Epigenetic inheritance and DNA replication in *Caenorhabditis elegans*

Kadri Reis

TESI DOCTORAL UPF / 2016

Director
Dr. Ben Lehner
Genetic Systems Group
Systems Biology Department
Centre for Genomic Regulation (CRG)

Universitat Pompeu Fabra Barcelona
GOBIERNO DE ESPAÑA MINISTERIO DE ECONOMÍA Y COMPETITIVIDAD
CRG Centre for Genomic Regulation

Acknowledgements

I feel that one long journey has come to an end, a journey that has introduced me to so many amazing people and also to new unknown aspects of myself.

The biggest thanks goes undoubtedly to Ben Lehner, the supervisor, who without any excessive pampering taught me how to believe in myself without always expecting and receiving praise. Thanks to him and his well-chosen lab members, the lab always felt like home even when the worms showed hostility. Never could have I wished for better co-workers than *the Lehners*, big diverse group of amazing minds that combine into a highly stimulating environment.

Over these beautiful yet sometimes bumpy PhD years, there were so many people who made even the hardest times feel sunny. And I want to thank all of them! Starting with my lovely family and friends back in Estonia who from afar always supported me fully - my parents, my brother, his family, Madli, Margit, and many others. And my “family” here: Adam and Marcos from the worm corner, always-helpful Jenny and Cristina, energetic Cici and André. Outside the lab, Kai my biggest friend and supporter-believer, Heidi with the best beer-talks, Alba, Sarah. And of course my standup comedy scene, who throughout the last

year were the ones to keep me more or less sane for sure! And many-many others with whom we have shared countless hikes, sport events, beers, parties, Impro theater, travels or what not, I am so thankful to all of you!

Abstract

A large proportion of the genome of most higher eukaryotes consists of transcriptionally-repressed repetitive DNA. To better understand how repressed chromatin states are inherited from one generation to the next, we performed a genome-wide RNA interference screen in *Caenorhabditis elegans* to identify genes required for the quantitative repression of an integrated multicopy transgene array in somatic cells. This revealed that inhibition of many components of the DNA replication machinery during early embryonic development leads to a global reduction in levels of the repressive histone post-translational modification H3K27me3 across the genome and a global increase in the levels of the active modifications H3K4me3 and H3K36me3. These results contribute to our understanding of inheritance of chromatin states.

Resumen

Una gran proporción del genoma de la mayoría de eucariotas superiores está formado por secuencias repetitivas de DNA que contienen señales de represión de la transcripción. Para entender mejor cómo funciona la herencia de una generación a otra de esta cromatina reprimida, llevamos a cabo un screening genómico de RNA de interferencia usando *Caernorhabditis elegans* con el objetivo de identificar los genes responsables de la represión cuantitativa de una secuencia integrada en el genoma de células somáticas formada por múltiples copias de un transgén. Así encontramos que la inhibición de muchos componentes de la maquinaria de replicación del DNA durante los primeros pasos del desarrollo embrionario lleva a una reducción de los niveles de metilación post-transcripcional de la histona 3 en su lisina 27 (H3K27me3) a lo largo de todo el genoma y a un aumento también global de los niveles de H3K4me3 y H3K36me3, que son modificaciones de la histona 3 relacionadas con un estado de cromatina activada. Estos resultados contribuyen a conocer mejor la herencia de los distintos estados de la cromatina.

Prologue

In this thesis I present my work carried out with the model organism *C. elegans* to understand how repressed chromatin states are inherited between generations. The main finding is that interfering with DNA replication in an early embryo results in global changes in the levels of histone modifications across the genome. *C. elegans* is a good model for this kind of study, because it is relatively straightforward to perform a genome wide RNA interference screen to identify genes and pathways required for a process. Using an increase in expression from a quantitatively repressed multicopy transgene array as a reporter we discovered that many components of the DNA replication machinery are required for the inheritance of a repressed chromatin state. Using chromatin immunoprecipitation (ChIP) and immunofluorescence revealed global changes in the levels of histone post-translational modifications in embryos after the inhibition of DNA replication with a global reduction in the levels of modifications associated with transcriptional repression and a global increase in the levels of modifications associated with transcription activation. This indicates that the inheritance of repressed chromatin between generations is tightly coupled to DNA replication in this species.

TABLE OF CONTENTS

1 - INTRODUCTION

With advances in techniques in biology and genetics, we keep rethinking and relearning about some historically well-established concepts and questions in evolutionary biology such as neo-Darwinian and Lamarckian theories of inheritance [1]. Lamarck proposed in 1802 [2] that the environment can directly alter phenotypes in a heritable manner. Since A. Weismann formulated the distinction between innate and acquired characteristics at the end of the 19th century, there has been continuous debate about the plausibility of the inheritance of acquired traits. Weismann argued that even though the environment can provoke adaptive responses in the somatic lineage, they could not be communicated to the germline (Weismann, 1891).

Evidence is accumulating, however, that epigenetic states can, at least in some cases, be inherited across generations (reviewed in [3]). In the following introduction I will review what is known about transgenerational epigenetic inheritance of gene expression states in different organisms, and in particular in the nematode *Caenorhabditis elegans*.

1.1. Epigenetics and the 'histone code'

Over time the consensus definition of "epigenetics" has transformed from describing the production of phenotypes from a particular genotype into the more concise "a stably heritable phenotype resulting from changes in a chromosome without alterations in the DNA sequence" as suggested by Berger, *et al*. [4]. Kelly *et al*. discuss that the current epigenetic definition containing the term "heritable" can encompass mitotic stability, meiotic stability or both [3]. They argue that as most mechanisms involved in epigenetic processes influence chromatin structure, it is an indication that chromatin structure (the "epigenome") like DNA sequence itself contains information that guides gene activity and is heritable.

In eukaryotes, DNA and the associated proteins form chromatin, which has to be compacted to fit in the nucleus. The basic repeating unit of chromatin, the nucleosome contains two copies of each of the four histones, H3, H4, H2A, H2B [5]. Nucleosomes are themselves further packaged into higher-order chromatin structures. Chromatin organization protects the genetic information and at the same time provides means for the cell to regulate gene activity The DNA itself and nucleosomal histones are considered the main building blocks of the epigenome. Both DNA and histones can be chemically modified. DNA can be methylated, primarily at CpG sites, which is considered an important epigenetic mechanism in mammals and other

species [6]. However, as in many invertebrates, CpG methylation levels are extremely low in *C. elegans* [7] so it is not considered further in this thesis. Post-translational modifications of amino acids in the N-terminal tails of the core histones (H2A, H2B, H3 and H4) are highly conserved and can obtain number of post-translational modifications [8], such as acetylation and methylation of Lysines (K), and Arginines (R), phosphorylation of Serines (S) and Threonines (T), ubiquitylation and sumolyation of lysines, and ribosylation (Figure 1). These “guidance” modifications are carried out by various chromatin-modifying complexes containing histone modification enzymes, such as histone acetyltransferase (HAT), histone deacetylase (HDAC), histone methyltransferase (HMT) and histone kinase [9, 10], reviewed at [11, 12]. These modifications serve as recognition sites for effector proteins and are able to influence the accessibility of DNA to other multi-protein complexes such as DNA and RNA polymerases. It is suggested that this ‘histone code’ both contributes to and reflects transcription and repression across the genome [9].

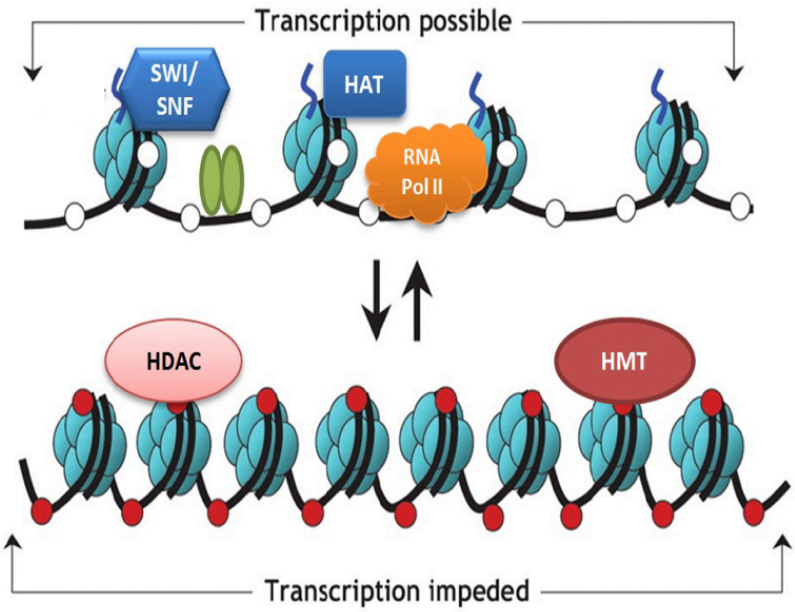


Figure 1. Chromatin remodeling complexes as players in transcriptional regulation, adapted from Luong, P., "Basic Principles of Genetics". In the "Transcription possible" scenario, through the interplay between the SWI/SNF (SWItch/Sucrose Non-Fermentable), HAT (histone acetyltransferase), HMT (histone methyltransferase) and HDAC (histone deacetylase)-mediated modifications, chromatin is loosely packaged and opens DNA regions where transcription machinery (RNA Pol II, transcription factors and co-activators) can bind for gene transcription to occur. In the "Transcription impeded" scenario, the chromatin is packaged more tightly and the transcription machinery is not associated to the chromatin and no transcription occurs.

1.1.1 Histone methylation in transcription repression and activation

In the context of epigenetic memory and changes in chromatin structure, the methylation patterns of the core histone H3 appear to be particularly important. The lysines whose methylation patterns are mostly associated with transcriptional activity and its heritability are H3K4, H3K9,

H3K27 and H3K36 [3]. Most of these lysine residues can be methylated to a different degree (mono-, bi-, trimethylation). For example, in the case of methylation levels at H3 on lysine 4, the mono-methylation is enriched at enhancers, bi- and tri-methylation more at the 5' end of the genes [13]. More importantly, methylation patterns at different lysines in H3 correlate with either activation or repression of transcription. In mice, for example, the transcriptionally inactive pericentric heterochromatin is enriched for the H3K9me3 mark [14-16]. The methylation status recruits specific methyl-binding proteins with specific activation/deactivation functions, for example HP1 proteins that specifically recognize H3K9me2 and H3K9me3 [17, 18]. The HP1 chromo domain is required for both targeting and transcriptional repression. In fission yeast the localization of Swi6, the ortholog of HP1, is dependent on the histone methylase Clr4, the ortholog of mammalian SUV39H1 suggesting a stepwise model for heterochromatin formation where SUV39H1 deposits specific methyl marks on histone H3 that are then recognized by the chromo domain of HP1 [17].

Many current models propose that the initial HP1 protein recruitment is independent of H3K9 methylation, but further spreading is H3K9 methylation-dependent [19-21]. Interestingly, it was recently shown that the distribution of the *C. elegans* HP1 ortholog HPL-2 can be achieved in an H3K9me2-independent manner [22] as HPL-2 persists on

chromatin in mutant embryos lacking H3K9me. This suggests that either the pattern of worm HPL-2 is already established by recruitment and does not involve significant spreading, or the spreading pattern itself is H3K9me1/2/3-independent. This work was done in *C. elegans* mutant strains lacking MET-2 and SET-25, the two H3K9 histone methyltransferases whose role in the heterochromatin formation and function is reviewed in the following chapter.

Additional histone H3 lysine methylations (e.g. mono-, di- and trimethylation of K4, K36 and K79) have been shown to correlate with transcriptional activation, and the enzymes responsible for H3K4me3 and H3K36me3 methylations are physically associated with RNA polymerase II (RNAPII) during elongation [23-25].

1.1.2 Histone variants and nucleosomal occupancy in the dynamics of gene expression

The use of alternate histones provides another level of chromatin modification. For example, *Drosophila* encodes three variants of histone H3, Cid is a structural component of centromeric chromatin, and two of them, H3 and H3.3 differ only at four amino acids [26]. While the major H3 is incorporated only during DNA replication, the variant H3.3 is deposited at particular active loci, is replication-independent, and the inclusion of these histone variants is expected to alter the functional properties [27-29].

Incorporation of H3.3 occurs in the dimer with H4 [30] and this might profoundly change chromatin state of the nucleosome as it brings along the enrichment in post-translational modifications associated with active chromatin and depletion in silent chromatin modifications [31, 32]. H3.3 is also incorporated during the decondensation of the sperm pronucleus in *Drosophila*, suggesting a direct role in chromatin remodeling before fertilization [33].

1.1.3 Histone modifying enzymes in *C. elegans*

The amino acid sequences of *C. elegans* H3 (CeHIS3) and H4 (CeHIS4) proteins are both 97-98% identical to their human counterparts and *C. elegans* has homologs of the mammalian histone modification enzymes, some of which have been studied with genetic and biochemical methodology [34, 35]. *C. elegans* has several chromatin modifiers with human orthologs, such as MES-2 (ortholog of human EZH2), histone methyltransferase (HMT) of histone H3 lysine 27 [34]. In yeast, all H3K36 methylation is carried out by one SET domain-containing protein Set2 [36], whereas in more complex eukaryotes, two different groups of enzymes are involved. One group includes MES-4-related proteins containing SET domain (NSD1, NSD2/WHSC1/MMSET, and NSD3/WHSC1L1 in mammals), with methyltransferase activity in vitro. In *C. elegans*, the transcription-independent MES-4 is required

for H3K36 bimethylation in germline nuclei (both in mitotic and early meiotic germline) [37] and contributes significantly to H3K36me3 as well [38]. The other group is Set2-related proteins, the transcription-dependent MET-1 in *C. elegans* (HYPB/Setd2 in mammals), that participates in transcription-dependent H3K36 methylation in embryos [39] and probably methylates H3K36 on newly expressed genes.

In *C. elegans* the MES proteins are the key chromatin regulators of the germline, where MES-4 participates in silencing of the X-linked genes [37], as in *mes-4* background several X-linked genes even in M+Z- (F2 progeny with maternal load but no expression from the zygotic genome) were up-regulated. They postulate that in wild type worms, MES-4 activates expression of an autosomal repressor that selectively represses genes on the X chromosome. Alternatively, the amount of MES-4/H3K36me2 mark repels a global repressor, thereby concentrating repression action on the X chromosome.

H3K36 methylations are incorporated during transcription in the parental germ cells where it is enriched in autosomes but depleted from X chromatin, correlating with the low levels of transcription on the X in the germ cells [37]. Immunofluorescence of MES-4 itself shows comparable distribution.

The maintenance of epigenetic marks is especially relevant to transmission through the germline, as this information

can potentially regulate the activity of genes across multiple generations. In primordial *C. elegans* germ cells, where transcriptional activity is dependent of H3K36 methylation levels, MES-4 appears to rather maintain than establish the specific H3K36 methylation pattern, that is itself independent on their transcriptional status. In the absence of MES-4 the introduced gametic H3K36me3 is rapidly diluted out. The MES-4 activity pattern in the embryo might reflect the maintenance of H3K36me3 at genetic loci that had been marked by transcription-coupled H3K36 methylation in the preceding generations [40].

Another possible role for MES-4 is to limit the spreading of the repressive modifications, such as H3K27me3 [41]. PRC2, a *C. elegans* Polycomb group repression complex 2 (including MES-2/-3/-6) is the main mediator of H3K27 methylation patterns [34]. ChIP-chip analysis of H3K36me3 and H3K27me3 patterns in early embryonic chromatin reveals that these inherited marks are in large scale mutually exclusive [41].

Another important component of epigenetic regulation is methylation of histone 3 lysine 4 (mono-, di-, trimethylation of H3K4me), that is a mark of transcriptional activity, but is also shown to participate in “epigenetic memory” – it can pass on the transcriptional memory of the loci to subsequent cell descendants or lineages [3]. There are several multi-protein complexes described in different organisms that control the H3K4me/me2/me3 patterns. In

yeast, the enzyme responsible for all H3K4 methylation is Set1 that is in the COMPASS (complex proteins associated with Set1p) complex [42]. Homologous proteins are found in *C. elegans*, referred to as Set1/MLL (after mixed-lineage leukemia from COMPASS mammal complex) [43], [44], [45]. As in mammals, the loss of Set1 homolog in *C. elegans* results in decrease in H3K4me3 [44]. In the germline, the maintenance of H3K4me2/me3 is dependent on the Set1/MLL complex components WDR-5.1 and RBBP-5 and it is proposed that H3K4 methylation is required to maintain the totipotent epigenome when passed through the germ line between generations [44].

1.1.4 Transgenerational epigenetic memory

Epigenetic information is by nature metastable and requires mechanisms to persist within and between generations. During germline development, chromatin structure undergoes changes as observed in meiotic chromosomes and during spermatogenesis. Also, as the gamete genomes are met by epigenetic reprogramming mechanisms upon fertilization, active mechanisms are required in order for the information from the parental epigenomes to persist in the germline and some of the offspring (Figure 2) [46].

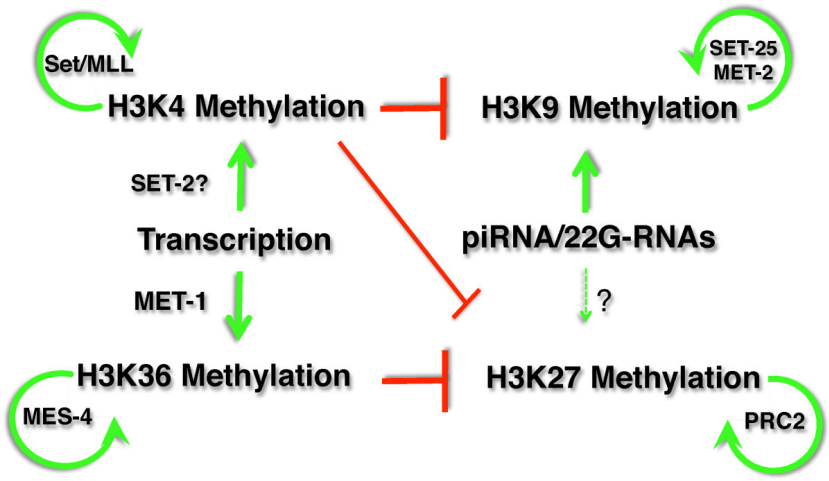


Figure 2. Methylations of histone H3 seem to play the key role in epigenetic memory and chromatin structural alterations. The methylation on one lysine can influence the modification on other lysines, resulting in a regulatory network. The main antagonistic relationships between histone H3 methylation on Lys4, Lys9, Lys27, and Lys 36 are depicted. In *C. elegans,* the antagonism between methylation levels of H3K36 and H3K27 (red lines) has been shown [41], whereas the antagonism between methylations of H3K4 and H3K9/H3K27 has been studied in many model organisms. The straight green arrows indicate mechanisms for establishment; the curved green arrows show mechanisms of maintenance (figure from [3]).

As argued in [3], the opposing activities of MES-4 and PRC might be in charge of maintaining heritable epigenomic patterns. When active repression is missing, germline-expressed genes might become active in any tissue. This line of reasoning suggests that the heritable epigenetic patterns are mainly generated by transcription-dependent H3K36 methylation in adult germ cells and can persist in the chromatin of the gametes and are maintained in the zygote by MES-4. The H3K36 methylation marks are also able to prevent invasion.

In this light, it is somewhat surprising that neither H3K36me nor H3K27me marks patterned by the opposing activities of MES-4/PRC2 are considerably affected by reprogramming mechanisms. It is possible that these modifications are resistant to the reprogramming or are continuously re-established. Another epigenetic memory-associated modification, H3K4 trimethylation is noticeably depleted during Z2/Z3 reprogramming [47] and the methylation patterns of H3K4 have been implicated in contributing to transgenerational phenotypes.

Similar to H3K36me3, the transcription-dependent H3K4-methylation requires mechanisms for its maintenance. As mentioned before, *C. elegans* has homologues of the complexes responsible for H3K4 methylation, such as COMPASS responsible for all H3K4 methylation in yeast, and MLL (mixed-linage leukemia) in mammals. Mutations in components of these complexes cause a substantial depletion of H3K4me2/me3 in the early embryo [43, 44]. H3K4 methylation in the early blastomeres appears to be largely transcription-independent as depletion of RNA Pol II does not cause a significant reduction in H3K4me2/me3 levels in early embryos [44], suggesting that H3K4 methylation in these early stages is rather due to the maintenance of this modification. Similar to MES-4 H3K36me3 patterns, the MLL-dependent H3K4me patterns in embryos seem to be inherited and maintained through gametes [44, 48].

The substantial MLL-dependent and transcription-independent H3K4 methylation is contributing to epigenetic information and is inherited by the offspring, as readily observable in sex-specific X-chromosome epigenetic profiles. In *C. elegans*, there is a clear X chromosome bias; genes that are expressed in germ cells of both sexes are not found on the X chromosome [49, 50]. The X chromosome is transcriptionally inactive during meiosis in both germlines, but becomes active during female gametogenesis, female germ cells showing a significant increase in H3K4me levels during oogenesis [50]. This chromatin memory is persistently transmitted to the next generation.

Although the components of the pathways covered here are highly conserved between organisms, their precise relevance in understanding transgenerational phenomena in other species remains unclear. All metazoans have MES-4 and PRC-2 related enzymes and MLL type complexes exist in all eukaryotes.

1.1.5 Transgenerational inheritance of acquired transcriptional patterns

In principle, germline reprogramming and epigenetic inheritance are two opposite phenomena. Reprogramming in the germline allows totipotency of the zygote, required to remove epigenetic signatures that have been acquired

during development. If the germline reprogramming does not occur, the epigenetic marks can be transmitted to the next generation, whereas these epialleles can be potentially neutral, deleterious or adaptive. For example, induced expression of a transgene expressing DNA from the Flock House virus in the *C. elegans* soma results in transmission of the silencing of the viral genome for many subsequent generations [51].

1.2. Inheritance of gene silencing and activation by small RNA pathways across generations in *C. elegans*

1.2.1 Small RNAs in *C. elegans*

C. elegans produces thousands of small RNAs that target coding genes, pseudogenes and other non-coding RNAs. These endogenous small RNAs are categorized into three main classes: microRNAs, endogenous small interfering RNAs (endo-siRNAs) and Piwi-acting RNAs (piRNAs). All of these bind Argonautes, effector proteins that regulate their bound targets mainly by inhibition. The antisense RNAs are able to silence genes very effectively. Already in 1991 [52] it was determined that the interfering agent is double-stranded RNA (dsRNA) from that gene, acting through degradation of targeted mRNA and that there is probably a

catalytic or amplification method that would explain its vast silencing potential and heritability. dsRNA, the trigger for RNAi is substrate for DICER, RNase III enzyme [53] that cleaves the dsRNA into primary short interfering RNAs (siRNA). Loss of DCR-1, the only ortholog of Dicer in *C. elegans*, leads to severe cell fate specification defects and germline abnormalities resulting in sterility and lethality [54], [55]. In addition to the primary siRNAs, the production of secondary siRNAs by RNA-dependent RNA polymerase (RdRP) directs the silencing to its target sequences [56]. The *C. elegans* RdRP EGO-1 is required for germline development [57]. These secondary RNAs interact with a number of Argonaute (AGO) proteins [58, 59].

Exogenous RNAi (exo-RNAi) may have a role as an anti-viral silencing response [60, 61]. The silencing mechanisms initiated by exo-RNAi engage a downstream endo-RNAi amplification pathway that is also shared with endogenous siRNAs (endo-siRNAs) and piwi-interacting siRNAs (piRNAs) (summarized in Figure 4). Before the discovery of endogenous silencing pathways, isolation of *C. elegans* mutants defective in both exo-RNAi and endogenous silencing phenomena (e.g. transposon silencing) suggested a shared mechanism [62], [63], [64]. There have been many studies investigating mechanisms of RNAi (reviewed in [65-68]) and many of the pathways are highly conserved between species.

1.2.1.1 Endogenous and exogenous small RNAs

The *C. elegans* small RNAs fall into three main classes, based on their size and function (Table 1) [65]:

(1) siRNAs (endo- and exo-siRNAs for endogenous and exogenous RNAs, respectively) that are 22-26 nucleotide sequences. The exo-siRNAs are induced by long dsRNA originating from virus-derived RNA, cellular transfections, microinjections or feeding with bacteria expressing dsRNA. The endo-siRNAs on the other hand, target RNAs produced by the worm genome itself and modify gene expression by degrading the transcripts by translational inhibition or chromatin modifications that will be discussed later.

(2) miRNAs that are also 22nt RNAs regulating gene expression during the development, differentiation and antiviral responses by altering the expression of other RNAs but are genomically encoded [69].

(3) PIWI-interacting RNAs (piRNAs), 21 and 22nt small RNAs, expressed in the germline that regulate germline development and transposon silencing during gametogenesis in order to protect genomic integrity and have been associated with transgenerational silencing [3, 69].

Small RNA	Argonautes and RISC components	Biogenesis factors
ERGO-1 26G	ERGO-1	DRH-3, ERI-5, DCR-1, ERI-1b, ERI-3, RDE-4
ALG-3/-4 26G	ALG-3/-4	DRH-3, ERI-5, DCR-1, ERI-1b, ERI-3, RDE-4
WAGO 22G	WAGOs	RRF-1/EGO-1, DRH-3, EKL-1
HRDE-1 22G	HRDE-1, NRDE-1, 2, 4	
CSR-1 22G	CDE-1	EGO-1
21U/piRNAs	PRG-1	FKH-3/-4/-5, UNC-130, PRDE-1
miRNAs	ALG-1/-2, AIN-1/-2, NHL-2, CGH-1, TSN-1, VIG-1	DCR-1, DRSH-1, PASH-1

Table 1 Table showing different classes of *C. elegans* small RNAs based on the size and function, showing only a subset of their diversity and biogenesis factors.

Interestingly studies have shown that the mechanisms through which these miRNAs, siRNAs and piRNAs act are very different in terms of their biogenesis, their protein co-factors and effector function [65]. All these classes of small

RNAs interact with Argonaute proteins to recognize their specific RNA transcript targets and through often-imperfect complementarity regulate their expression, usually though inhibition. In *C. elegans* 27 different Argonautes are described that are generally grouped into three classes: the Argonaute-like proteins, the PIWI-like proteins and worm-specific Argonautes (WAGO) [3]. What distinguishes the cytoplasmic siRNA pathway from that of the nuclear in somatic cells, is the Argonaute proteins involved: the 22 nucleotide siRNAs associated with WAGOs target mRNAs in the cytoplasm, and the siRNAs associated with NRDE-3 (in the soma [70]) or HRDE-1 (in the germline [59]) Argonaute function in the nucleus where they regulate silencing of their cytoplasmic targets through transcription inhibition or chromatin remodeling. The nuclear pathways will be discussed in the following chapters in more detail.

1.2.1.2 Diversity of small RNAs in C. elegans, the 26G, 22G, 21U RNAs

Thanks to the advances in deep sequencing techniques we can distinguish diverse sub-populations among the *C. elegans* small RNAs [71]. The pools identified dispose distinct 5' nucleotide bias and are categorized into different subpopulations depending on their length (22nt and 26nt). These subgroups are divided into smaller groups with overlapping requirements for biogenesis but different engagement pathways determined by the Argonautes to

which they are bound. There are 27 Argonaute proteins encoded in *C. elegans*, about half of them being worm-specific Argonautes (WAGOs) that bind the 22-nt RNAs. Another subgroup of small RNAs identified are the 21nt RNAs, later determined as the piRNAs (Piwi-interacting RNAs) of *C. elegans*. The next subsections will briefly cover these subgroups, their biogenesis and function.

1.2.1.2.1 26G RNAs

The distinct population of 26G RNAs was revealed by deep sequencing [71], these RNAs are anti-sense to annotated genes and carry a 5' guanosine bias. These 26G RNAs are enriched in male and female germline and are bound by Argonautes ALG-3 and ALG-4 in the spermatogenic gonad, while in oogenic gonad they are bound by the ERGO-1 Argonaute [72], [73], [74] that are also abundant in embryos. The biogenesis of the 26G RNAs is mediated by the ERI (enhanced RNAi) complex with a core RdRP (RNA-dependent polymerase) module consisting of RRF-3 (RNA-dependent RNA polymerase family), DRH-3 (Dicer Related Helicase) and ERI-5, with DCR-1 association as well. This ERI complex shares DCR-1 and RDE-4 with the RDE (exogenous RNAi defective) complex that mediates processing the exo-RNAi pathway dsRNAs (Figure 3).

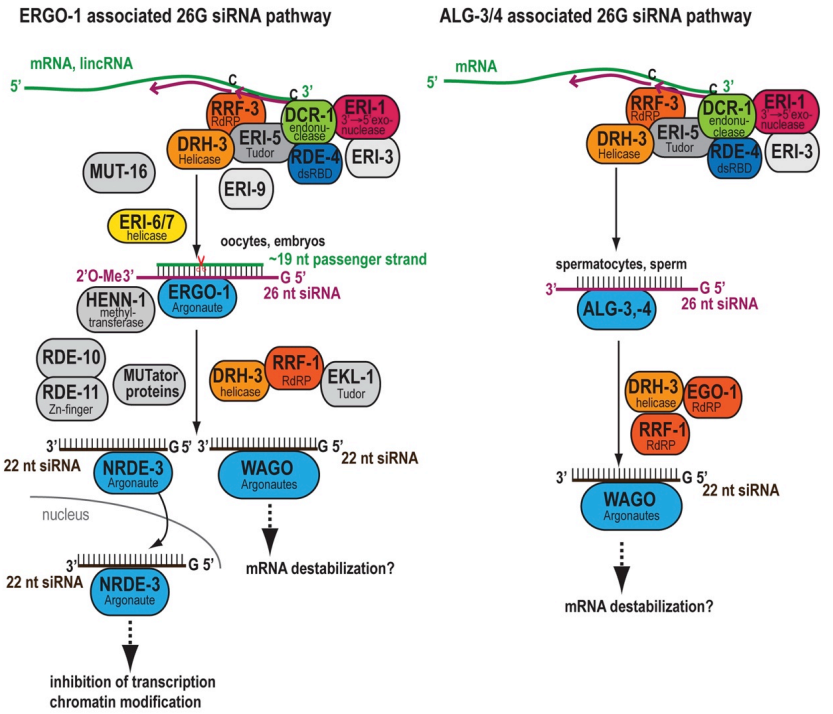


Figure 3. ERGO-1 and ALG-3/-4 siRNA pathways. The templates for the 26G siRNAs are mRNAs and lincRNAs. These 26G siRNAs, generated by the complex containing an RNA-dependent RNA polymerase (RdRP) and DCR-1 (*C. elegans* ortholog of Dicer) in both the oocytes/embryos and spermatocytes/sperm, associate with the Argonautes ERGO-1 and ALG-3/-4, respectively. This triggers the generation of the 22G siRNAs by another RdRP complex that then associate with WAGOs to silence their targets in the soma through nuclear cytoplasmic silencing pathways. Different colors denote the function of the proteins, the key players Argonautes in blue, RdRPs in red (adapted from Wormbook).

Even though during biogenesis of these 26G RNAs, DCR-1 catalyzes the cleavage of the mRNA template and reduces its levels by generating the dsRNA intermediate [75], the WAGO 22G RNAs triggered by 26G RNAs are the major effectors of the 26G RNA endo-RNAi pathway. So these secondary siRNAs are strictly required for 26G RNA target silencing [73], [74].

1.2.1.2.2 22G RNAs

In *C. elegans*, exo-RNAi and endo-RNAi converge at a common downstream pathway in silencing their targets, the WAGO 22G RNA pathway (Figure 4). The 22G RNAs are 22 nucleotide long RNAs showing 5' prominent guanosine bias [76], [71]. These 22G RNAs map antisense to mRNA [58], whereas most 22G RNAs are germline expressed and deposited into embryo. There are two distinct classes of small RNAs that engage unique pathways mediated by specific Argonautes to effect distinct outcomes. The WAGO-binding 22G RNAs mediate silencing of certain protein-coding genes, transposons, pseudogenes, and cryptic loci through transcriptional and post-transcriptional mechanisms [77], [58], [78], [79], [80], [59], [81], [82]. CSR-1, the other WAGO binds RNAs targeting germline-expressed genes promoting chromosome segregation, a specific role of these 22G RNAs that will be discussed later in more detail [69].

Two RdRPs, RRF-1 and EGO-1 contribute to the biogenesis of 22G RNAs that serves as an amplification pathway upon which most of the primary small RNAs in *C. elegans* converge. WAGO 22G RNAs with 5' triphosphate are generated downstream of the 26G RNA, primary exo-siRNA, RDE-1 small RNA and 21U RNA pathways.

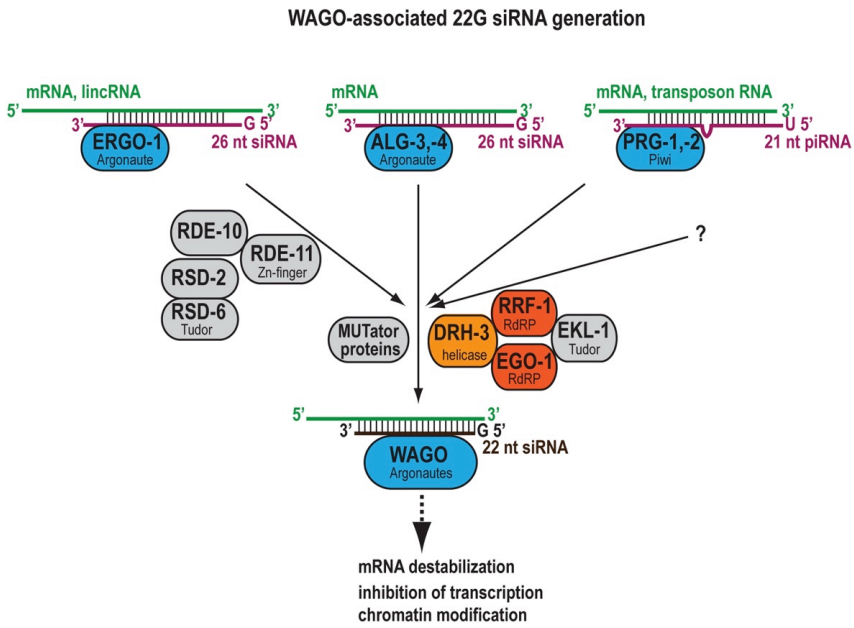


Figure 4 Different classes of small RNAs that trigger secondary siRNA generation. 26G siRNAs and 21U/piRNAs trigger the secondary 22G siRNA generation by the RdRPs RRF-1 and EGO-1 with the helicase DRH-3 and the Tudor domain protein EKL-1. The MUTator proteins are required for 22G siRNA production in the germline and the RDE-10/-11 complex is required to promote secondary 22G siRNA amplification in the soma. These secondary siRNAs are then able to destabilize mRNA, inhibit transcription and modify chromatin. Blue balloons depict the main Argonautes in these pathways (Adapted from Wormbook).

1.2.1.2.3 21U RNAs

Deep sequencing identified a third pool of small RNAs, the 21nt length small RNAs with 5' Uridine bias without further common sequence features [71]. Unlike endo-siRNAs, the 21U RNAs are not generated from mRNA templates but rather transcribed directly from genomic loci [71], [81] and are mostly depleted for overlap with exons and these 21U RNAs show no sequence conservation [71], [83]. These

RNAs are highly abundant in the germline and in embryos, with declining levels across development [84], [85], [86] and distinct subpopulations in the male and female germline. These 21U RNAs target transcripts by directing Argonuate PRG-1 to their targets that are depleted of protein-coding transcripts [87]. The 21U RNA targeting triggers production of WAGO 22G RNAs and these 21U RNAs are capable of triggering transgenerational silencing [80], [88]. The 21U RNAs are encoded as independent Pol II transcriptional units [81], [89] and associate with the Piwi protein PRG-1 to trigger secondary 22G siRNA production.

1.3. WAGO 22G RNA transcriptional silencing, the nuclear RNAi pathway

There appears to be two distinct forms of inherited RNAi: inheritance of somatic RNAi for one or a few generations and inheritance of germline RNAi, over multiple generations (Figure 5).

The WAGO Argonaute NRDE-3 (Nuclear RNAi defective) pathway is the sole mediator of transcriptional gene silencing in the soma. NRDE-3, which contains a nuclear localization signal (NLS) is triggered by WAGO 22G RNA binding to enter the nucleus and to bind its targets. Loss of

nrde-3 leads to accumulation of pre-mRNA and mRNA levels of its targets [77]. NRDE-3 is required for the inheritance of somatic silencing triggered by exo-RNAi for a single generation and in the progeny NRDE-3 is responsible for the accumulation of secondary WAGO 22G RNAs to re-establish H3K9me3 marks [79].

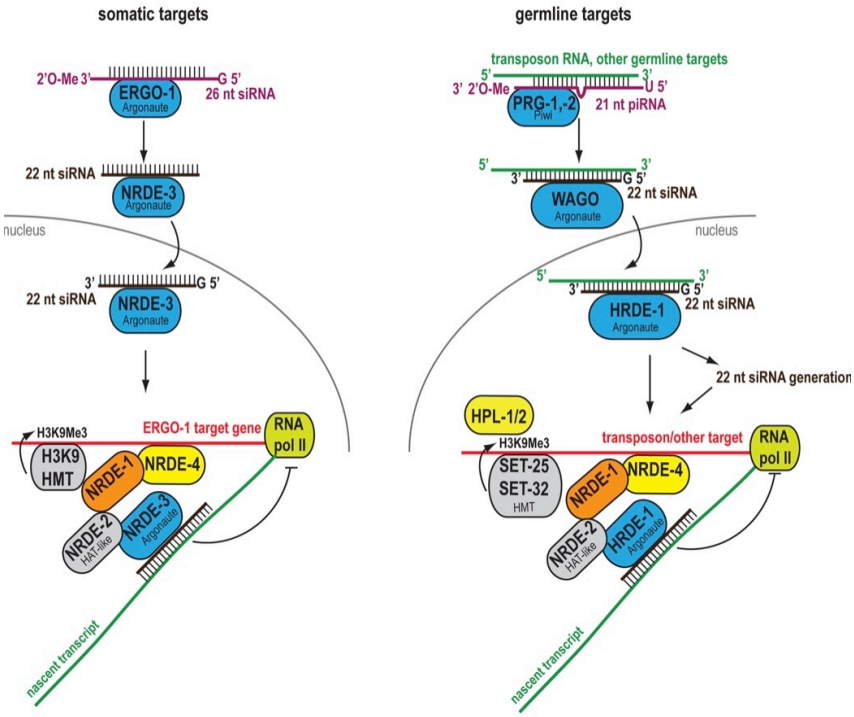


Figure 5 Nuclear RNAi pathways for trans-generational silencing. The somatic (left) and germline (right) silencing pathways converge on the same NRDE proteins but require distinct nuclear Argonautes (in blue): the NRDE-3 in the soma and HRDE-1 in the germline, both of which associate with the cytoplasmic secondary 22G siRNAs. In the nucleus the complex of Argonaute and siRNAs locate to the nascent pre-mRNA. NRDE-1 together with NRDE-4 promote H3K9 trimethylation (H3K9 HMT, in grey) and thus inhibit transcription. The NRDE proteins are also able inhibit RNA Polymerase II. This nuclear RNAi is heritable across generations: silencing signals are inherited (e.g. siRNAs generated in the germline of the parent) are inherited and direct the H3K9 trimethylation in the offspring (mod. from [65]).

