUMOylation acts as a switch to connect or disconnect many different functions of PIE-1 in a cell fate-dependent manner or a developmental timing-dependent manner for the differentiation of germline cells and specialized somatic cells.

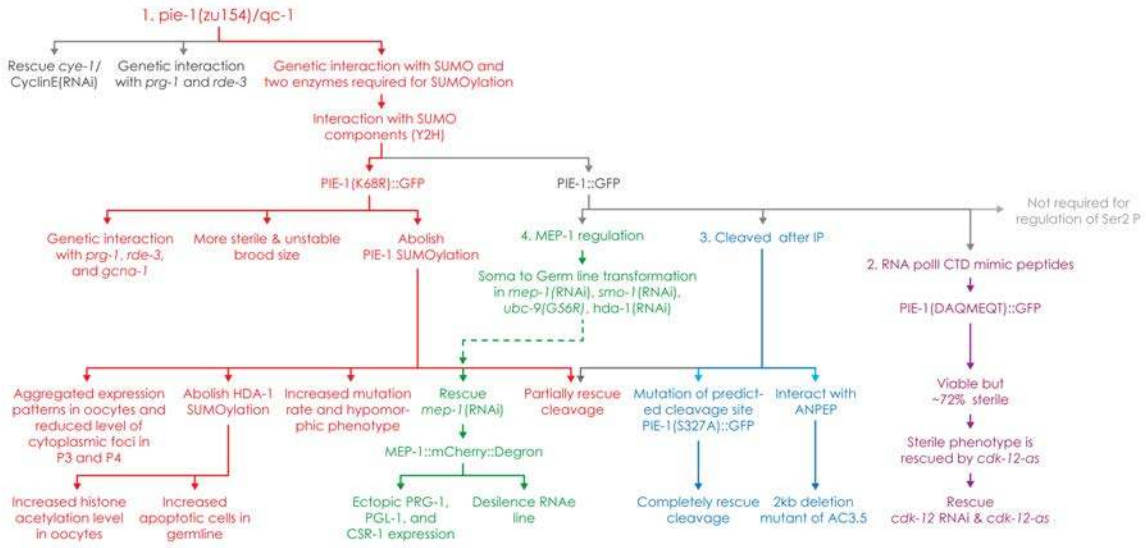


Figure 4.2 Schematic representation of the multiple-function of PIE-1

1. SUMOylation is required for PIE-1 function in germline specification (red). 2. A role for the YAPMAPT sequence in transcriptional regulation (purple). 3. PIE-1 cleavage event (blue). 4. MEP-1/NuRD function in maintaining germline-soma distinction (green).

REFERENCES

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

Ancelin, K., Lange, U.C., Hajkova, P., Schneider, R., Bannister, A.J., Kouzarides, T., and Surani, M.A. (2006). Blimp1 associates with Prmt5 and directs histone arginine methylation in mouse germ cells. Nat Cell Biol *8*, 623-630.

Arribere, J.A., Bell, R.T., Fu, B.X., Artiles, K.L., Hartman, P.S., and Fire, A.Z. (2014). Efficient marker-free recovery of custom genetic modifications with CRISPR/Cas9 in Caenorhabditis elegans. Genetics *198*, 837-846.

Bartkowiak, B., Liu, P., Phatnani, H.P., Fuda, N.J., Cooper, J.J., Price, D.H., Adelman, K., Lis, J.T., and Greenleaf, A.L. (2010). CDK12 is a transcription elongation-associated CTD kinase, the metazoan ortholog of yeast Ctk1. Genes Dev *24*, 2303-2316.

Bassett, A.R., Tibbit, C., Ponting, C.P., and Liu, J.L. (2013). Highly efficient targeted mutagenesis of Drosophila with the CRISPR/Cas9 system. Cell reports *4*, 220-228.

Batchelder, C., Dunn, M.A., Choy, B., Suh, Y., Cassie, C., Shim, E.Y., Shin, T.H., Mello, C., Seydoux, G., and Blackwell, T.K. (1999). Transcriptional repression by the Caenorhabditis elegans germ-line protein PIE-1. Genes Dev *13*, 202-212.

Belfiore, M., Mathies, L.D., Pugnale, P., Moulder, G., Barstead, R., Kimble, J., and Puoti, A. (2002). The MEP-1 zinc-finger protein acts with MOG DEAH box proteins to control gene expression via the fem-3 3' untranslated region in Caenorhabditis elegans. RNA *8*, 725-739.

Bender, L.B., Cao, R., Zhang, Y., and 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, L.B., Suh, J., Carroll, C.R., Fong, Y., Fingerman, I.M., Briggs, S.D., Cao, R., Zhang, Y., Reinke, V., and 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.

Benian, G.M., L'Hernault, S.W., and Morris, M.E. (1993). Additional sequence complexity in the muscle gene, unc-22, and its encoded protein, twitchin, of Caenorhabditis elegans. Genetics *134*, 1097-1104.

Betting, J., and Seufert, W. (1996). A yeast Ubc9 mutant protein with temperature-sensitive in vivo function is subject to conditional proteolysis by a ubiquitin- and proteasome-dependent pathway. J Biol Chem *271*, 25790-25796.

Bhaya, D., Davison, M., and Barrangou, R. (2011). CRISPR-Cas systems in bacteria and archaea: versatile small RNAs for adaptive defense and regulation. Annual review of genetics *45*, 273-297.

