Characterization of parental and zygotic contributions to siRNA mediated silencing in *C. elegans*

Dissertation

Zur Erlangung des Grades

Doktor der Naturwissenschaften

Am Fachbereich Biologie

der Johannes Gutenberg-Universität Mainz

Maria Placentino

geb. am 12/06/1988 in S. Giovanni Rotondo (Italien)

Mainz, 2019

Table of contents

List of abbreviations

DDR: DNA damage response

TEs: transposable elements

LTRs: long-terminal repeats

LINEs: non-LTR autonomous long interspersed elements

SINEs: non-LTR non-autonomous short interspersed elements

TIRs: terminal inverted repeats

RNAi: RNA interference

H3K9me3: histone H3 lysine 9 trimethylation

AGO: Argonaute

PIWI: P-element induced wimpy testis

WAGO: worm-specific Argonaute

RISC: RNA-induced silencing complex

RdRP: RNA-dependent RNA polymerase

nt: nucleotide

bp: base pairs

CRISPR/Cas9: clustered regularly interspaced short palindromic repeats/ CRISPR-associated protein 9

siRNAs: small interfering RNAs

miRNAs: microRNAs

piRNAs: PIWI interacting RNAs

dsRNA/DNA: double-stranded RNA/DNA

ssRNA/DNA: single-stranded RNA/DNA

UTR: untranslated region

mRNA: messenger RNA

lncRNAs: long non-coding RNAs

lincRNAs: long intergenic non-coding RNAs

RITS: RNA-induced transcriptional silencing

PTGS: post-transcriptional gene silencing

CTGS: co-transcriptional gene silencing

RdDM: RNA-directed DNA methylation

PRG: Piwi-related gene

RNA Pol: RNA polymerase

PID: piRNA-induced silencing defective

ERI: enhanced RNAi

RDE: RNAi defective

HRDE: heritable RNAi defective

NRDE: nuclear RNAi defective

RNAe: RNA-induced epigenetic silencing

CSR: chromosome segregation and RNAi deficient

Mrt: mortal germline phenotype

MosSCI: mos1-mediated single copy insertion

IP/MS: immunoprecipitation followed by mass spectrometry

PGCs: primordial germ cells

Summary

The Piwi pathway is a germline-specific defence mechanism that animals have evolved to silence transposable elements, in order to preserve genome integrity and ensure survival of a species. In *C. elegans*, the Piwi protein PRG-1 forms a complex together with its 21U RNA co-factor, in order to recognize and silence a target RNA. Upon target recognition, an RNA-dependent RNA polymerase is recruited to the target RNA to use it as a template and synthetize secondary siRNAs. These so-called 22G RNAs are then loaded onto secondary Argonaute proteins, such as WAGO-1 and HRDE-1, to amplify the silencing reaction. In some cases, the silencing mediated by HRDE-1 can become independent from PRG-1 and is accompanied by the deposition of heterochromatic marks at the targeted locus. This form of silencing is extremely stable and can be transmitted for several generations; it depends on HRDE-1 as well as mutator proteins, and it is called RNAe. RNAe establishes in a stochastic manner, but the mechanisms behind this decision are not known.

We therefore investigated how PRG-1 mediated silencing is connected to RNAe. In chapter 2, we show that maternally provided 21U RNAs are essential for establishing *de novo* 22G RNAs production. The parental contribution of both 21U RNAs and RNAe silencing memory is necessary to instruct the silencing machinery in the next generation, to ensure appropriate gene silencing as well as gene expression; this is an essential requirement to guarantee gonad development and fertility.

In chapter 3, we use a transgenic 21U RNA target and define that maternal 21U RNAs are not only necessary, but also sufficient to initiate *de novo* silencing. Moreover, in some cases, maternal 21U RNAs can trigger RNAe. This silencing can also affect endogenous targets, thereby causing variability in the transcriptome among different individuals.

In chapter 4, we describe the characterization of a novel factor, which we named PID-2. PID-2 is required to establish *de novo* target silencing initiated by maternal 21U RNAs, in particular to boost the production of secondary 22G RNAs and to establish RNAe. PID-2 is also involved in Tc1 silencing, acting just downstream of PRG-1. PID-2 interacts with two novel Tudor proteins, PID-4 and PID-5, and together they are required for maintenance of an immortal germline over generations. These factors are linking PRG-1 mediated

silencing to RNAe, therefore we started to unravel the requirements for establishing this very stable form of silencing.

Zusammenfassung

Der Piwi-Pathway ist ein keimbahnspezifischer Abwehrmechanismus, den Tiere entwickelt haben, um mobile Elemente der Erbinformation stillzulegen. Dies dient der Erhaltung der Genomintegrität und sichert die Fruchtbarkeit und somit das Überleben einer Art. In *C. elegans* bildet das Piwi-Protein PRG-1 zusammen mit seinem 21U RNA-Cofaktor einen Komplex, um RNAs zu erkennen und stillzulegen. Dabei wird eine RNA-abhängige RNA-Polymerase an die RNA rekrutiert um sekundäre siRNAs zu synthetisieren. Diese sogenannten 22G RNAs dienen als Cofaktoren für sekundäre Argonaute-Proteine wie WAGO-1 und HRDE-1, welche die Silencing-Reaktion verstärken. In einigen Fällen kann das von HRDE-1 vermittelte Silencing unabhängig von PRG-1 werden und wird von folglich von repressiven Histonmodifikationen an den Zielgenen begleitet. Diese Form des Silencing ist extrem stabil und kann über mehrere Generationen übertragen werden; sie hängt sowohl von HRDE-1 als auch von Mutatorproteinen ab und heißt RNAe. RNAe etabliert sich stochastisch, jedoch sind die Mechanismen hinter diesem Prozess nicht bekannt.

Wir haben daher untersucht, wie PRG-1 vermitteltes Silencing mit RNAe zusammenhängt. In Kapitel 2 zeigen wir, dass mütterlich vererbte 21U RNAs für die Etablierung der *de novo* 22G RNA Produktion unerlässlich sind. Sowohl der elterliche Beitrag von 21U RNAs als auch das RNAe Silencing Gedächtnis sind notwendig, um die Silencing-Maschinerie in der nächsten Generation anzuleiten und damit ein angemessenes Gen-Silencing zu erhalten. Dies ist eine wesentliche Voraussetzung, um die Entwicklung der Gonaden und die Fruchtbarkeit zu gewährleisten.

In Kapitel 3 verwenden wir ein Transgen, das von einer 21U RNA erkannt wird. Damit können wir nachweisen, dass mütterliche 21U RNAs nicht nur notwendig, sondern auch ausreichend sind um *de novo* Silencing einzuleiten. Darüber hinaus können in einigen Fällen mütterliche 21U RNAs RNAe auslösen. Diese Stillegung kann sich auch auf endogene Zielgene auswirken, was zu einer Variabilität des Transkriptoms zwischen verschiedenen Individuen führt.

In Kapitel 4 beschreiben wir die Charakterisierung eines bisher unbekannten Proteins, das wir PID-2 genannt haben. PID-2 ist erforderlich, um *de novo* Target Silencing zu

etablieren, das von mütterlichen 21U RNAs initiiert wird. In diesem Prozess ist PID-2 erforderlich um die Produktion von sekundären 22G RNAs zu steigern und RNAe zu etablieren. PID-2 ist auch am Tc1-Silencing beteiligt und agiert dabei nach PRG-1. PID-2 interagiert mit zwei neuartigen Tudor-Proteinen, PID-4 und PID-5, und zusammen werden sie für die transgenerationale Erhaltung einer unsterblichen Keimbahn benötigt. Die beschriebenen Faktoren verbinden das PRG-1 vermittelte Silencing mit RNAe und geben erste Aufschlüsse über die Anforderungen für die Etablierung dieser sehr stabilen Form des Silencing.

Chapter 1

General introduction

Since the discovery of DNA, the mechanisms behind the safeguard of genetic information have been thoroughly investigated. It is estimated that each cell can experience up to 10^5 spontaneous lesions per day, caused either by environmental agents, such as stress, mutagens, transposons, or created during DNA metabolism (Hoeijmakers, 2009). Therefore, it is essential for maintaining genome integrity that such threats are neutralized. Maintenance of genome integrity is particularly important in germ cells, as they are responsible for transmitting genetic information from one generation to the next, to guarantee the survival of a species.

To ensure appropriate genetic inheritance, organisms have evolved a series of defensive mechanisms in order to preserve genome integrity, collectively referred to as DNA damage response (DDR). To be effective, the damage needs to be sensed, leading then to the activation of downstream repairing activities, which need to be well controlled in time and space, to avoid inappropriate alterations of the DNA structure in the context of ordinary cellular processes, such as DNA replication and telomere maintenance. The DDR is constituted of a wide variety of repair mechanisms, each of which responds to a specific damage. Defects in DDR can cause genomic instability, which is the basis of several diseases, such as neurological and immune conditions and cancer predisposition, as well as infertility, when specifically germ cells are affected (Ciccia & Elledge, 2010).

Genome stability is maintained not only at the DNA level, but also by preserving chromatin structure as well as chromosomal stability. An important layer within maintenance of chromosomal stability, is the protection of telomeres, the ends of chromosomes, to avoid them being recognized as a damage to repair. Throughout the course of cell divisions, telomeres are progressively shortened until a critical length, that leads to cellular senescence. Shortening of telomeres is avoided by the action of a specific enzyme, namely telomerase, which is active only in a subset of cell types, such as stem cells, most of cancer cells, and germ cells (Armstrong & Tomita, 2017).

Finally, genome instability arises also from transposons. Transposons are mobile genetic elements that have the potential to move around within the host genome, thereby multiplying themselves. As a consequence of their transposition mechanisms, they can insert themselves within coding regions or cause chromosomal rearrangements (see below), therefore they represent a threat for genomic integrity and have a high impact on

gene expression of the host genome (Chuma & Nakano, 2012; Malone & Hannon, 2009; Slotkin & Martienssen, 2007). Particularly in germ cells, organisms have evolved dedicated pathways to maintain genome integrity, specifically to discriminate self from non-self. Such pathways act as genomic adaptive immune systems, in analogy to the immune system whose role is to protect the organism from pathogens, as they are constantly evolving to counteract the threats represented by selfish genetic elements, such as transposons. This thesis will focus on the latter aspect of genome defence.

Transposable elements

Transposons represent up to 50% of mammalian genomes. In comparison, the exonic sequences constitute only 1-2% of the genome (Chuma & Nakano, 2012; Lander et al., 2001; Slotkin & Martienssen, 2007). This simple fact clearly illustrates the replicative power of transposable elements (TEs), and also the need for proper TE-silencing mechanisms. Even though the amount of TEs increased during evolution, quite often truncated or mutated copies of TEs are generated, therefore only a minor part of TEs is able to move within the host genome (Chuma & Nakano, 2012; Kazazian, 2004; Russell & LaMarre, 2018). Eukaryotic genomes contain a wide variety of TEs; each of them has a peculiar mechanism of transposition and can have a different influence on the loci neighbouring the insertion site.

Classification of transposons

TEs can be divided in two main classes: class I, or retrotransposons, and class II, or DNA transposons (**Figure 1**). Each class can be further divided in superfamilies; each superfamily contains TEs that share the same replication strategy. A superfamily is composed of many families, whose members have conserved DNA sequences that resemble the original active transposon, that gave rise to the family (Kapitonov & Jurka, 2008; Wicker et al., 2007). Nonetheless, there is still no unique classification system of TEs, that allows transposons to be sorted according to their mobilization mechanisms as well as their evolutionary origin (Piégu et al., 2015).

Retrotransposons replicate via an intermediate RNA molecule that will be then retrotranscribed to a DNA molecule to be inserted into a novel location of the host genome.

At each mobilization event, a new copy of the retrotransposon is generated, so that the host genome contains multiple copies of the same retrotransposon. However, not every transposition event produces a full copy of the retrotransposon, due to premature termination of the reverse-transcription process. This leads to the presence of many copies of retrotransposons that are no longer functional, because they lack significant parts of their 5' regions (Chuma & Nakano, 2012; Kazazian, 2004; Russell & LaMarre, 2018).

Depending on their structure and mechanism of transposition, retrotransposons belong to three different groups: long-terminal repeats (LTRs), non-LTR autonomous long interspersed elements (LINEs) and non-LTR non-autonomous short interspersed elements (SINEs) (**Figure 1**). LTRs are the most similar to retroviruses, from which retrotransposons have initially originated, as they still are delimitated by long-terminal repeats, and they encode the reverse transcriptase as well as the endonuclease required for transposition; therefore, they are also classified as autonomous elements. LINEs encode also a reverse transcriptase, however they are no longer characterized by long-terminal repeats. SINEs are smaller than the TEs of the other two classes, they do not encode functional proteins, and they mostly exploit the machinery produced by LINEs. LINEs and SINEs are the most abundant retrotransposons in mammalian genomes and are therefore one of the main sources of individual variation within a species (Chuma & Nakano, 2012; Jurka et al., 2007; Kazazian, 2004; Malone & Hannon, 2009; Russell & LaMarre, 2018).

DNA transposons do not need a reverse transcription step, but they transpose by directly moving their genomic DNA, using a single- or double-stranded intermediate. In mammals, DNA transposons are mostly inactive and represent only 2-3% of the genome, whereas they represent the majority of TEs in the genome of the nematode *Caenorhabditis elegans* (Chuma & Nakano, 2012). DNA transposons can be further subdivided according to their replication mechanism: "cut-and-paste", rolling-circle and self-replicating (Feschotte & Pritham, 2007; Jurka et al., 2007) (**Figure 1**).

The cut-and-paste TEs are characterized by terminal inverted repeats (TIRs), that are recognized by a TE-encoded transposase protein (Feschotte & Pritham, 2007). The transposase cuts the double-stranded DNA transposon out, to reinsert it in a different genomic location (Craig et al., 2002). This mobilization mechanism leaves behind a portion of the transposon sequence, so that even after the transposon has excised itself, short

repeats can be found at the genomic location of the former insertion, upon repair of the double-strand break (Malone & Hannon, 2009). Even though this mechanism does not directly lead to an increase of the number of TE copies, such TEs can still multiply themselves; for instance, when such TEs move during DNA replication, from a chromosome region that has already been replicated to another that has not yet (Wicker et al., 2007). Alternatively, if the double-strand break is repaired by homologous recombination and the TE is present in the homologous chromosome, the TE will be then copied back (Feschotte & Pritham, 2007). The TEs that use a rolling-circle-like mechanism to transpose, also named *Helitrons*, do not have TIRs, but are instead delimitated by short conserved motifs. Some *Helitrons* can also be autonomous, if they code for a Rep/Hel protein, characterized by a replication initiator (Rep) and a helicase (Hel) domains. Specifically, Rep is responsible for cutting and ligating the DNA during the transposition reaction, in analogy to the transposase activity. Hel, on the other hand, is involved in unwinding the double-stranded DNA to facilitate the synthesis of a novel DNA strand, starting from the nick executed by Rep, until it has reached the starting site again, making thus a full circle. *Helitrons* exploit the host machinery to ensure their replication. Due to this transposition mechanism, *Helitrons* increase their copy number and, furthermore, have the ability of moving host genes during transposition, representing a potential tool for evolution (Jurka et al., 2007; Kapitonov & Jurka, 2001). The last type of class II transposons, *Maverick*/*Polintons*, are also characterized by long TIRs and use a yet unknown mechanism of transposition, probably mediated by a self-encoded DNA polymerase. Both *Helitrons* and *Mavericks* possibly use a single-stranded DNA intermediate to transpose, by its displacement or by its replication, respectively (Feschotte & Pritham, 2007).

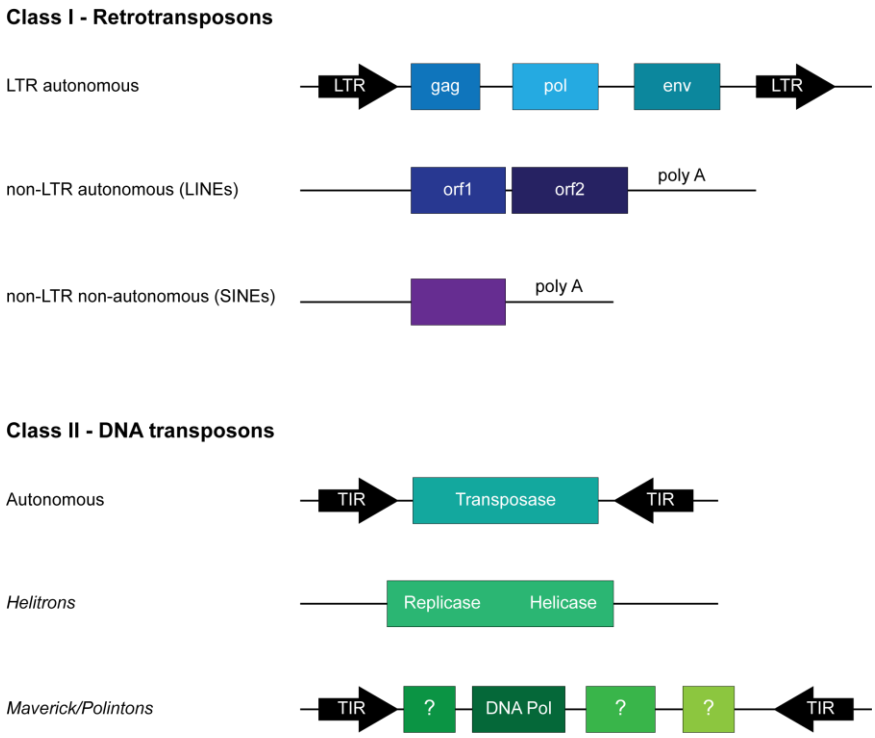


Figure 1. Transposable elements can be classified in two main classes: class I, or retrotransposons, and class II, or DNA transposons. Retrotransposons transpose via an RNA intermediate and can be autonomous, if they encode a reverse transcriptase and an endonuclease, such as LTRs, most similar to retroviruses, and LINEs. On the other hand, non-autonomous retrotransposons, like SINEs, do not code for functional proteins and exploit the machinery produced by autonomous retrotransposons. Class II, or DNA transposons, move directly their genomic DNA via diverse transposition mechanisms. The "cut-and-paste" transposons encode a transposase, therefore they are also classified as autonomous. The TEs that transpose using a rolling-circle-like mechanism, such as *Helitrons*, can also be autonomous, if they code for a functional Rep/Hel protein. *Maverick/Polintons* transpose via a yet uncharacterized mechanism, likely self-replicating, involving a TE-encoded DNA polymerase.

Transposons in *Caenorhabditis elegans*

The genome of *C. elegans* contains ~12% of transposons. Contrary to vertebrates, these are mainly class II TEs, that move via a DNA intermediate, and belong to the *IS630/Tc/mariner* superfamily (Bessereau, 2006; Das et al., 2008; Laricchia et al., 2017; Sijen & Plasterk, 2003). Transposons are actively jumping in somatic cells, whereas they are silenced in the germline by the RNA interference machinery (see below), as an obvious defence mechanism to preserve genome integrity and ensure accurate inheritance of genetic information to the next generation (Sijen & Plasterk, 2003; Vastenhouw et al., 2003).

As the majority of active TEs in *C. elegans* belong to the *Tc*/*mariner* superfamily, most of the genetic screens performed to identify genes involved in transposon silencing focussed on these TEs (Shirayama et al., 2012; Vastenhouw et al., 2003). Tc1 is one of the most active transposons in the genome of *C. elegans*. Its mobilization requires the activity of a transposase protein, named Tc1A, encoded by Tc1 itself. Tc1A recognizes the TIRs in a sequence-specific manner, and moves the Tc1 element to a different genomic location with a cut-and-paste mechanism: it excises Tc1 from the DNA using two double-strand breaks, and re-inserts it at a novel locus using a direct trans-esterification reaction. The remaining single-strand breaks at the new insertion site are repaired by cellular DNA repair mechanisms. Similar to other DNA TEs, Tc1 activity is silenced in the germline of most isolates of *C. elegans* by RNAi; in fact, mutants for any of the factors involved in the RNAi machinery causes reactivation of Tc1 in the germline. On the other hand, somatic activity of Tc1 is present in virtually all the isolates (Eide & Anderson, 1985; Fischer et al., 2003; Sijen & Plasterk, 2003).

Interestingly, as a result of transposon activity, there is high variability in copy numbers of single transposon families amongst wild isolates of *C. elegans*, reflecting potentially different regulatory mechanisms and/or various evolutionary advantages that have been selected over time (Laricchia et al., 2017). Extremely rare are TEs insertions within coding sequences, especially if loss of this gene results in a lethal phenotype, indicating a strong negative selection. Such insertions are often maintained if they are located within recently duplicated genes, or genes that have redundant functions with others, so that they do not really affect the functionality of the genome as a whole (Laricchia et al., 2017).

Consequences of transposon activity

Upon mobilization, transposons can cause DNA breaks and insert themselves in different genomic locations, potentially causing mutations. If a TE lands within a gene, it can disrupt the coding sequence, inducing loss-of-function phenotypes. Transposons can also act as enhancers or promoters, influencing the expression of neighbouring genes as well as epigenetic methylation patterns, such as on imprinted genes (Cordaux & Batzer, 2009; Jurka et al., 2007; Kazazian, 2004; Rebollo et al., 2012; Slotkin & Martienssen, 2007). Furthermore, they can serve as a substrate for non-allelic recombination, causing

chromosomal breakage or rearrangements, such as inversions and deletions (Feschotte & Pritham, 2007). Therefore, it is extremely important to keep their activity under control.

Since their discovery (McClintock, 1951), however, it became more and more clear that transposon activity can also be advantageous for the host; TEs can indeed drive speciation and contribute to genetic variation within the species (Jurka et al., 2007; Rebollo et al., 2012). DNA transposons, in particular, seem to favour an insertion site close to a gene and, afterwards, they can spontaneously excise themselves, often leaving "scars". As a consequence, they can generate easily different alleles, contributing to diversity within a species.

With increasing information from deep sequencing on genome composition, it is also becoming evident that quite a few host genes are in fact derived from TEs (Feschotte & Pritham, 2007). A striking example for evolution derived by domestication of TEs is represented by the V(D)J recombination in the adaptive immune system of jawed vertebrates (Carmona & Schatz, 2017). These recombination events share a lot of similarities with the transposition mechanisms of class II TEs, and are in fact driven by enzymes that clearly are derived from an ancient TE (Feschotte & Pritham, 2007; Jangam et al., 2017; Slotkin & Martienssen, 2007). Another example of co-option of a transposon sequence by the host genome is represented by the Arc gene, which encodes a protein derived from Gag proteins of retrotransposons. Arc is essential for synaptic plasticity in the neuronal system, from *Drosophila* to humans, and, intriguingly, it can form capsid-like structures that are packed into extracellular vesicles to transport RNA molecules through the synapsis, thereby retaining its initial virus-like property (Ashley et al., 2018; Pastuzyn et al., 2018).

On the other hand, though, several human diseases correlate with transposon activity. The first TE insertion to be linked to a disease was a novel LINE-1 insertion in the factor VIII gene, consequently causing haemophilia A (Kazazian et al., 1988). Since then, many more diseases, often X-chromosome-linked, have been associated to TEs (Ayarpadikannan & Kim, 2014; Belancio et al., 2009). For example, insertion of a 5' truncated LINE-1 retrotransposon within the dystrophin gene causes exon skipping during splicing, thereby leading to Duchenne muscular dystrophy (Narita et al., 1993). Also, a different LINE-1 insertion within the RP2 locus is responsible for causing the X-linked retinitis pigmentosa,

a progressive degeneration of the retina (Schwahn et al., 1998). More recently, there is evidence for LINE-1 activity in few brain regions, thus implying a possible role for LINE-1 in normal brain development and function as well as in psychiatric disorders (Guffanti et al., 2014).

Furthermore, retrotransposons are reactivated in many forms of cancer (Ayarpadikannan & Kim, 2014). For instance, a recent somatic Alu insertion has been found within the BRCA1/2 genes, which account for a high predisposition for familial breast cancer, when mutated (Teugels et al., 2005). Also, a somatic LINE-1 insertion disrupts the APC tumour suppressor gene, causing colon cancer (Miki et al., 1992). Another example is represented by acute myeloid leukaemia (AML), which is often associated with chromosomal translocations, mediated by recombination between two Alu sequences. Such chromosomal translocations involve the MLL1 gene and leads to the formation and expression of an oncogenic fusion gene, formed by MLL1 and the gene present on the partner chromosome at the site of translocation (Belancio et al., 2009).

In general, hypomethylation of the genome is a characteristic of cancer cells. In particular, the reduced level of 5-methyl cytosine is prevalent in intronic as well as intergenic regions, thereby affecting repetitive DNA sequences as well as transposons. Consequently, such global demethylation could cause an increased frequency of mutations and chromosomal rearrangements, promoting initiation and progression of cancer as well as of other pathological conditions (Wilson et al., 2007). Yet, DNA methylation is one of the defence mechanisms against transposons, ensuring their silencing. Upon demethylation, LINE-1, as well as other TEs, could thus become expressed and induce aberrant expression of neighbouring host genes (Ayarpadikannan & Kim, 2014; Chuong et al., 2017; Wilson et al., 2007). In fact, LINE-1 and Alu demethylation, and their consequent reactivation, have indeed been detected both in non-small cell lung carcinoma (Daskalos et al., 2009) and in colorectal cancer (Suter et al., 2004). Such events happen possibly very early during pathogenesis, as they can be detected also in the healthy tissues located in the vicinity of the tumour, and somehow contribute to the progression of the disease, as they positively correlate with increased genome instability (Daskalos et al., 2009; Suter et al., 2004).

Novel applications based on transposable elements

Transposons have been also extensively used as genetic tools, thanks to their mobile properties (Feschotte & Pritham, 2007; Kazazian, 2004; Slotkin & Martienssen, 2007). In particular, P element transposons have been widely used for mutagenesis studies in *Drosophila*, in order to generate knockout mutants for every coding sequence present in the genome (Spradling et al., 1999). Another transposon from *Drosophila*, the Mos1 element, has been inserted in the genome of *C. elegans* and largely employed for generation of transgenic lines as well as mutants (Bessereau et al., 2001; Frøkjær-Jensen et al., 2010, 2014, 2008; Robert & Bessereau, 2007).

More recently, class II TEs, in particular the Sleeping Beauty and PiggyBac transposons, started to be employed also for therapeutic applications (Kebriaei et al., 2017). The first application of a system based on the Sleeping Beauty transposon was in the field of immunotherapy, to generate a specific subset of T cells, that expresses a chimeric receptor, in order to specifically identify tumour cells and target them for elimination. Such modified T cells have been then reintroduced as adjuvant in patients affected by $CD19^{+}$ B-lymphoid malignancies, such as non-Hodgkin lymphoma and acute lymphoblastic leukaemia, in the context of hematopoietic stem cells (HSCs) transplant, which had a higher success rate (Kebriaei et al., 2016). Viral systems have been largely employed for gene therapy; they can efficiently infect the host cells, thereby delivering the genetic material, yet, often they cannot integrate in the host genome. A transposon-based system, on the other hand, can efficiently ensure insertion in the host genome (Kebriaei et al., 2017). The two systems can therefore be integrated to achieve an increased efficiency of gene therapy.

Mechanisms of transposon silencing

Transposons, as well as viruses, exploit the cellular machinery of their hosts to achieve their propagation and survival. Hence, the hosts have evolved a variety of defensive mechanisms to fight such parasites, which, on the other hand, need to overcome such challenges (Madhani, 2013). Transposition events are counteracted in different ways. Most commonly, defensive mechanisms are based on small RNAs (see below) and on chromatin modifications (Slotkin & Martienssen, 2007). Histones associated with TEs are often marked by histone H3 lysine 9 trimethylation (H3K9me3) repressive mark, to favour the

formation of heterochromatin at the loci of insertion. Therefore, TEs are eventually mostly localized in heterochromatic regions, such as pericentromeric or intergenic regions, promoters or introns, thus potentially affecting the expression of neighbouring genes as well as the chromatin landscape (Chuma & Nakano, 2012; Feschotte & Pritham, 2007; Laricchia et al., 2017; Malone & Hannon, 2009; Slotkin & Martienssen, 2007). Consequently, mutation of factors that are required for deposition of H3K9me3 results in TE activation (Slotkin & Martienssen, 2007). Also, DNA methylation on cytosine residues plays a role in suppressing TE activity. Therefore, proteins that are generally involved in establishing chromatin modifications, can also be required for silencing transposons (Slotkin & Martienssen, 2007). In mammals, also the transcription factors Kruppel-associated box zinc-finger proteins (KRAB-ZFPs) have their relevance in transposon silencing. They bind to TEs via specific DNA motifs, to repress them at the transcriptional level, thereby also regulating expression of adjacent genes (Yang et al., 2017). Finally, transposon sequences are often targeted by silencing mechanisms based on small RNA molecules, known as RNA interference (RNAi) (Feschotte & Pritham, 2007; Slotkin & Martienssen, 2007).

Despite all these defence mechanisms, genetic parasites are still present virtually in all domains of life. Viruses constantly evolve counter-defensive mechanisms to replicate, as they mutate very rapidly and also encode proteins to neutralize the host defence systems (Koonin & Dolja, 2013). On the other hand, it is completely unresolved, how TEs can counteract the defence mechanisms of the host genomes. Yet, as such defence mechanisms evolve relatively fast, it seems likely that TEs affect somehow such silencing mechanisms (Madhani, 2013).

RNA interference (RNAi)

RNA interference was discovered in 1998, when Fire and Mello observed that the introduction of double-stranded RNA (dsRNA) into the nematode *C. elegans* can induce silencing of a complementary target (Fire et al., 1998). RNA interference is more generally used to indicate a large variety of pathways in which a small RNA molecule is used by an Argonaute protein to form a complex, essential to recognize a target RNA, via base pair complementarity, and silence it via post-transcriptional gene silencing. These small RNA

silencing pathways have evolved to silence transposable elements (Ketting et al., 1999; Tabara et al., 1999) and to counteract viral infections (Berkhout, 2018; Hamilton & Baulcombe, 1999; Karlikow et al., 2014; Wilkins et al., 2005; Yang & Li, 2018). This defensive function is seemingly very ancient, as closely related key proteins responsible for silencing are found in plants, nematodes and flies (Fagard et al., 2000; Félix et al., 2011; Li et al., 2002; Mourrain et al., 2000). In the germline of *C. elegans*, the RNAi machinery can use the ample pool of small RNA molecules to protect the genome, not only from transposons, but also from other exogenous DNA sequences. For example, RNAi is also responsible for silencing multicopy transgenes. Furthermore, if the transgene has a sequence homologous to an endogenous gene, the latter will also be silenced via a mechanism known as co-suppression (Dernburg et al., 2000; Ketting & Plasterk, 2000; Robert et al., 2005). A silencing phenomenon similar to RNAi, named quelling, was already described in the fungus *Neurospora crassa* and it is also triggered by dsRNAs (Cogoni & Macino, 1997; Fulci & Macino, 2007; Romano & Macino, 1992).

Argonaute proteins

Argonaute proteins have a central role in RNAi pathways. They can be further divided in three clades: AGO, PIWI and WAGO. The AGO clade comprises the proteins that are more similar to *Arabidopsis* AGO1. The PIWI clade is mostly germline specific and the proteins that are included in this group are similar to *Drosophila* Piwi. The WAGO clade contains a subset of Argonaute proteins that are worm specific. Both AGO and WAGO proteins are ubiquitously expressed (Ender & Meister, 2010).

Argonaute proteins share the following domain composition: N-terminal, PAZ (Piwi-Argonaute-Zwille), MID (middle), PIWI. These domains are arranged to form two lobes, connected through the linkers L1 and L2; the N-terminal lobe contains the N and PAZ domains, whereas the MID and PIWI domains are part of the C-terminal lobe. A key function of Argonaute proteins is to bind a small RNA molecule (see next section) to form a functional RNA-induced silencing complex (RISC). The small RNA molecule serves as a guide for the Argonaute protein to identify target nucleic acids. The binding of this small RNA molecule takes place in the channel formed between the two lobes (Olina et al., 2018; Schirle & MacRae, 2012; Sheu-Gruttadauria & MacRae, 2017). The MID domain binds

the 5’ end of the small RNA molecule. The nucleotide at the very 5’ end of the small RNA never pairs with the target, as it is positioned at the interface between the MID and PAZ domain, and therefore it is not accessible for base pairing (Jinek & Doudna, 2009; Sheu-Gruttadauria & MacRae, 2017). The PAZ domain forms a pocket that binds the 3’ end of the small RNA (Jinek & Doudna, 2009). In order to ensure the base pairing between the whole small RNA molecule and its target, the PAZ domain needs to release the 3’ end of the small RNA (Sheng et al., 2014; Wang et al., 2009). The PIWI domain is characterized by an RNAse-H-like motif. It contains a catalytic tetrad (DEDX; X = D/H) that is responsible of positioning two Mg^{2+} ions in order to cleave the target RNA. However, not all AGO proteins are competent for target cleavage; many have lost the catalytic residues, therefore they cannot perform an endonucleolytic cleavage of the target (Matsumoto et al., 2016; Nakanishi et al., 2012; Olina et al., 2018; Ozata et al., 2018; Schirle & MacRae, 2012; Sheu-Gruttadauria & MacRae, 2017; Song et al., 2004).

After the RISC complex is formed, it can recognize the target RNA via base pair complementarity. Depending on the degree of complementarity between the small RNA and its target, the target RNA can encounter a variety of fates. If the complementarity between the small RNA and its target is perfect, the target RNA will be cleaved and degraded. On the other hand, if the pairing between the small RNA and its target presents mismatches, RISC can recruit additional downstream effectors, which could inhibit translation, or induce destabilization and eventually degradation of the target. In some organisms, upon target recognition, an RNA-dependent RNA polymerase (RdRP) is recruited to synthesize small RNAs, complementary to the target RNA, in order to amplify the silencing reaction. In some instances, such post-transcriptional silencing can be coupled to a nuclear form of silencing, which affects the chromatin landscape (Castel & Martienssen, 2013; Ghildiyal & Zamore, 2009; Jinek & Doudna, 2009; Ketting, 2011; van Wolfswinkel & Ketting, 2010). Nevertheless, in all cases the pairing of the target RNA with the so-called seed region of the small RNA (nucleotides 2-8) is fundamental to form a stable complex (Olina et al., 2018).

Argonaute proteins are very well conserved and also found in prokaryotes, indicating they have an ancient origin. Prokaryotic Argonautes can use either RNA or, interestingly, DNA guides, as shown from *in vitro* studies on *Thermus thermophilus*, *Pyrococcus furiosus* and

Methanocaldococcus jannaschii (Olina et al., 2018; Swarts et al., 2014). Their preference towards DNA guides is possibly due to the hydrophobic properties of the MID domain (Swarts et al., 2014). Even though most prokaryotic Argonautes do not have a characterized function yet, it has been shown, for the above-mentioned organisms, their ability to target exogenous DNA (e.g. plasmid DNA). This is strongly pointing towards the direction of a conserved genome defence function, suggesting that this is possibly the most ancient, original function of these prokaryotic Argonautes (Olina et al., 2018; Swarts et al., 2014). Nonetheless, the genome defence mechanism in prokaryotes is exerted mainly by the CRISPR/Cas interference system. Different types of Cas proteins can recognize and cleave exogenous DNA or RNA molecules, using as guides small RNA molecules, that have been taken up within the CRISPR loci, as a memory of the past infections (Olina et al., 2018). For more details on the prokaryotic systems, see previously published reviews (Marraffini, 2015; Olina et al., 2018; Sternberg et al., 2016; Swarts et al., 2014).

Classes of small RNAs

Small non-coding RNAs can be divided in three main classes, depending on their biogenesis and mechanism of action: small interfering RNAs (siRNAs), microRNAs (miRNAs) and PIWI interacting RNAs (piRNAs) (**Figure 2**).

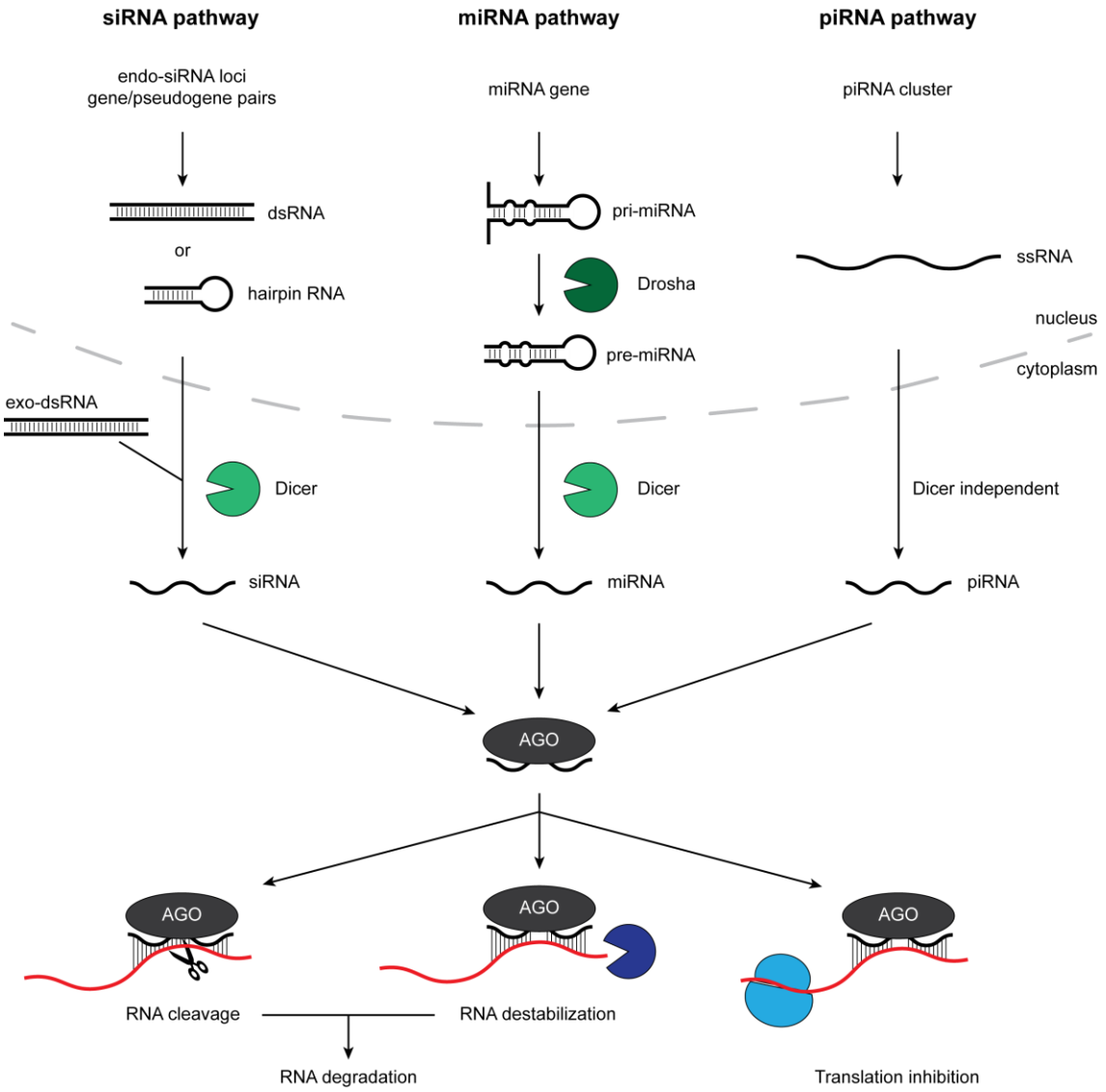


Figure 2. Schematic overview of the biogenesis of different classes of small RNAs. siRNAs (left) are produced from various sources of long dsRNAs, which are then cleaved by Dicer, to produce shorter dsRNAs that will give rise to a mature siRNA molecule. miRNAs (centre) derive mainly from stem-loop precursors, transcribed by RNA Pol II. The pri-miRNA is cleaved by Drosha to generate a pre-miRNA, which is later processed by Dicer, to eventually produce a mature miRNA. piRNAs (right) are transcribed as long ssRNA precursors and their maturation process is Dicer-independent. Mature small RNAs are then loaded onto a cognate Argonaute protein to recognize a target RNA. Depending on the degree of complementarity, the target RNA will encounter different fates.

siRNAs

Short interfering RNAs can be of exogenous (exo-siRNAs, e.g. viral) or endogenous origin (endo-siRNAs). They are usually generated from long double-stranded RNA (dsRNA) molecules (**Figure 2**). The production of such dsRNAs differs among species. For example, plants and *Schizosaccharomyces pombe* use an RNA-dependent RNA polymerase to synthesise long dsRNAs. On the other hand, in *Drosophila* as well as in testes of *Tupaia belangeri*, siRNAs derive either from hairpins, which are the result of self-folding of certain messenger RNAs (mRNAs), or from convergent transcription. In murine oocytes, dsRNAs can be generated by the pairing of a coding transcript with another transcript, produced from a related pseudogene (Carthew & Sontheimer, 2009; Ghildiyal & Zamore, 2009; Golden et al., 2008; Rosenkranz et al., 2015). Regardless of their origin, these precursors are then processed by enzymes of the RNase III family, such as Dicer, into shorter dsRNAs of 20-25 nucleotides (nt) (**Figure 2**). The cleavage performed by Dicer leaves a 2 nt 3'OH overhang and a 5' monophosphate on both ends of the dsRNA (Bernstein et al., 2001). Afterwards, the duplex will be loaded onto an Argonaute protein; one of the two siRNA strands will be retained (guide strand) and used to recognize its perfectly complementary target, while the other strand will be expelled and degraded (passenger strand). Usually, the guide strand is chosen depending on its thermodynamic stability: the strand whose 5' end requires less energy to be unwound will be retained within the Argonaute protein (Khvorova et al., 2003; Schwarz et al., 2003; Tomari et al., 2004).

After RISC formation, the target RNA is recognized with perfect complementarity and this results in endonucleolytic cleavage of the target RNA between the nucleotides 10 and 11, relative to the 5' end of the guide RNA (Olina et al., 2018; Sheu-Gruttadauria & MacRae, 2017) (**Figure 2**). In some cases though, siRNAs can also bind targets with imperfect complementarity. In this scenario, siRNAs initiate a miRNA-like response (see below), thereby targeting also RNA molecules that are not perfectly complementary. Fundamental to trigger such response is a perfect pairing between positions 2 and 8 of the siRNA with the target. Such effect is particularly relevant when using RNAi to silence a specific target in an experimental setup, as it can induce off-target, undesirable effects (Jackson & Linsley, 2010).

In plants, siRNAs are the mediators of transposon silencing, together with downstream DNA and histone methylation. siRNAs are also involved in transposon silencing in vertebrates, such as in the murine female germline (Tam et al., 2008; Watanabe et al., 2008), as well as in *Drosophila*, but only in somatic tissues (Ghildiyal et al., 2008; Kawamura et al., 2008; van Wolfswinkel & Ketting, 2010).

miRNAs

microRNAs are involved in practically all physiological processes and are responsible for regulating gene expression (Ender & Meister, 2010; Sheu-Gruttadauria & MacRae, 2017). Primary miRNA precursors (pri-miRNAs) are transcribed by RNA polymerase II and are characterized by a stem-loop that contains the mature miRNA(s). This stem-loop is characterized by bulges and mismatches, that affect its further processing. The stem-loop is recognized by a protein complex named "microprocessor", which will perform the first step of miRNA maturation: a cut in both RNA strands at the base of the stem-loop of the hairpin to produce a secondary miRNA precursor hairpin (pre-miRNA). The microprocessor contains an RNAse III enzyme, Drosha, and its cofactor Pasha/DGCR8, that is a double-stranded RNA binding protein. The pre-miRNA is then exported into the cytoplasm via Exportin 5, where it will be further processed by Dicer and its cofactor TRBP to generate a dsRNA of 21-23 nt (**Figure 2**). The duplex is then loaded onto an Argonaute protein, but only one of the two strands will be retained by the AGO (miRNA), while the complementary strand will be discarded (miRNA*), based on their relative thermodynamic stability. In fact, the non-perfect base pairing in this RNA duplex defines which of the two strands will be effectively loaded onto the Argonaute protein.

Contrary to siRNAs, miRNAs recognize their target allowing for some mismatches (Ender & Meister, 2010; Matsumoto et al., 2016; Nakanishi et al., 2012; Olina et al., 2018; Ozata et al., 2018; Schirle & MacRae, 2012; Sheu-Gruttadauria & MacRae, 2017; Song et al., 2004). In order to recognize its target, the seed region of the miRNA (nt 2-8) needs to pair perfectly to the mRNA target. Positions 10-11 are instead characterized by mismatches, in order to prevent the slicing of the target mRNA (**Figure 2**). Mismatches between the miRNA and its target are more tolerated towards the 3' end, rather than in the seed region. On the other end, it has been recently shown that the seed-distal region (nt 13-16) is also

important for target recognition, especially to reinforce target recognition when the seed region presents mismatches, and confers specificity to each member of miRNA families, as they share the same seed sequence (Brancati & Großhans, 2018; Broughton et al., 2016). The majority of miRNA binding sites resides on the 3' UTR of target mRNAs. Targeting could also occur at their 5' UTR or within the coding sequence, but these mechanisms seem to be less frequent and efficient for silencing, if compared to target sites residing within the 3' UTR (Bartel, 2018).

miRNAs mostly repress their targets by inhibition of translation initiation and/or elongation or by recruiting the CCR4-NOT deadenylase complex, that will then remove the polyA tail of the mRNA, inducing its destabilization and degradation (Arribas-Layton et al., 2013; Eulalio et al., 2008; Guo et al., 2010; Hendrickson et al., 2009; Wakiyama et al., 2007) (**Figure 2**). Such additional silencing complexes are recruited to the target mRNA by the Argonaute protein and its cofactor GW182 (Iwakawa & Tomari, 2015; Valencia-Sanchez et al., 2006; Wilczynska & Bushell, 2015). On the other hand, in plants, miRNAs recognize their target with (almost) perfect complementarity, which will result in target cleavage, similar to siRNA-mediated post-transcriptional gene silencing (Baumberger & Baulcombe, 2005; Rhoades et al., 2002; Schwab et al., 2005).

Another class of miRNAs is generated by an alternative pathway, independently of Drosha, from intronic sequences of mRNAs and are called mirtrons. Their precursors are produced via the splicing machinery. The first product of the splicing reaction is a lariat structure; after debranching, these mirtron precursors fold into a short stem-loop, resembling a pre-miRNA molecule, therefore they can bypass the requirement of Drosha for their biogenesis (Okamura et al., 2007; Ruby et al., 2007; Sibley et al., 2012). Additionally, other small non-coding RNAs, such as endogenous short hairpin RNAs, produced directly by transcription, transfer RNAs (tRNAs) or small nucleolar RNAs (snoRNAs), can be used as miRNA precursors (Ha & Kim, 2014). For example, some snoRNAs form a hairpin structure, resembling a pre-miRNA, therefore also in this scenario Drosha is dispensable, and the mature small RNAs can function as miRNAs (Babiarz et al., 2008; Chong et al., 2010; Ender et al., 2008; Ha & Kim, 2014). Dicer, though, is still required for the maturation of these miRNAs.

One particular miRNA, miR-451, is produced independently of Dicer, as after Drosha cleavage, its pre-miRNA is too short to be processed by Dicer. The biogenesis of miR-451 requires instead the catalytic activity of AGO2 (Cheloufi et al., 2010; Cifuentes et al., 2010; Ha & Kim, 2014; Yang et al., 2010), which cleaves the passenger strand. After the cleavage, the two resulting passenger strand fragments are removed from the RISC complex.

An additional non-conventional miRNA biogenesis pathway involves the terminal uridylyl transferases (TUTs). They are required for adding one nucleotide at the pre-miRNA molecules that carry 1 nt 3'OH overhang after Drosha cleavage, instead of the canonical 2 nt 3'OH overhang. After the addition of one extra terminal nucleotide, such pre-miRNAs are also suitable for further processing by Dicer (Ha & Kim, 2014; Heo et al., 2012).

piRNAs

The PIWI-interacting RNAs represent a germline specific class of small RNAs whose main duty is to protect the genome of germ cells from TE activity. Loss of PIWI proteins typically causes transposon upregulation and defects in gametogenesis, ultimately resulting in sterility (Ishizu et al., 2012; Siomi et al., 2011). Here, I describe some characteristics of piRNAs from *Drosophila* and mouse; more details about piRNAs in *C. elegans* will follow later.

In *Drosophila*, mature piRNAs are ~24-30 nt long and are loaded onto a PIWI protein, in order to recognize and silence transposon mRNAs. *Drosophila* has three PIWI proteins: Piwi, Aubergine (AUB) and AGO3. Piwi localizes in the nucleus, whereas AUB and AGO3 localize in the nuage (see below) and take part in the so-called "ping-pong" amplification cycle. piRNAs do not have only a role in transposon silencing, as a portion of the piRNA pool does not show complementarity to transposable elements, but rather could target endogenous protein coding genes and therefore regulate gene expression (Rojas-Ríos & Simonelig, 2018; Simonelig, 2014; Siomi et al., 2011).

In *Drosophila*, the biogenesis of piRNAs consists of two steps, primary biogenesis and ping-pong amplification. Both mechanisms are Dicer independent (Houwing et al., 2007; Vagin et al., 2006). piRNAs are produced from long single-stranded (ss) precursor

transcripts (**Figure 2**). Such precursors are transcribed by RNA polymerase II and can originate from uni-strand or dual-strand clusters, which contain remainders of active transposons as well as full sequences of TEs that, during evolution, landed in these genomic locations by transposition, thus representing a memory system. In order to fully protect germ cells from TEs, these clusters are conserved in localization but their sequence content evolves extremely fast. The piRNA clusters are usually localized in heterochromatic regions, such as pericentromeric and telomeric regions, adjacent to euchromatic regions (Czech et al., 2018; Ghildiyal & Zamore, 2009; Guzzardo et al., 2013; Yamashiro & Siomi, 2018).

The precursors that originate from uni-strand clusters are defined by specific promoters, whereas dual-strand clusters are not defined by specific promoters, and use a different transcription strategy. Uni-strand clusters, such as *flamenco*, produce transcripts that are virtually indistinguishable from any other cellular mRNA, as they also present a 5' cap, a polyA tail at their 3' end, and are spliced, before being exported into the cytoplasm (Czech et al., 2018; Ozata et al., 2018; Yamashiro & Siomi, 2018).

Dual-strand clusters are covered by the heterochromatic mark H3K9me3, which has a double function: it promotes heterochromatin formation to silence TEs, but also allows their basal transcription to generate a source to initiate piRNA biogenesis (Czech et al., 2018; Ozata et al., 2018; Yamashiro & Siomi, 2018). In *Drosophila* germ cells, an HP1 homolog protein, Rhino, forms a complex together with Cutoff and Deadlock, to bind the H3K9me3 mark on dual-strand clusters and stimulate their transcription, by recruiting yet two other cofactors, Moonshiner and TRF2 (Andersen et al., 2017; Klattenhoff et al., 2009; Pane et al., 2011; Zhang et al., 2014). This complex ensures the transcription of piRNA precursors, independently of promoters. Maelstrom, furthermore, represses the transcription from canonical promoters present within dual-strand clusters, to favour the transcription of piRNA precursors (Chang et al., 2019). Rhino is actually involved in *de novo* formation of dual-strand piRNA clusters (Akulenko et al., 2018) and counteracts splicing, whereas the DEAD-box helicase UAP56 binds the piRNA precursors, ensuring their export from the nucleus (Ozata et al., 2018). The definition of these piRNA clusters is mediated by the protein Piwi in *Drosophila* and needs to be established already during

embryogenesis, to ensure an appropriate chromatin state in adulthood, when its maintenance does not require Piwi anymore (Akkouche et al., 2017).

After transcription and transcript processing in the nucleus, these RNA molecules are exported into the cytoplasm for further processing, to generate mature piRNA molecules (Czech et al., 2018; Ozata et al., 2018) (**Figure 2**). An important factor in this maturation step is the endonuclease Zucchini. This enzyme is involved in generating both the 5' end as well as the 3' end of primary piRNAs (Han et al., 2015; Ipsaro et al., 2012; Mohn et al., 2015; Nishimasu et al., 2012). Additionally, another exonuclease, namely Nibbler, has also a role in removing the extra nucleotides at the 3' end (Feltzin et al., 2015; Hayashi et al., 2016; Wang et al., 2016). On the other hand, during the ping-pong cycle, PIWI proteins have a role in generating the 5' end of piRNAs, by cleaving the target transcript; the 3' end of these piRNAs is then generated by the combined action of Zucchini and Nibbler (Feltzin et al., 2015; Han et al., 2015; Hayashi et al., 2016; Ipsaro et al., 2012; Mohn et al., 2015; Nishimasu et al., 2012; Wang et al., 2016). After loading onto a PIWI protein, mature piRNAs are then 2'-O-methylated at their 3' end by HEN1 (Horwich et al., 2007; Houwing et al., 2007; Kirino & Mourelatos, 2007; Ohara et al., 2007).

Piwi binds preferentially to primary piRNAs, whose 5' end is specified by Zucchini. These are typically antisense to transposon mRNAs, and characterized by uracil at position 1 (1U). AGO3 binds mainly sense piRNAs, with a strong bias for adenine at position 10 (10A), whereas AUB binds secondary, antisense piRNAs with 1U bias, that have the same sequence as the primary piRNAs. As a result of the ping-pong cycle and the combined slicing activities of AUB and AGO3 on their target mRNAs, the first 10 nt of the secondary piRNAs loaded onto AGO3 and AUB are complementary to each other (Brennecke et al., 2007; Gunawardane et al., 2007; Yamashiro & Siomi, 2018).

Maternal contribution of small RNAs as well as their protein counterpart is fundamental for the fertility of the next generation, as it provides a memory of the targets that need to be silenced, priming the formation of a functional piRNA pathway in the embryo (Czech et al., 2018; Guzzardo et al., 2013; Le Thomas, Marinov, et al., 2014; Le Thomas, Stuwe, et al., 2014). A striking example of the importance of maternal contribution is represented by the hybrid dysgenesis in *Drosophila melanogaster*. The P-element transposon has been lost in the laboratory strain, therefore these flies do not produce piRNAs to silence it. On

the other hand, this transposon is still present in the wild. If a wild type male, carrying the P-element, is crossed with a laboratory female, their offspring will be sterile, as they do not produce piRNAs to counteract the P-element activity, nor they can inherit such piRNAs from the mother. However, if the reciprocal cross is performed, the wild type female produces piRNAs against the P-element and deposits them in the offspring, which is able to use the inherited piRNAs to silence the transposon and therefore is fertile (Brennecke et al., 2008).

Like *Drosophila*, mice have also three PIWI proteins: MIWI, MILI and MIWI2, expressed at different stages of spermatogenesis. In analogy to *Drosophila*, MILI binds the antisense piRNAs and is involved in the primary biogenesis, while MIWI2 is loaded with secondary piRNAs. MIWI2 is transported into the nucleus upon loading with secondary piRNAs and therefore is likely not taking part in the amplification loop (Aravin et al., 2008; De Fazio et al., 2011). There is no evidence for an heterotypic amplification reaction in mouse, although it is possible that MILI fuels an homotypic ping-pong cycle (De Fazio et al., 2011).

Different classes of piRNAs are present throughout murine spermatogenesis. During embryogenesis, uni-strand clusters are transcribed to produce piRNAs, antisense to TEs. After birth, only a fraction of piRNAs still targets TEs, yet the composition differs from the foetal piRNAs, and additionally some of them are derived from endogenous mRNAs; so, these so-called pre-pachytene piRNAs function not only to silence TEs, but possibly also to regulate gene expression. Later in spermatogenesis, in adulthood, at the pachytene stage, virtually no piRNAs target transposons, but the pool of piRNAs present at this stage is mostly derived from long non-coding RNAs (lncRNAs) and intergenic transcripts (Chuma & Nakano, 2012; Ozata et al., 2018).

The silencing activity of PIWI proteins does not always depend on their catalytic activity, as it has been shown that mutations of the catalytic residues of MIWI2 in mouse or Piwi in *Drosophila* do not affect transposon silencing nor fertility (De Fazio et al., 2011; Saito et al., 2010). These two PIWI proteins are both nuclear, and trigger chromatin modification upon recognition of a nascent transcript. It is therefore likely that, for this reaction, target cleavage is not required.

Despite their major role in germ cells, it has recently become clear that piRNAs do have also additional roles. Somatic piRNAs exert a defensive mechanism against TEs in arthropods, indicating how this function is not exclusively germline specific (Lewis et al., 2018). Some invertebrates, such as mosquitoes, also use piRNAs to exert an antiviral response in the soma (Miesen et al., 2015; Morazzani et al., 2012; Schnettler et al., 2013). As germ cells in several organisms originate from stem cells, it is not extremely surprising that piRNAs as well as PIWI pathway components are expressed also in stem cells of various animals. Furthermore, seemingly expression of PIWI proteins decreases as differentiation progresses, suggesting a role for the piRNA pathway in stem cell function, specifically in those organisms with regenerative potential, such as the flatworms planarians (van Wolfswinkel, 2014).

Phase separated structures and small RNA pathways

From studies in *Drosophila*, zebrafish, *C. elegans* and mammals, it is clear that an important fraction of the RNAi pathways in the germline takes place in the cytoplasm, in perinuclear membraneless compartments (Gao & Arkov, 2013; Hashimoto et al., 2004; Houwing et al., 2007; Seydoux, 2018; Siomi et al., 2011; Updike & Strome, 2010; Voronina, 2013). Such perinuclear structures take different names in different organisms, but they share subcellular localization as well as function. Moreover, they contain the majority of the factors involved in the ping-pong amplification loop and in post-transcriptional gene silencing, such as Argonaute and Tudor proteins as well as RNA helicases, besides many different RNA molecules (Gao & Arkov, 2013).

In *Drosophila* germ cells, electron-dense perinuclear granules form the nuage, localized in the vicinity of nuclear pores. The nuage functions as a surveillance machinery, scans the RNAs exported from the nucleus to identify and target transcripts complementary to piRNAs for silencing (Siomi et al., 2011).

In *C. elegans*, the perinuclear granules in the germline are called P granules. Their formation requires the proteins PGL-1 and PGL-3 (Hanazawa et al., 2011; Updike & Strome, 2010). The mRNA surveillance is likely taking place in P granules, as PRG-1, CSR-1 as well as WAGO proteins (see below) colocalize in these perinuclear structures (Batista et al., 2008; Claycomb et al., 2009; Gu et al., 2009; Kamminga et al., 2012).

Similarly to nuage, P granules are localized in the vicinity of nuclear pores, so that each transcript exported from the nucleus can be promptly identified as self or non-self and consequently sorted to be silenced or expressed, probably by routing it to the neighbouring mutator foci in the former case (see below) (Phillips et al., 2012).

Nuage as well as P granules are formed upon liquid-liquid phase separation (Brangwynne et al., 2009) and have the function of specifying germ cell lineage (Gao & Arkov, 2013; Voronina, 2013). Such structures have in fact liquid-like properties, yet, they separate from the surrounding cytoplasm. In this way, they can achieve a very high protein concentration. Also, the factors involved in RNAi are very close to each other, hence their interaction can take place extremely fast, ensuring a very rapid and efficient silence response against "foreign" transcripts.

Nuclear RNAi and epigenetic silencing

In some cases, RNA interference can affect the chromatin structure of the target locus, ensuring not only post-transcriptional, but also transcriptional gene silencing. Transcriptional gene silencing involves the deposition of silencing marks at the genomic level. These epigenetic marks, in certain cases, can also be transmitted transgenerationally.

The RNA-induced transcriptional silencing (RITS) complex was described in *Schizosaccharomyces pombe* (Bühler et al., 2006; Moazed et al., 2006; Verdel & Moazed, 2005). RITS is composed by AGO1, the GW domain protein Tas3 and the chromodomain protein Chp1. This complex is guided by siRNAs to highly repetitive, pericentromeric regions, in order to induce heterochromatin formation. Upon RITS association with the centromeres, the H3K9 methyltransferase Clr4 is recruited to the same genomic location and will in turn produce H3K9 methylation. H3K9 methylation is also required for Chp1 binding, suggesting that the transcriptional silencing mediated by RITS and heterochromatin formation are interdependent. Silencing of pericentromeric regions is established co-transcriptionally (co-transcriptional gene silencing, CTGS), however, the heterochromatic status can be later maintained without transcription (Bühler et al., 2006; Castel & Martienssen, 2013; Luteijn & Ketting, 2013; Martienssen & Moazed, 2015; van Wolfswinkel & Ketting, 2010).

Similar to CTGS of *S. pombe*, *Arabidopsis thaliana* also has a similar silencing mechanism named RNA-directed DNA methylation (RdDM). Repetitive DNA elements, such as transposons, are transcribed by RNA polymerase IV and used as substrate from the RNA-dependent RNA polymerase 2 (RDR2) to produce long dsRNAs. These dsRNAs are then processed into mature 24 nt siRNAs, that will be exported into the cytoplasm and loaded onto AGO4. After loading, AGO4 will shuttle back into the nucleus, where it will recognize nascent transcripts of RNA polymerase V, mostly intergenic non-coding transcripts. These loci will be *de novo* methylated by the cytosine methyltransferase DRM2 to induce heterochromatin formation (Castel & Martienssen, 2013; Malone & Hannon, 2009).

In animals, the nuclear RNAi pathway appears to operate mainly in the germline. Beyond post-transcriptional gene silencing, in fact, PIWI proteins can also induce *de novo* histone methylation at transposon loci to ensure transcriptional gene silencing (Aravin et al., 2008; Aravin et al., 2007; Kuramochi-Miyagawa et al., 2008). In *Drosophila*, loaded Piwi can recognize a nascent transcript and induce H3K9 trimethylation, helped by Panoramix/Silencio and Gtsf1/Asterix (Dönertas et al., 2013; Muerdter et al., 2013; Ohtani et al., 2013; Sienski et al., 2015; Sienski et al., 2012; Yu et al., 2015). Moreover, Piwi interacts with the histone H1 to influence its localization at the TE locus, to regulate chromatin accessibility, thus reinforcing its silencing (Iwasaki et al., 2016).