We now have accumulating evidence that some epigenetic phenomena involving RNA, histone modifications and/or DNA methylation that suggests that silenced allele could act in trans on a homologous sequence and cause a stable and heritable silencing, an example of paramutation. Several studies have now reported inheritance of environmental RNAi beyond the F1 generation [90], [91], [92], [93], [81]. Already the first experiments introducing foreign double stranded RNA into *C. elegans* showed effects of inference in both injected animals and their progeny [90]. It was especially surprising that this interference can persist into the next generation, even though many of the endogenous RNA transcripts are degraded in the early embryo [94].

In contrast to NRDE-3, HRDE-1 engages the nuclear RNAi pathway in the germ cells to direct the silencing inheritable across multiple generations (also called RNAe for RNA induced epigenetic silencing) [80], [87], [59], [95] [88], [82]. This transgenerational silencing can be triggered by exo-RNAi to establish transcript silencing that is stable for several succeeding generations without the necessity of the initial trigger exposure [80], [59], [95]. Endogenous small RNA pathway engage the germline nuclear RNAi as well, 22G RNAs that are bound by HRDE-1 and WAGO-1 overlap to a large extent [82], suggesting that nuclear and cytoplasmic WAGO Argonautes share common siRNA targets and cofactors. For example, the WAGO 22G RNAs

triggered by 21U RNAs and 26G RNAs that are in charge of maintaining germline integrity. These 21U RNAs encode fertility-promoting small RNAs that associate with HRDE-1 and are hypothesized to encode an epigenetic memory of 'non-self' required for genome surveillance [80], [87], [96], [82]. Loss of HRDE-1 results in progressive sterility – a moral germline (Mrt) phenotype [59].

Early experiments suggested that the exo-RNAi pathway acts mainly through post-transcriptional silencing, as injection of dsRNAs corresponding to introns or promoters did not result in efficient silencing [90]. Grishok *et al*. [91] examined the properties of inheritance associated with long-lasting effects of RNAi phenomenon triggered by exogenous double-stranded RNA. In the study they describe two different classes of genes, one of which involves genes needed for the production of the heritable factor of RNAi. A previous study had identified these two sets of *C. elegans* genes as requirement for RNAi [62], one class containing *rde-1* (an Argonaute protein of the PAZ-PIWI family) and *rde-4* (a dsRNA binding protein), that are deficient in RNAi but do not display other phenotypes. The second class, with *rde-2*, *rde-3*, *mut-2*, *mut-7* is essential for the response to the heritable component of RNA and show transposon mobilization, reduced fertility, and high incidence of chromosome loss. *mut-7* was already identified before [63], along with the first genetic screen for exo-RNAi-defective (Rde) mutants [62]. As *rde-1* and *rde-4* are

dispensable for the inheritance of the RNAi effects, RNAi itself does not seem to underlie the mechanism of inheritance. Another study using a candidate gene RNAi screen identified *hda-4* (a class II histone deacetylase), K03D10.3 (a histone acetyltransferase of the MYST family), *isw-1* (yeast chromatin-remodeling ATPase ISW1 homologue) and *mrg-1* (a chromo-domain protein) as genes affecting the maintenance of silencing [92]. Importantly, as these genes are all implicated in chromatin remodeling, the inheritance of RNAi-induced changes occurs at the transcriptional level.

Ashe *et al*. [80] (and others [59, 97, 98]) report how transgenerational inheritance of environmental RNAi and the piRNA pathway converge at one germline nuclear RNAi/chromatin pathway, whereas the nuclear RNAi factors and chromatin regulators are both essential for silencing (Figure 6). This epigenetic memory can last for many generations and once established, the initial silencing trigger in not required. They argue that as chromatin factors such as HPL-2 and the putative H3K9me3 methylatransferases SET-25 and SET-32 are required for the silencing, the chromatin changes observed in the transgenerational silencing are not merely correlative.

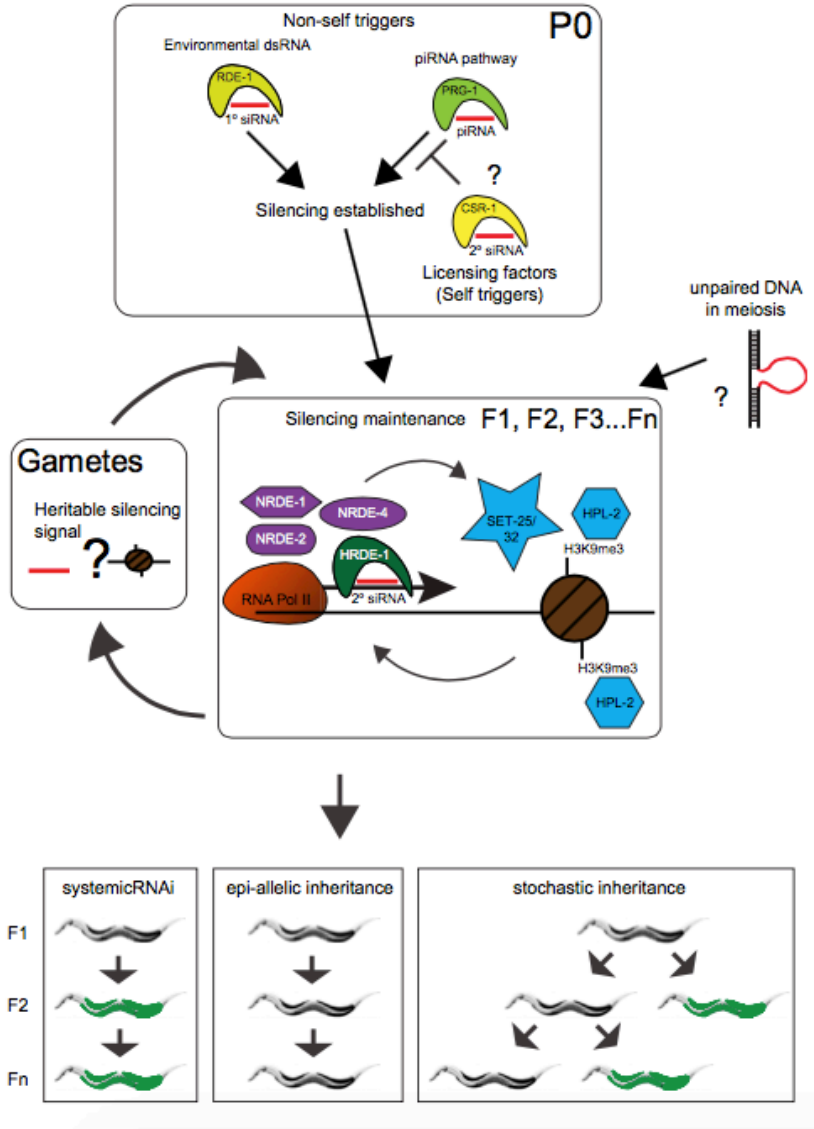


Figure 6 Model of transgenerational silencing in the *C. elegans* germline. Environmental RNAi and endogenous piRNAs establish a nuclear RNAi/chromatin pathway (in P0). Maintenance of the silencing depends on nuclear RNAi factors, such as germline-specific Argonaute HRDE-1/WAGO-9, and chromatin proteins, HPL-2 and histone methyltransferases SET-25/-32. Silencing can be transferred to many subsequent generations (F1, F2, Fn) via heritable silencing signal. Or it can become epi-allelic or multigenerational stochastic inheritance. Silencing can me suppressed by germline licensing pathway that recognizes germline transcripts (see the CSR-1 licensing chapter) [80].

1.3.1 Stable silencing through PIWI-interacting RNAs (piRNA)

Highly conserved PIWI-associated small RNAs (piRNAs) are suggested to be key players in the transmission of the repressive epigenetic memory across multiple generations. piRNAs in *C. elegans* are associated with orthologs of the Drosophila PIWI protein, an Argonaute mediating small RNA-associated processes [99]. PIWI proteins are a group of Argonaute family, mostly expressed in the germline [100, 101] that identify their targets by base-pairing. Some PIWI proteins function mostly in the cytoplasm and involve degradation of mRNAs [102-104] while others translocate to the nucleus, suggesting a transcriptional silencing similar to non-piRNA pathways described previously in fission yeast. *C. elegans* PIWI homolog *prg-1* mutants lack all detectable piRNAs, 21-nucleotide RNAs with 5'U (21U RNAs). piRNAs with known targets generate secondary 22G-RNAs (22-nucleotide small RNAs with 5'G) [96], produced by endo-siRNA pathways. PIWI pathway can affect the chromatin structure of its target loci [80, 82, 88]. Small secondary RNAs (22G RNAs) are synthesized by RNA-dependent RNA polymerases (RdRPs) that are then loaded onto worm-specific Argonautes as WAGO-9/HRDE-1 and transported together to the nucleus. With the help of nuclear RNAi factors NRDE-1/-2/-4 proteins they can trigger transcriptional silencing [59, 78-80, 82], followed by trimethylation of H3K9 at the target genomic loci [80, 82,

88]. This silenced state can be stably inherited across generations, process called RNA-induced epigenetic silencing (RNAe). This transgenerational effect has been shown to be dependent on the WAGO-9 and NRDE factors to be transmitted through meiosis [80, 82, 88]. Also, WAGO-9 mutants become sterile after several generations, a sign of germ cells losing their characteristic immortality, accompanied by the loss of H3K9me3 patterns at the target sites [59, 79].

The piRNA pathway overlaps with pathways that are essential for stable silencing of single copy transgenes in the germline across generations [80, 82, 88, 96]. Small RNAs are shown to be essential for the initiation of the silencing, but not for the heritable maintenance of the repression. Rather, nuclear RNAi pathway and chromatin modifying factors are essential for the stable multigenerational repression [88]. Regulation of transcriptional repression through chromatin structure is shown to involve RNAi mechanisms in many organisms [105, 106], initiating targeting of the repressive chromatin machinery to genomic loci and defects in these mechanisms often lead to derepression of transposons and repetitive elements. The model of transcriptional silencing guided by small-RNAs in *C. elegans* is summarized in Figure 7 [107].

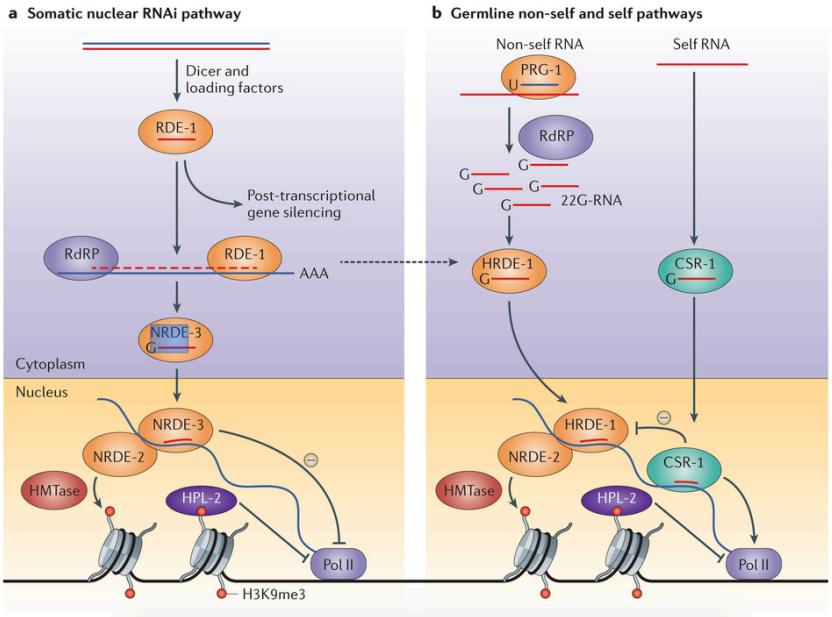


Figure 7 Transcriptional silencing guided by small RNAs in *C. elegans*, A) exogenous dsRNA is processed into primary small interfering RNAs (siRNAs) that are loaded onto Argonaute RDE-1 and amplified by RdRPs to generate secondary siRNAs, the 22G-RNAs. The somatic AGO NRDE-3 loads the 22G RNAs and translocates to the nucleus where with the silencing factor NRDE-2, it can silence the genes through targeting the nascent RNA transcripts. Gene expression is paused by NRDE-2 during transcriptional elongation and silencing is involving H3K9me3 and recruitment of heterochromatin-protein-like HPL-2. B) In the germ line, small RNA-guided transcriptional silencing is mediated by HRDE-1, an Argonaute that also acts with NRDE-2, H3K9me3 and HPL-2 in the nucleus. HRDE-1 receives 22G-RNA from both, the exogenous dsRNA pathway and from the PIWI- or 21U-RNA (piRNA) pathway. 21U RNAs with PRG-1 promote the RdRP-dependent generation of 22G-RNAs that are then loaded onto HRDE-1, that is then able to maintain a persistent transgenerational memory of silenced genes in the germline. Another Argonaute, the CSR-1 binds 22G RNAs that are complement to endogenous RNAs and protects the corresponding loci from silencing by HRDE-1. Thus, the 22G-RNAs with CSR-1 and HRDE-1 transmit a germline memory of 'self' and 'non-self' RNAs, to be silenced or licensed for expression (figure adapted from [107]).

1.3.2 Licensing of genetic activity by maternal transcripts

Whereas most of the described regulatory roles of RNA involve down-regulation, there is an alternative role for maternal transcripts in promoting zygotic activity of that gene, termed gene licensing. Maternal mRNA can regulate the expression of its cognate gene in the germline of the zygote independent of translation [108]. Maternal transcript of *C. elegans fem-1*, a messenger in the sex determination pathway, is essential to license the expression of the *fem-1* gene in the germline of the zygote [108]. Heterozygous offspring from the homozygous *fem-1* mutant mothers show reduced *fem-1* activity, and injection of *fem-1* RNA is capable of rescuing the defect in the progeny without previous translation into the protein in the maternal germline. As the defect of the zygotic *fem-1* expression is heritable, the *fem-1* gene requires prevention from epigenetic silencing by maternal *fem-1* transcripts.

1.3.2.1 Licensing of germline transcripts

It appears that *C. elegans* the germline is labeling most of the DNA as “bad” and only actively licensed genes are expressed. Small RNAs are proposed to be used as molecular memory for both the bad and the good. As mentioned previously, the foreign unwanted transcripts are

recognized by the Argonaute PRG-1 via Piwi-interacting RNAs that guide identifying its own targets. PRG-1 pathway triggers the recruitment of the RNA-dependent RNA Polymerase that generates a new population of small RNAs (22G RNAs) that are loaded on further Argonaute proteins WAGO-9/HRDE-1 and WAGO-10. This type of silencing is very stable and maintained for many generations as explained previously. This process is initiated by PRG-1 and then maintained by HRDE-1 [80, 82, 88]. There are more than 30,000 annotated piRNAs [81] and they are thought to be able to trigger repression via partial complementarity to mRNA targets. As such, almost any foreign sequence can be recognized. This raises an obvious problem, as this complex would be able to target most of *C. elegan's* own genes. What is the mechanism that allows endogenous germline transcripts to avoid silencing by the piRNA-mediated germline surveillance pathway that has such a vast silencing potential? One likely candidate is the Argonaute CSR-1 that interacts with 22G-RNAs antisense to most of the germline-expressed genes. While Seth *et al*. [97] showed that CSR-1, but not WAGO-1 or HRDE-1 was found to bind 22G RNAs from active transgenes, the experiments were more suggestive and correlative. Wedeles *et al*. [98] showed that the licensing recruitment of CSR-1 to a transcript protects it from piRNA-mediated silencing. They tethered CSR-1 onto a transcript, triggering its activation. In addition, this tethering of CSR-1

results in the buildup of a diffusible agent that targets CSR-1 to another homologous transgene. This heritable transcriptional licensing is able to protect germline transcripts from being silenced.

1.3.3 Extended lifespan in *C. elegans*

Greer *et al*. reported inheritance of lifespan extension in the descendants of worms deficient for complex depositing H3K4me3. There had been previous reports linking chromatin and aging, mostly through histone deacetylation by the Sir2 family [109, 110]. They looked at histone methylations for their role in development and maintenance of stem cell pluripotency in mammals. They showed that ASH-2 trithorax complex [111], trimethylating H3 at Lys4, is regulating *C. elegans* lifespan, and the extended longevity following trithorax complex depletion is inherited for several generations. Mutations in the ASH-2 complex (WDR-5, SET-2 and ASH-2 itself) extend the lifespan of the worms, with inhibition of the H3K4 demethylase RBR-2 suppressing this. So, the transgenerational inheritance of longevity could result from heritable depletion of H3K4me3 at some gene loci responsible for regulating aging. Interestingly, chromatin modifiers that regulate longevity via other pathways than H3K4 methylation or outside of the germline

(UTX-1, SET-9, SET-15) have effects on longevity that are not transgenerationally inherited.

1.3.4 The Immortal germline phenotype

Epigenetic modifications are able to limit the developmental potential of tissues during differentiation. During early embryogenesis the *C. elegans* germline (the P lineage), contains high levels of the “active” mark H3K4me2 [112], until dividing into two primordial germ cells (PGCs) Z2 and Z3. After the PGCs are generated and committed to the germ cell fate, the chromatin quickly loses the H3K4me2 marks, possibly to protect and maintain PGC fate via transcriptional repression. The erasure of the H3K4 methylation could include active processing by histone demethylases, three homologous of the mammalian LSD1/KDM1 demethylase being found in *C. elegans*. Mutants of *spr-5*, the *C. elegans* ortholog of LSD1, leads to germline mortality with the incidence of sterility increasing over generations (Mrt phenotype) [113]. This sterility correlates with the mis-regulation of genes expressed during spermatogenesis due to the stable accumulation of H3K4me2 mark at these loci, resulting in inappropriate maintenance of H3K4me2 in the PGCs and faulty oogenesis and spermatogenesis. This together suggests that the *spr-5* demethylase is required for removing

epigenetic information acquired from the parental germline and the failing to do so lead to accumulation of epigenetic memory [11]. The ability to epigenetically reprogram the germ cells' genome to maintain the totipotency might be the main distinction between the "mortal" soma and the "immortal" germline. Though the role of RNAi inheritance in transgenerational epigenetic inheritance is evident in many species [114], these RNAi pathways do not seem to play a role in the *spr-5*-induced epigenetic inheritance. The Argonautes needed for the main RNAi pathways in *C. elegans*, RDE-1 (exogenous RNAi pathway) [91] and ERGO-1 (endogenous) [115], are not required for the transgenerational inheritance of sterility of *spr-5* mutants [116].

Similar transgenerational replicative aging of germ cells was also observed with *prg-1* mutants [117], suggesting that PRG-1 and associated piRNAs possibly function upstream of nuclear RNAi factors promoting germ cell immortality and that *prg-1* is required for silencing of some endogenous loci.

1.4. RNA inheritance in other model organisms

Another common model organism for studying RNA interference is fission yeast, *S. pombe*, with single copies of the RNAi pathway components such as Argonaute, Dicer and RNA-dependent polymerase RdRP. The heterochromatic centromeres of fission yeast are a good

model for heterochromatic silencing studies [118] being subject to epigenetic regulation, involving the propagation of meta-stable chromatin states. The acetylation state of the histones may serve as a platform for the assembly of the centromeric factors in *S. pombe*. Already in 1977 Ekwall *et al*. [119] showed that treating the *S. pombe* cells with deacetylase inhibitor trichostatin A (TSA) results in hyperacetylated centromeric heterochromatin. This also correlated with derepressed marker genes at the centromeric heterochromatin and with defective chromosomal segregation. This highly acetylated state and defective repression was maintained up to 80-100 generations after the TSA was removed. Therefore, the acetylated histone state may act as a platform for the centromeric factors in the fission yeast and this state my act as a chromosomal imprint for its inheritance of the normal hypo-acetylated state of heterochromatin, both telomeric and centromeric. In addition to acetylation, other types of covalent modifications (phosphorylation of H3 at Ser10, H3-S10, and methylation of H3 and H4) support the histone code hypothesis through the interplay between these different covalent modifications [9]. This hypothesis suggests that a pre-existing modification affects subsequent modifications on histone tails and these consequently act as marks for the recruitment of protein complexes to regulate the chromatin functions such as gene expression, DNA replication, and chromosome segregation [11].

Deletions of the Argonaute (*ago1*), Dicer (*dcr1*), and RNA-dependent RNA polymerase (RdRP) (*rdr1*) in fission yeast exposed that centromeric silencing requires RNAi, resulting in accumulation of heterochromatic centromeric repeat transcripts [120]. This is accompanied by transcriptional de-repression of centromeric transgenes and reduction of H3 Lys9 methylation. Based on this, Volpe *et al.* proposed that dsRNAs from centromeric repeats target the formation and maintenance of heterochromatin via RNAi [120].

1.5. Other possible mechanisms of transcriptional inheritance

Even though there is clear evidence for the existence of the inheritance of epigenetic information, many questions regarding the extent and mechanisms of this inheritance remain open. Aside from the attractive hypotheses involving chromatin modifications and non-coding RNAs, other non-genetic models – such as prions, metabolites, and transcriptional loops - should be considered.

For example, several yeast strains have proteins that can behave as prions and aid these strains adapt to various environmental stresses such as osmotic, oxidative, pH stress and DNA damaging agents [121]. This suggests a possibility that endogenous prions could be transmitted through meiosis and thus be transgenerationally inherited.

Another possible mechanism for transgenerational epigenetic inheritance involves metabolites or other small molecules functioning as the reservoirs for the information. These metabolite level changes can either affect the chromatin states that could then be inherited over the generations [122, 123] or the metabolites in the oocyte cytoplasm could directly be inherited to the zygote and then directly affect the chromatin states or cellular physiology.

Or, the inheritance mechanisms might involve transcriptional feedback loops that are independent of histone modifications or other epigenetic players [124]. In this model, an extracellular signal activates transcription factor that then upregulates its own gene. Once activated, this feedback loop could alter gene expression persisting even in the absence of the initial activator [125].

1.6. DNA replication and inheritance of gene expression states

The multiple levels of chromatin regulations (histone modifications, histone variants, DNA modifications, non-coding RNAs) are the basis for the chromatin classification into heterochromatin and euchromatin. During mitotic cell division, in addition to the DNA replication, the chromatin structure as well must be propagated for transmission of epigenetic information [126].

1.6.1 Histone modifications can be inherited

Modifications of histones, called the 'histone code' can be coupled with specific transcriptional states that can at times carry the long-term transcriptional memory. Some of the modifications are mediating only short-term signaling function and are not able to carry epigenetic memory [127]. The marks that are able to carry this memory must be stable, and maintainable during the cell cycle, and they also have to be copied appropriately onto the new chromatin after the DNA replication cycle. It is therefore important to understand how these marks are propagated and dealt with by the replication machinery to maintain the stability of the epigenetic memory of the transcriptional states. Evidence that some histone modifications can be inherited during cell division comes from epigenetic model systems, such as *Drosophila*. During the development of the fruit fly, some proteins can switch on and off the expression of some genes (such as conserved *Hox* genes) responsible for its segment patterns in different regions of the embryo, and this expression can be maintained throughout the lifespan, without the initial signal [128]. The *Drosophila* HP1 homologue Su(var)2-5 has a role in transcriptional memory as loss of this gene results in loss of silencing. HP1 is able bind the H3K9 methylation and through its dimerization contributing to the compaction of chromatin [129]. HP1 can

also interact with the H3K9 methyltransferase SUV39, providing a spreading mechanism of the chromatin mark and supports the “self-recruitment” mechanism.

Next to *Drosophila*, fission yeast *Schizosaccharomyces pombe* and budding yeast *Saccharomyces cerevisiae* are other systems that has provided support that the chromatin modifications play a role in heritable gene expression states, through the studies with the mating type loci. In both of these yeasts, the mating type is determined by the expression of one of the two alleles of the mating factor, maintaining the other copy silent and inheriting this silent state during cell division [130, 131]. The Swi6, an essential structural component in the mating type locus of *S.pombe* [132] is involved in imprinting at the mating-type locus and contributes to the maintenance of the silenced state. Whereas in *S. cerevisiae*, the NAD^+-dependent deacetylase Sir2 is required for the maintenance of the silencing of the mating type loci [133, 134].

1.6.1.1 Histone dynamics during DNA replication

When DNA is replicated during S phase, the number of nucleosomes needs to be doubled as well and the new histones need to be deposited onto the DNA. At the same time, nucleosomes act as a barrier for the DNA replication machinery, so they need to be disassembled ahead of the DNA replication fork. After DNA replication, the new

nucleosomes are assembled, using both new and parental histones [135]. Failure to correctly assemble the newly synthesized DNA into chromatin following the replication fork leads to genomic instability and cell cycle arrest [136]). The new histones need then to carry the locus-specific information form the parental histones [137].

Several histone chaperones are described that assist the deposition of the new histones, such as chromatin assembly factor 1 (CAF-1), a conserved protein complex of three subunits [138]. CAF-1 binds H3-H4 and in interaction with PCNA assembles replicating DNA into nucleosomes. Mutations in CAF-1 result in challenged transcriptional silencing from yeast to mammals [139-142]. For example, in *Drosophila*, reduction of the largest subunit of CAF-1, p180, suppresses heterochromatic gene silencing [142], accompanied by the decrease in H3K9 methylation marks at pericentric heterochromatin and reduced recruitment of HP1. As well in mice, loss of CAF-1 p150 alters the structure of constitutive heterochromatin, implying CAF-1 role in heterochromatin silencing through preventing the spreading of heterochromatin. In yeast, the spreading of Sir (Silent Information Regulator, involved in organizing heterochromatin) proteins is challenged in CAF-1 knock-downs [143]. At the same time, CAF-1 interacts with proteins that are involved in heterochromatin silencing and maintenance and may be directly needed during the DNA replication to recruit these proteins. For example, CAF-1

interacts with HP1 in mammalian cells and HP1 on euchromatin leads to recruitment of Drosophila CAF-1 p180 at this site as well [142]. In addition, CAF-1 is suggested to recruit SETDB1 methyltransferase complex onto new histones and in complex with other proteins promotes trimethylation of H3K9me [144]. In summary, the model suggests that CAF-1 couples *de novo* nucleosome assembly with the recruitment of HP1/Swi6 and histone methyltransferase complexes for proper inheritance of the heterochromatic states.

Another well-studied histone chaperone involved in nucleosome assembly is anti-silencing factor 1 (Asf1), first identified in budding yeast where over-expression leads to gene de-silencing [145]. Asf1 in complex with HIRA (histone regulatory homolog A) spread the heterochromatin through the interaction with Swi6/HP1 [146]. MNase-Chip (micrococcal nuclease digestion with microarray analysis) in *S. pombe* showed that deletion of ASF1 results in reduced nucleosome occupancy at the heterochromatin, supporting the idea that Asf1 is regulating the silencing through ensuring proper occupancy at the heterochromatin loci. Yamane, K. *et al.* looked more specifically at the role of Asf1 in heterochromatic silencing and protective functions of chromatin in *S. pombe* [146]. They show that Asf1 functions in a parallel pathway with SHREC (Snf2/HDAC repressor complex that mediates heterochromatic transcriptional gene silencing), impacting nucleosome

occupancy at heterochromatic loci. As SHREC is promoting transcriptional gene silencing through H3K9me3-bound HP1 proteins [147], they wondered whether Asf1 affects the nucleosome occupancy at heterochromatic loci. They used microarrays containing probes from the major heterochromatin domains probed with mononucleosomal (MNase digested) DNA to measure nucleosome occupancy at heterochromatic loci.

In summary, while the precise function of any of these proteins in heterochromatin silencing is not clear, proposed roles include that these proteins facilitate DNA replication and nucleosome assembly of the newly replicated regions into heterochromatin, thus only effecting inheritance indirectly. These proteins also recruit factors that are needed for the establishment and maintenance of the heterochromatin and the corresponding modifications. Finally, these proteins may interact with siRNA machinery in the maintenance of the heterochromatin and ensure inheritance.

The main remaining question to answer is how these cell-specific gene-expression patterns are maintained and transferred through DNA replication cycles.

1.6.2 Epigenetic silencing of tumor suppressor genes

Tumor suppressor genes (TSGs) that inhibit normal cellular growth are shown to be frequently silenced epigenetically in

cancer [148]. It is generally thought to involve DNA cytosine methylation, histone modifications and the compaction of chromatin. Lin, J.C., *et al.* showed that silencing of the three transcription start sites of the MLH1 promoter CpG island is regulated by differential nucleosomal occupancy [149]. They showed that three nucleosomes that in normal cells are missing from the start sites are present on the silenced methylated promoter. This suggests that locating the nucleosomes to previously vacant positions regulates this epigenetic silencing.

1.7 Transgenesis in *C. elegans*

1.7.1 *C. elegans* as a model organism

All the work described in this thesis was performed using the model organism *C. elegans* [150] (Figure 8), a nematode worm that is about 1mm long in adulthood.

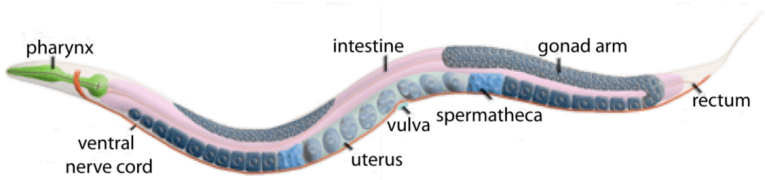


Figure 8 Drawing of *C. elegans* adult hermaphrodite anatomy (modified from figure by Altun and Hall, Wormatlas.org), left lateral side.

C. elegans has many characteristics that make it an attractive model organism, its strains are cheap and easy to

keep, the only requirements are humid environment, oxygen, cholesterol and bacteria as food. The strains are readily available to order from the *Caenorhabditis* Cenetics Center (CGC) and can be maintained on agar Petri dishes seeded with bacteria, most commonly with *E. coli* [150]. Another favorable characteristic beneficial for this particular study is the fact that the life cycle of *C. elegans* is about 3 days at 25C to 6 days at 15C [151], enabling us to study the phenotypes across many generations is short time. *C. elegans* has two sexes, a self-fertilizing hermaphrodite (XX) and male (X0). The nematode *C. elegans* is readily suitable for following gene expression, due to its well annotated and sequenced genome [152] and its transparency, allowing imagining during any developmental stage of the worm. The most commonly used strain is the wild type strain called N2, isolated from Bristol, UK [153].

1.7.2 Objectives for using transgenic worms

Transgenic DNA can be introduced into the worms via microinjection or microparticle bombardment and can be either integrated into the genomic DNA or inherited as an extrachromosomal array. Historically, there has been two main reporter constructs in *C. elegans*, transcriptional and translational. The more simple, transcriptional reporter consists only of a promoter fragment from the gene of interest, driving the reporter (e.g. LacZ or GFP). This

construct usually includes a portion of the *cis*-regulatory sequence of an endogenous gene. In the case of the more complex translational reporter, the whole genomic locus can be included.

1.7.2.1 Single copy and multi copy transgenes

DNA can be introduced into the *C. elegans* germline by microinjection [154], whereas the injected supercoiled DNA molecules form an extrachromosomal array composed of tandem repeats of the sequence. These repetitive arrays are usually unstable through cell division and can become heritable even without integrating into the chromosome. Expression of these transgenes in the lines created by germline injection is often mosaic due to uneven loss of the repetitive extrachromosomal arrays through mitotic instability [154]. Expression of these microinjected transgene sequences may not mimic the endogenous expression of the gene, as these tandemly repeated sequences can trigger gene silencing mechanisms (RNAi-like effects), especially in the *C. elegans* germline, where the arrays are rapidly silenced after a few generations [155], though remaining physically present in both soma and germline. “Loss” of the signal can be either due to the loss of the multicopy extrachromosomal array or the organism’s response to transgenic repetitive integrated arrays. Next I will cover different transgene silencing patterns, both for

single copy and multi copy transgenes and in germline tissue as well as in somatic tissue.

1.7.2.2 Germline transgene silencing

Non-integrated transgenes are transmitted as heritable linear extrachromosomal arrays and studies suggest that soma/germline “view” these transgenes differently and that gene expression may be uniquely repressed in the postembryonic germline. Many reporter gene constructs, e.g. LacZ, and GFP-tagged have been shown to be efficiently expressed in most somatic tissues, whereas these arrays are usually silenced in the germline [156, 157]. In addition, poor transgenic rescue of maternal effect mutations is common [158]. Kelly *et al*. first demonstrated that the *C. elegans* germline is able to silence the expression of genes in simple repetitive arrays [158]. This might be a way for the germline to control its gene expression - by preserving generalized silencing of DNA, it can prevent expression of somatic genes and suppress differentiation into specific somatic pathways to ensure the maintenance of the undifferentiated germline. Potentially, these prevalent silencing mechanisms in germline could help preventing proliferation of viruses and transposons. Heterochromatin, a highly repressive structure to prevent expression and activation of transposable elements and other possibly deleterious sequences is a classical cell

mechanism to control RNA polymerase. Stable inheritance of the silenced heterochromatic state requires specific protein complexes, such as Polycomb systems that will be discussed later in more detail.

This suggests that global silencing mechanisms play the key role in the maintenance and specification of germ line tissue. The degree of silencing can be measured with transgene arrays containing fluorescent reporter genes under the control of ubiquitous promoters. Transgenes are silenced in the germ cells and reactivated in the soma of each generation and the germline silencing is dependent on the MES proteins (maternal effect sterile) [159]. In particular, involvement of transcriptional repression in transgene silencing was shown with immunostaining experiments, where the silenced arrays are detectable with heterochromatin-specific antibodies, (e.g. histone H3 methylation on Lys9), while not detectable with antibodies specific for euchromatin (e.g. H3 methylation on Lys 4) [50]. Genes involved in this process are players in transcriptional regulation and include *mes-2*, *mes-6*, *his-24* (*C. elegans* linker histone variant H.1) [160, 161], *hpl-2* (one of the *C. elegans* heterochromatin 1 homologues) [162]. *mes-2* and *mes-6* encode worm homologs of the Drosophila Polycomb Group proteins, Enhancer of Zeste and Extra Sex Combs, respectively. Polycomb proteins are responsible for the maintenance of transcriptional repression of developmentally regulated genes through chromatin

conformation and both homologues of the Polycomb family of transcriptional repressors and histone H1.1 have been shown to be required for silencing repetitive transgene arrays in the germline [163]. Kelly and Fire [159] looked at the effects of maternal genotype on silencing of the ubiquitously expressed *let-858* tandem transgene array and showed that *mes-2*, *mes-3*, *mes-4* and *mes-6* mutant background results in desilencing of the transgene array. Additionally, the reporter construct was efficiently expressed in the soma but silenced in the germline lineage. Both *mes-2* and *mes-6* also contain a SET domain - known feature of chromatin-interacting proteins, supporting the view that gene silencing occurs through the regulation of chromatin conformation. The HP1 proteins have been implicated in somatic position-effect silencing in *Drosophila* and mammals [164], Couteau, F. *et al.* [162] tested the role of HPL-2 in the germline silencing. They showed that *hpl-2*-RNAi-fed worms failed to silence the afore-mentioned reporter *let-858* construct, concluding that as the MES proteins, HPL-2 is required for germline silencing. Silencing of the repetitive transgenic arrays in the germline is a stable phenomenon, the characteristics of which will be discussed in detail later.

1.7.2.3 Complex arrays

Germline expression of several transgenes has been reported when using more complex arrays where the linearized construct is mixed with *C. elegans* digested genomic DNA [158], this however could not be applied to all tested transgenes. The complex array approach attempts to minimize the effect of repetitive arrays on gene expression in the germline. In the complex arrays the construct could become dispersed and be imbedded into high sequence complexity, resembling euchromatin. Kelly, W.G *et al.* tested several transgenic lines from co-injecting *gfp*-tagged *let-858* with cleaved genomic DNA and obtained several lines showing robust expression in both soma and germline. However, not all tested transgenes could be desilenced using complex arrays and transgene germline silencing has been a re-occurring problem in studies involving expressing tagged proteins in germ cells and early embryos. Another group [165] reported successful germline de-silencing by combining the complex arrays with vector containing regulatory sequences from genes normally expressed in the germline such as *pie-1.* They reported successful robust germline expression of fusion genes encoding GFP-tagged versions of γ-tubulin, β-tubulin, and histone H2B.

1.7.2.4 Co-suppression

Interestingly, it was observed that when the repetitive transgenic DNA was similar to the endogenous germline genes, these endogenous loci were silenced as well, a phenomenon called co-suppression [166], [167], [168], [64]. Dernburg *et al.* [168] asked how do high-copy transgene arrays repress the endogenous chromosomal copies of the gene. Whereas in *Drosophila*, physical association between repetitive heterochromatic DNA and euchromatic gene can silence the gene expression [169], FISH experiments with *spo-11* loci failed to detect similar paring with the transgene and chromosomal loci in *C. elegans*. Furthermore, they suggest that cosuppression may involve a diffusible mediator instead, that transcription from the array is required for repression of the endogenous genes and that RNA molecules are involved in this process, relating the cosuppression mechanistically to RNAi. They show that although *rde-1*, the primary *C. elegans* Argonaute required for RNAi is not required for cosuppression, *rde-2* (novel *C. elegans* protein, functioning downstream of *rde-1* in the RNAi pathway) and *mut-7* (homolog of RnaseD repressing transposition of the main know transposons in *C. elegans* Tc1, Tc3, Tc4, Tc5,) are essential. Identifying key players in co-suppression facilitated linking the phenomenon to previously described post-transcriptional gene-silencing (PTGS) processes, transposon silencing and RNA interference [168], [64], [170]. These studies suggest that

yet unknown RNA mediators are required to establish and/or maintain co-suppression.

1.7.3 Integrated single copy transgenes

As we have learned, germ cells are much more efficient in silencing multicopy transgene arrays than somatic cells and that silencing also expands to endogenous loci of the same sequence. This silencing primarily affects expression in the germ cells, but also of early embryos that mostly depend on maternal mRNAs and proteins synthesized during oogenesis. Using strains with low-copy transgenes (or transgenes in complex arrays) helps overcome the potent germline silencing in some occasions. The transgene can be integrated into the genome, either at a random site by microparticle bombardment [171] that produces single- and low copy chromosomal insertions, providing stable transmission of the transgenic DNA over many generations. Using this microparticle bombardment method, Praitis *et al.* generated *C. elegans* lines that express GFP transgenes in reproducibly consistent patterns in somatic tissues as compared to extrachromosomal array lines that exhibit varying expression patterns across animals due to mosaic loss of the array [154]. They propose that in these low-copy transgenic lines, the number of transgene copies is insufficient to activate context-dependent gene silencing in soma. The microparticle bombardment technology has also

facilitated expressing the transgenes in the *C. elegans* germline. Out of five lines tested in the Praitis *et al.* pioneering microparticle bombardment study, they showed continuous expression of *pie-1::GFP* construct in the germline for >20 generations. Though this is obvious improvement, 2 out of 5 lines were still reported to silence the transgene expression in the germline, emphasizing the sensitivity of the germline to these exogenous transgenes. The ability to generate stable lines with consistent germline expression was an obvious improvement, but far from perfect.

Yet another improvement was introduced with the Mos1-mediated Single Copy Insertion (MosSCI) method [172] that inserts a single copy of a transgene into a defined locus (Figure 9, left panel). Mos1, is a *Drosophila* class II transposon and its mobilization generates a double strand break that is repaired through copying DNA from an extrachromosomal template into the chromosomal site. The insertions are likely to proceed via synthesis-dependent strand annealing (SDSA) [173]. The mechanism of MosSCI uses 'cut-and-paste' system and can be experimentally mobilized in the *C. elegans* germline [174]. In addition to facilitating transgene expression in the germline, the MosSCI technique also eliminates variability between different strains as the copy number and DNA context can be designed to be identical.