Boulton, S.J., Martin, J.S., Polanowska, J., Hill, D.E., Gartner, A., and Vidal, M. (2004). BRCA1/BARD1 orthologs required for DNA repair in Caenorhabditis elegans. Curr Biol *14*, 33-39.

Bowerman, B., Draper, B.W., Mello, C.C., and Priess, J.R. (1993). The maternal gene skn-1 encodes a protein that is distributed unequally in early C. elegans embryos. Cell *74*, 443-452.
Bowman, E.A., Bowman, C.R., Ahn, J.H., and Kelly, W.G. (2013). Phosphorylation of RNA polymerase II is independent of P-TEFb in the C. elegans germline. Development *140*, 3703-3713.
Brenner, S. (1974). The genetics of Caenorhabditis elegans. Genetics *77*, 71-94.
Capowski, E.E., Martin, P., Garvin, C., and Strome, S. (1991). Identification of grandchildless loci whose products are required for normal germ-line development in the nematode Caenorhabditis elegans. Genetics *129*, 1061-1072.
Carmell, M.A., Dokshin, G.A., Skaletsky, H., Hu, Y.C., van Wolfswinkel, J.C., Igarashi, K.J., Bellott, D.W., Nefedov, M., Reddien, P.W., Enders, G.C.*, et al.* (2016). A widely employed germ cell marker is an ancient disordered protein with reproductive functions in diverse eukaryotes. Elife *5*.
Chang, N., Sun, C., Gao, L., Zhu, D., Xu, X., Zhu, X., Xiong, J.W., and Xi, J.J. (2013). Genome editing with RNA-guided Cas9 nuclease in zebrafish embryos. Cell research *23*, 465-472.
Chen, C., Fenk, L.A., and de Bono, M. (2013). Efficient genome editing in Caenorhabditis elegans by CRISPR-targeted homologous recombination. Nucleic acids research *41*, e193.
Chen, C.C., Simard, M.J., Tabara, H., Brownell, D.R., McCollough, J.A., and Mello, C.C. (2005). A member of the polymerase beta nucleotidyltransferase superfamily is required for RNA interference in C. elegans. Curr Biol *15*, 378-383.
Cheng, A.W., Wang, H., Yang, H., Shi, L., Katz, Y., Theunissen, T.W., Rangarajan, S., Shivalila, C.S., Dadon, D.B., and Jaenisch, R. (2013). Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system. Cell research *23*, 1163-1171.
Cheng, J., Wang, D., Wang, Z., and Yeh, E.T. (2004). SENP1 enhances androgen receptor-dependent transcription through desumoylation of histone deacetylase 1. Mol Cell Biol *24*, 6021-6028.
Chiu, H., Schwartz, H.T., Antoshechkin, I., and Sternberg, P.W. (2013). Transgene-free genome editing in Caenorhabditis elegans using CRISPR-Cas. Genetics *195*, 1167-1171.
Cho, S.W., Kim, S., Kim, J.M., and Kim, J.S. (2013a). Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. Nature biotechnology *31*, 230-232.
Cho, S.W., Lee, J., Carroll, D., Kim, J.S., and Lee, J. (2013b). Heritable gene knockout in Caenorhabditis elegans by direct injection of Cas9-sgRNA ribonucleoproteins. Genetics *195*, 1177-1180.
Citro, S., Jaffray, E., Hay, R.T., Seiser, C., and Chiocca, S. (2013). A role for paralog-specific sumoylation in histone deacetylase 1 stability. J Mol Cell Biol *5*, 416-427.

Collins, J.J., and Anderson, P. (1994). The Tc5 family of transposable elements in Caenorhabditis elegans. Genetics *137*, 771-781.
Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A.*, et al.* (2013). Multiplex genome engineering using CRISPR/Cas systems. Science *339*, 819-823.
David, G., Neptune, M.A., and DePinho, R.A. (2002). SUMO-1 modification of histone deacetylase 1 (HDAC1) modulates its biological activities. J Biol Chem *277*, 23658-23663.
Dent, J.A., Smith, M.M., Vassilatis, D.K., and Avery, L. (2000). The genetics of ivermectin resistance in Caenorhabditis elegans. Proceedings of the National Academy of Sciences of the United States of America *97*, 2674-2679.
DeRenzo, C., Reese, K.J., and Seydoux, G. (2003). Exclusion of germ plasm proteins from somatic lineages by cullin-dependent degradation. Nature *424*, 685-689.
Detwiler, M.R., Reuben, M., Li, X., Rogers, E., and Lin, R. (2001). Two zinc finger proteins, OMA-1 and OMA-2, are redundantly required for oocyte maturation in C. elegans. Dev Cell *1*, 187-199.
Dicarlo, J.E., Conley, A.J., Penttila, M., Jantti, J., Wang, H.H., and Church, G.M. (2013). Yeast Oligo-Mediated Genome Engineering (YOGE). ACS synthetic biology.
Dickinson, D.J., Ward, J.D., Reiner, D.J., and Goldstein, B. (2013). Engineering the Caenorhabditis elegans genome using Cas9-triggered homologous recombination. Nature methods *10*, 1028-1034.
Dobbin, M.M., Madabhushi, R., Pan, L., Chen, Y., Kim, D., Gao, J., Ahanonu, B., Pao, P.C., Qiu, Y., Zhao, Y.*, et al.* (2013). SIRT1 collaborates with ATM and HDAC1 to maintain genomic stability in neurons. Nat Neurosci *16*, 1008-1015.
Duchaine, T.F., Wohlschlegel, J.A., Kennedy, S., Bei, Y., Conte, D., Jr., Pang, K., Brownell, D.R., Harding, S., Mitani, S., Ruvkun, G.*, et al.* (2006). Functional proteomics reveals the biochemical niche of C. elegans DCR-1 in multiple small-RNA-mediated pathways. Cell *124*, 343-354.
Ephrussi, A., and Lehmann, R. (1992). Induction of germ cell formation by oskar. Nature *358*, 387-392.
Feng, Z., Mao, Y., Xu, N., Zhang, B., Wei, P., Yang, D.L., Wang, Z., Zhang, Z., Zheng, R., Yang, L.*, et al.* (2014). Multigeneration analysis reveals the inheritance, specificity, and patterns of CRISPR/Cas-induced gene modifications in Arabidopsis. Proceedings of the National Academy of Sciences of the United States of America.
Feng, Z., Zhang, B., Ding, W., Liu, X., Yang, D.L., Wei, P., Cao, F., Zhu, S., Zhang, F., Mao, Y.*, et al.* (2013). Efficient genome editing in plants using a CRISPR/Cas system. Cell research *23*, 1229-1232.
Fong, Y., Bender, L., Wang, W., and Strome, S. (2002). Regulation of the different chromatin states of autosomes and X chromosomes in the germ line of C. elegans. Science *296*, 2235-2238.