In the ciliates *Tetrahymena* and *Paramecium*, there is a peculiar nuclear RNAi pathway, which leads to the so-called targeted genome elimination. These unicellular organisms are characterized by a somatic macronucleus and a germline micronucleus. During the vegetative state, the micronucleus is transcriptionally inactive and gene expression is derived only from the macronucleus. However, upon sexual reproduction, the micronucleus undergoes meiosis, the macronucleus is eliminated and, during this timeframe, small non-coding RNAs, named scanRNAs, are produced. ScanRNAs are 25-30 nt long, they associate with an Argonaute protein of the PIWI clade, Twi1, and target for elimination the former macronucleus, thereby helping establishing the new micronucleus in the zygote, as well as new macronucleus, following elimination of the micronucleus-specific sequences (Castel & Martienssen, 2013; Malone & Hannon, 2009; van Wolfswinkel & Ketting, 2010).

Other proteins in RNAi

Another class of proteins involved in many RNAi processes comprises Tudor proteins. Tudor proteins have a general role in RNA biology and participate also in small RNA pathways, particularly the piRNA pathway (Pek et al., 2012). Upon loss of a Tudor protein, the phenotype can also be quite mild, possibly because of redundancy with other proteins of the same family (Siomi et al., 2010; Pek et al., 2012).

The first Tudor protein to be identified was TUDOR (TUD) in *Drosophila*, which is required for formation of the germ plasm and for female fertility (Boswell & Mahowald, 1985). TUD also localizes to the nuage, where it interacts with AUB and AGO3, promoting piRNA association as well as their localization to the nuage (Arkov et al., 2006; Aravin et al., 2001).

The Tudor domain is a motif of approximately 60 amino acids, that recognizes (symmetrically di-) methylated arginine or lysine residues of their interacting proteins (Pek et al., 2012; Ponting, 1997). Many Tudor proteins contain multiple Tudor domains, so possibly they act as a platform to recruit PIWI proteins as well as other factors involved in the piRNA pathway to the nuage.

For example, in *Drosophila* germ cells, Piwi interacts with PRMT-5/Capsuleen, which is an arginine methyltransferase that, together with MEP-50/Valois, is responsible for symmetric methylation of arginine residues present in RG context at the N-terminus of Piwi, not only in *Drosophila* but also in several other organisms (Nishida et al., 2009; Pek et al., 2012; Siomi et al., 2010). In this way, Piwi can be recognized and bound by TUD (Anne & Mechler, 2005; Anne et al., 2007; Kirino et al., 2009; Nishida et al., 2009; Vagin et al., 2009). The Tudor domain protein Qin/Kumo is believed to be the main factor driving nuage assembly; it interacts with AUB, AGO3, the DEXD-box helicases Vasa and Spindle-E, all factors involved in the ping-pong cycle. Qin/Kumo in particular maintains the antisense bias of AUB, whereas an analogous function is exerted by yet another Tudor protein, Krimper, to maintain the sense bias of AGO3 (Czech et al., 2018; Sato et al., 2015; Webster et al., 2015).

In mouse, the Tudor proteins TDRD7 and TDRD1, and the latter also in zebrafish, localize to nuage and are required for transposon silencing, but not for piRNA biogenesis (Chen et

al., 2011; Pek et al., 2012; Siomi et al., 2010). It is therefore striking that Tudor proteins are essential for piRNA biogenesis and transposon silencing as well as for germ cells development in several organisms.

Small RNAs in *Caenorhabditis elegans*

The nematode *C. elegans* has been and is currently being widely used as a model organism to study RNAi and its related pathways, because it is easy to grow and manipulate and has a very fast life cycle (Brenner, 1974). Besides miRNAs, *C. elegans* has three main small RNA classes: 21U, 22G and 26G RNAs. Target silencing is usually initiated by a so-called primary Argonaute protein loaded with a small RNA molecule, either a 21U or 26G RNA, and then it is maintained by secondary Argonaute proteins loaded with 22G RNAs, that are responsible for amplifying the silencing reaction (Almeida et al., 2019; Billi et al., 2014). The genome of *C. elegans* encodes 27 Argonaute proteins, although some of them have redundant functions and not all of them have been extensively characterized yet. They can be classified not only according to their clade, but also according to the class of small RNAs that they bind (Yigit et al., 2006), which I will now describe.

21U RNAs

21U RNAs are considered the piRNAs of *C. elegans*. Despite being shorter than piRNAs in other organisms (21 versus 24-30 nt), 21U RNAs share all the conserved features of piRNAs. In fact, they are germline enriched, characterized by a 5' monophosphate and 1U bias and 2'-O-methylation at their 3' end, and are bound by a protein of the PIWI clade, PRG-1 (Batista et al., 2008; Billi et al., 2012; Das et al., 2008; Kamminga et al., 2012; Luteijn & Ketting, 2013; Montgomery et al., 2012; Ruby et al., 2006) (**Figure 3**).

The vast majority of 21U RNAs is transcribed from two large clusters on chromosome IV as single precursors by RNA Pol II, and belongs to the so-called type I (Batista et al., 2008; Ruby et al., 2006) (**Figure 3**). Upstream of the transcription start site (~40 nt), there is a 8 nt motif (Ruby motif, CTGTTTCA) that is recognized by transcription factors of the Forkhead family and that is required for transcription (Cecere et al., 2012; Ruby et al., 2006). Another small motif (YRNT) defines the transcription start site, being T the equivalent of U at position 1 of the mature 21U RNA, and transcription starts exactly two

nucleotides upstream (Ruby et al., 2006) (**Figure 3**). Recently, both forward genetic and RNAi-based screens have identified novel factors, required for transcription and biogenesis of these 21U RNAs (Goh et al., 2014; Kasper et al., 2014; Weick et al., 2014). Some of these novel factors assemble into a complex, the upstream sequence transcription complex (USTC), that is required for transcription of type I 21U RNA genes. The USTC is formed by PRDE-1, SNPC-4, TOFU-4 and TOFU-5 (Weng et al., 2018). PRDE-1 binds the piRNA clusters on chromosome IV, where it colocalizes with the transcription factor SNPC-4 (Kasper et al., 2014; Weick et al., 2014); TOFU-4 and TOFU-5 (Twenty-One-U Fouled-Ups-4 and -5) have been previously identified in an RNAi-based screen as novel factors involved in the biogenesis of 21U RNAs (Goh et al., 2014) (**Figure 3**).

Despite the majority of 21U RNAs belongs to the type I, later a different, minor class of 21U RNAs has been identified and classified as type II. Type II 21U RNAs originate from promoter sequences driving regular genes, and are interspersed throughout the whole genome. Their sequences mostly correspond to the 5' end of endogenous genes, probably due to premature transcription termination or pausing of RNA Pol II at the transcription start sites, and are transcribed bi-directionally. Such short transcripts will be further processed into mature 21U RNAs, if associated with the YRNT small motif (Gu et al., 2012) (**Figure 3**).

The precursors of 21U RNAs are capped at the 5' end and are mostly 26 nt long. They have two extra nucleotides at the 5' end, where transcription starts, and few more at the 3' end (Gu et al., 2012). After being exported from the nucleus to the cytoplasm, 21U RNAs will be further processed by decapping, exerted by a yet unknown factor, as well as exonucleolityc cleavage by PARN-1, to trim the extra nucleotides at the 3' end (Tang et al., 2016). Other factors involved in processing of the 21U RNA precursors are TOFU-1 and -2 (Goh et al., 2014), PID-1 (de Albuquerque et al., 2014) and its interacting proteins PID-3, ERH-2, TOFU-6 and IFE-3 (Goh et al., 2014; Rodrigues et al., 2018) (**Figure 3**).

21U RNAs are eventually methylated at their 3' ends by HENN-1, the nematode homolog of the methyltransferase HEN1, presumably after being loaded on the PIWI protein PRG-1 (Billi et al., 2012; Kamminga et al., 2012; Montgomery et al., 2012) (**Figure 3**).

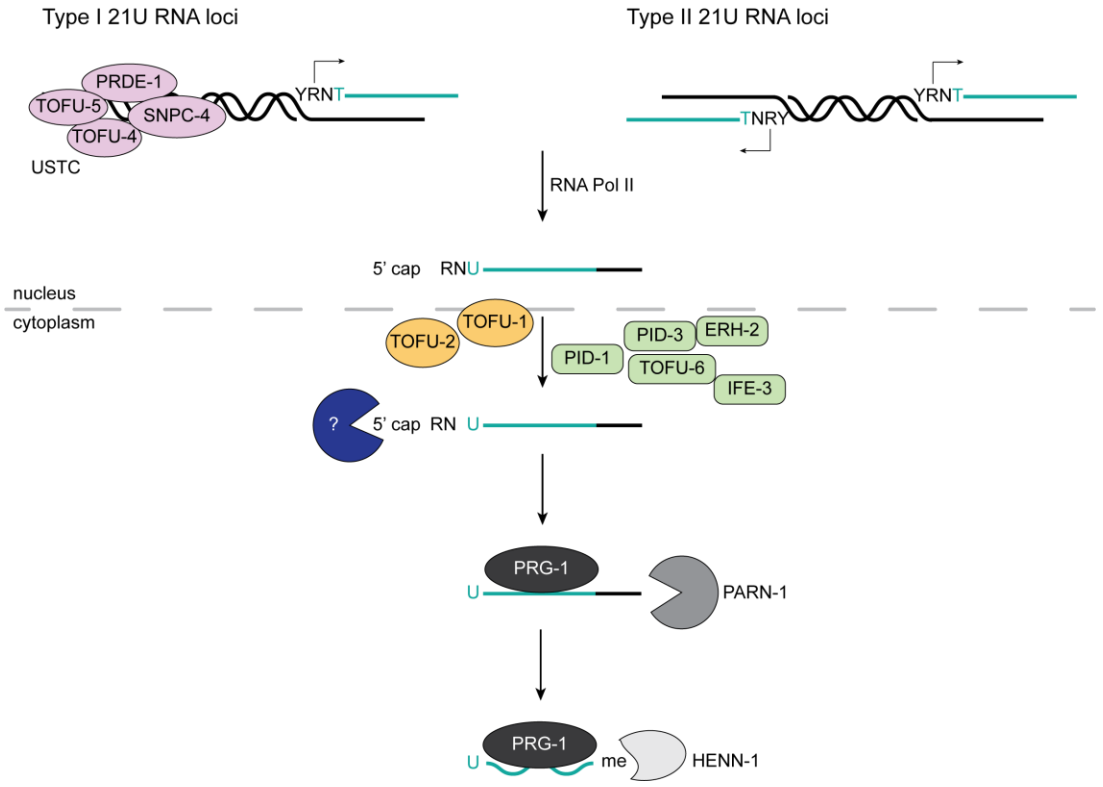


Figure 3. 21U RNAs are transcribed as single genes from RNA Pol II. The majority of 21U RNAs belongs to type I; they are produced from two main clusters on chromosome IV and are characterized by an upstream motif, recognized and bound by the USTC. Type II 21U RNAs are not preceded by this upstream motif, are transcribed bi-directionally and are found throughout the whole genome. Both types of 21U RNAs are characterized by a small motif (YRNT), marking the transcription start site. These precursors are 5' capped and ~26 nt long. After being exported into the cytoplasm, many factors are required to complete the maturation of these precursors. After 5' decapping, and trimming of the extra nucleotides at the 3' end by PARN-1, the mature 21U RNA is loaded onto the cognate PIWI protein, PRG-1, and methylated at the 3' end by HENN-1 to finalize the maturation.

26G RNAs

26G RNAs are produced by the so-called ERI complex. The ERI complex contains multiple proteins, essential for 26G RNA biogenesis, including the RdRP RRF-3 and the *C. elegans* Dicer homolog (DCR-1) (Billi et al., 2014). The Tudor protein ERI-5 interacts with a conserved CHHC zinc finger protein, GTSF-1, to efficiently bring RRF-3 to DCR-1 and ensure an effective production of 26G RNAs (Almeida et al., 2018; Duchaine et al., 2006; Thivierge et al., 2012). Although the biogenesis of 26G RNAs has not been well investigated yet, RRF-3 possibly produces a long dsRNA, which is eventually processed

by DCR-1 into shorter dsRNAs (Billi et al., 2014). The characteristics of 26G RNAs can be guessed from their name: they are 26 nt long and have a 5' bias for guanine at position 1 (1G) (Conine et al., 2010; Han et al., 2009; Yigit et al., 2006). 26G RNAs are monophosphorylated at their 5' ends and are germline enriched (Billi et al., 2014; Han et al., 2009; Ruby et al., 2006). Depending on which Argonaute they are loaded onto, 26G RNAs can be further classified in two groups. The 26G RNAs loaded onto the Argonaute protein ERGO-1 are enriched in oocytes and embryos, are also methylated by HENN-1 at their 3' end (Billi et al., 2012; Gent et al., 2010; Han et al., 2009; Kamminga et al., 2012; Montgomery et al., 2012; Vasale et al., 2010), and their accumulation depends on mutator proteins (see "22G RNAs" for more information on mutator proteins) (Zhang et al., 2011). On the other hand, 26G RNAs loaded onto the Argonaute proteins ALG-3/-4 are enriched in sperm, and are not 3' methylated (Conine et al., 2010, 2013; Han et al., 2009).

22G RNAs

After target recognition mediated by a 21U or 26G RNA and its cognate Argonaute protein, an RdRP, EGO-1 or RRF-1, is recruited to the target transcript, using it as a template to produce a class of antisense secondary siRNAs, perfectly complementary to the target. These secondary siRNAs are characterized by a 5' triphosphate and 1G bias and are mostly 22 nt long, therefore they are also known as 22G RNAs. 22G RNAs are mostly germline enriched and can have a dual function: depending on their partner Argonaute protein, they can either function to amplify and reinforce the silencing reaction, or they can promote target expression, thereby counteracting the silencing activities (Billi et al., 2014; Gu et al., 2009) (**Figure 4**).

The synthesis of 22G RNAs, that will be later loaded onto WAGO proteins, takes place in mutator foci (Phillips et al., 2012). The mutator foci are germline specific perinuclear granules, which contain several factors involved in the 22G RNA biology, and take their name from mutator proteins. Mutator proteins are a class of nematode specific proteins, that are essential for transposon silencing. In fact, upon loss of one of the mutator proteins, transposons in the germline get reactivated and the worms are resistant to RNAi, as the production of 22G RNAs, required for amplification of the silencing reaction is abolished (Ketting et al., 1999; Phillips et al., 2014; Zhang et al., 2011). Mutator proteins are

expressed both in soma and germ cells; in the latter, they localize to mutator foci, that are most prominent in the mitotic phase and transition zone (Phillips et al., 2012). The assembly of these foci specifically requires MUT-16, a Q/N rich protein, that recruits the other components: MUT-2/RDE-3, a nucleotidyl transferase; MUT-7, a 3'–5' exonuclease; MUT-14, a DEAD-box RNA helicase; RDE-2/MUT-8 and MUT-15/RDE-5, of unknown function (Billi et al., 2014; Chen et al., 2005; Ketting et al., 1999; Phillips et al., 2012; Tijsterman et al., 2002; Tops et al., 2005; Vastenhouw et al., 2003). Also the RdRP RRF-1 localizes to mutator foci, therefore they are likely the production centre of secondary 22G RNAs (Phillips et al., 2012). Mutator foci are neighbouring with P granules, the *C. elegans* equivalent of nuage, indicating a possible communication between the two types of granules, in order to efficiently recognize and target mRNAs for silencing (Phillips et al., 2012).

Small RNA-mediated silencing pathways in *C. elegans*

The PRG-1 pathway

21U RNAs are loaded onto the PIWI protein PRG-1, which localizes to perinuclear structures known as P granules, to form a complex to scan transcripts and silence them upon recognition. The genome of *C. elegans* codes for another PIWI protein, PRG-2, although it is probably not taking part in the 21U RNA pathway, as *prg-2* mutants do not show any defects (Batista et al., 2008; Wang & Reinke, 2008) and *prg-2* has been recently classified as a pseudogene.

21U RNAs do not show obvious complementarity to transposon transcripts. Indeed, only one class of transposons, Tc3 (Das et al., 2008), is directly targeted by 21U RNAs while the majority of the ~15.000 21U RNAs lacks obvious targets. It has been shown that the pairing between the 21U RNA and its target allows up to 4 mismatches, therefore it is difficult to predict the transcripts that will be recognized, although this represents definitely an advantage to recognize any foreign sequence and silence it (Bagijn et al., 2012). Nonetheless, 21U RNAs do require a perfect pairing between positions 2 and 7, supporting the existence of a piRNA seed sequence (Shen et al., 2018; Zhang et al., 2018), in analogy to the miRNA target recognition paradigm. Considering such targeting rules, it is now clear that 21U RNAs target also endogenous genes as well as other TEs (Bagijn et al., 2012; de

Albuquerque et al., 2015; Gu et al., 2012; Lee et al., 2012; Shen et al., 2018; Zhang et al., 2018). For instance, one of the 21U RNAs, *21ux-1*, targets an endogenous gene, *xol-1*, to ensure dosage compensation for a proper sexual development and hermaphrodite viability (Tang et al., 2018).

Contrary to other organisms, *prg-1* mutants are not sterile (Batista et al., 2008; Cox et al., 1998; Das et al., 2008; Wang & Reinke, 2008), although they lose germ cells over time (mortal germline phenotype, Mrt) (Simon et al., 2014), and they do not show dramatic transposon reactivation (Das et al., 2008). Hence, PRG-1 and 21U RNAs are only partially required for TEs silencing, whereas 22G RNAs play a major role in controlling transposon activity in the germline, together with mutator and WAGO proteins (Das et al., 2008; de Albuquerque et al., 2015; Kasper et al., 2014; Ketting et al., 1999; Sijen & Plasterk, 2003; Tabara et al., 1999). Furthermore, PRG-1 endonucleolityc activity is not required for target silencing (Bagijn et al., 2012), similar to what has been shown for Piwi in *Drosophila* (Saito et al., 2010) and MIWI2 in mouse (De Fazio et al., 2011).

Silencing mediated by secondary WAGOs and RNAe

After target recognition mediated by a primary Argonaute, both the 21U and the 26G RNA pathways trigger the downstream production of secondary siRNAs (22G RNAs). 22G RNAs are then loaded onto secondary Argonaute proteins, such as WAGO-1 and WAGO-9/HRDE-1, to amplify the silencing reaction (Gu et al., 2009; Pak & Fire, 2007; Yigit et al., 2006) (**Figure 4**). WAGO-1 localizes to cytoplasmic P granules and is responsible for the post-transcriptional gene silencing, while HRDE-1, a germline specific WAGO, shuttles into the nucleus after being loaded with a 22G RNA, where it is involved in transcriptional gene silencing (Buckley et al., 2012), helped by the nuclear RNAi machinery. HRDE-1, similar to the somatic nuclear WAGO protein NRDE-3/WAGO-12, uses the 22G RNAs to recognize and bind a perfectly complementary nascent transcript. Upon target recognition, NRDE-1 binds the chromatin of the transcribed locus, together with NRDE-2 and NRDE-4, to induce H3K9 trimethylation and heterochromatin formation (Burkhart et al., 2011; Burton et al., 2011; Guang et al., 2008; Guang et al., 2010; Kasper et al., 2014). To accomplish transcriptional silencing, also the HP1 ortholog HPL-2 and the histone methyltransferases SET-25 and SET-32 are required for H3K9 trimethylation. This

form of silencing is very stable across generations and can become independent of PRG-1. Thereafter, it is also referred to as RNAe (RNA induced epigenetic silencing) (**Figure 4**). RNAe can be transmitted for tens of generations without the initial silencing input carried by PRG-1, but it still depends on 22G RNAs and WAGO proteins, such as HRDE-1 and WAGO-1/-2/-3. RNAe is established in a rather stochastic way, as different individuals can either silence or express a certain 21U RNA target, despite being genetically identical (Ashe et al., 2012; Bagijn et al., 2012; Lee et al., 2012; Luteijn & Ketting, 2013; Luteijn et al., 2012; Shirayama et al., 2012). This is nicely exemplified in a strain in which a 21U RNA target, the so-called 21U sensor transgene, is mildly activated, due to absence of HENN-1. In a seemingly random manner, in isolated individuals, this mild activation is lost, and this silent state is then independent from PRG-1; hence, RNAe (Luteijn et al., 2012). The reasons behind this stochasticity are yet unclear.

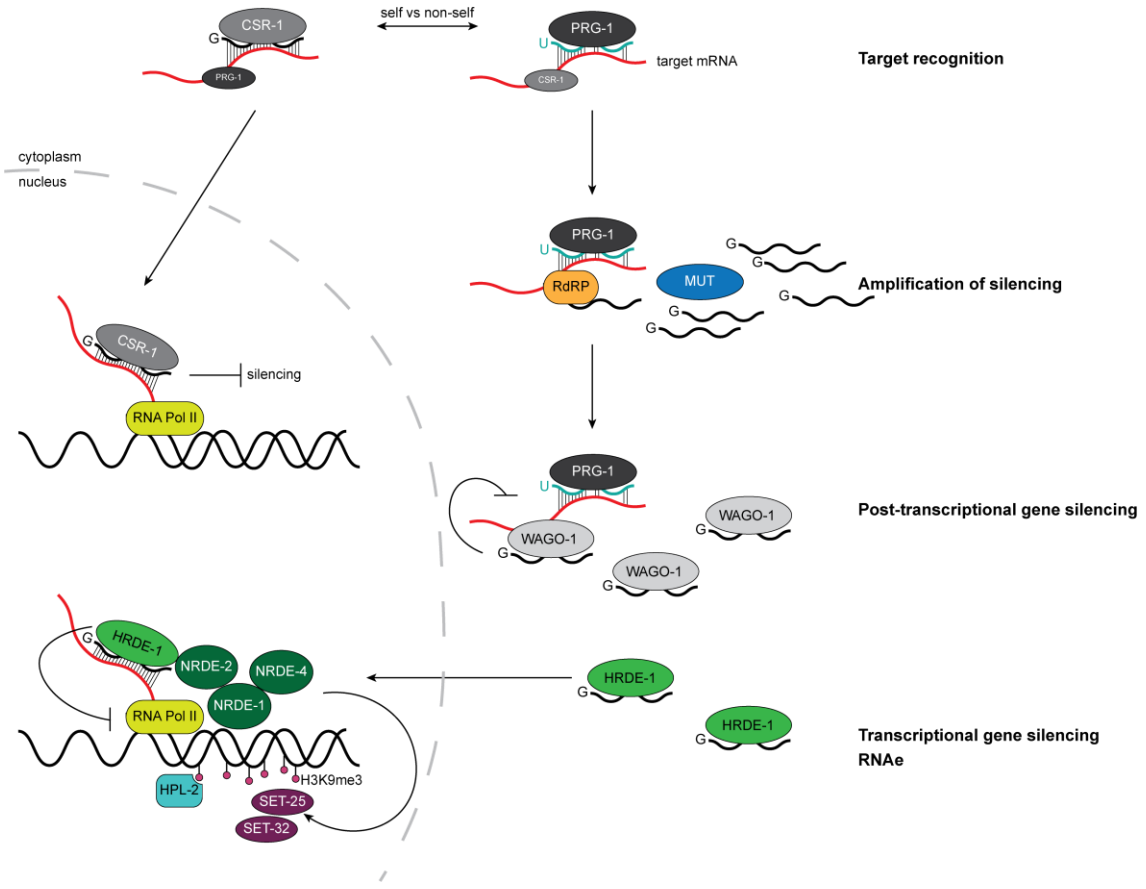


Figure 4. PRG-1, loaded with its 21U RNA co-factor, scans transcripts and induces production of secondary 22G RNAs, by an RdRP and mutator proteins, upon recognition of a foreign transcript. These 22G RNAs are then loaded onto WAGO proteins, such as WAGO-1, to ensure post-transcriptional gene silencing, and HRDE-1, which shuttles into the nucleus. HRDE-1 loaded with 22G RNAs can thus recognize a nascent

transcript and, together with the nuclear RNAi factors NRDE-1/-2/-4, triggers the deposition of heterochromatic marks at the genomic locus by the methyltransferases SET-25/-32. This transcriptional gene silencing is reinforced by HPL-2 and can become independent from PRG-1 and self-sustainable (RNAe). The silencing initiated by PRG-1 is counteracted by CSR-1, which is loaded with 22G RNAs, to recognize endogenous transcripts and ensure their expression.

Heritable epigenetic silencing can also be triggered by environmental stimuli, such as high temperature, starvation, growth conditions, even though the stability of these events seems to be limited to few generations (Klosin et al., 2017; Lev et al., 2018; Rechavi et al., 2014). Nevertheless, such heritable silencing is a remarkable feat, that brings potential for heritable adaptation to environmental conditions, in absence of genetic mutations.

The CSR-1 pathway

As a result of mismatch tolerance between 21U RNAs and their targets, the pool of 21U RNAs of *C. elegans* is potentially able to target almost any DNA sequence for silencing, including endogenous genes. Hence, in order to ensure gene expression, a mechanism to counteract this silencing activity must exist. The Argonaute protein CSR-1 is loaded with 22G RNAs complementary to expressed genes, implementing a surveillance memory to counteract the silencing activity mediated by PRG-1 and WAGO proteins (Claycomb et al., 2009; Conine et al., 2013; Gu et al., 2009; Lee et al., 2012; Seth et al., 2013; Shirayama et al., 2012; Wedeles et al., 2013) (**Figure 4**). Possibly, in the germline of *C. elegans,* gene expression is thus the result of a balance between PRG-1 and CSR-1 activities. Self-silencing is avoided by CSR-1, that protects transcripts of expressed genes from the silencing initiated by PRG-1 (Shen et al., 2018). This epigenetic memory is transmitted from the parents to the next generation and is mediated not only by chromatin marks, but also by the various small RNA populations (Almeida, 2019; de Albuquerque et al., 2014, 2015; Luteijn et al., 2012; Minkina & Hunter, 2018; Phillips et al., 2015; Xu et al., 2018). Indeed, if such memory of silenced and expressed genes is impaired, the balance between these two mechanisms is altered, resulting in erroneous gene silencing and defects in gonads development (de Albuquerque et al., 2015; Phillips et al., 2015). Moreover, upon depletion of CSR-1, the transcripts that would normally be targeted by the CSR-1 pathway,

are bound by additional 21U RNAs, suggesting that CSR-1 indeed directly antagonizes PRG-1 (Shen et al., 2018).

CSR-1 22G RNAs are produced by the RdRP EGO-1 (Claycomb et al., 2009; Gu et al., 2009) and are uridylated at their 3' end by the nucleotidyltransferase CDE-1. The latter targets them for degradation. This mechanism is probably required to avoid the accumulation of CSR-1 22G RNAs and the possible misrouting into the WAGO silencing proteins (van Wolfswinkel et al., 2009). 22G RNAs that are uridylated by CDE-1 are also loaded onto yet a different cytoplasmic WAGO protein, namely WAGO-4, to target germline-expressed genes. WAGO-4 also plays a role in inheritance of RNAi via the maternal lineage (Xu et al., 2018). On the other hand, CSR-1, together with ALG-3/-4 (see below), ensures gene expression in spermatocytes. While ALG-3/-4 are eliminated later during spermatogenesis, CSR-1 is present in mature sperm (Conine et al., 2010, 2013). Possibly, CSR-1 is responsible to transmit a memory of paternal gene expression to the next generation (Conine et al., 2013).

An additional role for the CSR-1 pathway is the chromosome segregation. In fact, CSR-1 is the only Argonaute protein essential for viability. *csr-1* mutants are almost completely sterile; yet, they can lay few embryos, that die because of chromosome segregation defect (Claycomb et al., 2009; Wedeles et al., 2013a; Yigit et al., 2006). Possibly, though, these chromosome segregation defects are an indirect effect of *csr-1* mutation. In fact, loss of CSR-1 causes severe defects in the microtubule assembly, due to altered expression of a kinesin-13 microtubule depolymerase, normally strongly targeted by CSR-1/22G RNAs (Gerson-Gurwitz et al., 2016).

As CSR-1 binds chromatin, it is possible that CSR-1, loaded with 22G RNAs, can directly recognize and bind a complementary nascent transcript, thereby promoting its expression (Seth et al., 2013).

Besides CSR-1 activity, periodic A_n/T_n-clusters (PATC) can also counteract the silencing mediated by PRG-1 (Frøkjær-Jensen et al., 2016). Indeed, PATC are particularly enriched in genes endogenously expressed in the germline, and this correlates with low production of 22G RNAs on their transcripts, supporting their role in promoting expression (Zhang et

al., 2018). This could be therefore used as a tool to favour transgene expression in the germline.

ALG-3/-4 and ERGO-1 pathways

As already mentioned, 26G RNAs exert different functions, depending on their partner Argonaute protein. ALG-3 and -4 are two seemingly redundant Argonaute proteins that, upon loading with 26G RNAs, target spermatogenesis-enriched transcripts, preferentially by binding at their 5' and 3' ends (Almeida et al., 2019; Conine et al., 2010). Loss of ALG-3/-4 26G RNAs causes sterility at 25 °C, due to failure in spermatogenesis and spermiogenesis, as well as a Him phenotype (high incidence of males), indicating X chromosome missegregation (Billi et al., 2014; Conine et al., 2010). This phenotype is somehow related with a role for ALG-3/-4 and their partner 26G RNAs in promoting spermatogenesis as well as transmitting a paternal memory of gene expression to the next generation (Conine et al., 2013). Such paternal memory is possibly also mediated by CSR-1/22G RNA complexes, which are also enriched in mature sperm and target expressed genes, acting downstream of ALG-3/-4 (Conine et al., 2013). However, these relations are built purely on genetics, and the molecular mechanisms involved have not been resolved yet.

ERGO-1 26G RNAs, on the other hand, target certain transcripts for silencing in the oogenic gonad as well as in embryos. In particular, ERGO-1 targets are repeat-rich, recently duplicated genes, pseudogenes and lincRNAs (Almeida et al., 2019; Billi et al., 2014; Vasale et al., 2010). Furthermore, ERGO-1 targets are generally not well conserved and have very few introns, often associated with weak splicing signals (Almeida et al., 2019; Newman et al., 2018). After target recognition, ERGO-1 triggers the downstream production of 22G RNAs, that are then loaded onto secondary Argonaute proteins, such as NRDE-3, to reinforce the silencing. NRDE-3 is a nuclear WAGO protein that ensures transcriptional silencing, together with its cofactors NRDE-1, -2 and -4, similarly to the activity of HRDE-1 downstream of PRG-1 (Billi et al., 2014).

Loss of ERGO-1 26G RNAs is responsible for enhanced RNAi response (Eri phenotype), due to a decrease in competition among WAGO proteins, for being loaded with secondary 22G RNAs (Billi et al., 2014; Duchaine et al., 2006; Gu et al., 2009; Yigit et al., 2006).

Given their germline enrichment and silencing function, it has been recently proposed that 26G RNAs could exert a function similar to piRNAs. Therefore, in *C. elegans*, piRNAs should not be unequivocally identified with the 21U RNAs, but also 26G RNAs should be considered (Almeida et al., 2019).

Open questions

The small RNA pathways are extremely relevant for many aspects of biology, as they can influence all physiological processes, from transposon silencing and maintenance of genome integrity in germ cells, to gene expression and their silencing. We use the nematode *Caenorhabditis elegans* as a model to illustrate the complexity of, and better understand, the mechanisms behind recognition of self versus non-self, which are essential to ensure survival of a species, as demonstrated by the sterility caused by mutations in pathways that are required to control TEs activity, such as PIWI mutants.

First of all, we asked: How important is the parental contribution of small RNAs? In chapter 2, we show that in *C. elegans*, parental small RNAs are necessary to enforce proper gene expression and to guarantee successful gonad development as well as fertility of the next generation (Almeida et al., 2019; de Albuquerque et al., 2015; Phillips et al., 2015). Our findings are in line with previous studies, that showed the importance of maternal proteins and RNAs in *Drosophila*, zebrafish and mammals (Brennecke et al., 2008; Czech et al., 2018; Despic & Neugebauer, 2018; Heard & Martienssen, 2014; Kaaij et al., 2013), but also extend beyond that, as we demonstrate that these small RNAs are not only needed to establish silencing of the proper targets, but also to prevent erroneous gene silencing.

Second, we wanted to know: Are maternally provided 21U RNAs sufficient to establish *de novo* silencing in the offspring? Our results in chapter 3 show that the answer is affirmative. Furthermore, we show that maternally deposited 21U RNAs can establish RNAe, the above described very stable form of heritable silencing. RNAe is associated with heterochromatin formation at the silenced locus and with production of 22G RNAs throughout the whole target transcript. Yet, in our experiments, we show that a portion of the silenced locus is not producing 22G RNAs. Possibly, this is the result of CSR-1 licensing, supporting its antagonizing function to PRG-1.

Third, in chapter 4, we address how the 21U RNA pathway and RNAe can be connected. To answer this question, we characterized a novel factor, PID-2, which is required to establish *de novo* silencing and to boost the production of downstream 22G RNAs. Interestingly, this function does not seem to be specific to the PRG-1 pathway. Moreover, we identified two novel Tudor proteins, PID-4 and PID-5; both interact with PID-2 and are required for PID-2 function. Finally, we find that PID-2, PID-4 and PID-5 are all required for normal germline health across generations, indicating the relevance of these proteins for reproduction in *C. elegans*.

References

Akkouche, A., Mugat, B., Barckmann, B., Varela-Chavez, C., Li, B., Raffel, R., … Chambeyron, S. (2017). Piwi Is Required during Drosophila Embryogenesis to License Dual-Strand piRNA Clusters for Transposon Repression in Adult Ovaries. *Molecular Cell*, *66*(3), 411–419.e4. https://doi.org/10.1016/j.molcel.2017.03.017

Akulenko, N., Ryazansky, S., Morgunova, V., Komarov, P. A., Olovnikov, I., Vaury, C., … Kalmykova, A. (2018). Transcriptional and chromatin changes accompanying de novo formation of transgenic piRNA clusters. *RNA*, *24*(4), 574–584. https://doi.org/10.1261/rna.062851.117

Almeida, M. V., Andrade-Navarro, M. A., & Ketting, R. F. (2019). Function and Evolution of Nematode RNAi Pathways. *Non-Coding RNA*, *5*(1), 8. https://doi.org/10.3390/ncrna5010008

Almeida, M. V., de Jesus Domingues, A. M., & Ketting, R. F. (2019). Maternal and zygotic gene regulatory effects of endogenous RNAi pathways. *PLOS Genetics*, *15*(2), e1007784. https://doi.org/10.1371/journal.pgen.1007784

Almeida, M. V., Dietz, S., Redl, S., Karaulanov, E., Hildebrandt, A., Renz, C., … Ketting, R. F. (2018). GTSF-1 is required for formation of a functional RNA-dependent RNA Polymerase complex in Caenorhabditis elegans. *The EMBO Journal*, *37*(12), e99325. https://doi.org/10.15252/embj.201899325

Andersen, P. R., Tirian, L., Vunjak, M., & Brennecke, J. (2017). A heterochromatin-dependent transcription machinery drives piRNA expression. *Nature*, *549*(7670), 54–59. https://doi.org/10.1038/nature23482

Anne, J., & Mechler, B. M. (2005). Valois, a component of the nuage and pole plasm, is involved in assembly of these structures, and binds to Tudor and the methyltransferase Capsuleen. *Development*, *132*(9), 2167–2177. https://doi.org/10.1242/dev.01809

Anne, J., Ollo, R., Ephrussi, A., & Mechler, B. M. (2007). Arginine methyltransferase Capsuleen is essential for methylation of spliceosomal Sm proteins and germ cell formation in Drosophila. *Development*, *134*(1), 137–146. https://doi.org/10.1242/dev.02687

Aravin, A. A., Naumova, N. M., Tulin, A. V, Vagin, V. V, Rozovsky, Y. M., & Gvozdev, V. A. (2001). Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the D. melanogaster germline. *Current Biology : CB*, *11*(13), 1017–1027. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11470406

Aravin, A. A., Sachidanandam, R., Bourc'his, D., Schaefer, C., Pezic, D., Toth, K. F., … Hannon, G. J. (2008). A piRNA Pathway Primed by Individual Transposons Is Linked to De Novo DNA Methylation in Mice. *Molecular Cell*, *31*(6), 785–799. https://doi.org/10.1016/J.MOLCEL.2008.09.003

Aravin, A. A., Sachidanandam, R., Girard, A., Fejes-Toth, K., & Hannon, G. J. (2007). Developmentally regulated piRNA clusters implicate MILI in transposon control. *Science (New York, N.Y.)*, *316*(5825), 744–747. https://doi.org/10.1126/science.1142612

Arkov, A. L., Wang, J.-Y. S., Ramos, A., & Lehmann, R. (2006). The role of Tudor domains in germline development and polar granule architecture. *Development (Cambridge, England)*, *133*(20), 4053–4062. https://doi.org/10.1242/dev.02572

Armstrong, C. A., & Tomita, K. (2017). Fundamental mechanisms of telomerase action in yeasts and mammals: understanding telomeres and telomerase in cancer cells. *Open Biology*, *7*(3), 160338. https://doi.org/10.1098/rsob.160338

Arribas-Layton, M., Wu, D., Lykke-Andersen, J., & Song, H. (2013). Structural and functional control of the eukaryotic mRNA decapping machinery. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, *1829*(6–7), 580–589. https://doi.org/10.1016/j.bbagrm.2012.12.006

Ashe, A., Sapetschnig, A., Weick, E. M., Mitchell, J., Bagijn, M. P., Cording, A. C., … Miska, E. A. (2012). PiRNAs can trigger a multigenerational epigenetic memory in the germline of C. elegans. *Cell*. https://doi.org/10.1016/j.cell.2012.06.018

Ashley, J., Cordy, B., Lucia, D., Fradkin, L. G., Budnik, V., & Thomson, T. (2018). Retrovirus-like Gag Protein Arc1 Binds RNA and Traffics across Synaptic Boutons. *Cell*, *172*(1–2), 262–274.e11. https://doi.org/10.1016/J.CELL.2017.12.022

Ayarpadikannan, S., & Kim, H.-S. (2014). The Impact of Transposable Elements in Genome Evolution and Genetic Instability and Their Implications in Various Diseases. *Genomics & Informatics*, *12*(3), 98. https://doi.org/10.5808/GI.2014.12.3.98

Babiarz, J. E., Ruby, J. G., Wang, Y., Bartel, D. P., & Blelloch, R. (2008). Mouse ES cells express endogenous shRNAs, siRNAs, and other Microprocessor-independent, Dicer-dependent small RNAs. *Genes & Development*, *22*(20), 2773–2785. https://doi.org/10.1101/gad.1705308

Bagijn, M. P., Goldstein, L. D., Sapetschnig, A., Weick, E.-M., Bouasker, S., Lehrbach, N. J., … Miska, E. A. (2012). Function, Targets, and Evolution of Caenorhabditis elegans piRNAs. *Science*.

Bartel, D. P. (2018). Metazoan MicroRNAs. *Cell*, *173*(1), 20–51. https://doi.org/10.1016/J.CELL.2018.03.006

Batista, P. J., Ruby, J. G., Claycomb, J. M., Chiang, R., Fahlgren, N., Kasschau, K. D., … Mello, C. C. (2008). PRG-1 and 21U-RNAs Interact to Form the piRNA Complex Required for Fertility in C. elegans. *Molecular Cell*. https://doi.org/10.1016/j.molcel.2008.06.002

Baumberger, N., & Baulcombe, D. C. (2005). Arabidopsis ARGONAUTE1 is an RNA Slicer that selectively recruits microRNAs and short interfering RNAs. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(33), 11928–11933. https://doi.org/10.1073/pnas.0505461102

Belancio, V. P., Deininger, P. L., & Roy-Engel, A. M. (2009). LINE dancing in the human genome: transposable elements and disease. *Genome Medicine*, *1*(10), 97. https://doi.org/10.1186/gm97

Berkhout, B. (2018). RNAi-mediated antiviral immunity in mammals. *Current Opinion in Virology*, *32*, 9–14. https://doi.org/10.1016/j.coviro.2018.07.008

Bernstein, E., Caudy, A. A., Hammond, S. M., & Hannon, G. J. (2001). Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature*, *409*(6818), 363–366. https://doi.org/10.1038/35053110

Bessereau, J.-L. (2006). Transposons in C. elegans. *WormBook : The Online Review of C. Elegans Biology*, 1–13. https://doi.org/10.1895/wormbook.1.70.1

Bessereau, J.-L., Wright, A., Williams, D. C., Schuske, K., Davis, M. W., & Jorgensen, E. M. (2001). Mobilization of a Drosophila transposon in the Caenorhabditis elegans germ line. *Nature*, *413*(6851), 70–74. https://doi.org/10.1038/35092567

Billi, A. C., Alessi, A. F., Khivansara, V., Han, T., Freeberg, M., Mitani, S., & Kim, J. K. (2012). The Caenorhabditis elegans HEN1 ortholog, HENN-1, methylates and stabilizes select subclasses of germline small RNAs. *PLoS Genetics*, *8*(4), e1002617. https://doi.org/10.1371/journal.pgen.1002617

Billi, A. C., Fischer, S. E. J., & Kim, J. K. (2014). Endogenous RNAi pathways in C. elegans. *WormBook*, 1–49. https://doi.org/10.1895/wormbook.1.170.1

Boswell, R. E., & Mahowald, A. P. (1985). tudor, a gene required for assembly of the germ plasm in Drosophila melanogaster. *Cell*, *43*(1), 97–104. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3935320

Brancati, G., & Großhans, H. (2018). An interplay of miRNA abundance and target site architecture determines miRNA activity and specificity. *Nucleic Acids Research*, *46*(7), 3259–3269. https://doi.org/10.1093/nar/gky201

Brangwynne, C. P., Eckmann, C. R., Courson, D. S., Rybarska, A., Hoege, C., Gharakhani, J., … Hyman, A. A. (2009). Germline P granules are liquid droplets that localize by controlled dissolution/condensation. *Science (New York, N.Y.)*, *324*(5935), 1729–1732. https://doi.org/10.1126/science.1172046

Brennecke, J., Aravin, A. A., Stark, A., Dus, M., Kellis, M., Sachidanandam, R., & Hannon, G. J. (2007). Discrete Small RNA-Generating Loci as Master Regulators of Transposon Activity in Drosophila. *Cell*, *128*(6), 1089–1103. https://doi.org/10.1016/J.CELL.2007.01.043

Brennecke, J., Malone, C. D., Aravin, A. A., Sachidanandam, R., Stark, A., & Hannon, G. J. (2008). An epigenetic role for maternally inherited piRNAs in transposon silencing. *Science (New York, N.Y.)*, *322*(5906), 1387–1392. https://doi.org/10.1126/science.1165171

Brenner, S. (1974). THE GENETICS OF CAENORHABDITIS ELEGANS. *Genetics*, *77*(1).

Broughton, J. P., Lovci, M. T., Huang, J. L., Yeo, G. W., & Pasquinelli, A. E. (2016). Pairing beyond the Seed Supports MicroRNA Targeting Specificity. *Molecular Cell*, *64*(2), 320–333. https://doi.org/10.1016/J.MOLCEL.2016.09.004

Buckley, B. A., Burkhart, K. B., Gu, S. G., Spracklin, G., Kershner, A., Fritz, H., … Kennedy, S. (2012). A nuclear Argonaute promotes multigenerational epigenetic inheritance and germline immortality. *Nature*, *489*(7416), 447–451. https://doi.org/10.1038/nature11352

Bühler, M., Verdel, A., & Moazed, D. (2006). Tethering RITS to a Nascent Transcript Initiates RNAi- and Heterochromatin-Dependent Gene Silencing. *Cell*, *125*(5), 873–886. https://doi.org/10.1016/J.CELL.2006.04.025

Burkhart, K. B., Guang, S., Buckley, B. A., Wong, L., Bochner, A. F., & Kennedy, S. (2011). A Pre-mRNA–Associating Factor Links Endogenous siRNAs to Chromatin Regulation. *PLoS Genetics*, *7*(8), e1002249. https://doi.org/10.1371/journal.pgen.1002249

Burton, N. O., Burkhart, K. B., & Kennedy, S. (2011). Nuclear RNAi maintains heritable gene silencing in Caenorhabditis elegans. *Proceedings of the National Academy of Sciences*, *108*(49), 19683–19688. https://doi.org/10.1073/pnas.1113310108

Carmona, L. M., & Schatz, D. G. (2017). New insights into the evolutionary origins of the recombination-activating gene proteins and V(D)J recombination. *The FEBS Journal*, *284*(11), 1590–1605. https://doi.org/10.1111/febs.13990

Carthew, R. W., & Sontheimer, E. J. (2009). Origins and Mechanisms of miRNAs and siRNAs. *Cell*, *136*(4), 642–655. https://doi.org/10.1016/J.CELL.2009.01.035

Castel, S. E., & Martienssen, R. A. (2013). RNA interference in the nucleus: roles for small RNAs in transcription, epigenetics and beyond. *Nature Reviews Genetics*, *14*(2), 100–112. https://doi.org/10.1038/nrg3355

Cecere, G., Zheng, G. X. Y., Mansisidor, A. R., Klymko, K. E., & Grishok, A. (2012). Promoters Recognized by Forkhead Proteins Exist for Individual 21U-RNAs. *Molecular Cell*, *47*(5), 734–745. https://doi.org/10.1016/j.molcel.2012.06.021

Chang, T. H., Mattei, E., Gainetdinov, I., Colpan, C., Weng, Z., & Zamore, P. D. (2019). Maelstrom Represses Canonical Polymerase II Transcription within Bi-directional piRNA Clusters in Drosophila melanogaster. *Molecular Cell*, *73*(2), 291–303.e6. https://doi.org/10.1016/J.MOLCEL.2018.10.038

Cheloufi, S., Dos Santos, C. O., Chong, M. M. W., & Hannon, G. J. (2010). A dicer-independent miRNA biogenesis pathway that requires Ago catalysis. *Nature*, *465*(7298), 584–589. https://doi.org/10.1038/nature09092

Chen, C.-C. G., Simard, M. J., Tabara, H., Brownell, D. R., McCollough, J. A., & Mello, C. C. (2005). A Member of the Polymerase β Nucleotidyltransferase Superfamily Is Required for RNA Interference in C. elegans. *Current Biology*, *15*(4), 378–383. https://doi.org/10.1016/J.CUB.2005.01.009

Chen, C., Nott, T. J., Jin, J., & Pawson, T. (2011). Deciphering arginine methylation: Tudor tells the tale. *Nature Reviews Molecular Cell Biology*, *12*(10), 629–642. https://doi.org/10.1038/nrm3185

Chong, M. M. W., Zhang, G., Cheloufi, S., Neubert, T. A., Hannon, G. J., & Littman, D. R. (2010). Canonical and alternate functions of the microRNA biogenesis machinery. *Genes & Development*, *24*(17), 1951–1960. https://doi.org/10.1101/gad.1953310

Chuma, S., & Nakano, T. (2012). piRNA and spermatogenesis in mice. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *368*(1609), 20110338–20110338. https://doi.org/10.1098/rstb.2011.0338

Chuong, E. B., Elde, N. C., & Feschotte, C. (2017). Regulatory activities of transposable elements: from conflicts to benefits. *Nature Reviews Genetics*, *18*(2), 71–86. https://doi.org/10.1038/nrg.2016.139

Ciccia, A., & Elledge, S. J. (2010). The DNA Damage Response: Making It Safe to Play with Knives. *Molecular Cell*, *40*(2), 179–204. https://doi.org/10.1016/J.MOLCEL.2010.09.019

Cifuentes, D., Xue, H., Taylor, D. W., Patnode, H., Mishima, Y., Cheloufi, S., … Giraldez, A. J. (2010). A Novel miRNA Processing Pathway Independent of Dicer Requires Argonaute2 Catalytic Activity. *Science*, *328*(5986), 1694–1698. https://doi.org/10.1126/science.1190809

Claycomb, J. M., Batista, P. J., Pang, K. M., Gu, W., Vasale, J. J., van Wolfswinkel, J. C., … Mello, C. C. (2009). The Argonaute CSR-1 and Its 22G-RNA Cofactors Are Required for Holocentric Chromosome Segregation. *Cell*. https://doi.org/10.1016/j.cell.2009.09.014

Cogoni, C., & Macino, G. (1997). Isolation of quelling-defective (qde) mutants impaired in posttranscriptional transgene-induced gene silencing in Neurospora crassa. *Proceedings of the National Academy of Sciences of the United States of America*, *94*(19), 10233–10238. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9294193

Conine, C. C., Batista, P. J., Gu, W., Claycomb, J. M., Chaves, D. A., Shirayama, M., & Mello, C. C. (2010). Argonautes ALG-3 and ALG-4 are required for spermatogenesis-specific 26G-RNAs and thermotolerant sperm in Caenorhabditis elegans. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(8), 3588–3593. https://doi.org/10.1073/pnas.0911685107

Conine, C. C., Moresco, J. J., Gu, W., Shirayama, M., Conte, D., Yates, J. R., & Mello, C. C. (2013). Argonautes promote male fertility and provide a paternal memory of germline gene expression in C. Elegans. *Cell*. https://doi.org/10.1016/j.cell.2013.11.032

Cordaux, R., & Batzer, M. A. (2009). The impact of retrotransposons on human genome evolution. *Nature Reviews Genetics*, *10*(10), 691–703. https://doi.org/10.1038/nrg2640

Cox, D. N., Chao, A., Baker, J., Chang, L., Qiao, D., & Lin, H. (1998). A novel class of evolutionarily conserved genes defined by piwi are essential for stem cell self-renewal. *Genes and Development*. https://doi.org/10.1101/gad.12.23.3715

Craig, N. L., Lambowitz, A. M., Craigie, R., & Gellert, M. (Eds.). (2002). *Mobile DNA II*. American Society of Microbiology. https://doi.org/10.1128/9781555817954

Czech, B., Munafò, M., Ciabrelli, F., Eastwood, E. L., Fabry, M. H., Kneuss, E., & Hannon, G. J. (2018). piRNA-Guided Genome Defense: From Biogenesis to Silencing. *Annual Review of Genetics*, *52*(1), 131–157. https://doi.org/10.1146/annurev-genet-120417-031441

Das, P. P., Bagijn, M. P., Goldstein, L. D., Woolford, J. R., Lehrbach, N. J., Sapetschnig, A., … Miska, E. A. (2008). Piwi and piRNAs Act Upstream of an Endogenous siRNA Pathway to Suppress Tc3 Transposon Mobility in the Caenorhabditis elegans Germline. *Molecular Cell*. https://doi.org/10.1016/j.molcel.2008.06.003

Daskalos, A., Nikolaidis, G., Xinarianos, G., Savvari, P., Cassidy, A., Zakopoulou, R., … Liloglou, T. (2009). Hypomethylation of retrotransposable elements correlates with genomic instability in non-small cell lung cancer. *International Journal of Cancer*, *124*(1), 81–87. https://doi.org/10.1002/ijc.23849

de Albuquerque, B. F. M., Luteijn, M. J., Cordeiro Rodrigues, R. J., van Bergeijk, P., Waaijers, S., Kaaij, L. J. T., … Ketting, R. F. (2014). PID-1 is a novel factor that operates during 21U-RNA biogenesis in Caenorhabditis elegans. *Genes and Development*. https://doi.org/10.1101/gad.238220.114

de Albuquerque, B. F. M., Placentino, M., & Ketting, R. F. (2015). Maternal piRNAs Are Essential for Germline Development following De Novo Establishment of Endo-siRNAs in Caenorhabditis elegans. *Developmental Cell*. https://doi.org/10.1016/j.devcel.2015.07.010

De Fazio, S., Bartonicek, N., Di Giacomo, M., Abreu-Goodger, C., Sankar, A., Funaya, C., … O'Carroll, D. (2011). The endonuclease activity of Mili fuels piRNA amplification that silences LINE1 elements. *Nature*, *480*(7376), 259–263. https://doi.org/10.1038/nature10547

Dernburg, A. F., Zalevsky, J., Colaiácovo, M. P., & Villeneuve, A. M. (2000). Transgene-mediated cosuppression in the C. elegans germ line. *Genes & Development*, *14*(13), 1578–1583. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10887151

Despic, V., & Neugebauer, K. M. (2018). RNA tales - how embryos read and discard messages from mom. *Journal of Cell Science*, *131*(5), jcs201996. https://doi.org/10.1242/jcs.201996

Dönertas, D., Sienski, G., & Brennecke, J. (2013). Drosophila Gtsf1 is an essential component of the Piwi-mediated transcriptional silencing complex. *Genes & Development*, *27*(15), 1693–1705. https://doi.org/10.1101/gad.221150.113

Duchaine, T. F., Wohlschlegel, J. A., Kennedy, S., Bei, Y., Conte, D., Pang, K., … Mello, C. C. (2006). Functional Proteomics Reveals the Biochemical Niche of C. elegans DCR-1 in Multiple Small-RNA-Mediated Pathways. *Cell*, *124*(2), 343–354. https://doi.org/10.1016/j.cell.2005.11.036

Eide, D., & Anderson, P. (1985). Transposition of Tc1 in the nematode Caenorhabditis elegans. *Proceedings of the National Academy of Sciences of the United States of America*, *82*(6), 1756–1760. https://doi.org/10.1073/PNAS.82.6.1756

Ender, C., Krek, A., Friedländer, M. R., Beitzinger, M., Weinmann, L., Chen, W., … Meister, G. (2008). A Human snoRNA with MicroRNA-Like Functions. *Molecular Cell*, *32*(4), 519–528. https://doi.org/10.1016/J.MOLCEL.2008.10.017

Ender, C., & Meister, G. (2010). Argonaute proteins at a glance. *Journal of Cell Science*, *123*(Pt 11), 1819–1823. https://doi.org/10.1242/jcs.055210

Eulalio, A., Huntzinger, E., Nishihara, T., Rehwinkel, J., Fauser, M., & Izaurralde, E. (2008). Deadenylation is a widespread effect of miRNA regulation. *RNA*, *15*(1), 21–32. https://doi.org/10.1261/rna.1399509

Fagard, M., Boutet, S., Morel, J. B., Bellini, C., & Vaucheret, H. (2000). AGO1, QDE-2, and RDE-1 are related proteins required for post-transcriptional gene silencing in plants, quelling in fungi, and RNA interference in animals. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(21), 11650–11654. https://doi.org/10.1073/pnas.200217597

Félix, M.-A., Ashe, A., Piffaretti, J., Wu, G., Nuez, I., Bélicard, T., … Wang, D. (2011). Natural and Experimental Infection of Caenorhabditis Nematodes by Novel Viruses Related to Nodaviruses. *PLoS Biology*, *9*(1), e1000586. https://doi.org/10.1371/journal.pbio.1000586

Feltzin, V. L., Khaladkar, M., Abe, M., Parisi, M., Hendriks, G.-J., Kim, J., & Bonini, N. M. (2015). The exonuclease Nibbler regulates age-associated traits and modulates piRNA length in *Drosophila*. *Aging Cell*, *14*(3), 443–452. https://doi.org/10.1111/acel.12323

Feschotte, C., & Pritham, E. J. (2007). DNA Transposons and the Evolution of Eukaryotic Genomes. *Annual Review of Genetics*, *41*(1), 331–368. https://doi.org/10.1146/annurev.genet.40.110405.090448

Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E., & Mello, C. C. (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature*, *391*(6669), 806–811. https://doi.org/10.1038/35888

Fischer, S. E. J., Wienholds, E., & Plasterk, R. H. A. (2003). Continuous exchange of sequence information between dispersed Tc1 transposons in the Caenorhabditis elegans genome. *Genetics*, *164*(1), 127–134. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12750326

Frøkjær-Jensen, C., Davis, M. W., Hollopeter, G., Taylor, J., Harris, T. W., Nix, P., … Jorgensen, E. M. (2010). Targeted gene deletions in C. elegans using transposon excision. *Nature Methods*, *7*(6), 451–453. https://doi.org/10.1038/nmeth.1454

Frøkjær-Jensen, C., Davis, M. W., Sarov, M., Taylor, J., Flibotte, S., LaBella, M., … Jorgensen, E. M. (2014). Random and targeted transgene insertion in Caenorhabditis elegans using a modified Mos1 transposon. *Nature Methods*. https://doi.org/10.1038/nmeth.2889

Frøkjær-Jensen, C., Jain, N., Hansen, L., Davis, M. W., Li, Y., Zhao, D., … Fire, A. Z. (2016). An Abundant Class of Non-coding DNA Can Prevent Stochastic Gene Silencing in the C. elegans Germline. *Cell*, *166*(2), 343–357. https://doi.org/10.1016/j.cell.2016.05.072

Frøkjær-Jensen, C., Wayne Davis, M., Hopkins, C. E., Newman, B. J., Thummel, J. M., Olesen, S. P., … Jorgensen, E. M. (2008). Single-copy insertion of transgenes in Caenorhabditis elegans. *Nature Genetics*. https://doi.org/10.1038/ng.248

Fulci, V., & Macino, G. (2007). Quelling: post-transcriptional gene silencing guided by small RNAs in Neurospora crassa. *Current Opinion in Microbiology*, *10*(2), 199–203. https://doi.org/10.1016/J.MIB.2007.03.016

Gao, M., & Arkov, A. L. (2013). Next generation organelles: Structure and role of germ granules in the germline. *Molecular Reproduction and Development*, *80*(8), 610–623. https://doi.org/10.1002/mrd.22115

Gent, J. I., Lamm, A. T., Pavelec, D. M., Maniar, J. M., Parameswaran, P., Tao, L., … Fire, A. Z. (2010). Distinct phases of siRNA synthesis in an endogenous RNAi pathway in C. elegans soma. *Molecular Cell*, *37*(5), 679–689. https://doi.org/10.1016/j.molcel.2010.01.012

Gerson-Gurwitz, A., Wang, S., Sathe, S., Green, R., Yeo, G. W., Oegema, K., & Desai, A. (2016). A Small RNA-Catalytic Argonaute Pathway Tunes Germline Transcript Levels to Ensure Embryonic Divisions. *Cell*, *165*(2), 396–409. https://doi.org/10.1016/J.CELL.2016.02.040

Ghildiyal, M., Seitz, H., Horwich, M. D., Li, C., Du, T., Lee, S., … Zamore, P. D. (2008). Endogenous siRNAs derived from transposons and mRNAs in Drosophila somatic cells. *Science (New York, N.Y.)*, *320*(5879), 1077–1081. https://doi.org/10.1126/science.1157396

Ghildiyal, M., & Zamore, P. D. (2009). Small silencing RNAs: An expanding universe. *Nature Reviews Genetics*. https://doi.org/10.1038/nrg2504

Goh, W.-S. S., Seah, J. W. E., Harrison, E. J., Chen, C., Hammell, C. M., & Hannon, G. J. (2014). A genome-wide RNAi screen identifies factors required for distinct stages of C. elegans piRNA biogenesis. *Genes & Development*, *28*(7), 797–807. https://doi.org/10.1101/gad.235622.113

Golden, D. E., Gerbasi, V. R., & Sontheimer, E. J. (2008). An Inside Job for siRNAs. *Molecular Cell*, *31*(3), 309–312. https://doi.org/10.1016/J.MOLCEL.2008.07.008

Gu, W., Lee, H.-C., Chaves, D., Youngman, E. M., Pazour, G. J., Conte, D., & Mello, C. C. (2012). CapSeq and CIP-TAP identify Pol II start sites and reveal capped small RNAs as C. elegans piRNA precursors. *Cell*, *151*(7), 1488–1500. https://doi.org/10.1016/j.cell.2012.11.023

Gu, W., Shirayama, M., Conte, D., Vasale, J., Batista, P. J., Claycomb, J. M., … Mello, C. C. (2009). Distinct argonaute-mediated 22G-RNA pathways direct genome surveillance in the C. elegans germline. *Molecular Cell*, *36*(2), 231–244. https://doi.org/10.1016/j.molcel.2009.09.020

Guang, S., Bochner, A. F., Burkhart, K. B., Burton, N., Pavelec, D. M., & Kennedy, S. (2010). Small regulatory RNAs inhibit RNA polymerase II during the elongation phase of transcription. *Nature*, *465*(7301), 1097–1101. https://doi.org/10.1038/nature09095

Guang, S., Bochner, A. F., Pavelec, D. M., Burkhart, K. B., Harding, S., Lachowiec, J., & Kennedy, S. (2008). An Argonaute Transports siRNAs from the Cytoplasm to the Nucleus. *Science*, *321*(5888), 537–541. https://doi.org/10.1126/science.1157647

Guffanti, G., Gaudi, S., Fallon, J. H., Sobell, J., Potkin, S. G., Pato, C., & Macciardi, F. (2014). Transposable elements and psychiatric disorders. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, *165*(3), 201–216. https://doi.org/10.1002/ajmg.b.32225

Gunawardane, L. S., Saito, K., Nishida, K. M., Miyoshi, K., Kawamura, Y., Nagami, T., … Siomi, M. C. (2007). A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in Drosophila. *Science (New York, N.Y.)*, *315*(5818), 1587–1590. https://doi.org/10.1126/science.1140494

Guo, H., Ingolia, N. T., Weissman, J. S., & Bartel, D. P. (2010). Mammalian microRNAs predominantly act to decrease target mRNA levels. *Nature*, *466*(7308), 835–840. https://doi.org/10.1038/nature09267

Guzzardo, P. M., Muerdter, F., & Hannon, G. J. (2013). The piRNA pathway in flies: highlights and future directions. *Current Opinion in Genetics & Development*, *23*(1), 44–52. https://doi.org/10.1016/j.gde.2012.12.003

Ha, M., & Kim, V. N. (2014). Regulation of microRNA biogenesis. *Nature Reviews Molecular Cell Biology*, *15*(8), 509–524. https://doi.org/10.1038/nrm3838

Hamilton, A. J., & Baulcombe, D. C. (1999). A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science (New York, N.Y.)*, *286*(5441), 950–952. https://doi.org/10.1126/SCIENCE.286.5441.950

Han, B. W., Wang, W., Li, C., Weng, Z., & Zamore, P. D. (2015). Noncoding RNA. piRNA-guided transposon cleavage initiates Zucchini-dependent, phased piRNA production. *Science (New York, N.Y.)*, *348*(6236), 817–821. https://doi.org/10.1126/science.aaa1264

Han, T., Manoharan, A. P., Harkins, T. T., Bouffard, P., Fitzpatrick, C., Chu, D. S., … Kim, J. K. (2009). 26G endo-siRNAs regulate spermatogenic and zygotic gene expression in Caenorhabditis elegans. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(44), 18674–18679. https://doi.org/10.1073/pnas.0906378106