One of the most recent tools in genetics adapted for use in *C. elegans* is the ability to edit the worm genome using clustered, regularly interspersed short palindromic repeats (CRISPR) RNA-guided Cas9 nuclease and homologous recombination [175, 176], (Figure 9, right panel). In prokaryotes, the Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) and CRISPR-associated (Cas) systems are an adaptive immune response against viruses and plasmids, where cells use RNA to guide cleavage of foreign DNA sequences. In genome editing, the synthetic single guide RNA (sgRNA) is used to target the DNA sequence, after which Cas9 introduces the double-stranded breaks that are efficiently repaired by homologous recombination. Goldstein *et al.* [177] were able to generate GFP knock-ins and targeted mutations that were previously impossible for C. *elegans* researchers. Generation of the sgRNA is relatively simple and transgenes can be assembled using Gibson cloning technique [178].

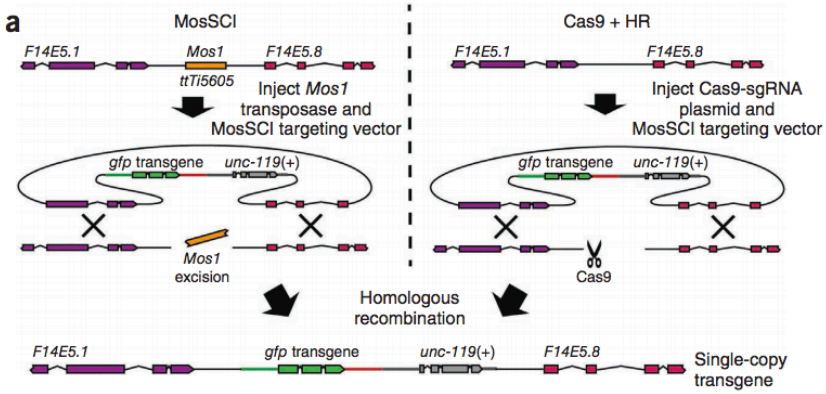


Figure 9 Newer and more precise genome modifying methods in *C. elegans* rely on homologous repair of double-strand breaks methods. (Left) Homologous recombination mediated by Mos1 transposon excision. The chromosome is broken at a chosen location by excising Mos1 transposon. In the presence of a DNA template with homology to the breakpoint the DNA from repair template is incorporated into the genome. Usually a positive selection marker (here *unc-119*) is used to select for successful events. (Right) Genome editing in *C. elegans* mediated by Cas9, a method using the clustered regularly interspersed short palindromic repeats (CRISPR) RNA-guided Cas9 nuclease and homologous recombination. The method relies on double-strand break repair using an engineered homologous template similar to *Mos1* transposon. Cas9 is able to induce DNA double-strand breaks with specificity for targeted sites and these breaks are repaired with high efficiency by homologous recombination [177].

1.7.4 Somatic expression and silencing

Heterochromatin often includes highly repetitive DNA. In many cases, the sequence that is active as a single copy locus in the genome can become inactivated when repeated, stronger effects with repeats at a single genomic site and in tandem repeat. Two models that could direct specific sequence for silencing are discussed next. In the first case, a weak *cis*-silencing element is present in each segment of the tandem repeat and the repetitive nature

would be the key to induce silencing by positioning several copies of the weak cis-silencing element in repetition. In the other model, the organism is simply able to recognize the repetitiveness and not any specific sequence.

In *C. elegans* the transgenes that are often maintained as long tandem extrachromosomal structures, and though expressed, there is a clear difference in transcriptional activity compared to the endogenous loci. The activity is often mosaic and expression level per copy commonly lower than for endogenous genes [179]. This difference cannot just be attributed to the extrachromosomal nature of these arrays, as integrated arrays producing stable transgenic lines can show variable expression as well [180]. This is consistent with the complex array approach discussed previously, where the expression of several germline and somatic constructs is relieved when the transgenic DNA is cotransformed with genomic DNA carrier, producing more complex array.

Hsieh, J. *et al.* [181] looked at context-dependent gene silencing, often used by organisms to stably modulate gene activity within large chromosomal regions. By using tandem array transgenes to screen for mutants in *C. elegans* that affect (trans)gene silencing in somatic tissues in context-dependent manner, they identified loss-of-function mutations in *tam-1* gene (tandem-array-modifier) that repressed the activity of several highly repetitive transgenes, whereas non-repetitive transgenes retained

their activity in the *tam-1* mutant background. They classified the previously uncharacterized *tam-1* as one of the synMuv class B genes. Synthetic multivulva class B proteins are a conserved group of transcriptional repressors that belong to the larger ABC SynMuv group of genes [182]. Some mutations in genes affecting vulval cell induction cause Multivulva phenotype (Muv) and in some cases two mutations at two discrete locations in the genome are required for the Muv phenotype to occur – synthetic multivulva (SynMuv) (Figure 10). Mutations in either class of SynMuv genes alone does not cause a Muv phenotype but a combination of any of the two groups does [183].

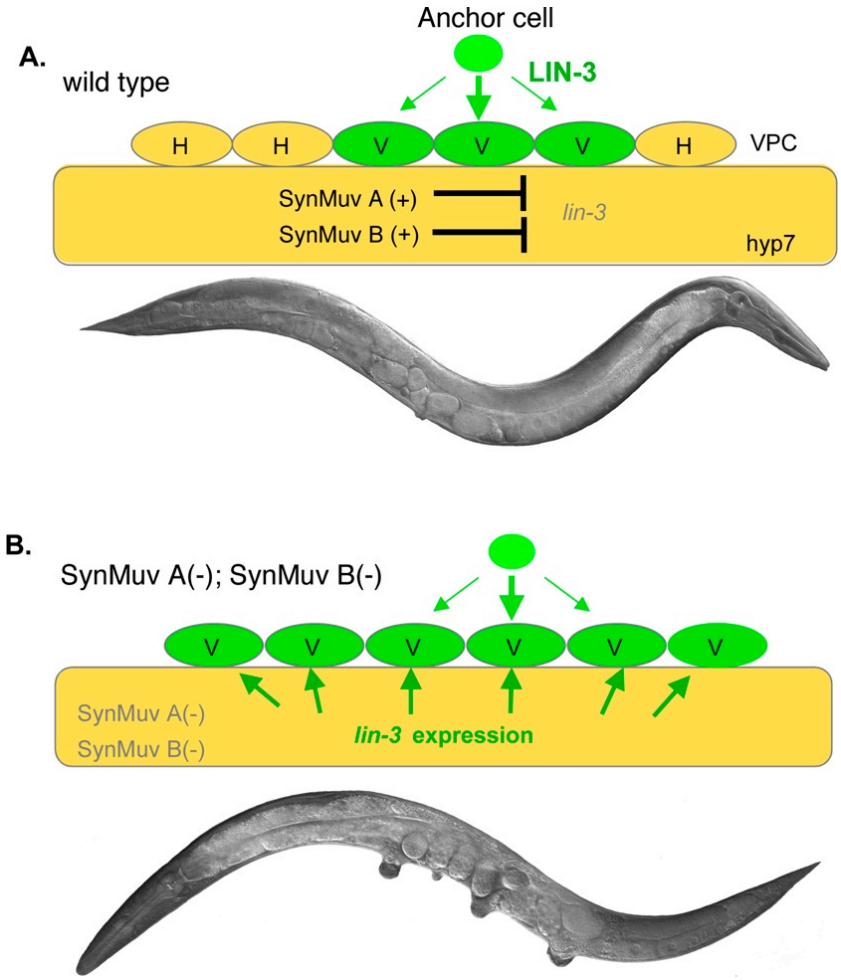


Figure 10 Model for the molecular basis of how SynMuv A and SynMuv B genes regulate vulval induction (modified from Cui *et al.*, 2006) within the hyp7 syncytium [184]. (A) In wild type the LIN-3 inhibited redundantly by SynMuv A and B genes, so three of the vulval precursor cells (VPCs) not receiving the signal divide to generate part of the hypodermis (H, yellow) while others (V, green) acquire the proper vulval cell fate. (B) In SynMuv A (-); SynMuv B (-) double mutants the de-repression of *lin-3* and *hyp-7* leads to the activation of RTK/RAS/MPK (Receptor Tyrosine Kinase Ras GTPase, MAP kinase) pathway (not shown) in all six VPCs (V, green), resulting in a Muv (multivulva) phenotype.

In *tam-1* mutants the silencing of tandem transgenes was enhanced and this silencing appeared to be specific to transgenes as expression of the corresponding endogenous

loci of *myo-3* locus were not altered. Also, they show that transgenes in complex array context were significantly less susceptible to silencing in *tam-1* null mutants [181].

The SynMuv class B genes have been shown to be important in the modulation of the EGFR-RAS pathway during vulval specification (Figure 10). But the class-B synMuv mutants have also been shown to exhibit defects in growth and fertility, suggesting additional roles for these genes in addition to vulval development [183]. In the Fire lab they also observed effects of other SynMuv class B genes (*lin-9*, *lin-15B*, *lin-35*, *lin-51* and *lin-52*) on transgene expression levels in mesoderm, it is likely that this gene family acts in most tissues of the animal, supported by the broad expression patterns of several SynMuv class B genes [182]. They suggest that the effects of the SynMuvB genes include changes in the acetylation level of the histones in the chromatin, supported by the observation that it is possible to induce SynMuvB phenotype by disrupting the function of histone deacetylase-1 (*hda-1*) [182], and histone deacetylase complexes are suggested to be negative regulators of gene expression [185]. Another possibility is that the synMuv genes act to direct silencing factors to sets of targets including tandem array transgenes.

Lehner *et al.* [186] analyzed the mutations in *lin-35* (the worm ortholog of the tumor suppressor gene *p105Rb*) and other related synMuv B family of chromatin-modifying genes, and reported enhanced somatic transgene silencing

via an RNAi-dependent pathway, in addition to higher penetrance of many RNAi phenotypes. They show that a subset of synMuv B genes negatively regulates RNAi and inactivation of those genes also results in somatic transgene silencing. As the observed somatic transgene silencing in these synMuv genes can be suppressed by inactivation of components of the RNAi machinery, they conclude that inactivation of these synMuv genes is inducing somatic transgene silencing via increasing RNAi. This is inconsistent with the data obtained with previously mentioned *tam-1* mutants, as in the latter case the somatic transgene silencing is enhanced without any observable changes in RNAi sensitivity.

Several of the multivulva-suppressing clones are annotated as chromatin factors. For example, MES-4 (homologous to the human MLL protein and *Drosophila* homologues of MES-4 interact with SWI/SNF and NuRD complexes and antagonize Polycomb complexes [187]). MES-4 coats the autosomes in the *C. elegans* germline and is retained from the X chromosome in the germline [188]. Wang *et al.* [189] discussed that in the absence of antagonistic pathways of MES-4 chromatin remodeling complex is active in the soma and activates several inappropriate genes such as P granule genes, including *pie-1*. Also, *mes-4* RNAi is able to suppress transgene silencing, vulval sell lineage defect and somatic expression of PGL-1 in the *lin-35* mutant that is one

of the *C. elegans* homologous of the retinoblastoma (Rb) tumor suppressor complex.

Wu *et al*. [190] looked more in depth into the somatic misexpression of germline-specific genes in synMuv B mutant animals and characterized three distinct chromatin complexes that prevent misexpression, including LIN-35/Rb-containing core complex (DRM), the SUMO-mediated Mec complex and synMuv B heterochromatin complex. The proteins of those three classes function to repress overlapping sets of P granule and RNAi genes, whereas misexpression can lead to different results. This contradicts the predicted possible positive correlation between RNAi efficiency and the ability to silence transgenes where enhanced RNAi leads to enhanced transgene silencing. The heterochromatin class of synMuv B genes seems to be required for transgene silencing, likely downstream or parallel to their effects on the efficiency of RNAi.

Whereas most of the gene silencing in the soma occurs at the post-transcriptional level, Grishok *et al.* [67] reported an example where silencing of transgene expression in the soma occurred at the level on transcription in *C. elegans*. This transgenic silencing is mediated through RNAi-dependent pathway [63], [62] and depends on several RNAi pathway genes such as *dcr-1*, *rde-1*, *rde-4*, *rrf-1*. In their study Grishok *et al.* showed that *elt-2::gfp/LacZ* transgene was silenced when the worms were fed with RNA produced

from the commonly used L4440 vector that shares the backbone sequence with the transgene. Different from previously described post-transcriptional gene silencing, this process is dependent on PAZ-PIWI protein ALG-1 and on the HP1 homolog HPL-2, a chromatin-silencing factor. The inhibition of transgene is occurring at the precursor mRNA level and is accompanied by a decrease in the acetylation of histones associated with the transgene. This silencing is distinguishable from transgene silencing in the germline as it cannot be stably transmitted to the next generation and it is dependent on the *rde-1* gene, the primary Argonaute protein in RNAi pathway in *C. elegans* [62]. They identified additional chromatin-modifying components affecting this RNAi-induced Transcriptional Gene Silencing (RNAi-TGS).

Involvement of RNAi pathways in gene silencing is most extensively studied in fission yeast where the small interfering RNAs corresponding to centromere repeats have been shown [191] and genes from the RNAi pathway (e.g. *dicer*, *argonaute*, RNA-dependent Polymerase) have been implicated in silencing at the centromeres [120]. The hallmark of the silenced chromatin is Histone 3 Lys 9 (H3K9) methylation and association of H3K9 with Swi6, a homologue of Heterochromatin Protein 1 (HP1). A connection between RNAi, histone methylation and DNA methylation has been shown in the context of transgene and transposon silencing in plants, and members of gene

RNAi-associated gene, such as *argonuate4* (*ago4*), *dicer-like3* (*dcl3*), *RNA-dependent RNA polymerase2* (*rdr2*) have been shown to be important for initiation of the silencing on the transcriptional level [192, 193].

As well in *Drosophila*, the role of the RNAi pathway for silencing of transgenes at the chromatin level has been shown. More specifically, mutations in the PAZ-PIWI proteins (*piwi* and *aubergine*) and *spindle-E* encoding DEAD-motif RNA helicase result in the reduction of H3K9 methylation and delocalization of HP1 and HP2 [194].

Aim of the thesis

The objective of this thesis was to use a genome-wide RNAi screen as the starting point to better understand how repressed chromatin is inherited in *C. elegans*.

2 - RESULTS

The project started with an interest in inter-individual phenotypic variation in isogenic organisms that live in the same shared environment, as in the case with *C. elegans* grown in the laboratory in the same conditions [195]. Especially interesting to explore is the case with mutations that show incomplete penetrance within an isogenic population. These processes are often described as an outcome of stochastic events in the expression of buffering systems that are able to influence the expression of these inherited mutations [196], [197]. Many mutations have outcomes that are dependent on the activity of molecular chaperones, a group of proteins whose main function is helping other proteins to fold that are involved in specific molecular mechanisms able to promote environmental robustness. Chaperone systems are suggested to be mediating the link between environment and genetic robustness [196], [197].

Chaperones can be either inducible, responding to the stress in the cell and functioning to help for example with the misfolded or aggregated proteins; or they can be constitutively expressed, and needed throughout the development, as for example Hsp90. The capacity of Hsp90 to buffer mutations was first demonstrated with *Drosophila melanogaster*: homozygous mutants are not viable but

heterozygotes are viable with a myriad of different phenotypic abnormalities [196].

Initially we decided to focus on chaperone systems and designed a screen to find regulators of Hsp90, a constitutively expressed chaperone whose function is important throughout development. It was shown previously that variation in the stress response relates to pre-existing variation in chaperone levels in *C. elegans* [198]. It was shown that pre-existing molecular variation in a population might lead to inter-individual variation in the heat stress response.

An RNAi screen was designed to find regulators of Hsp90, taking advantage of a fluorescent reporter for the *daf-21* gene that is the gene for Hsp90 in *C. elegans*. The level of Hsp90 could be assessed with the *pdaf-21::mCherry* multicopy reporter through the intensity of the fluorescence expressed in the worm. The expression of the reporter is highly variable even in isogenic individual worms reflecting the variation in the levels of Hsp90 in normal conditions. The screen was designed to detect both positive and negative regulators of Hsp90, identifiable by changes in the *daf-21::mCherry* expression readout.

2.1. An RNAi screen to identify regulators for Hsp90, using a high copy transcriptional *pdaf-21::mCherry* reporter

The RNAi screen was designed to identify regulators of zygotic Hsp90 expression using the *pdaf-21::mCherry* transgene where *pdaf-21* is the reporter for the *C. elegans* gene for HSP90. The transgene of the multi-copy *pdaf-21::mCherry* transgenic line used (Figure 11) is integrated somewhere into the 4th chromosome and qPCR of the genomic DNA from the different integrants of the same transgene estimated the copy number of the multi-copy transgene to be around 115-200 copies (Figure 12), the huge variation possibly indicating rearrangements that have occurred within the transgene The screen was designed to take advantage of the RNAi library with transformed *E. coli* for about 15,000 *C. elegans* genes that can be used to silence genes in worms [199]. The screen was performed in high throughput, feeding worms in liquid and in 96-well plates. This setup allowed the expression of the reporter gene to be monitored in the progeny of the RNAi-containing bacteria fed animals up to 96 treatments per plate after 4-5 days of feeding.

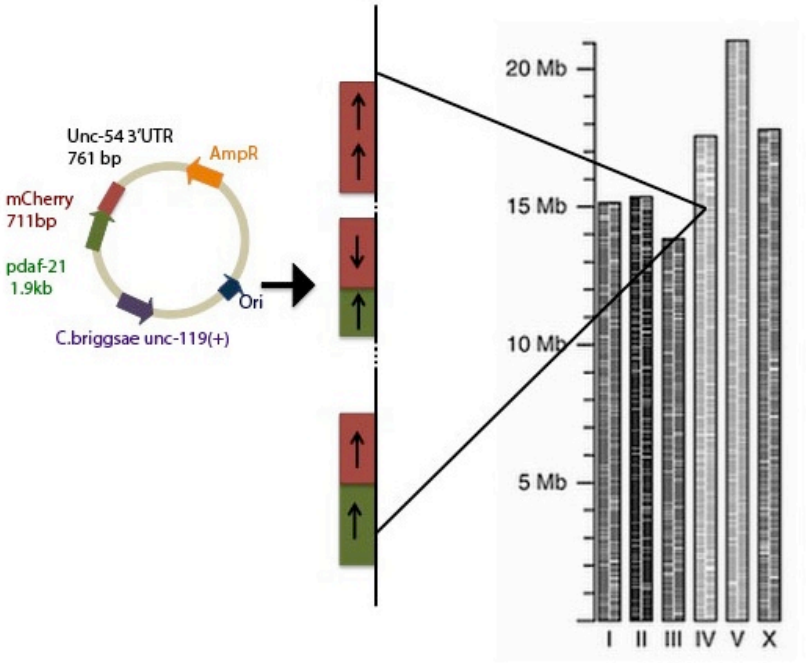


Figure 11 The *pdaf-21::mCherry* transcriptional reporter construct used in this study (left panel), generated previously in the lab by A. Burga [200]. The 1.9kb *daf-21* promoter region was amplified from N2 genomic DNA. The 3' entry vector pCM5.37 contains the *unc-54* 3'UTR, AmpR, *C. briggsae unc-119* rescuing fragment. The *C. elegans* strain was generated by bombardment in an *unc-119(ed3)* background [171, 201]. (Middle panel) Upon insertion, the transgenes are often reassembled into tandem arrays containing fragments of the transgene construct in inverted and dispersed organization (shown with arrows), *daf-21* promoter depicted green and *mCherry* in red, as in the vector. Whereas the precise location of the transgene insertion in the strain used is unknown, it is estimated to be somewhere in the 4th chromosome (I-V *C. elegans* autosomes, X – *C. elegans* sex chromosome).

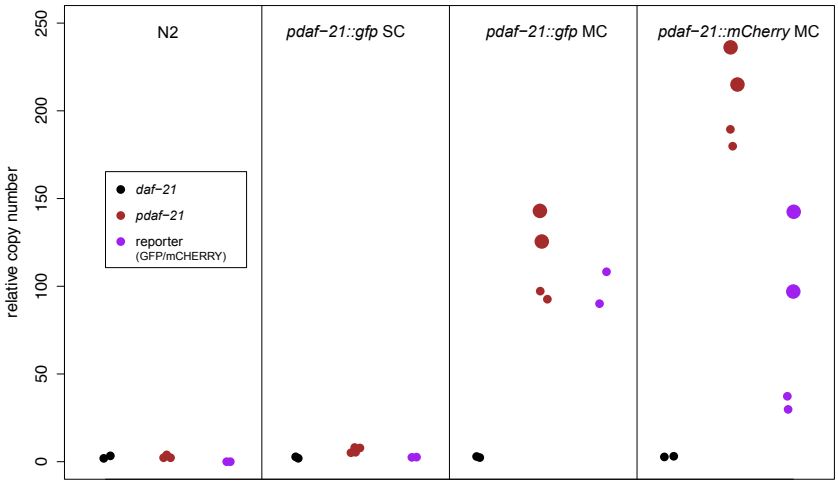


Figure 12 Transgene copy number in the *pdaf-21* reporter strains used in this study, measured in the *pdaf-21::gfp* SC (single copy) and MC (multi copy), and *pdaf-21::mCherry* MC strains. DNA from these three strains with the *pdaf-21*-reporter construct was analyzed by qPCR analysis to estimate the copy number of these transgenes. The results were normalized to a reference gene *cdc-42* and each point represents one biological replicate. The two different sizes of the dots represent two different primer pairs for that gene in the multi copy strains. For the N2 (wild type) strain, the primer pairs for the mCherry and GFP did not yield a detectable amplicon. For N2 the primers for *daf-21* exonic region and *daf-21* reporter region estimated similar value (average 2.84 with σ=0.8). That is comparable to *pdaf-21::gfp* SC and MC, as well as *pdaf21::mCherry* MC strain (average 2.64, σ=0.39) with the primer pair for the exonic *daf-21*. In the *pdaf-21::gfp* single copy strain the *pdaf-21* promoter primer pair estimated *daf-21* about 2x higher as the primer pair for the *daf-21* exonic region (6.59 (σ=1.22) to 2.64 (σ=0.38)), and the GFP primer pair 3.1 (σ=0.5). The transgene copy number estimated for the *pdaf-21::gfp* MC strain is 105.51 (σ=13.74), averaged over the promoter and reporter primer pairs of the transgene; and for the *pdaf-21::mCherry* MC the estimated copy number is 115.84 (averaging over the promoter and mCherry, σ=96), promoter region estimating 198.14 (σ=18.3) and mCHERRY primer pair estimating 33.55 (σ=3.74).

The primary screen was done by eye, possibly leading to a high number of false positives and false negatives. Also, due to the highly variable level of the expression of the reporter, it was not possible to identify positive regulators of the *pdaf-21::mCherry* reporter. Nevertheless, the initial

screen enabled us to identify 339 primary hits and these primary hits were subjected to a secondary more stringent screen, where a clone was called a confirmed hit when it was scored in 2 out of 3 replicates as up-regulating the expression of the transgene, resulting in 56 confirmed hits (Table 2).

Gene	Function/complex	*pdaf-21::mCherry* 'phenotype'
rpl-22	Large ribosomal subunit L22	High L4/adult Ste
rps-5	Small ribosomal subunit S5	High L3/L4
rps-11	Small ribosomal subunit S11	High L4
rpl-12	Large ribosomal subunit L12	High sick L3/L4
rpl-18	Large ribosomal subunit L18	High L4
rpl-39	Large ribosomal subunit L39	High L4
rpl-20	Large ribosomal subunit L18a	High L3
rps-8	Small ribosomal subunit S8	High L3/L4
rpl-13	Large ribosomal subunit L13	High Ste L3/L4
rpl-25.2	Large ribosomal subunit L23a	High Ste L3/L4
rpl-3	Large ribosomal subunit L3	High L3
rpl-23	Large ribosomal subunit L23	High L3
rpl-9	Large ribosomal subunit L9	High L3
rps-24	Small ribosomal subunit S24	High L3/L4
rpl-25.1	Large ribosomal subunit L23a	High P0, 50% Ste
R186.8	Ortholog of MRPL33 (mitochondrial ribosomal protein)	High Ste P0
C15H11.9	rrbs-1, ortholog of RRS1 (ribosome biogenesis regulator)	High Ste P0
rnr-1	Large subunit of ribonucleotide reductase	High P0, F1, 75% Ste
pri-2	Homolog of the DNA polymerase α-primase subunit C	***div-1*-like**
rnr-2	Small subunit of ribonucleotide reductase	High F1, 75% Emb
lrr-1	Leucine rich repeat-containing, with BC/Cul-2 box.	***div-1*-like**
div-1	**Homolog of the B subunit of the DNA polymerase α-primase complex**	**High F1; 30% Emb**
Y47D3A.29	DNA pol. α catalytic subunit	***div-1*-like**
ama-1	RNA Pol II subunit	75% Ste High P0, High F1
mel-47	Ortholog of SLIRP (SRA stem-loop interacting RNA binding protein)	***div-1*-like**
mes-6	Member of Polycomb-like chromatin repressive complex (MES-2/MES-3/MES-6)	High F1
polh-1	DNA polymerase eta	High F1
pole-2	Ortholog of POLE2 (polymerase-Σ 2 accessory subunit)	High P0, F1
rpa-2	Homolog of replication protein A (RPA) subunit	High F1
polh-1	DNA polymerase η	High F1
F33H2.5 (*pole-1*)	Ortholog of POLE (polymerase Σ) catalytic subunit	High F1

Histones (*his-1,-5,-10,-14,-18,-26,-28,-31,-37,-38,-46,-50,-56,-60,-64,-67*)	Nucleosome components	High L1, P0
rfc-1/-3	Ortholog of RFC1 (replication factor C (activator 1)	High P0 (L4)
cdc-42	RHO GTPase	High L1
C08B11.1 *zyg-11*	Leucine rich repeat-containing, CUL-2 associated	High L1
hel-1	Ortholog of UAP56	High L4/adults (75%Ste)
T13H5.4	Ortholog of SF3A3 (splicing factor subunit)	High L4/adult, Ste
M28.5	*phi-9*; ortholog of NHP2L1 (NHP2 non-histone chromosome protein 2-like 1)	High P0, Ste
nhr-61	DNA binding transcription factor	High P0, 50% Ste
F40G9.1	*sec-20*	High P0
unc-45	Muscle-specific chaperone for type II myosins, HSP90 co-chaperone	High P0, Ste
C35D10.7	-	High F1
rpb-2	Ortholog of POLR2B	High L3
mdt-15	Ortholog of MED15	High P0 Ste
npp-10	nucleoporin ortholog, nuclear core complex	High P0 (L3/L4)
snr-5	Ortholog of ribonucleoprotein SmF	High P0 (L3)
F58A4.5	*clec-1*, ortholog of MRC1	High P0, F1
W09G12.5	->F38A1.8	High P0 (L4)
F38A1.8	Ortholog of SRPR (signal recognition particle receptor)	High P0 L3/L4
pbs-1	Protease subunit, affecting fertility, embryonic/larval viability	Sick bright spotty P0
srz-23	Serpentine receptor clas Z	High Ste P0
E04A4.5	Ortholog of TIMM17B (intracellular transmembrane protein transport)	High Ste P0
icl-1	Isocitrate lyase, downstream of DAF-16 influencing lifespan	High spotty P0, F1
vha-17	ATPase subunit e	High P0 (L3, L4)
W08D2.7/*mtr-4*	RNA helicase homologue	High Ste P0
let-70	Class I E2 ubiquitination enzyme	High P0 L4/adult
spt-5	Ortholog of SUPT5H, SPT transcription factor	High P0 Ste/Emb
smo-1	Ortholog of SUMO	High P0, F1 not
C18E3.1	-	High P0, F1
H19N07.1	*erfa-3*, ortholog of human GSPT2&GSPT1	High Ste P0
syn-3	SYNtaxin family	High P0
snr-4	Ortholog of SNRPD2 (small nuclear ribonucleoprotein)	High P0
his-74	Ortholog of H3F3A	High P0, F1
B0250.7	-	High Ste P0

Table 2 Hits from the secondary stringent screen with the 56 confirmed hits. The 339 hits from the primary RNAi screen were subjected to a more stringent screen where all these *E. coli* dsRNA clones were screened in triplicates and a clone was called a 'true hit' only when 2/3 wells showed increased up-regulation as compared to control RNAi-fed worms. The screen was carried out as for the primary screen, where in every well of the 96-well plate about 75 L1 larval worms were fed with RNAi bacteria in liquid for 88-90 hours and then screened by eye observing the *mCherry* expression (at 10x

magnification). The columns in the table show the gene, its function/complex and the arbitrary *pdaf-21::mCherry* phenotype, where 'High' means the expression of the reporter compared to the control levels. P0 - parental generation, F1 - first offspring generation. Ste – sterility phenotype, Emb – embryonic lethal, *div-1*-like - upregulation of the reporter as in *div-1* RNAi fed worms. The colors in the table indicate: yellow – ribosomal genes, cyan - genes involved in DNA replication, orange – other genes.

After classifying the hits into groups by the level of the expression of the *pdaf-21::mCherry* reporter and their function, a subgroup of genes involved in DNA replication was apparent observable by eye during the screen and most of them are components of the DNA replication machinery (Table 2, Figure 13 and 14).

Other groups of genes identified as regulators of the transcriptional *pdaf-21::mCherry* reporter included proteasomal subunits and ribosomal proteins, most of which were inducing the reporter expression already in the P0 parental generation. The ribosomal genes included large ribosomal subunits (such as *rpl-22*, *rpl-12*, *rpl-23*, *rpl-9*) or small ribosomal subunits (*rps-5*, *rps-8*, *rps-24*) Many of the RNAi clones of the ribosomal subunits were causing highly penetrant Ste (sterile), Emb (embryonic lethal) and phenotypes or arresting the larval development at a very early stage, suggesting that the induction of the reporter was promoter-specific (*daf-21* encodes for Hsp90 in *C. elegans*), a response to disrupted proteostasis. In addition, some transcription factors were identified, often producing sick P0 animals, such as *nhr-61* and *spt-5*).

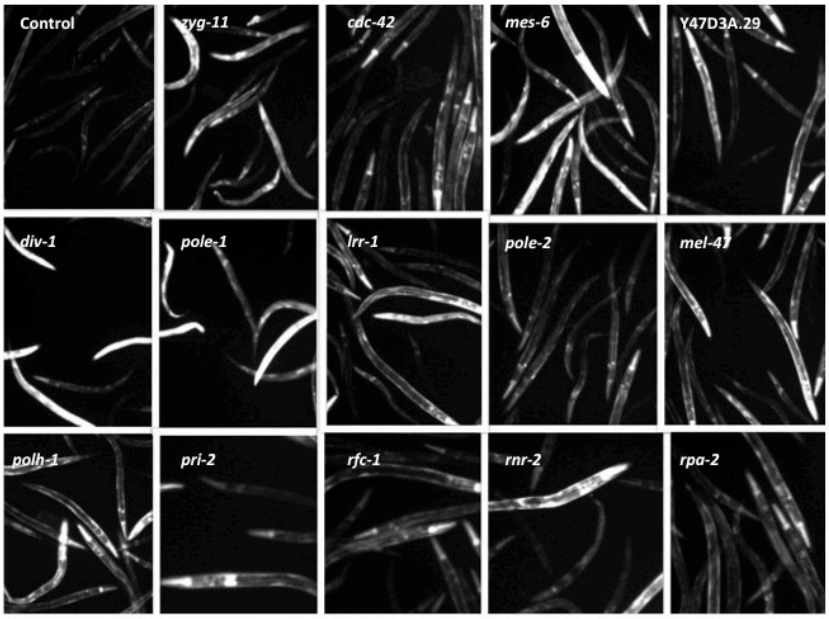


Figure 13 Field views observed during the screen depicting the hits from the DNA replication machinery category. The transgenic animals were subjected to genome wide RNAi screen, where the worms were fed for 88-90 hours with dsRNA-expressing *E. coli* in a 96 well plate after which the plate was subjected directly to screening by eye under fluorescent microscope (magnification 10x), when most of the worms were gravid adults with some wells with F1 L1 larvae. Only images of the DNA replication pathway category are shown. Control is *E. coli* RNAi strain HT115 expressing non-targeting dsRNA.

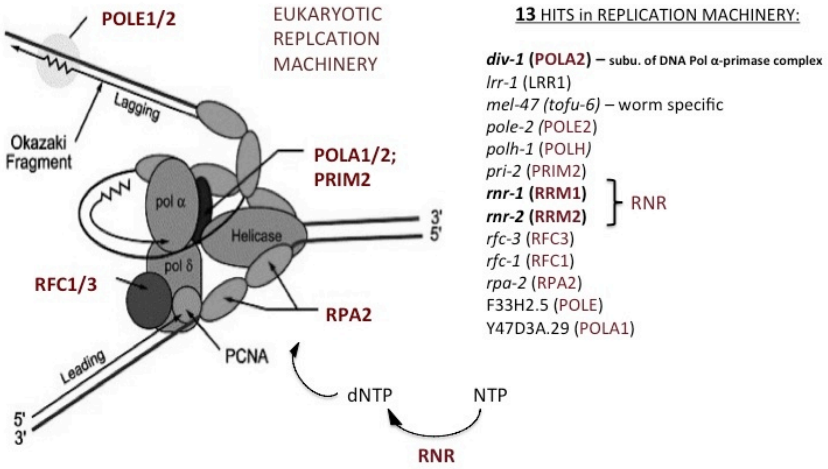


Figure 14 Many components of the DNA replication machinery were identified in the RNAi screen as causing up-regulation of the transcriptional *pdaf-21::mCherry* reporter when inhibited (See Figure 13). Most of the DNA replication category hits are components of the replication machinery, except for *rnr-1* and *rnr-2* that code for subunits of the ribonucleotide reductase complex catalyzing the biosynthesis of deoxyribonucleotides. Another exception not directly a component of the replication machinery is the *polh-1*, that encodes a DNA polymerase eta orthologous to human POLH, required for lesion bypass during DNA replication after DNA damage [202, 203]. The genes are shown on the model of an eukaryotic replication machinery, and listed on the right with their mammalian equivalents in brackets (Replication machinery adapted from [204]).

From the subgroup of replication machinery components, one of the hits, *div-1* (a homolog of the B subunit of the DNA polymerase alpha-primase complex), caused a very strong and consistent up-regulation of the *pdaf-21::mCherry* and was used in later experiments as a representative of the category of hits. To validate that *div-1* knockdown causes up-regulation of the transgene, we used the *div-1* hypomorphic mutant strain EU548 *div-1(or148)*. Indeed, when the *pdaf-21::mCherry* transgene was crossed into the *div-1* mutant background, the transgene became induced

and displayed developmental delay as with the *div-1* RNAi feeding (Figure 15).

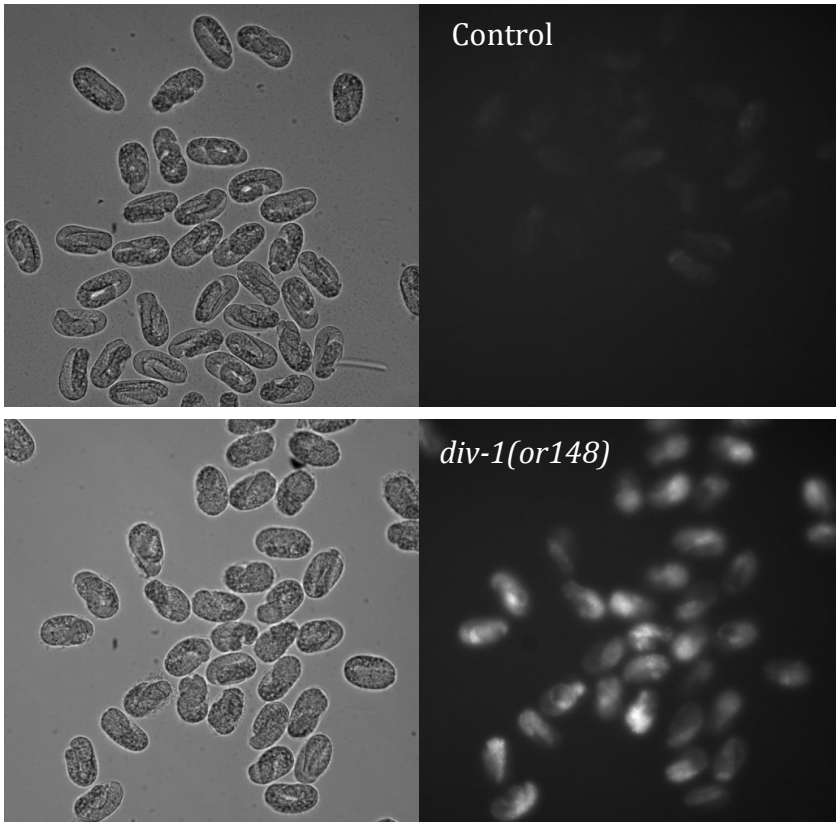


Figure 15 The *pdaf-21::mCherry* multi-copy transgene expression in the wild type background and in *div-1(or148)* background. The transgene in the *div-1* mutant is upregulated as seen before with the *div-1* RNAi fed worms (Figure 13). The worms were synchronized and images taken 400 min after the 2-cell stage. This synchronization also allows appreciation of the developmental delay in *div-1(or148)* worms, as the wild type worms at the chosen time-point are already at the 3 fold (Pretzel) stage, while in the *div-1(or148)* background many embryos are still at the comma/1.5 fold (Tadpole) stage. Also, the Ts sensitive allele of the *div-1(or148)* strain introduces up to 30% of embryonic lethality and penetrance of the phenotype increases with higher temperature. The images show brightfield (left) and the DsR (right, showing mCherry) channels with the same exposure and are adjusted to the same brightness for the comparison.

Importantly, there was no change in the copy number of the transgene in the worms fed with *div-1* RNAi compared to control RNAi fed worms in the second generation (Figure 16), suggesting that the increase in the expression of the transgene array was not due to increased repetition in the genome.

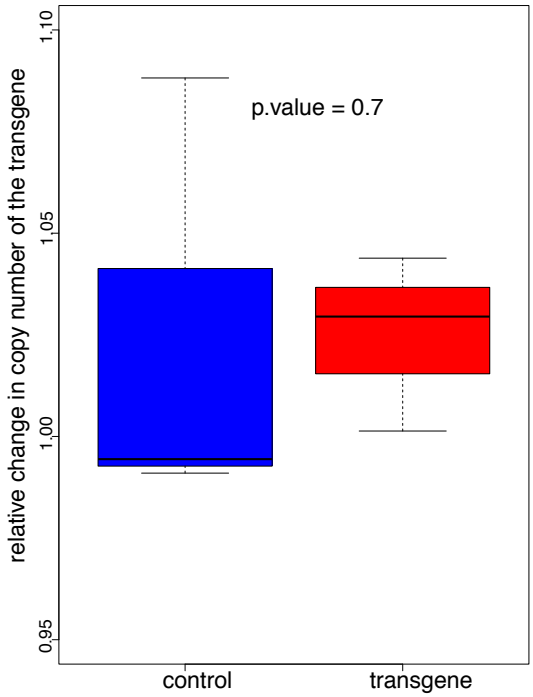


Figure 16 qPCR analysis of the genomic DNA showing no relative copy number change of the transgene in worms fed with *div-1* RNAi compared to control RNAi. The feeding was carried out as for the initial RNAi screen: L1 worms were fed for 88-90 hours with transformed *E. coli,* up-regulation of the transgene expression was confirmed by microscopy and the populations were then subjected to genomic DNA extraction followed by qPCR analysis. For the control regions, three housekeeping genes (*cdc-42*, *rpl-12* and *rpl-27*) and for transgene three different regions within the transgene were used (within the *daf-21* promoter and *mCherry* sequence, boxplots showing three regions each in two technical replicates).