Formstecher, E., Aresta, S., Collura, V., Hamburger, A., Meil, A., Trehin, A., Reverdy, C., Betin, V., Maire, S., Brun, C.*, et al.* (2005). Protein interaction mapping: a Drosophila case study. Genome Res *15*, 376-384.
Friedland, A.E., Tzur, Y.B., Esvelt, K.M., Colaiacovo, M.P., Church, G.M., and Calarco, J.A. (2013). Heritable genome editing in C. elegans via a CRISPR-Cas9 system. Nature methods *10*, 741-743.
Frokjaer-Jensen, C., Davis, M.W., Ailion, M., and Jorgensen, E.M. (2012). Improved Mos1-mediated transgenesis in C. elegans. Nature methods *9*, 117-118.
Frokjaer-Jensen, C., Davis, M.W., Hopkins, C.E., Newman, B.J., Thummel, J.M., Olesen, S.P., Grunnet, M., and Jorgensen, E.M. (2008). Single-copy insertion of transgenes in Caenorhabditis elegans. Nat Genet *40*, 1375-1383.
Fromont-Racine, M., Rain, J.C., and Legrain, P. (1997). Toward a functional analysis of the yeast genome through exhaustive two-hybrid screens. Nat Genet *16*, 277-282.
Furuhashi, H., Takasaki, T., Rechtsteiner, A., Li, T., Kimura, H., Checchi, P.M., Strome, S., and Kelly, W.G. (2010). Trans-generational epigenetic regulation of C. elegans primordial germ cells. Epigenetics Chromatin *3*, 15.
Gareau, J.R., and Lima, C.D. (2010). The SUMO pathway: emerging mechanisms that shape specificity, conjugation and recognition. Nat Rev Mol Cell Biol *11*, 861-871.
Geiss-Friedlander, R., and Melchior, F. (2007). Concepts in sumoylation: a decade on. Nat Rev Mol Cell Biol *8*, 947-956.
Ghosh, D., and Seydoux, G. (2008). Inhibition of transcription by the Caenorhabditis elegans germline protein PIE-1: genetic evidence for distinct mechanisms targeting initiation and elongation. Genetics *178*, 235-243.
Gilbert, L.A., Larson, M.H., Morsut, L., Liu, Z., Brar, G.A., Torres, S.E., Stern-Ginossar, N., Brandman, O., Whitehead, E.H., Doudna, J.A.*, et al.* (2013). CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. Cell *154*, 442-451.
Gill, G. (2005). Something about SUMO inhibits transcription. Curr Opin Genet Dev *15*, 536-541.
Giordano-Santini, R., Milstein, S., Svrzikapa, N., Tu, D., Johnsen, R., Baillie, D., Vidal, M., and Dupuy, D. (2010). An antibiotic selection marker for nematode transgenesis. Nature methods *7*, 721-723.
Golebiowski, F., Matic, I., Tatham, M.H., Cole, C., Yin, Y., Nakamura, A., Cox, J., Barton, G.J., Mann, M., and Hay, R.T. (2009). System-wide changes to SUMO modifications in response to heat shock. Sci Signal *2*, ra24.
Gong, F., and Miller, K.M. (2013). Mammalian DNA repair: HATs and HDACs make their mark through histone acetylation. Mutat Res *750*, 23-30.
Gratz, S.J., Cummings, A.M., Nguyen, J.N., Hamm, D.C., Donohue, L.K., Harrison, M.M., Wildonger, J., and O'Connor-Giles, K.M. (2013). Genome engineering of Drosophila with the CRISPR RNA-guided Cas9 nuclease. Genetics *194*, 1029-1035.