Hanazawa, M., Yonetani, M., & Sugimoto, A. (2011). PGL proteins self associate and bind RNPs to mediate germ granule assembly in *C. elegans*. *The Journal of Cell Biology*, *192*(6), 929–937. https://doi.org/10.1083/jcb.201010106

Hashimoto, Y., Maegawa, S., Nagai, T., Yamaha, E., Suzuki, H., Yasuda, K., & Inoue, K. (2004). Localized maternal factors are required for zebrafish germ cell formation. *Developmental Biology*, *268*(1), 152–161. https://doi.org/10.1016/J.YDBIO.2003.12.013

Hayashi, R., Schnabl, J., Handler, D., Mohn, F., Ameres, S. L., & Brennecke, J. (2016). Genetic and mechanistic diversity of piRNA 3′-end formation. *Nature*, *539*(7630), 588–592. https://doi.org/10.1038/nature20162

Heard, E., & Martienssen, R. A. (2014). Transgenerational Epigenetic Inheritance: Myths and Mechanisms. *Cell*, *157*(1), 95–109. https://doi.org/10.1016/j.cell.2014.02.045

Hendrickson, D. G., Hogan, D. J., McCullough, H. L., Myers, J. W., Herschlag, D., Ferrell, J. E., & Brown, P. O. (2009). Concordant Regulation of Translation and mRNA Abundance for Hundreds of Targets of a Human microRNA. *PLoS Biology*, *7*(11), e1000238. https://doi.org/10.1371/journal.pbio.1000238

Heo, I., Ha, M., Lim, J., Yoon, M.-J., Park, J.-E., Kwon, S. C., … Kim, V. N. (2012). Mono-Uridylation of Pre-MicroRNA as a Key Step in the Biogenesis of Group II let-7 MicroRNAs. *Cell*, *151*(3), 521–532. https://doi.org/10.1016/j.cell.2012.09.022

Hoeijmakers, J. H. J. (2009). DNA Damage, Aging, and Cancer. *New England Journal of Medicine*, *361*(15), 1475–1485. https://doi.org/10.1056/NEJMra0804615

Horwich, M. D., Li, C., Matranga, C., Vagin, V., Farley, G., Wang, P., & Zamore, P. D. (2007). The Drosophila RNA

Methyltransferase, DmHen1, Modifies Germline piRNAs and Single-Stranded siRNAs in RISC. *Current Biology*, *17*(14), 1265–1272. https://doi.org/10.1016/J.CUB.2007.06.030

Houwing, S., Kamminga, L. M., Berezikov, E., Cronembold, D., Girard, A., van den Elst, H., … Ketting, R. F. (2007). A Role for Piwi and piRNAs in Germ Cell Maintenance and Transposon Silencing in Zebrafish. *Cell*, *129*(1), 69–82. https://doi.org/10.1016/j.cell.2007.03.026

Ipsaro, J. J., Haase, A. D., Knott, S. R., Joshua-Tor, L., & Hannon, G. J. (2012). The structural biochemistry of Zucchini implicates it as a nuclease in piRNA biogenesis. *Nature*, *491*(7423), 279–283. https://doi.org/10.1038/nature11502

Ishizu, H., Siomi, H., & Siomi, M. C. (2012). Biology of PIWI-interacting RNAs: new insights into biogenesis and function inside and outside of germlines. *Genes & Development*, *26*(21), 2361–2373. https://doi.org/10.1101/gad.203786.112

Iwakawa, H., & Tomari, Y. (2015). The Functions of MicroRNAs: mRNA Decay and Translational Repression. *Trends in Cell Biology*, *25*(11), 651–665. https://doi.org/10.1016/j.tcb.2015.07.011

Iwasaki, Y. W., Murano, K., Ishizu, H., Shibuya, A., Iyoda, Y., Siomi, M. C., … Saito, K. (2016). Piwi Modulates Chromatin Accessibility by Regulating Multiple Factors Including Histone H1 to Repress Transposons. *Molecular Cell*, *63*(3), 408–419. https://doi.org/10.1016/j.molcel.2016.06.008

Jackson, A. L., & Linsley, P. S. (2010). Recognizing and avoiding siRNA off-target effects for target identification and therapeutic application. *Nature Reviews Drug Discovery*, *9*(1), 57–67. https://doi.org/10.1038/nrd3010

Jangam, D., Feschotte, C., & Betrán, E. (2017). Transposable Element Domestication As an Adaptation to Evolutionary Conflicts. *Trends in Genetics*, *33*(11), 817–831. https://doi.org/10.1016/j.tig.2017.07.011

Jinek, M., & Doudna, J. A. (2009). A three-dimensional view of the molecular machinery of RNA interference. *Nature*, *457*(7228), 405–412. https://doi.org/10.1038/nature07755

Jurka, J., Kapitonov, V. V., Kohany, O., & Jurka, M. V. (2007). Repetitive Sequences in Complex Genomes: Structure and Evolution. *Annual Review of Genomics and Human Genetics*, *8*(1), 241–259. https://doi.org/10.1146/annurev.genom.8.080706.092416

Kaaij, L. J. T., Hoogstrate, S. W., Berezikov, E., & Ketting, R. F. (2013). piRNA dynamics in divergent zebrafish strains reveal long-lasting maternal influence on zygotic piRNA profiles. *RNA*, *19*(3), 345–356. https://doi.org/10.1261/rna.036400.112

Kamminga, L. M., van Wolfswinkel, J. C., Luteijn, M. J., Kaaij, L. J. T., Bagijn, M. P., Sapetschnig, A., … Ketting, R. F. (2012). Differential Impact of the HEN1 Homolog HENN-1 on 21U and 26G RNAs in the Germline of Caenorhabditis elegans. *PLoS Genetics*, *8*(7), e1002702. https://doi.org/10.1371/journal.pgen.1002702

Kapitonov, V. V., & Jurka, J. (2008). A universal classification of eukaryotic transposable elements implemented in Repbase. *Nature Reviews Genetics*, *9*(5), 411–412. https://doi.org/10.1038/nrg2165-c1

Kapitonov, V. V, & Jurka, J. (2001). Rolling-circle transposons in eukaryotes. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(15), 8714–8719. https://doi.org/10.1073/pnas.151269298

Karlikow, M., Goic, B., & Saleh, M.-C. (2014). RNAi and antiviral defense in Drosophila: Setting up a systemic immune response. *Developmental & Comparative Immunology*, *42*(1), 85–92. https://doi.org/10.1016/J.DCI.2013.05.004

Kasper, D. M., Gardner, K. E., & Reinke, V. (2014). Homeland security in the *C. elegans* germ line. *Epigenetics*, *9*(1), 62–74. https://doi.org/10.4161/epi.26647

Kasper, D. M., Wang, G., Gardner, K. E., Johnstone, T. G., & Reinke, V. (2014). The C. elegans SNAPc Component SNPC-4 Coats piRNA Domains and Is Globally Required for piRNA Abundance. *Developmental Cell*, *31*(2), 145–158. https://doi.org/10.1016/j.devcel.2014.09.015

Kawamura, Y., Saito, K., Kin, T., Ono, Y., Asai, K., Sunohara, T., … Siomi, H. (2008). Drosophila endogenous small RNAs bind to Argonaute 2 in somatic cells. *Nature*, *453*(7196), 793–797. https://doi.org/10.1038/nature06938

Kazazian, H. H. (2004). Mobile elements: drivers of genome evolution. *Science (New York, N.Y.)*, *303*(5664), 1626–1632. https://doi.org/10.1126/science.1089670

Kazazian, H. H., Wong, C., Youssoufian, H., Scott, A. F., Phillips, D. G., & Antonarakis, S. E. (1988). Haemophilia A resulting from de novo insertion of L1 sequences represents a novel mechanism for mutation in man. *Nature*, *332*(6160), 164–166. https://doi.org/10.1038/332164a0

Kebriaei, P., Izsvák, Z., Narayanavari, S. A., Singh, H., & Ivics, Z. (2017). Gene Therapy with the Sleeping Beauty Transposon System. *Trends in Genetics*, *33*(11), 852–870. https://doi.org/10.1016/j.tig.2017.08.008

Kebriaei, P., Singh, H., Huls, M. H., Figliola, M. J., Bassett, R., Olivares, S., … Cooper, L. J. N. (2016). Phase I trials using Sleeping Beauty to generate CD19-specific CAR T cells. *The Journal of Clinical Investigation*, *126*(9), 3363–3376. https://doi.org/10.1172/JCI86721

Ketting, R. F. (2011). The Many Faces of RNAi. *Developmental Cell*. https://doi.org/10.1016/j.devcel.2011.01.012

Ketting, R. F., Haverkamp, T. H. A., Van Luenen, H. G. A. M., & Plasterk, R. H. A. (1999). mut-7 of C. elegans, required for transposon silencing and RNA interference, is a homolog of werner syndrome helicase and RNaseD. *Cell*. https://doi.org/10.1016/S0092-8674(00)81645-1

Ketting, R. F., & Plasterk, R. H. A. (2000). A genetic link between co-suppression and RNA interference in C. elegans. *Nature*, *404*(6775), 296–298. https://doi.org/10.1038/35005113

Khvorova, A., Reynolds, A., & Jayasena, S. D. (2003). Functional siRNAs and miRNAs exhibit strand bias. *Cell*, *115*(2), 209–216. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/14567918

Kirino, Y., Kim, N., de Planell-Saguer, M., Khandros, E., Chiorean, S., Klein, P. S., … Mourelatos, Z. (2009). Arginine methylation of Piwi proteins catalysed by dPRMT5 is required for Ago3 and Aub stability. *Nature Cell Biology*, *11*(5), 652–658. https://doi.org/10.1038/ncb1872

Kirino, Y., & Mourelatos, Z. (2007). Mouse Piwi-interacting RNAs are 2′-O-methylated at their 3′ termini. *Nature Structural & Molecular Biology*, *14*(4), 347–348. https://doi.org/10.1038/nsmb1218

Klattenhoff, C., Xi, H., Li, C., Lee, S., Xu, J., Khurana, J. S., … Theurkauf, W. E. (2009). The Drosophila HP1 homolog Rhino is required for transposon silencing and piRNA production by dual-strand clusters. *Cell*, *138*(6), 1137–1149. https://doi.org/10.1016/j.cell.2009.07.014

Klosin, A., Casas, E., Hidalgo-Carcedo, C., Vavouri, T., & Lehner, B. (2017). Transgenerational transmission of environmental information in *C. elegans*. *Science*, *356*(6335), 320–323. https://doi.org/10.1126/science.aah6412

Koonin, E. V, & Dolja, V. V. (2013). A virocentric perspective on the evolution of life. *Current Opinion in Virology*, *3*(5), 546–557. https://doi.org/10.1016/J.COVIRO.2013.06.008

Kuramochi-Miyagawa, S., Watanabe, T., Gotoh, K., Totoki, Y., Toyoda, A., Ikawa, M., … Nakano, T. (2008). DNA methylation of retrotransposon genes is regulated by Piwi family members MILI and MIWI2 in murine fetal testes. *Genes & Development*, *22*(7), 908–917. https://doi.org/10.1101/gad.1640708

Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., … International Human Genome Sequencing Consortium. (2001). Initial sequencing and analysis of the human genome. *Nature*, *409*(6822), 860–921. https://doi.org/10.1038/35057062

Laricchia, K. M., Zdraljevic, S., Cook, D. E., & Andersen, E. C. (2017). Natural Variation in the Distribution and Abundance of

Transposable Elements Across the Caenorhabditis elegans Species. *Molecular Biology and Evolution*, *34*(9), 2187–2202. https://doi.org/10.1093/molbev/msx155

Le Thomas, A., Marinov, G. K., & Aravin, A. A. (2014). A transgenerational process defines piRNA biogenesis in Drosophila virilis. *Cell Reports*, *8*(6), 1617–1623. https://doi.org/10.1016/j.celrep.2014.08.013

Le Thomas, A., Stuwe, E., Li, S., Du, J., Marinov, G., Rozhkov, N., … Aravin, A. A. (2014). Transgenerationally inherited piRNAs trigger piRNA biogenesis by changing the chromatin of piRNA clusters and inducing precursor processing. *Genes & Development*, *28*(15), 1667–1680. https://doi.org/10.1101/gad.245514.114

Lee, H. C., Gu, W., Shirayama, M., Youngman, E., Conte, D., & Mello, C. C. (2012). C. elegans piRNAs mediate the genome-wide surveillance of germline transcripts. *Cell*. https://doi.org/10.1016/j.cell.2012.06.016

Lev, I., Bril, R., Liu, Y., Cere, L. I., & Rechavi, O. (2018). Inter-generational Consequences for Growing C. elegans in Liquid. *BioRxiv*, 467837. https://doi.org/10.1101/467837

Lewis, S. H., Quarles, K. A., Yang, Y., Tanguy, M., Frézal, L., Smith, S. A., … Jiggins, F. M. (2018). Pan-arthropod analysis reveals somatic piRNAs as an ancestral defence against transposable elements. *Nature Ecology & Evolution*, *2*(1), 174–181. https://doi.org/10.1038/s41559-017-0403-4

Li, H., Li, W. X., & Ding, S. W. (2002). Induction and suppression of RNA silencing by an animal virus. *Science (New York, N.Y.)*, *296*(5571), 1319–1321. https://doi.org/10.1126/science.1070948

Luteijn, M. J., & Ketting, R. F. (2013). PIWI-interacting RNAs: From generation to transgenerational epigenetics. *Nature Reviews Genetics*. https://doi.org/10.1038/nrg3495

Luteijn, M. J., Van Bergeijk, P., Kaaij, L. J. T., Almeida, M. V., Roovers, E. F., Berezikov, E., & Ketting, R. F. (2012). Extremely stable Piwi-induced gene silencing in Caenorhabditis elegans. *EMBO Journal*. https://doi.org/10.1038/emboj.2012.213

Madhani, H. D. (2013). The frustrated gene: origins of eukaryotic gene expression. *Cell*, *155*(4), 744–749. https://doi.org/10.1016/j.cell.2013.10.003

Malone, C. D., & Hannon, G. J. (2009). Small RNAs as Guardians of the Genome. *Cell*. https://doi.org/10.1016/j.cell.2009.01.045

Marraffini, L. A. (2015). CRISPR-Cas immunity in prokaryotes. *Nature*, *526*(7571), 55–61. https://doi.org/10.1038/nature15386

Martienssen, R., & Moazed, D. (2015). RNAi and heterochromatin assembly. *Cold Spring Harbor Perspectives in Biology*, *7*(8), a019323. https://doi.org/10.1101/cshperspect.a019323

Matsumoto, N., Nishimasu, H., Sakakibara, K., Nishida, K. M., Hirano, T., Ishitani, R., … Nureki, O. (2016). Crystal Structure of Silkworm PIWI-Clade Argonaute Siwi Bound to piRNA. *Cell*, *167*(2), 484–497.e9. https://doi.org/10.1016/J.CELL.2016.09.002

McClintock, B. (1951). CHROMOSOME ORGANIZATION AND GENIC EXPRESSION. *Cold Spring Harbor Symposia on Quantitative Biology*, *16*(0), 13–47. https://doi.org/10.1101/SQB.1951.016.01.004

Miesen, P., Girardi, E., & van Rij, R. P. (2015). Distinct sets of PIWI proteins produce arbovirus and transposon-derived piRNAs in *Aedes aegypti* mosquito cells. *Nucleic Acids Research*, *43*(13), 6545–6556. https://doi.org/10.1093/nar/gkv590

Miki, Y., Nishisho, I., Horii, A., Miyoshi, Y., Utsunomiya, J., Kinzler, K. W., … Nakamura, Y. (1992). Disruption of the APC gene by a retrotransposal insertion of L1 sequence in a colon cancer. *Cancer Research*, *52*(3), 643–645. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1310068

Minkina, O., & Hunter, C. P. (2018). Intergenerational Transmission of Gene Regulatory Information in Caenorhabditis elegans. *Trends in Genetics*, *34*(1), 54–64. https://doi.org/10.1016/j.tig.2017.09.012

Moazed, D., Bühler, M., Buker, S. M., Colmenares, S. U., Gerace, E. L., Gerber, S. A., … Gygi, S. P. (2006). Studies on the mechanism of RNAi-dependent heterochromatin assembly. *Cold Spring Harbor Symposia on Quantitative Biology*, *71*, 461–471. https://doi.org/10.1101/sqb.2006.71.044

Mohn, F., Handler, D., & Brennecke, J. (2015). piRNA-guided slicing specifies transcripts for Zucchini-dependent, phased piRNA biogenesis. *Science*, *348*(6236), 812–817. https://doi.org/10.1126/science.aaa1039

Montgomery, T. A., Rim, Y.-S., Zhang, C., Dowen, R. H., Phillips, C. M., Fischer, S. E. J., & Ruvkun, G. (2012). PIWI associated siRNAs and piRNAs specifically require the Caenorhabditis elegans HEN1 ortholog henn-1. *PLoS Genetics*, *8*(4), e1002616. https://doi.org/10.1371/journal.pgen.1002616

Morazzani, E. M., Wiley, M. R., Murreddu, M. G., Adelman, Z. N., & Myles, K. M. (2012). Production of Virus-Derived Ping-Pong-Dependent piRNA-like Small RNAs in the Mosquito Soma. *PLoS Pathogens*, *8*(1), e1002470. https://doi.org/10.1371/journal.ppat.1002470

Mourrain, P., Béclin, C., Elmayan, T., Feuerbach, F., Godon, C., Morel, J.-B., … Vaucheret, H. (2000). Arabidopsis SGS2 and SGS3 Genes Are Required for Posttranscriptional Gene Silencing and Natural Virus Resistance. *Cell*, *101*(5), 533–542. https://doi.org/10.1016/S0092-8674(00)80863-6

Muerdter, F., Guzzardo, P. M., Gillis, J., Luo, Y., Yu, Y., Chen, C., … Hannon, G. J. (2013). A Genome-wide RNAi Screen Draws a Genetic Framework for Transposon Control and Primary piRNA Biogenesis in Drosophila. *Molecular Cell*, *50*(5), 736–748. https://doi.org/10.1016/j.molcel.2013.04.006

Nakanishi, K., Weinberg, D. E., Bartel, D. P., & Patel, D. J. (2012). Structure of yeast Argonaute with guide RNA. *Nature*, *486*(7403), 368–374. https://doi.org/10.1038/nature11211

Narita, N., Nishio, H., Kitoh, Y., Ishikawa, Y., Ishikawa, Y., Minami, R., … Matsuo, M. (1993). Insertion of a 5' truncated L1 element into the 3' end of exon 44 of the dystrophin gene resulted in skipping of the exon during splicing in a case of Duchenne muscular dystrophy. *Journal of Clinical Investigation*, *91*(5), 1862–1867. https://doi.org/10.1172/JCI116402

Newman, M. A., Ji, F., Fischer, S. E. J., Anselmo, A., Sadreyev, R. I., & Ruvkun, G. (2018). The surveillance of pre-mRNA splicing is an early step in *C. elegans* RNAi of endogenous genes. *Genes & Development*, *32*(9–10), 670–681. https://doi.org/10.1101/gad.311514.118

Nishida, K. M., Okada, T. N., Kawamura, T., Mituyama, T., Kawamura, Y., Inagaki, S., … Siomi, M. C. (2009). Functional involvement of Tudor and dPRMT5 in the piRNA processing pathway in Drosophila germlines. *The EMBO Journal*, *28*(24), 3820–3831. https://doi.org/10.1038/emboj.2009.365

Nishimasu, H., Ishizu, H., Saito, K., Fukuhara, S., Kamatani, M. K., Bonnefond, L., … Nureki, O. (2012). Structure and function of Zucchini endoribonuclease in piRNA biogenesis. *Nature*, *491*(7423), 284–287. https://doi.org/10.1038/nature11509

Ohara, T., Sakaguchi, Y., Suzuki, T., Ueda, H., Miyauchi, K., & Suzuki, T. (2007). The 3′ termini of mouse Piwi-interacting RNAs are 2′-O-methylated. *Nature Structural & Molecular Biology*, *14*(4), 349–350. https://doi.org/10.1038/nsmb1220

Ohtani, H., Iwasaki, Y. W., Shibuya, A., Siomi, H., Siomi, M. C., & Saito, K. (2013). DmGTSF1 is necessary for Piwi-piRISC-mediated transcriptional transposon silencing in the Drosophila ovary. *Genes & Development*, *27*(15), 1656–1661. https://doi.org/10.1101/gad.221515.113

Okamura, K., Hagen, J. W., Duan, H., Tyler, D. M., & Lai, E. C. (2007). The Mirtron Pathway Generates microRNA-Class Regulatory RNAs in Drosophila. *Cell*, *130*(1), 89–100. https://doi.org/10.1016/j.cell.2007.06.028

Olina, A. V., Kulbachinskiy, A. V., Aravin, A. A., & Esyunina, D. M. (2018). Argonaute Proteins and Mechanisms of RNA Interference in Eukaryotes and Prokaryotes. *Biochemistry (Moscow)*, *83*(5), 483–497. https://doi.org/10.1134/S0006297918050024

Ozata, D. M., Gainetdinov, I., Zoch, A., O'Carroll, D., & Zamore, P. D. (2018). PIWI-interacting RNAs: small RNAs with big functions. *Nature Reviews Genetics*, 1. https://doi.org/10.1038/s41576-018-0073-3

Pak, J., & Fire, A. (2007). Distinct populations of primary and secondary effectors during RNAi in C. elegans. *Science (New York, N.Y.)*, *315*(5809), 241–244. https://doi.org/10.1126/science.1132839

Pane, A., Jiang, P., Zhao, D. Y., Singh, M., & Schüpbach, T. (2011). The Cutoff protein regulates piRNA cluster expression and piRNA production in the *Drosophila* germline. *The EMBO Journal*, *30*(22), 4601–4615. https://doi.org/10.1038/emboj.2011.334

Pastuzyn, E. D., Day, C. E., Kearns, R. B., Kyrke-Smith, M., Taibi, A. V., McCormick, J., … Shepherd, J. D. (2018). The Neuronal Gene Arc Encodes a Repurposed Retrotransposon Gag Protein that Mediates Intercellular RNA Transfer. *Cell*, *172*(1–2), 275–288.e18. https://doi.org/10.1016/J.CELL.2017.12.024

Pek, J. W., Anand, A., & Kai, T. (2012). Tudor domain proteins in development. *Development*. https://doi.org/10.1242/dev.073304

Phillips, C. M., Brown, K. C., Montgomery, B. E., Ruvkun, G., & Montgomery, T. A. (2015). PiRNAs and piRNA-Dependent siRNAs Protect Conserved and Essential C. elegans Genes from Misrouting into the RNAi Pathway. *Developmental Cell*. https://doi.org/10.1016/j.devcel.2015.07.009

Phillips, C. M., Montgomery, B. E., Breen, P. C., Roovers, E. F., Rim, Y. S., Ohsumi, T. K., … Montgomery, T. A. (2014). MUT-14 and SMUT-1 DEAD box RNA helicases have overlapping roles in germline RNAi and endogenous siRNA formation. *Current Biology*. https://doi.org/10.1016/j.cub.2014.02.060

Phillips, C. M., Montgomery, T. A., Breen, P. C., & Ruvkun, G. (2012). MUT-16 promotes formation of perinuclear Mutator foci required for RNA silencing in the C. elegans germline. *Genes and Development*. https://doi.org/10.1101/gad.193904.112

Piégu, B., Bire, S., Arensburger, P., & Bigot, Y. (2015). A survey of transposable element classification systems – A call for a fundamental update to meet the challenge of their diversity and complexity. *Molecular Phylogenetics and Evolution*, *86*, 90–109. https://doi.org/10.1016/J.YMPEV.2015.03.009

Ponting, C. P. (1997). Tudor domains in proteins that interact with RNA. *Trends in Biochemical Sciences*, *22*(2), 51–52. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9048482

Rebollo, R., Romanish, M. T., & Mager, D. L. (2012). Transposable Elements: An Abundant and Natural Source of Regulatory Sequences for Host Genes. *Annual Review of Genetics*, *46*(1), 21–42. https://doi.org/10.1146/annurev-genet-110711-155621

Rechavi, O., Houri-Ze'evi, L., Anava, S., Goh, W. S. S., Kerk, S. Y., Hannon, G. J., & Hobert, O. (2014). Starvation-Induced Transgenerational Inheritance of Small RNAs in C. elegans. *Cell*, *158*(2), 277–287. https://doi.org/10.1016/j.cell.2014.06.020

Rhoades, M. W., Reinhart, B. J., Lim, L. P., Burge, C. B., Bartel, B., & Bartel, D. P. (2002). Prediction of plant microRNA targets. *Cell*, *110*(4), 513–520. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12202040

Robert, V., & Bessereau, J.-L. (2007). Targeted engineering of the Caenorhabditis elegans genome following Mos1-triggered chromosomal breaks. *The EMBO Journal*, *26*(1), 170–183. https://doi.org/10.1038/sj.emboj.7601463

Robert, V. J. P., Sijen, T., van Wolfswinkel, J., & Plasterk, R. H. A. (2005). Chromatin and RNAi factors protect the C. elegans germline against repetitive sequences. *Genes & Development*, *19*(7), 782–787. https://doi.org/10.1101/gad.332305

Rodrigues, R. J. C., Domingues, A. M. de J., Hellmann, S., Dietz, S., Albuquerque, B. de, Renz, C., … Ketting, R. (2018). PETISCO is a novel protein complex required for 21U RNA biogenesis and embryonic viability. *BioRxiv*, 463711. https://doi.org/10.1101/463711

Rojas-Ríos, P., & Simonelig, M. (2018). piRNAs and PIWI proteins: regulators of gene expression in development and stem cells. *Development (Cambridge, England)*, *145*(17), dev161786. https://doi.org/10.1242/dev.161786

Romano, N., & Macino, G. (1992). Quelling: transient inactivation of gene expression in Neurospora crassa by transformation with homologous sequences. *Molecular Microbiology*, *6*(22), 3343–3353. https://doi.org/10.1111/j.1365-2958.1992.tb02202.x

Rosenkranz, D., Rudloff, S., Bastuck, K., Ketting, R. F., & Zischler, H. (2015). Tupaia small RNAs provide insights into function and evolution of RNAi-based transposon defense in mammals. *RNA*, *21*(5), 911–922. https://doi.org/10.1261/rna.048603.114

Ruby, J. G., Jan, C. H., & Bartel, D. P. (2007). Intronic microRNA precursors that bypass Drosha processing. *Nature*, *448*(7149), 83–86. https://doi.org/10.1038/nature05983

Ruby, J. G., Jan, C., Player, C., Axtell, M. J., Lee, W., Nusbaum, C., … Bartel, D. P. (2006). Large-Scale Sequencing Reveals 21U-RNAs and Additional MicroRNAs and Endogenous siRNAs in C. elegans. *Cell*, *127*(6), 1193–1207. https://doi.org/10.1016/j.cell.2006.10.040

Russell, S. J., & LaMarre, J. (2018). Transposons and the PIWI pathway: genome defense in gametes and embryos. *Reproduction (Cambridge, England)*, *156*(4), R111–R124. https://doi.org/10.1530/REP-18-0218

Saito, K., Ishizu, H., Komai, M., Kotani, H., Kawamura, Y., Nishida, K. M., … Siomi, M. C. (2010). Roles for the Yb body components Armitage and Yb in primary piRNA biogenesis in Drosophila. *Genes & Development*, *24*(22), 2493–2498. https://doi.org/10.1101/gad.1989510

Sato, K., Iwasaki, Y. W., Shibuya, A., Carninci, P., Tsuchizawa, Y., Ishizu, H., … Siomi, H. (2015). Krimper Enforces an Antisense Bias on piRNA Pools by Binding AGO3 in the Drosophila Germline. *Molecular Cell*, *59*(4), 553–563. https://doi.org/10.1016/j.molcel.2015.06.024

Schirle, N. T., & MacRae, I. J. (2012). The crystal structure of human Argonaute2. *Science (New York, N.Y.)*, *336*(6084), 1037–1040. https://doi.org/10.1126/science.1221551

Schnettler, E., Donald, C. L., Human, S., Watson, M., Siu, R. W. C., McFarlane, M., … Fragkoudis, R. (2013). Knockdown of piRNA pathway proteins results in enhanced Semliki Forest virus production in mosquito cells. *Journal of General Virology*, *94*(Pt_7), 1680–1689. https://doi.org/10.1099/vir.0.053850-0

Schwab, R., Palatnik, J. F., Riester, M., Schommer, C., Schmid, M., & Weigel, D. (2005). Specific Effects of MicroRNAs on the Plant Transcriptome. *Developmental Cell*, *8*(4), 517–527. https://doi.org/10.1016/j.devcel.2005.01.018

Schwahn, U., Lenzner, S., Dong, J., Feil, S., Hinzmann, B., van Duijnhoven, G., … Berger, W. (1998). Positional cloning of the gene for X-linked retinitis pigmentosa 2. *Nature Genetics*, *19*(4), 327–332. https://doi.org/10.1038/1214

Schwarz, D. S., Hutvágner, G., Du, T., Xu, Z., Aronin, N., & Zamore, P. D. (2003). Asymmetry in the assembly of the RNAi enzyme complex. *Cell*, *115*(2), 199–208. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/14567917

Seth, M., Shirayama, M., Gu, W., Ishidate, T., Conte, D., & Mello, C. (2013). The C. elegans CSR-1 argonaute pathway counteracts epigenetic silencing to promote germline gene expression. *Developmental Cell*. https://doi.org/10.1016/j.devcel.2013.11.014

Seydoux, G. (2018). The P Granules of C. elegans: A Genetic Model for the Study of RNA–Protein Condensates. *Journal of Molecular Biology*. https://doi.org/10.1016/J.JMB.2018.08.007

Shen, E.-Z., Chen, H., Ozturk, A. R., Tu, S., Shirayama, M., Tang, W., … Mello, C. C. (2018). Identification of piRNA Binding Sites Reveals the Argonaute Regulatory Landscape of the C. elegans Germline. *Cell*, *172*(5), 937–951.e18. https://doi.org/10.1016/j.cell.2018.02.002

Sheng, G., Zhao, H., Wang, J., Rao, Y., Tian, W., Swarts, D. C., … Wang, Y. (2014). Structure-based cleavage mechanism of Thermus thermophilus Argonaute DNA guide strand-mediated DNA target cleavage. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(2), 652–657. https://doi.org/10.1073/pnas.1321032111

Sheu-Gruttadauria, J., & MacRae, I. J. (2017). Structural Foundations of RNA Silencing by Argonaute. *Journal of Molecular Biology*,

429(17), 2619–2639. https://doi.org/10.1016/J.JMB.2017.07.018

Shirayama, M., Seth, M., Lee, H. C., Gu, W., Ishidate, T., Conte, D., & Mello, C. C. (2012). PiRNAs initiate an epigenetic memory of nonself RNA in the C. elegans germline. *Cell*. https://doi.org/10.1016/j.cell.2012.06.015

Sibley, C. R., Seow, Y., Saayman, S., Dijkstra, K. K., El Andaloussi, S., Weinberg, M. S., & Wood, M. J. A. (2012). The biogenesis and characterization of mammalian microRNAs of mirtron origin. *Nucleic Acids Research*, *40*(1), 438–448. https://doi.org/10.1093/nar/gkr722

Sienski, G., Batki, J., Senti, K.-A., Dönertas, D., Tirian, L., Meixner, K., & Brennecke, J. (2015). Silencio/CG9754 connects the Piwi–piRNA complex to the cellular heterochromatin machinery. *Genes & Development*, *29*(21), 2258–2271. https://doi.org/10.1101/gad.271908.115

Sienski, G., Dönertas, D., & Brennecke, J. (2012). Transcriptional Silencing of Transposons by Piwi and Maelstrom and Its Impact on Chromatin State and Gene Expression. *Cell*, *151*(5), 964–980. https://doi.org/10.1016/J.CELL.2012.10.040

Sijen, T., & Plasterk, R. H. A. (2003). Transposon silencing in the Caenorhabditis elegans germ line by natural RNAi. *Nature*, *426*(6964), 310–314. https://doi.org/10.1038/nature02107

Simon, M., Sarkies, P., Ikegami, K., Doebley, A.-L., Goldstein, L. D., Mitchell, J., … Ahmed, S. (2014). Reduced insulin/IGF-1 signaling restores germ cell immortality to Caenorhabditis elegans Piwi mutants. *Cell Reports*, *7*(3), 762–773. https://doi.org/10.1016/j.celrep.2014.03.056

Simonelig, M. (2014). piRNAs, master regulators of gene expression. *Cell Research*, *24*(7), 779–780. https://doi.org/10.1038/cr.2014.78

Siomi, M. C., Mannen, T., & Siomi, H. (2010). How does the royal family of Tudor rule the PIWI-interacting RNA pathway? *Genes & Development*, *24*(7), 636–646. https://doi.org/10.1101/gad.1899210

Siomi, M. C., Sato, K., Pezic, D., & Aravin, A. A. (2011). PIWI-interacting small RNAs: the vanguard of genome defence. *Nature Reviews Molecular Cell Biology*, *12*(4), 246–258. https://doi.org/10.1038/nrm3089

Slotkin, R. K., & Martienssen, R. (2007). Transposable elements and the epigenetic regulation of the genome. *Nature Reviews Genetics*, *8*(4), 272–285. https://doi.org/10.1038/nrg2072

Song, J.-J., Smith, S. K., Hannon, G. J., & Joshua-Tor, L. (2004). Crystal structure of Argonaute and its implications for RISC slicer activity. *Science (New York, N.Y.)*, *305*(5689), 1434–1437. https://doi.org/10.1126/science.1102514

Spradling, A. C., Stern, D., Beaton, A., Rhem, E. J., Laverty, T., Mozden, N., … Rubin, G. M. (1999). The Berkeley Drosophila Genome Project gene disruption project: Single P-element insertions mutating 25% of vital Drosophila genes. *Genetics*, *153*(1), 135–177. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10471706

Sternberg, S. H., Richter, H., Charpentier, E., & Qimron, U. (2016). Adaptation in CRISPR-Cas Systems. *Molecular Cell*, *61*(6), 797–808. https://doi.org/10.1016/J.MOLCEL.2016.01.030

Suter, C. M., Martin, D. I., & Ward, R. L. (2004). Hypomethylation of L1 retrotransposons in colorectal cancer and adjacent normal tissue. *International Journal of Colorectal Disease*, *19*(2), 95–101. https://doi.org/10.1007/s00384-003-0539-3

Swarts, D. C., Makarova, K., Wang, Y., Nakanishi, K., Ketting, R. F., Koonin, E. V, … van der Oost, J. (2014). The evolutionary journey of Argonaute proteins. *Nature Structural & Molecular Biology*, *21*(9), 743–753. https://doi.org/10.1038/nsmb.2879

Tabara, H., Sarkissian, M., Kelly, W. G., Fleenor, J., Grishok, A., Timmons, L., … Mello, C. C. (1999). The rde-1 gene, RNA interference, and transposon silencing in C. elegans. *Cell*. https://doi.org/10.1016/S0092-8674(00)81644-X

Tam, O. H., Aravin, A. A., Stein, P., Girard, A., Murchison, E. P., Cheloufi, S., … Hannon, G. J. (2008). Pseudogene-derived small

interfering RNAs regulate gene expression in mouse oocytes. *Nature*, *453*(7194), 534–538. https://doi.org/10.1038/nature06904

Tang, W., Seth, M., Tu, S., Shen, E.-Z., Li, Q., Shirayama, M., … Mello, C. C. (2018). A Sex Chromosome piRNA Promotes Robust Dosage Compensation and Sex Determination in C. elegans. *Developmental Cell*, *44*(6), 762–770.e3. https://doi.org/10.1016/j.devcel.2018.01.025

Tang, W., Tu, S., Lee, H.-C., Weng, Z., & Mello, C. C. (2016). The RNase PARN-1 Trims piRNA 3′ Ends to Promote Transcriptome Surveillance in C. elegans. *Cell*, *164*(5), 974–984. https://doi.org/10.1016/j.cell.2016.02.008

Teugels, E., De Brakeleer, S., Goelen, G., Lissens, W., Sermijn, E., & De Grève, J. (2005). De novo Alu element insertions targeted to a sequence common to the BRCA1 and BRCA2 genes. *Human Mutation*, *26*(3), 284–284. https://doi.org/10.1002/humu.9366

Thivierge, C., Makil, N., Flamand, M., Vasale, J. J., Mello, C. C., Wohlschlegel, J., … Duchaine, T. F. (2012). Tudor domain ERI-5 tethers an RNA-dependent RNA polymerase to DCR-1 to potentiate endo-RNAi. *Nature Structural & Molecular Biology*, *19*(1), 90–97. https://doi.org/10.1038/nsmb.2186

Tijsterman, M., Ketting, R. F., Okihara, K. L., Sijen, T., & Plasterk, R. H. A. (2002). RNA helicase MUT-14-dependent gene silencing triggered in C. elegans by short antisense RNAs. *Science (New York, N.Y.)*, *295*(5555), 694–697. https://doi.org/10.1126/science.1067534

Tomari, Y., Matranga, C., Haley, B., Martinez, N., & Zamore, P. D. (2004). A Protein Sensor for siRNA Asymmetry. *Science*, *306*(5700), 1377–1380. https://doi.org/10.1126/science.1102755

Tops, B. B. J., Tabara, H., Sijen, T., Simmer, F., Mello, C. C., Plasterk, R. H. A., & Ketting, R. F. (2005). RDE-2 interacts with MUT-7 to mediate RNA interference in Caenorhabditis elegans. *Nucleic Acids Research*, *33*(1), 347–355. https://doi.org/10.1093/nar/gki183

Updike, D., & Strome, S. (2010). P Granule Assembly and Function in Caenorhabditis elegans Germ Cells. *Journal of Andrology*, *31*(1), 53–60. https://doi.org/10.2164/jandrol.109.008292

Vagin, V. V., Sigova, A., Li, C., Seitz, H., Gvozdev, V., & Zamore, P. D. (2006). A Distinct Small RNA Pathway Silences Selfish Genetic Elements in the Germline. *Science*, *313*(5785), 320–324. https://doi.org/10.1126/science.1129333

Vagin, V. V, Wohlschlegel, J., Qu, J., Jonsson, Z., Huang, X., Chuma, S., … Aravin, A. A. (2009). Proteomic analysis of murine Piwi proteins reveals a role for arginine methylation in specifying interaction with Tudor family members. *Genes & Development*, *23*(15), 1749–1762. https://doi.org/10.1101/gad.1814809

Valencia-Sanchez, M. A., Liu, J., Hannon, G. J., & Parker, R. (2006). Control of translation and mRNA degradation by miRNAs and siRNAs. *Genes & Development*, *20*(5), 515–524. https://doi.org/10.1101/gad.1399806

van Wolfswinkel, J. C. (2014). Piwi and Potency: PIWI Proteins in Animal Stem Cells and Regeneration. *Integrative and Comparative Biology*, *54*(4), 700–713. https://doi.org/10.1093/icb/icu084

van Wolfswinkel, J. C., Claycomb, J. M., Batista, P. J., Mello, C. C., Berezikov, E., & Ketting, R. F. (2009). CDE-1 Affects Chromosome Segregation through Uridylation of CSR-1-Bound siRNAs. *Cell*, *139*(1), 135–148. https://doi.org/10.1016/j.cell.2009.09.012

van Wolfswinkel, J. C., & Ketting, R. F. (2010). The role of small non-coding RNAs in genome stability and chromatin organization. *Journal of Cell Science*, *123*(Pt 11), 1825–1839. https://doi.org/10.1242/jcs.061713

Vasale, J. J., Gu, W., Thivierge, C., Batista, P. J., Claycomb, J. M., Youngman, E. M., … Conte, D. (2010). Sequential rounds of RNA-dependent RNA transcription drive endogenous small-RNA biogenesis in the ERGO-1/Argonaute pathway. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(8), 3582–3587.

https://doi.org/10.1073/pnas.0911908107

Vastenhouw, N. L., Fischer, S. E. J., Robert, V. J. P., Thijssen, K. L., Fraser, A. G., Kamath, R. S., … Plasterk, R. H. A. (2003). A Genome-Wide Screen Identifies 27 Genes Involved in Transposon Silencing in C. elegans. *Current Biology*, *13*(15), 1311–1316. https://doi.org/10.1016/S0960-9822(03)00539-6

Verdel, A., & Moazed, D. (2005). RNAi-directed assembly of heterochromatin in fission yeast. *FEBS Letters*, *579*(26), 5872–5878. https://doi.org/10.1016/j.febslet.2005.08.083

Voronina, E. (2013). The diverse functions of germline P-granules in *Caenorhabditis elegans*. *Molecular Reproduction and Development*, *80*(8), 624–631. https://doi.org/10.1002/mrd.22136

Wakiyama, M., Takimoto, K., Ohara, O., & Yokoyama, S. (2007). Let-7 microRNA-mediated mRNA deadenylation and translational repression in a mammalian cell-free system. *Genes & Development*, *21*(15), 1857–1862. https://doi.org/10.1101/gad.1566707

Wang, B., Li, S., Qi, H. H., Chowdhury, D., Shi, Y., & Novina, C. D. (2009). Distinct passenger strand and mRNA cleavage activities of human Argonaute proteins. *Nature Structural & Molecular Biology*, *16*(12), 1259–1266. https://doi.org/10.1038/nsmb.1712

Wang, G., & Reinke, V. (2008). A C. elegans Piwi, PRG-1, Regulates 21U-RNAs during Spermatogenesis. *Current Biology*. https://doi.org/10.1016/j.cub.2008.05.009

Wang, H., Ma, Z., Niu, K., Xiao, Y., Wu, X., Pan, C., … Liu, N. (2016). Antagonistic roles of Nibbler and Hen1 in modulating piRNA 3′ ends in *Drosophila*. *Development*, *143*(3), 530–539. https://doi.org/10.1242/dev.128116

Watanabe, T., Totoki, Y., Toyoda, A., Kaneda, M., Kuramochi-Miyagawa, S., Obata, Y., … Sasaki, H. (2008). Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. *Nature*, *453*(7194), 539–543. https://doi.org/10.1038/nature06908

Webster, A., Li, S., Hur, J. K., Wachsmuth, M., Bois, J. S., Perkins, E. M., … Aravin, A. A. (2015). Aub and Ago3 Are Recruited to Nuage through Two Mechanisms to Form a Ping-Pong Complex Assembled by Krimper. *Molecular Cell*, *59*(4), 564–575. https://doi.org/10.1016/j.molcel.2015.07.017

Wedeles, C. J., Wu, M. Z., & Claycomb, J. M. (2013a). A multitasking Argonaute: exploring the many facets of C. elegans CSR-1. *Chromosome Research*, *21*(6–7), 573–586. https://doi.org/10.1007/s10577-013-9383-7

Wedeles, C. J., Wu, M. Z., & Claycomb, J. M. (2013b). Protection of germline gene expression by the C. elegans argonaute CSR-1. *Developmental Cell*. https://doi.org/10.1016/j.devcel.2013.11.016

Weick, E.-M., Sarkies, P., Silva, N., Chen, R. A., Moss, S. M. M., Cording, A. C., … Miska, E. A. (2014). PRDE-1 is a nuclear factor essential for the biogenesis of Ruby motif-dependent piRNAs in C. elegans. *Genes & Development*, *28*(7), 783–796. https://doi.org/10.1101/gad.238105.114

Weng, C., Kosalka, J., Berkyurek, A. C., Stempor, P., Feng, X., Mao, H., … Miska, E. A. (2018). The USTC co-opts an ancient machinery to drive piRNA transcription in C. elegans. *Genes & Development*. https://doi.org/10.1101/gad.319293.118

Wicker, T., Sabot, F., Hua-Van, A., Bennetzen, J. L., Capy, P., Chalhoub, B., … Schulman, A. H. (2007). A unified classification system for eukaryotic transposable elements. *Nature Reviews Genetics*, *8*(12), 973–982. https://doi.org/10.1038/nrg2165

Wilczynska, A., & Bushell, M. (2015). The complexity of miRNA-mediated repression. *Cell Death & Differentiation*, *22*(1), 22–33. https://doi.org/10.1038/cdd.2014.112

Wilkins, C., Dishongh, R., Moore, S. C., Whitt, M. A., Chow, M., & Machaca, K. (2005). RNA interference is an antiviral defence mechanism in Caenorhabditis elegans. *Nature*, *436*(7053), 1044–1047. https://doi.org/10.1038/nature03957

Wilson, A. S., Power, B. E., & Molloy, P. L. (2007). DNA hypomethylation and human diseases. *Biochimica et Biophysica Acta*

(BBA) - Reviews on Cancer, *1775*(1), 138–162. https://doi.org/10.1016/J.BBCAN.2006.08.007

Xu, F., Feng, X., Chen, X., Weng, C., Yan, Q., Xu, T., … Guang, S. (2018). A Cytoplasmic Argonaute Protein Promotes the Inheritance of RNAi. *Cell Reports*, *23*(8), 2482–2494. https://doi.org/10.1016/J.CELREP.2018.04.072

Yamashiro, H., & Siomi, M. C. (2018). PIWI-Interacting RNA in *Drosophila* : Biogenesis, Transposon Regulation, and Beyond. *Chemical Reviews*, *118*(8), 4404–4421. https://doi.org/10.1021/acs.chemrev.7b00393

Yang, J.-S., Maurin, T., Robine, N., Rasmussen, K. D., Jeffrey, K. L., Chandwani, R., … Lai, E. C. (2010). Conserved vertebrate mir-451 provides a platform for Dicer-independent, Ago2-mediated microRNA biogenesis. *Proceedings of the National Academy of Sciences*, *107*(34), 15163–15168. https://doi.org/10.1073/pnas.1006432107

Yang, P., Wang, Y., & Macfarlan, T. S. (2017). The Role of KRAB-ZFPs in Transposable Element Repression and Mammalian Evolution. *Trends in Genetics : TIG*, *33*(11), 871–881. https://doi.org/10.1016/j.tig.2017.08.006

Yang, Z., & Li, Y. (2018). Dissection of RNAi-based antiviral immunity in plants. *Current Opinion in Virology*, *32*, 88–99. https://doi.org/10.1016/J.COVIRO.2018.08.003

Yigit, E., Batista, P. J., Bei, Y., Pang, K. M., Chen, C.-C. G., Tolia, N. H., … Mello, C. C. (2006). Analysis of the C. elegans Argonaute Family Reveals that Distinct Argonautes Act Sequentially during RNAi. *Cell*, *127*(4), 747–757. https://doi.org/10.1016/J.CELL.2006.09.033

Yu, Y., Gu, J., Jin, Y., Luo, Y., Preall, J. B., Ma, J., … Hannon, G. J. (2015). Panoramix enforces piRNA-dependent cotranscriptional silencing. *Science*, *350*(6258), 339–342. https://doi.org/10.1126/science.aab0700

Zhang, C., Montgomery, T. A., Gabel, H. W., Fischer, S. E. J., Phillips, C. M., Fahlgren, N., … Ruvkun, G. (2011). mut-16 and other mutator class genes modulate 22G and 26G siRNA pathways in Caenorhabditis elegans. *Proceedings of the National Academy of Sciences*. https://doi.org/10.1073/pnas.1018695108

Zhang, D., Tu, S., Stubna, M., Wu, W.-S., Huang, W.-C., Weng, Z., & Lee, H.-C. (2018). The piRNA targeting rules and the resistance to piRNA silencing in endogenous genes. *Science*, *359*(6375), 587–592. https://doi.org/10.1126/science.aao2840

Zhang, Z., Wang, J., Schultz, N., Zhang, F., Parhad, S. S., Tu, S., … Theurkauf, W. E. (2014). The HP1 homolog rhino anchors a nuclear complex that suppresses piRNA precursor splicing. *Cell*, *157*(6), 1353–1363. https://doi.org/10.1016/j.cell.2014.04.030

Chapter 2

Maternal piRNAs are essential for germline development following de novo establishment of endo-siRNAs in *Caenorhabditis elegans*

Bruno F.M de Albuquerque, Maria Placentino, René F. Ketting

Developmental Cell 2015 Aug 24;34(4):448-56. doi: 10.1016/j.devcel.2015.07.010

Summary

The Piwi-piRNA pathway represents a germline-specific transposon-defence system. *C.elegans* Piwi, *prg-1*, is a non-essential gene and triggers a secondary RNAi response that depends on mutator genes, endo-siRNAs (22G RNAs), and the 22G RNA-binding Argonaute protein HRDE-1. Interestingly, silencing of PRG-1 targets can become PRG-1 independent. This state, known as RNAe, is heritable and depends on mutator genes and HRDE-1. We studied how the transgenerational memory of RNAe and the piRNA pathway interact. We find that maternally provided PRG-1 is required for de novo establishment of 22G RNA populations, especially those targeting transposons. Strikingly, attempts to re-establish 22G RNAs in absence of both PRG-1 and RNAe memory result in severe germline proliferation defects. This is accompanied by a disturbed balance between gene-activating and -repressing 22G RNA pathways. We propose a model in which CSR-1 prevents the loading of HRDE-1 and in which both PRG-1 and HRDE-1 help to keep mutator activity focused on the proper targets.

Introduction

The Piwi-piRNA pathway is an RNAi-related mechanism that is essential for germ cell development in most organisms (Ghildiyal & Zamore, 2009; Ketting, 2011; Malone & Hannon, 2009). Loss of this pathway results in strong upregulation of transposon activity, apoptosis, and blocks at various stages of meiosis. In contrast, the *C.elegans* Piwi pathway has been shown to be not acutely required in germ cells (Batista et al., 2008; Cox et al., 1998; Das et al., 2008; Wang & Reinke, 2008). Also the impact of PRG-1 on transposon silencing is very limited (Das et al., 2008). Intriguingly, *prg-1* mutant animals have a so-called mortal germline (Mrt) phenotype (Simon et al., 2014), meaning that the germline deteriorates over generations.

PRG-1 uses Piwi-interacting RNAs (piRNAs; in *C.elegans* named 21U RNAs) to identify targets, on which it triggers the production of endogenous short-interfering RNAs (endo-siRNAs; 22G RNAs) (Bagijn et al., 2012; Lee et al., 2012). This occurs in a process that depends on an RNA-dependent RNA polymerase (RdRP) and mutator proteins (Zhang et al., 2011), in so-called mutator foci (Phillips et al., 2012) that flank bigger peri-nuclear aggregates (P granules). Animals lacking mutator activity display defects in RNAi and activation of various transposable elements (Ketting et al., 1999; Tabara et al., 1999). Different Argonaute proteins, including WAGO-1 (Gu et al., 2009), PPW-1 (Simon et al., 2014), and HRDE-1 (Ashe et al., 2012; Buckley et al., 2012; Luteijn et al., 2012; Shirayama et al., 2012), act as recipients for the 22G RNA output of mutators. Interestingly, whereas *hrde-1* mutants also display a Mrt phenotype (Buckley et al., 2012), mutator mutants do not (Simon et al., 2014). However, mutator genes do affect germline mortality because they are required for the suppression of the Mrt phenotype of *prg-1* by *daf-2* (Simon et al., 2014).

As mentioned, PRG-1 can silence target genes through the involvement of mutator genes. Interestingly, such PRG-1-initiated silencing can become PRG-1 independent (Ashe et al., 2012; Luteijn et al., 2012; Shirayama et al., 2012). This state, referred to as RNAe, can be faithfully inherited across many generations and depends on mutators, the nuclear 22G RNA- binding Argonaute protein HRDE-1, and chromatin factors (Ashe et al., 2012; Luteijn et al., 2012; Shirayama et al., 2012). In this light it is interesting to note that transposon activation is much stronger in mutator mutants than in *prg-1* mutants, as one of the most active transposons in *C.elegans*, Tc1, is still mostly inactive in *prg-1* mutants (Das et al., 2008). Possibly, transposon silencing depends for

a large extent on the PRG-1-independent, but mutator-dependent RNAe-related silencing memory.

In parallel to a memory that transmits silencing, *C.elegans* gametes also transmit information on genes that are active (Conine et al., 2013; Seth et al., 2013; Wedeles et al., 2013). This requires the Argonaute proteins ALG-3, ALG-4, and CSR-1. In fact, CSR-1 can reactivate genes that have been silenced through PRG-1 and mutator activity (Seth et al., 2013). The molecular mechanisms behind this activation are currently not clear. These may involve chromatin-related effects (Claycomb et al., 2009; Wedeles et al., 2013) but could also relate to 22G RNA turnover, since we previously described an enzyme named CDE-1 that is required for CSR-1-bound 22G RNA turnover through non-templated uridylation of CSR-1-bound 22G RNAs (van Wolfswinkel et al., 2009). Whatever its mechanism of action, the CSR-1 pathway is extremely important since loss of its activity leads to embryonic lethality and sterility (Claycomb et al., 2009; Qiao et al., 1995; Rocheleau et al., 2002; Yigit et al., 2006).

We set out to test the idea that transposons are kept silenced through RNAe-related memory and that PRG-1 is required specifically for the initiation of transposon silencing and not for maintenance. To do this we first erased the mutator-mediated silencing memory from *prg-1* mutant animals, and then reactivated this memory system. Indeed, we find that transposon-targeting 22G RNAs require PRG-1 to re-establish, and we demonstrate that PRG-1 and HRDE-1 synergistically act to silence Tc1 transposition. In addition, our experiments reveal an acute requirement for PRG-1 for proper germ cell development. We propose that this defect is related to mis-targeted mutator complexes that start to act, through HRDE-1 and WAGO-1/WAGO-2/WAGO-3, on transcripts that are normally protected from silencing through CSR-1.

Results

Mutator-induced sterility in *prg-1* mutants

To erase RNAe memory from *prg-1* mutants, in the presence of intact mutator activity, we created two strains (*prg-1;mut-7* and *prg-1;mut-16*). Both strains lack 21U RNAs and RNAe- related memory. These lines exhibit strong transposon mobilization (see below) and are fertile. Cross offspring of these two lines will remain *prg-1* defective but will have mutator activity, allowing one to address whether *prg-1* is required to initiate transposon silencing (**Figure 1A**). Unexpectedly, however, the offspring of

these crosses are completely sterile (**Figures 1B**, **1C**, and **S1A**–**S1C**). We first checked the generality of this phenotype by combining *prg-1* mutation with other mutator alleles. For this we used *mut-14* and *smut-1*, two redundant RNA helicases (Phillips et al., 2014). Animals lacking both *mut-14* and *smut-1* were shown to be RNAi defective, and we show here that *mut-14;smut-1* double mutants cannot maintain RNAe (**Figure S1D**). When *prg-1;smut-1;mut-14* mutant hermaphrodites are crossed with *mut-7* animals that lack PRG-1 activity, again a strong sterility phenotype develops (**Figure 1C**). If the sterility really comes from the re-establishment of mutator activity, the above crosses should yield fertile offspring if a third mutator protein is kept inactive. Indeed, *prg-1* mutant animals in which *mut-7* and *mut-16* are complemented while *mut-14* and *smut-1* are kept homozygous mutant are fertile (**Figure 1C**). Consistent with the fact that PRG-1 acts through 21U RNAs, re-establishment of mutator activity in absence of PID-1, a factor required for 21U RNA biogenesis (de Albuquerque et al., 2014), results in the same phenotype (**Figure 1C**).

Overall, the germ cell count in these animals is strongly reduced, and no or only limited numbers of mature germ cells are present. Both gonad arms tend to contain similar numbers of germ cells (**Figure S1E**). The germ cells that are still present in these animals express variable levels of the germ cell marker PGL-1 (**Figure S1F**) (Kawasaki et al., 1998). We did not detect expression of a somatic, neuronal gene *unc-119* in the remaining germ cells (data not shown), indicating that these germ cells are not subject to gross germ cell-to-soma transformation. Interestingly, the precise gonadal phenotype varies strongly among individuals (**Figure 1B**). Collectively, our data show that the *prg-1* pathway is required for normal germ cell development upon de novo establishment of mutator activity. The variable nature of the developmental defect would be consistent with rather stochastic molecular defects underlying the sterility phenotype.

Parental effects of mutator-induced sterility

PRG-1-mediated silencing has a strong maternal component (de Albuquerque et al., 2014). Consistent with this, we find that *prg-1*/+ L1-stage offspring from *prg-1* mutant hermaphrodites have as few 21U RNAs as straight *prg-1* mutant L1 larvae (**Figure 1D**), showing that the vast majority of 21U RNAs detected in wild-type L1 larvae are of maternal origin. We then tested whether maternal or paternal PRG-1 could rescue the sterility and found that loss of maternal PRG-1 is sufficient to trigger sterility upon re-

establishment of mutator activity (**Figure 2B**, right panel). Loss of only maternal (**Figures 1C** and **1E**) or paternal (data not shown) mutator activity does not result in sterility. We conclude that PRG-1 affects fertility mainly through the maternal lineage, whereas mutator activity acts both maternally as well as paternally.

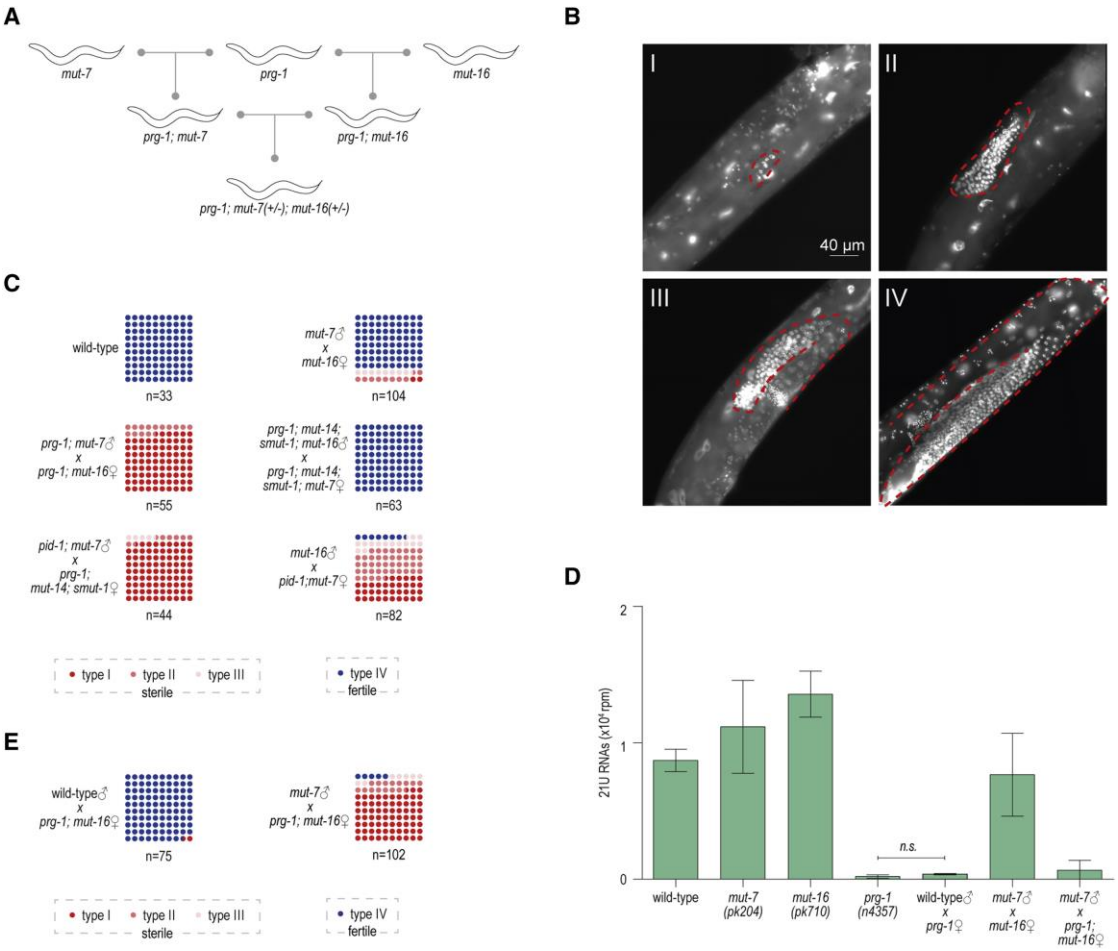


Figure 1. Mutator genes induce sterility in absence of maternal 21U RNAs. **A)** Schematic depicting the crosses performed to erase RNAe memory from *prg-1* mutant animals. **B)** DAPI staining showing the range of germline defects in worms whose mutator activity was restored in the absence of PRG-1. Gonads are outlined by the dashed line. The germlines of individual animals were classified into one of four categories: type I - none or very few germline cells; type II - some germ cells, but no apparent differentiation; type III - some germ cells with apparent differentiation; type IV - wild-type germline. Types I-III animals are sterile. **C, E)** Quantification of the observed germline defects in wild-type animals where the mutator pathway was reactivated in different genetic backgrounds. Each dot represents 1%. **D)** Levels of 21U RNA determined by small RNA-seq of total RNA from L1 larvae of the indicated genotype. Error bars represent SD between biological duplicates. See also **Figure S1**.

PRG-1 and HRDE-1 cooperate to silence Tc1 transposition

We next analysed small RNAs isolated from animals displaying mutator-induced sterility. In order to reduce potential secondary effects stemming from the developing germ cell phenotype, we focused our sequencing efforts on L1 and L2 larvae. We used homozygous wild-type and corresponding mutant strains as controls as well as offspring from a cross between two mutator mutants that have wild-type PRG-1 activity. For all samples, two or three biological replicates were processed, and progeny from crosses were hand-picked

to make sure cross-progeny were analysed. Finally, we included random barcodes in the small RNA libraries for the identification of unique ligation events. We first looked at the abundance of transposon-targeting 22G RNAs. As expected, these 22G RNAs are largely missing in mutator mutants, and their levels stay stable in wild-type and in *prg-1* mutants (**Figure 2A**). This shows that the majority of these 22G RNAs are inherited in a PRG-1-independent manner, consistent with ongoing transposon silencing in *prg-1* mutants. When mutator activity is de novo established, transposon-targeting 22G RNAs start to build up in L2 stage, in a process that depends on PRG-1 (**Figures 2A** and **2B**). This strongly suggests that PRG-1 is required to initiate transposon silencing in *C.elegans*, whereas it is not required for the 22G-RNA-mediated memory of it.