2.1.1 Genes involved in DNA replication as mediators of the repressed state of the multi-copy transgene

Our screen identified genes that, when depleted, cause a strong up-regulation of the *pdaf-21::mCherry* MC reporter in the F1 generation after RNAi feeding.

We next asked what other gene expression changes can be observed after disrupting the DNA replication machinery. To address this, worms (P0 generation) were fed with control RNAi and *div-1* RNAi and the total RNA was collected for mRNA-Seq (Illumina) from the offspring (F1 generation). To address the possible gene expression differences due to the developmental delay caused by *div-1* knockdown, 4 different time-points (from L3 to L4) were chosen where control was collected earlier and *div-1* RNAi fed worms later, with 3 overlapping time points (See Methods for details). Interestingly, the principle component analysis shows that the samples do not cluster according to the treatment (*div-1* compared to control RNAi) (Figure 17), suggesting that the biggest contribution to the variance between the samples is the time-point chosen. Differential expression analysis revealed only a few (*rol-1*, *col-63*, F35E12.5, *rrn-2.1*, Y53G8AM.5, M01H9.2) genes with more than 2-fold (p-value < 10^{-5}) expression level differences between the control and *div-1* RNAi fed worms (DESeq R package). Analysis of the expression of repeat sequences (RepeatMasker) revealed few other repeated elements with more than 2-fold increase (SSU-rRNA_Cel and LSU-

rRNA_Cel) compared to control RNAi. We tested a few of the rRNA genes (such as *rrn-2.1* encoding a 5.8s rRNA, a non-coding component of the large ribosomal subunit; *rrn-1.2* encoding a 18s rRNA, a small structural non-coding RNA component of the small ribosomal subunit; *rrn-3.1* encoding a 26s/28s rRNA, a non-coding RNA component of the large ribosomal subunit; and *rrn-3.56* encoding a 26s/28s rRNA fragment, a partial copy of non-coding RNA component of the large ribosomal subunit) and estimated their expression in control RNAi and *div-1* RNAi fed worms with qPCR. The qPCR analysis (Figure 18) of these rRNA genes showed no up- or down-regulation in worms fed with *div-1* RNA compared to the control, suggesting that the differential expression signal could be due to the incomplete and variable removal of rRNA during the library preparation.

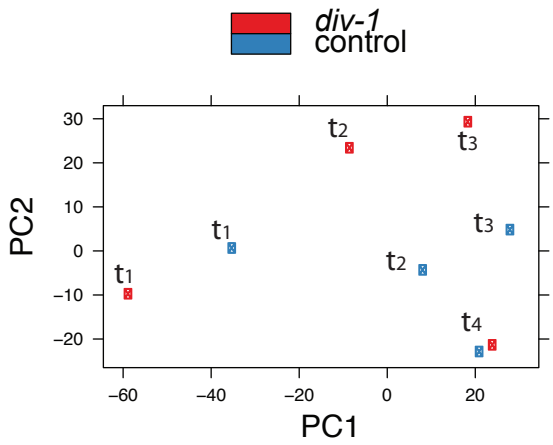


Figure 17. Scatterplot showing the samples along the first two principal components. The plot shows the samples in a 2-dimensional plot of the first two axes of a principal component analysis (PCA, DESeq package). The PCA analysis suggests that the highest contribution to the variance between the samples is the time point. The samples: control RNAi fed worms (t1-t4,

collected 34, 38, 41 and 44 hours post-L1) and *div-1* RNAi-fed worms (t1-t4, collected 38, 41, 44 and 47 hours post-L1, accounting for the developmental delay of the *div-1* RNAi fed worms compared to control).

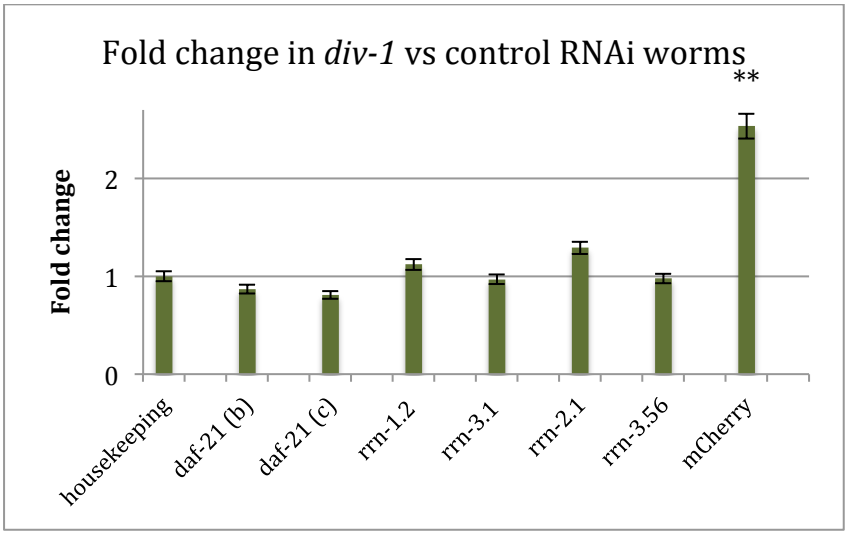


Figure 18 qPCR analysis from the genomic DNA showing gene expression differences between *div-1* and control RNAi fed *pdaf-21::mCherry* worms. Two housekeeping genes (*cdc-42* and *pmp-3*) were used and their average fold change value was set to 1. For endogenous *daf-21* region two sets of primers were used ((b) and (c)). The ribosomal genes (*rrn-1.2*, *rrn-3.1*, *rrn-3.56)* were selected based on the mRNA-Seq analysis. Significant fold change was observed only for the *mCherry* amplicon (marked with **, p-value < 0.001), each bar represents three biological in two technical replicates.

2.1.1.1 Where does the inhibition of the DNA replication machinery interfere with the inheritance of the repression of the transgene?

To establish where the inhibition of the DNA replication machinery interferes with the (inherited) repression of the transgene, we analyzed the expression of the transgene in the parental P0 (after *div-1* RNAi), as well as in the next, F1

generation, in *div-1(or148)* background (and previously during the screen with RNAi). The L1 larval worms were fed with *div-1* RNAi, and the expression of the *pdaf-21::mCherry* transgene was measured in the late L4/young adult worms. No significant up-regulation of the transgene was observed compared to the worms fed with control RNAi, except for the two posterior gut cells of the worm (Figure 19), possibly attributed to endoreduplication of the DNA, resulting in nuclei that are 32-ploid in the adult intestine [205]. This suggests that knocking down *div-1* in the P0 generation is not sufficient to induce the expression of the multicopy transgenic array widely within the P0 generation itself, supported by data from genetic crosses presented in the next section. Though based solely on the RNAi feeding experiments, it is possible that the knockdown of the *div-1* RNAi requires more than just one generation for sufficient depletion of the gene product required to induce the expression of the array.

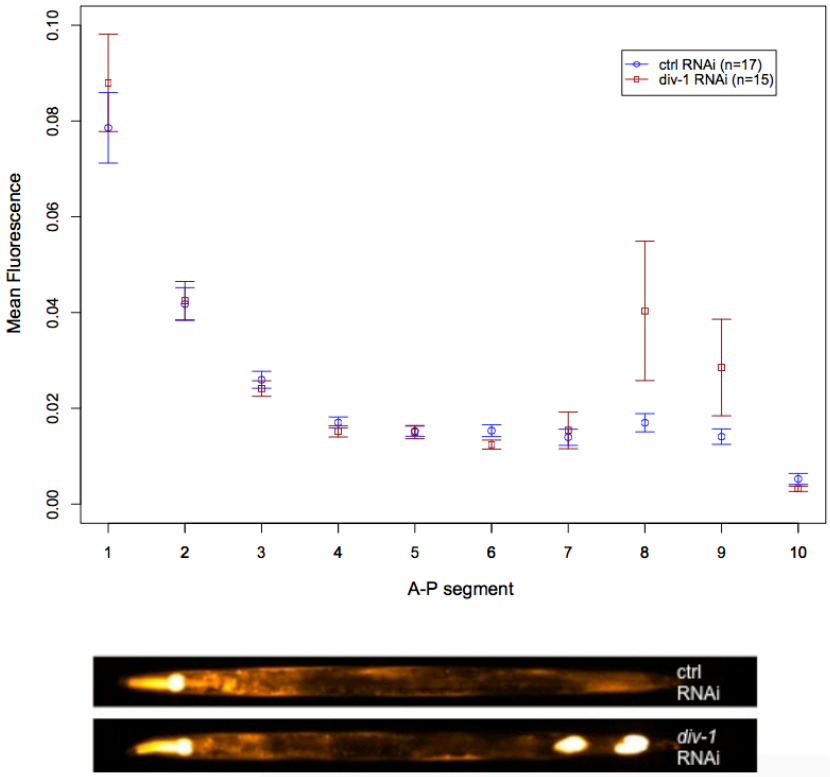


Figure 19 There is no over-all induction of the *pdaf-21::mCherry* transgene expression in the P0 worms fed with *div-1* RNAi compared to control RNAi (upper panel showing the mean fluorescence). The P0 generation was fed from L1 to L4/young adult with *div-1* and control RNAi and the *daf-21::mCherry* transgene expression was measured. The worms were divided into 10 equal segments (with CellProfiler) and the average fluorescence for each segment is shown (error bars show SEM, control RNAi (n=15), *div-1* RNAi (n=17)). Lower panel shows the straightened worms (CellProfiler) of control and div-1 RNAi fed worms, where in *div-1* RNAi fed worms the increased transgene expression can be appreciated.

2.1.1.2 Genetic crosses to determine the contribution of the maternal and paternal germline to the transgene expression in the offspring

Since there is no derepression of the transgene in P0 *div-1* RNAi-fed generation of worms (Figure 19), we asked whether the state of the replication machinery in the maternal or paternal germline is accountable for the repression of the multi-copy transgene in the progeny. For that, genetic crosses (summarized in the Table 3 and Figure 20) were carried out. In summary, the *pdaf-21::mCherry* transgene expression was measured in the offspring (F1 generation) of the parental P0 generation where either of the parents contributed the *div-1(or148)* mutant allele.

Paternal germline (homozygous for the transgene)	**Maternal germline (homozygous for the transgene)**	**F1 genotype (heterozygous for the transgene)**
pdaf-21::mCherry div-1 +/+	*div-1 +/+*	*pdaf-21::mCherry; div-1 +/+*
pdaf-21::mCherry; div-1 -/-	*div-1 +/+*	*pdaf-21::mCherry; div-1 +/-*
pdaf-21::mCherry; div-1 +/+	*div-1 -/-*	*pdaf-21::mCherry; div-1 +/-*

Table 3 Set-up of the genetic crosses to identify the maternal and paternal contribution on the multi-copy transgenic array. For every strain, the presence of the *pdaf-21::mCherry* array has been shown, as well as the 'state' of the replication machinery (*div-1* + or -) on both alleles. (*)To be able to distinguish between the self-progeny of the hermaphrodites and the progeny from the cross, an additional reporter was used (*myo-2::gfp*) when the *pdaf-21::mCherry* transgene was maternally contributed.

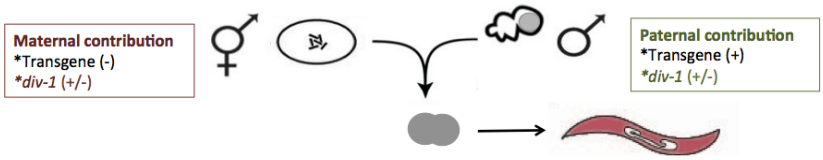


Figure 20 Setup of the genetic cross to estimate the induction of the paternally introduced transgene in the F1 generation during embryonic development (based on Table 3). For clarity, only paternally supplied transgene is shown. *div-1* +/- marks the presence or absence of the wild type *div-1* allele. The transgene in the early 2-cell embryo is not expressed (depicted gray); expression is turned on during embryonic development (red larva), the induction of which was measured with timelapse microscopy (Figure 21).

The progeny (F1) were analyzed following the expression of the transgene during embryonic development. Significant up-regulation of the transgene was measured in the embryos from parents carrying the transgene in the wild type *div-1* alleles in sperm and *div-1* homozygous mutation with no transgene from the oocyte (Figure 21), suggesting that the reduced replication machinery in the early embryo is interfering with the inheritance of the repressive marks during embryonic development and that maternally provided proteins are required for proper epigenetic inheritance.

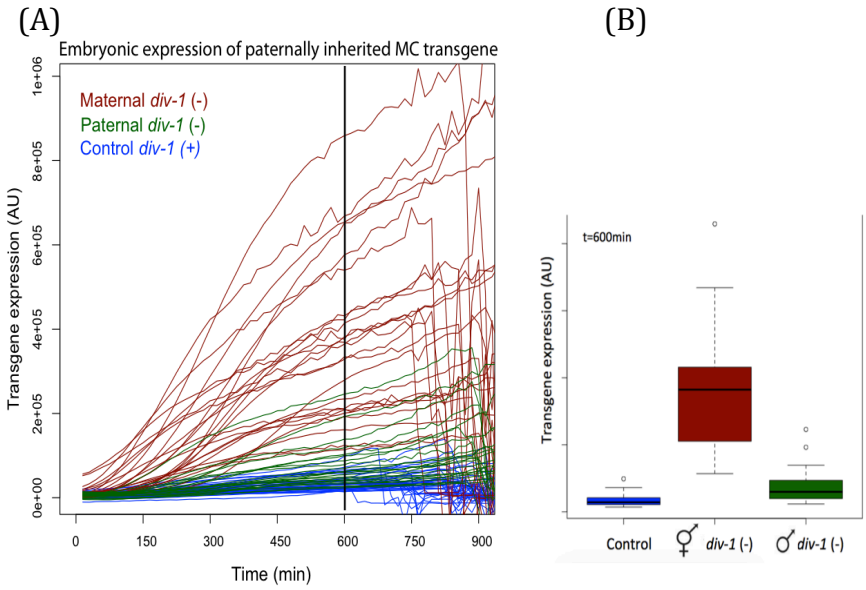


Figure 21 (A) Timelapse analysis of the induction of the multi-copy *pdaf-21::mCherry* transgene expression during the embryonic development (from early embryo until hatching) of the progeny from the specified genetic cross. *div-1*(+/-) denotes the presence/absence of the wild type *div-1* allele. The plot shows the expression of the transgene in embryos from 2-4 cell stage until hatching (messy lines from 750 min timepoint indicate hatching). Each line is one embryo and the colors of the lines are representing contribution of the *div-1* hypomorphic mutation. (B) Boxplot showing one selected timepoint (at 600min) for the transgene expression, where maternally contributed *div-1* hypomorphic mutation induced the paternally inherited transgene expression more than the paternally inherited *div-1(-)*, color-coding as in (A). Control worms (N2) had wild *div-1* alleles and no transgene.

2.1.2.2 Are other repetitive transgenes de-repressed when the DNA replication machinery is dysfunctional, and can we de-repress the pdaf-21::mCherry array with other methods than RNAi feeding?

The RNAi screen was carried out with *pdaf-21::mCherry* transgene that is a repetitive integrated array of the inserted DNA or its recombined fragments in the reporter strain

used. These heritable repetitive tandem arrays containing hundreds of copies of the DNA are often assembled into condensed transcriptionally silent chromatin in the germline in *C. elegans* [154]. We speculated that this silencing is relieved when the replication machinery is challenged, causing up-regulation of the multi-copy *mCherry* reporter by challenging the proper acquisition of the silencing patterns of these repetitive transgenic arrays. We then asked whether other multi-copy and/or stress reporter transgenes in *C. elegans* can be de-repressed in the same way and whether we could induce this de-silencing by using distinct agents other than RNAi and mutations.

2.1.2.2.1 De-repression of other multi-copy and/or stress reporter transgenes

Melo and Ruvkun [206] had previously shown that RNAi and toxin-mediated disruption of core cellular activities stimulate behavioral avoidance of normally attractive bacteria. This also induces expression of detoxification and innate immune effectors, even in the absence of toxins and pathogens. They showed that surveillance pathways in *C. elegans* monitor core cellular activities and when these pathways are disrupted, specific behavioral, immune, and detoxification responses are encaged, similar to those activated during pathogen attack. Melo and Ruvkun took advantage of several GFP reporters previously shown to represent activation of innate immune responses and we

tested a set of these cellular stress reporters feeding these strains with *div-1* RNAi and monitoring the expression of these stress reporters. The reporters tested included: *hsp-4::GFP* for endoplasmic reticulum unfolded protein response (ER^{UPR}) [207], *gst-4::GFP* (glutathione S-transferase detoxification reporter), *hsp-6::GFP* for mitochondrial unfolded protein response (UPR^{mito}) [208], *sod-3::GFP* for oxidative stress [209], F35E12.5::GFP for pathogen response such as *Y. pestis, M. nematophilum, P. aerugionsa* [210], [211], [212], *cyp-35::GFP* as intestinally expressed cytochrome P450 oxidase [213] and *nlp-29::GFP* as conserved glycine/tyrosine-rich antimicrobial peptide induced by *D. coniospora, S. marcescens*, and wounding [214]. We tested the expression of this selection of reporters in *div-1* RNAi-fed worms alongside the RNAi clones reported by Melo, J. and Ruvkun, G. to induce the expression of these reporters. Where available, agents known to induce appropriate stress were used as well (Table 2). Interestingly, there was no over-all consistency regarding which cellular stress reporters could be induced (for example oxidative stress reporter *gst-4::gfp* was induced by *div-1* RNAi, but *sod-3::gfp* reporter was not). In summary, 5 out of 7 cellular stress reporters tested were not induced by *div-1* RNAi, suggesting that inhibition of *div-1* does not up-regulate the transgene because it triggers a stress response.

Reporter	RNAi (by Melo, J., Ruvkun, G. 2012)	Cellular stress	Other agents	Up-regul. by agent	Up-regul. by RNAi
hsp-4::gfp	*sams-1*	ER	Tunicamycin	Yes	*sams-1(+), div-1(+)*
gst-4::gfp	*cco-1*	oxidative	Paraquat	-	*cco-1(+), div-1(+)*
hsp-6::gfp	*cco-1*	mitochond. protein handling	EtBr	-	*div-1* (-)
sod-3::gfp	*rpl-1*	oxidative	Salt	Yes	*rpl-1(+), div-1(-)*
F35E12.5::gfp	*sca-1, eft-2*	pathogen	*Y. pestis*	-	*sca-1(+), div-1(-)*
cyp-35::gfp	*prp-21, eat-6*	detox	Xenobiotics	-	*prp-21(+), div-1(-)*
nlp-29::gfp	*pept-1, nhr-25*	antimicrobial	Starvation, M9	Yes	*pept-1(+), div-1(-)*

Table 4 Reporters of innate immune programs, genes encoding xenobiotic detoxification enzymes and general cellular stress tested for induction by *div-1* RNAi (reporters from [206]). The table lists the set of cellular stress reporters tested along with their specified function, and other agents shown to induce these reporters. The last two columns show conditions tested in this study, whether these transgenes were inducible by the exogenous agents or RNAi clones previously confirmed to induce the expression of the transgene, alongside with the *div-1* RNAi from our study ('+' depicts induction, '-'no induction). The RNAi feeding was carried out as for the initial RNAi screen: the L1 worms were fed 88-90 hours with RNAi and F1 embryos and L1 larvae were monitored for the transgene expression.

We also compared the induction of low- vs. high-copy number transgenes, for the same *pdaf-21* construct. Interestingly, even though there was a quantifiable increase in the expression of a single copy reporter with the same *daf-21* promoter construct (Figure 22), the induction was much smaller than for the multi copy reporter (1.8 fold with the single copy and 20.5 fold with the multicopy).

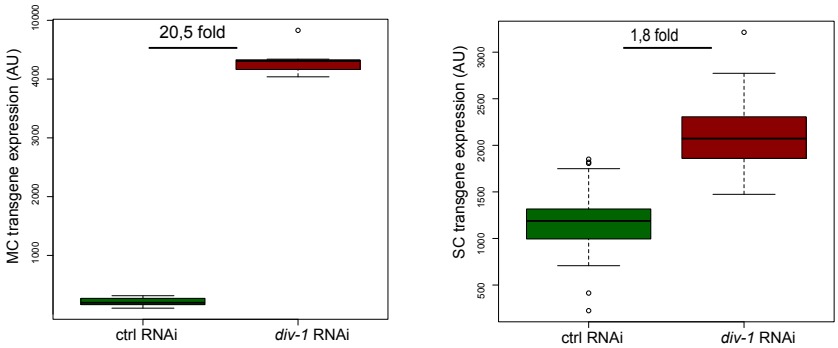


Figure 22 Multi-copy (MC) and single-copy (SC) strains of the same *pdaf-21::GFP* construct. The multi-copy strain displayed 20.5 fold increase in worms fed with *div-1* RNAi compared to control, whereas the single copy strain displayed only 1.8 fold increase in the transgene expression (p-values < 0.0005, Wilcoxon rank test). The P0 animals were fed with the control and *div-1* RNAi and the expression was measured in the F1 embryos.

Lastly, the effect of *div*-1(RNAi) on the expression of additional multicopy transgenes was tested. The selected transgenic strains were a strain carrying *let-858::gfp*, previously shown to become partially derepressed in class-B synMuv mutant backgrounds [181] and worms with the *scm::gfp* [215] and *sur-5::sur-5::gfp* arrays that ubiquitously express GFP in all cells [216] and show enhanced silencing upon *lin-35* or *lin-15B* inhibition [186].

Two of these transgenic strains (*pdaf-858::let* and *pscm::gfp*) showed a slight increase in the expression of the transgene (Figure 23), while one (*psur::sur-5::gfp*) showed a slight decrease. This suggests that the induction of the *pdaf-21::mCherry* array is due to a different mechanism than those previously described.

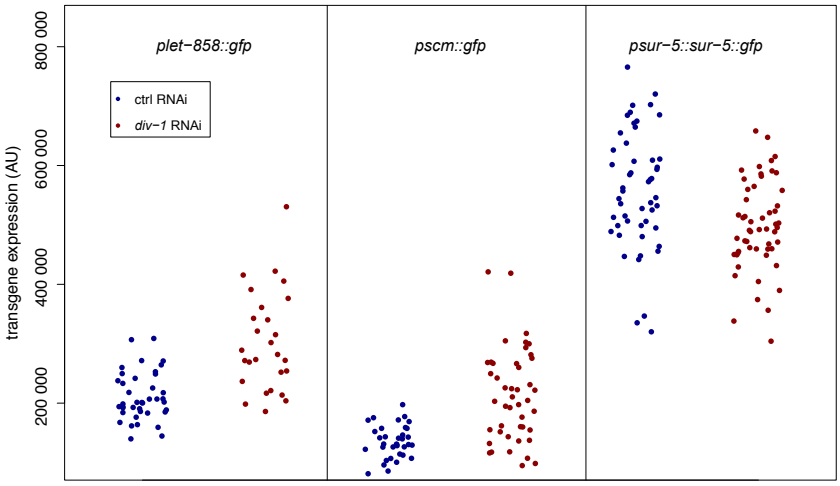


Figure 23 Other multi-copy transgenes previously shown to become partially de-repressed in other studies with multi copy transgenic arrays. The P0 generation was fed with *div-1* RNAi and control RNAi after which the expression of the transgene was measured in the L1 larvae of the F1 generation. Each dot represents one individual worm. There was only a modest increase in the *plet-858::gfp* and *pscm::gfp* expression (p-values *plet-858::gfp* and *pscm::gfp* < 0.0005) and a mild decrease in the *psur-5::sur-5::gfp* expression (p-value 0.002, Wilcoxon rank test).

2.1.2.2.2 De-repression of the multi-copy transgene using other agents than RNAi

To test the hypothesis that the de-silencing of the repressed transgenes can be coupled to the replication stress and/or due to a stalled replication fork, we used a set of available techniques to generate DNA damage and replication stress in worms.

In worms, the developing embryo and germline cells have different responses to DNA damaging agents, due to varying activations of the checkpoint pathways [217, 218]. The mitotic germline nuclei arrest their proliferation in

response to the DNA damage to allow time for DNA repair. In the mitotic part of the germ line, the cells with DNA damage are removed before oogenesis by apoptosis. In the early embryo, where the cell divisions are very fast, this DNA damage pathway is not activated by the environmental stimuli, but is rather actively silenced as the translesion synthesis pathway ensures that DNA damage does not slow down the early cell cycles [219].

2.1.2.2.2.1 Hydroxyurea, an agent causing replication stalling in C. elegans weakly de-represses the multi-copy transgenic array

Hydroxyurea (HU) depletes the cellular pools of dNTPs and thus directly inhibits DNA synthesis [220]. It specifically inhibits the ribonucleotide reductase, an enzyme whose two subunits in *C. elegans* are encoded by *rnr-1* and *rnr-2*, two genes that emerged from the de-repressing RNAi screen. Also, defects in *div-1*, a gene encoding DNA polymerase alpha and one of the main hits from the RNAi screen, delays cell division and the defects are accompanied by defects in the early asymmetric cleavages that produce the embryonic founder cells [221]. Hypomorphic mutations in *div-1* cause replication problems that lead to inappropriate activation checkpoint pathways of the *chk-1* pathway, the *C. elegans* homologue of Chk1. Checkpoint-mediated asynchrony in cell division is essential for embryonic patterning and thus must occur only in response to

developmental signals and not in response to unscheduled events such as replication defects, for example caused by DNA damage. Encalada *et al.* [221] showed that mutations in *div-1* cause delay in early embryonic cell divisions, and similar effects are caused by the use of HU [222], [219]. In these cases they speculated that embryonic sensitivity causes checkpoint activation and extends the natural asynchrony of cell division. Though, in these cases this increased asynchrony leads to cell death. Embryonic cells in some *div-1* mutants appear to attempt to go through mitosis without completing DNA replication, resulting is severe cell division defects [221]. Interestingly, some less severely defective mutants, the embryonic cell cycle checkpoint might delay the onset of mitosis until the completion of DNA replication [221].

We tested the effect of HU on our multi-copy transgenic worms in order to test whether the *div-1* knockdown causes the de-repression through the same mechanism. Even though treatment with HU caused up-regulation of the array (Figure 24, A), compared to *div-1* RNAi the effect was much less strong (Figure 24, B). The strength of the HU treatment was tested with a dilution series and monitoring the Emb and Ste (embryonic lethality and sterility) phenotypes, with the concentration of HU used (25mM), we scored 20% Emb and 40% Ste worms.

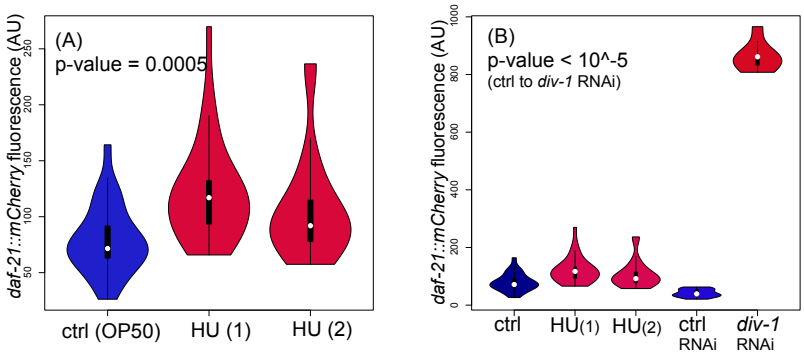


Figure 24 Induction of the *pdaf-21::mCherry* multi-copy transgenic array after growing the worms on plates with hydroxyurea (25mM) (panel A) and induction by *div-1* RNAi (panel B). Significant up-regulation was measured in L1 worms after treatment with HU (HU(1) and HU(2) represent independent experiments). In comparison to HU (compare panel A and B), *div-1* RNAi increased the transgene expression 13-15x more (1.35-1.5 x increase with HU compared to 20.5x increase with *div-1* RNAi, (p-values 0.0005 and < 10^5, Wilcoxon rank test)). The control worms in the HU experiment were fed with the standard *E. coli* strain OP50, whereas in the RNAi experiment control RNAi is used (HT115 *E.coli* strain).

2.1.2.2.3 UV-C radiation and EMS do not significantly de-repress the multi-copy pdaf-21::mCherry transgenic array

UV irradiation is a common means to induce a DNA damage response in *C. elegans* in order to study DNA repair and DNA damage response pathways. Importantly, the various DNA damaging agents have different consequences for the embryonic cell divisions and the germline cells –embryonic cells are able to tolerate relatively high levels of DNA damage using error prone polymerases as the embryonic divisions are very rapid at the expense of genomic integrity [223]. In contrast, the

germline has much longer cell cycles and is consequently more sensitive to DNA damaging agents. In the germline, DNA damage leads to prolonged G2 cell cycle arrest of mitotic germ cells, towards the distal end of the gonad [218]. To distinguish between these two distinct responses of the different cell cycle types, both the germline and the embryos were subjected to the damaging UV irradiation. To accomplish this, both the larval worms with developing germline (young adults) and early embryos were treated with UV irradiation and the expression of the *pdaf-21::mCherry* transgene was followed either throughout the embryonic development until hatching or measured at a certain developmental stage in the embryos (comma stage). In neither of the cases, significant increase in the fluorescence could be measured (Figure 25).

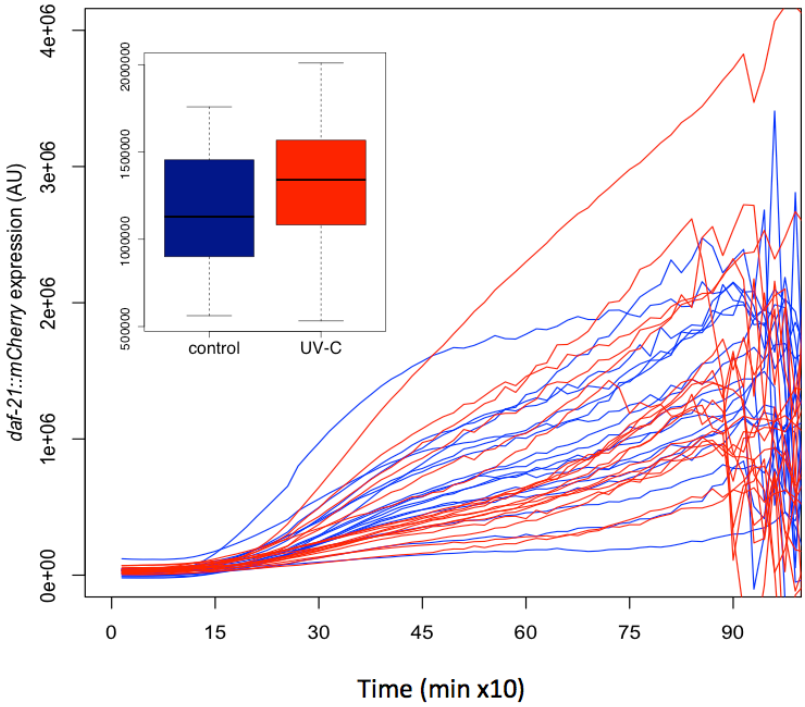


Figure 25 The expression of the *pdaf-21::mCherry* transgene was not increased after the UV treatment. Timelapse microscopy following expression of *pdaf-21::mCherry* transgene of developing embryos after irradiating the adult worms with UV-C. The graph shows expression of the transgene during embryonic development from 2-4-cell stage until hatching (observable as messy lines), blue lines show the control worms and red lines the UV-treated worms. Boxplots: fluorescence in the embryos measured after UV-C treatment of the embryos (measured at comma stage, 5-6 hours after UV-C), the difference was non-significant (p = 0.24, Wilcoxon rank test).

Ethyl methanesulfonate (EMS), one of the most commonly used and most potent mutagens used in *C. elegans* [150, 224] is an ethylating agent causing DNA damage that results in double-stranded breaks. We tested the potential of EMS to induce the expression of the *pdaf-21::mCherry* array. No induction was observed (Figure 26), though up to 20% increased lethality and sterility were detected, suggesting that the dose of EMS was sufficient.

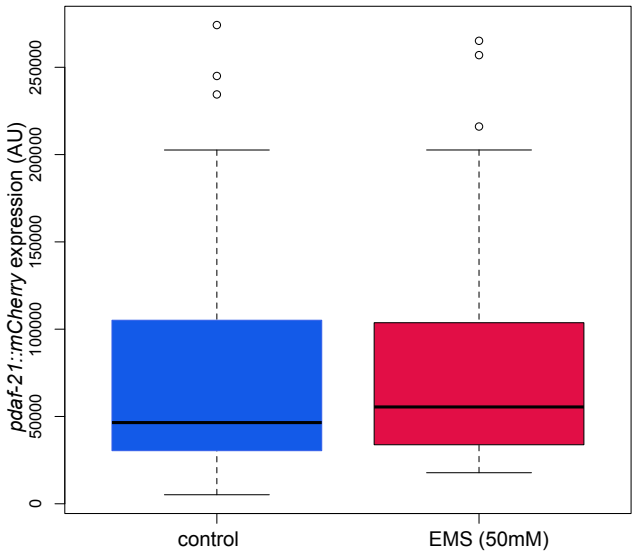


Figure 26 Treatment with EMS was not able to induce the expression of the *pdaf-21::mCherry* transgene array as compared to control worms. Synchronized L4 worms were washed into 50ml tube and incubated in 50mM EMS (in M9) for 4hours. The animals are then thoroughly washed and plated onto regular NGM feeding plates and the expression of the *pdaf-21::mCherry* multi-copy transgene was measured in their offspring in L1 larvae, n(control)=70 and n(EMS)=75. p = 0.353 , Wilcoxon rank test. The efficiency of the EMS treatment was estimated with increase in the penetrance of the Let (lethality), Emb (embryonic lethality) and Ste (sterility) phenotypes after the treatment compared to the control condition.

2.1.2.2.3 Treatment with HU and UV-C, but not *div-1* RNAi induces HUS-1::GFP foci in the adult germline.

To evaluate whether the methods we selected are indeed able to cause DNA damage, we used a *hus-1::GFP* strain as an estimation of the range of DNA damage caused. HUS-1 is a nuclear protein that is expressed in early embryos and the adult germlines and accumulates into distinct foci at putative sites of DNA damage, overlapping with chromatin, whereas it is diffusely distributed in nuclei in the absence of damage [225].

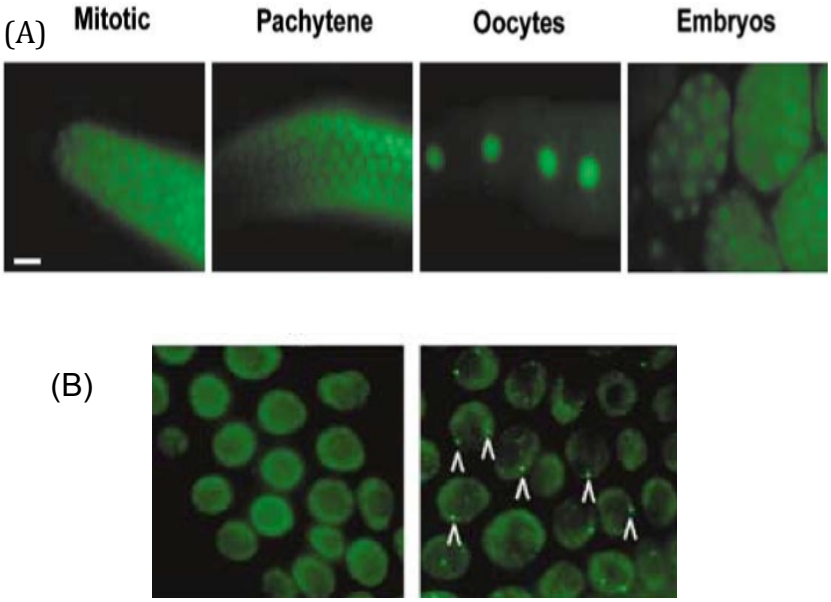


Figure 27 HUS-1::GFP localization in the nuclei of proliferating germ cells, meiotic germ cells, mature oocytes and embryos (panel A, in green), while it forms foci at sites of DNA damage overlapping with chromatin (panel B, white arrows pointing the foci, figure from [225]).

We evaluated the expression of the *hus-1::GFP* transgene in worms treated with HU, UV and *div-1* RNAi. The HUS-1::GFP foci are readily observable in the germline of animals treated with DNA damaging agents (Figure 27 and 28), whereas we could not detect any distinct foci in the embryos. There was an interesting variation between the ability of the treatments to cause DNA damage as estimated by the number of HUS-1::GFP foci. The *div-1* RNAi-fed worms having the lowest number, while efficiently being able to desilence the transgene array. In contrast, worms that were grown on plates with 25mM HU exhibited a high number of HUS-1 foci in the germline, though only mildly de-repressing the transgene (Figure 24). In case of the UV treatment, there was no up-regulation of the transgenic array (Figure 25), whereas up to 10 times more HUS-1::GFP foci were observed (Figure 28).

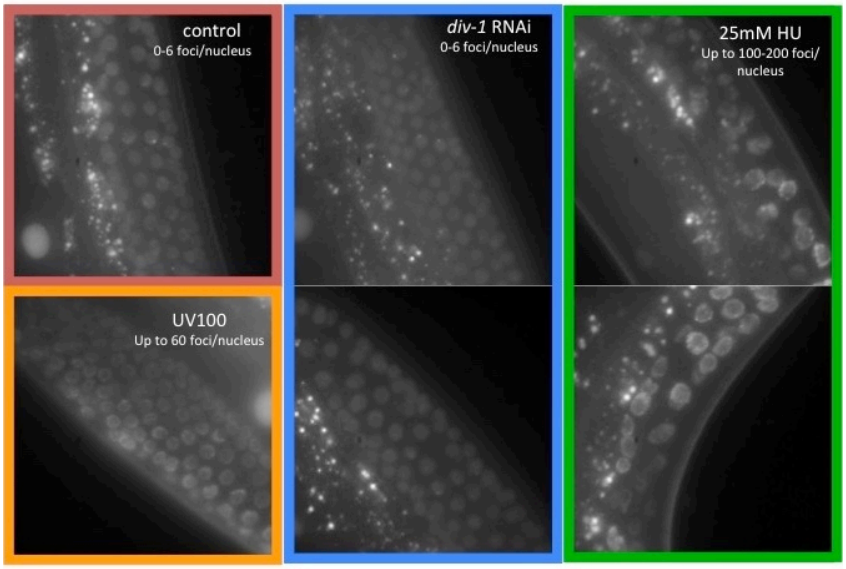


Figure 28 The HUS-1::GFP foci are observable at the double-stranded breaks in the nuclei co-localizing with the DNA. The control worms (grown on usual OP50 feeding plates) show 0-6 foci per nucleus. Similar estimation of the count was made for foci in the germlines of worms fed with *div-1* RNAi. The worms radiated with UV-C (100 W/m^2 and 254 nm) presented up to 60 foci per nucleus, whereas worms grown on NGM plates with 25mM hydroxyurea (HU) had up to 200 foci and showed enlarged nuclei as characteristic to the HU treatment (the counts of the foci are an estimation due to lack of the resolution).