Grishok, A., and Mello, C.C. (2002). RNAi (Nematodes: Caenorhabditis elegans). Advances in genetics *46*, 339-360.
Gruidl, M.E., Smith, P.A., Kuznicki, K.A., McCrone, J.S., Kirchner, J., Roussell, D.L., Strome, S., and Bennett, K.L. (1996). Multiple potential germ-line helicases are components of the germ-line-specific P granules of Caenorhabditis elegans. Proc Natl Acad Sci U S A *93*, 13837-13842.
Guo, D., Li, M., Zhang, Y., Yang, P., Eckenrode, S., Hopkins, D., Zheng, W., Purohit, S., Podolsky, R.H., Muir, A.*, et al.* (2004). A functional variant of SUMO4, a new I kappa B alpha modifier, is associated with type 1 diabetes. Nat Genet *36*, 837-841.
Guven-Ozkan, T., Nishi, Y., Robertson, S.M., and Lin, R. (2008). Global transcriptional repression in C. elegans germline precursors by regulated sequestration of TAF-4. Cell *135*, 149-160.
Gyory, I., Wu, J., Fejer, G., Seto, E., and Wright, K.L. (2004). PRDI-BF1 recruits the histone H3 methyltransferase G9a in transcriptional silencing. Nat Immunol *5*, 299-308.
Hannon, G.J. (2002). RNA interference. Nature *418*, 244-251.
Hanyu-Nakamura, K., Sonobe-Nojima, H., Tanigawa, A., Lasko, P., and Nakamura, A. (2008). Drosophila Pgc protein inhibits P-TEFb recruitment to chromatin in primordial germ cells. Nature *451*, 730-733.
Hardeland, U., Steinacher, R., Jiricny, J., and Schar, P. (2002). Modification of the human thymine-DNA glycosylase by ubiquitin-like proteins facilitates enzymatic turnover. EMBO J *21*, 1456-1464.
Harris, J., Lowden, M., Clejan, I., Tzoneva, M., Thomas, J.H., Hodgkin, J., and Ahmed, S. (2006). Mutator phenotype of Caenorhabditis elegans DNA damage checkpoint mutants. Genetics *174*, 601-616.
Hay, R.T. (2005). SUMO: a history of modification. Mol Cell *18*, 1-12.
Hodgkin, J., Horvitz, H.R., and Brenner, S. (1979). Nondisjunction Mutants of the Nematode CAENORHABDITIS ELEGANS. Genetics *91*, 67-94.
Hoeller, D., Hecker, C.M., Wagner, S., Rogov, V., Dotsch, V., and Dikic, I. (2007). E3-independent monoubiquitination of ubiquitin-binding proteins. Mol Cell *26*, 891-898.
Holway, A.H., Hung, C., and Michael, W.M. (2005). Systematic, RNA-interference-mediated identification of mus-101 modifier genes in Caenorhabditis elegans. Genetics *169*, 1451-1460.
Holway, A.H., Kim, S.H., La Volpe, A., and Michael, W.M. (2006). Checkpoint silencing during the DNA damage response in Caenorhabditis elegans embryos. J Cell Biol *172*, 999-1008.
Horii, T., Morita, S., Kimura, M., Kobayashi, R., Tamura, D., Takahashi, R.U., Kimura, H., Suetake, I., Ohata, H., Okamoto, K.*, et al.* (2013). Genome engineering of mammalian haploid embryonic stem cells using the Cas9/RNA system. PeerJ *1*, e230.
Horvath, P., and Barrangou, R. (2010). CRISPR/Cas, the immune system of bacteria and archaea. Science *327*, 167-170.

Hsiao, K.Y., and Mizzen, C.A. (2013). Histone H4 deacetylation facilitates 53BP1 DNA damage signaling and double-strand break repair. J Mol Cell Biol *5*, 157-165.
Ikenishi, K. (1998). Germ plasm in Caenorhabditis elegans, Drosophila and Xenopus. Development, growth & differentiation *40*, 1-10.
Illmensee, K., and Mahowald, A.P. (1974). Transplantation of posterior polar plasm in Drosophila. Induction of germ cells at the anterior pole of the egg. Proc Natl Acad Sci U S A *71*, 1016-1020.
Jiang, W., Bikard, D., Cox, D., Zhang, F., and Marraffini, L.A. (2013). RNA-guided editing of bacterial genomes using CRISPR-Cas systems. Nature biotechnology *31*, 233-239.
Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science *337*, 816-821.
Jones, D., Crowe, E., Stevens, T.A., and Candido, E.P. (2002). Functional and phylogenetic analysis of the ubiquitylation system in Caenorhabditis elegans: ubiquitin-conjugating enzymes, ubiquitin-activating enzymes, and ubiquitin-like proteins. Genome Biol *3*, RESEARCH0002.
Joung, H., Kwon, S., Kim, K.H., Lee, Y.G., Shin, S., Kwon, D.H., Lee, Y.U., Kook, T., Choe, N., Kim, J.C*., et al.* (2018). Sumoylation of histone deacetylase 1 regulates MyoD signaling during myogenesis. Exp Mol Med *50*, e427.
Kamath, R.S., and Ahringer, J. (2003). Genome-wide RNAi screening in Caenorhabditis elegans. Methods *30*, 313-321.
Katic, I., and Grosshans, H. (2013). Targeted heritable mutation and gene conversion by Cas9-CRISPR in Caenorhabditis elegans. Genetics *195*, 1173-1176.
Kawasaki, I., Amiri, A., Fan, Y., Meyer, N., Dunkelbarger, S., Motohashi, T., Karashima, T., Bossinger, O., and Strome, S. (2004). The PGL family proteins associate with germ granules and function redundantly in Caenorhabditis elegans germline development. Genetics *167*, 645-661.
Kawasaki, I., Shim, Y.H., Kirchner, J., Kaminker, J., Wood, W.B., and Strome, S. (1998). PGL-1, a predicted RNA-binding component of germ granules, is essential for fertility in C. elegans. Cell *94*, 635-645.
Kaymak, E., and Ryder, S.P. (2013). RNA recognition by the Caenorhabditis elegans oocyte maturation determinant OMA-1. J Biol Chem *288*, 30463-30472.
Kearns, N.A., Genga, R.M., Enuameh, M.S., Garber, M., Wolfe, S.A., and Maehr, R. (2014). Cas9 effector-mediated regulation of transcription and differentiation in human pluripotent stem cells. Development *141*, 219-223.
Kerscher, O., Felberbaum, R., and Hochstrasser, M. (2006). Modification of proteins by ubiquitin and ubiquitin-like proteins. Annu Rev Cell Dev Biol *22*, 159-180.
Kim, H., Ishidate, T., Ghanta, K.S., Seth, M., Conte, D., Jr., Shirayama, M., and Mello, C.C. (2014). A co-CRISPR strategy for efficient genome editing in Caenorhabditis elegans. Genetics *197*, 1069-1080.