To obtain more direct evidence of this idea, we determined the germline reversion frequency of the *unc-22::Tc1(st136)* allele. We confirmed that in *prg-1* mutants this allele reverts at a very low (Das et al., 2008) but reproducible frequency of ~10^{-5} (**Figure 2C**). In contrast, we could not detect reproducible reversion events in *hrde-1* mutants. Strikingly, in *prg-1;hrde-1* double mutants we observe a 100-fold increase in Tc1 excision frequencies, comparable with those in *mut-16* mutants (**Figure 2C**). Loss of PRG-1 does not further increase Tc1 activity in *mut-16* mutants (**Figure 2C**), consistent with the idea that PRG-1-mediated silencing is fully dependent on mutator activity (Bagijn et al., 2012). Since PRG-1 seems to act primarily via secondary Argonaute proteins, and Tc1 silencing is still intact in *hrde-1* mutants, it is likely that additional Argonaute proteins participate in Tc1 silencing. Indeed, we find that *wago-1;wago-2;wago-3* triple-mutant animals display activation of Tc1 at levels similar to those observed in *prg-1;hrde-1* double mutants or *mut-16* single mutants (**Figure 2C**).

These data suggest that both HRDE-1 and PRG-1 act through WAGO-1/WAGO-2/WAGO-3 to silence Tc1.

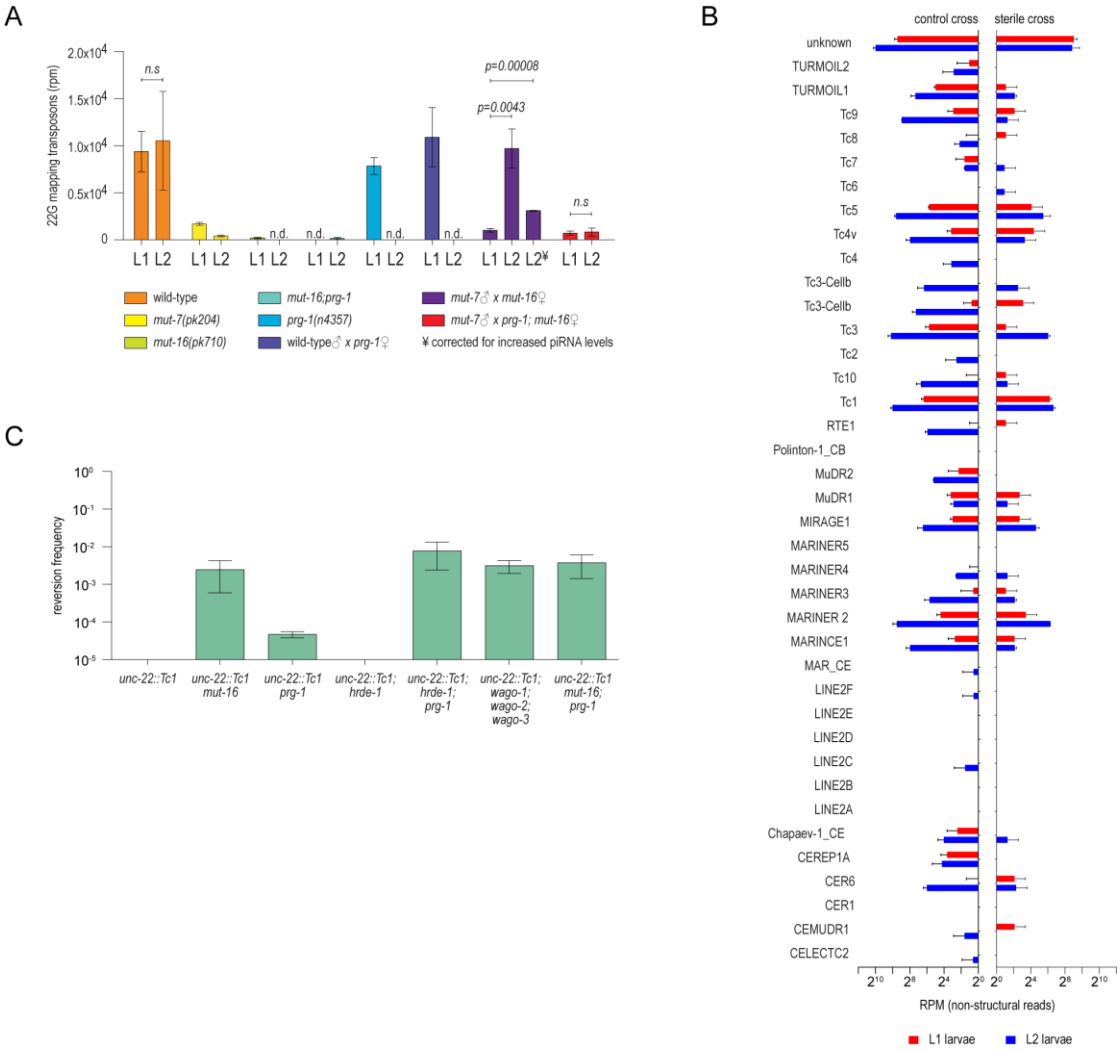


Figure 2. Maternal 21U RNAs are required to re-establish transposon silencing. **A)** Column chart showing levels of 22G small RNAs mapping anti-sense to transposons in L1 and L2 larvae of the indicated genotypes. Error bars represent SD between at least two biological duplicates. ¥: corrected for 21U RNA levels (see Experimental Procedures). **B)** Bar chart showing levels of 22G RNAs targeting various transposon families in larvae where the mutator pathway was reactivated in the presence (control cross; *mut-7* ♂ x *mut-16* ♀) or absence (sterile cross; *mut-7* ♂ x *mut-16;prg-1* ♀) of maternal 21U RNAs. **C)** Column chart depicting the reversion frequency of *unc-22::Tc1* in animals of the indicated genotype. Error bars represent SD among the values obtained from three experiments.

Inappropriate gene silencing in animals with mutator-induced sterility

Since transposon activation per se does not result in sterility in *C. elegans*, the question remains: What causes sterility upon reactivation of mutator activity in absence of 21U RNAs? Within the so-called WAGO-clade, CSR-1 is the only Argonaute required by

itself for fertility (Yigit et al., 2006), and the RdRP enzyme that makes CSR-1-bound 22G RNAs, EGO-1, has been identified as an enhancer of a germline proliferation defect (Qiao et al., 1995). Interestingly, CSR-1 stimulates gene expression (Claycomb et al., 2009), suggesting that loss of gene activity is more detrimental to germ cells than loss of silencing. We hypothesized that mutator-induced sterility may stem from inappropriate silencing of germ-cell-expressed genes. To test this we checked whether expression of endogenous genes is affected through single-worm RT-PCR analysis of a random set of germ-cell-specific transcripts. We found that our ability to detect transcripts from these genes differs among individual sterile animals (**Figure S2A**), suggesting that these genes may be stochastically, inappropriately silenced. Since this assay is blind to the specific germ-cell-defect individuals and may report on secondary defects, we addressed this issue also through visual analysis of a germ-cell-specific fluorescent reporter transgene. We crossed a 21U-targeted mCherry transgene from a *mut-7* mutant male into *prg-1* or *prg-1;mut-16* double-mutant hermaphrodites. Consistent with what we published before (de Albuquerque et al., 2014), all cross offspring of the *prg-1* mutant hermaphrodites showed mCherry expression in the germline. Cross offspring of the *prg-1;mut-16* double-mutant hermaphrodites displayed stochastic silencing of mCherry, even in animals with apparently almost wild-type germline morphology (**Figures 3A** and **S2B**). We conclude that genes that are normally not targeted for silencing can be silenced by mutator activity in absence of both maternal 21U RNAs and mutator information from both parents, possibly triggering the observed sterility phenotype that develops in these animals.

CSR-1 and WAGO pathways recognize mutual targets

The above-described mutator-driven, ectopic silencing of genes suggests that Argonaute proteins that induce silencing, such as WAGO-1 and HRDE-1, have the potential to be effectively loaded with 22G RNAs derived from expressed genes. In other words, WAGO-1 and HRDE-1 should be able to accept 22G RNAs that are normally found enriched in CSR-1. To address this, we re-analyzed published HRDE-1 (Shirayama et al., 2012) and WAGO-1 (Gu et al., 2009) immunoprecipitation (IP) data and found significant amounts of 22G RNAs from typical CSR-1 target genes. To check whether these 22G RNAs represent truly WAGO-1- or HRDE-1-bound 22G RNAs, we made use of the fact that CSR-1-bound 22G RNAs show higher 3' non-templated uridylation than 22G RNAs bound by other Argonautes (van Wolfswinkel et

al., 2009). The uridylation frequencies of typical CSR-1-bound 22G RNAs from the WAGO-1 and HRDE-1 IP datasets are lower than in CSR-1 IPs (**Figure S3**) (Claycomb et al., 2009), suggesting that these 22G RNAs do not stem from CSR-1 contamination but reflect genuine WAGO-1- and HRDE-1-bound 22G RNAs. Consistent with this, a moderate but significant drop of 22G RNAs derived from CSR-1 target genes can be observed upon loss of either *mut-16* or *prg-1* (**Figure 3B**), accompanied by increased uridylation frequencies of CSR-1-pathway 22G RNAs (**Figure 3C**). Given that CSR-1 operates independently of mutator genes (Gu et al., 2009), these data suggest that a non-CSR-1-bound pool of 22G RNAs derived from typical CSR-1 target genes is lost upon loss of PRG-1 or MUT-16.

The reverse is also true. In previously described CSR-1-bound 22G RNA populations (Claycomb et al., 2009), we detect significant amounts of 22G RNAs from genes that are not considered to be typical CSR-1 target genes (**Figure S3**). Importantly, these 22G RNAs show similar uridylation rates compared to 22G RNAs derived from genes considered to be true CSR-1 targets (**Figure S3**), indicating that they are indeed bound by CSR-1 and do not stem from non-CSR-1-bound 22G RNA contaminations. These findings indicate that CSR-1 and WAGO-Argonautes do not bind 22G RNAs from unique genes. Rather, a gene is characterized by a certain ratio in which its 22G RNAs are represented in CSR-1 and WAGO-Argonaute proteins.

To further probe this balance, we checked whether increasing the amount of 22G RNAs from typical CSR-1 targets can result in increased loading of HRDE-1 with such 22G RNAs. Since disruption of *cde-1* leads to an over-accumulation of CSR-1-type 22G RNAs (van Wolfswinkel et al., 2009), we sequenced 22G RNAs from HRDE-1 immunoprecipitates from wild-type and *cde-1* mutant animals. This revealed that in *cde-1* mutant animals, the fraction of CSR-1-target- derived 22G RNAs in HRDE-1 is higher than in wild-type animals (**Figure 3D**). Interestingly, loss of MUT-16 in a *cde-1* mutant animal triggers sterility that can be partially rescued by loss of HRDE-1 (**Figure 3E**), strongly suggesting that in a *cde-1;mut-16* double mutant, HRDE-1 may be effectively silencing CSR-1 target genes. We conclude that CSR-1, WAGO-1, and HRDE-1 are loaded with 22G RNAs from each other's target genes and that a disturbed loading balance among the various Argonaute proteins can have significant effects on germ cell development.

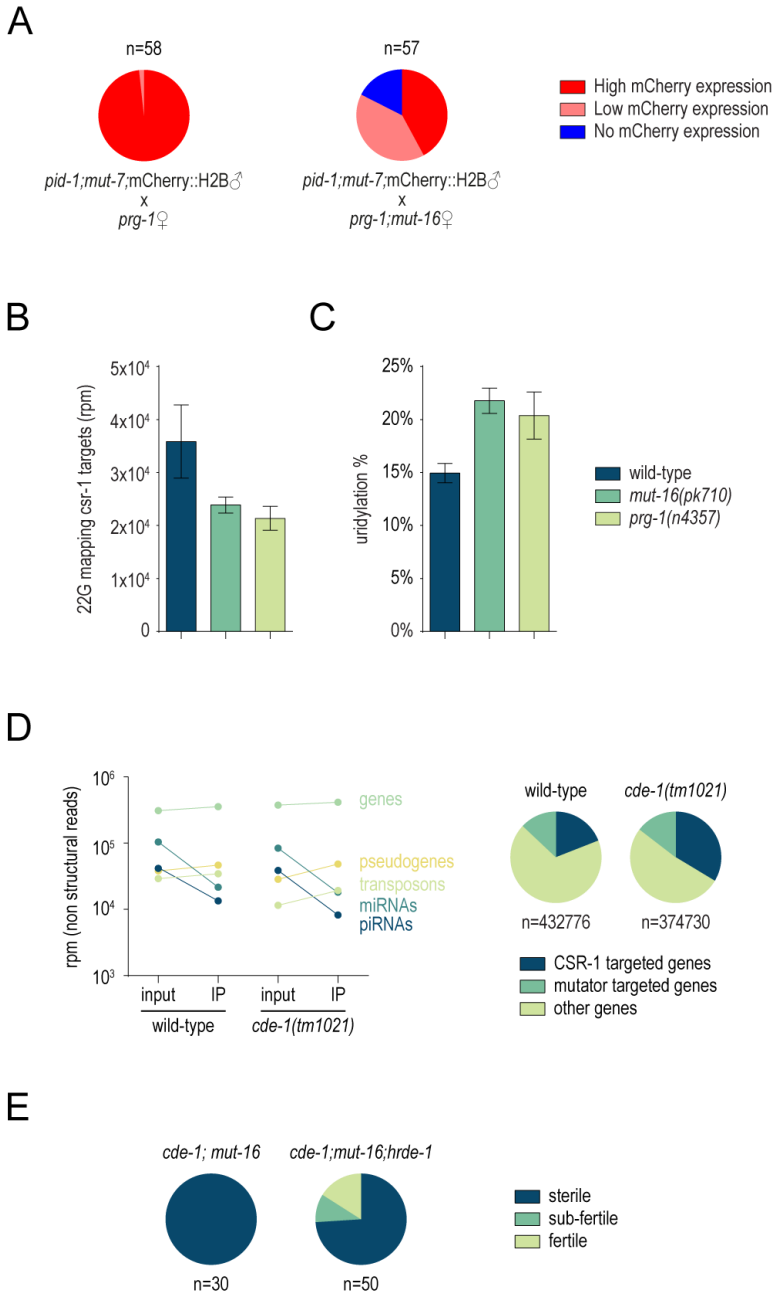


Figure 3. Loading of CSR-1-Type 22G RNAs into HRDE-1 and inappropriate gene silencing in the germline. **A)** Pie chart showing the fraction of offspring that expresses a germline-specific mCherry::H2B transgene, from a control cross (left) and cross that results in sterile offspring (right). **B)** Levels of 22G RNAs mapping to protein-coding genes targeted by CSR-1 (**Table S1**). Error bars represent SD between at least two biological replicates. **C)** Fraction of 22G RNAs mapping to protein-coding genes targeted by CSR-1 that contain non-templated 3' uridylation. **D)** Left panel depicts levels of 21U RNAs, miRNAs, and 22G RNAs anti-sense to genes, pseudogenes, and transposons in input and HRDE-1 immunoprecipitates from either wild- type or *cde-1(tm1021)* animals. Right panel shows the proportion of 22G RNAs mapping to genes annotated as CSR-1 targets and genes annotated as mutator targets in HRDE-1 immunoprecipitates. **E)** Pie chart depicting fertility of *cde-1;mut-16* and *cde-1;mut-16;hrde-1* mutant animals. Also see **Figures S2** and **S3**.

Impact of mutator-induced sterility on 22G RNAs from CSR-1 target genes

We next analysed the 22G RNA content of L1 and L2 larvae that develop mutator-induced sterility. The various small RNA pools, including 21U RNAs and different 22G RNA classes, overall behaved as can be expected from a loss of mutator activity or PRG-1 (**Figure S4**). We then focused on 22G RNAs derived from typical CSR-1 target genes and used uridylation frequencies of 22G RNA pools as a proxy for their physical association with either CSR-1 (high uridylation) or other Argonautes (low uridylation).

In wild-type animals, uridylation frequencies of typical CSR-1 22G RNAs increase when animals develop from L1 into L2 larvae (**Figure 4A**). At the same time, the abundance of these 22G RNAs drops strongly (by ~70%) (**Figure 4B**). In contrast, during the same developmental step, a significant decrease in uridylation of CSR-1 22G RNAs is observed in animals that display mutator-induced sterility, accompanied by only a small drop in 22G RNA abundance (~35%) (**Figures 4A** and **4B**). A possible explanation for these observations could be that in animals that develop mutator-induced sterility, the 22G RNAs from typical CSR-1 targets are in fact not bound by CSR-1 but by another Argonaute protein, triggering ectopic gene silencing and sterility.

HRDE-1 is required to trigger mutator-induced sterility

The nuclear protein HRDE-1 is a good candidate to be loaded with such CSR-1-target-derived 22G RNAs, because HRDE-1 has been shown to be downstream of mutator proteins and PRG-1 (Ashe et al., 2012; Buckley et al., 2012; Luteijn et al., 2012; Shirayama et al., 2012). Furthermore, we have shown that HRDE-1 can be readily loaded with 22G RNAs that normally load into CSR-1 (**Figure S3**) and that mutation of *hrde-1* can partially rescue the sterility of *mut-16;cde-1* mutants (**Figure 3E**). We thus tested whether HRDE-1 is required for the above- described, undue, stochastic silencing of a germline-expressed mCherry reporter transgene during mutator-induced sterility. Indeed, in absence of HRDE-1, this silencing is no longer observed (**Figure 4C**). In fact, the sterility phenotype itself is also largely rescued by loss of HRDE-1 (**Figure 4D**), and, consistent with these data, (Phillips et al., 2015) demonstrate that immuno-purified HRDE-1 binds more 22G RNAs from CSR-1 targets during mutator-induced sterility. Interestingly, a *wago-1/wago-2/wago-3* triple mutation also rescues the sterility (**Figure 4D**). These combined results provide strong evidence that Argonaute-driven silencing drives the mutator-induced sterility phenotype.

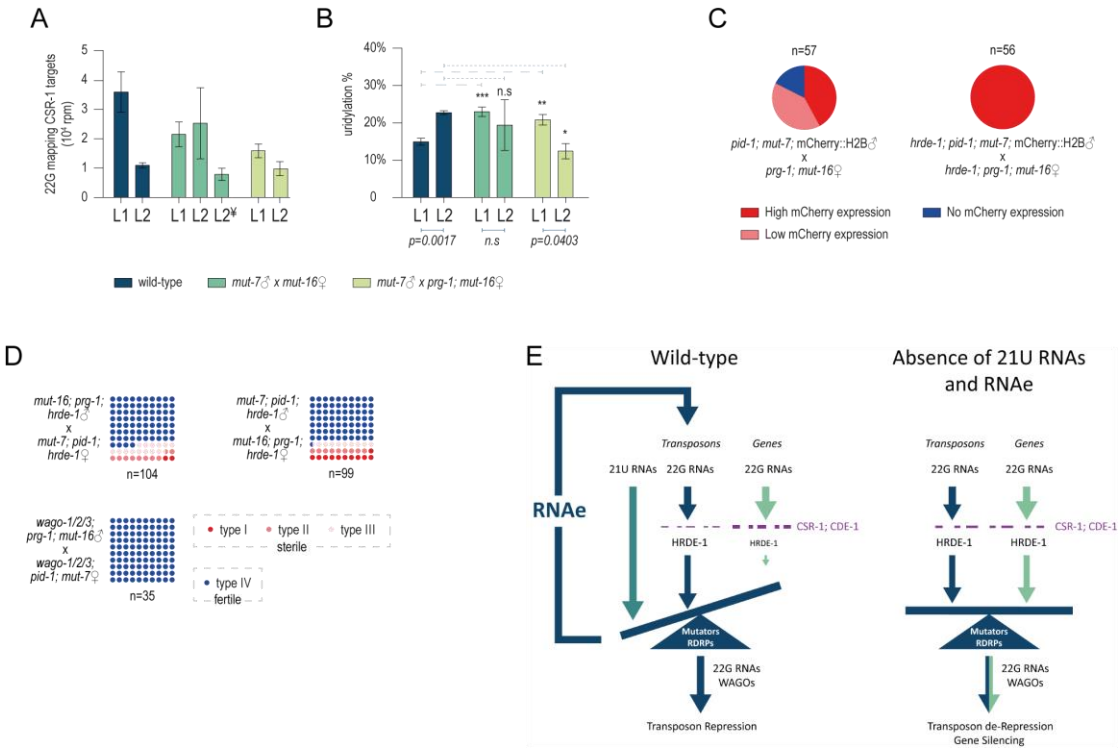


Figure 4. HRDE-1 and WAGO-1/WAGO-2/WAGO-3 drive mutator-induced sterility. **A)** Levels of 22G RNAs mapping to protein-coding genes targeted by CSR-1 (**Table S1**). Error bars represent SD between at least two biological replicates. ¥: corrected for 21U RNA levels (see Experimental Procedures). **B)** Fraction of 22G RNAs mapping to protein-coding genes targeted by CSR-1 that contain non-templated 3' uridylation in L1 and L2 larvae of the indicated genotype. p values where calculated using the two-tailed t test and assuming a normal distribution. Error bars represent SD between at least two biological replicates. **C)** Pie chart showing the fraction of the offspring expressing a germline-specific mCherry::H2B transgene, from the two indicated crosses. The left panel reflects the same data as depicted in **Figure 3A**, right panel. **D)** Quantitation of the observed germline defects in the *hrde-1* mutants and in *wago-1/wago-2/wago-3* triple mutants, where the mutator pathway was reactivated in the absence of functional 21U RNAs. Each dot represents 1%. **E)** Model describing how mutator-induced sterility can develop. See main text for more detailed description. Also see **Figure S4**.

Discussion

Multiple Argonaute proteins drive Tc1 silencing in *C. elegans*

We demonstrate a requirement for PRG-1 in the establishment of de novo transposon 22G RNAs and show that *prg-1*, *hrde-1*, and *wago-1/wago-2/wago-3* are all involved in executing Tc1 silencing. An interesting possibility is that the WAGO-1/WAGO-2/WAGO-3 proteins reflect the actual silencing Argonautes and that PRG-1 and HRDE-1 only bring targeting specificity into the mutator foci that drive 22G RNA biogenesis. In this light, PRG-1 could be seen as the provider of hard-wired (i.e.,

genome-encoded) silencing information, while HRDE-1 provides epigenetic memory of silencing. These findings extend the functional parallel between the *C.elegans* PRG-1 pathway and piRNA activity in other animal species, including the importance of the maternal piRNA pool (Le Thomas, Marinov, & Aravin, 2014; Le Thomas, Stuwe, et al., 2014).

Parental memory of RNAe

This study, as well as previous studies (Alcazar et al., 2008; Ashe et al., 2012; Grishok et al., 2000; Luteijn et al., 2012; Shirayama et al., 2012; Stoeckius et al., 2014), clearly demonstrates that mutator-dependent silencing information inherits through both the paternal and maternal lineages. It is, however, not clear how this memory is precisely transmitted. It seems likely that this occurs in the form of 22G RNAs, but how these small RNAs drive self-renewal is unclear. Given that mutator proteins are clustered in foci (Phillips et al., 2012), one scenario is that certain Argonaute proteins, including PRG-1 and HRDE-1, are capable of targeting mRNAs to these mutator foci, even though at steady state neither PRG-1 nor HRDE-1 has been reported to be in these foci. Since mutator foci are very small, this could be due to a lack of resolution in the experiments. Alternately, such targeting of transcripts to mutator foci may be indirect. For example, chromatin changes induced by HRDE-1 may lead to routing of transcripts from HRDE-1-targeted loci into mutator foci. Such mechanisms have been proposed to act in small- RNA-related chromatin pathways in *Drosophila* and fission yeast (Keller et al., 2012; Klattenhoff et al., 2009; Zhang et al., 2014).

Molecular mechanism behind mutator-induced sterility

Our results suggest that inappropriate targeting of CSR-1 target transcripts, i.e., mRNAs that are expressed in germ cells, by mutator activity, followed by loading of silencing-inducing Argonaute proteins with the resulting 22G RNAs, leads to sterility (see **Figure 4E** for a model). Interestingly, our data indicate that also in wild-type animals, mutators act on CSR-1 targets, although apparently this does not result in silencing of these targets. The easiest explanation for this is that in wild-type animals the small number of CSR-1-type 22G RNAs that is loaded into Argonaute proteins that drive silencing does not suffice to trigger silencing (also see next section). We propose that sterility develops only when 22G RNA production from CSR-1 target genes is amplified, for example, through mis-directed mutator activity, triggered by absence of

both parental 22G-RNA populations and maternal 21U RNAs. Various Argonaute proteins can execute the silencing leading to sterility. Whether these serve redundantly or act in more specialized settings is currently unclear. A second mechanism through which typical CSR-1-target 22G RNAs can be increased is through loss of CDE-1. These can trigger sterility, through HRDE-1, when "regular" mutator-driven 22G RNAs are removed, indicating that the ectopic mutator activity is but one mechanism to disrupt the balance between gene silencing and gene activity.

CSR-1: gene activation or protection from silencing?

How does CSR-1 counteract silencing? This issue has not been fully resolved. Some experiments have indicated that CSR-1 associates with chromatin (Claycomb et al., 2009; Wedeles et al., 2013), and hence a role for CSR-1 in maintaining open chromatin seems plausible. However, whether this is a direct effect of CSR-1 on chromatin or whether such chromatin effects are secondary remains unresolved. CSR-1 is also found on P granules, which are cytoplasmic structures. Therefore, CSR-1's role in maintaining gene expression might also depend on cytoplasmic activities. We propose that an important function of CSR-1 may be to de-stabilize 22G RNAs in a target RNA-dependent manner, similar to what has been described for miRNAs (Ameres et al., 2010). This is inspired by our finding that the number of CSR-1-bound 22G RNAs drop abruptly when animals develop from L1 to L2 stage, when the primordial germ cells (PGCs) become transcriptionally more active, accompanied by a rise in uridylation frequency. We demonstrated before that uridylation of these 22G RNAs by CDE-1 suppresses their abundance (van Wolfswinkel et al., 2009), and we now show that at least a fraction of this expanded pool of 22G RNAs is bound by HRDE-1. Hence, target-dependent 22G RNA de-stabilization might prevent the loading of 22G RNAs into HRDE-1, or similar Argonaute proteins, thus repressing silencing activities on CSR-1-targets. Such a mechanism could provide a silencing threshold, ensuring that a minimum level of 22G RNA production needs to be achieved before silencing takes effect. Mutator activity may have evolved for that very purpose. Future experiments aimed at the loading of individual Argonaute proteins under different experimental conditions will be required to further test these hypotheses.

Experimental procedures

Sample preparation for small RNA sequencing

C. elegans L1 larvae were obtained by bleaching gravid adults and hatching the eggs in M9. L1 cross-offspring larvae were obtained by single picking 200 eggs to an unseeded nematode growth media (NGM) plate, bleaching those eggs to remove any bacteria that were carried along, and allowing the eggs to hatch overnight. L1 larvae were then re-suspended in M9. Cross offspring were identified through the p*unc-119*::GFP transgene brought in via the male. L2 larvae were picked and washed in M9 buffer for each sample.

Total RNA isolation

One hundred fifty L1 or 50 L2 *C. elegans* larvae were washed in M9 buffer (22 mM Na_2HPO_4, 33 mM KH_2PO_4, 86 mM NaCl, and 1 mM $MgSO_4$) and digested in lysis buffer (200 mM NaCl, 100 mM Tris [pH 8.5], 50 mM EDTA, 0.5% SDS, and 200 µg/mL Proteinase-K) for 3 hours at 65 °C followed by 15 min at 95 °C to denature the Prot-K. Lysate was then incubated with DNase I (NEB) for 30 min at 37 °C. Total RNA was then isolated using TRIZOL-LS according to manufacturer instructions and dissolved in 8 µL of H_2O.

Library preparation, sequencing, and data analysis

Detailed procedures are described in the Supplemental Experimental Procedures. Sequencing data are available at GEO: GSE68988.

WAGOs IP data analysis

Sequencing data from CSR-1 IP, WAGO-1 IP, and WAGO-9 IP were obtained from (Claycomb et al., 2009), (Gu et al., 2009), and (Shirayama et al., 2012), respectively. The raw reads in FastQ format were filtered from 5’ barcodes 3’ adaptor sequences using a custom python script, mapped, and processed to the *C.elegans* genome reference WS224, as mentioned before, with the exception that any gene was allowed to be in several “target categories”.

Transposon excision analysis

For each analysed genotype, mutant worms carrying the *unc-22::Tc1* insertion were singled into a 6 cm NGM plate seeded with 100 µl of OP50 and grown at 20 °C. Plates were scored for wild-type moving worms at three different time points: when the total number of worms per plate was around 50, when the total number of worms per plate was around 100, and when the plate was starved, to which we estimated the total number of worms per plate to be 1,000. Transposition frequencies at each time point were calculated using the following formula: $f = -\ln[(T - R) / T] / N$, where T = total number of plates scored, R = number of plates with revertants, and N = number of worms in the plate. Each time point was considered as a biological replicate.

Author contributions

B.F.M.A. designed, executed, and interpreted experiments and performed computational analysis. M.P. planned, executed, and interpreted experiments. R.F.K. designed the study, interpreted results, and wrote the manuscript with input from B.F.M.A. and M.P.

Acknowledgments

We thank IMB Core Facilities, in particular Emil Karaulanov and Chung-Ting Han, for support in library preparation, sequencing, bioinformatic support, and media preparation. We thank members of the Ketting group for stimulating discussions. M.P. was supported by a Boehringer Ingelheim Fonds PhD Fellowship. This work was further supported by a Deutsche Forschungsgemeinschaft grant KE 1888/1-1 (Project Funding Programme; R.F.K.) and a grant from Fundação para a Ciência e Tecnologia ([FCT]SFRH/BD/51001/2010; B.F.M.A.).

References

Alcazar, R. M., Lin, R., & Fire, A. Z. (2008). Transmission dynamics of heritable silencing induced by double-stranded RNA in Caenorhabditis elegans. *Genetics*, *180*(3), 1275–1288. https://doi.org/10.1534/genetics.108.089433

Ameres, S. L., Horwich, M. D., Hung, J.-H., Xu, J., Ghildiyal, M., Weng, Z., & Zamore, P. D. (2010). Target RNA-directed trimming and tailing of small silencing RNAs. *Science (New York, N.Y.)*, *328*(5985), 1534–1539. https://doi.org/10.1126/science.1187058

Ashe, A., Sapetschnig, A., Weick, E. M., Mitchell, J., Bagijn, M. P., Cording, A. C., … Miska, E. A. (2012). PiRNAs can trigger a multigenerational epigenetic memory in the germline of C. elegans. *Cell*. https://doi.org/10.1016/j.cell.2012.06.018

Bagijn, M. P., Goldstein, L. D., Sapetschnig, A., Weick, E.-M., Bouasker, S., Lehrbach, N. J., … Miska, E. A. (2012). Function, Targets, and Evolution of Caenorhabditis elegans piRNAs. *Science*.

Batista, P. J., Ruby, J. G., Claycomb, J. M., Chiang, R., Fahlgren, N., Kasschau, K. D., … Mello, C. C. (2008). PRG-1 and 21U-RNAs Interact to Form the piRNA Complex Required for Fertility in C. elegans. *Molecular Cell*. https://doi.org/10.1016/j.molcel.2008.06.002

Buckley, B. A., Burkhart, K. B., Gu, S. G., Spracklin, G., Kershner, A., Fritz, H., … Kennedy, S. (2012). A nuclear Argonaute promotes multigenerational epigenetic inheritance and germline immortality. *Nature*, *489*(7416), 447–451. https://doi.org/10.1038/nature11352

Claycomb, J. M., Batista, P. J., Pang, K. M., Gu, W., Vasale, J. J., van Wolfswinkel, J. C., … Mello, C. C. (2009). The Argonaute CSR-1 and Its 22G-RNA Cofactors Are Required for Holocentric Chromosome Segregation. *Cell*. https://doi.org/10.1016/j.cell.2009.09.014

Conine, C. C., Batista, P. J., Gu, W., Claycomb, J. M., Chaves, D. A., Shirayama, M., & Mello, C. C. (2010). Argonautes ALG-3 and ALG-4 are required for spermatogenesis-specific 26G-RNAs and thermotolerant sperm in Caenorhabditis elegans. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(8), 3588–3593. https://doi.org/10.1073/pnas.0911685107

Conine, C. C., Moresco, J. J., Gu, W., Shirayama, M., Conte, D., Yates, J. R., & Mello, C. C. (2013). Argonautes promote male fertility and provide a paternal memory of germline gene expression in C. Elegans. *Cell*. https://doi.org/10.1016/j.cell.2013.11.032

Cox, D. N., Chao, A., Baker, J., Chang, L., Qiao, D., & Lin, H. (1998). A novel class of evolutionarily conserved genes defined by piwi are essential for stem cell self-renewal. *Genes and Development*. https://doi.org/10.1101/gad.12.23.3715

Das, P. P., Bagijn, M. P., Goldstein, L. D., Woolford, J. R., Lehrbach, N. J., Sapetschnig, A., … Miska, E. A. (2008). Piwi and piRNAs Act Upstream of an Endogenous siRNA Pathway to Suppress Tc3 Transposon Mobility in the Caenorhabditis elegans Germline. *Molecular Cell*. https://doi.org/10.1016/j.molcel.2008.06.003

de Albuquerque, B. F. M., Luteijn, M. J., Cordeiro Rodrigues, R. J., van Bergeijk, P., Waaijers, S., Kaaij, L. J. T., … Ketting, R. F. (2014). PID-1 is a novel factor that operates during 21U-RNA biogenesis in Caenorhabditis elegans. *Genes and Development*. https://doi.org/10.1101/gad.238220.114

Ghildiyal, M., & Zamore, P. D. (2009). Small silencing RNAs: An expanding universe. *Nature Reviews Genetics*. https://doi.org/10.1038/nrg2504

Grishok, A., Tabara, H., & Mello, C. C. (2000). Genetic requirements for inheritance of RNAi in C. elegans. *Science (New York, N.Y.)*, *287*(5462), 2494–2497. https://doi.org/10.1126/SCIENCE.287.5462.2494

Gu, W., Shirayama, M., Conte, D., Vasale, J., Batista, P. J., Claycomb, J. M., … Mello, C. C. (2009). Distinct argonaute-mediated 22G-RNA pathways direct genome surveillance in the C. elegans germline. *Molecular Cell*, *36*(2), 231–244. https://doi.org/10.1016/j.molcel.2009.09.020

Kawasaki, I., Shim, Y. H., Kirchner, J., Kaminker, J., Wood, W. B., & Strome, S. (1998). PGL-1, a predicted RNA-binding component of germ granules, is essential for fertility in C. elegans. *Cell*, *94*(5), 635–645. https://doi.org/10.1016/S0092-8674(00)81605-0

Keller, C., Adaixo, R., Stunnenberg, R., Woolcock, K. J., Hiller, S., & Bühler, M. (2012). HP1(Swi6) mediates the recognition and destruction of heterochromatic RNA transcripts. *Molecular Cell*, *47*(2), 215–227. https://doi.org/10.1016/j.molcel.2012.05.009

Ketting, R. F. (2011). The Many Faces of RNAi. *Developmental Cell*. https://doi.org/10.1016/j.devcel.2011.01.012

Ketting, R. F., Haverkamp, T. H. A., Van Luenen, H. G. A. M., & Plasterk, R. H. A. (1999). mut-7 of C. elegans, required for transposon silencing and RNA interference, is a homolog of werner syndrome helicase and RNaseD. *Cell*. https://doi.org/10.1016/S0092-8674(00)81645-1

Klattenhoff, C., Xi, H., Li, C., Lee, S., Xu, J., Khurana, J. S., … Theurkauf, W. E. (2009). The Drosophila HP1 homolog Rhino is required for transposon silencing and piRNA production by dual-strand clusters. *Cell*, *138*(6), 1137–1149. https://doi.org/10.1016/j.cell.2009.07.014

Le Thomas, A., Marinov, G. K., & Aravin, A. A. (2014). A transgenerational process defines piRNA biogenesis in Drosophila virilis. *Cell Reports*, *8*(6), 1617–1623. https://doi.org/10.1016/j.celrep.2014.08.013

Le Thomas, A., Stuwe, E., Li, S., Du, J., Marinov, G., Rozhkov, N., … Aravin, A. A. (2014). Transgenerationally inherited piRNAs trigger piRNA biogenesis by changing the chromatin of piRNA clusters and inducing precursor processing. *Genes & Development*, *28*(15), 1667–1680. https://doi.org/10.1101/gad.245514.114

Lee, H. C., Gu, W., Shirayama, M., Youngman, E., Conte, D., & Mello, C. C. (2012). C. elegans piRNAs mediate the genome-wide surveillance of germline transcripts. *Cell*. https://doi.org/10.1016/j.cell.2012.06.016

Luteijn, M. J., Van Bergeijk, P., Kaaij, L. J. T., Almeida, M. V., Roovers, E. F., Berezikov, E., & Ketting, R. F. (2012). Extremely stable Piwi-induced gene silencing in Caenorhabditis elegans. *EMBO Journal*. https://doi.org/10.1038/emboj.2012.213

Malone, C. D., & Hannon, G. J. (2009). Small RNAs as Guardians of the Genome. *Cell*. https://doi.org/10.1016/j.cell.2009.01.045

Phillips, C. M., Brown, K. C., Montgomery, B. E., Ruvkun, G., & Montgomery, T. A. (2015). PiRNAs and piRNA-Dependent siRNAs Protect Conserved and Essential C. elegans Genes from Misrouting into the RNAi Pathway. *Developmental Cell*. https://doi.org/10.1016/j.devcel.2015.07.009

Phillips, C. M., Montgomery, B. E., Breen, P. C., Roovers, E. F., Rim, Y. S., Ohsumi, T. K., … Montgomery, T. A. (2014). MUT-14 and SMUT-1 DEAD box RNA helicases have overlapping roles in germline RNAi and endogenous siRNA formation. *Current Biology*. https://doi.org/10.1016/j.cub.2014.02.060

Phillips, C. M., Montgomery, T. A., Breen, P. C., & Ruvkun, G. (2012). MUT-16 promotes formation of perinuclear Mutator foci required for RNA silencing in the C. elegans germline. *Genes and Development*. https://doi.org/10.1101/gad.193904.112

Qiao, L., Lissemore, J. L., Shu, P., Smardon, A., Gelber, M. B., & Maine, E. M. (1995). Enhancers of glp-1, a gene required for cell-signaling in Caenorhabditis elegans, define a set of genes required for germline development. *Genetics*, *141*(2), 551–569. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8647392

Rocheleau, C. E., Howard, R. M., Goldman, A. P., Volk, M. L., Girard, L. J., & Sundaram, M. V. (2002). A lin-45 raf enhancer screen identifies eor-1, eor-2 and unusual alleles of Ras pathway genes in Caenorhabditis elegans. *Genetics*, *161*(1), 121–131. https://doi.org/10.1051/medsci/20031967725

Seth, M., Shirayama, M., Gu, W., Ishidate, T., Conte, D., & Mello, C. (2013). The C. elegans CSR-1 argonaute pathway counteracts epigenetic silencing to promote germline gene expression. *Developmental Cell*. https://doi.org/10.1016/j.devcel.2013.11.014

Shirayama, M., Seth, M., Lee, H. C., Gu, W., Ishidate, T., Conte, D., & Mello, C. C. (2012). PiRNAs initiate an epigenetic memory of nonself RNA in the C. elegans germline. *Cell*. https://doi.org/10.1016/j.cell.2012.06.015

Simon, M., Sarkies, P., Ikegami, K., Doebley, A.-L., Goldstein, L. D., Mitchell, J., … Ahmed, S. (2014). Reduced insulin/IGF-1 signaling restores germ cell immortality to Caenorhabditis elegans Piwi mutants. *Cell Reports*, *7*(3), 762–773. https://doi.org/10.1016/j.celrep.2014.03.056

Stoeckius, M., Grün, D., & Rajewsky, N. (2014). Paternal RNA contributions in the Caenorhabditis elegans zygote. *The EMBO Journal*, *33*(16), 1740–1750. https://doi.org/10.15252/embj.201488117

Tabara, H., Sarkissian, M., Kelly, W. G., Fleenor, J., Grishok, A., Timmons, L., … Mello, C. C. (1999). The rde-1 gene, RNA interference, and transposon silencing in C. elegans. *Cell*. https://doi.org/10.1016/S0092-8674(00)81644-X

van Wolfswinkel, J. C., Claycomb, J. M., Batista, P. J., Mello, C. C., Berezikov, E., & Ketting, R. F. (2009). CDE-1 Affects Chromosome Segregation through Uridylation of CSR-1-Bound siRNAs. *Cell*, *139*(1), 135–148. https://doi.org/10.1016/j.cell.2009.09.012

Vasale, J. J., Gu, W., Thivierge, C., Batista, P. J., Claycomb, J. M., Youngman, E. M., … Conte, D. (2010). Sequential rounds of RNA-dependent RNA transcription drive endogenous small-RNA biogenesis in the ERGO-1/Argonaute pathway. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(8), 3582–3587. https://doi.org/10.1073/pnas.0911908107

Wang, G., & Reinke, V. (2008). A C. elegans Piwi, PRG-1, Regulates 21U-RNAs during Spermatogenesis. *Current Biology*. https://doi.org/10.1016/j.cub.2008.05.009

Wedeles, C. J., Wu, M. Z., & Claycomb, J. M. (2013). Protection of germline gene expression by the C. elegans argonaute CSR-1. *Developmental Cell*. https://doi.org/10.1016/j.devcel.2013.11.016

Yigit, E., Batista, P. J., Bei, Y., Pang, K. M., Chen, C.-C. G., Tolia, N. H., … Mello, C. C. (2006). Analysis of the C. elegans Argonaute Family Reveals that Distinct Argonautes Act Sequentially during RNAi. *Cell*, *127*(4), 747–757. https://doi.org/10.1016/J.CELL.2006.09.033

Zhang, C., Montgomery, T. A., Gabel, H. W., Fischer, S. E. J., Phillips, C. M., Fahlgren, N., … Ruvkun, G. (2011). mut-16 and other mutator class genes modulate 22G and 26G siRNA pathways in Caenorhabditis elegans. *Proceedings of the National Academy of Sciences*. https://doi.org/10.1073/pnas.1018695108

Zhang, Z., Wang, J., Schultz, N., Zhang, F., Parhad, S. S., Tu, S., … Theurkauf, W. E. (2014). The HP1 homolog rhino anchors a nuclear complex that suppresses piRNA precursor splicing. *Cell*, *157*(6), 1353–1363. https://doi.org/10.1016/j.cell.2014.04.030

Supplemental information

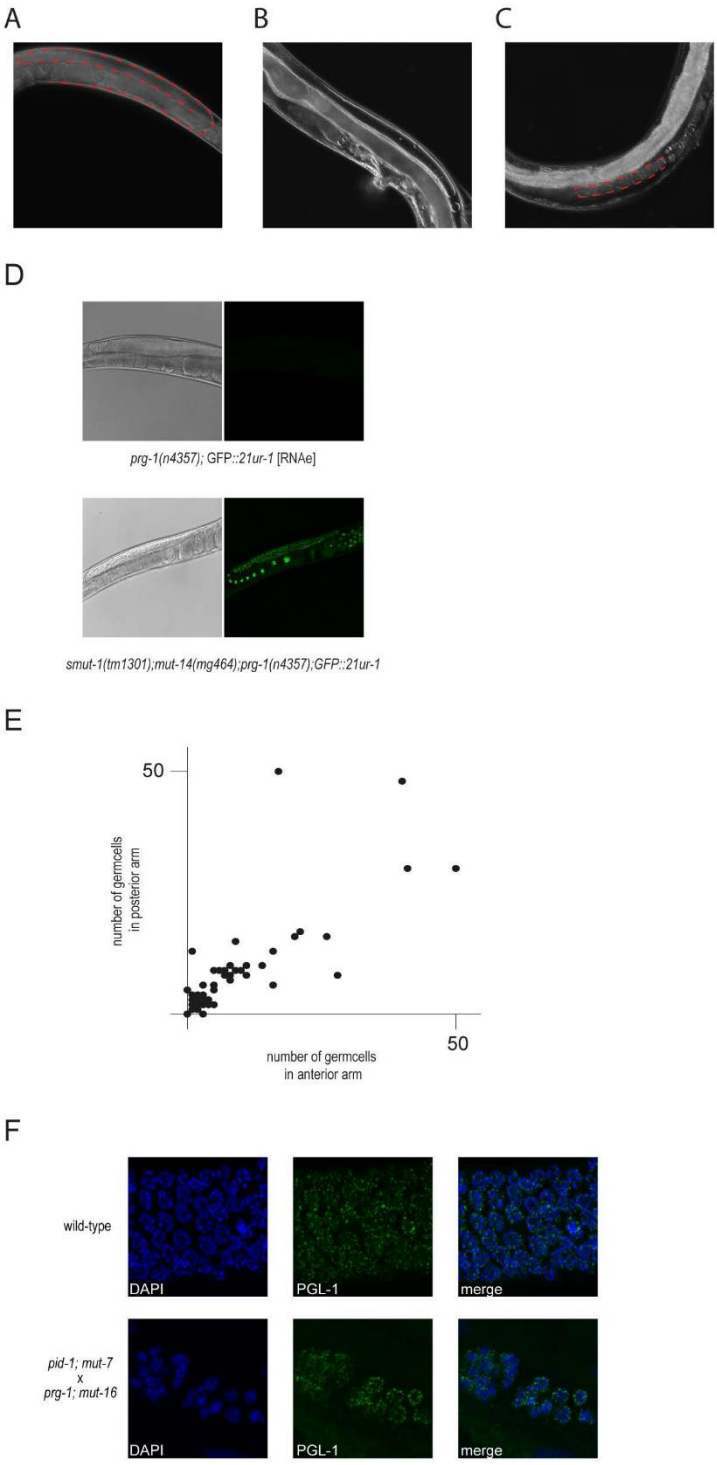
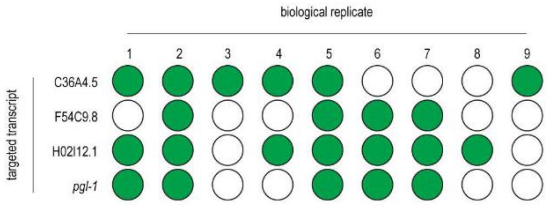


Figure S1. (Related to Figure 1) Mutator-induced sterility phenotypes. DIC picture of: **(A)** wild-type animal and **(B, C)** *prg-1(n4357);mut-7(+/pk204);mut-16(+/pk710)* from a cross between *mut-16(pk710);prg-1(n4357)* hermaphrodites and *mut-7(pk204);prg- 1(n4357)* males. **D)** Confocal image of an epigenetically silenced 21U sensor reactivated in a *smut-1;mut-14* double mutant background. *smut-1* and *mut-14* mutant alleles were crossed into animals deficient for *prg-1* that carried an RNAe silenced 21U sensor. The resulting triple mutant is able to reactivate the sensor, whose silencing had previously

become PRG-1 independent. **E)** Number of germ cells in the posterior an anterior arm of worms displaying Mutator Induced Sterility (n=57). Mutator Induced Sterility was obtained by crossing *pid-1;mut-7* males with *prg-1;mut-16* hermaphrodites. **F)** Confocal image of PGL-1 localization in adult animals displaying Mutator-Induced-Sterility. *prg-1;mut-16* hermaphrodites were crossed with *pid-1;mut-7* males and adult F1s were stained for PGL-1 (green) and DAPI (blue).

A



B

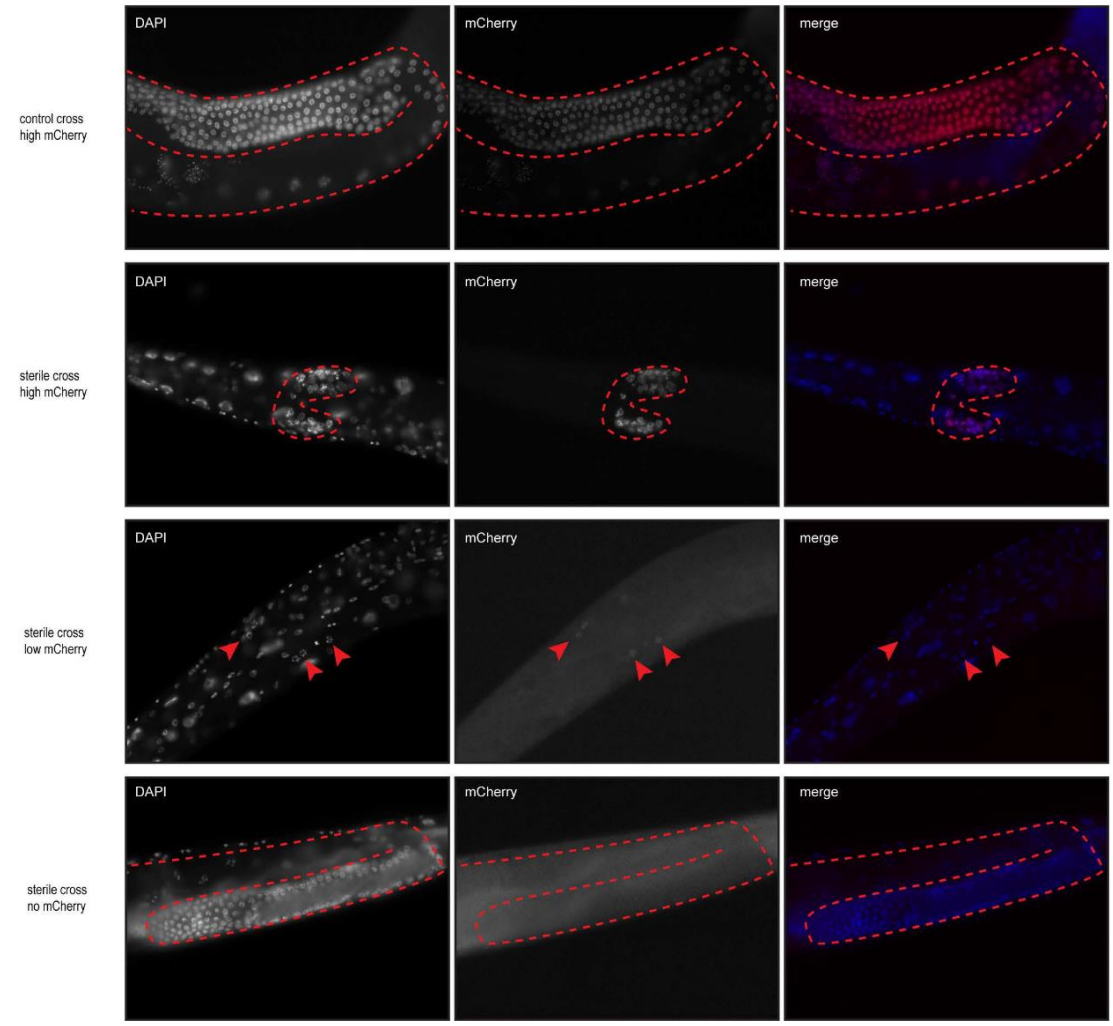


Figure S2. (Related to Figure 3) Gene silencing during Mutator-Induced Sterility. **A)** Graphical representation of the presence or absence of four germline specific genes in animals displaying Mutator

Induced Sterility. Green dots represent presence of the transcript while white dots represent absence of the transcript after 40 cycles of PCR. Each biological replicate consists of a single F1 from a cross between *pid-1;mut-7* males and *prg-1;mut-16* hermaphrodites. **B)** Widefield Fluorescence images of animals carrying a germline specific mCherry::H2B transgene. Germ cells are outlined by the red dashed line or indicated by red arrows. "Control cross" animals were obtained by crossing *mut-7;pid-1;mCherry::H2B* males with *prg-1* hermaphrodites. "Sterile worms" were obtained by crossing *mut-7;pid-1;mCherry::H2B* males with *mut-16;prg-1* hermaphrodites. The mCherry panels depicting animals with low or no mCherry expression have their contrast enhanced for easier visualization.

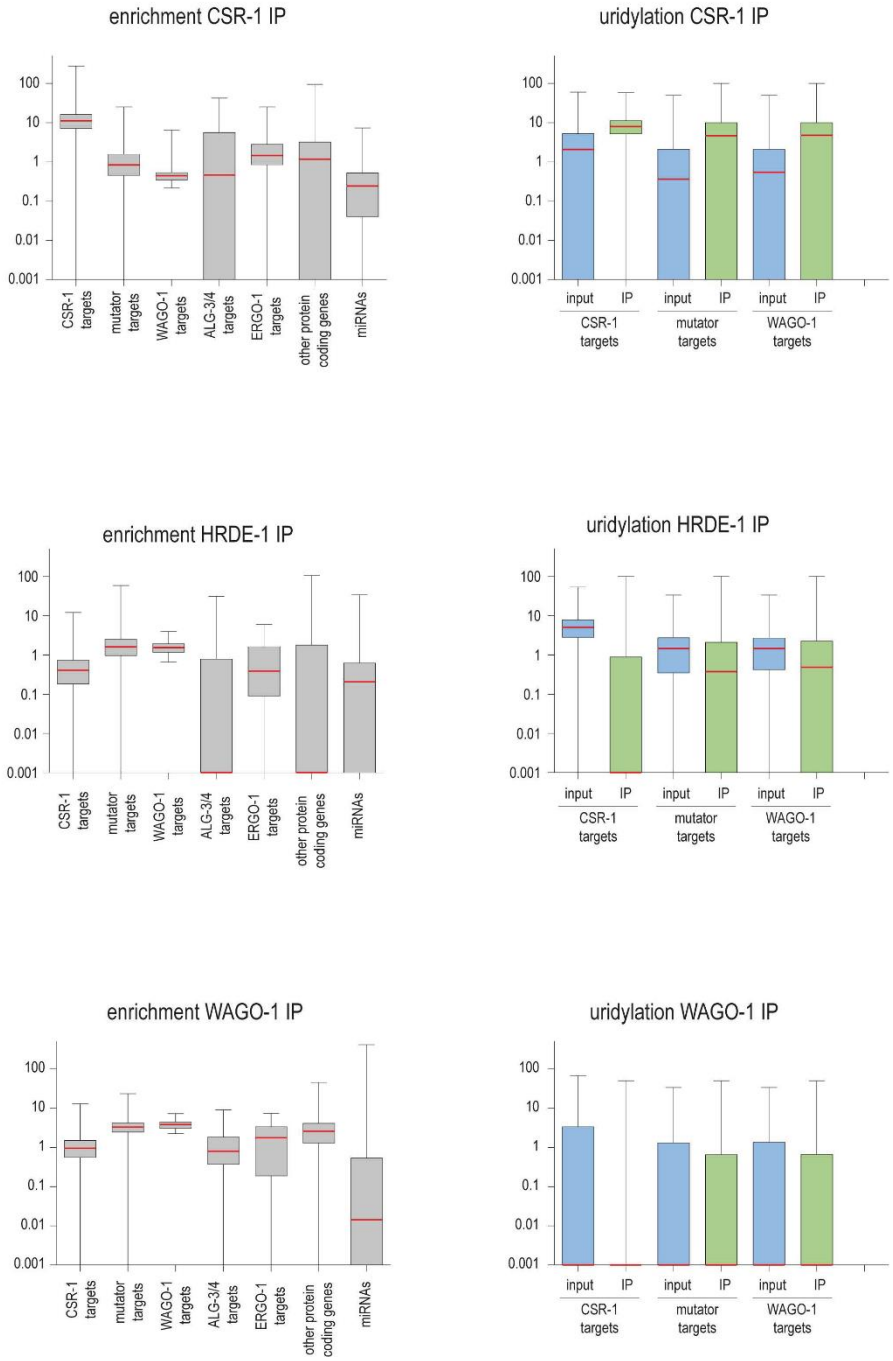


Figure S3. (Related to Figure 3) 22G RNA analysis of different Argonaute IPs. Enrichment and uridylation levels for 22G RNAs antisense to protein coding genes that co-immunoprecipitate with CSR-1, HRDE-1 and WAGO-1. Reads were considered as 22G RNAs if they were 20-23 nucleotides long, started with a G and were anti-sense to annotated genes. Protein coding genes were subdivided into groups as indicated on the X-axis. These sub-divisions were derived from previously published work (see methods and Table S1). Levels are indicated in 'reads per million' of non-structural reads (rpm).

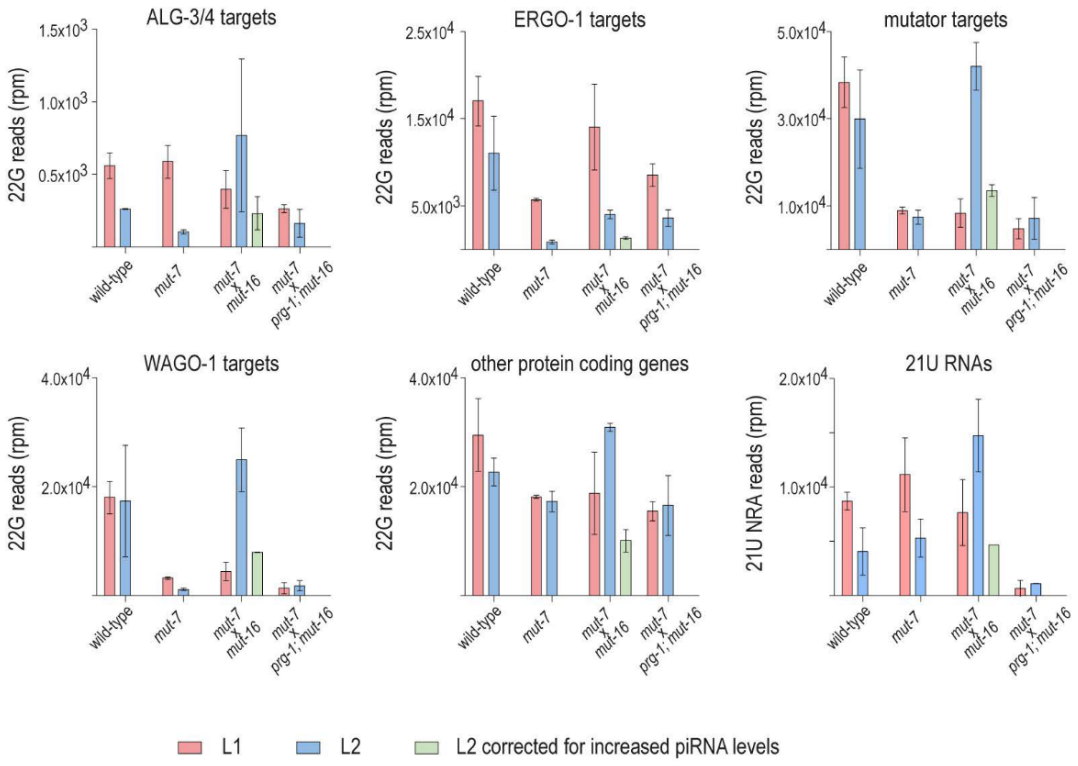


Figure S4. (Related to Figure 4) 22G RNA dynamics in larvae. Column chart showing 22G RNAs to **A)** ALG-3/4 target genes, **B)** ERGO-1 target genes, **C)** mutator target genes, **D)** WAGO-1 target genes, **E)** remaining protein coding genes. **F)** Levels of 21U RNAs. Values are in "reads per million" of non-structural reads (rpm) and errors represent standard deviation between at least two biological replicates.

Table S1. (Related to Figures 3 and 4) This table is an excel sheet containing information on which genes were categorized as a target for which Argonaute. The excel file is available online. http://dx.doi.org/10.1016/j.devcel.2015.07.010

	Sample name	Genotype	Total mapped	Mapped 18-30	Non-structural reads
L2	wild-type_L2_A	N2	2325113	1590416	361940
	wild-type_L2_B	N2	740650	506737	79221
	mut-7_L2_A	RFK179	993562	614892	106718
	mut-7_L2_B	RFK179	1962593	1238411	192970
	mut-16_prg-1_L2_A	RFK232	1152744	767006	165669
	mut-16_prg-1_L2_B	RFK232	1206209	790264	153039
	control_cross_L2_A	RFK258♂ x RFK240♀	1928221	1243183	343456
	control_cross_L2_B	RFK258♂ x RFK240♀	2729489	1752633	326982
	sterile_cross_L2_A	RFK231♂ x RFK232♀	1924652	1261777	260325
	sterile_cross_L2_B	RFK231♂ x RFK232♀	1816006	1075126	203344
L1	wild-type_L1_A	OH441	1529601	1158242	339308
	wild-type_L1_B	OH441	2631423	2008094	495152
	mut-16_L1_A	RFK240	2913868	2135282	525812
	mut-16_L1_B	RFK240	3786140	2577927	569112
	mut-7_L1_A	RFK231	293937	187826	47302
	mut-7_L1_B	RFK231	1178626	818640	198765
	mut-7_L1_C	RFK231	2395818	1803061	404776
	prg-1_L1_A	SX922	1551145	1021268	224469
	prg-1_L1_B	SX922	2803688	1961449	471299
	prg-1_L1_C	SX922	2878124	1658862	405842
	prg-1_cross_L1_A	N2♂ x SX922♀	566551	336655	123145
	prg-1_cross_L1_B	N2♂ x SX922♀	1268714	957918	285274
	prg-1_cross_L1_C	N2♂ x SX922♀	1464043	993611	233270
	control_cross_L1_A	RFK231♂ x RFK240♀	2145206	1502899	448209
	control_cross_L1_B	RFK231♂ x RFK240♀	2965104	1773575	344584
	control_cross_L1_C	RFK231♂ x RFK240♀	824708	562172	160774
	sterile_cross_L1_B	RFK231♂ x RFK232♀	192668	115722	24414
	sterile_cross_L1_C	RFK231♂ x RFK232♀	1819925	1112448	232654
	wild-type_L1_C	N2	647601	519557	105177
HRDE-1 IP	wt_input	N2 input	2163135	784329	542433
	wt_hrde-1_IP	N2 IP	3175071	1493533	1214210
	cde-1_input	RFK90 input	2518793	1075068	845926
	cde-1_hrde-1_IP	RFK90 IP	2414157	1055071	904139

Table S2. (Related to Figures 1, 2, 3 and 4) Sequencing Statistics. Sequencing datasets are available at GEO (GSE68988).

Experimental procedures

Library preparation

All 8 µl of total RNA was treated with 5 U of tobacco acid phosphatase (Epicenter) at 37°C for 2 h to digest 5' tri- and di-phosphates to mono-phosphates. RNA was size-selected between 15- to 35-nt on 15% TBE-urea gel. Gel-purified RNA was eluted overnight in 300 mM NaCl and then precipitated with 100% isopropanol and Glycoblue for 1 h at -20°C. The pellet was washed once with 75% ethanol and dissolved in nuclease-free water. Then, this purified fraction was confirmed by Bioanalyzer Small RNA chip (Agilent). Library preparation was based on the NEBNext Multiplex Small RNA Library Prep Set for Illumina (New England BioLabs) with slight modification. In brief, small RNA was first ligated to the 3' adapter and then the 5' adapter, both of which contained 4 random bases and were chemically synthesized by Bio Scientific.