2.1.1.3 Interfering with regulators of chromatin structure further derepresses the multi-copy pdaf-21::mCherry transgenic array in div-1(or148) background

To test the hypothesis that inhibition of the DNA replication machinery alters the chromatin modifications we inhibited genes responsible for depositing methylation marks on histone 3 in *C. elegans* (H3K9 and H3K36 methylations), as well as other genes previously identified in our lab as increasing expression from the *pdaf-21::mCherry* transgene when inhibited. The genes tested were:

- *lin-53* – a class B Syn Muv gene that antagonizes the Ras pathway, negatively regulating vulval development. *lin-53* encodes protein similar to Rb-associated protein p48.
- *mes-4* – SET domain protein, involved in germline silencing of repetitive arrays. MES-4 generated H3K36 methylation serves an epigenetic role by marking germline genes and carrying the gene expression memory to the next generation of germ cells.
- *mrg-1* – mortality factor related gene, contains a chromodomain and associates with methylated histone tails. MRG-1 is important for genomic integrity in meiosis.
- *mut-7* - encodes a homolog of RnaseD that represses transposition of Tc1, Tc3, Tc4, Tc5, possibly by degrading transposon-specific messages.
- *nrde-2* – functions in the nuclear RNAi pathway to regulate gene expression via inhibition of RNA Pol II via enrichment of H3K9 methylation at sites targeted by RNAi.
- *set-25* – a histone methyltransferase, solely responsible for H3K9me3 deposition in *C. elegans*.

Inhibition of all of the genes further de-repressed the transgene in the *div-1* mutant background (Figure 29).

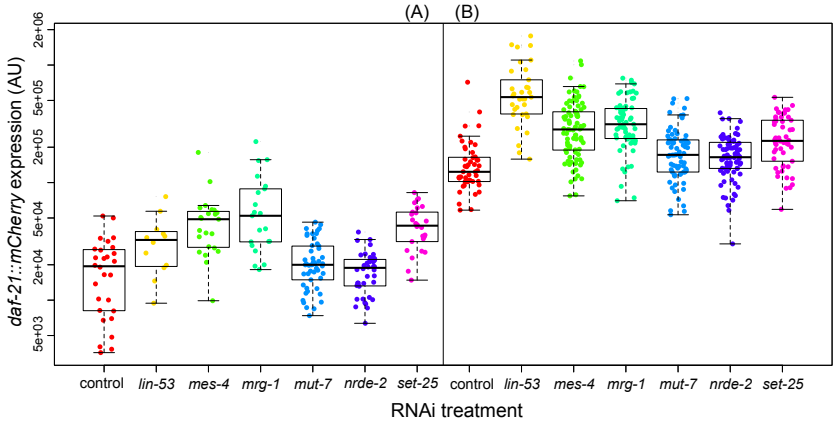


Figure 29 Expression of the multi-copy *pdaf-21::mCherry* transgene in the wild type background (panel A) or *div-1(or148)* mutant background after RNAi feeding (panel B) (clones presented on the X-axis). The parental P0 worms were fed with RNAi bacteria and the expression of the transgene was measured in the subsequent generation, in L1 larvae. The knockdown of the genes studied (see text for details) further up-regulates the transgene even in the temperature sensitive *div-1(or148)* mutant background at 20C. Fold changes for mean values of measured expression of the transgene along with the p-values compared to the control RNAi are as follows: WT *lin-53* (1.39, p-value 0.058), *mes-4* (2.26, p-value<10^{-5}), *mrg-1* (3.07, p-value <10^{-5}), *mut-7* (1.0, p-value 0.28), *nrde-2* (0.83, p-value 0.92), *set-25* (2.01, p-value < 10^{-6}) and in *div-1* background with additional fold change to wild type: *lin-53* (4.21, p-value <10^{-15}, 29.62), *mes-4* (2.08, p-value <10^{-10}, 14.56), *mrg-1* (2.22, p-value <10^{-11}, 15.58), *mut-7*(1.24, p-value 0.005, 8.72), *nrde-2* (1.12, p-value 0.005, 7.87), *set-25* (1.61, p-value <10^{-5}, 11.26) and *div-1* background to WT 7.01 fold, p-value <10^{-20}. Y-axis in log scale.

2.1.2 The CSR-1 small RNA pathway has a protective role in regulating the expression of the *pdaf-21::mCherry* transgenic array

In *C. elegans* there are thousands of small RNAs produced, most involved in gene silencing, like Piwi-interacting small RNA (piRNA) mediated pathway in the germline. The piRNA pathway acts as a germline surveillance system, through the 21U-RNAs [59, 80, 82, 87, 88, 96]. In the Piwi pathway, the Argonuate PRG-1 coupled with the 21U-RNAs

is able to identify foreign sequences though incomplete complementarity and induces the production of the secondary 22G-RNAs [71, 80, 82, 84, 85, 87, 95, 96]. These can be bound with worm Argonautes (WAGOs) and silence transcriptionally as well as post-transcriptionally the foreign nucleic acids. Here emerges an obvious question, with this vast silencing potential, how do endogenous genes manage to avoid this vast silencing potential? One possible candidate is another Argonaute CSR-1 (Chromosome Segregation and RNAi deficient) [98, 226]. This Argonuate CSR-1 is able to bind almost all germline-expressed genes and is recruited to the target loci to chromatin and is suggested to oppose piRNA silencing to protect germline transcription. The main components of this pathway in addition to CSR-1 are EGO-1 (Enhancer of glp-1(one)), DRH-3 (Dicer Related Helicase) and EKL-1 (Enhancer of ksr-1 Lethality), whereas DRH-3 and EKL-1 are not germline restricted, functioning in WAGO 22G RNA biogenesis both in germline and soma [58, 226]. We asked, whether the CSR-1 pathway is able to prevent increase in the expression of the transgenic array induced by hypomorphic *div-1* mutation. The *pdaf-21::mCherry* transgene was down-regulated in the *drh-3(ne4253)* background (Figure 30), and in the *drh-3(ne4253)/div-1(or148)* double mutant the induction of the reporter is significantly reduced compared to *div-1(or148)* background.

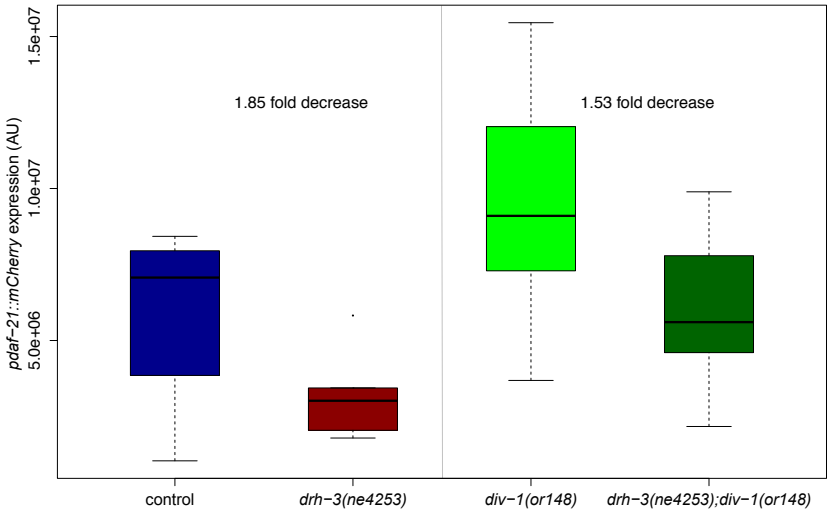


Figure 30 The missense allele of *drh-3* (coding for Dicer related helicase) is able to reduce the expression of the *pdaf-21::mCherry* transgene and reduce the induction of the transgene caused by the *div-1(or148)* hypomorphic mutation. The *pdaf-21::mCherry* transgene in the wild type background was crossed with *div-1(or148),* with *drh-3(ne4253)* or with *div-1(or148)/drh-3(ne4253)* double mutant and the expression was measured in the embryos directly from the cross. The naïve transgene was always introduced paternally and the mutation was introduced (shown on X-axis) maternally (using hermaphrodites). Embryos from the cross were collected and expression of the transgene was measured at comma stage. As a control, transgenic males were also crossed with the N2 hermaphrodites, carrying wild type alleles of both *div-1* and *drh-3*. The decrease of the expression of the transgene in *drh-3(ne4253)* background was 1.85 fold in the *drh-3(ne4253)* background and 1.53 fold in the *div-1(or148)/drh-3(ne4253)* double mutant (p-values 0.047 and 0.0018 respectively, Student's t-test).

In parallel, we also looked at the *pdaf-21::mCherry* transgene in the *rde-1(ne219)* background. *rde-1* encodes an Argonaute in *C. elegans*, acting in the exogenous RNAi pathway as opposed to the CSR-1 endogenous pathway. The transgene array was not down-regulated in the *rde-1(ne219)* background (Figure 31) and in the *rde-*

1(ne219)/div-1(or148) double mutant the induction of the transgene was not reduced as compared to *div-1(or148)* background, but rather induced further.

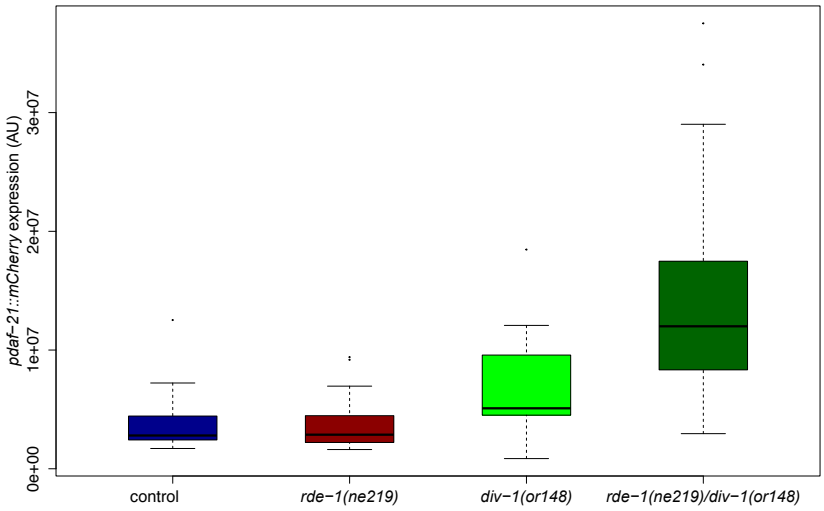


Figure 31 The substitution allele of the *rde-1* (coding for RNAi-DEfective 1, a primary Argonuate) is not able to decrease the expression of the *pdaf-21::mCherry* transgene and not able to reduce the induction of the transgene caused by the *div-1* hypomorphic mutation (compare to Figure 30). The *pdaf-21::mCherry* transgene in the wild type background was crossed with *div-1(or148), rde-1(ne219)* or with the *div-1(or148)/rde-1(ne219)* double mutant and the expression was measured in the L1 larvae directly from that cross. The naïve transgene was always introduced paternally and the mutation was introduced (shown on X-axis) maternally (using hermaphrodites). As a control, the transgenic males were also crossed with the N2 hermaphrodites, carrying wild type alleles of both *div-1* and *rde-1*. P-values (Student's t-test) 0.87 (control to *rde-1(ne219),* 0.009 (control to *div-1(or148),* 10^{-10} (control to *rde-1(ne219)/div-1(or148);* 0.0001 (*div-1(or148)* to *rde-1(ne219)/div-1(or148)*, 10^{-11} *(rde-1(ne219)* to *rde-1(ne219)/div-1(or148)* and 0.006 *rde-1(ne219)* to *div-1(or148).*

2.2. Both repressive and active Histone 3 methylation levels exhibit global changes after the inhibition of DNA replication

2.2.1 Heterochromatic mark H3K27me3 is globally reduced in the worms after inhibition of DNA replication whereas H3K9me3 is less affected

The germline of *C. elegans* shows a remarkable ability to specifically and reliably silence transgenic DNA, even in low to medium copy number. Chromatin factors are required for the maintenance of transgene repression as mutations in MES-2 and MES-6 (Polycomb related proteins in *C. elegans*) are able to disrupt the silencing [159]. Also, Kelly *et al*. showed that heterochromatic histone 3 lysine 9 (H3K9) dimethylation marks are enriched on repetitive arrays [50]. In the germline, the silencing mechanisms can be disrupted in mutant animals of the maternal effect genes such as *mes-2*, *mes-3*, *mes-4*, *mes-6*. Interestingly, whereas in the *div-1(or148)* background the *pdaf-21::mCherry* transgene was very effectively desilenced (Figure 15), it was still not expressed in the germline (Figure 32).

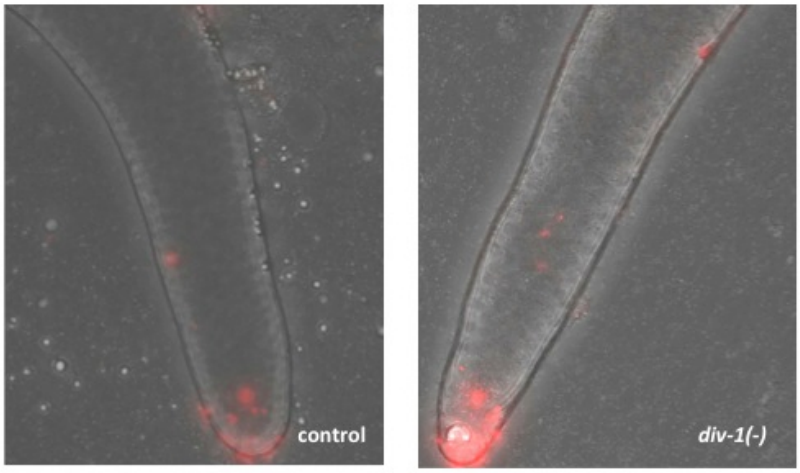
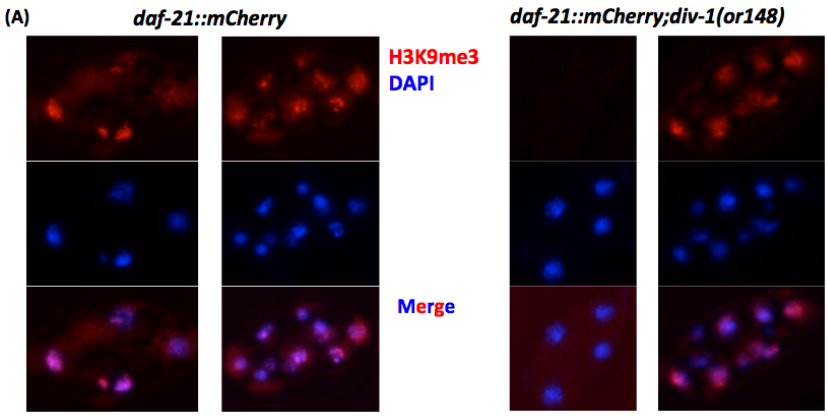


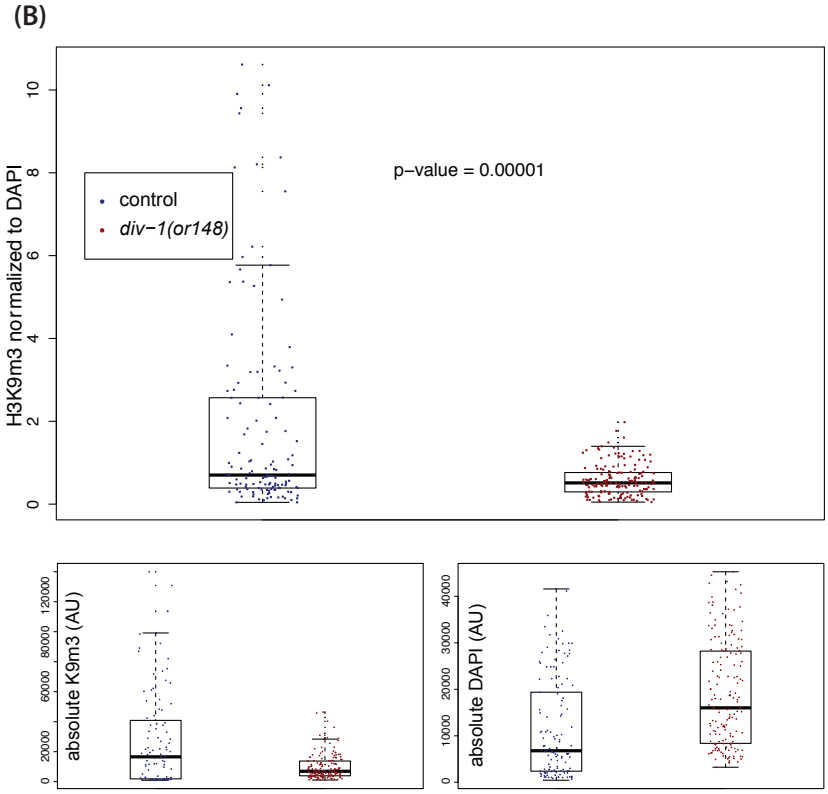
Figure 32 The *pdaf-21::mCherry* multi copy transgene is silenced in the germline in control RNAi-fed worms as well as *div-1* RNAi-fed worms. In both of the conditions, the transgene expression in the germline is only seen in the DTC, a somatic cell at the tip of each gonad arm (shown in red).

Based on the literature, we initially tested for changes of the repressive marks H3K9me3 and H3K27me3 specifically on the transgenic multi-copy array in worms after replication stress (*div-1* RNAi or *div-1* Ts mutant), using both immunofluorescence and ChIP-qPCR. Interestingly, immunofluorescence staining with H3K9me3 and H3K27me3 specific antibodies showed a decrease over the whole genomic DNA in the nuclei of the embryos studied (Figures 33, 34). This observation was confirmed by ChIP-qPCR analysis of gravid worms with the same antibodies (Figures 33 and 34, panel C), although for H3K9me3 the decrease was less consistent.

There was a large technical variation in the immunofluorescence signal both within the same

experiment and between batches, requiring a normalization for which we used DAPI staining marking the DNA of the Immunofluorescence-stained embryos. The quantification of the antibodies is presented for the non-normalized antibody signal and also as normalized to the DAPI signal from the same nucleus.


(A)
daf-21::mCherry
daf-21::mCherry;div-1(or148)
H3K9me3
DAPI
Merge


(B)
H3K9m3 normalized to DAPI
0
2
4
6
8
10
control
div−1(or148)
p−value = 0.00001
absolute K9m3 (AU)
0
20000
60000
80000
120000
absolute DAPI (AU)
0
10000
20000
30000
40000

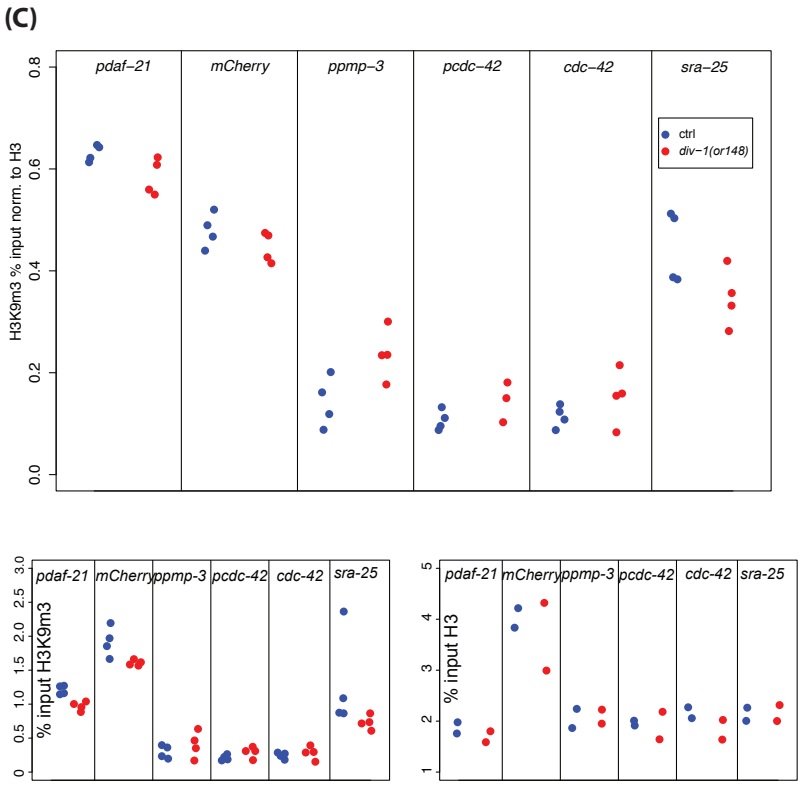


Figure 33 Repressive chromatin mark H3K9me3 is reduced after inhibition of the DNA replication machinery in *div-1(or148)* worms, in the immunostaining experiments.

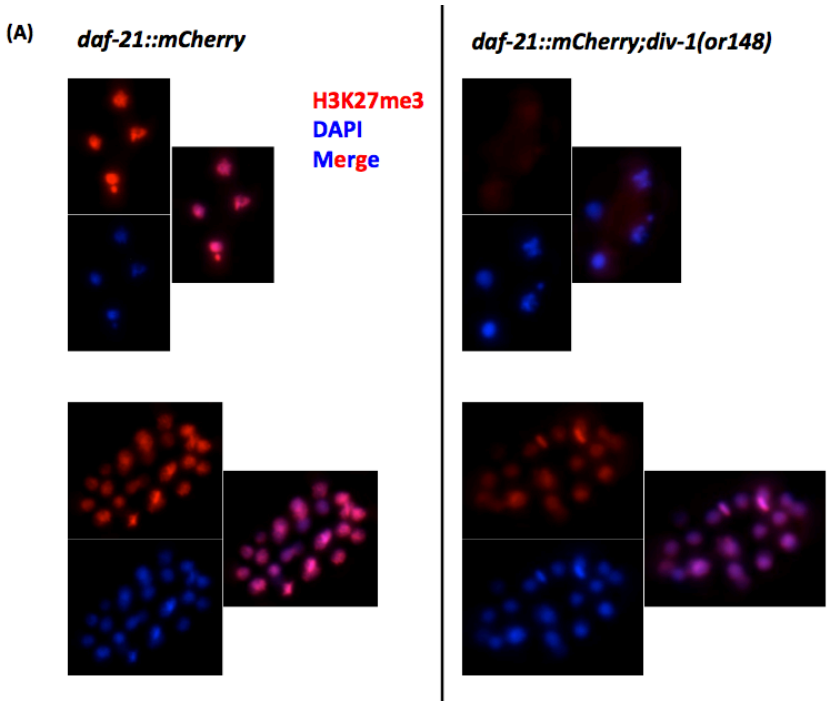
Panel A: Immunostaining experiments with anti-H3K9me3 antibody. Young gravid control and *div-1(or148)* worms were bleached and the extracted young populations of embryos were prepared with freeze-cracking method and subjected to methylation-specific primary antibody (anti-H3K9me3 (ab 07-442)) and incubated overnight. The samples were incubated with the secondary antibody and with a DNA-FISH probe for the transgene (mCherry). The samples were mounted using mounting medium containing DAPI and imaged at 40x maginification. For both of the conditions, one representative young embryo is shown. Magenta is marking the H3K9me3, blue is DAPI staining of the DNA and green marks the FISH probe for the transgene. Control – transgenic *pdaf-21::mCherry* strain, *div-1(or148)* – *div-1* hypomorphic mutant worms.

Panel B: Quantification of the immunostaining with anti-H3K9me3. To count for the slide- and/or embryo-specific variation, H3K9me3 signal from each nucleus was normalized to the DAPI signal from the same nucleus. The graph shows H3K9me3 signal normalized to DAPI and two smaller graphs show the un-normalized signals for both H3K9me3 (left) and DAPI (right). Each dot represents one nucleus.

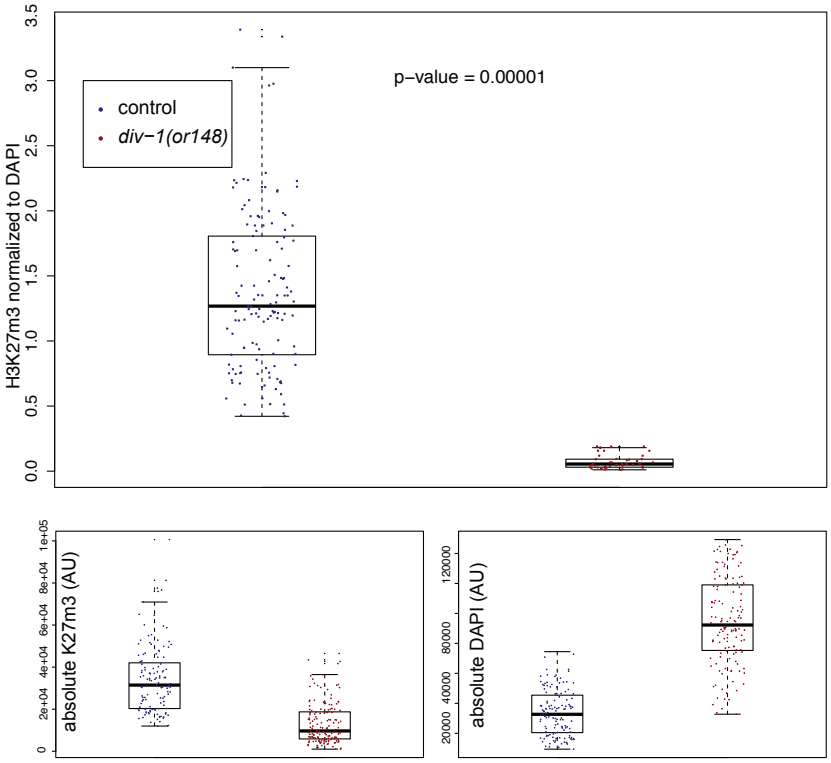
Panel C: Chromatin immunoprecipitation followed by qPCR analysis with gravid control and *div-1(or148)* worms. The samples were fixed in 2% formaldehyde and after sonication subjected to precipitation with the anti-

H3K9me3 (ab 07-442). Primer pairs for the qPCR represent the transgene (*mCherry*, *pdaf-21*), two control promoter regions (*ppmp-3* and *pcdc-42*) and two exonic primer pairs (*sra-25* and *cdc-42*). The values are presented as % input after normalization to the H3, each dot is one biological replicate. P-values (Student's t-test) for normalized % input: *pdaf-21* 0.08, *mCherry* 0.34, *ppmp-3* 0.047, *pcdc-42* 0.08, *cdc-42* 0.25, *sra-25* 0.07. P-values for non-normalized % input: *pdaf-21* 0.02, *mCherry* 0.063, *ppmp-3* 0.37, *pcdc-42* 0.115, *cdc-42* 0.53, *sra-25* 0.21. P-values for % input for H3: *pdaf-21* 0.37, *mCherry* 0.68, *ppmp-3* 0.89, *pcdc-42* 0.86, *cdc-42* 0.3, *sra-25* 0.92.

The levels of H3K27me3 were significantly reduced in the *div-1(or148)* worms, in most cases to almost undetectable levels (Figure 34, panel A, B) and this reduction was confirmed with ChIP-qPCR analysis (Figure 34, panel C).



(B)


H3K27m3 normalized to DAPI
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
control
div−1(or148)
p−value = 0.00001
absolute K27m3 (AU)
0
2e+04
4e+04
6e+04
8e+04
1e+05
absolute DAPI (AU)
20000
40000
80000
120000

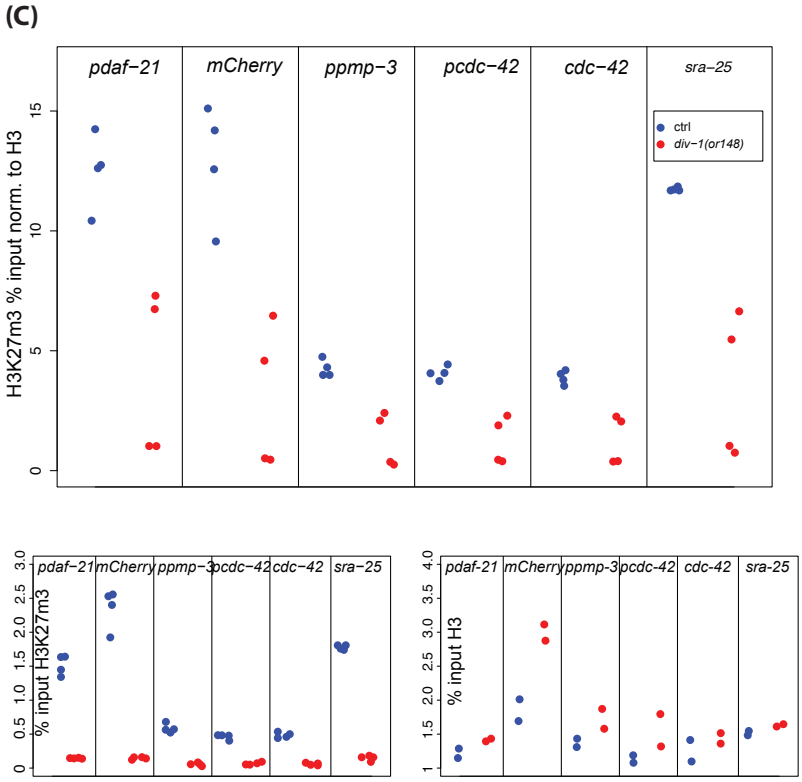


Figure 34 Immunostaining and ChIP-qPCR analysis of H3K27me3 revealed that this repressive mark was reduced in the *div-1(or148)* worms compared to control worms. All the Immunostaining and ChIP-qPCR experiments for the antibodies were carried out as specified for Figure 33.

Panel A: Immunostaining experiments done as in Figure 33, (with H3K27me3 antibody (ab 07-449)). Young gravid control and *div-1(or148)* worms were bleached and the extracted young populations of embryos were prepared with freeze-cracking method and subjected to methylation-specific primary antibody (anti-H3K27me3 (ab 07-449)) and incubated overnight. The samples were then incubated with the secondary antibody, followed by DNA-FISH probe for the transgene (mCherry). The samples were mounted using mounting medium containing DAPI and imaged at 40x magnification. For both of the conditions, one representative young embryo is shown. Magenta is marking the H3K27me3, blue is DAPI staining of the DNA and green marks the FISH probe for the transgene. Control – transgenic *pdaf-21::mCherry* strain, *div-1(or148)* – *div-1* hypomorphic mutant worms. The *div-1(or148)* worms appear highly devoid of the H3K27me3 signal.

Panel B: Quantification of the immunostaining with anti-H3K27me3. To count for the slide- and/or embryo variation, H3K27me3 signal from each embryo was normalized to the DAPI signal from the same nucleus. The graph shows H3K27me3 signal normalized to DAPI and two smaller graphs show

the un-normalized signals for both H3K27me3 (left) and DAPI (right). Each dot represents one nucleus.
Panel C: Chromatin immunoprecipitation as in Figure 33, with anti-H3K27me3 antibody. The values are presented as % input after normalization to the H3, each dot is one biological replicate. Primer pairs for the qPCR represent the transgene (*mCherry*, *pdaf-21*), two control promoter regions (*ppmp-3* and *pcdc-42*) and two exonic primer pairs (*sra-25* and *cdc-42*). p-values (Student's t-test) for normalized % input: *pdaf-21* 0.010, *mCherry* 0.0025, *ppmp-3* 0.0096, *pcdc-42* 0.013, *cdc-42* 0.011, *sra-25* 0.012. P-values for non-normalized % input: *pdaf-21* 0.0003, *mCherry* 0.0006, *ppmp-3* 0.0002, *pcdc-42* $2.75*10^{-5}$, *cdc-42* $4.3*10^{-5}$, *sra-25* $1.5*10^{-9}$. P-values for H3 % input: *pdaf-21* 0.2, *mCherry* 0.03, *ppmp-3* 0.21, *pcdc-42* 0.32, *cdc-42* 0.44, *sra-25* 0.12.

As a control for the ChIP efficiency, we checked for the enrichment of the H3K9me3 modification over the Tc1 and *gpd-1* sequences (Figure 35). *C. elegans* transposon Tc1 is enriched for the repressive mark H3K9me3, whereas the active *gpd-1*, a gene coding for Glyceraldehyde 3-Phosphate Dehydrogenase, is shown to be depleted from this mark [88].

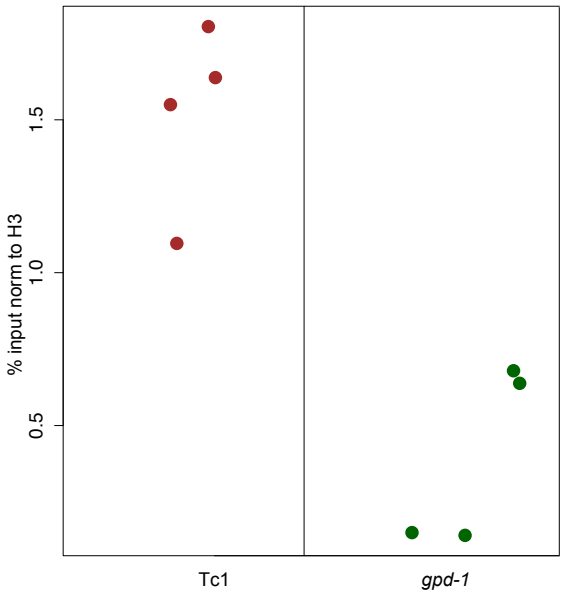
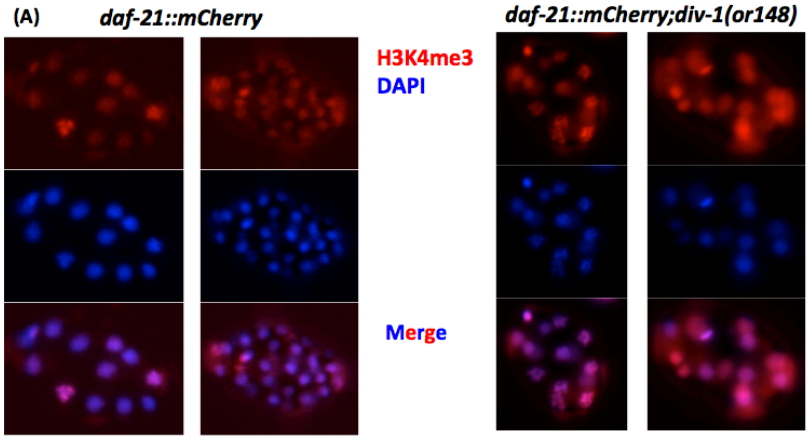
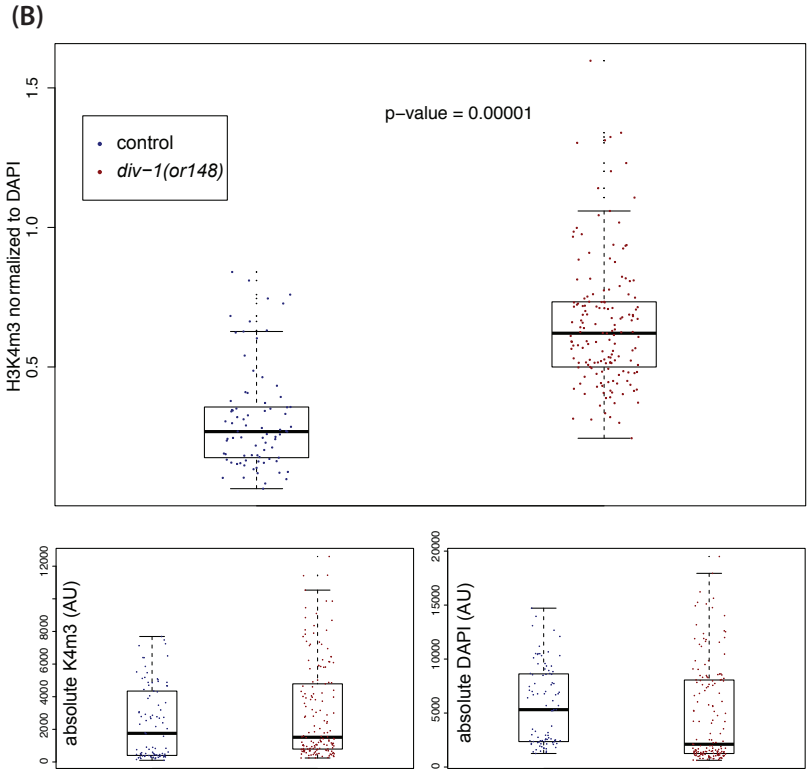


Figure 35 Control for the ChIP-qPCR specificity and efficiency using the H3K9me3 (07-442) antibody. The ChIP-qPCR was carried out as in Figure 33 with wild type gravid worms. The samples were fixed in 2% formaldehyde and after sonication subjected to precipitation with the anti-H3K9me3 (ab 07-442). Values are presented as % input after normalization to total H3. The *C. elegans* transposon Tc1 is shown to be enriched for the repressive mark H3K9me3 as opposed to the active region *pgpd-1* [88].

2.2.2 H3K4me3 and H3K36me3 are globally increased in the embryos with replication stress

We looked at the suggestive euchromatic marks, shown to correlate with active DNA, such as H3K4me3 and H3K36me3 [228-230]. As with the repressive marks (Figure 33 and 34), we observed changes of these modifications rather globally than just restricted to the transgene (Figure 36 and 37), whereas the H3K36me3 modification in embryos of the control worms was depleted in early embryos (Figure 37).