Kim, H.M., and Colaiacovo, M.P. (2014). ZTF-8 interacts with the 9-1-1 complex and is required for DNA damage response and double-strand break repair in the C. elegans germline. PLoS Genet *10*, e1004723.
Kim, H.M., and Colaiacovo, M.P. (2015). New Insights into the Post-Translational Regulation of DNA Damage Response and Double-Strand Break Repair in Caenorhabditis elegans. Genetics *200*, 495-504.
Kim, S.H., and Michael, W.M. (2008). Regulated proteolysis of DNA polymerase eta during the DNA-damage response in C. elegans. Mol Cell *32*, 757-766.
Kobayashi, S., Yamada, M., Asaoka, M., and Kitamura, T. (1996). Essential role of the posterior morphogen nanos for germline development in Drosophila. Nature *380*, 708-711.
Kuznicki, K.A., Smith, P.A., Leung-Chiu, W.M., Estevez, A.O., Scott, H.C., and Bennett, K.L. (2000). Combinatorial RNA interference indicates GLH-4 can compensate for GLH-1; these two P granule components are critical for fertility in C. elegans. Development *127*, 2907-2916.
Larson, M.H., Gilbert, L.A., Wang, X., Lim, W.A., Weissman, J.S., and Qi, L.S. (2013). CRISPR interference (CRISPRi) for sequence-specific control of gene expression. Nature protocols *8*, 2180-2196.
Lawson, K.A., Dunn, N.R., Roelen, B.A., Zeinstra, L.M., Davis, A.M., Wright, C.V., Korving, J.P., and Hogan, B.L. (1999). Bmp4 is required for the generation of primordial germ cells in the mouse embryo. Genes Dev *13*, 424-436.
Leatherman, J.L., Levin, L., Boero, J., and Jongens, T.A. (2002). germ cell-less acts to repress transcription during the establishment of the Drosophila germ cell lineage. Curr Biol *12*, 1681-1685.
Leight, E.R., Glossip, D., and Kornfeld, K. (2005). Sumoylation of LIN-1 promotes transcriptional repression and inhibition of vulval cell fates. Development *132*, 1047-1056.
Lera, R.F., and Burkard, M.E. (2012). The final link: tapping the power of chemical genetics to connect the molecular and biologic functions of mitotic protein kinases. Molecules *17*, 12172-12186.
Lin, R. (2003). A gain-of-function mutation in oma-1, a C. elegans gene required for oocyte maturation, results in delayed degradation of maternal proteins and embryonic lethality. Dev Biol *258*, 226-239.
Lo, T.W., Pickle, C.S., Lin, S., Ralston, E.J., Gurling, M., Schartner, C.M., Bian, Q., Doudna, J.A., and Meyer, B.J. (2013). Precise and heritable genome editing in evolutionarily diverse nematodes using TALENs and CRISPR/Cas9 to engineer insertions and deletions. Genetics *195*, 331-348.
Lu, N., Yu, X., He, X., and Zhou, Z. (2009). Detecting apoptotic cells and monitoring their clearance in the nematode Caenorhabditis elegans. Methods Mol Biol *559*, 357-370.
Ma, Y., Shen, B., Zhang, X., Lu, Y., Chen, W., Ma, J., Huang, X., and Zhang, L. (2014). Heritable Multiplex Genetic Engineering in Rats Using CRISPR/Cas9. PloS one *9*, e89413.