Adapter-ligated RNA was reverse-transcribed and PCR-amplified for 14 cycles using index primers. The PCR-amplified cDNA construct was purified using AMPure XP beads (Beckman Coulter). The purified PCR reaction was checked on the Bioanalyzer using High Sensitivity DNA chip (Agilent). Size selection of the small RNA library was done on LabChip XT instrument (Perkin Elmer) using DNA 300 assay kit. Only the fraction containing 140-165 bp was pooled in equal molar ratio. The resulting 10 nM pool was denatured and diluted to 10 pM with 5% PhiX spike-in and sequenced as single-read on HiSeq 2500 (Illumina) in rapid mode for 50 cycles using on-board cluster generation. Sample "L1 N2" was sequenced as single-read on Miseq (Illumina). After demultiplexing, on average 35 million passing filter reads were obtained per sample.

Data analysis

The raw reads in FastQ format were filtered from 3' adapter sequences and size-selected in the range 15-35 bases (plus 8 bases random barcodes) using cutadapt v.1.2.1 (Martin, 2011) using parameters: -a AGATCGGAAGAGCACACGTCT -O 8 -m 23 -M 43. Subsequently, PCR clonal reads were deduplicated using Bash and Awk commands. All reads containing low-quality (Phred+33 score less than 20) bases were filtered with the FastX toolkit (http://hannonlab.cshl.edu/fastx_toolkit/; fastq_quality_filter -q 20 -p 100 -Q 33), then the files were reformatted from FastQ into tabular format, sorted and deduplicated based on full sequence identity (library insert

plus 5' and 3' random barcodes of 4 nucleotides), and finally converted back to FastQ format for mapping. Quality assessment of the raw and processed data was done with FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc).

Mapping to the *C.elegans* genome reference WS244 was performed using Bowtie v1 (Langmead et al., 2009) with parameters: -v0 -M1 --best --strata --nomaqround --tryhard -- trimm5 4 --trim3 4. Unmapped reads where then filtered for the ones ending with a "T" using custom python scripts and remapped using bowtie using the same parameter but trimming one extra 3' base. These newly mapped reads were considered mono-uridylated. This filtering/mapping cycle was performed 10 times in order to map up to penta-uridylated reads. All mapped reads where then annotated using bedtools intersect (Quinlan & Hall, 2010) with a customized WS244.gff3 (ftp://ftp.wormbase.org/pub/wormbase/releases/WS244/species/ c_elegans/PRJNA13758/c_elegans/PRJNA13758.WS244.annotations.gff3.gz); using the following parameters: -abam –b [custom annotated gff3] -bed -wa -wb. Mapped and annotated reads were subsequently filtered for size and starting nucleotide using custom made python scripts and normalized to total of non-structural reads between 18 and 30 nucleotides. Structural reads were considered as reads that mapped rRNAs, tRNAs and snoRNAs. During the analysis we considered miRNA reads that were 22 to 24 nt long sense to annotated miRNAs, we considered piRNA reads that were 21 nt long, started with a "T" and were sense to annotated piRNAs, and 22G reads that were 20 to 23 nt long, started with a "G" and map antisense to genes, transposable elements or pseudogenes. In order to categorize genes, we retrieved WAGO-1 targets, ERGO-1 targets, mutator targets, ALG-3/4 targets and CSR-1 targets from (Gu et al., 2009), (Vasale et al., 2010), (Phillips et al., 2014), (Conine et al., 2010), (Claycomb et al., 2009) respectively (also see **Table S1**). For each protein coding gene we saw if it was a WAGO-1 target, if not we saw if it was an ERGO-1 target, if not we saw if it was a mutator target and so on, in the previously mention order. The validation of a gene in a category would automatically exclude it from the following categories; this was to insure that there would be no duplication, since some categories overlap partially. In the L2-samples for the *mut-7* x *mut-16* cross we noted an exceptionally high abundance of 21U RNAs that we believe does not reflect a significant change in piRNA levels, but rather a staging problem. During L2 stage the germline starts to proliferate, and hence this can result in apparent increases in small RNA levels. We therefor also present 'corrected' values for this particular L2 sample where we have scaled down the piRNA

levels to those seen on average in wild-type and mutator-mutant animals. The conversion factor was 3,65 for control_cross_L2_A and 2,63 for cross_L2_B.

Immunohistochemistry

Adult worm gonads were dissected in a poli-L lysine (Sigma P0425) and freeze cracked. Worms were then fixed for 2 minutes in ice cold methanol followed by 2 minute in acetone, and washed 3 times 10 min with PBS-tween (0.05%). Blocking was done with 10% lamb serum in PBT-Tween20 (0.05%). Samples where then incubated O.N. at 4 °C with primary antibody (mouse anti-PGL-1 (Capowski et al., 1991) in block buffer). After O.N. incubation, samples were washed 3 x 10 min with PBS-Tween20 (0.05%), incubated with secondary antibody in block buffer for 1h at room temperature, washed 3 x 10 min with PBS-Tween20 (0.05%) and mounted Fluoroshield with DAPI (Sigma F6057). Samples were then visualized using a Leica TCS SPE confocal.

Single Worm Single Step RT-PCR

RFK257 (*pid-1(xf35)II; mut-7(pk204)III; mjSi22[Pmex-5::mCherry::his-58::21UR-1_as::tbb- 2(3'UTR)]I; otls45 [unc119::GFP]V*) males were crossed with RFK232 (*mut-16(pk710)I;prg-1(n4357)I*) hermaphrodites to obtain sterile cross-offspring. Single adult cross progeny worms were lysed in single worm lysis buffer (25 mM KCl, 1.25 mM MgCl2, 5 mM Tris pH 8.3, 0.2% IGEPAL CA-630, 0.2% Tween20, 0.05% gelatin, 200 µg/ml Prot-K) for 60 min at 65 °C followed by Prot-K denaturation at 95 °C for 15 min. Each lysate was treated with 1U of DNase I (NEB) for 30 min at 37 °C followed by DNA denaturation at 75 °C for 15 min. RNA was precipitated with 100% isopropanol, washed with 75% ethanol and re-dissolved in 30 µl of H_2O. Each RT-PCR reaction mix consisted of 1.5 µl of total RNA, 1 µl of 200 nM primer mix, 2.5 µl of iTaq reaction mix (Biorad) and 0.06 µl of M-MuLV Enzyme Mix (NEB). The RT-PCR program was the following: 15 min at 50 °C, 1 min at 95 °C and 45 cycles of 10 sec at 95 °C and 1 min at 60 °C. The PCR products were resolved in a 4% agarose gel and scored for the presence of absence of amplicon.

Primer list:

C36A4.5 forward – CTCCTGGGTTTGAGGAACCA

C36A4.5 reverse – GGCCTTGGTAAGATCCGC

F54C9.8 forward – TGCAGATGTCGATCCAGAAATTC

F54C9.8 reverse – AGGTTGTCATCGCGACACAG

H02I12.1 forward – GCTGCTTCTACTAACATTTTTTTCG

H02I12.1 reverse – CTGCATGCATTTCCTGCTATGT

pgl-1 forward – CGTCGATAGCTTCAAGAAATTTG

pgl-1 reverse – TCTGCCACTTGGTGTATCGG

Immuno-precipitation

HRDE-1 was precipitated from synchronized adult worms. 50 µl of packed worms were washed with M9 and sonicated in 200 µl of lysis buffer (25 mM Tris pH 7.5, 150 mM NaCl, 1.5 mM MgCl, 1 mM DTT and cOmplete mini, EDTA free (Roche, 1 tablet per 100 mL)) 5 times 30 sec ON, 30 sec OFF at the maximum energy setting. Lysates were centrifuged at 21000 x g for 5 min. 50 µl of the supernatant was used as input and 100 µl was incubated with 1 mg of anti-HRDE-1 antibody (Ashe et al., 2012) and 10 µl of Dynabeads Protein G (Life Technologies) for 2 hours. The beads were washed 5 times 10 min with wash buffer (25 mM Tris pH 7.5, 150 mM NaCl, 1.5 mM MgCl, 1 mM DTT and cOmplete mini, EDTA free (Roche, 1 tablet per 100 mL), 0.2 % Triton X-100). RNA was extracted from both input and IP with TRIZOL-LS (Life Technologies) using the manufacturer instructions and dissolved in 8 µl of H_2O.

Immunohistochemistry

Adult worm gonads were dissected in a poly-L lysine (Sigma P0425) and freeze cracked. Worms were then fixed for 2 minutes in ice cold methanol followed by 2 minute in acetone, and washed 3 times 10 min with PBS-tween (0.05%). Blocking was done with 10% lamb serum in PBT-Tween20 (0.05%). Samples where then incubated O.N. at 4 °C with primary antibody (mouse anti-PGL-1 (Capowski et al., 1991) in block buffer). After O.N. incubation, samples were washed 3 x 10 min with PBS-Tween20 (0.05%), incubated with secondary antibody in block buffer for 1 h at room temperature, washed 3 x 10 min with PBS-Tween20 (0.05) and mounted Fluoroshield with DAPI (Sigma F6057). Samples were then visualized using a Leica TCS SPE-confocal.

Statistical analysis

Samples were compared using two-tailed t-test, assuming a normal distribution. L1 samples OH441 (2x) and N2 (1x) were considered wild-type triplicates.

List of strains used

Strain	Genotype
OH441	*otls45[Punc119::GFP] V*
RFK162	*hrde-1(tm1200) III*
RFK179	*mut-7(pk204) III*
RFK231	*mut-7(pk204) III; mjSi22[Pmex-5::mCherry::his-58::21UR-1_as::tbb-2(3'UTR)] I; otls45[Punc119::GFP] V*
RFK232	*mut-16(pk710) I;prg-1(n4357) I*
RFK233	*mut-16(pk710) I; prg-1(n4357) I; unc-22(st136::Tc1) IV*
RFK234	*mut-7(pk204) III; pid-1(xf35) II*
RFK235	*hrde-1(tm1200) III;mut-7(pk204) III; pid-1(xf35) II*
RFK236	*hrde-1(tm1200) III; mut-16(pk710) I; prg-1(n4357) I*
RFK237	*mut-7(pk204) III; prg-1(n4357) I*
RFK238	*mut-14(mg464) V; smut-1(tm1301) V; mut-16(pk710) I; prg-1(n4357) I*
RFK239	*mut-14(mg464) V; smut-1(tm1301) V; mut-7(pk204) III; prg-1(n4357) I*
RFK240	*mut-16(pk710) I*
RFK241	*hrde-1(tm1200) III; mut-16(pk710)I*
RFK242	*prg-1(n4357) I; mjls144[Pmex-5::egfp::his-58::21UR-1_as::tbb-2(3'UTR)] [RNAe] II*
RFK243	*mut-14(mg464) V; smut-1(tm1301) V; prg-1(n4357) I; mjls144[Pmex-5::egfp::his-58::21UR-1_as::tbb-2(3'UTR)] II*
RFK244	*mut-14(mg464) V; smut-1(tm1301) V; prg-1(n4357) I*
SX922	*prg-1(n4357) I*
RFK245	*mut-16(pk710) I; mjSi22[Pmex-5::mCherry::his-58::21UR-1_as::tbb-2(3'UTR)] I*
RFK246	*prg-1(n4357) I; hrde-1(tm1200) III*
RFK247	*prg-1(n4357) I; unc-22(st136::Tc1) IV*
RFK248	*hrde-1(tm1200) III; unc-22(xf49) IV*
RFK249	*prg-1(n4357) I; hrde-1(tm1200) III; unc-22(xf49) IV*
NL3643	*unc-22(st136::Tc1) IV*
RFK250	*wago-1(tm1414) I; wago-2(tm2686) I; wago-3(tm1120) I; unc-22::Tc1(st136::Tc1) IV*
RFK251	*pid-1(xf35) II; hrde-1(tm1200) III; mut-7(pk204) III; mjSi22[Pmex-5::mCherry::his-58::21UR-1_as::tbb-2(3'UTR)] I; otls45[unc119::GFP] V*
RFK252	*mut-16(pk710) I; unc-22(st136::Tc1) IV*
RFK253	*hrde-1(tm1200) III; unc-22(st136::Tc1) IV*
RFK254	*prg-1(n4357) I; hrde-1(tm1200) III; unc-22(st136::Tc1) IV*
RFK255	*mut-16(pk710) I;prg-1(n4357) I; unc-22(st136::Tc1) IV*

RFK90	*cde-1(tm1021) III*
RFK256	*hrde-1(tm1200) III; cde-1(tm1021) III; mut-16(pk710) I*
RFK257	*pid-1(xf35) II; mut-7(pk204) III; mjSi22[Pmex-5::mCherry::his-58::21UR-1_as::tbb-2(3'UTR)] I; otls45[unc119::GFP] V*
RFK258	*mut-7(pk204) III; mjSi22[Pmex-5::mCherry::his-58::21UR-1_as::tbb-2(3'UTR)] I*

References

Ashe, A., Sapetschnig, A., Weick, E. M., Mitchell, J., Bagijn, M. P., Cording, A. C., … Miska, E. A. (2012). PiRNAs can trigger a multigenerational epigenetic memory in the germline of C. elegans. *Cell*. https://doi.org/10.1016/j.cell.2012.06.018

Capowski, E. E., Martin, P., Garvin, C., & Strome, S. (1991). Identification of grandchildless loci whose products are required for normal germ-line development in the nematode Caenorhabditis elegans. *Genetics*, *129*(4), 1061. Retrieved from http://www.genetics.org/content/129/4/1061.abstract

Claycomb, J. M., Batista, P. J., Pang, K. M., Gu, W., Vasale, J. J., van Wolfswinkel, J. C., … Mello, C. C. (2009). The Argonaute CSR-1 and Its 22G-RNA Cofactors Are Required for Holocentric Chromosome Segregation. *Cell*. https://doi.org/10.1016/j.cell.2009.09.014

Conine, C. C., Batista, P. J., Gu, W., Claycomb, J. M., Chaves, D. A., Shirayama, M., & Mello, C. C. (2010). Argonautes ALG-3 and ALG-4 are required for spermatogenesis-specific 26G-RNAs and thermotolerant sperm in Caenorhabditis elegans. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(8), 3588–3593. https://doi.org/10.1073/pnas.0911685107

Gu, W., Shirayama, M., Conte, D., Vasale, J., Batista, P. J., Claycomb, J. M., … Mello, C. C. (2009). Distinct argonaute-mediated 22G-RNA pathways direct genome surveillance in the C. elegans germline. *Molecular Cell*, *36*(2), 231–244. https://doi.org/10.1016/j.molcel.2009.09.020

Langmead, B., Trapnell, C., Pop, M., & Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology*, *10*(3), R25. https://doi.org/10.1186/gb-2009-10-3-r25

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.Journal*, *17*(1), 10. https://doi.org/10.14806/ej.17.1.200

Phillips, C. M., Montgomery, B. E., Breen, P. C., Roovers, E. F., Rim, Y. S., Ohsumi, T. K., … Montgomery, T. A. (2014). MUT-14 and SMUT-1 DEAD box RNA helicases have overlapping roles in germline RNAi and endogenous siRNA formation. *Current Biology*. https://doi.org/10.1016/j.cub.2014.02.060

Quinlan, A. R., & Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics (Oxford, England)*, *26*(6), 841–842. https://doi.org/10.1093/bioinformatics/btq033

Vasale, J. J., Gu, W., Thivierge, C., Batista, P. J., Claycomb, J. M., Youngman, E. M., … Conte, D. (2010). Sequential rounds of RNA-dependent RNA transcription drive endogenous small-RNA biogenesis in the ERGO-1/Argonaute pathway. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(8), 3582–3587. https://doi.org/10.1073/pnas.0911908107

Contribution to the paper

In the context of this manuscript, I took part in the following experiments:

- Generation of worm strains, maintenance and freezing.
- Microscopy for classification of fertility defects.
- Preparation of samples for small RNA sequencing.
- Tc1 reversion assay.

Chapter 3

Maternal 21U RNAs are sufficient to establish *de novo* silencing and RNAe

Maria Placentino, António Miguel de Jesus Domingues, René F. Ketting

Summary

The Piwi pathway is a germline specific mechanism that organisms have evolved to silence transposable elements. In *C. elegans*, the Piwi protein PRG-1 is guided by its small RNA cofactor, a 21U RNA molecule, to a target mRNA, which is recognised via base pair complementarity. After target recognition, an RNA-dependent RNA polymerase is recruited to the 3' end of the target mRNA to produce secondary antisense siRNAs (22G RNAs), complementary to the target. These 22G RNAs are then loaded onto secondary Argonaute proteins in order to amplify the silencing reaction. This silencing can stochastically become independent of PRG-1 and self-sustainable and it is named RNAe. Using a 21U sensor transgene, we show here that maternally deposited 21U RNAs are not only necessary, but also sufficient for *de novo* target silencing, and that they can in fact establish a stably inherited RNAe state. We also show that maternal 21U RNAs can initiate silencing stochastically at endogenous genomic locations, revealing the relevance of maternal 21U RNAs for endogenous loci as well.

Introduction

The transmission of genetic information from one generation to the next one is the main task of germ cells. In order to accomplish this goal, the genome of germ cells has to be maintained intact and protected from agents that threaten its integrity, such as transposons. Transposons are mobile genetic elements that can insert themselves in the host genome, potentially causing mutations as well as genomic instability. In order to preserve the genome integrity of germ cells, organisms have evolved the Piwi/piRNA pathway to silence transposable elements.

The Piwi pathway is a germline specific pathway in which an Argonaute protein of the Piwi subclade is loaded with a small RNA molecule, called piRNA (Piwi-interacting RNA), to form a complex that recognizes a target RNA via base pair complementarity to silence it. The best established targets of the Piwi/piRNA complexes are transposon-derived transcripts.

Given its role in transposon silencing, the Piwi pathway is essential in most organisms, not only for transposon control, but also for fertility and germ cells development (Ghildiyal & Zamore, 2009; Ketting, 2011; Malone & Hannon, 2009). Consequently, the absence of the Piwi pathway results in transposon upregulation and sterility.

In *Caenorhabditis elegans*, the main Piwi protein is PRG-1 (Piwi-related gene 1) and, together with its small RNA cofactors (21U RNAs), is responsible for silencing one single class of transposons, Tc3 (Das et al., 2008). Nonetheless, PRG-1 is involved in Tc1 transposon silencing as well, yet it is not essential, consequently Tc1 transposons are not reactivated in *prg-1* mutants (de Albuquerque et al., 2015). Different from most organisms, in *C. elegans*, the Piwi pathway is not required for germ cell integrity (Batista et al., 2008; Cox et al., 1998; Das et al., 2008; Wang & Reinke, 2008), although *prg-1* mutants do lose germ cells gradually over generations (mortal germline phenotype, Mrt) (Simon et al., 2014). The PRG-1/21U RNA pathway is also involved in the recognition of self versus non-self, given the ample pool of 21U RNAs and that the target can be recognized with up to 4 mismatches (Bagijn et al., 2012; Lee et al., 2012). As a consequence, the pool of 21U RNAs is potentially able to target for silencing every exogenous DNA sequence (Gu et al., 2012).

After target recognition, an RNA-dependent RNA polymerase (RdRP) is recruited to the target mRNA to produce antisense secondary siRNAs (22G RNAs), to ensure post-

transcriptional gene silencing. The production of 22G RNAs takes place in perinuclear granules called mutator foci, that reside adjacent to P granules (Phillips et al., 2014). Mutator foci contain mutator proteins that are also essential for the production of 22G RNAs (Phillips et al., 2012; Zhang et al., 2011). Loss of mutator proteins results in massive transposon upregulation, suggesting that mutator proteins as well as 22G RNAs are the main responsible for transposon silencing in *C. elegans*, in contrast to other organisms (Ketting et al., 1999; Tabara et al., 1999).

22G RNAs are then loaded onto secondary Argonaute proteins, such as HRDE-1 (Heritable RNAi Deficient-1), to exert post-transcriptional target silencing (Ashe et al., 2012; Buckley et al., 2012; Shirayama et al., 2012). In some cases, this silencing can become extremely stable and self-sustainable, independent of PRG-1, and transmitted for several generations in absence of the initial input. This form of silencing is called RNAe (RNA-induced epigenetic gene silencing) and depends on HRDE-1, mutator proteins and it is characterized by the deposition of heterochromatin marks at the targeted locus (Ashe et al., 2012; Buckley et al., 2012; Lee et al., 2012; Luteijn et al., 2012; Shirayama et al., 2012). However, the molecular mechanisms behind the establishment of RNAe are not yet understood.

We have previously shown that 22G RNAs provided both from the paternal and maternal lineages, as well as 21U RNAs that are maternally contributed, are required to establish appropriate *de novo* gene silencing in the next generation and for germ cells development (see chapter 2) (de Albuquerque et al., 2015; Phillips et al., 2015). We sought therefore to further study specifically the silencing potential of the maternally deposited 21U RNAs. Using a genetic approach, we show that maternally inherited 21U RNAs are not only necessary, but can in fact be sufficient to establish *de novo* silencing. Moreover, we show through small RNA sequencing that the silencing thus established has the characteristics of RNAe, and that it can be stochastically initiated by maternal 21U RNAs also at endogenous loci.

Results

Maternal 21U RNAs are sufficient for *de novo* silencing

The 21U sensor is a germline expressed transgene that codes for a mCherry::H2B fusion protein and has, at its 3' UTR, a sequence antisense to one of the most abundant 21U RNAs of *C. elegans*, *21ur-1* (Bagijn et al., 2012). The *21ur-1* sequence can be

recognized by PRG-1 and the 21U sensor will be silenced. If the silencing pathway is impaired, this transgene will not be silenced by the complex PRG-1/21U RNA and the mCherry::H2B fusion protein will be expressed and can be visualized by microscopy. Hence, this transgene can be used as a reporter of the integrity of the 21U RNA pathway; therefore, it is also named 21U sensor.

It was previously shown that 21U RNAs provided by the maternal lineage are necessary to establish *de novo* silencing of the 21U sensor (de Albuquerque et al., 2014; see also chapter 2). More precisely, if males that lack 22G RNAs, and consequently express the 21U sensor (e.g. *mut-7* mutants), are crossed with wild type hermaphrodites, their offspring will be able to silence the 21U sensor. On the other hand, the same males, expressing the 21U sensor, mated with hermaphrodites lacking 21U RNAs (e.g. *pid-1* or *prg-1* mutants), sire offspring that are no longer able to silence the 21U sensor effectively. The offspring of both crosses are able to produce 21U RNAs themselves, but in the latter case, they do not receive any 21U RNAs from the hermaphrodites, proving that maternally provided 21U RNAs are necessary for *de novo* target silencing (de Albuquerque et al., 2014). We repeated these crosses, and could confirm these results (**Figure 1**).

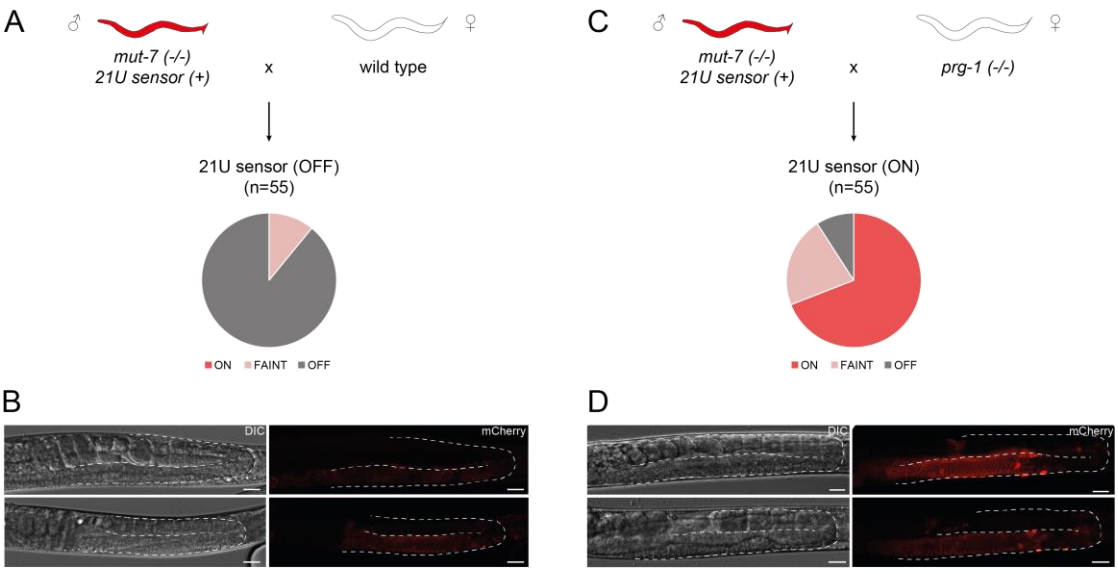


Figure 1. A, C) Scheme representing the cross strategy to address the re-silencing of the 21U sensor. A *mut-7* mutant male expressing the 21U sensor, because of depletion of 22G RNAs, is crossed either with a wild type (**A**) or with a *prg-1* mutant (**C**) hermaphrodite. If their offspring inherit a pool of 21U RNAs from the hermaphrodite, they are able to establish *de novo* silencing on the 21U sensor, in 90% of the worms (see pie chart, **A**). If the hermaphrodite cannot transmit any 21U RNAs to the next generation, the offspring are not able to silence the 21U sensor (90%; see pie chart, **C**). **B, D)** Representative images

of the 21U sensor that is either silenced, very faintly expressed or expressed (left, DIC pictures of one gonad; right: mCherry signal). Gonads are outlined by a dashed line. Scale bar = 25 µm.

Next, we tested if maternally provided 21U RNAs are not only required, but also sufficient for inducing *de novo* silencing. To do so, we crossed *pid-1* mutant males that express the 21U sensor, due to defects in 21U RNA production (de Albuquerque et al., 2014), with *pid-1* heterozygous hermaphrodites. In this situation, all offspring will inherit 21U RNAs from the hermaphrodite, but 50% of the offspring will be *pid-1* homozygous mutant, implying they cannot produce 21U RNAs themselves. In this setup, we can address specifically the impact of maternally provided 21U RNAs in establishing *de novo* target silencing, using the 21U sensor as a readout.

Interestingly, in the offspring that is *pid-1* mutant, the 21U sensor could also be silenced, even though not to a complete extent (**Figure 2**). The sensor was silenced in a similar fraction compared to the *pid-1* heterozygous offspring (**Figure 2**). These results support the hypothesis that maternally inherited 21U RNAs can be sufficient to establish *de novo* silencing, and also that zygotic 21U RNAs do not provide a major contribution to the efficiency of this process.

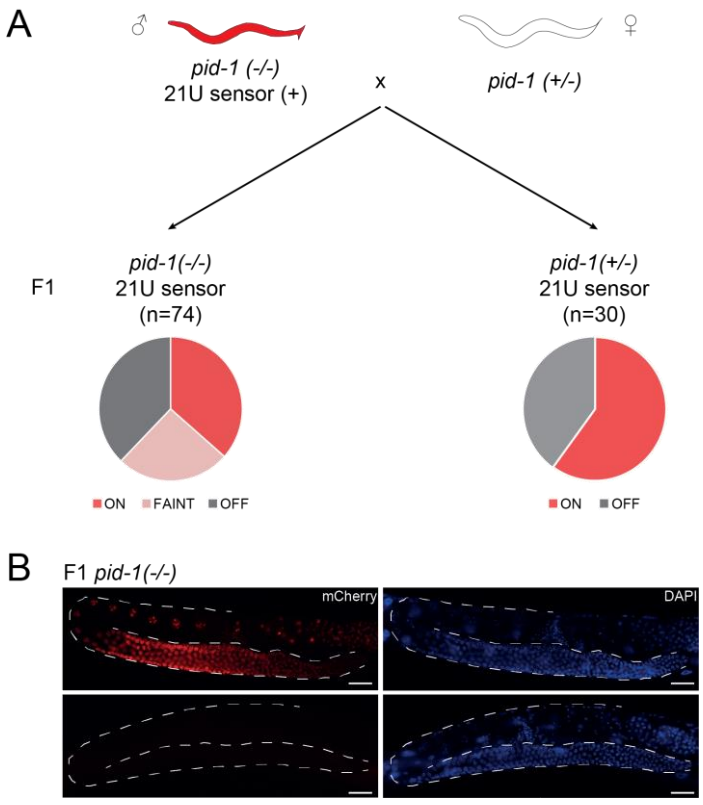


Figure 2. A) Scheme representing the cross strategy to address if maternal 21U RNAs are sufficient to re-silence the 21U sensor. A *pid-1* mutant male expressing the 21U sensor, because of depletion of 21U

RNAs, is crossed with a hermaphrodite that is heterozygote for the same mutation. Their offspring inherit a pool of 21U RNAs from the hermaphrodite and this is sufficient to establish *de novo* silencing of the 21U sensor in a fraction of the *pid-1* mutant offspring (left), similarly to the *pid-1* heterozygous offspring (right). **B)** Representative images of the 21U sensor that is either expressed (upper panels) or silenced (lower panels: left, mCherry signal; right: DAPI staining) in *pid-1* mutant offspring. Gonads are outlined by a dashed line. Scale bar = 25 µm.

Maternal 21U RNAs can initiate RNAe

Next, we tested the transgenerational stability of the silencing induced by the maternal 21U RNAs. Therefore, *pid-1* mutant F1 animals, from the cross depicted in **Figure 2A**, carrying a silenced sensor, were singled out and cultured for additional generations. This revealed that the induced silencing could be stably maintained over months of culturing, with only rare events of sensor reactivation observed within the first few generations. With a generation time of roughly three days, this means that inheritance of silencing can be stable for dozens of generations. This situation, that is inheritance of 21U RNA-induced silencing in absence of a functional 21U RNA pathway, strongly resembles the previously described RNAe status of the 21U sensor (Ashe et al., 2012; Lee et al., 2012; Luteijn et al., 2012; Shirayama et al., 2012). We therefore sought to test whether strains carrying a silenced sensor, isolated from the experiment in **Figure 2**, showed other features known to be characteristic of RNAe.

We know that, amongst other factors, the worm specific Argonaute protein HRDE-1 is required for RNAe maintenance (Ashe et al., 2012; Lee et al., 2012; Luteijn et al., 2012; Shirayama et al., 2012). We therefore introduced an *hrde-1* mutation into the *pid-1* mutant strains carrying a silenced sensor, and found that, upon *hrde-1* mutation, the 21U sensor is reactivated (**Figure 3**).

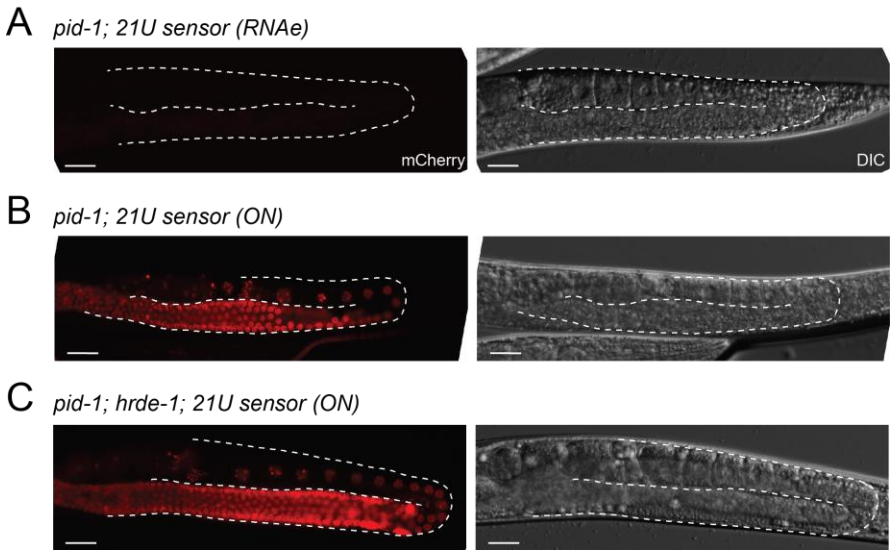


Figure 3. 21U sensor is silenced (**A**) or expressed (**B**) in a strain isolated from *pid-1* offspring that originally have inherited maternal 21U RNAs only. **C)** Upon additional mutation of *hrde-1* in the strain represented in panel **A**, the 21U sensor is reactivated. Left panels represent mCherry expression of the 21U sensor; right panels are DIC images of the same gonads represented on the left. Gonads are outlined by a dashed line. Scale bar = 25 µm.

Another characteristic of RNAe is the production of 22G RNAs along the coding sequence of the target, spreading towards the 5' part of the transcript, relative to the 21U RNA recognition site. Therefore, we performed small RNA sequencing of these *de novo* isolated, *pid-1* mutant strains. In addition, we also sequenced the 22G RNA populations of *pid-1* mutant strains that have been grown in the laboratory for few years already, and have the 21U sensor in a stably active or silenced (RNAe) state (Ashe et al., 2012; Lee et al., 2012; Luteijn et al., 2012; Shirayama et al., 2012). This experiment shows that the 22G RNAs complementary to the 21U sensor have almost identical profiles in both sets of strains, indicating that maternal 21U RNAs themselves, without additional zygotic 21U RNAs, can induce a stably inherited 22G RNA population indistinguishable from that found in previously established strains carrying the 21U sensor in an RNAe state (**Figure 4**).

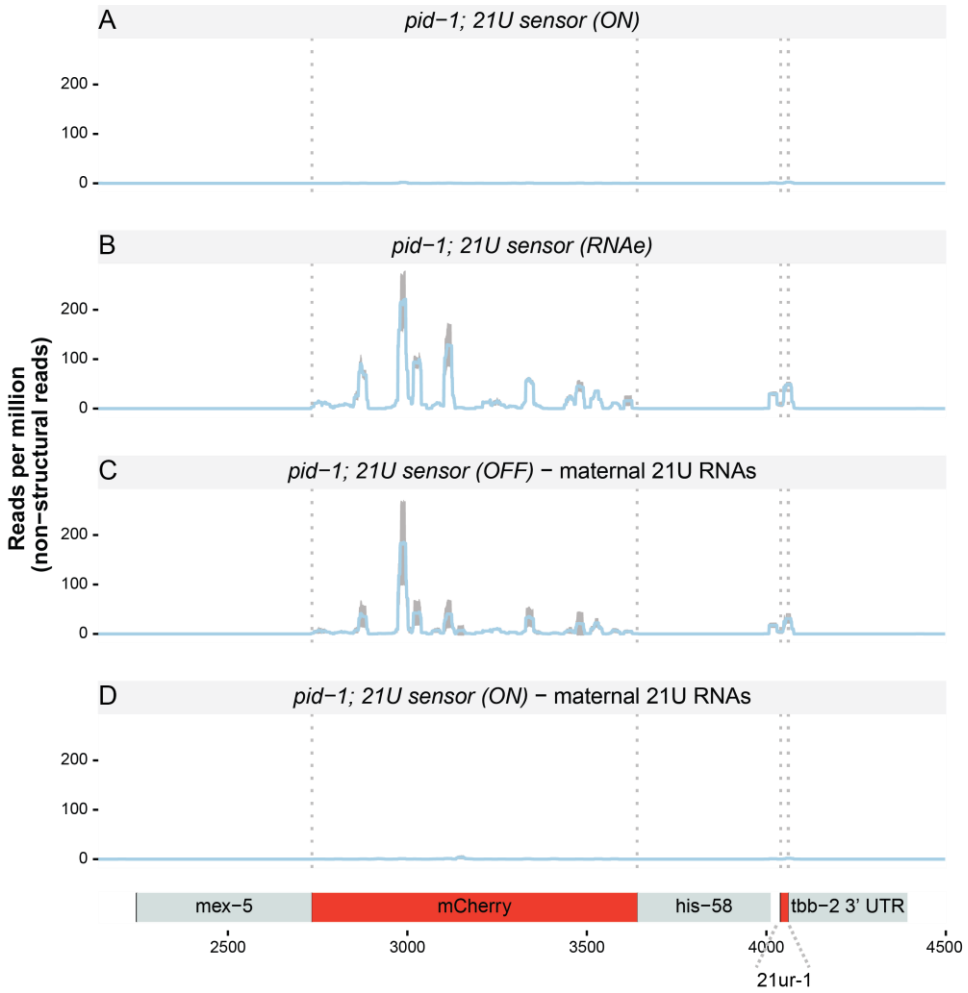


Figure 4. 22G RNAs mapping antisense to the mCherry::21U sensor locus. In the strains that silence the 21U sensor, two distinct populations of 22G RNAs can be observed: a local production around the *21ur-1* recognition site (secondary 22G RNAs) and along the mCherry coding sequence, spreading along the gene towards the 5' end (tertiary 22G RNAs) (Sapetschnig et al., 2015) (**B, C**). When the 21U sensor is expressed, no 22G RNAs mapping complementary to the genomic locus are detected (**A, D**). Sequencing data of the strains that have been originally isolated from the cross in which the offspring have only 21U RNAs inherited from the maternal lineage (**C, D**) show no differences compared to strains of the same genotype that we have grown in the laboratory for few years (**A, B**). In each plot, the average of three biological replicates is represented and the shading represents the standard deviation among the replicates.

Maternally provided 21U RNAs initiate *de novo* silencing at endogenous loci in a stochastic manner

We have shown that maternally provided 21U RNAs are necessary and sufficient to establish *de novo* silencing as well as RNAe on the 21U sensor. The 21U sensor is, however, a transgene, representing a non-physiological situation. Hence, we decided to investigate whether maternally deposited 21U RNAs are also sufficient to establish *de*

novo silencing on endogenous targets. To test this, we performed mRNA sequencing of wild type worms as well as of the *pid-1* strains, originally isolated from the offspring that inherited maternal 21U RNAs only. In particular, we were interested in comparing gene expression variability among the various strains, and whether the status of the 21U sensor (expressed or not) correlates with activity of endogenous targets. In other words, is the activity status of the 21U sensor predictive for effects of maternal 21U RNAs on endogenous transcripts or not?

Principle component (PC) analysis of the genes showing the most variance shows that whilst the gene expression patterns of *pid-1* mutant strains are different compared to the wild type control. Furthermore, gene expression has also a very high variability among *pid-1* mutants; this variability does not reflect the status of the 21U sensor (expressed or silenced), as the strains with the same 21U sensor status do not cluster together (**Figure 5A**). We also observed that hundreds of genes are either up- or down-regulated in *pid-1* mutants, compared to wild type (**Figure 5B**, **C**). The set of genes that is differentially expressed is not the same among the *pid-1* mutant strains, but shows some variability (**Figure 5D**, **E**).

These data suggest that indeed maternally provided 21U RNAs are able to affect the expression levels of endogenous genes; these effects appear to be rather stochastic, possibly reflecting the pool of 21U RNAs that has been inherited. Some of the most variable genes are transposons, pointing out the fundamental role of 21U RNAs in initiating transposon silencing.

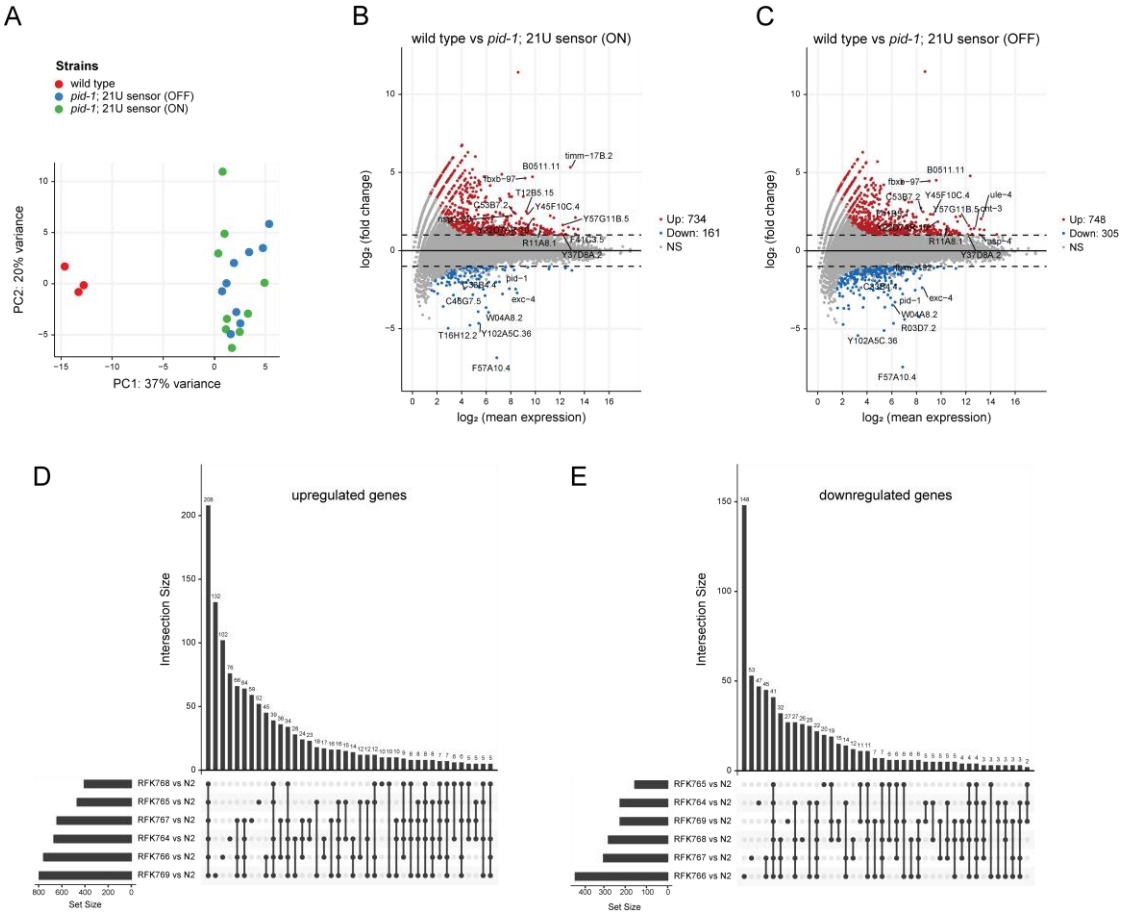


Figure 5. A) PC analysis for the 500 genes that show the largest variance across all samples, comparing wild type and *pid-1* mutant strains. The three wild type replicates cluster together; whereas the *pid-1* mutants show high variability and they do not cluster according to the expression status of the 21U sensor. **B**, **C)** Differential gene expression of *pid-1* mutants versus wild type. Independently of the status of the 21U sensor, several hundreds of genes are up- and down-regulated in *pid-1* mutants, in comparison to wild type worms. **D**, **E)** Upset plots representing the number of genes that are either up- (**D**) or down-regulated (**E**), in comparison to wild type worms, in each of the *pid-1* mutant strains or in multiple strains. It is clear that, even though many genes are commonly up- or down-regulated in all the *pid-1* mutant strains, some changes in gene expression are exclusive to single strains.

Discussion and conclusion

How and when is the decision made: To silence or not to silence?

We have shown that 21U RNAs that are maternally deposited are necessary (chapter 2) (de Albuquerque et al., 2014) and sufficient (this chapter) to initiate *de novo* silencing on a 21U RNA target in *C. elegans*. Maternally deposited 21U RNAs are, however, not sufficient to establish the silencing of the 21U sensor to an extent of 100%. In fact, 60% of the offspring that inherited maternal 21U RNAs is able to escape the silencing, despite being genetically identical to the remaining 40% that is instead able to establish

silencing on the 21U sensor (**Figure 2**). Similar to the different activity status observed for the 21U RNA sensor, also at the endogenous level, we have obtained evidence that the silencing can be induced rather stochastically, indicating that this partial response is not specific to the 21U sensor transgene (**Figure 5**). Given that *C. elegans* is very much inbred, the basis for these differences in expression between individuals cannot be genetic. How then, can these differences arise?

Possibly, this could either reflect the diversity of the original pool of 21U RNAs deposited, that drives silencing of different endogenous targets in different individuals, or be the consequence of an unreached threshold. In order to induce silencing, a certain amount of a targeting 21U RNA species has to be present. Animals that did not silence the 21U sensor, perhaps did not receive a sufficient amount of *21ur-1* to trigger the silencing. Similar scenarios could be sketched for endogenous 21U RNA targets. However, it is also clear that zygotic 21U RNA do contribute to increase silencing efficiency, even though they are not sufficient by themselves. This may indicate that silencing is established early during development, at a time-point when the production of zygotic 21U RNAs does take place, but their levels are simply not sufficient. During later development, when 21U RNA production is rather strong, other factors may be missing to induce silencing.

This stochastic silencing phenomenon reflects what has been already described for RNAe. In fact, we have previously observed that the 21U sensor, that is expressed at low levels in *henn-1* mutants, can become silenced in some individuals and that the silencing is then maintained for several generations (Luteijn et al., 2012). *henn-1* mutants cannot methylate the 21U RNAs, affecting their stability (Billi et al., 2012; Kamminga et al., 2012; Montgomery et al., 2012). Interestingly, we did not observe a decrease of 21U RNAs in *henn-1* mutant adults (Kamminga et al., 2012), whereas others have described a reduction of 21U RNAs in *henn-1* mutant embryos (Billi et al., 2012; Montgomery et al., 2012), indicating a partially functional 21U RNA pathway. Possibly, *henn-1* mutants can provide their offspring with a pool of 21U RNAs, varying in quantity and composition between different individuals. Consequently, only in some individuals, that happen to receive many relative molecules, the silencing of the 21U sensor can be established, while others do not have enough *21ur-1* molecules to trigger RNAe. Finally, even if all the individuals would receive identical amounts of small RNAs, it is still possible to trigger silencing only in some and not in others purely based

on thermodynamics, in case the amount of 21U RNAs is close to a threshold level for inducing silencing.

Is RNAe always established upon *de novo* silencing?

From the above, it is clear that the establishment of RNAe appears to be a rather stochastic process. However, we cannot exclude that in the experiment that addresses silencing by maternal 21U RNAs only (**Figure 2**), some F1 animals initially established silencing, but lost it during development to the adult stage, which is when we analysed the animals. In support of such an idea, even though we could easily isolate strains in which the maternally induced silencing was extremely stable, we did observe a few populations in which silencing was lost in part of the offspring (not shown). It seems that the establishment of a stably inherited form of silencing (RNAe) depends on factors we do not yet understand. Given the involvement of the nuclear Argonaute protein HRDE-1 in RNAe, possibly a certain threshold of 22G RNA production needs to be reached in order to load sufficient amount of HRDE-1 proteins with the required 22G RNAs. We will later describe other experiments that suggest that the 22G RNA response can indeed be tuned at different levels (see chapter 4).

Do maternal 21U RNAs affect chromatin structure?

Another characteristic of RNAe is the deposition of heterochromatic marks, namely H3K9me3, at the silenced locus. In this chapter, we did not check for H3K9 methylation status at the 21U sensor locus nor at the genomic level, to investigate the chromatin landscape at endogenous loci. Nevertheless, as our data support that the establishment of RNAe is initiated by maternally provided 21U RNAs (**Figures 3** and **4**), it is likely that they also affect the chromatin structure of the 21U sensor locus as well as of those endogenous loci that are silenced. Further experiments are required to prove this hypothesis and conclusively confirm that maternally inherited 21U RNAs can drive RNAe.

Which other factors are co-responsible for the stochastic silencing triggered by maternal 21U RNAs?

Different small RNA pathways are interconnected to guarantee a correct balance between gene expression and silencing (Claycomb et al., 2009; Conine et al., 2013; de

Albuquerque et al., 2015; Phillips et al., 2015; Seth et al., 2013; Wedeles et al., 2013). It has been proposed that the silencing activity exerted by PRG-1/21U RNAs and its downstream WAGO/22G RNAs effectors is counteracted by CSR-1, another WAGO protein that targets transcripts that are, and need to be, expressed (Claycomb et al., 2009; Conine et al., 2013; Seth et al., 2013). Whether a transcript will be expressed or silenced has been proposed to depend on the relative amount of PRG-1 or WAGO silencing- and CSR-1 activating-complexes that cover the mRNA molecule (Shen et al., 2018). Hence, the 21U sensor activity may well be affected by the CSR-1 pathway. Indeed, the 21U sensor codes for a mCherry::H2B fusion protein, and the gene coding for histone H2B (*his-58*) is targeted by the CSR-1 pathway. Obviously, the histone H2B needs to be expressed. In fact, we never observed strong production of 22G RNAs on the portion of the 21U sensor mRNA encoding the histone H2B (**Figure 4**; see also chapter 4), suggesting that the CSR-1 pathway is counteracting 22G RNAs production from this part of the 21U sensor. Given that this effect is very local, possibly CSR-1 only acts at the level of individual 22G RNAs, and not of the complete transcript or locus. This may relate to the fact that 22G RNAs loaded on CSR-1 are destabilized by 3' uridylation by CDE-1, to counteract their accumulation and mis-loading on other WAGO proteins, thereby inducing their degradation (van Wolfswinkel et al., 2009).

The 21U RNAs are not the sole molecules involved in promoting gene silencing. WAGO as well as mutator proteins are in fact required for RNAe establishment and maintenance, together with 22G RNAs. We also recently identified novel factors (PID-2, PID-4 and PID-5) that are involved in promoting RNAe initiation and stimulate 22G RNAs production, following 21U RNA target recognition (see chapter 4). Their abundance and/or activity could be the key for modulating the silencing pathway, contributing to the stochastic nature of RNAe establishment. In order to understand the molecular mechanisms behind RNAe initiation, we need to start from the characterization of these novel factors and their interactions.

We have already shown that parentally inherited 22G RNAs as well as RNAe memory (see also chapter 2) (de Albuquerque et al., 2015; Phillips et al., 2015) are essential for establishing *de novo* gene silencing and to ensure appropriate gene expression as well as gonad development. This study reinforces the importance of parental deposition of small RNAs in zygotes in order to guide the silencing machinery to its targets, such as transposable elements, and to avoid the silencing of endogenous genes, which could be detrimental for the individual.

Methods

Strains maintenance

Worm strains have been grown according to standard laboratory conditions on NGM plates seeded with *Escherichia coli* OP50 and grown at 20 °C, unless otherwise stated (Brenner, 1974). We used the N2 Bristol strain as wild type strain.

List of strains

Strain	Genotype
	wild type N2
RFK231	*mjSi22 [Pmex-5::mCherry::his-58::21UR-1_as::tbb-2(3'UTR)] I; mut-7(pk204) III; otls45 [Punc119::GFP] V*
SX922	*prg-1(n4357) I*
RFK315	*mjSi22 [Pmex-5::mCherry::his-58::21UR-1_as::tbb-2(3'UTR)] I; pid-1(xf35) II; otls45 [unc119::GFP] V*
RFK182	*pid-1(xf35) II*
NL3643	*unc-22(st136::Tc1) IV*
RFK184	*mjSi22 [Pmex-5::mCherry::his-58::21UR-1_as::tbb-2(3'UTR)] I; pid-1(xf35) II*
RFK422	*mjSi22 [Pmex-5::mCherry::his-58::21UR-1_as::tbb-2(3'UTR)] [RNAe] I; pid-1(xf35) II*
RFK770	*mjSi22 [Pmex-5::mCherry::his-58::21UR-1_as::tbb-2(3'UTR)] I; pid-1(xf35) II; hrde-1 (tm1200) III*
RFK764	*mjSi22 [Pmex-5::mCherry::his-58::21UR-1_as::tbb-2(3'UTR)] I; pid-1(xf35) II*
RFK765	*mjSi22 [Pmex-5::mCherry::his-58::21UR-1_as::tbb-2(3'UTR)] I; pid-1(xf35) II*
RFK766	*mjSi22 [Pmex-5::mCherry::his-58::21UR-1_as::tbb-2(3'UTR)] I; pid-1(xf35) II*
RFK767	*mjSi22 [Pmex-5::mCherry::his-58::21UR-1_as::tbb-2(3'UTR)] [RNAe] I; pid-1(xf35) II*
RFK768	*mjSi22 [Pmex-5::mCherry::his-58::21UR-1_as::tbb-2(3'UTR)] [RNAe] I; pid-1(xf35) II*
RFK769	*mjSi22 [Pmex-5::mCherry::his-58::21UR-1_as::tbb-2(3'UTR)] [RNAe] I; pid-1(xf35) II*
RFK771	*mjSi22 [Pmex-5::mCherry::his-58::21UR-1_as::tbb-2(3'UTR)] I; pid-1(xf35) II; hrde-1 (tm1200) III*
RFK772	*mjSi22 [Pmex-5::mCherry::his-58::21UR-1_as::tbb-2(3'UTR)] I; pid-1(xf35) II; hrde-1 (tm1200) III*
RFK773	*mjSi22 [Pmex-5::mCherry::his-58::21UR-1_as::tbb-2(3'UTR)] I; pid-1(xf35) II; hrde-1 (tm1200) III*

Microscopy

20-25 worms have been picked to a drop of M9 (80 µl) on a slide, washed and then fixed with acetone (2 x 80 µl). After acetone has evaporated, worms have been washed 2 x 10 minutes with 80 µl of PBS-Triton X100 0,1%. After removing the excess of

PBS-Triton X100 0,1%, the worms have been mounted on a coverslip with Fluoroshield™ with DAPI (5 µl) (Art. No. F6057, Sigma).

Alternatively, for live imaging, 20-25 worms have been picked to a drop of M9 (80 µl) on a slide, washed and then 2 µl of 1 M NaN_3 have been added for 10 minutes to paralyze the worms. After removing the excess of M9, a slide prepared with 2% agarose (in water) has been placed on top of the coverslip and worms have been imaged directly.

Images in **Figures 1**, **2** and **3** have been acquired at a Leica DM6000B microscope. Images have then been processed with Leica LAS software and ImageJ.

For scoring the 21U sensor as active or silenced (**Figures 1**, **2**, **3**), we have used a Leica M165FC widefield microscope to look at live worms on plate. The 21U sensor has been scored as: active, if the fluorescence was easily visible with a lower magnification (Plan APO 1.0x, Art. No. 10450028; Leica); faint, if the fluorescence was only visible with a higher magnification (Plan APO 5.0x/0.50 LWD, Art. No. 10447243; Leica); silenced, if no fluorescence was visible. The worms have been later used also for imaging with a Leica DM6000B microscope as described above.

Small RNA sequencing

RNA extraction

Synchronized gravid adults have been collected with M9 and fast frozen on dry ice in 250 µl of Worm Lysis Buffer (200 mM NaCl; 100 mM Tris HCl pH=8.5; 50 mM EDTA pH=8; 0,5% SDS). 30 µl of Proteinase K (20 mg/ml; Art. No. 7528.1, Carl Roth) have been added to dissolve the worms for 90 minutes at 65 °C with gentle shaking. Lysate has been centrifuged at maximum speed for 5 minutes at room temperature (RT) and the supernatant was transferred on a Phase Lock Gel tube (Art.No. 2302830, QuantaBio). 750 µl TRIzol LS (Art. No. 10296028, Invitrogen™) have been added per 250 µl of sample and, after homogenization, the samples have been incubated for 5 minutes at RT to allow complete dissociation of the nucleoprotein complex. Then 300 µl of chloroform (Art. No. 288306, Sigma-Aldrich) were added per 750 µl of TRIzol LS and the samples were incubated for 15 minutes at room temperature after mixing. Samples have been centrifuged at 12000 x g for 5 minutes at RT and another round of chloroform extraction has been performed. The aqueous phase has been then transferred to an Eppendorf tube and 500 µl of cold isopropanol was added to precipitate the RNA;

samples have been mixed vigorously, incubated at RT for 10 minutes and spun down at maximum speed for at least 10 minutes at 4 °C. The pellet was then washed twice with 1 ml of 75% ethanol and centrifuged for 5 minutes at 7500 x g at 4 °C. The pellet has been dried and diluted in 50 µl of nuclease-free water with gentle shaking for 10 minutes at 42 °C. In order to remove any contamination of genomic DNA, 5 µl of 10X TURBO™ DNase Buffer and 1 µl of TURBO™ DNase (Art. No. AM2238, Invitrogen™) were added to the RNA and incubated at 37 °C for 30 minutes with gentle shaking. The reaction has been stopped by adding 5 µl of 10X TURBO™ DNase Inactivation Reagent. Samples have been centrifuged at 10000 x g for 90 seconds and RNA transferred to a fresh tube. RNA quality has been assessed at Nanodrop and on agarose gel and then samples have been further processed for enrichment of small RNA populations.

Small RNAs enrichment

In order to enrich for small RNAs, we used the *mir*Vana™ kit (Art. No. AM1561, Invitrogen™). 400 µl of *mir*Vana™ Lysis/Binding buffer and 48 µl of *mir*Vana™ Homogenate Additive have been added to the total RNA (80 µl). The mix has been incubated at RT for 5 minutes to denature RNA, then 1/3 of volume of 100% ethanol has been added and after mixing, samples have been spun down at 2500 x g for 4 minutes at RT to pellet large RNAs (>200 nt). The supernatant has been transferred to a new Eppendorf tube and RNA has been precipitated at -80 °C for 1 h with isopropanol (1:1). Samples have been centrifuged at maximum speed for at least 10 minutes at 4 °C to pellet small RNAs. The pellet has then been washed twice with 75% ethanol and spun down at maximum speed for 5 minutes at 4 °C. Pellet has been dried and resuspended in 16 µl of nuclease-free water. RNA quality has been checked at Nanodrop and on agarose gel and further processed for library preparation and deep sequencing.

Library preparation and sequencing

For each strain, three biological replicates have been used for RNA extraction and library preparation. Total RNA was treated with RppH (RNA 5' Pyrophosphohydrolase, Art. No. M0356S, New England Biolabs) to dephosphorylate small RNAs and specifically enrich for 22G RNAs, as previously described (Almeida et al., 2019). For each sample, 1 µg of RNA was incubated for 1 hour at 37 °C with 5 units of RppH and 10X NEB Buffer 2. After dephosphorylation, 500 mM EDTA was added and samples

were incubated for 5 minutes at 65°C to stop the RppH treatment. Small RNAs (15-30 nt) were enriched by gel size selection of RppH-treated RNA prior to library preparation. RNA samples were run on a 15% TBE-Urea polyacrylamide gel (BioRad), and the 15-30 nt fraction was resected from the gel and purified with sodium chloride/isopropanol precipitation.

NGS library preparation was performed with NEBNext´s Small RNA Library Prep Kit for Illumina following instructions of manual, with a modification of the adaptors, for which custom made random barcodes for both 3' SR Adaptor and 5' SR Adaptor were used (HISS Diagnostics GmbH, 5' rApprnrnrnrnAGATCGGAAGAGCACACGTCT-NH2- 3', and 5' rGrUrUrCrArGrArGrUrUrCrUrArCrArGrUrCrCrGrArCrGrArUrC rnrnrnrn- 3', respectively). Libraries were amplified in 14 PCR cycles. Libraries were size selected for the 135-170 bp fraction on a DNA 300 LabChip with a LabChip XT (Caliper). Libraries were profiled in a High Sensitivity DNA on a 2100 Bioanalyzer (Agilent technologies) and quantified using the Qubit dsDNA HS Assay Kit, in a Qubit 2.0 Fluorometer (Life technologies). All 12 samples were pooled in equimolar ratio and sequenced on 2 HiSeq 2500 rapid lanes, single read 51 bp (1x 51 cycles for read 1 plus 7 cycles for the index read).

Read procession and mapping

Before mapping to the genome, reads were processed in the following manner: (i) trimming of sequencing adapters with cutadapt v1.9 (-a TGGAATTCTCGGGTGCCAAGG -O 5 -m 26 -M 38) (Martin, 2011); (ii) removal of reads with low-quality calls with the FASTX-Toolkit v0.0.14 (fastq_quality_filter -q 20 -p 100 -Q 33); (iii) collapsing of PCR duplicates (custom bash script), making use of the unique molecule identifiers (UMIs) added during library preparation; (iv) trimming of UMIs with seqtk v1.2 (trimfq -b 4 -e 4); and (v) removal of very short sequences with seqtk v1.2 (seq -L 15). Read quality was assessed before and after these processing steps with FastQC v0.11.5 (http://www.bioinformatics. babraham.ac.uk/projects/fastqc).

Reads that passed the above filtering steps were mapped to a custom *C. elegans* genome (WBcel235) to which the 21U sensor sequence (Bagijn et al., 2012) was added as an extra contig. The mapping was done with bowtie v0.12.8 (-q –sam –phred33-quals –tryhard –best –strata –chunkmbs 256 -v 0 -M 1) (Langmead et al., 2009). To generate genome browser tracks we used a combination of Bedtools v2.25.0

(genomeCoverageBed -bg -split -scale -ibam -g) (Quinlan & Hall, 2010), to summarize genome coverage normalized to mapped non-structural reads (rRNA/tRNA/snoRNA/snRNA) * 1 million (RPM, Reads Per Million), and bedGraphToBigWig to finally create the bigwig tracks.

Small RNA classification and quantification

Gene annotation was retrieved from Ensembl (release-38) and merged with transposon coordinates retrieved from Wormbase (PRJNA13758.WS264), creating a custom annotation used for the analysis. Mapped reads were categorized in small RNA classes as follows: 21U RNAs are 21 nt long sequences mapping sense to annotated 21U RNA loci; 22G RNAs are 20-23 nt long and map antisense to protein-coding/pseudogenes/lincRNA/transposons; 26G RNAs, are those which are 26 nt long, and map antisense to annotated protein-coding/pseudogenes/lincRNA; and miRNAs are 20-24 nt long mapping sense to annotated miRNA loci. Read filtering was done with a python script (https://github.com/adomingues/filterReads/blob/master/filterReads/filterSmallRNAclasses.py) based on pysam v0.8.1 / htslib (Li et al., 2009), in combination with Bedtools intersect. Reads belonging to each class were then counted for each library (total levels).

22G RNAs coverage on 21U sensor

For targeting of the 21U sensor by 22G RNAs, we considered only sequences that were 22 nt long mapping unambiguously to the 21U sensor sequence. Coverage was calculated with Bedtools v2.25.0 (genomeCoverageBed -ibam - -d) (Quinlan & Hall, 2010). Visualization was generated with the R/Bioconductor package ggbio (Lawrence & Morgan, 2014) [Yinggbiopackageextending2012].