(A)
daf-21::mCherry
daf-21::mCherry;div-1(or148)
H3K4me3
DAPI
Merge


(B)
p−value = 0.00001
control
div−1(or148)
H3K4m3 normalized to DAPI
0.5
1.0
1.5
absolute K4m3 (AU)
0
2000
4000
6000
8000
12000
absolute DAPI (AU)
0
5000
10000
15000
20000

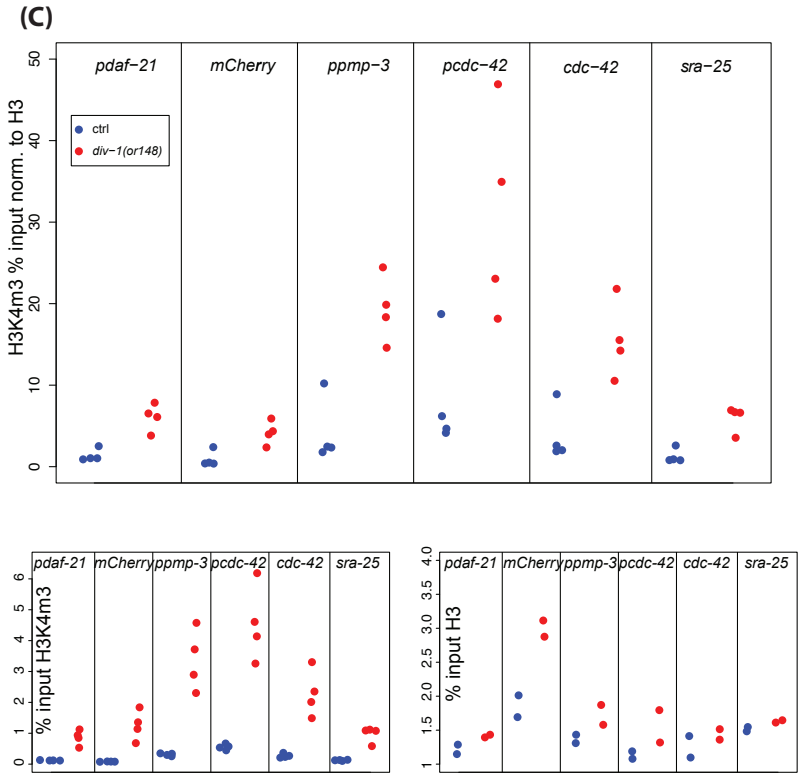
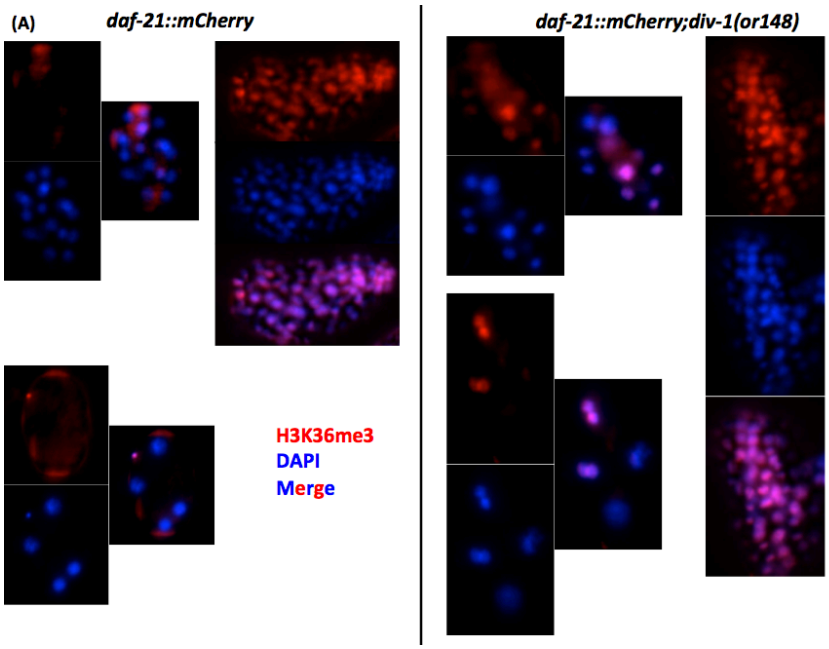


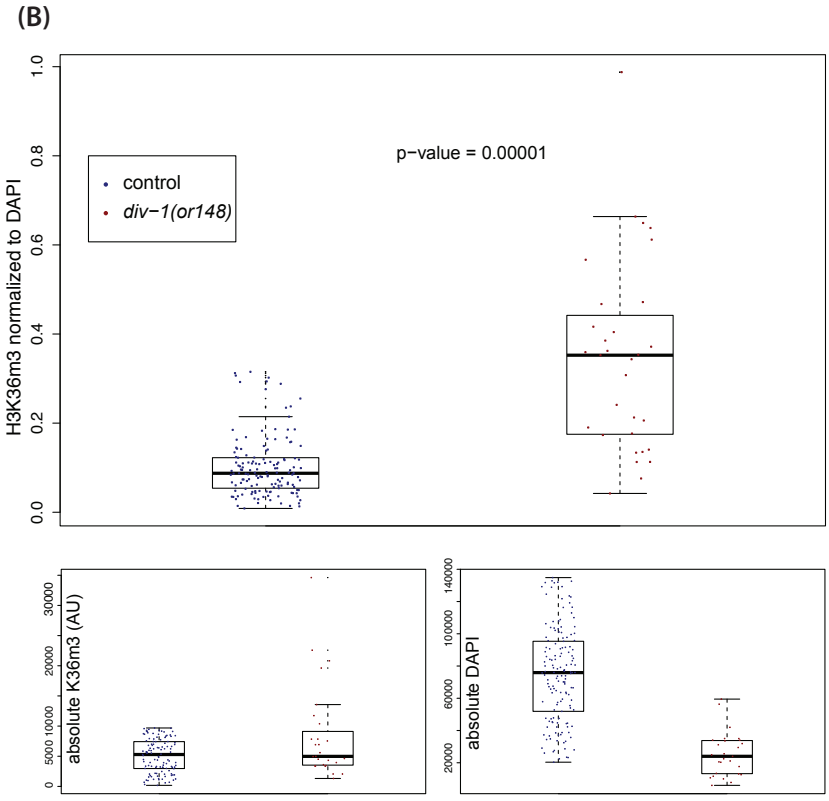
Figure 36 Immunostaining and ChIP-qPCR analysis of H3K4me3 revealed that this active mark was increased in the *div-1(or148)* worms compared to control worms. All the Immunostaining and ChIP-qPCR experiments for the antibodies were carried out as in Figure 33.

<u>Panel A</u>: Immunostaining experiments done as in Figure 33 (with H3K4me3 antibody (ab 8580)). Young gravid control and *div-1(or148)* worms were bleached and the extracted young populations of embryos were prepared with freeze-cracking method and subjected to methylation-specific primary antibody (anti-H3K4me3 (ab 8580)) and incubated overnight. The samples were then incubated with the secondary antibody followed by DNA-FISH probe for the transgene (mCherry). The samples were mounted with mounting medium containing DAPI. For both of the conditions, one representative young embryo is shown. Magenta is marking the H3K4me3, blue is DAPI staining of the DNA and green marks the FISH probe for the transgene. Control – transgenic *pdaf-21::mCherry* strain, *div-1(or148)* – *div-1* hypomorphic mutant worms. Images taken at 40x magnification.

<u>Panel B</u>: Quantification of the immunostaining with anti-H3K4me3. To count for the slide- and/or embryo variation, H3K4me3 signal from each embryo was normalized to the DAPI signal from the same nucleus. The graph shows H3K4me3 signal normalized to DAPI and two smaller graphs show the un-normalized signals for both H3K4me3 (left) and DAPI (right). Each dot represents one nucleus.

<u>Panel C</u>: Chromatin immunoprecipitation with H3K4me3 (ab 8580) as in Figure 33. The samples were fixed in 2% formaldehyde and after sonication subjected to precipitation with the anti-H3K4me3 (ab 8580) as previously for immunocytochemistry. Primer pairs for the qPCR represent the transgene (*mCherry*, *pdaf-21*), two control promoter regions (*ppmp-3* and *pcdc-42*) and two exonic primer pairs (*sra-25* and *cdc-42*). The values are presented as % input after normalization to the H3, each dot is one biological replicate. p-values (Student's t-test) for the normalized % input: *pdaf-21* 0.006, *mCherry* 0.013, *ppmp-3* 0.001, *pcdc-42* 0.03, *cdc-42* 0.008, *sra-25* 0.005. p-values for the non-normalized % input *pdaf-21* 0.009, *mCherry* 0.02, *ppmp-3* 0.008, *pcdc-42* 0.001, *cdc-42* 0.013, *sra-25* 0.007. P-values for H3 % input: *pdaf-21* 0.2, *mCherry* 0.03, *ppmp-3* 0.21, *pcdc-42* 0.32, *cdc-42* 0.44, *sra-25* 0.12.


(A)
daf-21::mCherry
daf-21::mCherry;div-1(or148)
H3K36me3
DAPI
Merge


(B)
H3K36m3 normalized to DAPI
0.0
0.2
0.4
0.6
0.8
1.0
control
div−1(or148)
p−value = 0.00001
absolute K36m3 (AU)
0
5000
10000
20000
30000
absolute DAPI
20000
60000
100000
140000

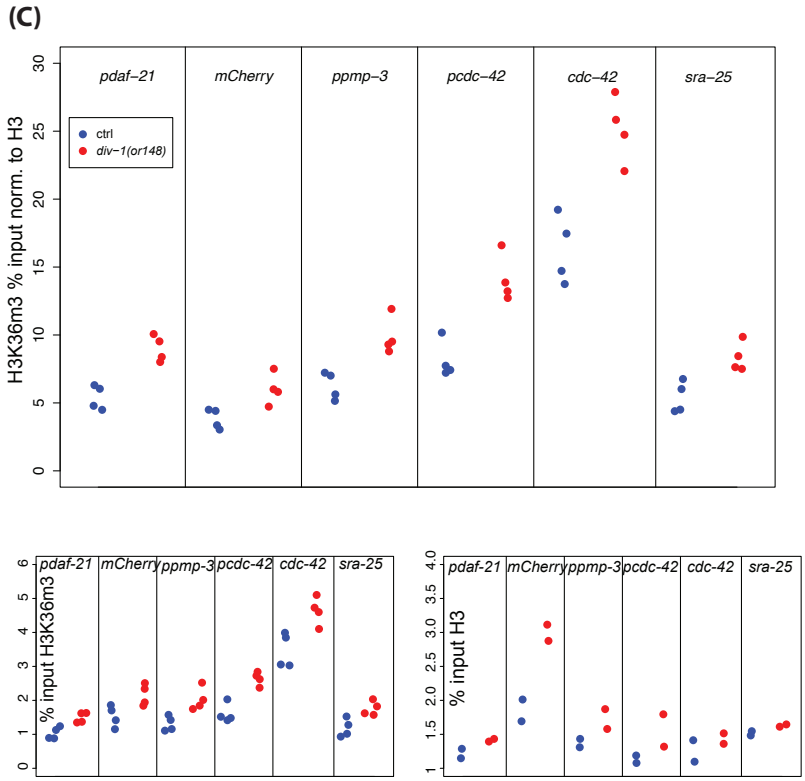


Figure 37 Immunostaining and ChIP-qPCR analysis of H3K36me3 revealed that this active mark was increased in the *div-1(or148)* worms compared to control worms. All the Immunostaining and ChIP-qPCR experiments for the antibodies were carried out as specified Figure 33.

<u>Panel A</u>: Immunostaining experiments done as in Figure 33, with the H3K36me3 (ab9050) antibody Young gravid control and *div-1(or148)* worms were bleached and the extracted young populations of embryos were prepared with freeze-cracking method and subjected to methylation-specific primary antibody (anti-H3K36me3 (ab 9050)) and incubated overnight. The samples were then incubated with the secondary antibody followed by DNA-FISH probe for the transgene (mCherry). The samples were mounted with mounting medium containing DAPI. For both of the conditions, one representative young embryo is shown. Magenta is marking the H3K36me3, blue is DAPI staining of the DNA and green marks the FISH probe for the transgene. Control – transgenic *pdaf-21::mCherry* strain, *div-1(or148)* – *div-1* hypomorphic mutant worms. Images taken at 40x magnification.

<u>Panel B</u>: Quantification of the immunostaining with anti-H3K36me3. To count for the slide- and/or embryo variation, H3K36me3 signal from each embryo was normalized to the DAPI signal from the same nucleus. The graph shows H3K36me3 signal normalized to DAPI and two smaller graphs show

the un-normalized signals for both H3K36me3 (left) and DAPI (right). Each dot represents one nucleus.

<u>Panel C</u>: Chromatin immunoprecipitation followed by qPCR with H3K36me3 (ab 9050) as in Figure 33. The samples were fixed in 2% formaldehyde and after sonication subjected to precipitation with the anti-H3K36me3 (ab 9050) as previously for immunocytochemistry. Primer pairs for the qPCR represent the transgene (*mCherry*, *pdaf-21*), two control promoter regions (*ppmp-3* and *pcdc-42*) and two exonic primer pairs (*sra-25* and *cdc-42*). p-values (Student's t-test) for the % input normalized to H3: *pdaf-21* 0.0016, *mCherry* 0.023, *ppmp-3* 0.007, *pcdc-42* 0.002, *cdc-42* 0.002, *sra-25* 0.010. P-values for the non-normalized % input: *pdaf-21* 0.008, *mCherry* 0.03, *ppmp-3* 0.017, *pcdc-42* 0.001, *cdc-42* 0.014, *sra-25* 0.016. P-values for % input for H3: *pdaf-21* 0.37, *mCherry* 0.68, *ppmp-3* 0.89, *pcdc-42* 0.86, *cdc-42* 0.3, *sra-25* 0.92.

3 - METHODS

3.1. Worm strains and culture conditions

All strains used are listed in the table with the specification about their genotype. N2 (Bristol) was used as a wild type strain and the transgenic strains (if not stated otherwise) are all derived from that strain. Worms were cultured in standard conditions [231], fed at 20C on NGM plates seeded with *E. coli* OP-50 except when stated otherwise for a particular experiment. Standard method to yield synchronized (offspring) generation, the worms were treated with hypochlorite solution [231] and after 3 washes with M9 + MgSO4, let over-night to hatch, resulting in a semi-synchronized L1 larval population. Strains used in this study:

strain	**genotype**
N2	Wild type
AU133	*agIs17 [myo-2p::mCherry + irg-1p::GFP] IV*
AU185	*agIs26 [myo-2::mCherry,clec-60::gfp]*
AY101	*acIs101[pDB09.1(pF35E12.5::gfp)*
BCN1050	*crgIs1002[daf-21p::mCherry::unc-54 3'UTR; unc-119(+)]*
BCN6101	*div-1(or148) III crgIs1002 IV*
BCN8011	drh-3(ne4253);;pdaf-21::mCherry
C4573	*cyp-35B::GFP*
CF1553	*muIs84[pAD76(sod-3p::GFP)]*
CL2166	*dvIs19[pAF15(gst-4p::GFP::NLS)]*
DW101	*atl-1(tm853) V/nT1 [unc-?(n754) let-? qIs50] (IV;V)*

EU548	*div-1(or148)*
IG274	*frIs7 [nlp-29p::GFP + col-12p::dsRed] IV*
JR667	*wIs51(SCM::GFP)*
MH1113	*dpy-20(e1282) IV; sur-5(ku74) X*
MH1870	*kuIs54[sur-5::gfp]*
MT17463	*set-25(n5021) III*
NL2507	*pkIs1582[let-858::GFP + rol-6(su1006)]*
PD4793	*mIs10[myo-2p::GFP;pes-10p::GFP]*
SJ4005	*zcIs4[hsp-4p::GFP] V*
SJ4103	*zcIs14[myo-3::GFP(mit)]*
TJ375	*gpIs1 [hsp-16.2p::GFP]*
WM206	*drh-3(ne4253) I*
WM27	*rde-1(ne219) V*

3.2. Strain construction and microscopy

The transgenic strains BCN1049 (*pdaf-21::GFP*) and BCN1050 (*pdaf-21::mCherry*) were generated in our laboratory by a previous PhD student Alejandro Burga [200], and the strain BCN6101 (*pdaf-21::mCherry;div-1(or148)*) was generated by Adam Klosin (personal communication). When the goal was to analyse the progeny after the cross, the crosses were carried out at 20C by picking about 1 male per hermaphrodite (L4 larvae). In case the F1 progeny was directly subjected to analysis of the transgene expression (as in case of the male-female *div-1(or148)* cross), the embryos were released the following day by dissecting the worms in 10mM levamisole. The

embryos were washed quickly with 10% hypochlorite solution and washed briefly with M9. Young embryos of 2-4 cell stage were collected with an eyelash and transferred with the mouth-pipette to the 96-well plates (Nunc, optical bottom) with 150ul of PBS. The embryos were carefully moved with an eyelash-pick to the center of the field view and imaged with Leica DMI6000 B microscope. If there was more than 2 samples, the delay in preparing the embryos was recorded and later added to the analysis pipeline with ImageJ to account for any additional time-delay. For the time-lapse either 10x or 20x objective was used. For time-lapse, the images were taken after every 10min for 17-19 hours (there is a time-delay with *div-1(or148)* embryonic development) in BrightField, Green (GFP) and dsR (mCHERRY) channel. The images were analysed with ImageJ where the embryos were first selected in the brightfield and the selection was transferred to the fluorescent images, from where the level of expression was measured. For each embryo, a background area was selected from close proximity that the intensity ('integrated intensity' in imageJ) was then normalized to. The same process was carried out for all of the time points, giving a transgene expression intensity curve that was visualized with R (version 2.15.3). All subsequent analysis was carried out with R.

3.3. RNAi screen 96-well liquid format

The screen was carried out in high throughput liquid feeding format in 96-well plates. For feeding, the library with transformed *E. coli* with about 15,000 *C. elegans* genes was used [199]. For the screen, a large number of embryos were harvested by bleaching and the worms were let hatch overnight in M9 to acquire large population of synchronized L1s. In the feeding plates, every well contained culture of one transformed *E. coli* clone, grown in 800ul of LB + Amp overnight at 37°C, 220rpm. To set up the feeding, the worms were counted and diluted in order to have 5 worms/ul and 10ul of worms were dispensed into each well to have 50-75 worms per well. 1 hour before adding the *E. coli* cultures to the wells, RNAi synthesis was induced by adding 4ul of 1M IPTG to the cultures and grown 1h at 37°C at 220rpm. The bacteria was pelleted at 2500g for 5 min and resuspended in 100ul NGM + Amp and IPTG. 40ul of resuspended bacteria was added to each well of 50-75 worms. The worms were let grow until most of the food was gone and worms were gravid with some L1 larvae around (88h-90h). Each 96-well plate included several wells of feeding with control RNAi to be used as a reference well for screening. The primary screen was carried out by eye with Leica DMI6000 B microscope with Lumen 200 metal arc lamp, observing the intensity of the transgene expression under the microscope with 10x magnification. The

secondary more stringent screen was carried out feeding the worms with triplicates with all the primary hits.

3.4. ChIP-qPCR analysis

Worms were grown for two generations on RNAi (to avoid starvation, 2x concentrated RNAi culture was used) or, in the case of mutants, worms were grown on regular OP50 feeding plates. Synchronized cultures were bleached, hatched overnight and up to 1000 worms per plate (60x15mm) were grown to obtain gravid adults (about 65-70 hours post-L1). The worms were collected in M9, washed 2-3 times to get rid of the remaining bacteria.

The samples were fixed in 1.5% formaldehyde (280ul of 37% formaldehyde in 7ml of M9, adding 3ml per sample) with mild shaking at RT for 30' and quenched by adding 3ml of 0.5M glycine (to final concentration of 0.25M) for 15min at RT. The samples were washed twice with M9 followed by final wash with FA buffer (50 mM HEPES/KOH pH 7.5, 1 mM EDTA, 1% Triton™ X-100, 0.1% sodium deoxycholate,150 mM NaCl) with Protease Inhibitors (Roche *cOmplete*, Mini EDTA-free). After aspirating the buffer, the remaining 300-400ul of worm pellet was frozen at -80C for longer or snap frozen and proceeded immediately with chromatin preparation.

To prepare for the ChIP, FA buffer with Protease Inhibitors was added to the samples until 0.9ml and distributed into 3 eppendorfs, up to 300ul each. The samples were sonicated using the Bioraptor® Sonication System Diagenode v1.1, at 4°C at intensity settings: High power, 30 sec on + 30 sec off. For gravid worms 20 cycles was applied, adding ice after 10 cycles if needed, followed by centrifugation for 25min at 4C, 13,000rpm. The three supernatants were pooled together and centrifuged again if another round of purification from the floating material was needed. Next, the generated fragment size was checked with 1% Agarose gel. For that, 20ul of supernatant was reverse crosslinked with 80ul of FA buffer (without Protease inhibitors), incubated at 65°C at 1100rpm for 3 hours. The fragments (100ul) were purified with QIAquick PCR Purification Kit, eluted in 30ul and 1/3 and 2/3 of the elute was run on the 1% Agarose gel to check the chromatin fragment size (preferably homogenous 200-600bp). If the fragment size was too big, the supernatant was re-added to the chromatin pellet and re-sonicated. To set up the ChIP pull-down with the primary antibody, the protein was quantified with Bradford method and 0.1mg of protein was added into the primary antibody reaction for histone modification ChIP in FA buffer (with Protease Inhibitors) up to 500ul with 2ul of each antibody. The antibodies used in the experiment were: H3 (Abcam ab1791), H3K9me3 (Millipore 07-442), H3K27me3 (Millipore 07-449), H3K4me3 (Abcam ab8580), H3K36me3

(Abcam ab9050). From each of the experiment, 1% of the volume was stored as an input control at -20°C just before adding the antibody. The 1st antibody reaction was rotated at 4°C over-night.

Next day the unblocked protein A beads were prepared (for rabbit antibodies, Diagenode Cat. No: kch-503-880) by adding 30ul of the beads with cut pipette tip and washing them with 500ul of FA buffer, centrifuging at 3000rpm for 3 minutes at 4°C and discarding the supernatant. The chromatin/1st-antibody mix was added to the beads, followed by 2 hour incubation on rotating wheel at 4°C. Following, the 2ndary antibody reaction was centrifuged for 3 minutes at 2000rpm 4°C and the supernatant was discarded. The beads were washed 3x with 1ml low salt buffer (50mM Hepes-KOH pH 7.5, 150mM NaCl, 1% Triton X100), inverting tubes 5 times. All centrifugations were carried out at 4°C 3000rpm for 3min. The samples were washed 1x with 1ml high salt buffer (50mM Hepes-KOH pH 7.5, 500mM NaCl, 1% Triton X100), inverting the tubes 5 times. After the last wash the beads were left dry and the samples were eluted from the beads along with the input samples (from -20°C). Elution buffer was always prepared fresh (1%SDS with 0,1M NaHCO3) and 100ul was added to the samples and the inputs, and incubated at 65°C for 3 hours at 1100rpm. After the incubation, the beads were centrifuged at 3000rpm and supernatant was collected and

purified with QIAquick PCR Purification Kit (#28104), and the samples were eluted in 200ul of PCR-grade water.
For qPCR reactions: 2ul of each of the sample was used and all the samples were run in duplicates. The reactions were run in the LightCycler 480 Multiwell Plate 384 (#04729749001), each well containing: 2ul of the sample, 5ul of the 2xLightCycler 480 SYBR Green I Master Mix (#04707526001) and 1uM of reverse and forward primer.

Primers used in qPCR analyses:
for the transgenic region:
pdaf-21 fwd: GCAGCATCTTCTTCGTCCTC,
rev GAAAAATTGAGGGCAGGTGA
mCherry_5' fwd: AAGGGCGAGGAGGATAACAT
rev: ACATGAACTGAGGGGACAGG
For control regions:
ppmp-3 (previously recommended as a control for qPCR analysis in *C. elegans* [232] as one of the most stably expressed gene),
ppmp-3 fwd: TGTTCACTCACAGCCAGCTC,
ppmp-3 rev: ACCATCCCATTCAAACCAAA.
cdc-42 (recommended as a control for qPCR analysis [232] as one of the most stable reference gene [233]),
pcdc-42 fwd: AGTTGTTTTGGCCATTTTGC,
pcdc-42 rev: TGAAAAACGAATTGCGAAACA.
cdc-42 3rd exonic region
cdc-42 fwd: GCCTGAAATTTCGCATCATT,

cdc-42 rev: TCCTTTGCCAACTTCTCTCC.

Primer pairs for controlling efficiency and specificity of the ChIP reaction ([88], Tc1 is shown to be enriched in H3K9me3 marks and have many repeats in the *C. elegans* genome, and *gpd-1* as a region for a low level of H3K9me3 [88]).

Tc1 fwd: AACCGTTAAGCATGGAGGTG

Tc1 rev: CACACGACGACGTTGAAACC.

pgpd-1 fwd GCGCAAGTTTCTGCTGTTTT

pgpd-1 rev CGGAAGATTCACAAGAAGCAA.

sra-25 as additional control region, selected randomly as a control for the mRNA-Seq:

sra-25 (3rd) exonic region fwd: ATCCCACTACAACCCAGGT,

rev: GACTACCGTGCGGAAATCAT.

rpl-12 fwd ACCCAAGACTGGAAGGGTCT,

rev GCCATCGATCTTGGTCTCAT

rrn-3.56 fwd GAACAGCGGGTTCAAACATT,

rev GATAGAGATGCCTCCCGACA

rrn-2.1 fwd CGTACTAGCTTCAGCGATGG,

rev ACCCTGAACCAGACGTACCA

3.4.1 qPCR normalization

From the qPCR analysis the Ct was first normalized to the input and with the delta-delta method [234] the percent input was calculated. For each antibody, this value was

then normalized to the % input value of H3 total histone. For each biological replicates two technical replicates were analysed.

3.5. Immunofluorescence

One OP50 feeding plate (60x15mm) of freshly gravid worms was bleached to acquire mostly young embryos. 10ul of worms were transferred to a poly-lysine coated slide and directly covered with 60x24mm coverslip. For freeze-cracking, the slide was introduced into liquid nitrogen for 10 seconds or the slides were frozen on metal blocks on dry ice, the coverslip was then removed with a swift movement. The slide was emerged into ice-cold MeOH for 5min, followed by emerging it into 1% paraformaldehyde for 2 minutes. The slides were then washed 3 times with PBS+0.25% Triton X-100 for 5 minutes and incubated in blocking solution (0.5% BSA in PBS with 0.25% Triton X-100) for 30min. 100ul of the first antibody (1:500 dilution in PBS+0.25 Triton X-100) was then added, covered with coverslip and the slides were incubated at 4°C overnight. The next day the slides were submerged into PBS+0.25 Triton for 15min, followed by additional 2 washes. The secondary antibody was then added (Alexa-555 anti-rabbit, Invitrogen), followed by 3 washes of PBS + 0.25% Triton X-100. The slides were then directly mounted with

Fluoroshield with DAPI mounting medium (Sigma). The slides were imaged using an oil immersion 40X objective on Leica DMI6000 B inverted microscope equipped with Hamamatsu Orca Flash 4.0 digital camera and a Lumen metal arc lamp (Prior Scientific).

The images were imaged with Leica DMI6000 B microscope, with 40x oil-immersion objective. The images were taken in Z-stack, and analyzed with the program ImageJ. The mean values for each nucleus from the background-corrected stacks were obtained and these values were used to estimate the changes between the intensity of the signals from the antibody-staining and DAPI.

3.6. Genomic DNA extraction

Several plates (60x15mm) of worms were grown with appropriate bacteria (either OP50 or RNAi bacteria on plates with IPTG and Amp). The worms were washed off the plates into M9, getting rid of as much bacteria as possible and pelleted at 1000rpm for 1min, followed by flash freezing the pellet in liquid nitrogen or dry ice with EtOH. 5 volumes of worm genomic DNA lysis buffer with Proteinase K (0.1mg/ml) was added and the samples were incubated at 65°C for 1-2 hours, followed by 20-30min deactivation of Proteinase K at 95°C. RNAse A (0.1mg/ml) was added to the samples and incubated at 37°C for 1h. In the fume

hood, 1 volume of phenol/chloroform was added and the samples were span at 4,000rpm for 5min. The aqueous phase was transferred to a new tube and another round of phenol/chloroform was added in case still significant amount of white precipitate was visible at the aqueous/organic interface. 0.1 volumes of 3M sodium acetate was added and 2 volumes of 100% ethanol. After 1 hour of incubation the DNA pellet was centrifuged at 14,000rpm for 15minutes, and the pellet is washed with 70% of ethanol. The pellet was air-dried and resuspended in water and the concentration was measured with Nanodrop.

3.7. Total RNA isolation and mRNA-Seq

To account for the developmental delay of the *div-1* RNAi worms compared to the worms fed with control RNAi, the time points were chosen as follows:

I – control RNAi: L1+34hs, *div-1* RNAi L1+38hs;

II – control RNAi: L1 + 38hs, *div-1* RNAi: L1 + 41hs,

III – control RNAi: L1 + 41hs, *div-1* RNAi L1 + 44hs,

IV – control RNAi: L1 + 44hs, *div-1* RNAi L1 + 47hs.

The worms are washed off from the plates and pelleted and cleaned from excessive bacteria. 100ul of supernatant (M9) was left to cover the worms, 400ul of Trizol (Invitrogen) was added and the samples were vortexed for 2min. The

samples were frozen, followed by 5 freeze-thaw cycles (between water bath at 37°C and dry ice/ethanol). 200ul of Trizol was added and after 5min of incubation, 140ul of chloroform was added. After 15 sec of vigorous shaking the samples were incubated for 2min, followed by 15min of centrifugation at 12,000g at 4°C and the aqueous phase was collected to a new eppendorf. An equal volume of 70% EtOH was added and the mixture was transferred to a RNeasy (Qiagen) spin column and the RNA was eluted with 30ul of RNase free water. The concentration was measured on the spectrophotometer and the samples were stored at -80C.

The mRNA sequencing libraries were prepared with the Illumina Sequencing Kit TruSeq v3 and were sequenced in paired end mode (read length of 50bp) with the Illumina HiSeq sequencer. Transcript abundance was estimated with the RSEM program [235], using Bowtie (default aligner used by RSEM) to align reads to the *C. elegans* genome (Wormbase WS220). Less than 1% of reads were mapping to ncRNA species. The reads were then normalized using the trimmed mean of M-values (TMM) method that estimates scale factors between samples [236]. The reads were quality controlled with the ArrayQualityMetrics package. A principal component analysis (PCA) was performed in all samples to try to visualize whether the condition (control vs. *div-1* RNAi) or the time points

influence the most how the samples relate to each other. As the samples clustered rather time-dependently (along PCA1 axis) than condition-dependently, the analysis was next focused on the time points rather than the experimental condition. The following analysis was carried out treating each sample as experiment without replicates. The differential gene expression was performed using DESeq (R/Bioconductor package). Additional annotation was obtained with the biomaRt package [237]. Repetitive elements were identified using the program RepeatMasker (www.repeatmasker.org), the gene and tandem repeat annotations were retrieved from Wormbase (WS220).

4 - Discussion

4.1. An RNAi screen to identify regulators for a highly variable *pdaf-21::mCherry* reporter

In our laboratory the phenotypic variation of isogenic organisms grown in the same environment has always been the question of heart. The worm *C. elegans* is a great model organism to study this variation, as the worm is a hermaphrodite, so can be studied in large isogenic populations. The population can be grown on the same feeding plate, reducing the impact of environmental factors in the study. We have been studying the variability of the expression of Hsp90, using an ubiquitously expressed chaperone with a transcriptional reporter for *daf-21*, the gene coding for HSP90 in *C. elegans*. Interestingly, worms grown in the same environment can express the reporter gene in highly variable levels. We were interested in understanding the premises for this variability and also in how the difference in the levels of chaperone account for the survivability and health of the worms. It has been previously shown that some chaperones, including HSP90 can act as buffers and influence the penetrance of mutations, as shown with Drosophila [196]. Our *pdaf-21::mCherry* reporter was potentially a great tool to study this variation, as the level of CHERRY expression in those worms was indeed very variable, even on the same feeding

plates in normal conditions. We were interested in the underlying mechanisms for this variation, so we designed a screen to study this. The rationale behind this was to knock down genes in *C. elegans* in high throughput and by observing the reporter expression, identify regulators of the chaperone. Interestingly, we instead identified a group of genes whose function was essential for the acquisition of the proper repressive chromatin state on the transgenic array containing our multi copy *pdaf-21::mCherry*. We observed that when the DNA replication machinery is inhibited, multiple chromatin marks are altered on the array, including the H3K27me3 repressive mark (Figure 38). These changes are likely responsible for the increased expression of the transgene.

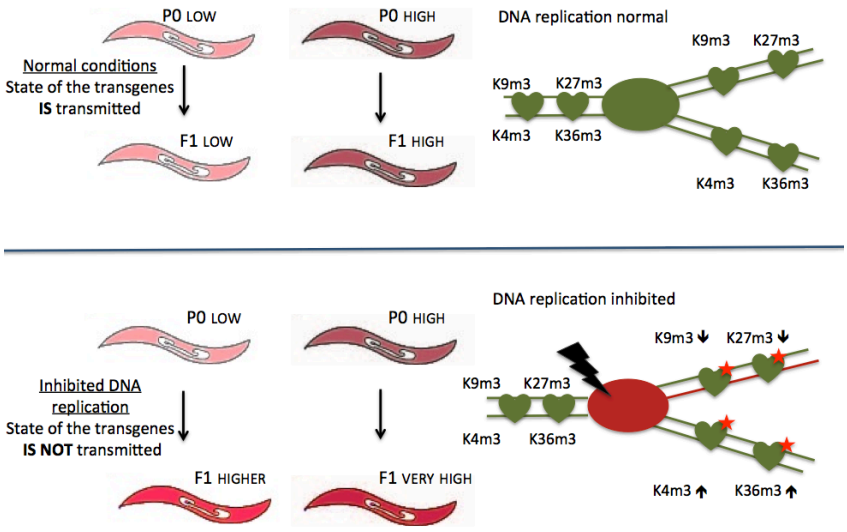


Figure 38 Model summarizing the hypothesis what we think occurs when the DNA replication machinery is inhibited by either RNAi or in appropriate mutant background. In normal condition the expression of the multi-copy *pdaf-21::mCherry* array is highly variable but faithfully transmitted to the progeny, 'low' P0 worms giving arise to 'low' F1 offspring and 'high' P0 to

'high' F1 (shown with different shades of red). The undisrupted DNA replication (green oval) gives rise to expected chromatin modifications (green hearts) showing histone methylations tested in this study. In case of the inhibited DNA replication machinery (red oval with thunderbolt), the expression level of the parents is not transmitted to the progeny and the expression is always increased, even in the case of 'High' worms, suggesting that in normal conditions the arrays are repressed and this repression is challenged when the replication machinery is dysfunctional. As shown with this study, after inhibition of the DNA replication, global levels of tested histone methylations are altered (red stars) and arrows showing either increase or decrease of the particular histone modification.

4.1.1. Components of the replication machinery are required for repression of the *pdaf-21::mCherry* transgene

The screen was carried out with a strain with transcriptional reporter for *daf-21* carrying the transgene in multiple copies. The screen enabled us to identify a number of genes that altered expression from the transcriptional reporter, when they were inhibited. While the inhibition of many genes were able to induce the expression of the transgene significantly, the highest and most consistent expression was seen with a subgroup of genes encoding subunits of the DNA replication machinery, further analysis was therefore focused on these genes. Though initially we were in search for regulators of the highly variable chaperone HSP90, we soon realized that we might have been studying the level of silencing of high copy number transgenes instead. It has been shown that transgenic worms can have the integrated DNA in tandem repetitive arrays and due to this nature, they are often organized into highly ordered heterochromatin-like

constructs that are silenced in the germline. Indeed, the qPCR analysis on the genomic DNA from the *pdaf-21::mCherry* worms showed that the transgenic DNA was present in high copy number, and possibly in re-arranged repetitive arrays as suggested by the copy number differences within the transgenic region with different primer sets (Figure 12).

The fact that the level of endogenous *daf-21* expression was not changed (Figure 18) supported our hypothesis that we had rather identified a set of genes required for acquisition of proper repressive chromatin state of the transgenic array. And as all these RNAi knock-downs led to an increase in the expression of the transgene, we hypothesized that these genes were required for proper silencing/repressive patterns on the transgene. Due to the fact that the *pdaf-21::mCherry* expression was highly variable and that the screen was carried out by eye, it was not possible to identify any negative regulators. One of the genes with the largest effect on array expression, *div-1* (a homolog of the B subunit of the DNA polymerase alpha-primase complex in *C. elegans*) was also confirmed using a mutant with a Ts mutant allele *or148* (Figure 15). Other hits included genes coding for ribosomal subunits and components of the proteasomal pathway, which can be accounted for by the function of the Hsp90 gene as a cellular chaperone that is induced in response to misfolded

proteins. Single genes without any emerging subgroup were discarded due to the scope of the study.

4.1.2 Inhibition of the replication machinery in the embryo is sufficient to interfere with the proper acquisition of the repressed chromatin state of the *pdaf-21::mCherry* multi copy array

We wanted to know where does the inhibition of replication interfere with the acquisition of the silenced state and derepression of the transgene. We show that the inhibition of *div-1* (from here on used as representative of the DNA replication subgroup from the RNAi screen hits) needs one generation to have an effect on the expression of the transgenic array (Figure 19-21), suggesting that it is the processes in the early embryo that establish the proper repressed state of the transgenic array. There is still possibility that the RNAi itself needs a generation to become effective, to sufficiently deplete the HSP90 protein levels. Interestingly, when we looked at the expression of the transgene in the RNAi-fed parental generation directly, we did not observe an increase in the expression compared to the control worms, except for a few posterior intestinal cells (Figure 19). It has been previously reported that some transgenic worms accumulate the reporter protein in the gut cells, observable as an accumulation of the fluorescent signal. Previous studies have suggested it to be similar to immune response [238]. In *C. elegans*, the 14 posterior-

most cells undergo nuclear division at the L1 stage and become binucleate, and all the total intestinal 20 cells undergo endoreduplication of the DNA at each larval stage, resulting in nuclei that are 32-ploid in the adult intestine [205]. Previous analysis of several *div-1* mutants [221] revealed that whereas homozygous *div-1* mutant hermaphrodites generally differentiate well, they frequently fail to produce intestinal and pharyngeal cells, reminding *skn-1* mutants that also lack pharynx and intestine and produce excess skin instead [239], leading these more actively dividing cells to increased expression of the transgenic array.

We then showed that when the transgene passes through the germline of *div-1* mutant worms, it becomes desilenced. We compared the induction of the transgene expression in different scenarios, introducing the naïve transgenic arrays coming from wild type worms into either *div-1* mutant female (non-selfed hermaphrodites) or male germline (Table 3 and Figures 20, 21). The transgene introduced via mutant male germline was weakly up-regulated compared to the wild type background, but when wild type males with the transgenic array were crossed with *div-1(or148)* mutant hermaphrodites, the transgene expression was increased much more significantly (Figure 21). This suggests that when the transgene was introduced via *div-1* mutant males, the mother's contribution was able to rescue the proper acquisition of the repressed state of the transgene by

providing the wild type DIV-1 protein to the oocyte. In contrast, when the 'female' (from the hermaphrodite) contribution was from the *div-1(or148)* mutant background, the male-contributed wild type *div-1* allele was not able to rescue the expression of DIV-1, suggesting that the establishment of the heterochromatic state of the transgene is established during the early embryonic development when the genome is still under maternal control.

For this experiment we took advantage of another transgene, a *myo-2::gfp* multi copy array to identify cross progeny and it is possible that the presence of multiple multi copy transgenic arrays could interfere with the regulation of each other's chromatic state for example by targeting the repetitive sequences by the pool RNAs produced in response to one or the other transgene. Repeating these crosses without using a transgene marker would test this possibility.