Magnusdottir, E., Dietmann, S., Murakami, K., Gunesdogan, U., Tang, F., Bao, S., Diamanti, E., Lao, K., Gottgens, B., and Azim Surani, M. (2013). A tripartite transcription factor network regulates primordial germ cell specification in mice. Nat Cell Biol *15*, 905-915.
Mali, P., Aach, J., Stranges, P.B., Esvelt, K.M., Moosburner, M., Kosuri, S., Yang, L., and Church, G.M. (2013a). CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. Nature biotechnology *31*, 833-838.
Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., and Church, G.M. (2013b). RNA-guided human genome engineering via Cas9. Science *339*, 823-826.
Matunis, M.J., Zhang, X.D., and Ellis, N.A. (2006). SUMO: the glue that binds. Dev Cell *11*, 596-597.
Melchior, F. (2000). SUMO--nonclassical ubiquitin. Annu Rev Cell Dev Biol *16*, 591-626.
Mello, C.C., Draper, B.W., Krause, M., Weintraub, H., and Priess, J.R. (1992). The pie-1 and mex-1 genes and maternal control of blastomere identity in early C. elegans embryos. Cell *70*, 163-176.
Mello, C.C., Schubert, C., Draper, B., Zhang, W., Lobel, R., and Priess, J.R. (1996). The PIE-1 protein and germline specification in C. elegans embryos. Nature *382*, 710-712.
Miller, K.M., Tjeertes, J.V., Coates, J., Legube, G., Polo, S.E., Britton, S., and Jackson, S.P. (2010). Human HDAC1 and HDAC2 function in the DNA-damage response to promote DNA nonhomologous end-joining. Nat Struct Mol Biol *17*, 1144-1151.
Moerman, D.G., and Baillie, D.L. (1979). Genetic Organization in CAENORHABDITIS ELEGANS: Fine-Structure Analysis of the unc-22 Gene. Genetics *91*, 95-103.
Nikolova, T., Kiweler, N., and Kramer, O.H. (2017). Interstrand Crosslink Repair as a Target for HDAC Inhibition. Trends Pharmacol Sci *38*, 822-836.
Nishi, Y., and Lin, R. (2005). DYRK2 and GSK-3 phosphorylate and promote the timely degradation of OMA-1, a key regulator of the oocyte-to-embryo transition in C. elegans. Dev Biol *288*, 139-149.
Ohinata, Y., Ohta, H., Shigeta, M., Yamanaka, K., Wakayama, T., and Saitou, M. (2009). A signaling principle for the specification of the germ cell lineage in mice. Cell *137*, 571-584.
Ohinata, Y., Payer, B., O'Carroll, D., Ancelin, K., Ono, Y., Sano, M., Barton, S.C., Obukhanych, T., Nussenzweig, M., Tarakhovsky, A.*, et al.* (2005). Blimp1 is a critical determinant of the germ cell lineage in mice. Nature *436*, 207-213.
Paques, F., and Haber, J.E. (1999). Multiple pathways of recombination induced by double-strand breaks in Saccharomyces cerevisiae. Microbiology and molecular biology reviews : MMBR *63*, 349-404.
Praitis, V., Casey, E., Collar, D., and Austin, J. (2001). Creation of low-copy integrated transgenic lines in Caenorhabditis elegans. Genetics *157*, 1217-1226.

Prendergast, J.A., Ptak, C., Arnason, T.G., and Ellison, M.J. (1995). Increased ubiquitin expression suppresses the cell cycle defect associated with the yeast ubiquitin conjugating enzyme, CDC34 (UBC3). Evidence for a noncovalent interaction between CDC34 and ubiquitin. J Biol Chem *270*, 9347-9352.
Psakhye, I., and Jentsch, S. (2012). Protein group modification and synergy in the SUMO pathway as exemplified in DNA repair. Cell *151*, 807-820.
Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P., and Lim, W.A. (2013). Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell *152*, 1173-1183.
Radman, I., Greiss, S., and Chin, J.W. (2013). Efficient and rapid C. elegans transgenesis by bombardment and hygromycin B selection. PloS one *8*, e76019.
Ran, F.A., Hsu, P.D., Lin, C.Y., Gootenberg, J.S., Konermann, S., Trevino, A.E., Scott, D.A., Inoue, A., Matoba, S., Zhang, Y.*, et al.* (2013). Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. Cell *154*, 1380-1389.
Rechtsteiner, A., Ercan, S., Takasaki, T., Phippen, T.M., Egelhofer, T.A., Wang, W., Kimura, H., Lieb, J.D., and Strome, S. (2010). The histone H3K36 methyltransferase MES-4 acts epigenetically to transmit the memory of germline gene expression to progeny. PLoS Genet *6*, e1001091.
Reese, K.J., Dunn, M.A., Waddle, J.A., and Seydoux, G. (2000). Asymmetric segregation of PIE-1 in C. elegans is mediated by two complementary mechanisms that act through separate PIE-1 protein domains. Mol Cell *6*, 445-455.
Reichman, R., Shi, Z., Malone, R., and Smolikove, S. (2018). Mitotic and Meiotic Functions for the SUMOylation Pathway in the Caenorhabditis elegans Germline. Genetics *208*, 1421-1441.
Ren, X., Sun, J., Housden, B.E., Hu, Y., Roesel, C., Lin, S., Liu, L.P., Yang, Z., Mao, D., Sun, L.*, et al.* (2013). Optimized gene editing technology for Drosophila melanogaster using germ line-specific Cas9. Proceedings of the National Academy of Sciences of the United States of America *110*, 19012-19017.
Robert, C., and Rassool, F.V. (2012). HDAC inhibitors: roles of DNA damage and repair. Adv Cancer Res *116*, 87-129.
Rodriguez, M.S., Dargemont, C., and Hay, R.T. (2001). SUMO-1 conjugation in vivo requires both a consensus modification motif and nuclear targeting. J Biol Chem *276*, 12654-12659.
Sarangi, P., and Zhao, X. (2015). SUMO-mediated regulation of DNA damage repair and responses. Trends Biochem Sci *40*, 233-242.
Sato, K., Hayashi, Y., Ninomiya, Y., Shigenobu, S., Arita, K., Mukai, M., and Kobayashi, S. (2007). Maternal Nanos represses hid/skl-dependent apoptosis to maintain the germ line in Drosophila embryos. Proc Natl Acad Sci U S A *104*, 7455-7460.
Schaner, C.E., Deshpande, G., Schedl, P.D., and Kelly, W.G. (2003). A conserved chromatin architecture marks and maintains the restricted germ cell lineage in worms and flies. Dev Cell *5*, 747-757.