Sequencing statistics

Strain	Sample ID	Sequenced Reads	Mapped Reads	Non-structural Reads
RFK767	01_xf35_new_sensor_off_adult_r1	24757402	9287694	2100603
RFK768	02_xf35_new_sensor_off_adult_r2	24195018	4808964	1360405
RFK769	03_xf35_new_sensor_off_adult_r3	25029484	10305409	2654722
RFK764	04_xf35_new_sensor_on_adult_r1	24587675	6543218	2074545
RFK765	05_xf35_new_sensor_on_adult_r2	23699483	9438805	3121632
RFK766	06_xf35_new_sensor_on_adult_r3	24223060	9666942	2873549
RFK422	07_xf35_old_sensor_off_adult_r1	21521030	8304205	3046244
RFK422	08_xf35_old_sensor_off_adult_r2	21125672	9431545	3375720
RFK422	09_xf35_old_sensor_off_adult_r3	25486251	9586297	2582408

RFK184	10_xf35_old_sensor_on_adult_r1	25551143	10804478	3915949
RFK184	11_xf35_old_sensor_on_adult_r2	27319695	9102517	2567677
RFK184	12_xf35_old_sensor_on_adult_r3	22939937	6309871	1439481

mRNA sequencing

RNA extraction

Synchronized gravid adults have been collected with M9 and fast frozen on dry ice in 250 µl of Worm Lysis Buffer (200 mM NaCl; 100 mM Tris HCl pH=8.5; 50 mM EDTA pH=8; 0,5% SDS). After thawing worms, samples have been spun down at 1100 x g for 2 minutes in order to remove the supernatant and proceed to RNA extraction with a small pellet (~50 µl). 500 µl of Trizol LS (Art. No. 10296028, Invitrogen™) have been added to each sample and, after mixing, worms have been fast frozen for 30 seconds in liquid nitrogen and then thawn at 37 °C for 2 minutes. Six cycles of fast freezing and thawing have been performed to dissolve the worms, until no intact worms are visible under the microscope. Samples have been then spun down at maximum speed for 10 minutes to eliminate debris and the supernatant has been transferred to a new tube. 100% ethanol has been added to the samples (1:1) and RNA has been extracted with the Direct-zol™ RNA Kit (Art. No. R2060, Zymo Research), according to the manufacturer instructions. RNA has been then dissolved in 15 µl of nuclease-free water.

Library preparation and sequencing

NGS library preparation was performed with Illumina's TruSeq stranded mRNA LT Sample Prep Kit following Illumina's standard protocol (Part # 15031047 Rev, E), which includes a poly-A selection step. Libraries were prepared with a starting amount of 1 µg and amplified in 10 PCR cycles. Libraries were profiled in a DNA 1000 Chip on a 2100 Bioanalyzer (Agilent technologies) and quantified using the Qubit dsDNA HS Assay Kit, in a Qubit 2,0 Fluorometer (Life technologies),

All 21 samples were pooled in equimolar ratio and sequenced on 1 NextSeq 500 High Output FC, single read 85bp (1x 85 cycles for read 1 plus 7 cycles for the index read).

mRNA read processing and mapping

Prior to mapping the raw read quality was assessed with FastQC. Alignment to the *C. elegans* genome / transcriptome, assembly WBcel235, and the custom GTF described above including transposons and sensor sequence, was performed with STAR v2.5.2b

(–runMode alignReads –outSAMattributes Standard –outSJfilterReads Unique –outSAMunmapped Within –outReadsUnmapped None –outFilterMismatchNmax 2 –outFilterMultimapNmax 10 –alignIntronMin 21 –sjdbOverhang 83). Reads mapping to annotated features in the custom GTF were counted with subread featureCounts v1.5.1 (-s 2 -p -F GTF –donotsort -t exon -g gene_id). Coverage tracks were generated with deepTools v2.4.3 (bamCoverage –smoothLength 60 –binSize 20 –normalizeUsingRPKM) (Ramírez et al., 2016).

Differential gene expression

Reads mapping to annotated features in the custom GTF were counted with subread featureCounts v1.5.1 (Liao et al., 2014) (-s 2 -p -F GTF –donotsort -t exon -g gene_id). Differential expression comparisons were performed with DESeq2 v.1.18.1 (Love et al., 2014). For the selection of genes differentially expressed (mRNA), a cut-off of at least a 2 fold-change difference between conditions and an adjusted p-value (FDR) less than 0.1 were applied. rlog normalized counts were used as an input for hierarchical clustering and PCA analysis.

Sequencing statistics

Strain	Sample ID	Sequenced Reads	Mapped Reads
N2	01_N2_WT_nosensor_rep1_polyA	17272162	17160438
N2	02_N2_WT_nosensor_rep2_polyA	18453192	18348860
N2	03_N2_WT_nosensor_rep3_polyA	16950026	16854983
RFK764	04_RFK764_xf35_sensorON_rep1_polyA	19070305	18966055
RFK764	05_RFK764_xf35_sensorON_rep2_polyA	16568899	16471088
RFK764	06_RFK764_xf35_sensorON_rep3_polyA	18175628	18068953
RFK765	07_RFK765_xf35_sensorON_rep1_polyA	18988961	18890291
RFK765	08_RFK765_xf35_sensorON_rep2_polyA	18877366	18778365
RFK765	09_RFK765_xf35_sensorON_rep3_polyA	18950543	18851280
RFK766	10_RFK766_xf35_sensorON_rep1_polyA	19678706	19559535
RFK766	11_RFK766_xf35_sensorON_rep2_polyA	18365198	18273009
RFK766	12_RFK766_xf35_sensorON_rep3_polyA	17799968	17687948
RFK767	13_RFK767_xf35_sensorOFF_rep1_polyA	19229237	19108647
RFK767	14_RFK767_xf35_sensorOFF_rep2_polyA	19336875	19203577
RFK767	15_RFK767_xf35_sensorOFF_rep3_polyA	20318451	20198687
RFK768	16_RFK768_xf35_sensorOFF_rep1_polyA	19441951	19344330
RFK768	17_RFK768_xf35_sensorOFF_rep2_polyA	18205421	18101519
RFK768	18_RFK768_xf35_sensorOFF_rep3_polyA	18430369	18327772
RFK769	19_RFK769_xf35_sensorOFF_rep1_polyA	18257650	18154959
RFK769	20_RFK769_xf35_sensorOFF_rep2_polyA	18158595	18052803
RFK769	21_RFK769_xf35_sensorOFF_rep3_polyA	17764280	17662565

References

Almeida, M. V., de Jesus Domingues, A. M., Lukas, H., Mendez-Lago, M., & Ketting, R. F. (2019). RppH can faithfully replace TAP to allow cloning of 5'-triphosphate carrying small RNAs. *MethodsX*, *6*, 265–272. https://doi.org/10.1016/j.mex.2019.01.009

Ashe, A., Sapetschnig, A., Weick, E. M., Mitchell, J., Bagijn, M. P., Cording, A. C., … Miska, E. A. (2012). PiRNAs can trigger a multigenerational epigenetic memory in the germline of C. elegans. *Cell*. https://doi.org/10.1016/j.cell.2012.06.018

Bagijn, M. P., Goldstein, L. D., Sapetschnig, A., Weick, E.-M., Bouasker, S., Lehrbach, N. J., … Miska, E. A. (2012). Function, Targets, and Evolution of Caenorhabditis elegans piRNAs. *Science*.

Batista, P. J., Ruby, J. G., Claycomb, J. M., Chiang, R., Fahlgren, N., Kasschau, K. D., … Mello, C. C. (2008). PRG-1 and 21U-RNAs Interact to Form the piRNA Complex Required for Fertility in C. elegans. *Molecular Cell*. https://doi.org/10.1016/j.molcel.2008.06.002

Billi, A. C., Alessi, A. F., Khivansara, V., Han, T., Freeberg, M., Mitani, S., & Kim, J. K. (2012). The Caenorhabditis elegans HEN1 ortholog, HENN-1, methylates and stabilizes select subclasses of germline small RNAs. *PLoS Genetics*, *8*(4), e1002617. https://doi.org/10.1371/journal.pgen.1002617

Brenner, S. (1974). THE GENETICS OF CAENORHABDITIS ELEGANS. *Genetics*, *77*(1).

Buckley, B. A., Burkhart, K. B., Gu, S. G., Spracklin, G., Kershner, A., Fritz, H., … Kennedy, S. (2012). A nuclear Argonaute promotes multigenerational epigenetic inheritance and germline immortality. *Nature*, *489*(7416), 447–451. https://doi.org/10.1038/nature11352

Claycomb, J. M., Batista, P. J., Pang, K. M., Gu, W., Vasale, J. J., van Wolfswinkel, J. C., … Mello, C. C. (2009). The Argonaute CSR-1 and Its 22G-RNA Cofactors Are Required for Holocentric Chromosome Segregation. *Cell*. https://doi.org/10.1016/j.cell.2009.09.014

Conine, C. C., Moresco, J. J., Gu, W., Shirayama, M., Conte, D., Yates, J. R., & Mello, C. C. (2013). Argonautes promote male fertility and provide a paternal memory of germline gene expression in C. Elegans. *Cell*. https://doi.org/10.1016/j.cell.2013.11.032

Cox, D. N., Chao, A., Baker, J., Chang, L., Qiao, D., & Lin, H. (1998). A novel class of evolutionarily conserved genes defined by piwi are essential for stem cell self-renewal. *Genes and Development*. https://doi.org/10.1101/gad.12.23.3715

Das, P. P., Bagijn, M. P., Goldstein, L. D., Woolford, J. R., Lehrbach, N. J., Sapetschnig, A., … Miska, E. A. (2008). Piwi and piRNAs Act Upstream of an Endogenous siRNA Pathway to Suppress Tc3 Transposon Mobility in the Caenorhabditis elegans Germline. *Molecular Cell*. https://doi.org/10.1016/j.molcel.2008.06.003

de Albuquerque, B. F. M., Luteijn, M. J., Cordeiro Rodrigues, R. J., van Bergeijk, P., Waaijers, S., Kaaij, L. J. T., … Ketting, R. F. (2014). PID-1 is a novel factor that operates during 21U-RNA biogenesis in Caenorhabditis elegans. *Genes and Development*. https://doi.org/10.1101/gad.238220.114

de Albuquerque, B. F. M., Placentino, M., & Ketting, R. F. (2015). Maternal piRNAs Are Essential for Germline Development following De Novo Establishment of Endo-siRNAs in Caenorhabditis elegans. *Developmental Cell*. https://doi.org/10.1016/j.devcel.2015.07.010

Ghildiyal, M., & Zamore, P. D. (2009). Small silencing RNAs: An expanding universe. *Nature Reviews Genetics*. https://doi.org/10.1038/nrg2504

Gu, W., Lee, H.-C., Chaves, D., Youngman, E. M., Pazour, G. J., Conte, D., & Mello, C. C. (2012). CapSeq and CIP-TAP identify Pol II start sites and reveal capped small RNAs as C. elegans piRNA precursors. *Cell*, *151*(7), 1488–1500. https://doi.org/10.1016/j.cell.2012.11.023

Kamminga, L. M., van Wolfswinkel, J. C., Luteijn, M. J., Kaaij, L. J. T., Bagijn, M. P., Sapetschnig, A., … Ketting, R. F. (2012). Differential Impact of the HEN1 Homolog HENN-1 on 21U and 26G RNAs in the Germline of Caenorhabditis elegans.

PLoS Genetics, *8*(7), e1002702. https://doi.org/10.1371/journal.pgen.1002702

Ketting, R. F. (2011). The Many Faces of RNAi. *Developmental Cell*. https://doi.org/10.1016/j.devcel.2011.01.012

Ketting, R. F., Haverkamp, T. H. A., Van Luenen, H. G. A. M., & Plasterk, R. H. A. (1999). mut-7 of C. elegans, required for transposon silencing and RNA interference, is a homolog of werner syndrome helicase and RNaseD. *Cell*. https://doi.org/10.1016/S0092-8674(00)81645-1

Langmead, B., Trapnell, C., Pop, M., & Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology*, *10*(3), R25. https://doi.org/10.1186/gb-2009-10-3-r25

Lawrence, M., & Morgan, M. (2014). Scalable Genomics with R and Bioconductor. *Statistical Science : A Review Journal of the Institute of Mathematical Statistics*, *29*(2), 214–226. https://doi.org/10.1214/14-STS476

Lee, H. C., Gu, W., Shirayama, M., Youngman, E., Conte, D., & Mello, C. C. (2012). C. elegans piRNAs mediate the genome-wide surveillance of germline transcripts. *Cell*. https://doi.org/10.1016/j.cell.2012.06.016

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., … 1000 Genome Project Data Processing Subgroup. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Liao, Y., Smyth, G. K., & Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, *30*(7), 923–930. https://doi.org/10.1093/bioinformatics/btt656

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*(12), 550. https://doi.org/10.1186/s13059-014-0550-8

Luteijn, M. J., Van Bergeijk, P., Kaaij, L. J. T., Almeida, M. V., Roovers, E. F., Berezikov, E., & Ketting, R. F. (2012). Extremely stable Piwi-induced gene silencing in Caenorhabditis elegans. *EMBO Journal*. https://doi.org/10.1038/emboj.2012.213

Malone, C. D., & Hannon, G. J. (2009). Small RNAs as Guardians of the Genome. *Cell*. https://doi.org/10.1016/j.cell.2009.01.045

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.Journal*, *17*(1), 10. https://doi.org/10.14806/ej.17.1.200

Montgomery, T. A., Rim, Y.-S., Zhang, C., Dowen, R. H., Phillips, C. M., Fischer, S. E. J., & Ruvkun, G. (2012). PIWI associated siRNAs and piRNAs specifically require the Caenorhabditis elegans HEN1 ortholog henn-1. *PLoS Genetics*, *8*(4), e1002616. https://doi.org/10.1371/journal.pgen.1002616

Phillips, C. M., Brown, K. C., Montgomery, B. E., Ruvkun, G., & Montgomery, T. A. (2015). PiRNAs and piRNA-Dependent siRNAs Protect Conserved and Essential C. elegans Genes from Misrouting into the RNAi Pathway. *Developmental Cell*. https://doi.org/10.1016/j.devcel.2015.07.009

Phillips, C. M., Montgomery, B. E., Breen, P. C., Roovers, E. F., Rim, Y. S., Ohsumi, T. K., … Montgomery, T. A. (2014). MUT-14 and SMUT-1 DEAD box RNA helicases have overlapping roles in germline RNAi and endogenous siRNA formation. *Current Biology*. https://doi.org/10.1016/j.cub.2014.02.060

Phillips, C. M., Montgomery, T. A., Breen, P. C., & Ruvkun, G. (2012). MUT-16 promotes formation of perinuclear Mutator foci required for RNA silencing in the C. elegans germline. *Genes and Development*. https://doi.org/10.1101/gad.193904.112

Quinlan, A. R., & Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics (Oxford, England)*, *26*(6), 841–842. https://doi.org/10.1093/bioinformatics/btq033

Ramírez, F., Ryan, D. P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A. S., … Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Research*, *44*(W1), W160–W165. https://doi.org/10.1093/nar/gkw257

Sapetschnig, A., Sarkies, P., Lehrbach, N. J., & Miska, E. A. (2015). Tertiary siRNAs Mediate Paramutation in C. elegans. *PLOS Genetics*, *11*(3), e1005078. https://doi.org/10.1371/journal.pgen.1005078

Seth, M., Shirayama, M., Gu, W., Ishidate, T., Conte, D., & Mello, C. (2013). The C. elegans CSR-1 argonaute pathway counteracts epigenetic silencing to promote germline gene expression. *Developmental Cell*. https://doi.org/10.1016/j.devcel.2013.11.014

Shen, E.-Z., Chen, H., Ozturk, A. R., Tu, S., Shirayama, M., Tang, W., … Mello, C. C. (2018). Identification of piRNA Binding Sites Reveals the Argonaute Regulatory Landscape of the C. elegans Germline. *Cell*, *172*(5), 937–951.e18. https://doi.org/10.1016/j.cell.2018.02.002

Shirayama, M., Seth, M., Lee, H. C., Gu, W., Ishidate, T., Conte, D., & Mello, C. C. (2012). PiRNAs initiate an epigenetic memory of nonself RNA in the C. elegans germline. *Cell*. https://doi.org/10.1016/j.cell.2012.06.015

Simon, M., Sarkies, P., Ikegami, K., Doebley, A.-L., Goldstein, L. D., Mitchell, J., … Ahmed, S. (2014). Reduced insulin/IGF-1 signaling restores germ cell immortality to Caenorhabditis elegans Piwi mutants. *Cell Reports*, *7*(3), 762–773. https://doi.org/10.1016/j.celrep.2014.03.056

Tabara, H., Sarkissian, M., Kelly, W. G., Fleenor, J., Grishok, A., Timmons, L., … Mello, C. C. (1999). The rde-1 gene, RNA interference, and transposon silencing in C. elegans. *Cell*. https://doi.org/10.1016/S0092-8674(00)81644-X

van Wolfswinkel, J. C., Claycomb, J. M., Batista, P. J., Mello, C. C., Berezikov, E., & Ketting, R. F. (2009). CDE-1 Affects Chromosome Segregation through Uridylation of CSR-1-Bound siRNAs. *Cell*, *139*(1), 135–148. https://doi.org/10.1016/j.cell.2009.09.012

Wang, G., & Reinke, V. (2008). A C. elegans Piwi, PRG-1, Regulates 21U-RNAs during Spermatogenesis. *Current Biology*. https://doi.org/10.1016/j.cub.2008.05.009

Wedeles, C. J., Wu, M. Z., & Claycomb, J. M. (2013). Protection of germline gene expression by the C. elegans argonaute CSR-1. *Developmental Cell*. https://doi.org/10.1016/j.devcel.2013.11.016

Zhang, C., Montgomery, T. A., Gabel, H. W., Fischer, S. E. J., Phillips, C. M., Fahlgren, N., … Ruvkun, G. (2011). mut-16 and other mutator class genes modulate 22G and 26G siRNA pathways in Caenorhabditis elegans. *Proceedings of the National Academy of Sciences*. https://doi.org/10.1073/pnas.1018695108

Chapter 4

PID-2 forms a complex with the two Tudor domain proteins PID-4 and PID-5, to enhance PRG-1 mediated silencing

Maria Placentino, António Miguel de Jesus Domingues, Jan Schreier, Svenja Hellmann, Bruno F. M. de Albuquerque, Sabrina Dietz, Falk Butter, René F. Ketting

Summary

The Piwi pathway is a germline specific silencing mechanism that serves to control transposable elements, in order to preserve genome integrity of germ cells. In *C. elegans*, the Piwi protein PRG-1 forms a complex with its small RNA cofactor (21U RNA) to recognise and silence a target mRNA, which is specified by the 21U RNA sequence. After recognition by PRG-1, the target mRNA is used by an RNA-dependent RNA polymerase as a template to produce secondary antisense 22G RNAs. These 22G RNAs are loaded onto secondary Argonaute proteins, such as HRDE-1, to amplify the silencing reaction. HRDE-1 mediates a form of silencing that can become self-sustainable, independent of PRG-1 and it is characterized by heterochromatic marks at the targeted locus, a status known as RNAe. RNAe establishes rather stochastically, and mechanisms behind this choice are completely unknown. We identify a novel factor, PID-2, which is required to establish target silencing and RNAe, but is not essential for its maintenance. PID-2 acts downstream of PRG-1 and has a role in Tc1 silencing. PID-2 localizes to perinuclear granules and interacts with two, partially redundant, Tudor proteins, PID-4 and PID-5, to form a complex that helps to initiate target silencing as well as RNAe. Like *prg-1* mutants, *pid-2* and *pid-4/-5* mutants display a mortal germline phenotype. These studies reveal the first proteins that genetically act between PRG-1 mediated target recognition and the 22G RNA amplification response, and open the door to mechanistic studies on how RNAe can be established.

Introduction

Germ cells are responsible for transmitting genetic information to the next generation, therefore it is extremely important for the organism to preserve their genome. Genome integrity is constantly threatened by external factors, such as stress, mutagenic agents, and by transposable elements. One of the mechanisms that organisms have evolved to ensure genome integrity of germ cells is the Piwi pathway.

Piwi proteins are a germline specific subclade of Argonaute proteins that, upon loading with their small RNA cofactor, that is a piRNA (Piwi-interacting RNA), form a complex to silence transposons. As a consequence, the Piwi/piRNA pathway is essential in most organisms not only for transposon silencing, but also for germ cell development and fertility (Ghildiyal & Zamore, 2009; Ketting, 2011; Malone & Hannon, 2009). In contrast to other organisms, in *Caenorhabditis elegans* the Piwi pathway is only partially needed for transposon silencing (Das et al., 2008) and is not required for integrity of the germline (Batista et al., 2008; Cox et al., 1998; Das et al., 2008; Wang & Reinke, 2008), even though the germline is progressively lost over generations (mortal germline phenotype, Mrt) (Simon et al., 2014).

The main Piwi protein of *C. elegans* is PRG-1 (Piwi Related Gene-1); it binds to piRNAs (21U RNAs), that then guide the complex to endogenous targets via base-pairing, allowing up to 4 mismatches (Bagijn et al., 2012; Lee et al., 2012). As a consequence of mismatch tolerance, the PRG-1/21U RNA complexes are potentially able to recognize and silence every exogenous DNA sequence (Gu et al., 2012), and also have been shown to silence endogenous genes through such incomplete base-pairing between 21U RNA and mRNA (Bagijn et al., 2012).

After target recognition, an RNA-dependent RNA polymerase (RdRP) is recruited to the target mRNA to produce antisense small RNAs (22G RNAs) in order to amplify the silencing reaction. The synthesis of these secondary 22G RNAs depends not only on RdRP, but also on mutator proteins (Phillips et al., 2012; Zhang et al., 2011) and it takes place in small granules, called mutator foci, that reside adjacent to P granules and nuclear pores on the outer part of germ cells nuclei (Phillips et al., 2014). These 22G RNAs are then loaded onto secondary Argonaute proteins, such as the nuclear protein HRDE-1 (Heritable RNAi Deficient-1), to execute post-transcriptional target silencing (Ashe et al., 2012; Buckley et al., 2012; Shirayama et al., 2012). In some animals, in an apparently stochastic manner, the silencing that is thus established can become

independent of PRG-1 itself and self-sustainable. This kind of silencing is extremely stable and can be transmitted for tens of generations in absence of PRG-1; it depends on mutator activity, HRDE-1 and is accompanied by heterochromatin formation at targeted loci, and it is called RNAe (RNA-induced epigenetic silencing) (Ashe et al., 2012; Luteijn et al., 2012; Shirayama et al., 2012).

Despite all the current knowledge, 21U RNA biogenesis as well as RNAe establishment are still not completely characterized. Therefore, we performed a forward mutagenesis screen to identify novel factors involved in these pathways (de Albuquerque et al., 2014). Amongst other factors that we have identified, we describe here the characterization of a novel protein, that we named PID-2. PID-2 is necessary to establish RNAe, but not for its maintenance. Our data show that PID-2 acts downstream of PRG-1, to stimulate secondary 22G RNA production, and to ensure long-term fertility. Furthermore, we show that PID-2 interacts with two Tudor domain proteins, PID-4 and PID-5. Absence of PID-4 and PID-5 together phenocopies the *pid-2* mutant, both at the level of 21U RNA activity and fertility. We propose that PID-2, PID-4 and PID-5 function together as a platform to bring together factors that are required to establish 21U RNA-mediated gene silencing and to maintain germline health across generations.

Results

Identification of *pid-2*

To identify novel factors of the 21U RNA silencing machinery, we previously performed a forward mutagenesis screen in which we used EMS (Ethyl methanesulfonate) to mutagenize a strain carrying a 21U sensor (de Albuquerque et al., 2014). The 21U sensor is a single copy transgene, specifically expressed in the germline, that encodes a GFP::H2B fusion protein; its 3' UTR contains a stretch of nucleotides perfectly complementary to one of the most abundant 21U RNAs of *C. elegans*, *21ur-1* (Bagijn et al., 2012). As a consequence, the 21U sensor can be used as a reporter of the integrity of the 21U RNA pathway. If the pathway is functional, PRG-1 loaded with the 21U RNA complementary to the target sequence will recognize the 21U sensor and silence it. On the other hand, if the pathway is not functioning properly, PRG-1 will not be able to recognize its target and this will result in expression of the GFP::H2B in the germline of the worm.

We performed the screen in a *henn-1* mutant background, in which the 21U sensor is partially reactivated (Kamminga et al., 2012). This allowed us to mutagenize animals in which RNAe had not been established yet on the 21U sensor, something that is not possible in a wild type background. Therefore, we were able to recover mutants that are defective in the establishment of RNAe, but still proficient in its maintenance.

From this screen we have isolated several mutants that show defects in 21U RNA-induced silencing (piRNA induced silencing defective: *pid* mutants). Some of these mutants carry novel alleles of factors that were already known to be involved in the 21U RNA silencing and/or RNAe pathways (e.g. *mut-7*, *hrde-1*, *rde-3*), and we also identified a novel factor involved in 21U RNA biogenesis, PID-1 (de Albuquerque et al., 2014). In this chapter, we will describe the characterization of a gene identified by another allele isolated from this screen, *xf23*.

Through genome sequencing of the strain carrying the *xf23* allele, and conservation analysis of mutated genes, we identified a point mutation (tgg → tga) that introduces a premature STOP codon (W122X) in the gene Y48G1C.1. Y48G1C.1 encodes a protein of 454 amino acids with no known domains or functions. It is enriched in primordial germ cells (PGCs) and conserved throughout the *Caenorhabditis* genera, which are common features of other factors involved in the 21U RNA pathway, making this mutation a good candidate for causing the phenotype of *xf23* ("Genetics and dynamics of piRNA induced silencing", PhD thesis, de Albuquerque, 2015).

In order to confirm that the causal mutation of *xf23* is within the Y48G1C.1 gene, we tested whether a different mutant allele of the same gene shows the same phenotype. Therefore, we used a publicly available deletion allele of Y48G1C.1, *tm1614*, which removes 451 base pairs of the coding sequence, starting from its 5' UTR through almost the end of the second exon (obtained from the National Bioresource Project, Tokyo, Japan, which is part of the International *C. elegans* Gene Knockout Consortium) (*C. elegans* Deletion Mutant Consortium, 2012). Both mutant strains show low expression of the 21U sensor and have similar phenotypes, as shown in **Figure 2**.

Finally, we generated transgenic lines expressing a tagged version of Y48G1C.1. We utilized the MosSCI system (Frøkjær-Jensen et al., 2008) to generate transgenic lines, using the endogenous promoter and 3' UTR of Y48G1C.1. This system allows for transgene insertion in a defined genomic location, that is stable and ensures transgene expression at levels comparable to the endogenous gene (Frøkjær-Jensen et al., 2008).

We have thus generated transgenes encoding 3xFLAG::Y48G1C.1 (*xfIs146*), and eGFP tagged versions of Y48G1C.1, both N-terminal (*xfIs144*) and C-terminal (*xfIs145*).

To check if these transgenes can rescue *xf23*, we introduced this mutation into the transgenic lines and then crossed these hermaphrodites with males carrying an active 21U sensor in a *xf23* mutant background (**Figure 1A**). If the transgenes are functional, they should rescue the absence of the endogenous Y48G1C.1 protein and silence the 21U sensor. We observed a full rescue for the 3xFLAG::Y48G1C.1 (*xfIs146*) transgene (**Figure 1B**, **C**). On the other hand, N-terminally eGFP tagged Y48G1C.1 did not rescue the 21U sensor silencing (**Figure 1B**, **C**). Also, we were not able to visualize the N-terminally tagged Y48G1C.1 (*xfIs144*) at the confocal microscope nor to detect it in the immunoprecipitation followed by mass spectrometry (IP/MS) analysis (not shown). When the eGFP was positioned at the C-terminal (*xfIs145*), the transgene was partially able to rescue (**Figure 1B**, **C**). Likely the eGFP interferes with Y48G1C.1 folding and/or functionality, and this effect is more prominent when the fluorescent tag is positioned at the N-terminal end, while this effect is weaker when present at the C-terminal end. Everything considered, we conclude that *xf23*, like *tm1614*, is a mutant allele of Y48G1C.1 and we renamed it *pid-2*.

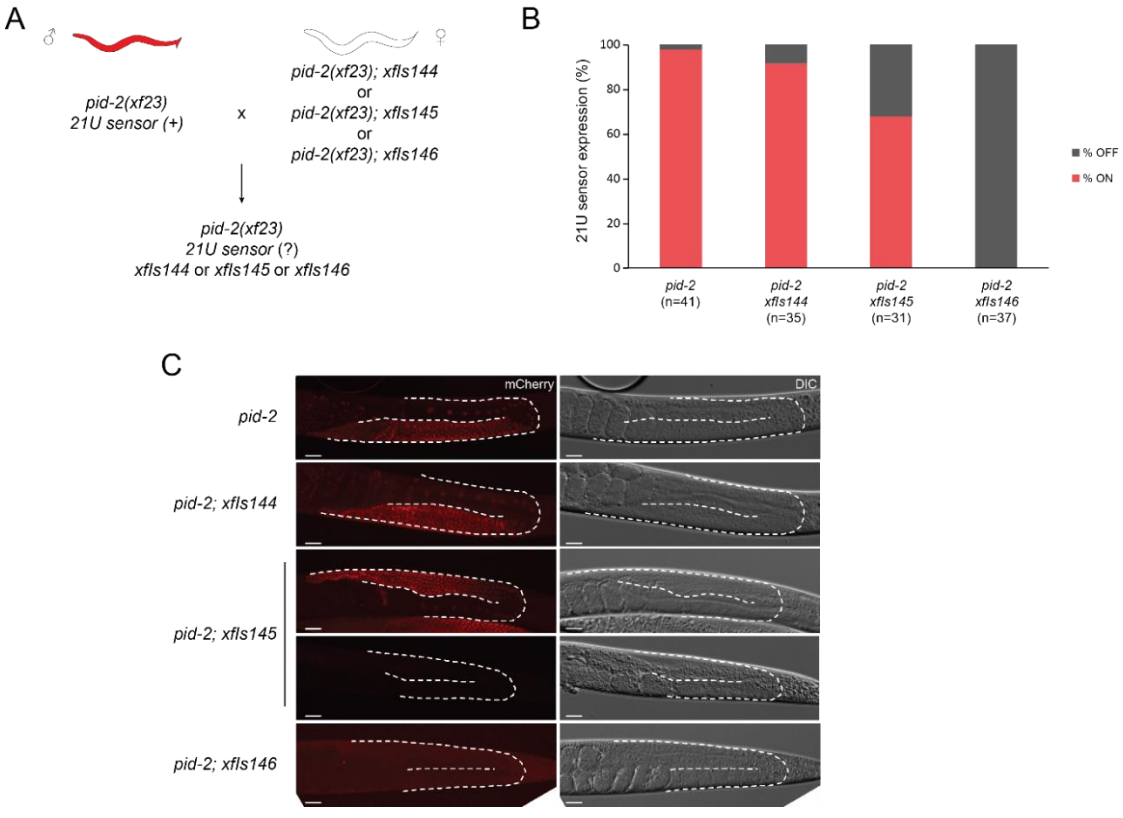


Figure 1. A) Crossing strategy to test the rescuing potential of the Y48G1C.1/PID-2 transgenes. **B)** The expression of the 21U sensor has been scored via microscopy and quantified. The rescue by *xfIs146* is complete, by *xfIs145* is partial (~40%), however *xfIs144* does not rescue. **C)** Representative images of the 21U sensor expression quantified in **B**. Genotypes are indicated on the left side. Left panels: mCherry signal; right panels: DIC of the same gonad arms. Gonads are outlined by a dashed line. Scale bar = 25 µm.

pid-2 mutants are not able to silence a 21U RNA target *de novo*, but can maintain RNAe

We then investigated in more detail how 21U RNA target silencing is affected upon *pid-2* mutation using a 21U sensor. In this experiment, we have used another variant of the 21U sensor, in which mCherry is used instead of GFP (**Figure 2A**) (Bagijn et al., 2012). When we introduce a non-silenced version of this 21U sensor into a *pid-2* mutant background, the sensor is maintained active, although its expression is visibly lower compared to its expression level in a *mut-7* background (**Figure 2B**). This indicates that without PID-2, full silencing of this sensor cannot be achieved. Nevertheless, it is also clear that in absence of PID-2, a significant degree of silencing can still be established. In contrast, when we use a 21U sensor that is silenced already through RNAe, the absence of PID-2 has no effect (**Figure 2B**), meaning that the transgene remains fully silenced.

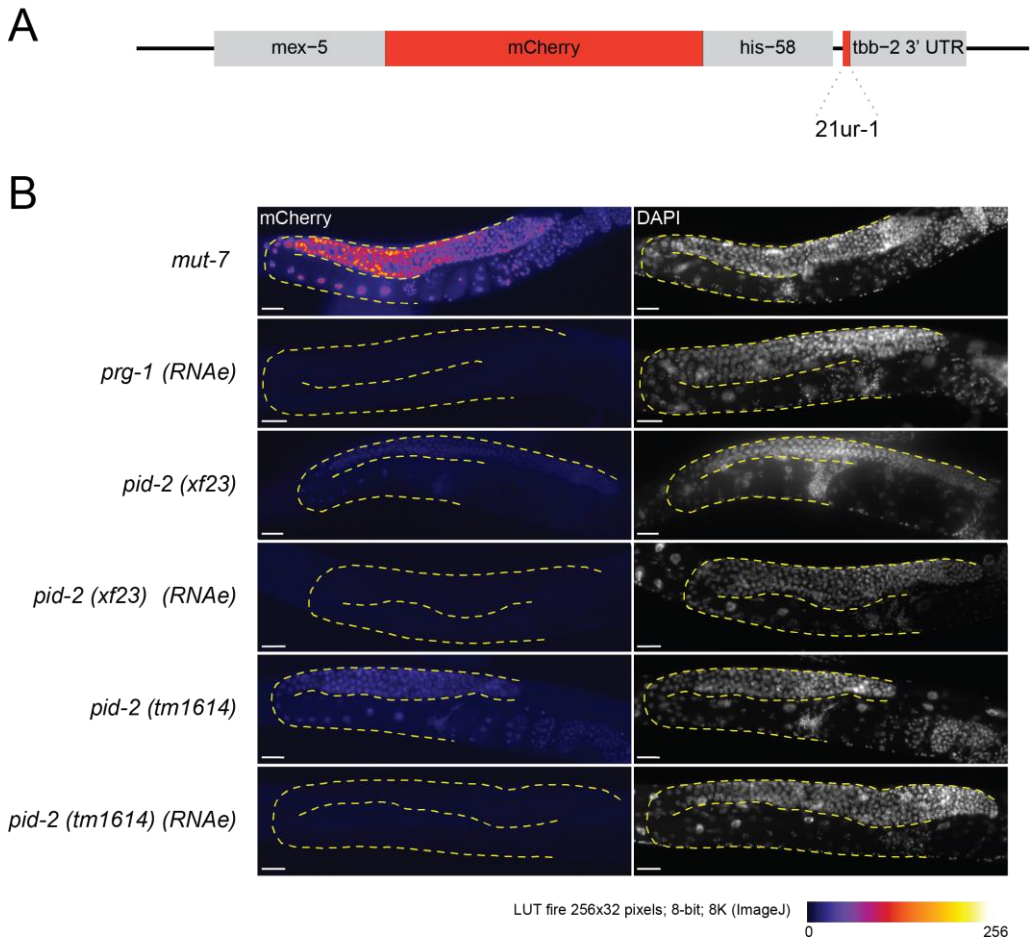


Figure 2. A) Schematic representation of the 21U sensor transgene. **B)** Expression of the 21U sensor in the indicated genetic backgrounds: the absence of PID-2 does not reactivate the 21U sensor, that has been previously silenced through RNAe. If the sensor has not been previously silenced though, the absence of PID-2 does not allow for establishing of the silencing, although the expression of the 21U sensor is lower compared to e.g. *mut-7* mutants, that lack 22G RNAs as well as RNAe. Scale bar = 25 μm. The mCherry signal is represented in pseudo-colours [LUT fire (ImageJ)] to reflect differences in the intensity of the signal.

These data show that PID-2 is not essential for RNAe maintenance, but that it is required to establish full *de novo* silencing by 21U RNAs.

PID-2 is required to establish *de novo* silencing mediated by maternally inherited 21U RNAs and for germline development

We have previously shown that maternal 21U RNAs are necessary to re-silence the 21U sensor in the next generation (see chapter 3) (de Albuquerque et al., 2014). As described in chapter 3, if we cross *mut-7* mutant males, expressing the 21U sensor, with either wild type or *prg-1* mutant hermaphrodites, the offspring will silence or fail to silence

the 21U sensor, respectively, reflecting the maternal contribution of 21U RNAs (**Figure 3A**, **B**, **E**, **F**: these are the same panels of **Figure 1**, chapter 3). We also tested what happens to the silencing of the 21U sensor if we use *pid-2* mutant hermaphrodites. In this situation, the offspring show defects in initiating the silencing of the 21U sensor to some extent, although a fraction (42-45%) of the offspring is still able to silence it (**Figure 3C**, **D**, **G**, **H**).

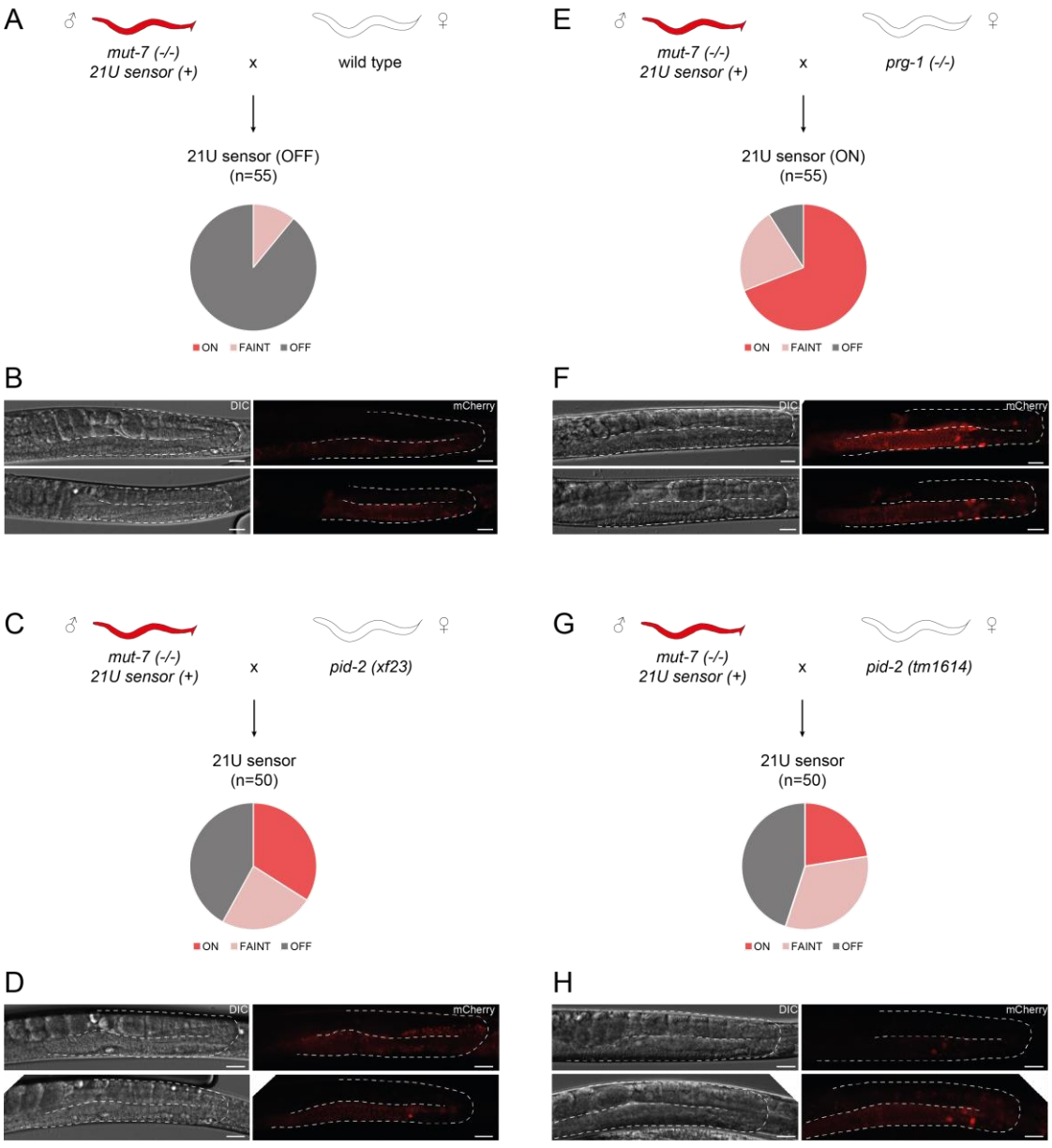


Figure 3. **A**, **C**, **E**, **G)** Scheme representing the cross strategy to address the re-silencing of the 21U sensor. A *mut-7* mutant male expressing the 21U sensor, because of depletion of 22G RNAs, is crossed either with a wild type (**A**), *prg-1* (**E**) or *pid-2* (**C**, **G**) mutant hermaphrodite. Their offspring are able to establish *de novo* silencing on the 21U sensor, if they inherit a pool of 21U RNAs from the hermaphrodite, in 90% of the worms (see pie chart: **A**). If the hermaphrodite cannot transmit any 21U RNAs to the next generation, the offspring is not able to silence the 21U sensor (90%; see pie chart; **E**).

The offspring of *pid-2* mutant hermaphrodites are partially able to silence the 21U sensor, because 21U RNAs are deposited from the maternal lineage. However, *pid-2* mutation does not allow the silencing of the 21U sensor in the whole population (see pie chart; **C**, **G**). **B**, **D**, **E**, **H)** Representative images of the 21U sensor that is either silenced, faintly expressed or expressed (left, DIC pictures of one gonad; right: mCherry signal). Gonads are outlined by a dashed line. Scale bar = 25 µm. Note: Panels **A**, **B**, **E**, **F** are the same represented in **Figure 1** of chapter 3.

This result reinforces the idea that PID-2 is involved in establishing *de novo* silencing of a 21U RNA target, as supported by our first observation of the 21U sensor silencing defects in *pid-2* mutants (**Figure 2**). Nevertheless, the offspring of *pid-2* mutant hermaphrodites are able to re-initiate the silencing of the 21U sensor in approximately half of the population. This observation indicates that PID-2 is not as stringently required for silencing initiation as PRG-1.

To further pinpoint the role of PID-2 in establishing 21U RNA target silencing, we set up a genetic system in which we are able to address the contribution of PID-2 to the silencing that is induced by solely maternally provided 21U RNAs. To this end, we crossed males that express the 21U sensor due to absence of 21U RNAs (*pid-1;pid-2* double mutant) with hermaphrodites that are heterozygous for *pid-1* (*pid-1(+/-);pid-2*). As both parents are *pid-2* mutant, 50% of the offspring will be *pid-1;pid-2* double mutant, and will have received maternal 21U RNAs, without being able to make them themselves. We scored 106 *pid-1;pid-2* double mutant progeny for activity of the 21U sensor and, in line with our hypothesis, none of them is able to silence the 21U sensor (**Figure 4A, C**). We can therefore conclude that PID-2 is absolutely required to establish *de novo* target silencing mediated by maternal 21U RNAs. This is in line with our findings that the silencing established in these crosses has all the characteristics of RNAe (see chapter 3), and our other indications that PID-2 is required for RNAe establishment (**Figures 2**, **3**, **11**, **12**).

Interestingly, not only was the 21U sensor always expressed in *pid-1;pid-2* double mutants, but we also observed that the *pid-1;pid-2* double mutant offspring display defects in germline development and are producing no (84%) or very few offspring (16%; N offspring = 1-14) (**Figure 4B**). The *pid-1;pid-2* double mutant offspring had feminized gonads, as indicated by the arrayed appearance of the oocytes (**Figure 4C**), and the fact that we could rescue the apparent sterility by crossing these offspring with wild type males (not shown).

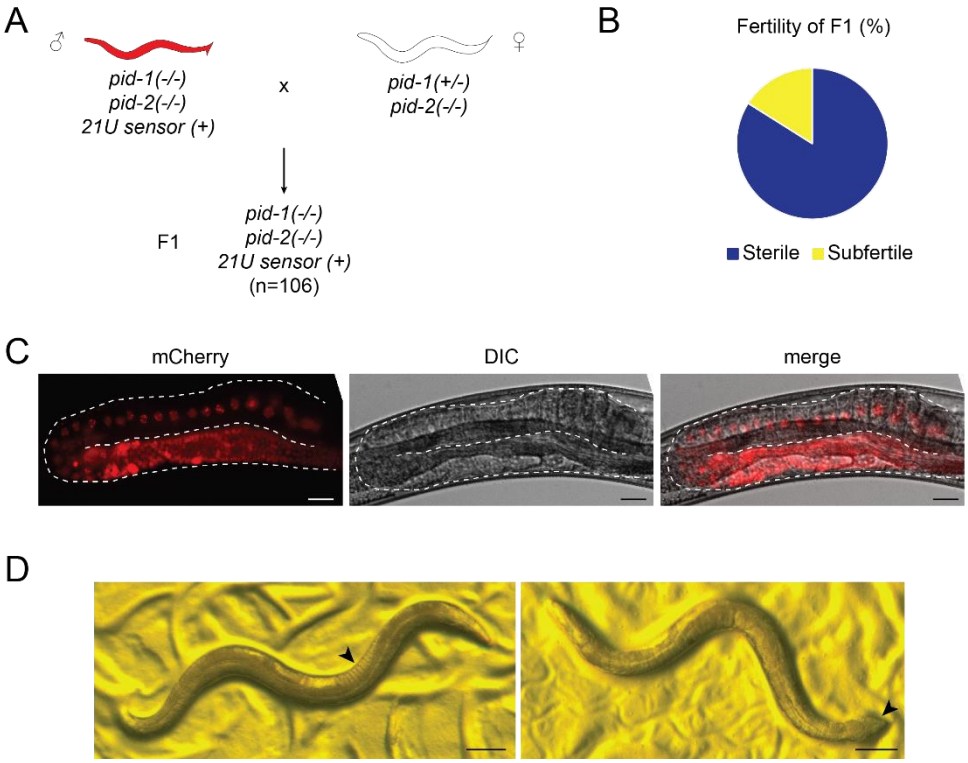


Figure 4. A) Schematic crossing strategy to investigate the contribution of maternally provided 21U RNAs to silencing in F1, in a *pid-2* mutant background. **B)** Pie chart representing the fertility (%) of the F1 from the cross represented in **A**. 84% of the offspring does not produce any embryos (N=89); 16% of the offspring is subfertile (brood size = 1 – 14 worms) (N=17). **C)** Feminized germline of F1. The gonad is outlined by the dashed line. The 21U sensor is always expressed in F1 that are double mutant *pid-1;pid-2*. Scale bar = 25 µm. **D)** *pid-1;pid-2* double mutants are mostly fertile and the strain can be easily maintained, however there is a certain frequency of masculinized (right) and feminized worms (left). Right: the arrow indicates a male-like tail, despite the worm shows other characteristics of hermaphrodites. Left: the arrow indicates the arrayed appearance of oocytes (see also panel **C**). Scale bar = 100 µm.

The fact that the *pid-1;pid-2* mutant offspring was feminized was surprizing, since we generated a *pid-1;pid-2* double mutant strain for this experiment in the first place. We therefore took a closer look at this strain that we had in culture. Despite the fact that the strain grows relatively normally and can be maintained without problems, we did indeed observe animals with a feminized germline (occurrence = 3%). In addition, we also detected animals with a masculinized appearance (occurrence = 10%) (**Figure 4D**). These findings support a role for PID-1 and PID-2 in the sex determination pathway. Recently published data also support a role for a specific 21U RNA in the sex determination pathway (Tang et al., 2018), although a feminization phenotype was not reported. However, it is unclear why the penetrance of this phenotype was so much

stronger in the cross depicted in **Figure 4**, compared to the *pid-1;pid-2* mutant strain that we have in culture.

pid-2 mutants do not cause, or rescue, the mutator induced sterility phenotype

We have previously shown that in absence of parental 22G RNAs (or, in other words, RNAe memory) as well as maternal 21U RNAs, animals that are proficient for 22G RNAs become sterile. The sterility is caused by erroneous silencing of endogenous genes that should be expressed (see chapter 2) (de Albuquerque et al., 2015; Phillips et al., 2015). Because the animals themselves need to have a functional mutator pathway driving 22G RNA biogenesis in order to cause fertility defects, this phenotype was named "mutator induced sterility".

As we found PID-2 to be active in the PRG-1 pathway, we decided to check if re-establishing *de novo* silencing, in absence of both RNAe memory and PID-2, causes mutator induced sterility. We crossed two mutants that lack 22G RNAs because of mutation in two different mutator genes (*mut-7* and *mut-16*) and that are both also *pid-2* mutant. The offspring of this cross will remain *pid-2* mutant, but will be able to produce *de novo* 22G RNAs, because they will inherit a wild type copy of each of the two mutator genes from their parents. These offspring are mostly fertile, comparable to the offspring of *mut-7* and *mut-16* single mutants (**Figure 5**). In comparison, if the *pid-2* mutation would be replaced with a *prg-1* mutation, all the F1 would be sterile (de Albuquerque et al., 2015; Phillips et al., 2015) (see chapter 2).

Next, we tested if lack of PID-2, and hence a hampered 22G RNA pathway downstream of PRG-1, could rescue the mutator induced sterility. To do this, we crossed two *pid-2* mutants that additionally both lack 21U RNAs (due to a *pid-1* or *prg-1* mutation) as well as RNAe memory (due to mutation of *mut-7* in the father and of *mut-16* in the hermaphrodite). Upon crossing these triple mutants, their offspring will stay *pid-2* mutant and have no 21U RNAs, but will have functional mutator activity. If *pid-2* absence would impede the establishment of *de novo* silencing, as in the case of *hrde-1* or *wago-1/-2/-3* mutation (de Albuquerque et al., 2015; Phillips et al., 2015), the sterility of these offspring should be rescued (see chapter 2). However, the offspring are sterile, suggesting that even in absence of PID-2, the production of *de novo* 22G RNAs is sufficient to induce erroneous target silencing (**Figure 5**).

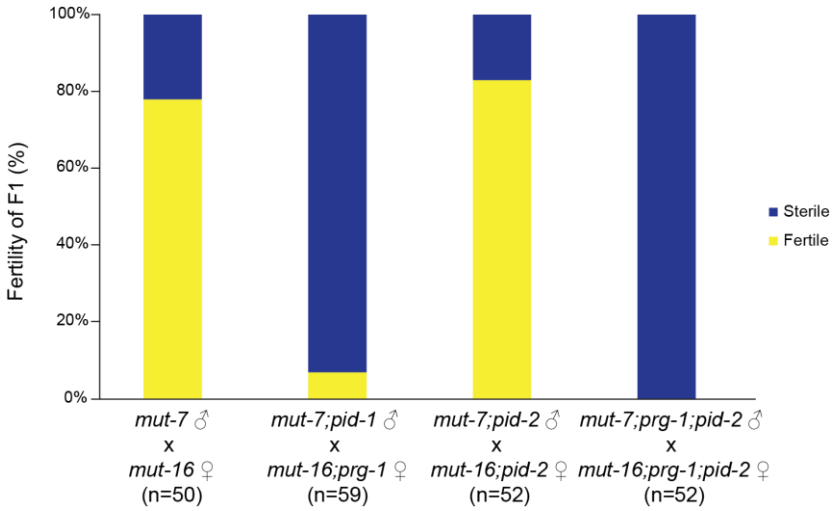


Figure 5. F1 offspring of different crosses were easily identified because of the pan-neuronal GFP marker (punc-119::GFP) transmitted by the males and singled out. The plates were scored two days later for the presence of F2 offspring. The F1 from *mut-7* mutant males and *mut-16* mutant hermaphrodites were fertile in 80% of the cases, comparable to the F1 of the double mutants *mut-7;pid-2* and *mut-16;pid-2*, indicating that reactivation of the mutator pathway does not cause major fertility defects, also in absence of PID-2. On the other hand, the F1 offspring of mutants that lack both mutator activity as well as 21U RNAs (e.g. *mut-7;pid-1* and *mut-16;prg-1*) are mostly sterile, due to the reactivation of the mutator pathway in absence of parental 21U RNAs and RNAe memory. This sterility is not rescued by the addition of *pid-2* mutation (e.g. *mut-7;prg-1;pid-2* and *mut-16;prg-1;pid-2*), as the absence of PID-2 does not prevent the re-establishment of *de novo* silencing in absence of parental 21U RNAs as well as RNAe.

We conclude that lack of PID-2 in the parents cannot trigger mutator induced sterility, and absence of PID-2 in the embryos cannot rescue it. Given the mild phenotypes shown by *pid-2* mutants in the other assays, this result may mean that loss of PID-2 simply does not induce a sufficiently strong effect to reveal a clear phenotype in these sterility assays.

PID-2 is required for prolonged maintenance of a healthy germline

In contrast to other organisms, *prg-1* mutants do not show transposon reactivation nor acute sterility, however they gradually lose germ cells over generations. This is named a mortal germline phenotype, or Mrt (Simon et al., 2014). The Mrt phenotype is not only a characteristic of *prg-1* mutants, but also of mutants for other factors that are involved in the RNAe machinery, such as *hrde-1, nrde-1/-2/-4* (Buckley et al., 2012), the histone H3 lysine 4 (H3K4) methyltransferase *set-2* (Xiao et al., 2011) and the H3K9 methyltransferase homolog *set-32* (Spracklin et al., 2017). We therefore decided

to test if *pid-2* mutants also show a Mrt phenotype. For this, we cultured *pid-2* mutants at 25 °C and followed them over time. As control, we cultured at the same time wild type worms, which do not have a Mrt phenotype, as well as *prg-1* and *hrde-1* mutants, which are known to have a Mrt phenotype (Buckley et al., 2012; Simon et al., 2014).

As expected, wild type worms do not show a Mrt phenotype (after 21 generations), whereas *prg-1* and *hrde-1* mutants become sterile between 6 and 9 generations. *pid-2* mutants also start to become sterile after 9 generations, but it takes up to 18-20 generations for the Mrt phenotype to be fully established (**Figure 6**). We conclude that *pid-2* mutants have a Mrt phenotype, and that this is slightly weaker than that of *prg-1* and *hrde-1* mutants. Possibly, this is connected to the fact that the 21U sensor in *pid-2* mutants is clearly not as strongly activated as in *prg-1* mutants (**Figure 2**).

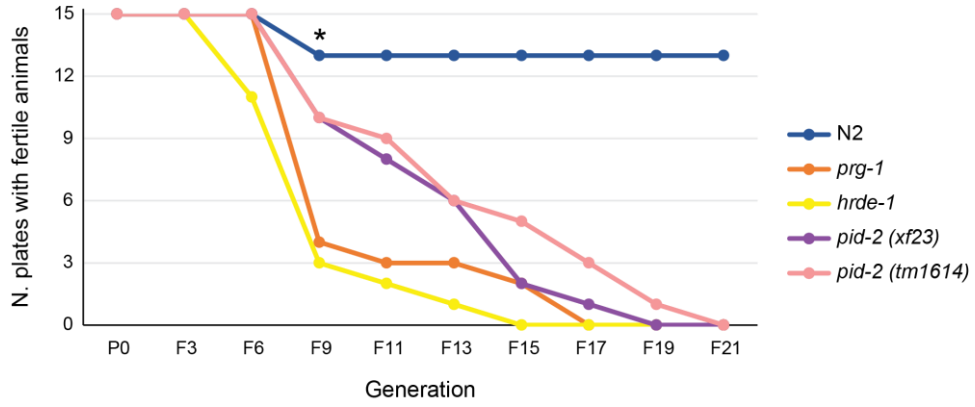


Figure 6. Mrt assay representing the number of fertile plates per generation. Wild type N2 worms grow steadily throughout the assay, whereas *prg-1*, *hrde-1* and *pid-2* mutants gradually become sterile. *: two plates from N2 were contaminated and were excluded from the assay.

PID-2 has a role in Tc1 transposon silencing

As PID-2 has a role in promoting germ cells immortality over time, similarly to PRG-1 and HRDE-1 (**Figure 6**), and it is involved in establishing *de novo* silencing of the 21U sensor (**Figures 1-4**), we tested if PID-2 has also a role in Tc1 transposon silencing.

In order to test the involvement of a particular gene in Tc1 transposon silencing, we use animals that carry a Tc1 insertion (*st136*) in the coding sequence of the *unc-22* gene, which codes for UNC-22/twitchin, a protein required for muscle function and morphology. As a consequence of this Tc1 insertion, the worms are mostly paralyzed and show a twitching phenotype. Whenever Tc1 silencing is impaired, Tc1 will be able to excise itself from the *unc-22* locus. Depending on how the double-strand break is

repaired, this can lead to a restoration of the reading frame, and function of the UNC-22 protein. These so-called reversion events can be easily scored, as the worms start to move normally again (revertants) (Ketting et al., 1999).

We set out to test the role of PID-2 in this transposon-regulation system. On its own, loss of PID-2 causes only very rare reversion events, in a frequency comparable to *prg-1* mutants (10^{-5}) (Das et al., 2008). As published before, also loss of the RNAe factor HRDE-1 does not induce high Tc1 activity (de Albuquerque et al., 2015). However, in combination with loss of HRDE-1, both *prg-1*, as expected (de Albuquerque et al., 2015), and *pid*-2 mutants show a higher frequency of reversion events. On the contrary, no revertants are observed in *pid-2;prg-1* double mutants (**Figure 7**). These data show that PID-2, like PRG-1, has a role in Tc1 silencing, suggesting that probably PID-2 and PRG-1 act in the same branch of the Tc1 silencing network.

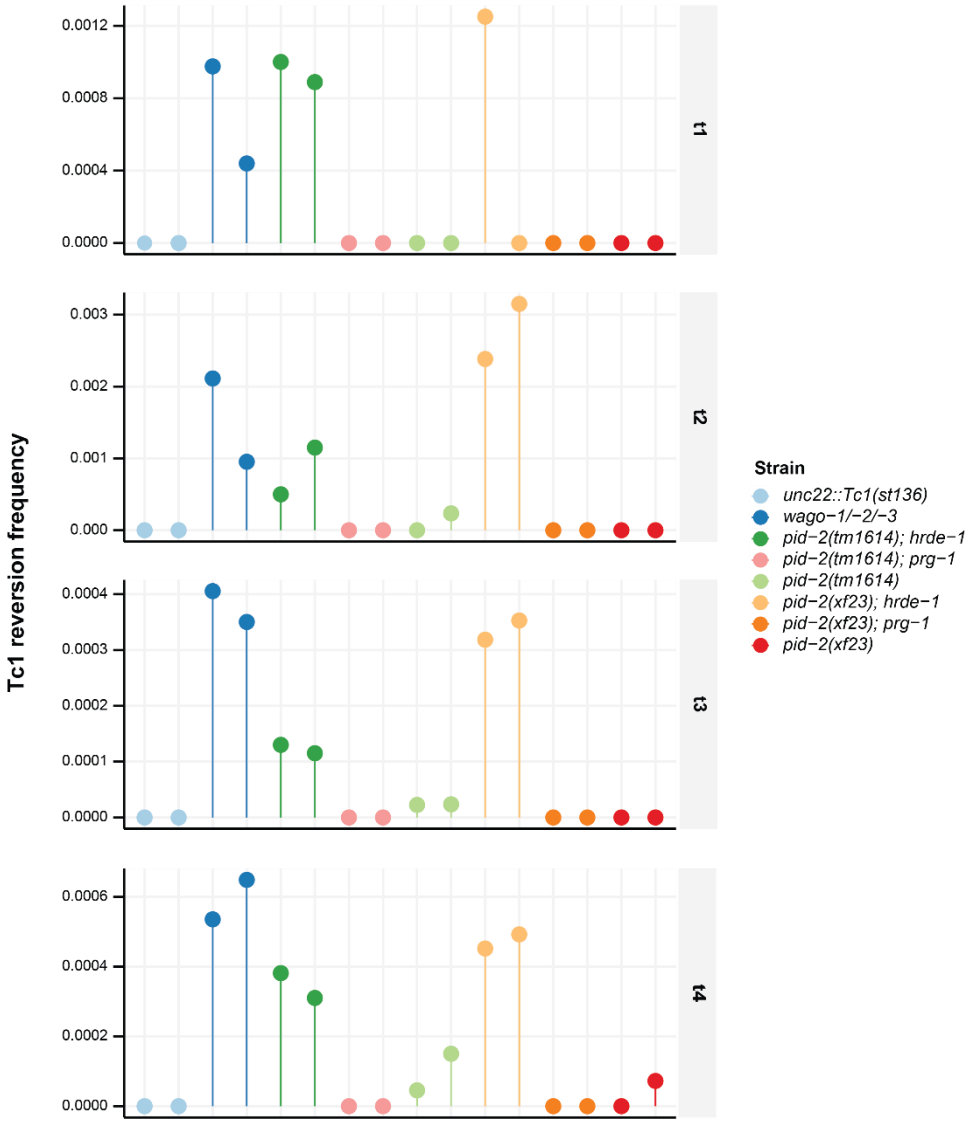


Figure 7. Tc1 reversion assay in different genetic backgrounds at different time points. All the strain tested carry the *unc-22::Tc1(st136)* allele. Negative control = *unc-22::Tc1(st136)* in a wild type background; positive control = *wago-1/-2/-3*. *pid-2* mutants produce very few revertants (frequency 10^{-5} – 10^{-4}), however *pid-2;hrde-1* double mutants produce revertants to levels comparable to the positive control (frequency 10^{-4}). No reversion events have been observed in *pid-2;prg-1* double mutants. t1 = 50 worms/plate; t2 = 100 worms/plate; t3 = 1000 worms/plate; t4 = starvation (1000 worms/plate). For each experiment, 50 plates/strain have been tested. Two independent experiments are represented.

Loss of PID-2 does not heavily affect small RNA populations

In order to see how PID-2 may function, we looked in more detail at the different small RNA classes in *pid-2* mutants by performing small RNA sequencing to uncover any defects caused by the absence of PID-2.