4.1.3 None of the common DNA damaging agents tested could induce upregulation of the transgenic array as much as *div-1* RNAi or the *div-1(or148)* mutant background

In most cells, the DNA damage checkpoint delays cell division when replication is stalled for example due to DNA damage. *C. elegans* early embryos are different though, as the checkpoint here only responds to developmental signals that are responsible for regulating the timing of cell divisions

[218, 222]. Here, any non-developmental input would disrupt the timing of the cell cycles and cause embryonic lethality. For that reason, a checkpoints response to DNA damage is non-active in embryos. The check-point mediated asynchrony in cell division is crucial for embryonic patterning in *C. elegans* and when reduced, the germline of these worms is not able to develop, and the survivors are usually sterile. At the same time, extending the asynchrony where the time difference between AB and P1 is further increased, has also detrimental effects, as shown with hypomorphic *div-1* mutants, where the asynchrony in cell division is extended resulting in mislocalization of developmental regulators, defects in embryonic patterning and lethality [239]. There has been a thorough analysis of cell-cycle timings in the *div-1* mutants, including the *div-1(or148)* allele that we have used throughout our study. In wild-type embryos the posterior blastomere P1 of the 2-cell stage is still in mitosis by the time the anterior blastomere AB finishes dividing, leaving only a very brief 3-cell stage. In *div-1* mutant embryos there is longer time interval between P1 and AB divisions, resulting in a prominent 3-cell stage [221]. Moreover, all the blastomeres exhibit longer divisions, in *div-1* mutant embryos during the first 3 cell cycles, mostly observed as an increase in the duration of the interphase, while mitosis appears relatively normal. As the cell cycles in early *C. elegans* embryos consist only of DNA replication and mitosis without any gap phases [240], [241], the delays

in cell divisions in *div-1* mutant embryos are consistent with the defects in DNA replication. In conclusion, though checkpoint activation is important for proper development, it can only happen in response to developmental signals and not in response to unscheduled events such as replication problems. A common source of replication problems is DNA damage and thus, our attempt was to replicate the increase in the expression of the transgenic array using DNA damaging agents available such as UV irradiation, MMS (methyl methanesulfonate) and EMS (ethyl methanesulfonate). Also, mutational inactivation of processes other than DNA replication, for example, nucleotide metabolism can result in the same delayed division phenotype. This hypothesis is supported by the fact that RNAi of both *rnr-1* and *rnr-2*, the genes encoding the subunits of the RNR – ribonucleotide reductase in *C. elegans*, an enzyme directly responsible for generating substrates for the replication machinery caused an increase in the expression of the multi copy transgenic array. Thus, another agent we tried was hydroxyurea (HU) as this agent is the direct inhibitor of the ribonucleotide reductase. Plus, with our *div-1* RNAi and *div-1(or148)* mutants we were able to observe similar cell cycle delays as reported with HU before [219] where the authors showed that HU treatment doubled the time between pronuclear migration and nuclear envelope breakdown (indication of mitosis initiation). Interestingly with none of the agents used to induce a DNA

replication defect or stalling of the replication machinery were we able to induce expression of the transgenic array as much as with *div-1* RNAi or with the *div-1(or148)* mutation (Figures 24-26). After treating the worms with EMS or UV irradiation, no significant induction was detected (Figures 26 and 25, respectively). We then confirmed that our chosen methods were able to efficiently induce DNA damage and/or cell cycle delay. The *div-1* knockdowns were able to cause delay in early embryonic development (Figure 15) and for the DNA damage effect we used a reporter strain for DNA damage, HUS-1::GFP. In normal conditions, the HUS-1 protein is diffusely distributed in the cell, but upon DNA damage accumulates into foci marking the locations for double stranded breaks [225]. HUS-1 is one of the checkpoint proteins, and member of the 9-1-1 complex (HUS-1/MRT-2/HPR-9) that is activated during DNA damage along with cell cycle arrest and apoptosis. The HUS-1 complex acts downstream of NER (nucleotide excision repair) to promote DNA-damage induced apoptosis [242]. This is different to our case, as the early embryonic cells with *div-1(-)* background are not undergoing apoptosis. Others [221] have also reported that *div-1* and other *div* mutants exhibit delayed embryonic cell divisions but appear otherwise normal.

Nevertheless we used the HUS-1::GFP strain as a reporter for DNA damage observable in the germline (Figure 27) and the DNA-damage-associated foci were readily observable in

div-1(or148), UV-C and HU treated worms (Figure 28). Interestingly, the *div-1* worms had foci count similar to wild type (0-6 foci per field view) whereas UV-C treated worms had up to 200 foci, while being a condition where no induction of the transgenic array could be seen (Figure 25). Thus, external treatments that cause extensive DNA damage in the germline and delayed cell division in the early embryo do not result in the induction of the transgene expression in the early embryo. In contrast, inhibition of the replication machinery or of ribonucleotide reductase causes only mild induction of the DNA damage response in the germline but upregulates the transgene in the early embryo and causes cell cycle delay. This suggests that the induction of the transgenic array may not be caused by replication fork stalling or replication stress *per se*. Instead, there may be a non-monotonic relationship between the extent of replication stress and induction of the transgene.

4.1.4 Not all multicopy arrays are de-repressed after inhibiting the DNA replication machinery

We were interested in understanding whether the de-repression of the transgenic tandem arrays is a general phenomenon coupled to DNA replication defects or if it is specific to some transgenes and dependent on the promoters driving the reporter genes. For that we looked at a small selection of transgenes, either: 1) driven by the

same *pdaf-21* reporter but in single copy, 2) other stress-responsive reporters (*hsp-4::gfp, gst-4::gfp*, *hsp-6::gfp*, *sod-3::gfp*, *cyp-35::gfp* and *nlp-29::gfp,)*, 3) other multi copy transgenic arrays (*plet-858::gfp, pscm::gfp* and *psur-5::sur-5::gfp)*. Interestingly, we did not see consistency amongst any of these selected subgroups (Table 4, Figures 22, 23). Some of the reporters (*hsp-4::gfp, gst::gfp* (Table 4), *plet-858::gfp* and *pscm::gfp, (*Figure 23)) did show increased expression after inhibition of the replication machinery, and though a single copy strain of *pdaf-21:gfp* reporter exhibited an increase in the expression of the transgene (Figure 22), it was about 10 times less than with multi copy transgenic array (1.8 fold increase compared to 10.5 fold). This suggests that some, but not all of the induction of the expression is linked to the reporter being present in the multi-copy array in the genome. As the *daf-21* gene codes for HSP90 chaperone in *C. elegans*, the reporter can respond to stress, including misfolded proteins [243]. The fact that some of the multi copy transgenes are more induced than others, may suggest that the location to where the transgene is inserted may play a role in the level of repression of the transgene. Common example of this is the PEV (position effect variegation), classical example from *Drosophila* known from 1930s, where a gene becomes repressed when it is placed close to pericentric heterochromatin [244]. Having reporter strains with the same construct inserted into different locations would test

this, enabling to distinguish between contributions of the promoter and the location. In a similar manner, the location of these transgenes can have an effect on the level of the expression depending on the surrounding chromatin.

4.2. The repressive chromosomal architecture associated with multicopy transgenic arrays changes after inhibiting the DNA replication machinery in *C. elegans* accompanied by global chromatin changes

Until recently [245], there was no evidence of *C. elegans* having any DNA modifications such as cytosine methylation to regulate its transcriptional activity and relies mostly on histone modifications to dictate the accessibility to DNA through its chromatin structure [246]. In *C. elegans*, the somatic cells and germline cells see the transgenes differently, especially the repetitive tandem arrays. Unlike somatic cells, germ cells are very efficient in silencing genes present in high copy number, an effective strategy for germline surveillance. In several cases, the transgene silencing can even affect the endogenous loci, a mechanism called co-suppression [247]. Interestingly in our case, there is no expression of the *pdaf-21::mCherry* in the germline cells (Figure 32), even in the *div-1(-)* background

where the transgene in somatic cells is highly de-repressed and expressed.

4.2.1 Interfering with DNA replication interferes with the acquisition of the proper chromatin marks in the early embryo and causes derepression of the multi copy transgenic arrays

Transcriptional control of the genome in eukaryotes involves interplay between different posttranslational modifications of the core histones. It has been shown that repetitive genetic arrays are subject to transcriptional repression [181] and accumulate repressive histone marks, especially H3K9me3 and H3K27me3 [248], [249]. Thus, the integrated repetitive transgenes are probably similar to endogenous heterochromatin in the germline [227]. The histone modifications are all part of a complex interplay and involved in various levels of crosstalk. In general, histone modifications regulate the transcriptional activity of genomic regions as suggested by apparent correlation between histone modifications and gene expression (summarized in [250]), H3K9 and H3K27 (mono-, bi- or tri-) methylations correlating with silenced DNA, or active DNA with H3K4 and H3K36 methylations. Based on these and other observations with transgenic arrays, we hypothesized that the mechanism through which inhibiting DNA replication machinery could increase the expression of our tandem array is through altering its chromatin structure and

compaction. We assayed the methylation levels of H3K9me3, H3K27me3, H3K4me3 and H3K36me3 over the transgene, modifications associated with repressive and active regions, respectively. H3K9me3 enrichment has been shown over silenced transgenes [158], accompanied with the absence of H3K4me3, and other active chromatin marks [50, 251]. Interestingly, rather than observing an increase or decrease in these chromatin marks locally, specifically over the transgene, the ChIP-qPCR analysis revealed similar changes at additional loci (Figure 33 (H3K9me3), Figure 34 (H3K27me3), Figure 36 (H3K4me3), Figure 37 (H3K36me3). That this is a global change in chromatin modifications across the genome was confirmed by immunofluorescence experiments which revealed that the decrease in the of H3K9me3 (Figure 33, panels A and B, though less significant that suggested by ChIP-qPCR, panel C) and H3K27me3 signals (Figure 34, panels A and B) and increase in H3K4me3 (Figure 36, panels A and B) and in H3K36me3 signals (Figure 37, panels A and B) was spread over the chromatin in the nucleus, stained with DAPI. The images acquired from the immunostaining experiment exhibited variable levels of signal intensities, thus for the analysis we normalized the mean signal intensity from the antibodies to that of the DAPI signal from the same nucleus and analyzed each experiment separately. Better control than DAPI-staining would be to use another antibody to count for the variability in the

permeability of the embryo to the antibodies introduced by the freeze-cracking method.

Previous studies have identified that nuclei from worms with high copy transgenic arrays exhibit regions of high occupancy of H3K9me3 [248], where they demonstrated that a multi-copy array *mIs10* FISH and high H3K9me3 immunofluorescence signals co-localize. In our study we could not specifically determine the location of the H3K9me3 mark in the nucleus, but it appears diffusely distributed within the nucleus rather than forming specific foci marking the transgene (Figure 33, panel A). Also, there was only slightly significant decrease in the total H3K9me3 levels in *div-1* mutant worms as measured by the anti-H3K9me3 antibody levels (Figure 33, panel A). Unfortunately, for both the ChIP-qPCR and the immunostaining experiments, the H3K9me3 antibodies used were giving the lowest signal, making it difficult to estimate the reliability of the change in enrichment in *div-1* mutants compared to control. The specificity and efficiency of the immunostaining with the anti- H3K9me3 antibody was validated with the *set-25* deficient worms (data not shown), a gene coding for the protein solely responsible for depositing the H3K9me3 mark in *C. elegans* [227]. As expected, there was no H3K9me3 signal observed in the *set-25(n5021)* background as compared to control worms. Bessler *et al.* also looked at the localization and enrichment of H3K27me3 over a transgene array, a mark that has been

shown to be enriched on X chromosomes in the germ cells of *C. elegans* hermaphrodites [34]. They reported no enrichment of H3K27me3 on the *mIs10* transgenic array, but rather a broad distribution on all chromosomes. Similarly, we saw no enrichment of H3K27me3 on the transgene (Figure 34, C). At the same time, this repressive modification H3K27me3 showed the most significant change in the *div-1* mutant animals (Figure 34).

From the active chromatin marks we looked at tri-methylations over H3K4 and H3K36, marks shown to be enriched in transcriptionally competent euchromatin, H3K4me3 as the hallmark of actively transcribed proteins and H3K36me enriched in gene bodies [17, 252]. Again, we observed a global increase in these active marks, rather than just a local change over the transgene, with H3K4me3 showing increased enrichment over promoter regions of the studied control sequences, *pmp-3* and *cdc-42* (Figure 36 (H3K4me3) and Figure 37 (H3K36me3)).

Thus, interfering with the DNA replication during the cell cycles of the early embryo, the levels of repressive marks H3K9me3 and H3K27me3 are decreased, whereas H3K4me3 and H3K36me3 levels are increased in the *div-1* mutant worms. Interestingly, the worms with perturbed DNA replication appear phenotypically quite normal and are able to cope well with the altered chromatin organization, as confirmed by phenotypically comparing the worms before and after the inhibition of DNA replication. Also the mRNA

analysis of these worms revealed no significant differential expression of endogenous genes (Figure 18).

Our results reveal a new interesting phenomenon where a perturbation of the DNA replication machinery causes a global re-organization of the chromatin, whereas the change is uni-directional, reducing the heterochromatic marks and increasing the euchromatic on histone H3 tails. Previously global changes in the histone methylation levels were shown in worms with depleted S-adenosylmethionine (SAM) synthetase, the methyl-group donor [227], causing a reduction of methylation levels in all studied trimethylation levels of K4, K9, K27 and K36 on histone H3 [227]. Reduction of heterochromatic marks has been shown before by inhibiting the enzymes specifically responsible for depositing these marks [39, 227, 248] whereas to our knowledge, no global systematic change has been reported in *C. elegans* before.

Saccharomyces cerevisiae adapts to environmental stresses by fast changes in global gene expression patterns [253]. Accompanying change of chromatin regulators in stress conditions in yeast reveals how chromatin plasticity is tightly regulated on a global level to mediate transcriptional response to stress [254]. The role of chromatin marks may depend on the cellular context, for example H3K4me3 and H3K36me3 show unexpected patterns in stress conditions, the latter generally involved in regulating downstream areas of genes can also be found at promoters under stress

conditions [255]. DNA damage-induced chromatin relaxation has been shown, such as heterochromatic loss in tau transgenic *Drosophila* in cells after DNA damage, suggesting that heterochromatically silenced genes are transcriptionally more sensitive to changes in the chromatin environment than are active genes. Role of Argonaute proteins in this regulation has been suggested in *S. pombe* for example, the RNA-induced transcriptional silencing (RITS) that associates with nascent transcripts and DNA-dependent RNA polymerase leads to heterochromatin formation [105]. Ago3, a homolog to the human PIWIL1, regulates the PIWI-associated RNAs (piRNAs) [102] is required for post-transcriptional silencing [99]. Ago3 levels are increased fourfold in tauopathy-associated chromatin relaxation in *Drosophila*, suggesting a regulative function for Ago3-mediated piRNA biogenesis in chromatin modifications [256].

Sarkies *et al*. [137] have previously discussed in a review how problems encountered by the replication fork might disturb the histone mark propagation. The histones deposited on newly synthesized DNA are also deposited during DNA replication, and nucleosomes ahead of the fork are removed for the DNA to unwind. This nucleosome displacement-replacement is tightly controlled and disrupting DNA replication interferes with histone recycling and leads to a loss of epigenetic information. For example, when the replicative helicase is uncoupled from the DNA

synthesis, long segments of single-stranded DNA are exposed, that is an inadequate substrate for nucleosome assembly [257]). Replication stress has shown to increase the number of histones with parental modifications, suggesting that histones were indeed displaced from the template prior to the fork but not replaced [258]. To further support the idea of requirement for coupling between the displacement of parental histones ahead of the replication fork and replacement on the newly synthesized DNA, Sarkies *et al*. [137] looked at histone modifications around G quadruplex (G4) DNA, sites in DNA that have capability to stall replication [259] in chicken DT40 cells lacking the translesion synthesis polymerase REV1 [137]. Replication forks in REV1(-) cells stall more often at sites of DNA damage and they show that G4 DNA was associated with loss of gene repression, including loss of H3K9me2 modification and accumulation of modifications of newly synthesized histones. So, the G4 DNA site can interfere with the fork progression, jeopardizing the maintenance of the chromatin through the loss of parental histone modifications and challenging the epigenetic silencing of the region [137].

Also Shachar, S. *et al*. [260] suggest a role for replication in determining the position of genome regions. They developed HIPMap (High-throughput imaging position mapping), an in situ hybridization-imaging pipeline to map spatial location of genome regions at large scale. With this

method they carried out an unbiased siRNA screen for factors involved in genome organization in human cells. They identify 50 factors (including chromatin modifiers, histone modifiers, nuclear envelope and pore proteins) that are required for appropriate positioning diverse genomic loci. Interestingly, many of the genome positioning factors identified included many DNA replication-associated proteins, such as post-replication histone chaperones. They then tested whether DNA replication is required for re-positioning in response to knockdown, and show that the S phase progression was required for the proper positioning by these identified factors, even if they were not directly involved in DNA replication, suggesting that the replication process, rather than the individual factors, determine gene positioning. As many of the replication-associated repositioning factors are involved in chromatin assembly, they speculate that the post-replication chromatin assembly is a critical for establishing and maintaining gene position.

4.2.2 Directly interfering with regulators of chromatin structure further derepresses the multi-copy transgenic array in *div-1(or148)* background

We looked at the expression of the transgenic array in *div-1(or148)* background after feeding them with RNAi bacteria for genes directly responsible for depositing the H3K9 and H3K36 methylation marks along with other genes that we

observed to increase the expression of the *pdaf-21::mCherry* transgenic array (Figure 29). The RNAi feeding of *lin-53*, *mes-4*, *mrg-1*, *mut-7*, *set-25* were all desilencing the transgenic array, and in the *div-1(or148)* the induction was further increased. This could suggest that the *div-1* caused inhibition of the DNA replication machinery does not directly act through these pathways. If the induction of the transgene caused by *div-1* mutation was caused merely by interfering with the pathways depositing H3K9me3 (*set-25* RNAi) or H3K36me3 (*mes-4* RNAi) marks on the chromatin, reducing the levels of these proteins would reverse the induction caused by *div-1* mutation. At the same time, caution has to be taken interpreting these results; as for example *nrde-2* itself is involved in the RNAi pathways, possibly affecting the efficiency of the RNAi knockdown.

4.2.3 The CSR-1 licensing pathway is able to partially suppress the induction of the multi copy transgenic array caused by inhibiting the DNA replication machinery

In *C. elegans*, the small RNA producing factory has an intriguing small RNAi pathway, the CSR-1 Argonaute pathway. Most of the other small RNAs have an interfering function [59, 80, 82, 87, 88, 96], where they silence the transcription of their targets, whereas CSR-1 is thought to act as a licensing pathway [98]. The small RNAs from this pathway bind their target sequences and protect these

normally endogenous germline transcripts from the vast silencing potential of the piRNAs (Piwi-interacting RNAs) in *C. elegans* germ cells. We showed (Figure 30) that the expression of the multi-copy *pdaf-21::mCherry* transgenic array is reduced in DRH-3 mutants, that is a Dicer-related helicase in *C. elegans*, a component of CSR-1 pathway, compared to the array in the wild type background. More interestingly, the *div-1(or148) drh-3(ne4253)* double mutant exhibits almost as low level of the transgene expression as does the wild type strain. This could suggest that by eliminating the CSR-1 pathway function, we are able to reverse the effect of the replication stress caused by *div-1* mutation. Using the *drh-3(ne4253)* strain is not optimal though as DRH-3 also acts in other 22G small RNA pathways. But eliminating any other essential component of the CSR-1 pathway leads to lethality, constraining us to the *drh-3*. What supports the hypothesis for the specificity for the CSR-1 pathway though, is the fact that no other small RNAi pathway mutants showed a decrease of the transgenic array (data not shown). To support the specificity of the *drh-3* to the CSR-1 pathway, we carried out the same genetic crosses with *rde-1* mutant. RDE-1 (RNAi Defective 1) is a primary Argonaute required for RNAi in *C. elegans*, being part of the exogenous RNAi pathway as opposed to *drh-3* and *csr-1.* In the *rde-1(ne219)* mutant background we did not observe reduced expression of the *pdaf-21::mCherry* array (Figure 31), nor was the *div-1(or148)*

induced derepression reduced in the *rde-1(ne219);div-1(or148)* double mutant. This could suggest that other small RNAi pathways acting through DRH-3 (and RDE-1) do not have an affect on the transgene expression as does the CSR-1 pathway. To further investigate the role of different pools of small RNAs in the regulation of the transgene expression and silencing, pools of small RNAs from control as well as *div-1(or148)* worms could be sequenced, at the L1 larval stage and from gravid adults to subject these to small RNA-Seq. Caution should be taken though when analyzing the small RNA pathways in the strain with *div-1(or148)* background. It has been previously shown that a soma-specific RNAi-defective allele of *mut-16 (mg461)* is present in the background of many *C. elegans* strains commonly used in the laboratory [261]. Strains with *mut-16(mg461)* allele might interfere with the endogenous siRNA pathways, as the strains with this background mutation have reduced response to RNAi pathways. As we discovered (A.Klosin, personal communication), the *div-1(or148)* also has this background mutation and its presence may confound some of the results from the RNAi pathway studies [58, 262].

4.3 Future directions

In addition to the directions discussed in the previous section, following questions remain open that were out of the scope of this study.

To better understand the global changes in the histone methylations after replication stress, ChIP-Seq analysis antibodies against these marks could be done. Immunostaining provides only a quantitative estimate of the decreases/increases in these modifications per whole genome and ChIP-qPCR is restricted to only a few chosen regions.

One of the most perplexing question throughout the project has been the specificity of the response of the *pdaf-21::mCherry* transgenic strain to the tested perturbations. More specifically, with the help of MosSCI [172, 174] generated transgenic worms could provide more information how does the location and copy number of a specific transgene affect the magnitude of the increase.

Additionally, change in the total histone levels and/or nucleosomal occupancy globally or over the transgene after inhibition of DNA replication machinery could be studied. In eukaryotes, next to the H3 that is one of the core histones, histone variant H3.3 is universal and is incorporated at sites of active transcription throughout the cell cycle [27, 263] and the H3.3 in chromatin is associated with histone turnover and it has been shown that the H3.3 patterns are similar to H3K4 methylation patterns [251]. The

repressive/active modifications are enriched on different histones of the H3/H4 tetrad of the nucleosome (activation marks enriched on the H3.3 rather than H3). Unfortunately, H3.3 difference to H3 is only few amino acids at the active site and cannot be detected with the anti-H3 and anti-H3.3 antibodies available [264]. Another possibility would be to measure the level of H4 and compare it to the change of H3 as both H3.3 and H3 make complexes with H4. Here, if the H4 level is constant and only the amount of H3 changes, we could estimate whether there is change in the ratio between H3.3 and H3. Also, we could hypothesize that the level of H3 reflects the overall nucleosomal occupancy, and this could be studied with further methods, for example with restriction enzymes whose efficiency depends on the packaging of the nucleosome, again giving us an estimate of the structure of the nucleosome. One could also measure the level of the Histone 3 with more standard methods, as Western blot, to test whether the change decrease is in cellular level or restricted to H3 associated with DNA.

As discussed previously (section 1.6.1.1), reduction of a p180 subunit of CAF-1 (chromatin assembly factor 1), in *Drosophila* suppresses heterochromatic gene silencing [142], accompanied by the decrease in H3K9 methylation marks at pericentric heterochromatin and reduced recruitment of HP1. Would then inhibiting the CAF-1-mediated nucleosome formation thus have profound effects on the expression of the repetitive transgenes?

5. References

1. Penny D: **Epigenetics, Darwin and Lamarck.** *Genome Biology and Evolution* 2015.
2. Rappaport R: **Hydrogeology. J. B. Lamarck. Translated from the French (Paris, 1802) by Albert V. Carozzi. University of Illinois Press, Urbana, 1964. viii + 152 pp. Illus. $4.75.** *Science* 1965, **148:**67-68.
3. Kelly WG: **Transgenerational epigenetics in the germline cycle of Caenorhabditis elegans.** *Epigenetics Chromatin* 2014, **7:**6.
4. Berger SL, Kouzarides T, Shiekhattar R, Shilatifard A: **An operational definition of epigenetics.** *Genes Dev* 2009, **23:**781-783.
5. Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ: **Crystal structure of the nucleosome core particle at 2.8 A resolution.** *Nature* 1997, **389:**251-260.
6. Bird AP: **CpG islands as gene markers in the vertebrate nucleus.** *Trends in Genetics*, **3:**342-347.
7. Rae PM, Steele RE: **Absence of cytosine methylation at C-C-G-G and G-C-G-C sites in the rDNA coding regions and intervening sequences of Drosophila and the rDNA of other insects.** *Nucleic Acids Res* 1979, **6:**2987-2995.
8. Cheung P, Allis CD, Sassone-Corsi P: **Signaling to chromatin through histone modifications.** *Cell* 2000, **103:**263-271.
9. Strahl BD, Allis CD: **The language of covalent histone modifications.** *Nature* 2000, **403:**41-45.
10. Turner BM: **Histone acetylation and an epigenetic code.** *Bioessays* 2000, **22:**836-845.
11. Jenuwein T, Allis CD: **Translating the histone code.** *Science* 2001, **293:**1074-1080.

12. Peterson CL, Laniel MA: **Histones and histone modifications.** *Curr Biol* 2004, **14:**R546-551.
13. Schneider R, Bannister AJ, Myers FA, Thorne AW, Crane-Robinson C, Kouzarides T: **Histone H3 lysine 4 methylation patterns in higher eukaryotic genes.** *Nat Cell Biol* 2004, **6:**73-77.
14. Peters AH, Kubicek S, Mechtler K, O'Sullivan RJ, Derijck AA, Perez-Burgos L, Kohlmaier A, Opravil S, Tachibana M, Shinkai Y, et al: **Partitioning and plasticity of repressive histone methylation states in mammalian chromatin.** *Mol Cell* 2003, **12:**1577-1589.
15. Rice JC, Briggs SD, Ueberheide B, Barber CM, Shabanowitz J, Hunt DF, Shinkai Y, Allis CD: **Histone methyltransferases direct different degrees of methylation to define distinct chromatin domains.** *Mol Cell* 2003, **12:**1591-1598.
16. Schotta G, Lachner M, Sarma K, Ebert A, Sengupta R, Reuter G, Reinberg D, Jenuwein T: **A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin.** *Genes Dev* 2004, **18:**1251-1262.
17. Bannister AJ, Zegerman P, Partridge JF, Miska EA, Thomas JO, Allshire RC, Kouzarides T: **Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain.** *Nature* 2001, **410:**120-124.
18. Lachner M, O'Carroll D, Rea S, Mechtler K, Jenuwein T: **Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins.** *Nature* 2001, **410:**116-120.
19. Grewal SI, Moazed D: **Heterochromatin and epigenetic control of gene expression.** *Science* 2003, **301:**798-802.
20. Dialynas GK, Makatsori D, Kourmouli N, Theodoropoulos PA, McLean K, Terjung S, Singh PB, Georgatos SD: **Methylation-independent binding to histone H3 and cell cycle-dependent incorporation of HP1beta into heterochromatin.** *J Biol Chem* 2006, **281:**14350-14360.

21. Maison C, Bailly D, Roche D, Montes de Oca R, Probst AV, Vassias I, Dingli F, Lombard B, Loew D, Quivy JP, Almouzni G: **SUMOylation promotes de novo targeting of HP1alpha to pericentric heterochromatin.** *Nat Genet* 2011, **43:**220-227.
22. Garrigues JM, Sidoli S, Garcia BA, Strome S: **Defining heterochromatin in C. elegans through genome-wide analysis of the heterochromatin protein 1 homolog HPL-2.** *Genome Res* 2015, **25:**76-88.
23. Ng HH, Robert F, Young RA, Struhl K: **Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity.** *Mol Cell* 2003, **11:**709-719.
24. Krogan NJ, Kim M, Tong A, Golshani A, Cagney G, Canadien V, Richards DP, Beattie BK, Emili A, Boone C, et al: **Methylation of histone H3 by Set2 in Saccharomyces cerevisiae is linked to transcriptional elongation by RNA polymerase II.** *Mol Cell Biol* 2003, **23:**4207-4218.
25. Li B, Howe L, Anderson S, Yates JR, 3rd, Workman JL: **The Set2 histone methyltransferase functions through the phosphorylated carboxyl-terminal domain of RNA polymerase II.** *J Biol Chem* 2003, **278:**8897-8903.
26. Akhmanova AS, Bindels PC, Xu J, Miedema K, Kremer H, Hennig W: **Structure and expression of histone H3.3 genes in Drosophila melanogaster and Drosophila hydei.** *Genome* 1995, **38:**586-600.
27. Ahmad K, Henikoff S: **The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly.** *Mol Cell* 2002, **9:**1191-1200.
28. Mito Y, Henikoff JG, Henikoff S: **Genome-scale profiling of histone H3.3 replacement patterns.** *Nat Genet* 2005, **37:**1090-1097.
29. Chow CM, Georgiou A, Szutorisz H, Maia e Silva A, Pombo A, Barahona I, Dargelos E, Canzonetta C, Dillon N: **Variant histone H3.3 marks promoters of**

transcriptionally active genes during mammalian cell division. *EMBO Rep* 2005, **6:**354-360.

30. Tagami H, Ray-Gallet D, Almouzni G, Nakatani Y: **Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis.** *Cell* 2004, **116:**51-61.
31. McKittrick E, Gafken PR, Ahmad K, Henikoff S: **Histone H3.3 is enriched in covalent modifications associated with active chromatin.** *Proc Natl Acad Sci U S A* 2004, **101:**1525-1530.
32. Hake SB, Garcia BA, Duncan EM, Kauer M, Dellaire G, Shabanowitz J, Bazett-Jones DP, Allis CD, Hunt DF: **Expression patterns and post-translational modifications associated with mammalian histone H3 variants.** *J Biol Chem* 2006, **281:**559-568.
33. Loppin B, Bonnefoy E, Anselme C, Laurencon A, Karr TL, Couble P: **The histone H3.3 chaperone HIRA is essential for chromatin assembly in the male pronucleus.** *Nature* 2005, **437:**1386-1390.
34. Bender LB, Cao R, Zhang Y, Strome S: **The MES-2/MES-3/MES-6 complex and regulation of histone H3 methylation in C. elegans.** *Curr Biol* 2004, **14:**1639-1643.
35. Bender AM, Wells O, Fay DS: **lin-35/Rb and xnp-1/ATR-X function redundantly to control somatic gonad development in C. elegans.** *Dev Biol* 2004, **273:**335-349.
36. Strahl BD, Grant PA, Briggs SD, Sun ZW, Bone JR, Caldwell JA, Mollah S, Cook RG, Shabanowitz J, Hunt DF, Allis CD: **Set2 is a nucleosomal histone H3-selective methyltransferase that mediates transcriptional repression.** *Mol Cell Biol* 2002, **22:**1298-1306.
37. Bender LB, Suh J, Carroll CR, Fong Y, Fingerman IM, Briggs SD, Cao R, Zhang Y, Reinke V, Strome S: **MES-4: an autosome-associated histone methyltransferase that participates in silencing the X chromosomes in the C. elegans germ line.** *Development* 2006, **133:**3907-3917.
38. Rechtsteiner A, Ercan S, Takasaki T, Phippen TM, Egelhofer TA, Wang W, Kimura H, Lieb JD, Strome S:

The histone H3K36 methyltransferase MES-4 acts epigenetically to transmit the memory of germline gene expression to progeny. *PLoS Genet* 2010, **6:**e1001091.

39. Andersen EC, Horvitz HR: **Two C. elegans histone methyltransferases repress lin-3 EGF transcription to inhibit vulval development.** *Development* 2007, **134:**2991-2999.
40. Furuhashi H, Takasaki T, Rechtsteiner A, Li T, Kimura H, Checchi PM, Strome S, Kelly WG: **Trans-generational epigenetic regulation of C. elegans primordial germ cells.** *Epigenetics Chromatin* 2010, **3:**15.
41. Gaydos LJ, Rechtsteiner A, Egelhofer TA, Carroll CR, Strome S: **Antagonism between MES-4 and Polycomb repressive complex 2 promotes appropriate gene expression in C. elegans germ cells.** *Cell Rep* 2012, **2:**1169-1177.
42. Shilatifard A: **The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis.** *Annu Rev Biochem* 2012, **81:**65-95.
43. Simonet T, Dulermo R, Schott S, Palladino F: **Antagonistic functions of SET-2/SET1 and HPL/HP1 proteins in C. elegans development.** *Dev Biol* 2007, **312:**367-383.
44. Li T, Kelly WG: **A role for Set1/MLL-related components in epigenetic regulation of the Caenorhabditis elegans germ line.** *PLoS Genet* 2011, **7:**e1001349.
45. Xiao Y, Bedet C, Robert VJ, Simonet T, Dunkelbarger S, Rakotomalala C, Soete G, Korswagen HC, Strome S, Palladino F: **Caenorhabditis elegans chromatin-associated proteins SET-2 and ASH-2 are differentially required for histone H3 Lys 4 methylation in embryos and adult germ cells.** *Proc Natl Acad Sci U S A* 2011, **108:**8305-8310.
46. Feng S, Jacobsen SE, Reik W: **Epigenetic reprogramming in plant and animal development.** *Science* 2010, **330:**622-627.

47. Ng RK, Gurdon JB: **Epigenetic memory of an active gene state depends on histone H3.3 incorporation into chromatin in the absence of transcription.** *Nat Cell Biol* 2008, **10:**102-109.
48. Arico JK, Katz DJ, van der Vlag J, Kelly WG: **Epigenetic patterns maintained in early Caenorhabditis elegans embryos can be established by gene activity in the parental germ cells.** *PLoS Genet* 2011, **7:**e1001391.
49. Reinke V, Gil IS, Ward S, Kazmer K: **Genome-wide germline-enriched and sex-biased expression profiles in Caenorhabditis elegans.** *Development* 2004, **131:**311-323.
50. Kelly WG, Schaner CE, Dernburg AF, Lee MH, Kim SK, Villeneuve AM, Reinke V: **X-chromosome silencing in the germline of C. elegans.** *Development* 2002, **129:**479-492.
51. Rechavi O, Minevich G, Hobert O: **Transgenerational inheritance of an acquired small RNA-based antiviral response in C. elegans.** *Cell* 2011, **147:**1248-1256.
52. Fire A, Albertson D, Harrison SW, Moerman DG: **Production of antisense RNA leads to effective and specific inhibition of gene expression in C. elegans muscle.** *Development* 1991, **113:**503-514.
53. Zamore PD, Tuschl T, Sharp PA, Bartel DP: **RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals.** *Cell* 2000, **101:**25-33.
54. Grishok A, Pasquinelli AE, Conte D, Li N, Parrish S, Ha I, Baillie DL, Fire A, Ruvkun G, Mello CC: **Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control C. elegans developmental timing.** *Cell* 2001, **106:**23-34.
55. Ketting RF, Fischer SE, Bernstein E, Sijen T, Hannon GJ, Plasterk RH: **Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans.** *Genes Dev* 2001, **15:**2654-2659.

56. Sijen T, Fleenor J, Simmer F, Thijssen KL, Parrish S, Timmons L, Plasterk RH, Fire A: **On the role of RNA amplification in dsRNA-triggered gene silencing.** *Cell* 2001, **107:**465-476.
57. Smardon A, Spoerke JM, Stacey SC, Klein ME, Mackin N, Maine EM: **EGO-1 is related to RNA-directed RNA polymerase and functions in germ-line development and RNA interference in C. elegans.** *Curr Biol* 2000, **10:**169-178.
58. Gu W, Shirayama M, Conte D, Jr., Vasale J, Batista PJ, Claycomb JM, Moresco JJ, Youngman EM, Keys J, Stoltz MJ, et al: **Distinct argonaute-mediated 22G-RNA pathways direct genome surveillance in the C. elegans germline.** *Mol Cell* 2009, **36:**231-244.
59. Buckley BA, Burkhart KB, Gu SG, Spracklin G, Kershner A, Fritz H, Kimble J, Fire A, Kennedy S: **A nuclear Argonaute promotes multigenerational epigenetic inheritance and germline immortality.** *Nature* 2012, **489:**447-451.
60. Zambon RA, Vakharia VN, Wu LP: **RNAi is an antiviral immune response against a dsRNA virus in Drosophila melanogaster.** *Cell Microbiol* 2006, **8:**880-889.
61. Wang XH, Aliyari R, Li WX, Li HW, Kim K, Carthew R, Atkinson P, Ding SW: **RNA interference directs innate immunity against viruses in adult Drosophila.** *Science* 2006, **312:**452-454.
62. Tabara H, Sarkissian M, Kelly WG, Fleenor J, Grishok A, Timmons L, Fire A, Mello CC: **The rde-1 gene, RNA interference, and transposon silencing in C. elegans.** *Cell* 1999, **99:**123-132.
63. Ketting RF, Haverkamp TH, van Luenen HG, Plasterk RH: **Mut-7 of C. elegans, required for transposon silencing and RNA interference, is a homolog of Werner syndrome helicase and RNaseD.** *Cell* 1999, **99:**133-141.
64. Ketting RF, Plasterk RH: **A genetic link between co-suppression and RNA interference in C. elegans.** *Nature* 2000, **404:**296-298.

65. Billi AC, Fischer SE, Kim JK: **Endogenous RNAi pathways in C. elegans.** *WormBook* 2014**:**1-49.
66. Fischer SE: **Small RNA-mediated gene silencing pathways in C. elegans.** *Int J Biochem Cell Biol* 2010, **42:**1306-1315.
67. Grishok A, Sinskey JL, Sharp PA: **Transcriptional silencing of a transgene by RNAi in the soma of C. elegans.** *Genes Dev* 2005, **19:**683-696.
68. Feng X, Guang S: **Small RNAs, RNAi and the inheritance of gene silencing in Caenorhabditis elegans.** *J Genet Genomics* 2013, **40:**153-160.
69. Claycomb JM: **Caenorhabditis elegans small RNA pathways make their mark on chromatin.** *DNA Cell Biol* 2012, **31 Suppl 1:**S17-33.
70. Zhuang JJ, Banse SA, Hunter CP: **The nuclear argonaute NRDE-3 contributes to transitive RNAi in Caenorhabditis elegans.** *Genetics* 2013, **194:**117-131.
71. Ruby JG, Jan C, Player C, Axtell MJ, Lee W, Nusbaum C, Ge H, Bartel DP: **Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in C. elegans.** *Cell* 2006, **127:**1193-1207.
72. Han T, Manoharan AP, Harkins TT, Bouffard P, Fitzpatrick C, Chu DS, Thierry-Mieg D, Thierry-Mieg J, Kim JK: **26G endo-siRNAs regulate spermatogenic and zygotic gene expression in Caenorhabditis elegans.** *Proc Natl Acad Sci U S A* 2009, **106:**18674-18679.
73. Conine CC, Batista PJ, Gu W, Claycomb JM, Chaves DA, Shirayama M, Mello CC: **Argonautes ALG-3 and ALG-4 are required for spermatogenesis-specific 26G-RNAs and thermotolerant sperm in Caenorhabditis elegans.** *Proc Natl Acad Sci U S A* 2010, **107:**3588-3593.
74. Vasale JJ, Gu W, Thivierge C, Batista PJ, Claycomb JM, Youngman EM, Duchaine TF, Mello CC, Conte D, Jr.: **Sequential rounds of RNA-dependent RNA transcription drive endogenous small-RNA biogenesis in the ERGO-1/Argonaute pathway.** *Proc Natl Acad Sci U S A* 2010, **107:**3582-3587.

75. Welker NC, Maity TS, Ye X, Aruscavage PJ, Krauchuk AA, Liu Q, Bass BL: **Dicer's helicase domain discriminates dsRNA termini to promote an altered reaction mode.** *Mol Cell* 2011, **41:**589-599.
76. Ambros V, Lee RC, Lavanway A, Williams PT, Jewell D: **MicroRNAs and other tiny endogenous RNAs in C. elegans.** *Curr Biol* 2003, **13:**807-818.
77. Guang S, Bochner AF, Pavelec DM, Burkhart KB, Harding S, Lachowiec J, Kennedy S: **An Argonaute transports siRNAs from the cytoplasm to the nucleus.** *Science* 2008, **321:**537-541.
78. Burkhart KB, Guang S, Buckley BA, Wong L, Bochner AF, Kennedy S: **A pre-mRNA-associating factor links endogenous siRNAs to chromatin regulation.** *PLoS Genet* 2011, **7:**e1002249.
79. Burton NO, Burkhart KB, Kennedy S: **Nuclear RNAi maintains heritable gene silencing in Caenorhabditis elegans.** *Proc Natl Acad Sci U S A* 2011, **108:**19683-19688.
80. Ashe A, Sapetschnig A, Weick EM, Mitchell J, Bagijn MP, Cording AC, Doebley AL, Goldstein LD, Lehrbach NJ, Le Pen J, et al: **piRNAs can trigger a multigenerational epigenetic memory in the germline of C. elegans.** *Cell* 2012, **150:**88-99.
81. Gu W, Lee HC, Chaves D, Youngman EM, Pazour GJ, Conte D, Jr., Mello CC: **CapSeq and CIP-TAP identify Pol II start sites and reveal capped small RNAs as C. elegans piRNA precursors.** *Cell* 2012, **151:**1488-1500.
82. Shirayama M, Seth M, Lee HC, Gu W, Ishidate T, Conte D, Jr., Mello CC: **piRNAs initiate an epigenetic memory of nonself RNA in the C. elegans germline.** *Cell* 2012, **150:**65-77.
83. de Wit E, Linsen SE, Cuppen E, Berezikov E: **Repertoire and evolution of miRNA genes in four divergent nematode species.** *Genome Res* 2009, **19:**2064-2074.
84. Batista PJ, Ruby JG, Claycomb JM, Chiang R, Fahlgren N, Kasschau KD, Chaves DA, Gu W, Vasale JJ, Duan S, et al: **PRG-1 and 21U-RNAs interact to form the piRNA**

complex required for fertility in C. elegans. *Mol Cell* 2008, **31:**67-78.