Semple, J.I., Garcia-Verdugo, R., and Lehner, B. (2010). Rapid selection of transgenic C. elegans using antibiotic resistance. Nature methods *7*, 725-727.
Seydoux, G., and Dunn, M.A. (1997). Transcriptionally repressed germ cells lack a subpopulation of phosphorylated RNA polymerase II in early embryos of Caenorhabditis elegans and Drosophila melanogaster. Development *124*, 2191-2201.
Seydoux, G., Mello, C.C., Pettitt, J., Wood, W.B., Priess, J.R., and Fire, A. (1996). Repression of gene expression in the embryonic germ lineage of C. elegans. Nature *382*, 713-716.
Shen, T.H., Lin, H.K., Scaglioni, P.P., Yung, T.M., and Pandolfi, P.P. (2006). The mechanisms of PML-nuclear body formation. Mol Cell *24*, 331-339.
Shi, Y., and Mello, C. (1998). A CBP/p300 homolog specifies multiple differentiation pathways in Caenorhabditis elegans. Genes Dev *12*, 943-955.
Shirayama, M., Seth, M., Lee, H.C., Gu, W., Ishidate, T., Conte, D., Jr., and Mello, C.C. (2012). piRNAs initiate an epigenetic memory of nonself RNA in the C. elegans germline. Cell *150*, 65-77.
Spike, C., Meyer, N., Racen, E., Orsborn, A., Kirchner, J., Kuznicki, K., Yee, C., Bennett, K., and Strome, S. (2008). Genetic analysis of the Caenorhabditis elegans GLH family of P-granule proteins. Genetics *178*, 1973-1987.
Stengel, K.R., and Hiebert, S.W. (2015). Class I HDACs Affect DNA Replication, Repair, and Chromatin Structure: Implications for Cancer Therapy. Antioxid Redox Signal *23*, 51-65.
Sternberg, S.H., Redding, S., Jinek, M., Greene, E.C., and Doudna, J.A. (2014). DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. Nature *507*, 62-67.
Strome, S. (2005). Specification of the germ line. WormBook, 1-10.
Sulston, J.E., Schierenberg, E., White, J.G., and Thomson, J.N. (1983). The embryonic cell lineage of the nematode Caenorhabditis elegans. Developmental biology *100*, 64-119.
Tammsalu, T., Matic, I., Jaffray, E.G., Ibrahim, A.F.M., Tatham, M.H., and Hay, R.T. (2014). Proteome-wide identification of SUMO2 modification sites. Sci Signal *7*, rs2.
Tao, C.C., Hsu, W.L., Ma, Y.L., Cheng, S.J., and Lee, E.H. (2017). Epigenetic regulation of HDAC1 SUMOylation as an endogenous neuroprotection against Abeta toxicity in a mouse model of Alzheimer's disease. Cell Death Differ *24*, 597-614.
Tatham, M.H., Matic, I., Mann, M., and Hay, R.T. (2011). Comparative proteomic analysis identifies a role for SUMO in protein quality control. Sci Signal *4*, rs4.
Tatham, M.H., Rodriguez, M.S., Xirodimas, D.P., and Hay, R.T. (2009). Detection of protein SUMOylation in vivo. Nat Protoc *4*, 1363-1371.
Tenenhaus, C., Schubert, C., and Seydoux, G. (1998). Genetic requirements for PIE-1 localization and inhibition of gene expression in the embryonic germ lineage of Caenorhabditis elegans. Dev Biol *200*, 212-224.