First of all, we looked at the levels of 21U RNAs, as we have identified *pid-2* in a forward genetic screen for 21U RNA silencing defective mutants. Interestingly, 21U RNAs were present in amounts comparable to wild type, while in *prg-1* mutants, they were absent as expected (**Figure 8**). Therefore, we conclude that PID-2 is not involved in 21U RNA biogenesis.

We went on to look for potential defects in 26G or 22G RNA levels, as the latter are responsible for the amplification of the silencing reaction. Regarding 26G RNAs, there is a big variability among replicates, which could be explained by the low numbers of 26G RNAs themselves. Seemingly, *pid-2* mutants have reduced levels of 26G RNAs compared to wild type and to *prg-1* mutants (**Figure 8**). Yet, overall, *pid-2* mutants did not show any phenotype that could be related to the 26G RNA pathway, and we have not investigated this aspect further. At a global level, 22G RNAs are not severely affected by the absence of PID-2, although *pid-2* mutants have significantly less 22G RNAs than wild type (**Figure 8**).

As control, we also checked the levels of miRNAs, which indeed are comparable among the different strains analysed (**Figure 8**).

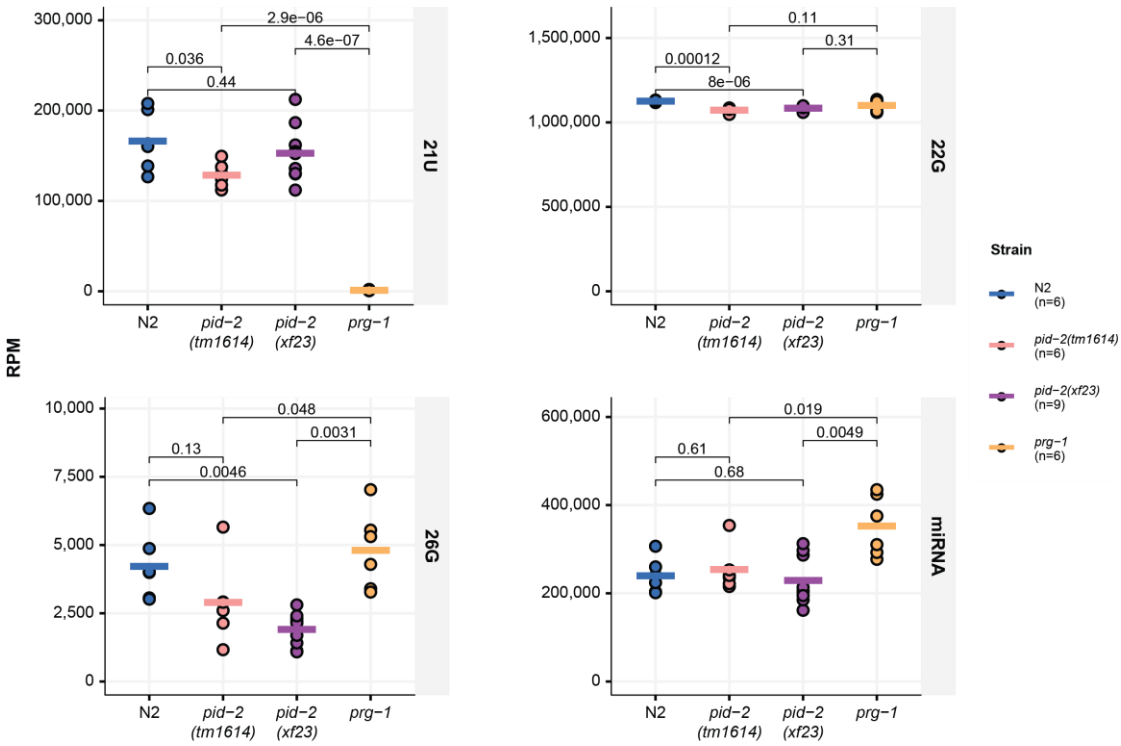


Figure 8. Small RNA levels of 21U, 22G, 26G RNAs and miRNAs in wild type, *pid-2* and *prg-1* mutants. Despite some variation among the replicates, there is no striking difference between wild type and *pid-2* mutants, except for a slight but significant reduction in 22G and 26G RNA levels. *prg-1* mutants are devoid of 21U RNAs, as expected. Each dot represents one replicate and the median is represented. P values are calculated with a two-tailed unpaired t-test.

As *pid-2* mutants showed a slight reduction of global 22G RNA levels, which could account for their inability to establish *de novo* silencing (**Figures 1-4**) as well as for their defects in Tc1 silencing (**Figure 7**), we sought to pinpoint which 22G RNA populations are specifically affected. We looked at 22G RNAs complementary to genes that have been previously reported to be targeted by a particular pathway: CSR-1 (Conine et al., 2013); ALG-3/-4; ERGO-1 (Almeida et al., 2019); mutator (Phillips et al., 2014); NRDE-3 (Zhou et al., 2014); PRG-1 (Bagijn et al., 2012) and WAGO-1 (Gu et al., 2009) targets. We observed that the 22G RNA levels are reduced in all the classes; most affected by loss of PID-2 are those mapping to CSR-1, mutator, NRDE-3 and PRG-1 targets (**Figure 9**). We conclude that loss of PID-2 leads to a general, but very modest, reduction of 22G RNAs from endogenous genes.

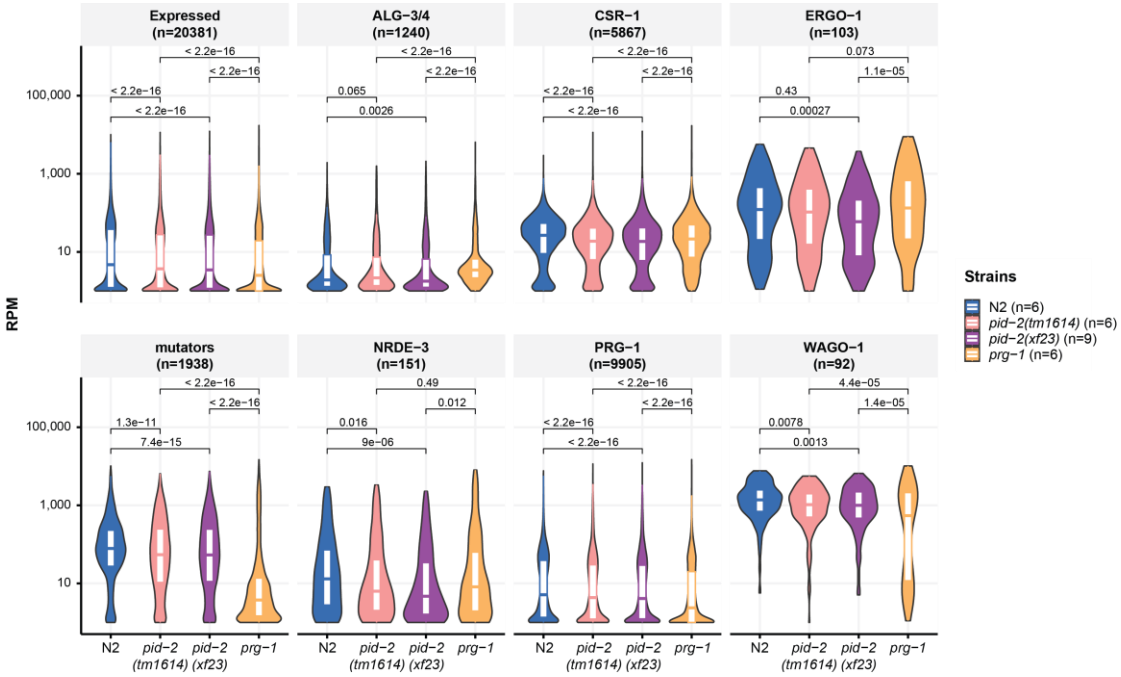


Figure 9. Violin plots representing the distribution of 22G RNAs mapping to targets of different small RNA pathways in wild type, *pid-2* and *prg-1* mutants. There is a global reduction of 22G RNAs in *pid-2* mutants compared to wild type (expressed). The classes of 22G RNAs that are most affected upon loss of PID-2 are mapping to CSR-1, mutator, NRDE-3 and PRG-1 targets. The white boxes inside each of the violin plot represent the 75th and the 25th percentile of the distribution, top and bottom respectively. The median of the distribution is represented by the line in each box. P values are calculated with a two-sided unpaired Mann-Whitney/Wilcoxon rank-sum test, indicating the differences between *pid-2* mutants and either wild type or *prg-1* mutant as reference.

We then analysed further the production of 22G RNAs on endogenous, previously identified 21U RNA target sites (Bagijn et al., 2012; Lee et al., 2012). We focused specifically on a 100 bp window, surrounding the 21U RNA target sites, to analyse the synthesis of secondary 22G RNAs. It is clear that in *pid-2* mutants the production of antisense 22G RNAs is lower compared to a wild type background, although still ongoing. As expected, the production of 22G RNAs is completely abolished in *prg-1* mutants, as the target sites cannot be recognized and therefore the downstream production of 22G RNAs is not triggered (**Figure 10**).

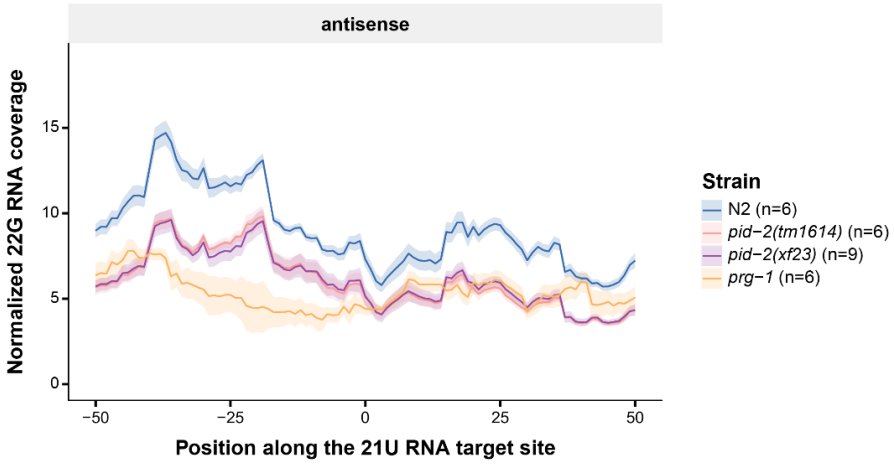


Figure 10. Profile of antisense 22G RNAs produced in a 100 bp window around endogenous 21U RNA target sites (50 bp upstream, 50 bp downstream) of WAGO-1 target genes, centred on position 10 of the 21U RNA sequence. Each line represents a specific genotype and the shading represents the standard deviation for each genotype, sequenced in triplicate. Although the levels of 22G RNAs are generally low, there is a slight reduction in *pid-2* mutants compared to wild type, whereas in *prg-1* mutants the production of secondary 22G RNAs is practically abolished.

This result further supports the idea that PID-2 is not essential for the recognition of 21U RNA target sites per se, but generally stimulates the production of 22G RNAs.

Loss of PID-2 causes a reduced production of 22G RNAs on the 21U sensor and affects secondary 22G RNA production on endogenous 21U RNA target sites

Next, we analysed specifically the 22G RNAs that are derived from the 21U sensor, which was included in all the sequenced strains. When possible, we created and sequenced strains with this transgene either in an active or in a silenced (RNAe) state.

In a wild type background, the 21U sensor is silenced via RNAe, therefore, we can identify two distinct populations of 22G RNAs: the first is produced locally around the *21ur-1* recognition site and depends on PRG-1 recognition of the target (secondary 22G RNAs); the second instead covers the whole mCherry coding sequence (tertiary 22G RNAs) and is independent of PRG-1, one of the characteristics of RNAe (**Figure 11A**) (Sapetschnig et al., 2015).

In *prg-1* mutants, we can observe two different scenarios. If the silencing of the 21U sensor depends on PRG-1, the transgene cannot be recognized nor silenced, therefore there are no 22G RNAs produced on the 21U sensor transcript, which is in turn

expressed (**Figure 11B**). On the other hand, if the silencing of the 21U sensor has become independent of PRG-1, the transgene is silenced by RNAe and we observe production of 22G RNAs along the 21U sensor mRNA, recapitulating the wild type situation (**Figure 11C**) (Bagijn et al., 2012).

If the silencing mediated by 22G RNAs is abolished by mutation of *hrde-1* or *mut-7*, RNAe cannot be established. As expected, *hrde-1* mutants still display production of secondary 22G RNAs on the 21U sensor (**Figure 11D**). In fact, HRDE-1 is specifically involved in RNAe establishment and maintenance (Buckley et al., 2012; Sapetschnig et al., 2015). On the other hand, the synthesis of 22G RNAs is practically abolished upon loss of MUT-7, therefore no 22G RNAs, secondary or tertiary, can be detected along the 21U sensor mRNA (**Figure 11E**) (Gu et al., 2009; Zhang et al., 2011).

In *pid-2* mutants, there is still production of secondary 22G RNAs on an active 21U sensor, but at lower levels than in wild type or *prg-1* mutant strains (**Figure 11F**, **H**), recapitulating the scenario of the production of secondary 22G RNAs on endogenous 21U RNA target sites (**Figure 10**). If the 21U sensor was already stably silenced via RNAe, also tertiary 22G RNAs were still produced, although their levels were also reduced compared to wild type and *prg-1* mutants (**Figure 11G**, **I**).

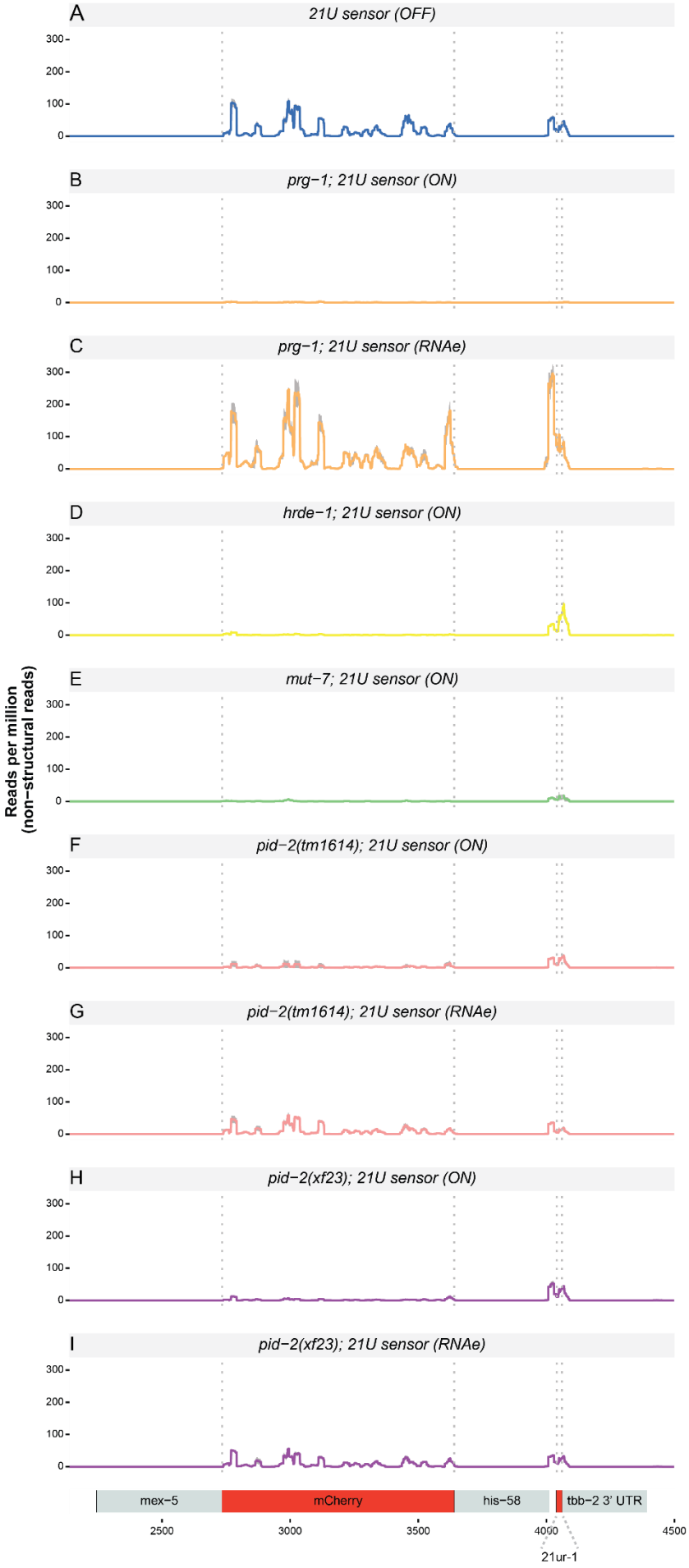

A
21U sensor (OFF)
B
prg−1; 21U sensor (ON)
C
prg−1; 21U sensor (RNAe)
D
hrde−1; 21U sensor (ON)
E
mut−7; 21U sensor (ON)
F
pid−2(tm1614); 21U sensor (ON)
G
pid−2(tm1614); 21U sensor (RNAe)
H
pid−2(xf23); 21U sensor (ON)
I
pid−2(xf23); 21U sensor (RNAe)
Reads per million (non−structural reads)
300
200
100
0
mex−5
mCherry
his−58
tbb−2 3' UTR
2500
3000
3500
4000
4500
21ur-1

Figure 11. 22G RNAs mapping against the 21U sensor locus, schematically represented at the bottom. *pid-2* mutants still display local production of secondary 22G RNAs around the *21ur-1* recognition site triggered by PRG-1 and mutator activity (**F**, **H**); these secondary 22G RNAs are in fact lost in *prg-1* (**B**) and *mut-7* (**E**) mutants. *pid-2* mutants are not able to produce the 22G RNAs upstream of the *21ur-1* recognition site, therefore they cannot establish RNAe, similarly to *hrde-1* mutants (**D**). On the contrary, if the 21U sensor was already silenced by RNAe (**C**), the 22G RNAs mapping throughout the mCherry coding sequence are still present (**G**, **I**), although their levels are lower compared to a *prg-1* mutant strain (**C**) or to a wild type strain (**A**). In each plot, the average of three biological replicates is represented and the shading represents the standard deviation among the replicates.

We then quantified the total levels of secondary and tertiary 22G RNAs produced on the 21U sensor in the different strains represented in **Figure 11** (**Figure 12**). Again, we clearly see that the production of secondary and tertiary 22G RNAs in *pid-2* mutants is lower than the wild type situation, and more similar to *mut-7* mutants. Moreover, *pid-2* mutants having a 21U sensor silenced via RNAe, produce higher levels of tertiary 22G RNAs than the same mutants with an active 21U sensor (**Figure 12**). Yet, the levels of tertiary 22G RNAs are lower than a wild type or *prg-1* mutant background (**Figure 12**). These data indicate that PID-2 has a role in stimulating both secondary and tertiary 22G RNA production, without being essential for it.

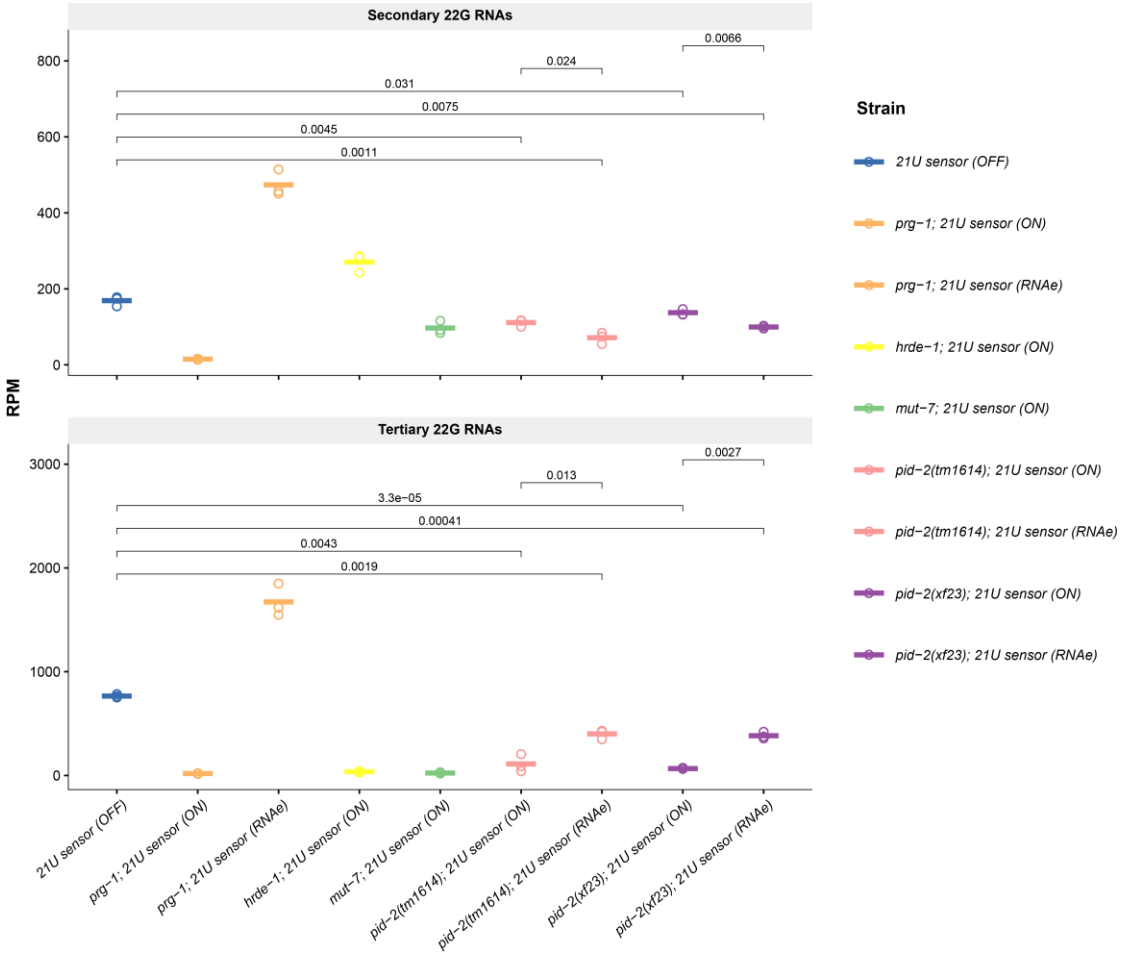


Figure 12. Dot plot for quantification of the production of secondary (upper panel) and tertiary 22G RNAs (lower panel) on the 21U sensor in different genetic backgrounds. The quantification of the levels of 22G RNAs reflects the plots represented in **Figure 11**, as the strains used are the same. Each replicate is represented by a dot and the median is represented. P values are calculated with a two-tailed unpaired t-test.

Together, these results argue that PID-2 is not required for the recognition of the 21U RNA target site per se, as secondary 22G RNAs are being produced, independently of the activity status of the 21U sensor. Possibly, their production in absence of PID-2 is below a threshold level, not sufficient to trigger RNAe on an active 21U sensor. Additionally, also tertiary 22G RNA levels go down in *pid-2* mutants, even though at the phenotypic level, the 21U sensor under RNAe, and hence producing tertiary 22G RNAs, is not affected.

PID-2 interacts with two predicted Tudor domain proteins

We then focussed on the characterization of the interactome of PID-2. We used both the polyclonal antibody αPID-2 we raised (**Figure 13B**) as well as the transgenic lines

we generated (**Figure 13A**) to perform immunoprecipitation followed by quantitative label-free mass spectrometry (IP/MS). Independently of the antibody used, that is, either αPID-2 to enrich for the endogenous protein in a wild type background, or αGFP or αFLAG (not shown) antibodies to enrich for the tagged PID-2 in the transgenic lines, two interactors were found in all the experiments: W03G9.2 and Y45G5AM.2, both non-characterized proteins. Based on the fact that they also affect 21U RNA mediated silencing (see below), we renamed W03G9.2 and Y45G5AM.2 as *pid-4* and *pid-5*, respectively.

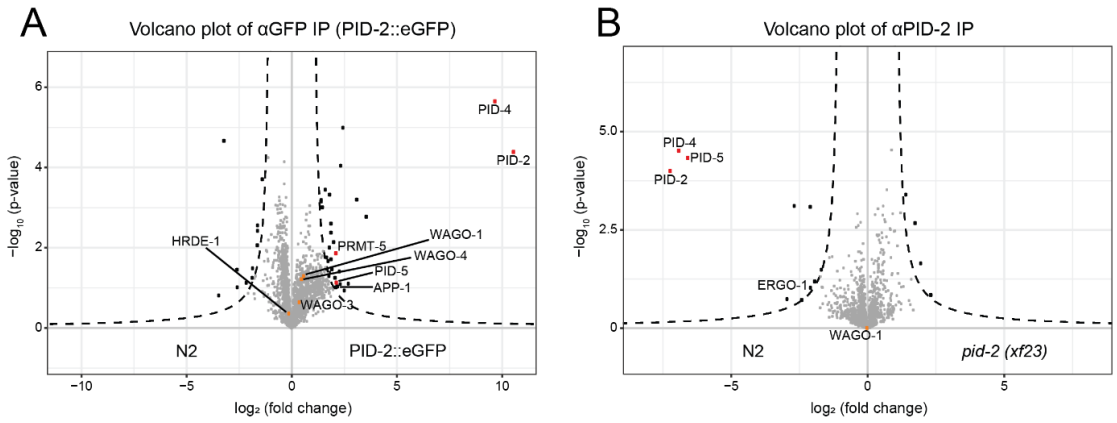


Figure 13. A) Volcano plot representing the enrichment of proteins that are interacting with PID-2. We immunoprecipitated the eGFP::PID-2 protein, which is highly enriched, and we could identify some interactors. As control, we used wild type protein extracts, not containing any eGFP tagged protein. For each set of IP, we used 4 replicates for every strain. **B)** Volcano plot representing the enrichment of proteins that are interacting with PID-2. We immunoprecipitated PID-2 using the polyclonal antibody against the endogenous protein. PID-2 is highly enriched in wild type protein extracts, compared to *pid-2* mutant protein extracts. For each set of IP, we used 4 replicates for every strain. Two proteins are clearly interacting with PID-2: W03G9.2/PID-4 and Y45G5AM.2/PID-5.

Both PID-4 and PID-5 have a predicted Tudor domain (**Figure 14C**). Tudor domains recognize (symmetrically di-) methylated arginine or lysine of their substrates and very often have a role in RNA metabolism, in particular in the piRNA pathway, and are required for fertility (Pek et al., 2012). In many organisms, Tudor domains act to recruit effector proteins. It is very exciting to speculate that PID-4 and PID-5 could serve as a platform for the formation of protein complexes with the aim of boosting the 22G RNAs production.

PID-5 has in addition a predicted Xaa-Pro aminopeptidase domain, which is very homologous to that found in APP-1. This domain cleaves the most N-terminal amino acid from a peptide, provided that the second amino acid is a proline. APP-1 is a

strongly conserved enzyme, from yeast to human, with roles in the catabolism of proline-containing polypeptides. Interestingly, the catalytic site of PID-5 appears to be catalytically inactive, as the residues required to coordinate the two Zn^{2+} ions for enzymatic activity are not conserved (**Figure 14D**).

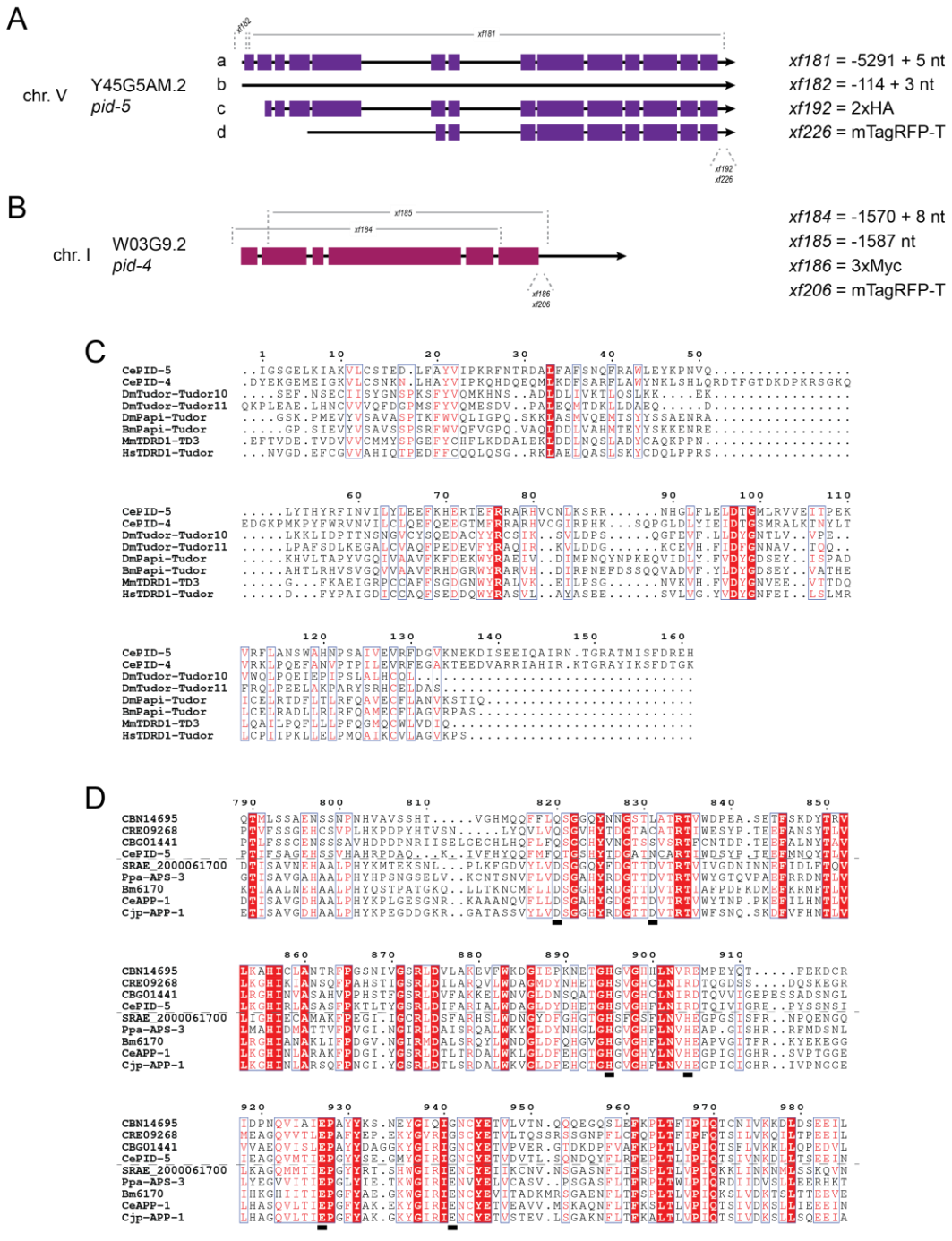


Figure 14. A) Genomic locus of *pid-5* and predicted transcripts. The alleles generated using CRISPR/Cas9 technology are indicated. **B)** Genomic locus of *pid-4* and predicted transcript. The alleles generated using CRISPR/Cas9 technology are indicated. **C)** Multiple sequence alignment, using ESPript 3 (http://espript.ibcp.fr) (Robert & Gouet, 2014), of the predicted Tudor domains of PID-5/Y45G5AM.2 and PID-4/W03G9.2 with the more closely related Tudor domains of other organisms. Ce,

Caenorhabditis elegans. Dm, *Drosophila melanogaster*. Bm, *Bombyx mori*. Mm, *Mus musculus*. Hs, *Homo sapiens*. **D)** Multiple sequence alignment of the Xaa-Pro aminopeptidase domain of PID-5 (CePID-5) with its orthologs and with APP-1 (CeAPP-1) and its orthologs in other nematods: *C. brenneri* CBN14695; *C. remanei* CRE09268; *C. briggsae* CBG01441; *S. ratti* SRAE_2000061700; *P. pacificus* Ppa-APS-3 PPA08577; *B. malayi* Bm6170; *C. japonica* Cjp-APP-1. If the conservation is >70%, the columns are framed in blue. The residues perfectly conserved are highlighted in red. The residues written in black are not conserved. Dots (.) indicate gaps in the protein sequence generated through the alignment. The dashed line separates the protein according to the presence or absence of the catalytic residues, which are underlined.

PID-2 is expressed in the adult germline and in embryos and localizes to perinuclear granules together with PGL-1 and PID-4

Taking advantage of the transgenic lines we generated, we performed confocal microscopy to investigate the localization and expression pattern of PID-2 tagged with eGFP.

PID-2::eGFP (*xfIs145*) is expressed in the adult germline, as well as in embryos. We could not detect significant signal in other tissues. Both in the germline and in the embryo, PID-2::eGFP is expressed in the cytoplasm where it localizes to granules, possibly P granules.

In the germline, expression of PID-2 is not homogeneous. PID-2 signal is very faint, but present, throughout the whole germline. Nevertheless, the highest expression is observed from the transition zone, throughout the meiotic zone, specifically at the pachytene stage, where it colocalizes with PGL-1, a P granule marker (Kawasaki et al., 1998) and then it gradually dissolves during oogenesis (**Figure 15A**). This pattern is also observed in a gonad that is spermatogenic.

Curiously, PGL-1 signal is present from the mitotic region of the germline throughout meiosis, but it becomes fainter during pachytene, exactly at the stage when PID-2 is most strongly expressed. This could be due to a FRET (Fluorescence Resonance Energy Transfer) phenomenon (Forster, 1946); in fact, PGL-1 is endogenously tagged with mTurquoise2, which is known to be a FRET partner of eGFP (Bajar et al., 2016). In fact, in absence of any eGFP signal, PGL-1 is homogeneously expressed throughout the whole germline and we do not observe such decreased intensity of the signal. This effect is especially apparent during the fourth larval stage (L4) (not shown). Nonetheless, this observation supports even more strongly the colocalization between

PID-2 and PGL-1, as PGL-1 signal becomes weaker exactly when PID-2 signal is stronger.

We also introduced a fluorescent tag at the endogenous loci of *pid-4* and *pid-5* (**Figure 14A**, **B**). For both loci, we introduced at the C-terminal position a red fluorescent protein (RFP) to visualize the endogenous expression pattern of PID-4 and PID-5. PID-4::RFP (*xf206*) and PID-5::RFP (*xf226*) show very similar expression to PID-2::eGFP. They are expressed throughout the germline and they are most highly expressed during the pachytene stage, when they are present in perinuclear granules. Furthermore, PID-4 perfectly colocalizes with PID-2, confirming the interaction identified by IP/MS, and with PGL-1 (**Figure 15A**). Even though we do not have colocalization data for PID-5 yet, from its expression pattern (not shown) and from the IP/MS results, we can assume that it also colocalizes with PID-2, PID-4 and PGL-1 and that the PID-2/PID-4/PID-5 complex is localized in P granules.

In embryos, we could also detect the highest expression of PID-2 in perinuclear granules, specifically in the P cell lineage, and later in Z2/Z3 cells, which are the primordial germ cells (PGCs) that will then give rise to the germline during development. Also in the embryos, PID-2 colocalizes with PGL-1, supporting its localization to P granules (**Figure 15B**). Furthermore, PID-4 and PID-5 are also specifically expressed in PGCs and form perinuclear granules. PID-4 colocalizes with PID-2 and PGL-1 also in embryos, supporting indeed the localization of this novel complex in P granules (**Figure 15B**).

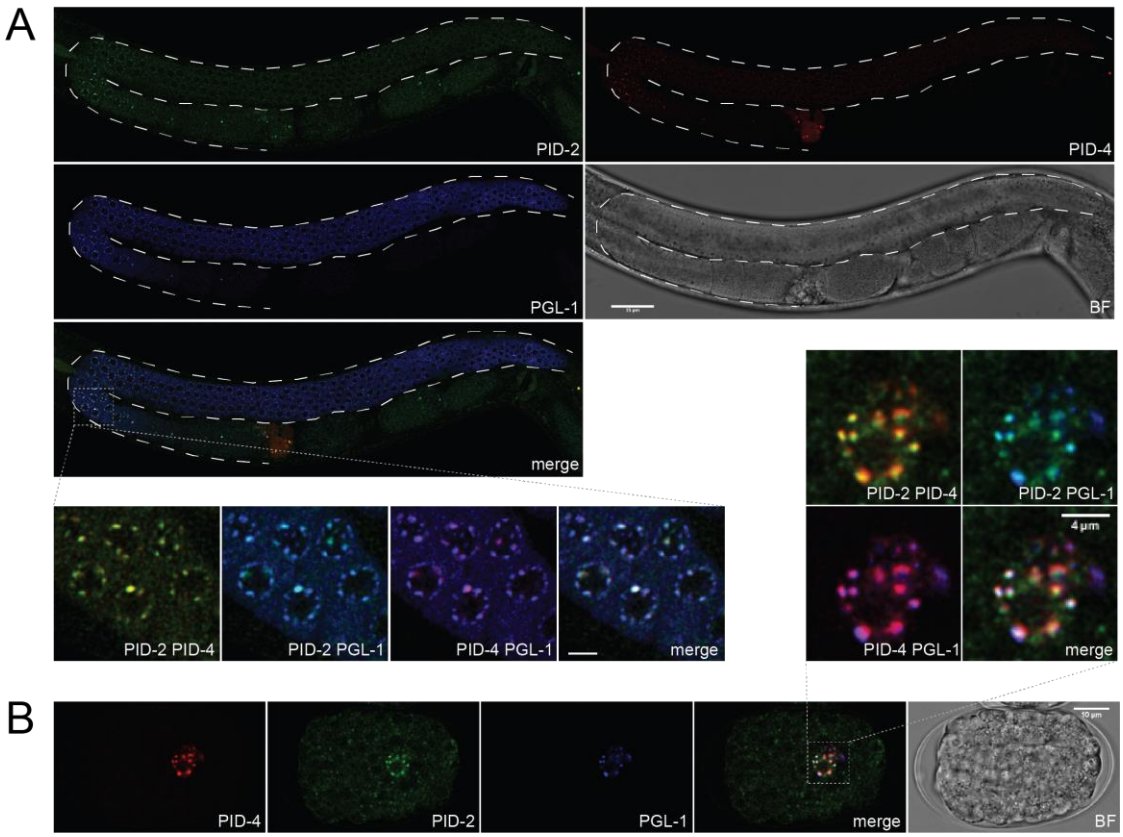


Figure 15. Confocal imaging of PID-2, PID-4 and PGL-1 in the adult germline (**A**) and in the embryo (**B**). **A**) In the adult germline, PID-2 and PID-4 show a similar expression pattern; they are localized in perinuclear granules, with higher expression through the transition zone to the pachytene stage of meiosis. The white dashed line is highlighting the gonad. Scale bar = 25 µm, and 4 µm for the magnification on the lower left. **B)** PID-2, PID-4 and PGL-1 are specifically expressed in Z2/Z3 cells and also show a perinuclear localization. Scale bar = 10 µm, and 4 µm for the magnification on the upper right. Both in the adult germline (**A**) and in embryos (**B**), PID-2 and PGL-1 colocalize to P granules.

pid-4 and *pid-5* mutants do not show any major defects

In order to further characterize *pid-4* and *pid-5*, we have used CRISPR/Cas9 technology to generate mutants as well as to tag the endogenous loci with an epitope and with a fluorescent protein (Arribere et al., 2014; Mouridi et al., 2017; Paix et al., 2014; Ward, 2015) (**Figure 14A**, **B**).

We have generated *pid-4* and *pid-5* mutants in which we have deleted either the 5' UTR together with the start codon or the almost complete coding sequence. As *pid-5* encodes four different transcripts, of which three are coding, according to Wormbase (https://www.wormbase.org/species/c_elegans/gene/WBGene00021555#0c4-9gd-3), we aimed at deleting the whole genetic locus in order to affect all the different

transcripts (**Figure 14A**). We then used the alleles *xf184* and *xf181* as *pid-4* and *pid-5* mutants, respectively, throughout the study.

pid-4 and *pid-5* mutants do not show any apparent defects; they phenotypically behave like wild type worms. *pid-2* mutants do not show any obvious phenotypes either, so it is not completely unexpected that its interactors do not display gross developmental defects.

Loss of PID-4 or PID-5 does not affect small RNA populations

We first checked if loss of PID-4 or PID-5 affects any of the small RNA classes. Similar to what we previously described for *pid-2* mutants, we first looked at the levels of 21U RNAs: they are not affected and are comparable to wild type. As expected, *prg-1* mutants are devoid of 21U RNAs (**Figure 16**). Therefore, we can conclude that PID-4 and PID-5 are not participating in 21U RNA biogenesis.

We also checked the levels of 26G and 22G RNAs, but neither of these classes are affected in *pid-4* and *pid-5* mutants (**Figure 16**). To exclude any biases in our analysis, we also made sure that the levels of miRNAs are similar among the different strains sequenced (**Figure 16**).

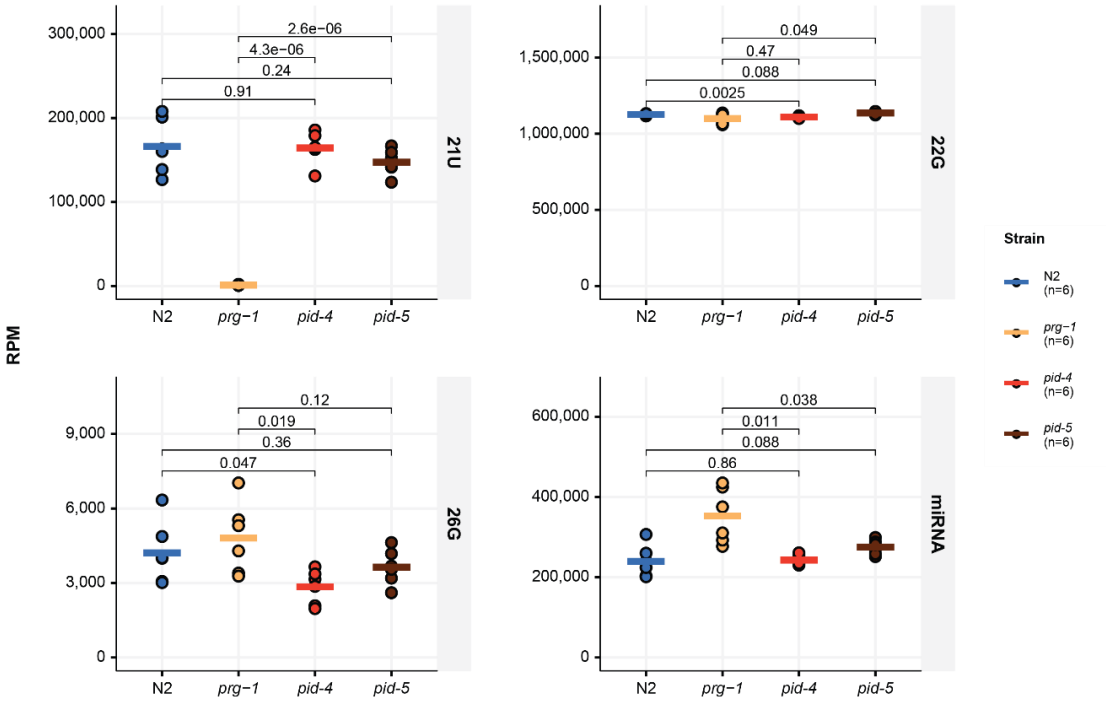


Figure 16. Analysis of 21U, 22G, 26G RNA and miRNA levels in *pid-4* and *pid-5* mutants. None of the small RNA classes shows differences compared to wild type. *prg-1* mutants are devoid of 21U RNAs,

as expected. Each dot represents one replicate and the median is represented. P values are calculated with a two-tailed unpaired t-test.

PID-4 and PID-5 act redundantly in 21U sensor silencing

As we have shown that *pid-2* mutants have defects in the silencing of the 21U sensor, we next investigated the behaviour of the 21U sensor in *pid-4* and *pid-5* mutants. We therefore introduced a 21U sensor, either active or under RNAe, into the *pid-4* and *pid-5* mutant backgrounds, and probed the 21U sensor expression status. We observed that the 21U sensor was always silenced, both in *pid-4* as well as in *pid-5* mutants, independently of the initial expression status (**Figure 17**).

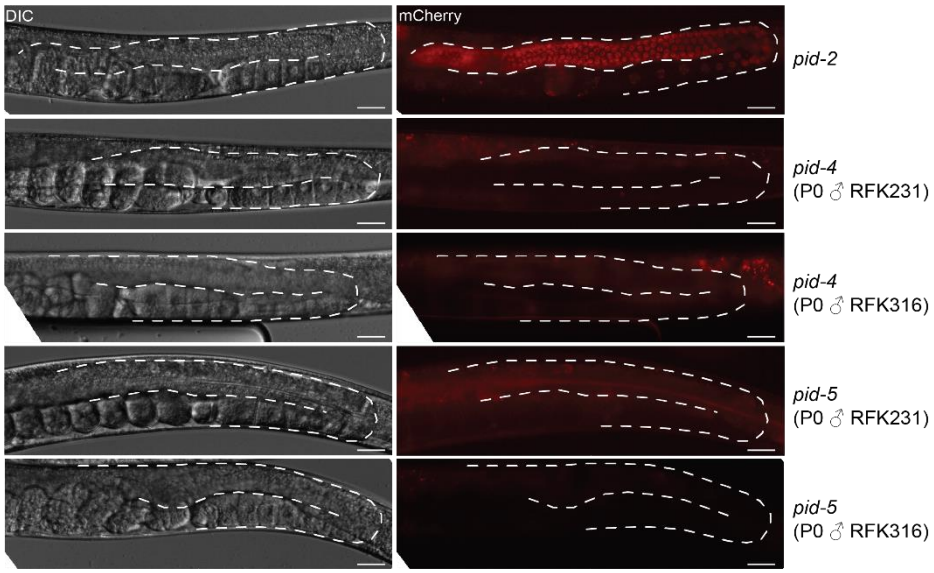


Figure 17. Expression of the 21U sensor in the indicated genetic backgrounds. *pid-2* mutants express the sensor, however *pid-4* and *pid-5* mutants are able to silence the 21U sensor, if it has been introduced in an active state (P0 ♂ RFK231). Furthermore, *pid-4* and *pid-5* mutants keep the silenced state of a 21U sensor previously silenced via RNAe (P0 ♂ RFK316). Gonads are outlined by a dashed line. Scale bar = 25 μm.

Despite the small RNA sequencing did not show any major defects on endogenous small RNAs (**Figure 16**) and the 21U sensor activity does not seem to be impaired (**Figure 17**), we analysed further the 22G RNAs that are derived from the 21U sensor in *pid-4* and *pid-5* mutant backgrounds to uncover possible mild defects. As described above, the 21U sensor is silenced by RNAe in a wild type background, therefore both secondary and tertiary 22G RNAs are produced (**Figure 18A**: same panel as in **Figure 11A**). In a *prg-1* mutant background, the 21U sensor can be either active or in an RNAe state. In the first scenario, no 22G RNAs are produced on the 21U sensor transcript

(**Figure 18B**: same panel as in **Figure 11B**), whereas in the latter case, we can observe production of both secondary and tertiary 22G RNAs (**Figure 18C**: same panel as in **Figure 11C**). In accordance with the microscopy data (**Figure 17**), independently of the initial activity status of the 21U sensor, *pid-4* and *pid-5* mutants produce both secondary as well as tertiary 22G RNAs (**Figure 18D-F**), recapitulating the wild type situation (**Figure 18A**). In conclusion, our sequencing results show that *pid-4* and *pid-5* single mutants do not have defects in any of the small RNA populations (**Figure 16**, **18**).

Figure 18. 22G RNAs mapping antisense to the mCherry::21U sensor locus in *pid-4* and *pid-5* mutants. Two distinct populations of 22G RNAs can be observed: a local production around the *21ur-1* recognition site (secondary 22G RNAs) and along the mCherry coding sequence, spreading along the gene towards the 5' end (tertiary 22G RNAs) (Sapetschnig et al., 2015). Independently of whether the 21U sensor was initially silenced via RNAe (**E**, **G**) or active (**F**, **G**), in both *pid-4* (**D**, **E**) and *pid-5* (**F**, **G**) mutants the silencing is maintained or re-initiated, respectively. The 22G RNAs population show the pattern characteristic of RNAe, although the levels of 22G RNAs are lower than in the control (**A**, **C**). In each

plot, the average of three biological replicates are represented and the shading represents the standard deviation among replicates. Note: panels **A**, **B**, **C** are the same represented in **Figure 11**.

Given the fact that both PID-4 and PID-5 contain a very homologous Tudor domain, and that the single mutants do not show any major silencing defects, we hypothesised that they may be partially redundant. We therefore tested whether a *pid-4;pid-5* double mutant affects the silencing of the 21U sensor.

We have shown that maternally provided 21U RNAs are required to establish *de novo* silencing of the 21U sensor (**Figure 19A**, **B**, **E**, **F**) (de Albuquerque et al., 2014) (see also chapter 3). Thus we repeated the same experiment and tested whether *pid-4;pid-5* double mutants have defects in establishing *de novo* target silencing. To this end, we crossed *mut-7* mutant males expressing the 21U sensor, as devoid of 22G RNAs, with *pid-4;pid-5* double mutant hermaphrodites. The *pid-4;pid-5* double mutant hermaphrodites behave similarly to wild type hermaphrodites, in this assay: their offspring inherit 21U RNAs that are sufficient to re-establish the silencing of the 21U sensor (**Figure 19A**, **B**, **C**, **D**). This result differs from that obtained with *pid-2* mutant hermaphrodites. In their F1, following this cross, the sensor is silenced in less than half of the population, whereas the majority of the offspring expresses the 21U sensor, indicating that *pid-2* hermaphrodites fail to provide their offspring with all the factors required to establish silencing (**Figure 3C**, **D**, **G**, **H**).

Following up on this cross, we isolated F2 animals that were homozygous mutant for both *pid-4* and *pid-5*. In these animals we observed that the 21U sensor is expressed, at levels that are very similar to what we observe in *pid-2* mutants (**Figure 19G**).

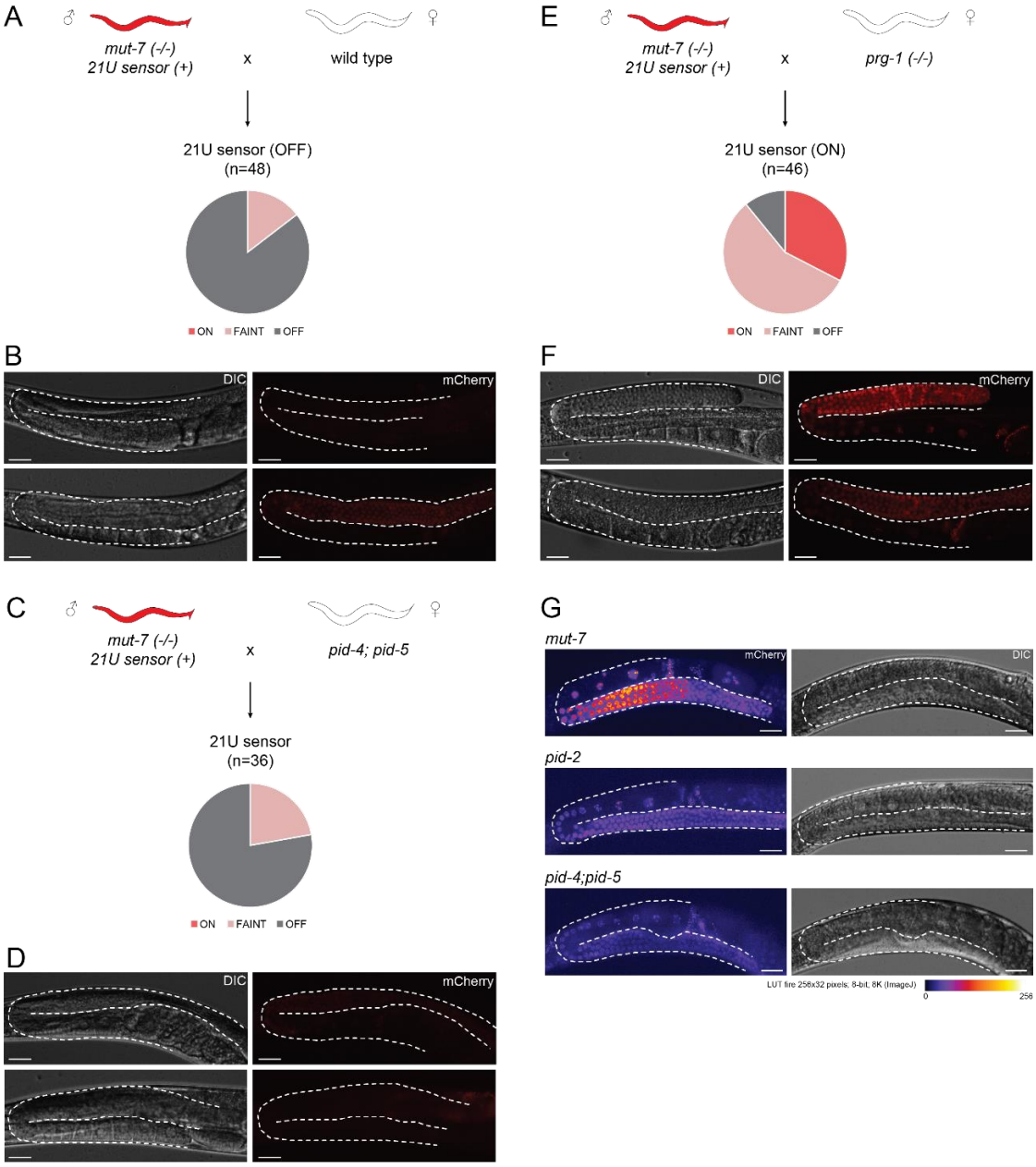


Figure 19. A) Scheme representing the cross strategy to address the re-silencing of the 21U sensor. A *mut-7* mutant male expressing the 21U sensor, because of depletion of 22G RNAs, is crossed either with a wild type (**A**), *prg-1* (**E**) or *pid-4;pid-5* mutant (**C**) hermaphrodite. The offspring are able to establish *de novo* silencing on the 21U sensor, if they inherit a pool of 21U RNAs from the hermaphrodite, either wild type (**A**) or *pid-4;pid-5* (**C**) mutant, in 85% or 78% of the population, respectively (see pie chart). If the hermaphrodite cannot transmit any 21U RNAs to the next generation (*prg-1* mutant, **E**), the offspring is not able to silence the 21U sensor (89%; see pie chart). **B**, **D**, **F)** Representative images of the 21U sensor that is either silenced, faintly expressed or expressed (left, DIC pictures of one gonad; right: mCherry signal). Gonads are outlined by a dashed line. Scale bar = 25 µm. **G)** Expression of the 21U sensor in the indicated genetic backgrounds. In *mut-7* mutants, the 21U sensor is strongly expressed because of absence of 22G RNAs as well as RNAe. In *pid-2* and in *pid-4;pid-5* double mutants, the 21U sensor is also expressed, although at lower levels. Scale bar = 25 µm. The mCherry signal is represented in pseudo-colours [LUT fire (ImageJ)] to reflect differences in the intensity of the signal.

These data indicate that PID-4 and PID-5 individually do not have an essential role in establishing *de novo* target silencing, but that together they are required for efficient 21U silencing, supporting our hypothesis that PID-4 and PID-5 can act, at least partially, redundantly.

PID-4 and PID-5 are required for germline integrity over time

Thus far, our data show that *pid-4;pid-5* double mutants phenocopy *pid-2* mutants. We therefore tested if *pid-4* or *pid-5* mutants or the double mutants *pid-4;pid-5* also show a Mrt phenotype. Again, we cultured different mutant strains at 25 °C and followed them over time. As control, we cultured at the same time wild type worms, which do not have a Mrt phenotype, as well as *prg-1* and *hrde-1* mutants, which are known to have a Mrt phenotype (Buckley et al., 2012; Simon et al., 2014).

As expected, wild type worms do not show a Mrt phenotype (after 44 generations). Nonetheless, few plates became sterile after 30 generations, possibly because they have been cultured for many years and the temperature of 25 °C is stressful. As expected, *prg-1* and *hrde-1* mutants show the Mrt phenotype, already after 6-8 generations. The double mutant *pid-4;pid-5* animals also start to become sterile after 8 generations, but the Mrt phenotype is fully established only after 24 generations, similar to *pid-2* mutants. Strikingly, also *pid-4* and *pid-5* single mutants show a Mrt phenotype, though it becomes evident only after 16 generations and it is established very slowly. Yet, we noticed that, already after 6 generations, *pid-4* and *pid-5* mutants produce much less progeny than wild type animals (**Figure 20**).

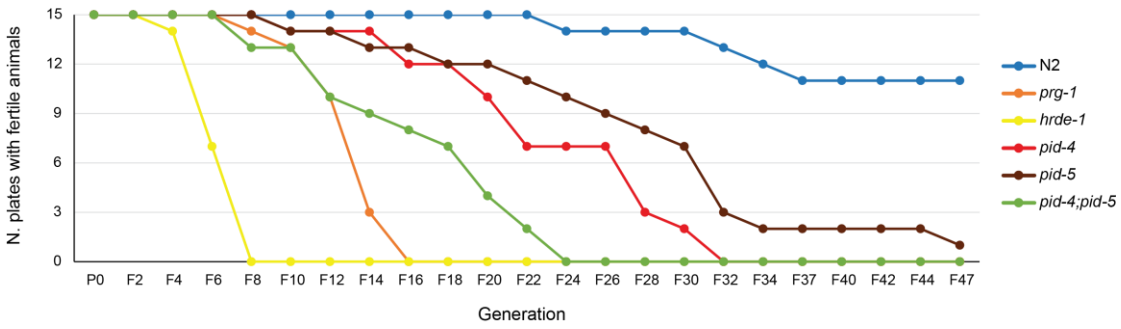


Figure 20. Mrt assay representing the number of fertile plates per generation. Wild type N2 worms grow steadily throughout the assay, although after 30 generations they are also affected by the stressful temperature of 25 °C. After 6 generations, *prg-1*, *hrde-1* and *pid-4;pid-5* mutants start to gradually become sterile. After 16-18 generations, also *pid-4* and *pid-5* single mutants start to show a Mrt phenotype.

We conclude that *pid-4* and *pid-5* mutants have a Mrt phenotype, even though not very severe. On the other hand, *pid-4;pid-5* double mutants have a Mrt phenotype comparable to *pid-2* mutants, yet weaker than *prg-1* and *hrde-1* mutants. Together these results indicate that PID-4 and PID-5 are partially redundant, as also the single mutant animals become sterile over time, and that they, like PID-2, are required for maintaining an immortal germline.

PID-4 and PID-5 interact with PID-2, but not with each other

We introduced a 3xMyc tag at the C-terminus of PID-4 (*xf186*) and a 2xHA tag within the PID-5 locus, also at the C-terminal end (*xf192*), using the CRISPR/Cas9 technology (**Figure 14A**, **B**) (Chen et al., 2013; Mouridi et al., 2017; Paix et al., 2014). We then used the epitopes to perform immunoprecipitation followed by quantitative label-free mass spectrometry (IP/MS) (**Figure 21A**, **B**). The results from these experiments are very clear: the only strong interactor of both PID-4 and PID-5 is PID-2. This indicates that their interaction with PID-2 is very specific and that the two proteins, PID-4 and PID-5, do not interact with each other, although they could be part of the same complex. Interestingly, PID-5 did bring down APP-1, which has been shown to dimerize (Iyer et al., 2015; Laurent et al., 2001), indicating that possibly PID-5:APP-1 heterodimers may exist, as well as PID-5 homodimers. Finally, we found PRMT-5 to be enriched in the pull down for PID-4 (**Figure 21A**); this is quite interesting, as PRMT-5 is the enzyme responsible for methylation of arginine or lysine residues of the partner proteins of Tudor domain proteins (Anne et al. 2007; Hirota et al., 2017; Siomi et al. 2010). We also observed PRMT-5 to be present in the interactome of PID-2::eGFP (**Figure 13A**).

Next, we used the polyclonal antibody αPID-2 to investigate how the interactome of PID-2 is affected upon *pid-4* or *pid-5* mutation. We repeated the IP/MS analysis both in a wild type background (not shown) and in *pid-2* mutants and compared them to *pid-4* and *pid-5* mutants (**Figure 21C**, **D**). Our analysis confirms that PID-2 interacts with both PID-4 and PID-5 and that the interaction is maintained in absence of either of these two proteins. When comparing *pid-4* and *pid-5* mutants with wild type, the only enriched protein in wild type is PID-4 or PID-5, respectively, indicating that the interactome of PID-2 is not affected in these mutants, except for loss of the interaction with the mutated protein itself (not shown).