85. Das PP, Bagijn MP, Goldstein LD, Woolford JR, Lehrbach NJ, Sapetschnig A, Buhecha HR, Gilchrist MJ, Howe KL, Stark R, et al: **Piwi and piRNAs act upstream of an endogenous siRNA pathway to suppress Tc3 transposon mobility in the Caenorhabditis elegans germline.** *Mol Cell* 2008, **31:**79-90.
86. Stoeckius M, Maaskola J, Colombo T, Rahn HP, Friedlander MR, Li N, Chen W, Piano F, Rajewsky N: **Large-scale sorting of C. elegans embryos reveals the dynamics of small RNA expression.** *Nat Methods* 2009, **6:**745-751.
87. Bagijn MP, Goldstein LD, Sapetschnig A, Weick EM, Bouasker S, Lehrbach NJ, Simard MJ, Miska EA: **Function, targets, and evolution of Caenorhabditis elegans piRNAs.** *Science* 2012, **337:**574-578.
88. Luteijn MJ, van Bergeijk P, Kaaij LJ, Almeida MV, Roovers EF, Berezikov E, Ketting RF: **Extremely stable Piwi-induced gene silencing in Caenorhabditis elegans.** *EMBO J* 2012, **31:**3422-3430.
89. Billi AC, Freeberg MA, Day AM, Chun SY, Khivansara V, Kim JK: **A conserved upstream motif orchestrates autonomous, germline-enriched expression of Caenorhabditis elegans piRNAs.** *PLoS Genet* 2013, **9:**e1003392.
90. Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC: **Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans.** *Nature* 1998, **391:**806-811.
91. Grishok A, Tabara H, Mello CC: **Genetic requirements for inheritance of RNAi in C. elegans.** *Science* 2000, **287:**2494-2497.
92. Vastenhouw NL, Brunschwig K, Okihara KL, Muller F, Tijsterman M, Plasterk RH: **Gene expression: long-term gene silencing by RNAi.** *Nature* 2006, **442:**882.
93. Alcazar RM, Lin R, Fire AZ: **Transmission dynamics of heritable silencing induced by double-stranded**

RNA in Caenorhabditis elegans. *Genetics* 2008, **180:**1275-1288.

94. Seydoux G, Fire A: **Soma-germline asymmetry in the distributions of embryonic RNAs in Caenorhabditis elegans.** *Development* 1994, **120:**2823-2834.
95. Gu SG, Pak J, Guang S, Maniar JM, Kennedy S, Fire A: **Amplification of siRNA in Caenorhabditis elegans generates a transgenerational sequence-targeted histone H3 lysine 9 methylation footprint.** *Nat Genet* 2012, **44:**157-164.
96. Lee HC, Gu W, Shirayama M, Youngman E, Conte D, Jr., Mello CC: **C. elegans piRNAs mediate the genome-wide surveillance of germline transcripts.** *Cell* 2012, **150:**78-87.
97. Seth M, Shirayama M, Gu W, Ishidate T, Conte D, Jr., Mello CC: **The C. elegans CSR-1 argonaute pathway counteracts epigenetic silencing to promote germline gene expression.** *Dev Cell* 2013, **27:**656-663.
98. Wedeles CJ, Wu MZ, Claycomb JM: **Protection of germline gene expression by the C. elegans Argonaute CSR-1.** *Dev Cell* 2013, **27:**664-671.
99. Girard A, Sachidanandam R, Hannon GJ, Carmell MA: **A germline-specific class of small RNAs binds mammalian Piwi proteins.** *Nature* 2006, **442:**199-202.
100. Ghildiyal M, Zamore PD: **Small silencing RNAs: an expanding universe.** *Nat Rev Genet* 2009, **10:**94-108.
101. Tolia NH, Joshua-Tor L: **Slicer and the argonautes.** *Nat Chem Biol* 2007, **3:**36-43.
102. Brennecke J, Aravin AA, Stark A, Dus M, Kellis M, Sachidanandam R, Hannon GJ: **Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila.** *Cell* 2007, **128:**1089-1103.
103. Gunawardane LS, Saito K, Nishida KM, Miyoshi K, Kawamura Y, Nagami T, Siomi H, Siomi MC: **A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in Drosophila.** *Science* 2007, **315:**1587-1590.

104. Li C, Vagin VV, Lee S, Xu J, Ma S, Xi H, Seitz H, Horwich MD, Syrzycka M, Honda BM, et al: **Collapse of germline piRNAs in the absence of Argonaute3 reveals somatic piRNAs in flies.** *Cell* 2009, **137:**509-521.
105. Moazed D: **Small RNAs in transcriptional gene silencing and genome defence.** *Nature* 2009, **457:**413-420.
106. Beisel C, Paro R: **Silencing chromatin: comparing modes and mechanisms.** *Nat Rev Genet* 2011, **12:**123-135.
107. Holoch D, Moazed D: **RNA-mediated epigenetic regulation of gene expression.** *Nat Rev Genet* 2015, **16:**71-84.
108. Johnson CL, Spence AM: **Epigenetic licensing of germline gene expression by maternal RNA in C. elegans.** *Science* 2011, **333:**1311-1314.
109. Blander G, Guarente L: **The Sir2 family of protein deacetylases.** *Annu Rev Biochem* 2004, **73:**417-435.
110. Dang W, Steffen KK, Perry R, Dorsey JA, Johnson FB, Shilatifard A, Kaeberlein M, Kennedy BK, Berger SL: **Histone H4 lysine 16 acetylation regulates cellular lifespan.** *Nature* 2009, **459:**802-807.
111. Wysocka J, Myers MP, Laherty CD, Eisenman RN, Herr W: **Human Sin3 deacetylase and trithorax-related Set1/Ash2 histone H3-K4 methyltransferase are tethered together selectively by the cell-proliferation factor HCF-1.** *Genes Dev* 2003, **17:**896-911.
112. Schaner CE, Deshpande G, Schedl PD, Kelly WG: **A conserved chromatin architecture marks and maintains the restricted germ cell lineage in worms and flies.** *Dev Cell* 2003, **5:**747-757.
113. Katz DJ, Edwards TM, Reinke V, Kelly WG: **A C. elegans LSD1 demethylase contributes to germline immortality by reprogramming epigenetic memory.** *Cell* 2009, **137:**308-320.
114. Moazed D: **Mechanisms for the inheritance of chromatin states.** *Cell* 2011, **146:**510-518.

115. Yigit E, Batista PJ, Bei Y, Pang KM, Chen CC, Tolia NH, Joshua-Tor L, Mitani S, Simard MJ, Mello CC: **Analysis of the C. elegans Argonaute family reveals that distinct Argonautes act sequentially during RNAi.** *Cell* 2006, **127:**747-757.
116. Greer EL, Beese-Sims SE, Brookes E, Spadafora R, Zhu Y, Rothbart SB, Aristizabal-Corrales D, Chen S, Badeaux AI, Jin Q, et al: **A histone methylation network regulates transgenerational epigenetic memory in C. elegans.** *Cell Rep* 2014, **7:**113-126.
117. Simon M, Sarkies P, Ikegami K, Doebley AL, Goldstein LD, Mitchell J, Sakaguchi A, Miska EA, Ahmed S: **Reduced insulin/IGF-1 signaling restores germ cell immortality to Caenorhabditis elegans Piwi mutants.** *Cell Rep* 2014, **7:**762-773.
118. Pidoux AL, Allshire RC: **Kinetochore and heterochromatin domains of the fission yeast centromere.** *Chromosome Res* 2004, **12:**521-534.
119. Ekwall K, Olsson T, Turner BM, Cranston G, Allshire RC: **Transient inhibition of histone deacetylation alters the structural and functional imprint at fission yeast centromeres.** *Cell* 1997, **91:**1021-1032.
120. Volpe TA, Kidner C, Hall IM, Teng G, Grewal SI, Martienssen RA: **Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi.** *Science* 2002, **297:**1833-1837.
121. Halfmann R, Jarosz DF, Jones SK, Chang A, Lancaster AK, Lindquist S: **Prions are a common mechanism for phenotypic inheritance in wild yeasts.** *Nature* 2012, **482:**363-368.
122. Ladurner AG: **Rheostat control of gene expression by metabolites.** *Mol Cell* 2006, **24:**1-11.
123. Teperino R, Schoonjans K, Auwerx J: **Histone methyl transferases and demethylases; can they link metabolism and transcription?** *Cell Metab* 2010, **12:**321-327.
124. Ptashne M: **On the use of the word 'epigenetic'.** *Curr Biol* 2007, **17:**R233-236.
125. Zordan RE, Galgoczy DJ, Johnson AD: **Epigenetic properties of white-opaque switching in Candida**

albicans are based on a self-sustaining transcriptional feedback loop. *Proc Natl Acad Sci U S A* 2006, **103:**12807-12812.

126. Groth A, Corpet A, Cook AJ, Roche D, Bartek J, Lukas J, Almouzni G: **Regulation of replication fork progression through histone supply and demand.** *Science* 2007, **318:**1928-1931.
127. Turner BM: **Defining an epigenetic code.** *Nat Cell Biol* 2007, **9:**2-6.
128. Ringrose L, Paro R: **Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins.** *Annu Rev Genet* 2004, **38:**413-443.
129. Nakayama J, Rice JC, Strahl BD, Allis CD, Grewal SI: **Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly.** *Science* 2001, **292:**110-113.
130. Haber JE: **Mating-type gene switching in Saccharomyces cerevisiae.** *Annu Rev Genet* 1998, **32:**561-599.
131. Klar AJ: **Lessons learned from studies of fission yeast mating-type switching and silencing.** *Annu Rev Genet* 2007, **41:**213-236.
132. Lorentz A, Heim L, Schmidt H: **The switching gene swi6 affects recombination and gene expression in the mating-type region of Schizosaccharomyces pombe.** *Molecular and General Genetics MGG*, **233:**436-442.
133. Pillus L, Rine J: **Epigenetic inheritance of transcriptional states in S. cerevisiae.** *Cell* 1989, **59:**637-647.
134. Rusche LN, Kirchmaier AL, Rine J: **The establishment, inheritance, and function of silenced chromatin in Saccharomyces cerevisiae.** *Annu Rev Biochem* 2003, **72:**481-516.
135. Stillman B: **Chromatin assembly during SV40 DNA replication in vitro.** *Cell* 1986, **45:**555-565.
136. Ye X, Franco AA, Santos H, Nelson DM, Kaufman PD, Adams PD: **Defective S phase chromatin assembly causes DNA damage, activation of the S phase**

checkpoint, and S phase arrest. *Mol Cell* 2003, **11:**341-351.

137. Sarkies P, Reams C, Simpson LJ, Sale JE: **Epigenetic instability due to defective replication of structured DNA.** *Mol Cell* 2010, **40:**703-713.
138. Smith S, Stillman B: **Purification and characterization of CAF-I, a human cell factor required for chromatin assembly during DNA replication in vitro.** *Cell* 1989, **58:**15-25.
139. Monson EK, de Bruin D, Zakian VA: **The yeast Cac1 protein is required for the stable inheritance of transcriptionally repressed chromatin at telomeres.** *Proc Natl Acad Sci U S A* 1997, **94:**13081-13086.
140. Ono T, Kaya H, Takeda S, Abe M, Ogawa Y, Kato M, Kakutani T, Mittelsten Scheid O, Araki T, Shibahara K: **Chromatin assembly factor 1 ensures the stable maintenance of silent chromatin states in Arabidopsis.** *Genes Cells* 2006, **11:**153-162.
141. Houlard M, Berlivet S, Probst AV, Quivy JP, Hery P, Almouzni G, Gerard M: **CAF-1 is essential for heterochromatin organization in pluripotent embryonic cells.** *PLoS Genet* 2006, **2:**e181.
142. Huang H, Yu Z, Zhang S, Liang X, Chen J, Li C, Ma J, Jiao R: **Drosophila CAF-1 regulates HP1-mediated epigenetic silencing and pericentric heterochromatin stability.** *J Cell Sci* 2010, **123:**2853-2861.
143. Huang S, Zhou H, Tarara J, Zhang Z: **A novel role for histone chaperones CAF-1 and Rtt106p in heterochromatin silencing.** *EMBO J* 2007, **26:**2274-2283.
144. Loyola A, Tagami H, Bonaldi T, Roche D, Quivy JP, Imhof A, Nakatani Y, Dent SY, Almouzni G: **The HP1alpha-CAF1-SetDB1-containing complex provides H3K9me1 for Suv39-mediated K9me3 in pericentric heterochromatin.** *EMBO Rep* 2009, **10:**769-775.
145. Singer MS, Kahana A, Wolf AJ, Meisinger LL, Peterson SE, Goggin C, Mahowald M, Gottschling DE:

Identification of high-copy disruptors of telomeric silencing in Saccharomyces cerevisiae. *Genetics* 1998, **150:**613-632.

146. Yamane K, Mizuguchi T, Cui B, Zofall M, Noma K, Grewal SI: **Asf1/HIRA facilitate global histone deacetylation and associate with HP1 to promote nucleosome occupancy at heterochromatic loci.** *Mol Cell* 2011, **41:**56-66.
147. Sugiyama T, Cam HP, Sugiyama R, Noma K, Zofall M, Kobayashi R, Grewal SI: **SHREC, an effector complex for heterochromatic transcriptional silencing.** *Cell* 2007, **128:**491-504.
148. Jones PA, Baylin SB: **The epigenomics of cancer.** *Cell* 2007, **128:**683-692.
149. Lin JC, Jeong S, Liang G, Takai D, Fatemi M, Tsai YC, Egger G, Gal-Yam EN, Jones PA: **Role of nucleosomal occupancy in the epigenetic silencing of the MLH1 CpG island.** *Cancer Cell* 2007, **12:**432-444.
150. Brenner S: **The genetics of Caenorhabditis elegans.** *Genetics* 1974, **77:**71-94.
151. Byerly L, Cassada RC, Russell RL: **The life cycle of the nematode Caenorhabditis elegans. I. Wild-type growth and reproduction.** *Dev Biol* 1976, **51:**23-33.
152. **Genome sequence of the nematode C. elegans: a platform for investigating biology.** *Science* 1998, **282:**2012-2018.
153. Nicholas WL, Dougherty EC, Hansen EL: **AXENIC CULTIVATION OF CAENORHARDITIS BRIGGSAE (NEMATODA: RHABDITIDAE) WITH CHEMICALLY UNDEFINED SUPPLEMENTS; COMPARATIVE STUDIES WITH RELATED NEMATODES*.** *Annals of the New York Academy of Sciences* 1959, **77:**218-236.
154. Stinchcomb DT, Shaw JE, Carr SH, Hirsh D: **Extrachromosomal DNA transformation of Caenorhabditis elegans.** *Mol Cell Biol* 1985, **5:**3484-3496.
155. Okkema PG, Harrison SW, Plunger V, Aryana A, Fire A: **Sequence requirements for myosin gene expression and regulation in Caenorhabditis elegans.** *Genetics* 1993, **135:**385-404.

156. Fire A, Harrison SW, Dixon D: **A modular set of lacZ fusion vectors for studying gene expression in Caenorhabditis elegans.** *Gene* 1990, **93:**189-198.
157. Hope IA: **'Promoter trapping' in Caenorhabditis elegans.** *Development* 1991, **113:**399-408.
158. Kelly WG, Xu S, Montgomery MK, Fire A: **Distinct requirements for somatic and germline expression of a generally expressed Caernorhabditis elegans gene.** *Genetics* 1997, **146:**227-238.
159. Kelly WG, Fire A: **Chromatin silencing and the maintenance of a functional germline in Caenorhabditis elegans.** *Development* 1998, **125:**2451-2456.
160. Jedrusik MA, Schulze E: **A single histone H1 isoform (H1.1) is essential for chromatin silencing and germline development in Caenorhabditis elegans.** *Development* 2001, **128:**1069-1080.
161. Jedrusik MA, Schulze E: **Telomeric position effect variegation in Saccharomyces cerevisiae by Caenorhabditis elegans linker histones suggests a mechanistic connection between germ line and telomeric silencing.** *Mol Cell Biol* 2003, **23:**3681-3691.
162. Couteau F, Guerry F, Muller F, Palladino F: **A heterochromatin protein 1 homologue in Caenorhabditis elegans acts in germline and vulval development.** *EMBO Rep* 2002, **3:**235-241.
163. Kennison JA: **The Polycomb and trithorax group proteins of Drosophila: trans-regulators of homeotic gene function.** *Annu Rev Genet* 1995, **29:**289-303.
164. Jones DO, Cowell IG, Singh PB: **Mammalian chromodomain proteins: their role in genome organisation and expression.** *Bioessays* 2000, **22:**124-137.
165. Strome S, Powers J, Dunn M, Reese K, Malone CJ, White J, Seydoux G, Saxton W: **Spindle dynamics and the role of gamma-tubulin in early Caenorhabditis elegans embryos.** *Mol Biol Cell* 2001, **12:**1751-1764.

166. Jones AR, Schedl T: **Mutations in gld-1, a female germ cell-specific tumor suppressor gene in Caenorhabditis elegans, affect a conserved domain also found in Src-associated protein Sam68.** *Genes Dev* 1995, **9:**1491-1504.
167. Gaudet J, VanderElst I, Spence AM: **Post-transcriptional regulation of sex determination in Caenorhabditis elegans: widespread expression of the sex-determining gene fem-1 in both sexes.** *Mol Biol Cell* 1996, **7:**1107-1121.
168. Dernburg AF, Zalevsky J, Colaiacovo MP, Villeneuve AM: **Transgene-mediated cosuppression in the C. elegans germ line.** *Genes Dev* 2000, **14:**1578-1583.
169. Dernburg AF, Broman KW, Fung JC, Marshall WF, Philips J, Agard DA, Sedat JW: **Perturbation of nuclear architecture by long-distance chromosome interactions.** *Cell* 1996, **85:**745-759.
170. Vastenhouw NL, Fischer SE, Robert VJ, Thijssen KL, Fraser AG, Kamath RS, Ahringer J, Plasterk RH: **A genome-wide screen identifies 27 genes involved in transposon silencing in C. elegans.** *Curr Biol* 2003, **13:**1311-1316.
171. Praitis V, Casey E, Collar D, Austin J: **Creation of low-copy integrated transgenic lines in Caenorhabditis elegans.** *Genetics* 2001, **157:**1217-1226.
172. Frokjaer-Jensen C, Davis MW, Hopkins CE, Newman BJ, Thummel JM, Olesen SP, Grunnet M, Jorgensen EM: **Single-copy insertion of transgenes in Caenorhabditis elegans.** *Nat Genet* 2008, **40:**1375-1383.
173. Chen JM, Cooper DN, Chuzhanova N, Ferec C, Patrinos GP: **Gene conversion: mechanisms, evolution and human disease.** *Nat Rev Genet* 2007, **8:**762-775.
174. Robert V, Bessereau JL: **Targeted engineering of the Caenorhabditis elegans genome following Mos1-triggered chromosomal breaks.** *EMBO J* 2007, **26:**170-183.
175. Horvath P, Barrangou R: **CRISPR/Cas, the immune system of bacteria and archaea.** *Science* 2010, **327:**167-170.

176. Wiedenheft B, Sternberg SH, Doudna JA: **RNA-guided genetic silencing systems in bacteria and archaea.** *Nature* 2012, **482:**331-338.
177. Dickinson DJ, Ward JD, Reiner DJ, Goldstein B: **Engineering the Caenorhabditis elegans genome using Cas9-triggered homologous recombination.** *Nat Methods* 2013, **10:**1028-1034.
178. Gibson DG, Young L, Chuang RY, Venter JC, Hutchison CA, 3rd, Smith HO: **Enzymatic assembly of DNA molecules up to several hundred kilobases.** *Nat Methods* 2009, **6:**343-345.
179. Mello C, Fire A: **DNA transformation.** *Methods Cell Biol* 1995, **48:**451-482.
180. Krause M, Harrison SW, Xu SQ, Chen L, Fire A: **Elements regulating cell- and stage-specific expression of the C. elegans MyoD family homolog hlh-1.** *Dev Biol* 1994, **166:**133-148.
181. Hsieh J, Liu J, Kostas SA, Chang C, Sternberg PW, Fire A: **The RING finger/B-box factor TAM-1 and a retinoblastoma-like protein LIN-35 modulate context-dependent gene silencing in Caenorhabditis elegans.** *Genes Dev* 1999, **13:**2958-2970.
182. Lu X, Horvitz HR: **lin-35 and lin-53, two genes that antagonize a C. elegans Ras pathway, encode proteins similar to Rb and its binding protein RbAp48.** *Cell* 1998, **95:**981-991.
183. Ferguson EL, Horvitz HR: **The multivulva phenotype of certain Caenorhabditis elegans mutants results from defects in two functionally redundant pathways.** *Genetics* 1989, **123:**109-121.
184. Cui M, Kim EB, Han M: **Diverse chromatin remodeling genes antagonize the Rb-involved SynMuv pathways in C. elegans.** *PLoS Genet* 2006, **2:**e74.
185. Kuo MH, Allis CD: **Roles of histone acetyltransferases and deacetylases in gene regulation.** *Bioessays* 1998, **20:**615-626.
186. Lehner B, Calixto A, Crombie C, Tischler J, Fortunato A, Chalfie M, Fraser AG: **Loss of LIN-35, the**

Caenorhabditis elegans ortholog of the tumor suppressor p105Rb, results in enhanced RNA interference. *Genome Biol* 2006, **7:**R4.

187. Klymenko T, Muller J: **The histone methyltransferases Trithorax and Ash1 prevent transcriptional silencing by Polycomb group proteins.** *EMBO Rep* 2004, **5:**373-377.
188. Fong Y, Bender L, Wang W, Strome S: **Regulation of the different chromatin states of autosomes and X chromosomes in the germ line of C. elegans.** *Science* 2002, **296:**2235-2238.
189. Wang D, Kennedy S, Conte D, Jr., Kim JK, Gabel HW, Kamath RS, Mello CC, Ruvkun G: **Somatic misexpression of germline P granules and enhanced RNA interference in retinoblastoma pathway mutants.** *Nature* 2005, **436:**593-597.
190. Wu X, Shi Z, Cui M, Han M, Ruvkun G: **Repression of germline RNAi pathways in somatic cells by retinoblastoma pathway chromatin complexes.** *PLoS Genet* 2012, **8:**e1002542.
191. Reinhart BJ, Bartel DP: **Small RNAs correspond to centromere heterochromatic repeats.** *Science* 2002, **297:**1831.
192. Zilberman D, Cao X, Jacobsen SE: **ARGONAUTE4 control of locus-specific siRNA accumulation and DNA and histone methylation.** *Science* 2003, **299:**716-719.
193. Chan SW, Zilberman D, Xie Z, Johansen LK, Carrington JC, Jacobsen SE: **RNA silencing genes control de novo DNA methylation.** *Science* 2004, **303:**1336.
194. Pal-Bhadra M, Leibovitch BA, Gandhi SG, Chikka MR, Bhadra U, Birchler JA, Elgin SC: **Heterochromatic silencing and HP1 localization in Drosophila are dependent on the RNAi machinery.** *Science* 2004, **303:**669-672.
195. Horvitz HR, Sulston JE: **Isolation and genetic characterization of cell-lineage mutants of the nematode Caenorhabditis elegans.** *Genetics* 1980, **96:**435-454.

196. Rutherford SL, Lindquist S: **Hsp90 as a capacitor for morphological evolution.** *Nature* 1998, **396:**336-342.
197. Queitsch C, Sangster TA, Lindquist S: **Hsp90 as a capacitor of phenotypic variation.** *Nature* 2002, **417:**618-624.
198. Casanueva MO, Burga A, Lehner B: **Fitness trade-offs and environmentally induced mutation buffering in isogenic C. elegans.** *Science* 2012, **335:**82-85.
199. Kamath RS, Ahringer J: **Genome-wide RNAi screening in Caenorhabditis elegans.** *Methods* 2003, **30:**313-321.
200. Burga A, Casanueva MO, Lehner B: **Predicting mutation outcome from early stochastic variation in genetic interaction partners.** *Nature* 2011, **480:**250-253.
201. Wilm T, Demel P, Koop HU, Schnabel H, Schnabel R: **Ballistic transformation of Caenorhabditis elegans.** *Gene* 1999, **229:**31-35.
202. Burgers PM, Koonin EV, Bruford E, Blanco L, Burtis KC, Christman MF, Copeland WC, Friedberg EC, Hanaoka F, Hinkle DC, et al: **Eukaryotic DNA polymerases: proposal for a revised nomenclature.** *J Biol Chem* 2001, **276:**43487-43490.
203. Ohmori H, Friedberg EC, Fuchs RP, Goodman MF, Hanaoka F, Hinkle D, Kunkel TA, Lawrence CW, Livneh Z, Nohmi T, et al: **The Y-family of DNA polymerases.** *Mol Cell* 2001, **8:**7-8.
204. Srivastava VK, Busbee DL: **Replicative enzymes and ageing: importance of DNA polymerase alpha function to the events of cellular ageing.** *Ageing Res Rev* 2002, **1:**443-463.
205. Hedgecock EM, White JG: **Polyploid tissues in the nematode Caenorhabditis elegans.** *Dev Biol* 1985, **107:**128-133.
206. Melo JA, Ruvkun G: **Inactivation of conserved C. elegans genes engages pathogen- and xenobiotic-associated defenses.** *Cell* 2012, **149:**452-466.
207. Calfon M, Zeng H, Urano F, Till JH, Hubbard SR, Harding HP, Clark SG, Ron D: **IRE1 couples endoplasmic**

reticulum load to secretory capacity by processing the XBP-1 mRNA. *Nature* 2002, **415:**92-96.
208. Yoneda T, Benedetti C, Urano F, Clark SG, Harding HP, Ron D: **Compartment-specific perturbation of protein handling activates genes encoding mitochondrial chaperones.** *J Cell Sci* 2004, **117:**4055-4066.
209. Libina N, Berman JR, Kenyon C: **Tissue-specific activities of C. elegans DAF-16 in the regulation of lifespan.** *Cell* 2003, **115:**489-502.
210. Troemel ER, Chu SW, Reinke V, Lee SS, Ausubel FM, Kim DH: **p38 MAPK regulates expression of immune response genes and contributes to longevity in C. elegans.** *PLoS Genet* 2006, **2:**e183.
211. O'Rourke D, Baban D, Demidova M, Mott R, Hodgkin J: **Genomic clusters, putative pathogen recognition molecules, and antimicrobial genes are induced by infection of C. elegans with M. nematophilum.** *Genome Res* 2006, **16:**1005-1016.
212. Bolz DD, Tenor JL, Aballay A: **A conserved PMK-1/p38 MAPK is required in caenorhabditis elegans tissue-specific immune response to Yersinia pestis infection.** *J Biol Chem* 2010, **285:**10832-10840.
213. Iser WB, Wilson MA, Wood WH, 3rd, Becker K, Wolkow CA: **Co-regulation of the DAF-16 target gene, cyp-35B1/dod-13, by HSF-1 in C. elegans dauer larvae and daf-2 insulin pathway mutants.** *PLoS One* 2011, **6:**e17369.
214. Pujol N, Cypowyj S, Ziegler K, Millet A, Astrain A, Goncharov A, Jin Y, Chisholm AD, Ewbank JJ: **Distinct innate immune responses to infection and wounding in the C. elegans epidermis.** *Curr Biol* 2008, **18:**481-489.
215. Terns RM, Kroll-Conner P, Zhu J, Chung S, Rothman JH: **A deficiency screen for zygotic loci required for establishment and patterning of the epidermis in Caenorhabditis elegans.** *Genetics* 1997, **146:**185-206.
216. Gu T, Orita S, Han M: **Caenorhabditis elegans SUR-5, a novel but conserved protein, negatively regulates**

LET-60 Ras activity during vulval induction. *Mol Cell Biol* 1998, **18:**4556-4564.

217. Hartman PS: **UV irradiation of wild type and radiation-sensitive mutants of the nematode Caenorhabditis elegans: fertilities, survival, and parental effects.** *Photochem Photobiol* 1984, **39:**169-175.
218. Gartner A, Milstein S, Ahmed S, Hodgkin J, Hengartner MO: **A conserved checkpoint pathway mediates DNA damage--induced apoptosis and cell cycle arrest in C. elegans.** *Mol Cell* 2000, **5:**435-443.
219. Holway AH, Kim SH, La Volpe A, Michael WM: **Checkpoint silencing during the DNA damage response in Caenorhabditis elegans embryos.** *J Cell Biol* 2006, **172:**999-1008.
220. Koc A, Wheeler LJ, Mathews CK, Merrill GF: **Hydroxyurea arrests DNA replication by a mechanism that preserves basal dNTP pools.** *J Biol Chem* 2004, **279:**223-230.
221. Encalada SE, Martin PR, Phillips JB, Lyczak R, Hamill DR, Swan KA, Bowerman B: **DNA replication defects delay cell division and disrupt cell polarity in early Caenorhabditis elegans embryos.** *Dev Biol* 2000, **228:**225-238.
222. Brauchle M, Baumer K, Gonczy P: **Differential activation of the DNA replication checkpoint contributes to asynchrony of cell division in C. elegans embryos.** *Curr Biol* 2003, **13:**819-827.
223. Kim SH, Holway AH, Wolff S, Dillin A, Michael WM: **SMK-1/PPH-4.1-mediated silencing of the CHK-1 response to DNA damage in early C. elegans embryos.** *J Cell Biol* 2007, **179:**41-52.
224. Flibotte S, Edgley ML, Chaudhry I, Taylor J, Neil SE, Rogula A, Zapf R, Hirst M, Butterfield Y, Jones SJ, et al: **Whole-Genome Profiling of Mutagenesis in Caenorhabditis elegans.** *Genetics* 2010, **185:**431-441.
225. Hofmann ER, Milstein S, Boulton SJ, Ye M, Hofmann JJ, Stergiou L, Gartner A, Vidal M, Hengartner MO: **Caenorhabditis elegans HUS-1 is a DNA damage checkpoint protein required for genome stability**

and EGL-1-mediated apoptosis. *Curr Biol* 2002, **12:**1908-1918.

226. Claycomb JM, Batista PJ, Pang KM, Gu W, Vasale JJ, van Wolfswinkel JC, Chaves DA, Shirayama M, Mitani S, Ketting RF, et al: **The Argonaute CSR-1 and its 22G-RNA cofactors are required for holocentric chromosome segregation.** *Cell* 2009, **139:**123-134.
227. Towbin BD, Gonzalez-Aguilera C, Sack R, Gaidatzis D, Kalck V, Meister P, Askjaer P, Gasser SM: **Step-wise methylation of histone H3K9 positions heterochromatin at the nuclear periphery.** *Cell* 2012, **150:**934-947.
228. Bell JT, Pai AA, Pickrell JK, Gaffney DJ, Pique-Regi R, Degner JF, Gilad Y, Pritchard JK: **DNA methylation patterns associate with genetic and gene expression variation in HapMap cell lines.** *Genome Biol* 2011, **12:**R10.
229. Rosenfeld JA, Wang Z, Schones DE, Zhao K, DeSalle R, Zhang MQ: **Determination of enriched histone modifications in non-genic portions of the human genome.** *BMC Genomics* 2009, **10:**143.
230. Liu T, Rechtsteiner A, Egelhofer TA, Vielle A, Latorre I, Cheung MS, Ercan S, Ikegami K, Jensen M, Kolasinska-Zwierz P, et al: **Broad chromosomal domains of histone modification patterns in C. elegans.** *Genome Res* 2011, **21:**227-236.
231. Stiernagle T: **Maintenance of C. elegans.** *WormBook* 2006**:**1-11.
232. Zhang Y, Chen D, Smith MA, Zhang B, Pan X: **Selection of reliable reference genes in Caenorhabditis elegans for analysis of nanotoxicity.** *PLoS One* 2012, **7:**e31849.
233. Hoogewijs D, Houthoofd K, Matthijssens F, Vandesompele J, Vanfleteren JR: **Selection and validation of a set of reliable reference genes for quantitative sod gene expression analysis in C. elegans.** *BMC Mol Biol* 2008, **9:**9.
234. Livak KJ, Schmittgen TD: **Analysis of relative gene expression data using real-time quantitative PCR**

and the 2(-Delta Delta C(T)) Method. *Methods* 2001, **25:**402-408.
235. Li B, Dewey CN: **RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome.** *BMC Bioinformatics* 2011, **12:**323.
236. Robinson MD, Oshlack A: **A scaling normalization method for differential expression analysis of RNA-seq data.** *Genome Biol* 2010, **11:**R25.
237. Durinck S, Moreau Y, Kasprzyk A, Davis S, De Moor B, Brazma A, Huber W: **BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis.** *Bioinformatics* 2005, **21:**3439-3440.
238. Estes KA, Szumowski SC, Troemel ER: **Non-lytic, actin-based exit of intracellular parasites from C. elegans intestinal cells.** *PLoS Pathog* 2011, **7:**e1002227.
239. Bowerman B, Eaton BA, Priess JR: **skn-1, a maternally expressed gene required to specify the fate of ventral blastomeres in the early C. elegans embryo.** *Cell* 1992, **68:**1061-1075.
240. Edgar LG, McGhee JD: **DNA synthesis and the control of embryonic gene expression in C. elegans.** *Cell* 1988, **53:**589-599.
241. Azzaria M, McGhee JD: **DNA synthesis in the early embryo of the nematode Ascaris suum.** *Dev Biol* 1992, **152:**89-93.
242. Stergiou L, Doukoumetzidis K, Sendoel A, Hengartner MO: **The nucleotide excision repair pathway is required for UV-C-induced apoptosis in Caenorhabditis elegans.** *Cell Death Differ* 2007, **14:**1129-1138.
243. van Oosten-Hawle P, Porter RS, Morimoto RI: **Regulation of organismal proteostasis by transcellular chaperone signaling.** *Cell* 2013, **153:**1366-1378.
244. Aagaard L, Laible G, Selenko P, Schmid M, Dorn R, Schotta G, Kuhfittig S, Wolf A, Lebersorger A, Singh PB, et al: **Functional mammalian homologues of the Drosophila PEV-modifier Su(var)3-9 encode centromere-associated proteins which complex**

with the heterochromatin component M31. *EMBO J* 1999, **18:**1923-1938.
245. Greer Eric L, Blanco Mario A, Gu L, Sendinc E, Liu J, Aristizábal-Corrales D, Hsu C-H, Aravind L, He C, Shi Y: **DNA Methylation on N6-Adenine in C. elegans.** *Cell* 2015, **161:**868-878.
246. Simpson RT: **Nucleosome positioning in vivo and in vitro.** *Bioessays* 1986, **4:**172-176.
247. Robert VJ, Vastenhouw NL, Plasterk RH: **RNA interference, transposon silencing, and cosuppression in the Caenorhabditis elegans germ line: similarities and differences.** *Cold Spring Harb Symp Quant Biol* 2004, **69:**397-402.
248. Bessler JB, Andersen EC, Villeneuve AM: **Differential localization and independent acquisition of the H3K9me2 and H3K9me3 chromatin modifications in the Caenorhabditis elegans adult germ line.** *PLoS Genet* 2010, **6:**e1000830.
249. Meister P, Towbin BD, Pike BL, Ponti A, Gasser SM: **The spatial dynamics of tissue-specific promoters during C. elegans development.** *Genes Dev* 2010, **24:**766-782.
250. Dong X, Weng Z: **The correlation between histone modifications and gene expression.** *Epigenomics* 2013, **5:**113-116.
251. Reuben M, Lin R: **Germline X chromosomes exhibit contrasting patterns of histone H3 methylation in Caenorhabditis elegans.** *Dev Biol* 2002, **245:**71-82.
252. Lee DY, Teyssier C, Strahl BD, Stallcup MR: **Role of protein methylation in regulation of transcription.** *Endocr Rev* 2005, **26:**147-170.
253. Gasch AP, Spellman PT, Kao CM, Carmel-Harel O, Eisen MB, Storz G, Botstein D, Brown PO: **Genomic expression programs in the response of yeast cells to environmental changes.** *Mol Biol Cell* 2000, **11:**4241-4257.
254. Weiner A, Chen HV, Liu CL, Rahat A, Klien A, Soares L, Gudipati M, Pfeffner J, Regev A, Buratowski S, et al: **Systematic dissection of roles for chromatin**

regulators in a yeast stress response. *PLoS Biol* 2012, **10:**e1001369.
255. Wagner EJ, Carpenter PB: **Understanding the language of Lys36 methylation at histone H3.** *Nat Rev Mol Cell Biol* 2012, **13:**115-126.
256. Frost B, Hemberg M, Lewis J, Feany MB: **Tau promotes neurodegeneration through global chromatin relaxation.** *Nat Neurosci* 2014, **17:**357-366.
257. Almouzni G, Mechali M, Wolffe AP: **Transcription complex disruption caused by a transition in chromatin structure.** *Mol Cell Biol* 1991, **11:**655-665.
258. Jasencakova Z, Scharf AN, Ask K, Corpet A, Imhof A, Almouzni G, Groth A: **Replication stress interferes with histone recycling and predeposition marking of new histones.** *Mol Cell* 2010, **37:**736-743.
259. Mirkin EV, Mirkin SM: **Replication fork stalling at natural impediments.** *Microbiol Mol Biol Rev* 2007, **71:**13-35.
260. Shachar S, Voss TC, Pegoraro G, Sciascia N, Misteli T: **Identification of Gene Positioning Factors Using High-Throughput Imaging Mapping.** *Cell* 2015, **162:**911-923.
261. Zhang C, Montgomery TA, Gabel HW, Fischer SE, Phillips CM, Fahlgren N, Sullivan CM, Carrington JC, Ruvkun G: **mut-16 and other mutator class genes modulate 22G and 26G siRNA pathways in Caenorhabditis elegans.** *Proc Natl Acad Sci U S A* 2011, **108:**1201-1208.
262. Domeier ME, Morse DP, Knight SW, Portereiko M, Bass BL, Mango SE: **A link between RNA interference and nonsense-mediated decay in Caenorhabditis elegans.** *Science* 2000, **289:**1928-1931.
263. Ahmad K, Henikoff S: **Histone H3 variants specify modes of chromatin assembly.** *Proc Natl Acad Sci U S A* 2002, **99 Suppl 4:**16477-16484.
264. Ooi SL, Priess JR, Henikoff S: **Histone H3.3 variant dynamics in the germline of Caenorhabditis elegans.** *PLoS Genet* 2006, **2:**e97.