Tenenhaus, C., Subramaniam, K., Dunn, M.A., and Seydoux, G. (2001). PIE-1 is a bifunctional protein that regulates maternal and zygotic gene expression in the embryonic germ line of Caenorhabditis elegans. Genes & development *15*, 1031-1040.
Terns, M.P., and Terns, R.M. (2011). CRISPR-based adaptive immune systems. Current opinion in microbiology *14*, 321-327.
Thurn, K.T., Thomas, S., Raha, P., Qureshi, I., and Munster, P.N. (2013). Histone deacetylase regulation of ATM-mediated DNA damage signaling. Mol Cancer Ther *12*, 2078-2087.
Tzur, Y.B., Friedland, A.E., Nadarajan, S., Church, G.M., Calarco, J.A., and Colaiacovo, M.P. (2013). Heritable custom genomic modifications in Caenorhabditis elegans via a CRISPR-Cas9 system. Genetics *195*, 1181-1185.
Unhavaithaya, Y., Shin, T.H., Miliaras, N., Lee, J., Oyama, T., and Mello, C.C. (2002). MEP-1 and a homolog of the NURD complex component Mi-2 act together to maintain germline-soma distinctions in C. elegans. Cell *111*, 991-1002.
Vincent, S.D., Dunn, N.R., Sciammas, R., Shapiro-Shalef, M., Davis, M.M., Calame, K., Bikoff, E.K., and Robertson, E.J. (2005). The zinc finger transcriptional repressor Blimp1/Prdm1 is dispensable for early axis formation but is required for specification of primordial germ cells in the mouse. Development *132*, 1315-1325.
Voinnet, O. (2001). RNA silencing as a plant immune system against viruses. Trends in genetics : TIG *17*, 449-459.
von Zelewsky, T., Palladino, F., Brunschwig, K., Tobler, H., Hajnal, A., and Muller, F. (2000). The C. elegans Mi-2 chromatin-remodelling proteins function in vulval cell fate determination. Development *127*, 5277-5284.
Waaijers, S., Portegijs, V., Kerver, J., Lemmens, B.B., Tijsterman, M., van den Heuvel, S., and Boxem, M. (2013). CRISPR/Cas9-targeted mutagenesis in Caenorhabditis elegans. Genetics *195*, 1187-1191.
Wang, D., Kennedy, S., Conte, D., Jr., Kim, J.K., Gabel, H.W., Kamath, R.S., Mello, C.C., and Ruvkun, G. (2005). Somatic misexpression of germline P granules and enhanced RNA interference in retinoblastoma pathway mutants. Nature *436*, 593-597.
Wang, H., Yang, H., Shivalila, C.S., Dawlaty, M.M., Cheng, A.W., Zhang, F., and Jaenisch, R. (2013). One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. Cell *153*, 910-918.
Wang, J.T., and Seydoux, G. (2013). Germ cell specification. Adv Exp Med Biol *757*, 17-39.
Wang, T., Wei, J.J., Sabatini, D.M., and Lander, E.S. (2014). Genetic screens in human cells using the CRISPR-Cas9 system. Science *343*, 80-84.
Weber, S., Eckert, D., Nettersheim, D., Gillis, A.J., Schafer, S., Kuckenberg, P., Ehlermann, J., Werling, U., Biermann, K., Looijenga, L.H.*, et al.* (2010). Critical function of AP-2 gamma/TCFAP2C in mouse embryonic germ cell maintenance. Biol Reprod *82*, 214-223.

Wiedenheft, B., Sternberg, S.H., and Doudna, J.A. (2012). RNA-guided genetic silencing systems in bacteria and archaea. Nature *482*, 331-338.
Wood, A., and Shilatifard, A. (2006). Bur1/Bur2 and the Ctk complex in yeast: the split personality of mammalian P-TEFb. Cell Cycle *5*, 1066-1068.
Wu, X., Shi, Z., Cui, M., Han, M., and Ruvkun, G. (2012). Repression of germline RNAi pathways in somatic cells by retinoblastoma pathway chromatin complexes. PLoS Genet *8*, e1002542.
Xu, L., Fong, Y., and Strome, S. (2001). The Caenorhabditis elegans maternal-effect sterile proteins, MES-2, MES-3, and MES-6, are associated in a complex in embryos. Proc Natl Acad Sci U S A *98*, 5061-5066.
Yamaji, M., Seki, Y., Kurimoto, K., Yabuta, Y., Yuasa, M., Shigeta, M., Yamanaka, K., Ohinata, Y., and Saitou, M. (2008). Critical function of Prdm14 for the establishment of the germ cell lineage in mice. Nat Genet *40*, 1016-1022.
Ying, Y., Liu, X.M., Marble, A., Lawson, K.A., and Zhao, G.Q. (2000). Requirement of Bmp8b for the generation of primordial germ cells in the mouse. Mol Endocrinol *14*, 1053-1063.
Yu, J., Angelin-Duclos, C., Greenwood, J., Liao, J., and Calame, K. (2000). Transcriptional repression by blimp-1 (PRDI-BF1) involves recruitment of histone deacetylase. Mol Cell Biol *20*, 2592-2603.
Yu, X., Lu, N., and Zhou, Z. (2008). Phagocytic receptor CED-1 initiates a signaling pathway for degrading engulfed apoptotic cells. PLoS Biol *6*, e61.
Yu, Z., Chen, H., Liu, J., Zhang, H., Yan, Y., Zhu, N., Guo, Y., Yang, B., Chang, Y., Dai, F.*, et al.* (2014). Various applications of TALEN- and CRISPR/Cas9-mediated homologous recombination to modify the Drosophila genome. Biology open.
Zaborowska, J., Egloff, S., and Murphy, S. (2016). The pol II CTD: new twists in the tail. Nat Struct Mol Biol *23*, 771-777.
Zamore, P.D. (2001). RNA interference: listening to the sound of silence. Nature structural biology *8*, 746-750.
Zhang, F., Barboric, M., Blackwell, T.K., and Peterlin, B.M. (2003). A model of repression: CTD analogs and PIE-1 inhibit transcriptional elongation by P-TEFb. Genes Dev *17*, 748-758.
Zhao, P., Zhang, Z., Ke, H., Yue, Y., and Xue, D. (2014). Oligonucleotide-based targeted gene editing in C. elegans via the CRISPR/Cas9 system. Cell research *24*, 247-250.
Zhou, J., Shen, B., Zhang, W., Wang, J., Yang, J., Chen, L., Zhang, N., Zhu, K., Xu, J., Hu, B.*, et al.* (2014). One-step generation of different immunodeficient mice with multiple gene modifications by CRISPR/Cas9 mediated genome engineering. The international journal of biochemistry & cell biology *46*, 49-55.
Zhou, Z., Hartwieg, E., and Horvitz, H.R. (2001). CED-1 is a transmembrane receptor that mediates cell corpse engulfment in C. elegans. Cell *104*, 43-56.