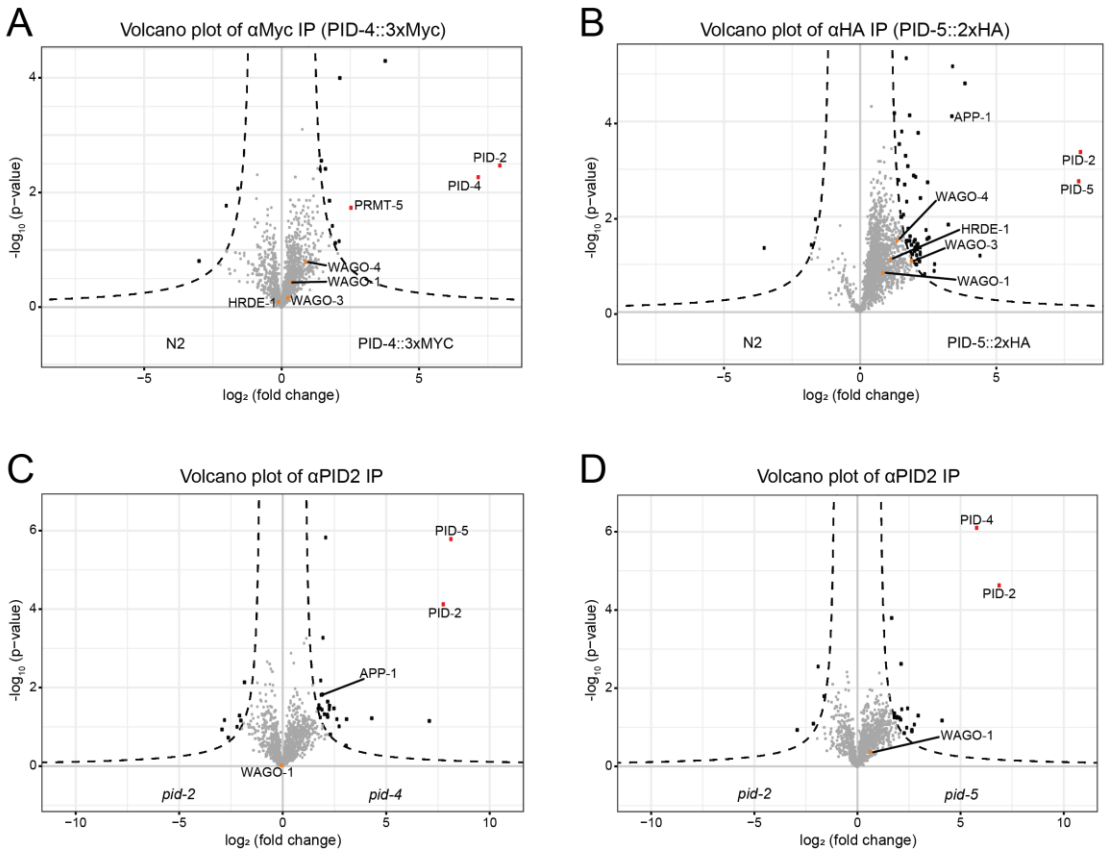


Figure 21. A) Volcano plot representing the enrichment of proteins that are interacting with PID-4. We immunoprecipitated the protein PID-4::3xMyc, which is highly enriched, and we could confirm the interaction with PID-2. As control, we used wild type protein extracts, not containing any Myc tagged protein. For each set of IP, we used 4 replicates for every strain. **B)** Volcano plot representing the enrichment of proteins that are interacting with PID-5. We immunoprecipitated the protein PID-5::2xHA, which is highly enriched, and we could confirm the interaction with PID-2. As control, we used wild type protein extracts, not containing any HA tagged protein. For each set of IP, we used 4 replicates for every strain. **C**, **D)** Volcano plot representing the enrichment of proteins that are interacting with PID-2 in mutant backgrounds in which either PID-4 or PID-5 have been deleted, respectively. We immunoprecipitated PID-2 using the polyclonal antibody against the endogenous protein. PID-2 is highly enriched in both mutants protein extracts, compared to *pid-2* mutant protein extracts. For each set of IP, we used 4 replicates for every strain.

Discussion and conclusion

To discover novel factors involved in the 21U RNA silencing pathway, we have isolated and characterized a novel gene from our previously published forward mutagenesis screen (de Albuquerque et al., 2014). This novel factor was named PID-2 (piRNA induced silencing defective-2). In turn, we used PID-2 to identify two additional proteins, PID-4 and PID-5, that act in a partially redundant manner in the 21U RNA pathway. Several aspects of our results will be discussed here.

PID-2, a mediator between the 21U RNA pathway and RNAe establishment?

Our data show that PID-2 has a role in the 21U RNA pathway. It appears to stimulate 22G RNA production, and not to affect 21U RNAs themselves. How could PID-2 act to stimulate 22G RNA production, or accumulation? One possibility is that PID-2 is required to stimulate the production of 22G RNAs by recruiting additional factors to the 21U RNA recognition site, such as an RdRP. Though we did not identify such interaction through the IP/MS experiment, it may simply be too transient to be detected. On the other hand, we did detect very mild enrichment for the Argonaute proteins WAGO-1, WAGO-3, WAGO-4 and HRDE-1, in the interactome of PID-2::eGFP (**Figure 13A**). Even though these interactors were below the set significance and at a level where many other, most likely non-specific, proteins were detected, they may hint to a role for PID-2 in recruiting these WAGO proteins, so that they can be loaded with 22G RNAs. Typically, small RNA co-factors are not stable without their Argonaute protein partner, possibly explaining the 22G RNA reduction, should WAGO-loading be affected by loss of PID-2. These interactions could indeed fit well within the P granules, where PID-2 as well as many other factors involved in mRNA surveillance localize.

Another observation on PID-2 activity was that initiation, but not maintenance, of RNAe is affected in *pid-2* mutants (**Figures 2**, **11**). Nonetheless also tertiary 22G RNAs, associated with RNAe, do go down, indicating that PID-2 does not specifically act on initiation; rather, more likely, initiation and maintenance of RNAe require different levels of 22G RNAs, being lower for the latter (**Figures 11**, **12**). Whether this links to a specific WAGO protein that needs to be loaded during initiation, potentially requiring higher 22G RNA levels, or to other, more downstream, effects that need to overcome a certain threshold, currently remains unclear. We already proposed the existence of a threshold level for establishing *de novo* silencing and RNAe, triggered by maternal 21U RNAs (see chapter 3). Potential thresholds would be, for example, the molecular link between small RNAs and heterochromatin formation, or the amount of mRNA that needs to be silenced, this being likely lower anyway during RNAe, due to H3K9 trimethylation of the genomic locus.

The Tudor proteins PID-4 and PID-5 are partially redundant, yet *pid-4;pid-5* double mutants do not fully phenocopy *pid-2* mutants

Immunoprecipitation of PID-2 revealed a complex formed by PID-2 and two Tudor domain proteins, namely W03G9.2 and Y45G5AM.2, which we named PID-4 and PID-5, respectively (**Figures 13**, **14**, **21**). The two Tudor proteins interact with PID-2 but not with each other, as shown by the reverse label-free quantitative mass spectrometry (**Figure 21**), and they could act redundantly, as suggested by our data on the 21U sensor silencing and from small RNA sequencing experiments (**Figures 16**, **17**, **18**).

The *pid-4;pid-5* double mutants phenocopy to some extent *pid-2* mutants, as the 21U sensor is expressed at similar levels in both strains (**Figure 19**) and the Mrt phenotype is comparable between them (**Figure 20**). Nevertheless, *pid-4;pid-5* double mutants show an important difference compared to *pid-2* mutants: heterozygous offspring of *pid-4;pid-5* double mutant hermaphrodites are proficient in establishing *de novo* silencing of the 21U sensor (**Figure 19**), whereas such offspring of *pid-2* mutant hermaphrodites show silencing defects (**Figure 3**). We currently do not understand this difference between *pid-2* and *pid-4;pid-5* mutants. Possibly, this could relate to the rate of zygotic production of PID-2 versus PID-4/-5, but it could also relate to the yet not well understood mutual relationship, and molecular functions of the proteins.

The Tudor domains: a platform for the assembly of a complex required to stimulate the production of 22G RNAs?

PID-4 and PID-5 have a Tudor domain (**Figure 14C**), which possibly serves as a platform to recruit RdRPs and/or other proteins to the target mRNA in order to promote the establishment of the silencing. Tudor domains generally recognize (symmetrically di-) methylated arginine or lysine residues of their partner protein, which is thought to be required to promote the formation of protein complexes within specific subcellular compartments (Pek et al., 2012). In fact, both PID-4 and PID-5 are specifically localized in perinuclear granules in the germline, where they could recruit factors required to promote RNAe establishment, such as PID-2 (**Figure 15**). Interestingly, we also co-immunoprecipitated, with PID-2::eGFP (**Figure 13A**) and with PID-4 (**Figure 21A**), PRMT-5, the enzyme responsible for methylating substrates, that are then recognized by the Tudor domain (Pek et al., 2012; Siomi et al., 2010). However, PID-2 has only two RG motifs and we could not detect the methylation of these arginine residues in

our experiments. On the other hand, Argonaute proteins have more RG sequences, especially HRDE-1 and PRG-1, as well as the RdRP RRF-1. Possibly, PRMT-5 methylates the arginine residues of such proteins, which are in turn recognized and bound by PID-4 and PID-5, forming a complex together with PID-2, that promotes the production of 22G RNAs as well as establishment of RNAe. For more insights into this, a hierarchy of PID-2 and PID-4/-5 localization would be very informative, as well as testing the effect of PRMT-5 loss on PID-2/-4/-5 localization.

The Xaa-Pro aminopeptidase domain of PID-5 could specifically interact with WAGO-3 and WAGO-4

Tudor domains are often associated with additional domains, that specify the function of the protein (Pek et al., 2012). PID-5 has an additional Xaa-Pro aminopeptidase domain, in which the catalytic residues, that would be required for the enzymatic activity, are lost (**Figure 14D**). Nevertheless, PID-5 could still use this domain to lock the interaction with such proteins, carrying a proline residue at position 2. For instance, WAGO-3 and WAGO-4 are such candidate proteins, which would fit with this idea. PID-5 could recruit these WAGO proteins to ensure the establishment of silencing on a 21U RNA target, whereas PID-4, and possibly the Tudor domain of PID-5, would additionally recruit other WAGOs or an RdRP, forming a complex required to stimulate 22G RNAs production and RNAe establishment. PID-2 would link such a complex to the silencing initiated by PRG-1, being just downstream of PRG-1 activity, but upstream of the silencing exerted by 22G RNAs and establishment of RNAe.

A very exciting hypothesis is that PID-5 can use its catalytically dead Xaa-Pro aminopeptidase domain to dimerize with APP-1, which would then bring the enzymatic activity missing from PID-5. It is known, in fact, that APP-1 forms homodimers (Iyer et al., 2015; Laurent et al., 2001) and it could also possibly form heterodimers with PID-5.

The Mrt phenotype indicates that the complex PID-2/PID-4/PID-5 is required for germline integrity

The Mrt phenotype is associated with loss of PRG-1, and consequently of the 21U RNA pathway (Simon et al., 2014), or with loss of factors that are involved in the RNAe machinery (Buckley et al., 2012; Spracklin et al., 2017; Xiao et al., 2011). The end

result of such defects is sterility, caused by loss of germ cells over generations. We observed the Mrt phenotype also for *pid-2* (**Figure 6**) and *pid-4;pid-5* mutants (**Figure 20**). In a milder way, also *pid-4* and *pid-5* mutants lose germ cells over time (**Figure 20**), yet this result supports our idea of PID-4 and PID-5 being partially redundant. There could be many explanations for this complicated phenotype. However, given the reduction of global 22G RNA levels in *pid-2*, and possibly in *pid-4;pid-5* double mutants, RNAe may not be established properly on endogenous loci, whereas some other loci could be erroneously silenced, and the combination of these events eventually results in sterility. Loss of RNAe is associated with loss of heterochromatin at endogenous loci, which should be maintained silenced. Though we did not test it yet, it is possible that H3K9 trimethylation is also gradually lost from heterochromatic loci in *pid-2* and *pid-4;pid-5* mutants, which could provoke erroneous gene expression.

Further experiments need to be performed in order to elucidate the function of this novel complex and how it contributes to the maintenance of a healthy germline over time.

Our study adds more factors in the complicated picture of the silencing machinery in *C. elegans*. We identified the first Tudor proteins to play a role in the 21U RNA pathway in *C. elegans*, in analogy to other organisms (Pek et al., 2012). More importantly, we identified the first factors that are essential for RNAe initiation, but not maintenance, indicating that these two steps of RNAe are different and mediated by different factors. Possibly, this specific role is related to the restricted expression of this novel complex at the pachytene stage (**Figures 15**). Previous studies, so far, have not taken expression specificity during germ cell development into account. In some specific cases, such as our study, it may be worthwhile to carefully investigate different developmental stages to gain more insight into the function of a particular factor or pathway (see chapter 5).

Methods

Strains maintenance

Worm strains have been grown according to standard laboratory conditions on NGM plates seeded with *Escherichia coli* OP50 and grown at 20 °C, unless otherwise stated (Brenner, 1974). We used the N2 Bristol strain as wild type strain.

List of strains

Strain	Genotype
	wild type N2
RFK507	*pid-2(xf23) I*
	pid-2 (tm1614) I
RFK231	*mjSi22 [Pmex-5::mCherry::his-58::21UR-1_as::tbb-2(3'UTR)] I; mut-7(pk204) III; otls45 [Punc119::GFP] V*
RFK316	*mjSi22 [Pmex-5::mCherry::his-58::21UR-1_as::tbb-2(3'UTR)] (RNAe) I; prg-1(n4357) I*
RFK851	*mjSi22 [Pmex-5::mCherry::his-58::21UR-1_as::tbb-2(3'UTR)] I; prg-1(n4357) I*
SX2078	*mjSi22 [Pmex-5::mCherry::his-58::21UR-1_as::tbb-2(3'UTR)] I*
RFK416	*mjSi22 [Pmex-5::mCherry::his-58::21UR-1_as::tbb-2(3'UTR)] I; hrde-1(tm1200) III*
RFK677	*pid-2(xf23) I; mjSi22 [Pmex-5::mCherry::his-58::21UR-1_as::tbb-2(3'UTR)] I*
RFK585	*pid-2(xf23) I; mjSi22 [Pmex-5::mCherry::his-58::21UR-1_as::tbb-2(3'UTR)] (RNAe) I*
RFK528	*pid-2(tm1614) I; mjSi22 [Pmex-5::mCherry::his-58::21UR-1_as::tbb-2(3'UTR)] I*
RFK586	*pid-2(tm1614); mjSi22 [Pmex-5::mCherry::his-58::21UR-1_as::tbb-2(3'UTR)] (RNAe) I*
SX922	*prg-1(n4357) I*
RFK621, RFK622	wild type (new isolate from outcrossing *pid-2(xf23) I; prg-1(n4357) I; unc22::Tc1 (st136::Tc1) IV*)
RFK613	*pid-2(xf23) I; prg-1(n4357) I*
RFK618	*pid-2(tm1614) I; prg-1(n4357) I*
RFK615	*pid-2(xf23) I; hrde-1(tm1200) III*
RFK620	*pid-2(tm1614) I; hrde-1(tm1200) III*
RFK709	*pid-2(xf23) I*
NL3643	*unc-22(st136::Tc1) IV*
RFK611	*wago-1(tm1414) I; wago-2(tm2686) I; ppw-2(tm1120) I; unc-22::Tc1(st136::Tc1) IV*
RFK610	*pid-2(xf23) I; unc-22::Tc1 (st136::Tc1) IV*
RFK612	*pid-2(xf23) I; prg-1(n4357) I; unc22::Tc1 (st136::Tc1) IV*
RFK614	*pid-2(xf23) I; hrde-1(tm1200) III; unc22::Tc1 (st136::Tc1) IV*
RFK616	*pid-2(tm1614) I; unc22::Tc1 (st136::Tc1) IV*
RFK617	*pid-2(tm1614) I; prg-1(n4357) I; unc22::Tc1 (st136::Tc1) IV*

RFK619	*pid-2(tm1614) I; hrde-1(tm1200) III; unc22::Tc1 (st136::Tc1) IV*
RFK240	*mut-16(pk710) I*
RFK257	*mjSi22[Pmex-5::mCherry::his-58::21UR-1_as::tbb-2(3'UTR)] I; pid-1(xf35) II; mut-7(pk204) III; otIs45 [unc119::GFP] V*
RFK232	*mut-16(pk710) I; prg-1(n4357) I*
RFK503	*pid-2(xf23) I; mut-7(pk204) III; otIs45 [unc-119::GFP] V*
RFK506	*pid-2(xf23) I; mut-16(pk710) I*
RFK707	*pid-2(xf23) I prg-1(n4357) I; mut-7(pk204) III; otIs45 [unc-119::GFP] V*
RFK708	*pid-2(xf23) I; mut-16(pk710) I; prg-1(n4357) I*
RFK804	*pid-2(xf23) I; mjSi22 [Pmex-5::mCherry::his-58::21UR-1_as::tbb-2(3'UTR)] I; pid-1(xf35) II; otIs45 [unc-119::GFP] V*
RFK774	*pid-2(xf23) I; pid-1(xf35) II*
RFK805	*pid-2(tm1614) I; mjSi22 [Pmex-5::mCherry::his-58::21UR-1_as::tbb-2(3'UTR)] I; pid-1(xf35) II; otIs45 [unc-119::GFP] V*
RFK776	*pid-2(tm1614) I; pid-1(xf35) II*
RFK978	*prg-1(n4357) I* (4x outcrossed)
RFK979	*hrde-1(tm1200) III* (4x outcrossed)
RFK980	*pid-2(xf23) I* (4x outcrossed)
RFK981	*pid-2 (tm1614) I* (4x outcrossed)
EG6699	*ttTi5605 II; unc-119(ed3) III; oxEx1578*
RFK654	*xfIs144 [pid-2(5'UTR)::eGFP::pid-2::pid-2(3'UTR); cb-unc119(+)] II*
RFK655	*xfIs145 [pid-2(5'UTR)::pid-2::eGFP::pid-2(3'UTR); cb-unc119(+)] II*
RFK656	*xfls146 [pid-2(5'UTR)::3xFLAG::pid-2::pid-2(3'UTR); cb-unc119(+)] II*
RFK693	*pid-2(xf23) I; xfls144 [pid-2(5'UTR)::eGFP::pid-2::pid-2(3'UTR); cb-unc119(+)] II*
RFK694	*pid-2(xf23) I; xfls145 [pid-2(5'UTR)::pid-2::eGFP::pid-2(3'UTR); cb-unc119(+)] II*
RFK695	*pid-2(xf23) I; xfls146 [pid-2(5'UTR)::3xFLAG::pid-2::pid-2(3'UTR); cb-unc119(+)] II*
RFK853	*mjSi22 [Pmex-5::mCherry::his-58::21UR-1_as::tbb-2(3'UTR)] I pid-2(xf23) I; xfls144 [pid-2(5'UTR)::eGFP::pid-2::pid-2(3'UTR); cb-unc119(+)] II*
RFK854	*mjSi22 [Pmex-5::mCherry::his-58::21UR-1_as::tbb-2(3'UTR)] I pid-2(xf23) I; xfIs145 [pid-2(5'UTR)::pid-2::eGFP::pid-2(3'UTR); cb-unc119(+)] II*
RFK855	*mjSi22 [Pmex-5::mCherry::his-58::21UR-1_as::tbb-2(3'UTR)] I pid-2(xf23) I; xfls146 [pid-2(5'UTR)::3xFLAG::pid-2::pid-2(3'UTR); cb-unc119(+)] II*
RFK857	*pid-4 (xf184) I*
RFK858	*pid-4 (xf185) I*
RFK859	*pid-5 (xf181) V*
RFK860	*pid-5 (xf182) V*
RFK982	*pid-4 (xf184) I; mjSi22 [Pmex-5::mCherry::his-58::21UR-1_as::tbb-2(3'UTR)] I*
RFK983	*pid-4 (xf184) I; mjSi22 [Pmex-5::mCherry::his-58::21UR-1_as::tbb-2(3'UTR)] (RNAe) I*
RFK984	*mjSi22 [Pmex-5::mCherry::his-58::21UR-1_as::tbb-2(3'UTR)] I; pid-5 (xf181) V*

RFK985	*mjSi22 [Pmex-5::mCherry::his-58::21UR-1_as::tbb-2(3'UTR)] I; pid-5 (xf181) (RNAe) V*
RFK986	*pid-4 (xf184) I; pid-5 (xf181) V*
RFK987	*pid-4 (xf184) I; mjSi22 [Pmex-5::mCherry::his-58::21UR-1_as::tbb-2(3'UTR)] I; pid-5 (xf181) V*
RFK847	*pid-4 (xf186[pid-4::3xMyc]) I*
RFK879	*pid-5 (xf192[pid-5::2xHA]) V*
RFK932	*pid-4 (xf204[pid-4::d10]) I*
RFK988	*pid-4 (xf206[pid-4::mTagRFP-T]) I*
RFK972	*pid-5 (xf221[pid-5::d10]) V*
RFK1025	*pid-5 (xf226[pid-5::mTagRFP-T]) V*
RFK1029	*pid-4 (xf206[pid-4::mTagRFP-T]) I; xfIs145[pid-2(5'UTR)::pid-2::eGFP::pid-2(3'UTR); cb-unc119(+)] II; pgl-1(xf222[pgl-1::mTurquoise2]) IV*

Single worm lysis and genotyping

After they have laid embryos, single adult worms were picked in 5 µl of single worm lysis buffer (25 mM KCl; 12,5 mM $MgCl_2$; 5 mM Tris pH=8.3; 2,25% NP40; 2,25% Tween-20; 0.05% gelatin; 0,4 mg/ml Proteinase K). Lysis was performed for 1 hour at 65 °C, followed by incubation for 15 minutes at 95 °C to inactivate the Proteinase K. Lysate was then diluted by adding 20 µl of nuclease-free water and 2 µl were used as template for the PCR, that was performed according to the manufacturer instructions (*Taq* DNA Polymerase, New England Biolabs, M0273L).

List of primers:

pid-2:

*wt*_FW	gtaccgcgatcacgatgtgt
*xf23*_FW	gtaccgcgatcacgatttat
*wt/xf23*_RV	aatctgagcactctcgctgg
*tm1614*_FW	ccgtggagtacacgacaatg
*tm1614*_RV	cagaccgtctccgatgattt
*ttTi5605*_wt_FW	aggcagaatgtgaacaagactcg
*ttTi5605*_RV	atcgggaggcgaacctaactg
*xfIs144/xfIs146*_FW	cgaagcacttatcgccggac
*xfIs145*_FW	acagctgctgggattacaca

pid-4:

*wt/xf184/xf185*_FW	ccacaccgcccataaatgtc

*wt/xf184/xf185*_RV	acggaccaacgggtcttaac
*wt*_FW2	caacccggactcgatttgta
*wt/xf186/xf206*_FW	aagccattcgtggacagatc
*wt/xf186/xf206*_RV	tcggtgcagagattcaagcc
*xf206*_RV2	ggattccttgggtgtggttg

pid-5:

*wt/xf181/xf182*_FW1	actgcgagaccattgcctac
*wt/xf181*_RV1	cgctggccacttgtgata
*wt*_FW2	ctacgagacggtcgatgtca
*xf182*_RV2	gctgtttctcctgcttgacc
*wt/xf192*_FW	gactggctcttcaatgcatg
*wt/xf192*_RV	ttgaaattcccttcccaccc

Microscopy

20-25 worms have been picked to a drop of M9 (80 µl) on a slide, washed and then fixed with acetone (2 x 80 µl). After acetone has evaporated, worms have been washed 2 x 10 minutes with 80 µl of PBS-Triton X100 0,1%. After removing the excess of PBS-Triton X100 0,1%, the worms have been mounted on a coverslip with Fluoroshield™ with DAPI (5 µl) (Art. No. F6057, Sigma).

Alternatively, for live imaging, 20-25 worms have been picked to a drop of M9 (80 µl) on a slide, washed and then 2 µl of 1 M NaN_3 have been added for 10 minutes to paralyze the worms. After removing the excess of M9, a slide prepared with 2% agarose (in water) has been placed on top of the coverslip and worms have been imaged directly.

Images in **Figures 1C**, **2B**, **3** (**B**, **D**, **F**, **H**), **4C**, **17**, **19** (**B**, **D**, **F**, **G**) have been acquired at a Leica DM6000B microscope; images in **Figure 4D** have been acquired at a Leica M165FC widefield microscope with a Leica DFC450C camera; images in **Figure 15** have been acquired at a Leica TCS SP5 confocal microscope. Images have then been processed with Leica LAS software and ImageJ.

For scoring the 21U sensor as active or silenced (**Figures 1**, **2**, **3**, **4**, **17**, **19**), we have used a Leica M165FC widefield microscope. The 21U sensor has been scored as: active, if the fluorescence was easily visible with a lower magnification (Plan APO 1.0x, Art. No. 10450028; Leica); faint, if the fluorescence was only visible with a higher

magnification (Plan APO 5.0x/0.50 LWD, Art. No. 10447243; Leica); silenced, if no fluorescence was visible. The worms have been later used also for live imaging with a Leica DM6000B microscope as described above.

Small RNA sequencing

RNA extraction

Synchronized gravid adults have been collected with M9 and fast frozen on dry ice in 250 µl of Worm Lysis Buffer (200 mM NaCl; 100 mM Tris HCl pH=8.5; 50 mM EDTA pH=8; 0,5% SDS). 30 µl of Proteinase K (20 mg/ml; Art. No. 7528.1, Carl Roth) have been added to dissolve the worms for 90 minutes at 65 °C with gentle shaking. Lysate has been centrifuged at maximum speed for 5 minutes at room temperature (RT) and the supernatant was transferred on a Phase Lock Gel tube (Art.No. 2302830, QuantaBio). 750 µl TRIzol LS (Art. No. 10296028, Invitrogen™) have been added per 250 µl of sample and, after homogenization, the samples have been incubated for 5 minutes at RT to allow complete dissociation of the nucleoprotein complex. Then 300 µl of chloroform (Art. No. 288306, Sigma-Aldrich) were added per 750 µl of TRIzol LS and the samples were incubated for 15 minutes at room temperature after mixing. Samples have been centrifuged at 12000 x g for 5 minutes at RT and another round of chloroform extraction has been performed. The aqueous phase has been then transferred to an Eppendorf tube and 500 µl of cold isopropanol was added to precipitate the RNA; samples have been mixed vigorously, incubated at RT for 10 minutes and spun down at maximum speed for at least 10 minutes at 4 °C. The pellet was then washed twice with 1 ml of 75% ethanol and centrifuged for 5 minutes at 7500 x g at 4 °C. The pellet has been dried and diluted in 50 µl of nuclease-free water with gentle shaking for 10 minutes at 42 °C. In order to remove any contamination of genomic DNA, 5 µl of 10X TURBO™ DNase Buffer and 1 µl of TURBO™ DNase (Art. No. AM2238, Invitrogen™) were added to the RNA and incubated at 37 °C for 30 minutes with gentle shaking. The reaction has been stopped by adding 5 µl of 10X TURBO™ DNase Inactivation Reagent. Samples have been centrifuged at 10000 x g for 90 seconds and RNA transferred to a fresh tube. RNA quality has been assessed at Nanodrop and on agarose gel and then samples have been further processed for enrichment of small RNA populations.

Small RNAs enrichment

In order to enrich for small RNAs, we used the *mir*Vana™ kit (Art. No. AM1561, Invitrogen™). 400 µl of *mir*Vana™ Lysis/Binding buffer and 48 µl of *mir*Vana™ Homogenate Additive have been added to the total RNA (80 µl). The mix has been incubated at RT for 5 minutes to denature RNA, then 1/3 of volume of 100% ethanol has been added and after mixing, samples have been spun down at 2500 x g for 4 minutes at RT to pellet large RNAs (>200 nt). The supernatant has been transferred to a new Eppendorf tube and RNA has been precipitated at -80 °C for 1 h with isopropanol (1:1). Samples have been centrifuged at maximum speed for at least 10 minutes at 4 °C to pellet small RNAs. The pellet has then been washed twice with 75% ethanol and spun down at maximum speed for 5 minutes at 4 °C. Pellet has been dried and resuspended in 16 µl of nuclease-free water. RNA quality has been checked at Nanodrop and on agarose gel and further processed for library preparation and deep sequencing.

Library preparation and sequencing

For each strain, three biological replicates have been used for RNA extraction and library preparation. RNA was treated with RppH (RNA 5' Pyrophosphohydrolase, Art. No. M0356S, New England Biolabs) to dephosphorylate small RNAs and specifically enrich for 22G RNAs, as previously described (Almeida et al., 2019). For each sample, 1 µg of RNA was incubated for 1 hour at 37 °C with 5 units of RppH and 10X NEB Buffer 2. After dephosphorylation, 500 mM EDTA was added and samples were incubated for 5 minutes at 65°C to stop the RppH treatment and RNA was purified with sodium chloride/isopropanol precipitation. After purification, the recovered total RNA was used for library preparation.

NGS library preparation and sequencing details are listed below.

Project ID	Kit	Starting RNA	PCR amplification	Flowcell	Sequencing
LIMS-861: imb_ketting_2017_10_placentino_pid-2 smRNA	NEXTFLEX® Small RNA-Seq Kit v3 for Illumina® Platforms (BIOO Scientific)	570-780ng	16 cycles	NextSeq 500 High Output	single read for 1x 85bp (85 cycles for Read 1 plus 7 cycles for the index read)
LIMS-1026: imb_ketting_2018_	NEXTFLEX® Small RNA-Seq Kit v3	1000ng	14 cycles	NextSeq 500 High Output	single read for 1x 85bp (85 cycles

15_placentino_ smRNA	(Step A to Step G; BIOO Scientific v16.06)				for Read 1 plus 7 cycles for the index read)
LIMS-1066: imb_ketting_2018_ 22_placentino_ smRNA	NEXTFLEX® Small RNA-Seq Kit v3 (Step A to Step G; BIOO Scientific v16.06)	640ng	15 cycles	NextSeq 500 Mid Output	single read for 1x 51bp (51 cycles for Read 1 plus 7 cycles for the index read)

Amplified libraries were purified by running an 8% TBE gel and size-selected for the 146-168bp fraction. Libraries were profiled in a High Sensitivity DNA on a 2100 Bioanalyzer (Agilent technologies) and quantified using the Qubit dsDNA HS Assay Kit, in a Qubit 2.0 Fluorometer (Life technologies).

Read procession and mapping

Before mapping to the genome, reads were processed in the following manner: (i) trimming of sequencing adapters with cutadapt v1.9 (-a TGGAATTCTCGGGTGCCAAGG -O 5 -m 26 -M 38) (Martin, 2011); (ii) removal of reads with low-quality calls with the FASTX-Toolkit v0.0.14 (fastq_quality_filter -q 20 -p 100 -Q 33); (iii) collapsing of PCR duplicates (custom bash script), making use of the unique molecule identifiers (UMIs) added during library preparation; (iv) trimming of UMIs with seqtk v1.2 (trimfq -b 4 -e 4); and (v) removal of very short sequences with seqtk v1.2 (seq -L 15). Read quality was assessed before and after these processing steps with FastQC v0.11.5 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc).

Reads that passed the above filtering steps were mapped to a custom *C. elegans* genome (WBcel235) to which the 21U sensor sequence (Bagijn et al., 2012) was added an extra contig. The mapping was done with bowtie v0.12.8 (-q –sam –phred33-quals –tryhard –best –strata –chunkmbs 256 -v 0 -M 1) (Langmead et al., 2009). To generate genome browser tracks we used a combination of Bedtools v2.25.0 (genomeCoverageBed -bg -split -scale -ibam -g) (Quinlan & Hall, 2010), to summarize genome coverage normalized to mapped non-structural reads (rRNA/tRNA/snoRNA/snRNA) * 1 million (RPM, Reads Per Million), and bedGraphToBigWig to finally create the bigwig tracks.

Small RNA classification and quantification

Gene annotation was retrieved from Ensembl (release-38) and merged with transposon coordinates, retrieved from Wormbase (PRJNA13758.WS264), creating a custom annotation used for the analysis. Mapped reads were categorized in small RNA classes as follows: 21U RNAs are 21 nt long sequences mapping sense to annotated 21U RNA loci; 22G RNAs are 20-23 nt long and map antisense to protein-coding/pseudogenes/lincRNA/transposons; 26G RNAs, are those which are 26 nt long, and map antisense to annotated protein-coding/pseudogenes/lincRNA; and miRNAs are 20-24 nt long mapping sense to annotated miRNA loci. Read filtering was done with a python script (https://github.com/adomingues/filterReads/blob/master/filterReads/filterSmallRNAclasses.py) based on pysam v0.8.1 / htslib (Li et al., 2009), in combination with Bedtools intersect. Reads belonging to each class were then counted for each library (total levels). For some analysis the number of small RNAs in certain sub-classes was summarized based on previously published lists: ALG-3/-4 (Almeida et al., 2019); ERGO-1 (Almeida et al., 2019); CSR-1 (Conine et al., 2013); NRDE-3 (Zhou et al., 2014); mutators (Phillips et al., 2014); and WAGO-1 (Gu et al., 2009). The genomic locations of 22G and 26G RNAs was then intersected with that of the genes, and counted for each library.

22G RNAs coverage on 21U sensor

For targeting of the 21U sensor by 22G RNAs, we considered only sequences that were 22 nt long mapping unambiguously to the 21U sensor sequence. Coverage was calculated with Bedtools v2.25.0 (genomeCoverageBed -ibam - -d) (Quinlan & Hall, 2010). Visualization was generated with the R/Bioconductor package ggbio (Lawrence & Morgan, 2014) [Yinggbiopackageextending2012].

Secondary and tertiary 22G RNAs on 21U sensor

The tertiary populations were defined as the 22G RNAs mapping antisense to the mCherry coding sequence, within the 21U sensor (Sapetschnig et al., 2015). To define the secondary 22G (Sapetschnig et al., 2015), those surrounding the *21ur-1* recognition site, the spread from this site was visually estimated on a genome browser (IGV) to be +/-200 bp. The reads mapping to these positions were counted and normalized to the number of uniquely mapping reads.

Coverage of local 22G RNAs on endogenous 21U RNA target sites

Following the analysis of (Lee et al., 2012), 21U RNA targets were identified by mapping the annotated 21U RNA sequences (WBCel235) to the genome with bowtie v1.2.1.1 with relatively flexible mapping (-n 1 -l 8 -e 300 -k 1000 –best), allowing for one mismatch in the 8 nucleotide seed and reporting up to 1000 valid alignments. The alignments were then filtered allowing for one T to G change in the seed, and two additional mismatches and one T to G change in the remaining alignment (see also (Bagijn et al., 2012)). These putative 21U RNA target coordinates were then intersected with the genomic location of WAGO-1 associated genes, to define a confident set of 21U RNA binding sites. Finally the 5' 21U RNA mapping site was shifted by 10 nt (bedtools shift -s 10), extended in both directions by 50 bp (slopBed -l 50 -r 50 -s), and finally the 22G RNA coverage was calculated (coverageBed -d -s) and normalized by the number of non-structural reads.

Sequencing statistics

Project ID	Strain	Sample ID	Sequenced reads	Mapped reads	Non-structural reads
LIMS-1026	SX2078	01_WT_sensorOFF_rep1_RppH	20495279	8970544	4698349
	SX2078	02_WT_sensorOFF_rep2_RppH	23815281	10530815	5168582
	SX2078	03_WT_sensorOFF_rep3_RppH	21670321	9332782	4490019
	RFK851	04_prg1_sensorON_rep1_RppH	23370846	9027741	3574552
	RFK851	05_prg1_sensorON_rep2_RppH	22830836	8503804	3135787
	RFK851	06_prg1_sensorON_rep3_RppH	20830943	7589448	2950844
LIMS-861	RFK621	01_WT_new_rep1_RppH	7890694	3957830	1776071
	RFK622	02_WT_new_rep2_RppH	13643668	6666541	3100674
	RFK622	03_WT_new_rep3_RppH	11228507	5870012	2979931
	RFK613	04_prg1_xf23_rep1_RppH	8513955	3389388	1349462
	RFK613	05_prg1_xf23_rep2_RppH	9716616	3596419	1396926
	RFK613	06_prg1_xf23_rep3_RppH	11397301	4569504	1728399
	RFK618	07_prg1_tm1614_rep1_RppH	7188058	2885907	1112652
	RFK618	08_prg1_tm1614_rep2_RppH	9692133	4031090	1727044
	RFK618	09_prg1_tm1614_rep3_RppH	8704135	4135054	1908477
	RFK615	10_hrde1_xf23_rep1_RppH	12434220	5190638	1965672
	RFK615	11_hrde1_xf23_rep2_RppH	10146233	4303389	1665587
	RFK615	12_hrde1_xf23_rep3_RppH	8498682	4113233	1827768
	RFK620	13_hrde1_tm1614_rep1_RppH	9544567	4580768	2151261
	RFK620	14_hrde1_tm1614_rep2_RppH	11589202	5626960	2667058
	RFK620	15_hrde1_tm1614_rep3_RppH	9330338	4751781	2525439
	N2	16_WT_old_rep1_RppH	9304276	4520823	1980662

	N2	17_WT_old_rep2_RppH	11634674	4634869	1640288
	N2	18_WT_old_rep3_RppH	8506375	4114787	1898297
	RFK416	19_hrde1_sensorON_rep1_RppH	8871279	3982792	1743751
	RFK416	20_hrde1_sensorON_rep2_RppH	8624676	3668484	1503875
	RFK416	21_hrde1_sensorON_rep3_RppH	7438750	3041676	1212914
	RFK709	22_xf23_new_rep1_RppH	8470889	3603405	1621307
	RFK709	23_xf23_new_rep2_RppH	9201649	4063242	1859684
	RFK709	24_xf23_new_rep3_RppH	8119819	3361799	1429503
	RFK677	25_xf23_sensorON_rep1_RppH	12463441	6204334	3545332
	RFK677	26_xf23_sensorON_rep2_RppH	9259083	4334493	2232378
	RFK677	27_xf23_sensorON_rep3_RppH	9689994	4596796	2419712
	RFK585	28_xf23_sensorOFF_rep1_RppH	9171771	4281527	2102983
	RFK585	29_xf23_sensorOFF_rep2_RppH	9111576	4401986	2109144
	RFK585	30_xf23_sensorOFF_rep3_RppH	9143831	4432833	2229756
	RFK528	31_tm1614_sensorON_rep1_RppH	9076058	4075760	1530474
	RFK528	32_tm1614_sensorON_rep2_RppH	9116519	4564336	2274887
	RFK528	33_tm1614_sensorON_rep3_RppH	9574243	4639933	2347228
	RFK586	34_tm1614_sensorOFF_rep1_RppH	7532106	3643866	1842216
	RFK586	35_tm1614_sensorOFF_rep2_RppH	8926929	4248914	2055841
	RFK586	36_tm1614_sensorOFF_rep3_RppH	7979048	3359705	1643397
	RFK231	37_mut7_sensorON_rep1_RppH	9307687	3299101	992818
	RFK231	38_mut7_sensorON_rep2_RppH	4982141	1558700	360911
	RFK231	39_mut7_sensorON_rep3_RppH	8427277	3019357	847422
	RFK316	40_prg1_sensorOFF_rep1_RppH	10668746	4123902	1587452
	RFK316	41_prg1_sensorOFF_rep2_RppH	9054859	3579894	1299603
	RFK316	42_prg1_sensorOFF_rep3_RppH	7790328	3096786	1069152
LIMS-1066	RFK984	01_Y45_xf181_sensorON_rep1_RppH	11640357	4621156	2001885
	RFK984	02_Y45_xf181_sensorON_rep2_RppH	12880188	4919411	2105695
	RFK984	03_Y45_xf181_sensorON_rep3_RppH	14840342	5752691	2544801
	RFK985	04_Y45_xf181_sensorOFF_rep1_RppH	13589568	5631852	2665844
	RFK985	05_Y45_xf181_sensorOFF_rep2_RppH	12221489	4923431	2171374
	RFK985	06_Y45_xf181_sensorOFF_rep3_RppH	13319756	5506522	2571844
	RFK982	07_W03_xf184_sensorON_rep1_RppH	12668231	5247465	2438664
	RFK982	08_W03_xf184_sensorON_rep2_RppH	11780576	4386659	2099021
	RFK982	09_W03_xf184_sensorON_rep3_RppH	9922450	3583457	1572692
	RFK983	10_W03_xf184_sensorOFF_rep1_RppH	11670225	4387000	1933200
	RFK983	11_W03_xf184_sensorOFF_rep2_RppH	12911219	5261012	2456414

	RFK983	12_W03_xf184_sensorOFF_rep3_RppH	11148585	4622215	2104743

Transposon excision analysis

For each analysed genotype, mutant worms carrying the *unc-22::Tc1(st136)* insertion were singled into a 6 cm^2 NGM plate seeded with 100 µl of OP50 and grown at 20 °C. For each genotype, 50 worms have been singled out. Plates were scored for wild type moving worms at three different time points: when the total number of worms per plate was ~50, when the total number of worms per plate was ~100, and when the plate was starved, to which we estimated the total number of worms per plate to be ~1000. Transposition frequencies at each time point were calculated using the following formula: f = -ln [(T - R) / T] / N, where T = total number of plates scored, R = number of plates with revertants, and N = number of worms on the plate. Each time point was considered as a biological replicate. The graph represents the average of two experiments.

Mortal germline assay

Before starting the experiment, mutants have been outcrossed four times. For the assay, N2 have been used as wild type strain and the desired mutant strains have been tested. For each strain, 6 L3 worms have been picked onto 15 NGM plates (10 cm^2) seeded with 300 µl OP50, grown at 25 °C and followed over time. Worms have been picked every 4-6 days, before starvation, and we assumed 2-3 generations, respectively, have passed. Worms have been passed to fresh plates every 4-6 days until all the mutants died.

Production of PID-2 protein for antibody generation

Cloning

In order to produce a polyclonal antibody against PID-2, we cloned the full-length coding sequence (CDS) of *pid-2* (UniProtID Q9N3P1) into a vector for recombinant protein overexpression in *E. coli*.

The cDNA, obtained from wild type worms with the ProtoScript® First Strand cDNA Synthesis Kit (Art. No. E6300S, New England Biolabs), has been used as template for the amplification of the CDS of *pid-2* using Q5® Hot Start High-Fidelity DNA

Polymerase (Art. No. M0493, New England Biolabs). PCR primers were designed to contain restriction sites (forward: NcoI ggccatgggcatgacagttattatagcgtcacact; reverse (without stop codon): XhoI ggctcgagaaatggg cactcgctgaat) for subsequent cloning into the pET-28a(+) expression vector (Art. No. 69864-3, Merck), which confers bacterial resistance to Kanamycin and adds a 6x histidine tag for affinity purification, either at the N- or at the C-terminus of the protein of interest. The PCR product was purified from agarose gel (QIAquick Gel Extraction Kit, Art. No. 28706, QIAgen). The vector and the PCR product (insert) have been digested with the restriction enzymes NcoI (Art. No. R3193, New England Biolabs) and XhoI (Art. No. R0146, New England Biolabs) for 2 hours at 37 °C. The insert has been additionally dephosphorylated using Antarctic Phosphatase (Art. No. M0289, New England Biolabs) for 15 minutes at 37 °C to avoid self-ligation. Ligation was then performed for 30 minutes at room temperature using a molar ratio 3:1 = vector:insert using T4 DNA Ligase (Art. No. M0202, New England Biolabs). The ligation reaction was then transformed in Subcloning Efficiency™ DH5α™ Competent Cells (Art. No. 18265017, Invitrogen™) and plated for selection on LB agar plates with Kanamycin (30 µg/ml). The plasmid was isolated from the bacterial culture (PureLink™ HiPure Plasmid Miniprep Kit, Art. No. K210011, Invitrogen™), checked by enzymatic digestion and sequencing.

Protein expression and purification

For protein expression, the pET-28a(+)-PID-2 construct has been transformed into Rosetta™ (DE3) Competent Cells (Art. No. 70954-3, Merck) and positive clones have been selected on LB agar plates with Kanamycin (30 µg/ml). A single colony has been inoculated into 20 ml of LB media supplied with 100 µg/ml Kanamycin and 35 µg/ml Chloramphenicol as pre-culture and grown overnight at 37 °C. The pre-culture has then been inoculated in 1 l LB media supplied with 100 µg/ml Kanamycin and 35 µg/ml Chloramphenicol and grown for 3 hours at 37 °C, until exponential phase (OD600 = 0,6). Protein expression has been induced by adding 0,5 mM IPTG (Art. No. V3953, Promega) and incubated overnight at 18 °C. The bacterial culture has been harvested by spinning down at 4000 x g for 15 minutes at 4 °C. Cells pellet has been collected and washed with HEPES buffer to remove excess of growth media. The bacteria were spun down again at 4500 x g for 20 minutes at 4 °C and the pellet was frozen at -20 °C. In order to purify PID-2 protein tagged with 6xHis-tag from inclusion bodies, bacterial pellet was thawn on ice and resuspended in 30 ml of lysis buffer (500 mM NaCl; 100 mM Tris HCl pH=8.5). The pellet has been homogenized, sonicated 3 x 3 minutes

(Branson Sonifier 450; output 4-5; duty cycle 2-3) and the cell lysate has been centrifuged at 19000 x g for 25 minutes at 4 °C. The pellet containing PID-2 has then been resuspended in 30 ml of denaturing buffer (500 mM NaCl; 100 mM Tris HCl pH=8.5; 8 M urea) to solubilize the inclusion bodies and centrifuged again at 19000 x g for 25 minutes at 4 °C to remove cells debris. The supernatant has then been loaded on a batch column containing 1 ml of Ni-NTA Agarose slurry (Art. No. 30210, QIAgen), previously equilibrated with the same buffer (500 mM NaCl; 100 mM Tris HCl pH=8.5; 8 M urea). After binding of the protein, the Ni-NTA Agarose beads have been washed with 10 ml of buffer and then eluted with 10 ml of elution buffer (500 mM NaCl; 100 mM TrisHCl pH=8.5; 4 M Urea; 250 mM imidazole). Eluate was stored at -80 °C.

Antibody production

After checking the purity of the protein by SDS-PAGE, the eluted protein PID-2::6xHis has been concentrated to a final concentration of 1,8 mg/ml and sent to Eurogentec for antibody production (two rabbits; 28-day Speedy protocol). We then received the serum from two rabbits (823 and 824) and used 823 for all the experiments (1:100 for immunoprecipitation).

Transgenic lines generation using the MosSCI system

In order to generate *pid-2* transgenic lines, we used the MosSCI system and targeted the locus *ttTi5605* on LGII, as previously described (Frøkjær-Jensen et al., 2008). The injection mix contains plasmids encoding for the co-injection markers (10 ng/µl pGH8; 2,5 ng/µl pCFJ90; 5 ng/µl pCFJ104), for the transposase (50 ng/µl pCFJ601) and for the desired transgene (50 ng/µl pRK1036, pRK1037 or pRK1038). We injected the mix in the strain EG6699, which carries a Mos insertion on the locus *ttTi5605* on LGII, and kept the worms at 25 °C until starvation. We first screened for mCherry expressing worms, indicative of a successful injection, and later on, for wild type moving worms that have no extrachromosomal array (no mCherry expression). From the templates pRK1036, pRK1037 and pRK1038, we have isolated *xfIs144* [eGFP::PID-2], *xfIs145* [PID-2::eGFP] and *xfIs146* [3xFLAG::PID-2], respectively. Worms were then lysed and genotyped to confirm the insertion of the transgene.

Generation of mutant and endogenously tagged lines using CRISPR/Cas9 technology

Cloning

All the sgRNAs have been cloned in the vector p46169 (a gift from John Calarco, (Friedland et al., 2013)), except for Y45G5AM.2_sgRNA7, which has been cloned in the vector pRK2412 (pDD162 backbone, Cas9 deleted with improved sgRNA(F+E) sequence, as described in (Chen et al., 2013)). We performed inverse PCR with Q5® Hot Start High-Fidelity DNA Polymerase (Art. No. M0493, New England Biolabs) using a reverse primer specific for the backbone and a forward primer annealing to the backbone containing the specific sequence of the sgRNA. To re-ligate the vector, 4 µl of PCR product were added to a mix containing 1X T4 DNA Ligase reaction buffer (Art. No. B0202S, New England Biolabs), 1 mM ATP (Art. No. P0756S, New England Biolabs), 400 units T4 DNA Ligase (Art. No. M0202S, New England Biolabs), 10 units T4 PNK (Art. No. M0201S, New England Biolabs) and 20 units DpnI (Art. No. R0176S, New England Biolabs) to digest the original plasmid and the reaction was performed overnight at 25 °C. 5 µl of the ligation reaction were then transformed in Subcloning Efficiency™ DH5α™ Competent Cells (Art. No. 18265017, Invitrogen™) and plated for selection on LB agar plates with Ampicillin (100 µg/ml). The plasmid was isolated from the bacterial culture (PureLink™ HiPure Plasmid Miniprep Kit, Art. No. K210011, Invitrogen™), checked by enzymatic digestion and sequencing.

List of primers and plasmids generated

Plasmid	Oligo	Sequence
	p46169_RV	aaacatttagatttgcaatt
pRK1047	W03G9.2_sgRNA2	gcatgaacacgcataacgcgagttttagagctagaaatagc
pRK1050	W03G9.2_sgRNA5	gcaaagtacgcgaaagagggttttagagctagaaatagc
pRK1052	W03G9.2_sgRNA7	gcttcatatagctgccgctcggttttagagctagaaatagc
pRK1053	W03G9.2_sgRNA8	gaatatcaagcactagttggcgttttagagctagaaatagc
pRK1054	Y45G5AM.2_sgRNA1	ggctgtcatcagcgctcgtgttttagagctagaaatagc
pRK1056	Y45G5AM.2_sgRNA3	gtttcgcggttcgaagctacggttttagagctagaaatagc
pRK1057	Y45G5AM.2_sgRNA4	ggattggaatcaacgtgaggttttagagctagaaatagc
pRK1059	Y45G5AM.2_sgRNA6	gaatgagatcgatcgaagccggttttagagctagaaatagc
pRK1060	Y45G5AM.2_sgRNA7	caattaaaaatgctctagatgtttaagagctatgctggaaac

Generation of mutant lines

Wild type worms have been injected with an injection mix containing 50 ng/µl pJW1259 (encoding for *Peft-3::cas9::tbb-2 3'UTR*, a gift from Jordan Ward, (Ward, 2015)), co-injection markers (10 ng/µl pGH8; 5 ng/µl pCFJ104; 2,5 ng/µl pCFJ90) and 30 ng/µl of each of the plasmids encoding for the sgRNAs, specifically pRK1054, pRK1056, pRK1057, pRK1059 to target the *Y45G5AM.2* locus and pRK1047, pRK1050, pRK1052, pRK1053 to target the *W03G9.2* locus. After injections, worms have been kept at 20 °C and F1 offspring expressing the co-injection markers have been singled out. After the F1 offspring have laid embryos, they have been picked to 5 µl of single worm lysis buffer and the lysate has been used as PCR template to screen for mutant alleles. We isolated two deletion alleles of *Y45G5AM.2/pid-5* (*xf181* and *xf182*) and two deletion alleles of *W03G9.2/pid-4* (*xf184* and *xf185*). Each allele has been sequenced to pinpoint the exact deletion at nucleotide resolution. The mutant strains have been outcrossed two times against wild type N2 strain to remove any potential off-targets effect of Cas9 and used for further experiments.

Generation of endogenously tagged lines

In order to introduce an epitope tag at the endogenous loci of *W03G9.2/pid-4* and *Y45G5AM.2/pid-5*, we used the co-conversion approach as previously described (Arribere et al., 2014). Wild type worms have been injected with an injection mix containing 50 ng/µl pJS164 (Cas9 + sgRNA *dpy-10*); 750 nM ssODN SJ665 (repair oligo for *dpy-10*(*cn64*)); 50 ng/µl pRK1053; 750 nM ssODN SJ969 (repair oligo for W03G9.2::3xMyc) to introduce a 3xMyc epitope tag at the endogenous *W03G9.2/pid-4* locus (*xf186*). Wild type worms have been injected with an injection mix containing 50 ng/µl pJS164; 500 nM ssODN SJ665; 50 ng/µl pRK1060; 1000 nM ssODN SJ964 (repair oligo for Y45G5AM.2::2xHA) to introduce a 2xHA epitope tag at the endogenous *Y45G5AM.2/pid-5* locus (*xf192*). After injections, worms have been kept at 20 °C and F1 offspring with a roller phenotype (*rol-6*) have been singled out. After the F1 offspring have laid embryos, they have been picked to 5 µl of single worm lysis buffer and the lysate has been used as PCR template to screen for edited alleles. Each allele has been sequenced to ensure that the insertion of the epitope is in frame. The tagged strains have been outcrossed two times against wild type N2 strain to remove any potential off-targets effect of Cas9 and used for further experiments.

To introduce a fluorescent protein at the endogenous locus, we have first used a *unc-58* co-conversion approach (Arribere et al., 2014) to introduce at the *W03G9.2/pid-4* and at the *Y45G5AM.2/pid-5* loci a sequence of 20 nucleotides of *dpy-10* gene that serves as efficient protospacer sequence for subsequent edits (*xf204* and *xf221*, respectively), as previously described (Mouridi et al., 2017). We then used the generated strains (RFK932 and RFK972) as reference for the injection of a mix containing 50 ng/µl pJS164 (Cas9 + sgRNA *dpy-10*); 1000 nM SJ665 (repair oligo for *dpy-10*(*cn64*)); 300 ng/µl SJP010, a PCR product that was amplified from plasmid pDD286 (a gift from Bob Goldstein, Addgene plasmid # 70684) and used as a donor for the insertion of the mTagRFP-T sequence at the endogenous *W03G9.2/pid-4* and *Y45G5AM.2/pid-5* loci (Paix et al., 2014). After injections, worms have been kept at 20 °C and F1 offspring with a roller phenotype have been singled out. After the F1 offspring have laid embryos, they have been picked to 5 µl of single worm lysis buffer and the lysate has been used as PCR template to screen for edited alleles. We isolated the *xf206* and *xf226* alleles, which have been sequenced to ensure that the insertion of the mTagRFP-T sequence is in frame. The tagged strains have been outcrossed two times against wild type N2 strain to remove any potential off-targets effect of Cas9 and used for further experiments.

List of ssODN repair templates

SJ665

cacttgaacttcaatacggcaagatgagaatgactggaaaccgtaccgcatgcggtgcctatggtagcggagcttcacatg
gcttcagaccaacagcctat

SJ969

gatgaaaaataattaaagcttgaatatcaagcactacaagtcttcctcgctgatcaacttctgctcgaggtcctcctcggagat
gagcttttgctcaagatcctcttcagaaataagtttttgttcacctccacctccggatccgttggccggcttcatatagctgccg
ctcgcgggaa

SJ964

cgaaaataacttaaaaacaattaaaaatgctctaggcatagtctggaacgtcatatgggtaagcgtaatctgggacatcgtat
ggataacctccacctccggatccgatcggctggcatgcattgaagagccagtcatattc

Protein extraction and immunoprecipitation using GFP-Trap®

Synchronized worms have been grown until adulthood, then washed with cold M9 buffer, collected in a final volume of 200 µl water and fast frozen on dry ice. 350 µl of

2X Lysis Buffer (20 mM Tris pH=7.5; 300 mM NaCl; 1 mM EDTA; 1% NP40; cOmplete™ Mini EDTA-free Protease Inhibitor Cocktail, Art. No. 11836170001, Roche) have been added to each sample. Worms have then been sonicated using Bioruptor Plus (30 seconds ON – 30 seconds OFF, 10 cycles, high) and lysates were spun down at maximum speed for 10 min at 4 °C. The cleared protein extracts (500 µl) have been transferred to a new tube and 30 µl of GFP-Trap®_M (Art. No. gtm-20, ChromoTek), previously equilibrated with Dilution/Wash Buffer (3 x 5 minutes) (10 mM Tris pH=7.5; 150 mM NaCl; 0,5 mM EDTA; cOmplete™ Mini EDTA-free Protease Inhibitor Cocktail), have been added to each sample and the immunoprecipitation was performed for 3 hours at 4 °C. Then, beads have been washed 3 x 5 minutes with 500 µl of Dilution/Wash Buffer. After the last wash, the beads have been resuspended in 25 µl of NuPAGE® LDS Sample Buffer 1x (Art. No. NP0007, Life technologies) with 100 mM DTT and boiled at 95 °C for 10 minutes.

Protein extraction and immunoprecipitation using Dynabeads™ Protein G

Synchronized worms have been grown until adulthood, then washed with cold M9 buffer, collected in a final volume of 200 µl water and fast frozen on dry ice. 350 µl of 2X Lysis Buffer (50 mM Tris HCl pH=7.5; 300 mM NaCl; 3 mM $MgCl_2$; 2 mM DTT; 0,2% Triton-X100; cOmplete™ Mini EDTA-free Protease Inhibitor Cocktail, Art. No. 11836170001, Roche) have been added to each sample. Worms have then been sonicated using Bioruptor Plus (30 seconds ON – 30 seconds OFF, 10 cycles, high) and lysates were spun down spun down at maximum speed for 10 min at 4 °C. The cleared protein extracts (500 µl) have been transferred to a new tube and 2 µg of antibody (αHA clone HA-7, Art. No. H3663, Sigma; αMYC 9B11, Art. No. 2276, Cell Signalling Technology) or 1:100 of serum (823 αPID-2) have been added to each sample. The immunoprecipitation was performed for 2 hours at 4 °C. Then, 30 µl of Dynabeads™ Protein G (Art.No. 10004D, Invitrogen™), previously equilibrated with Wash Buffer (3 x 5 minutes) (25 mM Tris HCl pH=7.5; 150 mM NaCl; 1,5 mM $MgCl_2$; 1 mM DTT; cOmplete™ Mini EDTA-free Protease Inhibitor Cocktail), have been added and incubate for 1 additional hour at 4 °C. Then, beads have been washed 3 x 5 minutes with 500 µl of Wash Buffer. After the last wash, the beads have been resuspended in 25 µl of NuPAGE® LDS Sample Buffer 1x (Art. No. NP0007, Life technologies) with 100 mM DTT and boiled at 95 °C for 10 minutes.

Mass Spectrometry

Label-free quantitative mass spectrometry has been performed as described in (Almeida et al., 2018). After boiling (see above), the samples were separated on a 4–12% gradient Bis-Tris gel (NuPAGE Bis-Tris gels, 1.0 mm, 10 well; Art. No. NP0321; Life Technologies) in 1X MOPS (NuPAGE 20X MOPS SDS running buffer; Art. No. NP0001; Life Technologies) at 180 V for 10 min, afterwards processed by in-gel digest (Kappei et al., 2013; Shevchenko et al., 2007) and desalted using a C18 StageTip (Rappsilber et al., 2007). The digested peptides were separated on a 25-cm reverse-phase capillary (75 µM inner diameter) packed with Reprosil C18 material (Dr. Maisch) with a 2 h gradient from 2 to 40% Buffer B (see Stage tip purification) with the EASY-nLC 1,000 system (Thermo Scientific). Measurement was done on a Q Exactive Plus mass spectrometer (Thermo Scientific) operated with a Top10 data-dependent MS/MS acquisition method per full scan (Bluhm et al., 2016). The measurements were processed with the MaxQuant software, version 1.5.2.8 (Cox & Mann, 2008) against the UniProt *C. elegans* database (version of May, 2016) for quantitation.

References

Almeida, M. V., de Jesus Domingues, A. M., & Ketting, R. F. (2019). Maternal and zygotic gene regulatory effects of endogenous RNAi pathways. *PLOS Genetics*, *15*(2), e1007784. https://doi.org/10.1371/journal.pgen.1007784

Almeida, M. V., de Jesus Domingues, A. M., Lukas, H., Mendez-Lago, M., & Ketting, R. F. (2019). RppH can faithfully replace TAP to allow cloning of 5'-triphosphate carrying small RNAs. *MethodsX*, *6*, 265–272. https://doi.org/10.1016/j.mex.2019.01.009

Almeida, M. V., Dietz, S., Redl, S., Karaulanov, E., Hildebrandt, A., Renz, C., … Ketting, R. F. (2018). GTSF-1 is required for formation of a functional RNA-dependent RNA Polymerase complex in Caenorhabditis elegans. *The EMBO Journal*, *37*(12), e99325. https://doi.org/10.15252/embj.201899325

Anne, J., Ollo, R., Ephrussi, A., & Mechler, B. M. (2007). Arginine methyltransferase Capsuleen is essential for methylation of spliceosomal Sm proteins and germ cell formation in Drosophila. *Development*, *134*(1), 137–146. https://doi.org/10.1242/dev.02687

Arribere, J. A., Bell, R. T., Fu, B. X. H., Artiles, K. L., Hartman, P. S., & Fire, A. Z. (2014). Efficient marker-free recovery of custom genetic modifications with CRISPR/Cas9 in Caenorhabditis elegans. *Genetics*, *198*(3), 837–846. https://doi.org/10.1534/genetics.114.169730

Ashe, A., Sapetschnig, A., Weick, E. M., Mitchell, J., Bagijn, M. P., Cording, A. C., … Miska, E. A. (2012). PiRNAs can trigger a multigenerational epigenetic memory in the germline of C. elegans. *Cell*. https://doi.org/10.1016/j.cell.2012.06.018

Bagijn, M. P., Goldstein, L. D., Sapetschnig, A., Weick, E.-M., Bouasker, S., Lehrbach, N. J., … Miska, E. A. (2012). Function, Targets, and Evolution of Caenorhabditis elegans piRNAs. *Science*.

Bajar, B., Wang, E., Zhang, S., Lin, M., Chu, J., Bajar, B. T., … Chu, J. (2016). A Guide to Fluorescent Protein FRET Pairs. *Sensors*, *16*(9), 1488. https://doi.org/10.3390/s16091488

Batista, P. J., Ruby, J. G., Claycomb, J. M., Chiang, R., Fahlgren, N., Kasschau, K. D., … Mello, C. C. (2008). PRG-1 and 21U-RNAs Interact to Form the piRNA Complex Required for Fertility in C. elegans. *Molecular Cell*. https://doi.org/10.1016/j.molcel.2008.06.002

Bluhm, A., Casas-Vila, N., Scheibe, M., & Butter, F. (2016). Reader interactome of epigenetic histone marks in birds. *PROTEOMICS*, *16*(3), 427–436. https://doi.org/10.1002/pmic.201500217

Brenner, S. (1974). THE GENETICS OF CAENORHABDITIS ELEGANS. *Genetics*, *77*(1).

Buckley, B. A., Burkhart, K. B., Gu, S. G., Spracklin, G., Kershner, A., Fritz, H., … Kennedy, S. (2012). A nuclear Argonaute promotes multigenerational epigenetic inheritance and germline immortality. *Nature*, *489*(7416), 447–451. https://doi.org/10.1038/nature11352

C. elegans Deletion Mutant Consortium, T. C. elegans D. M. (2012). large-scale screening for targeted knockouts in the Caenorhabditis elegans genome. *G3 (Bethesda, Md.)*, *2*(11), 1415–1425. https://doi.org/10.1534/g3.112.003830

Chen, B., Gilbert, L. A., Cimini, B. A., Schnitzbauer, J., Zhang, W., Li, G. W., … Huang, B. (2013). Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell*, *155*(7), 1479–1491. https://doi.org/10.1016/j.cell.2013.12.001

Conine, C. C., Moresco, J. J., Gu, W., Shirayama, M., Conte, D., Yates, J. R., & Mello, C. C. (2013). Argonautes promote male fertility and provide a paternal memory of germline gene expression in C. Elegans. *Cell*. https://doi.org/10.1016/j.cell.2013.11.032

Cox, D. N., Chao, A., Baker, J., Chang, L., Qiao, D., & Lin, H. (1998). A novel class of evolutionarily conserved genes defined by piwi are essential for stem cell self-renewal. *Genes and Development*. https://doi.org/10.1101/gad.12.23.3715

Cox, J., & Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and

proteome-wide protein quantification. *Nature Biotechnology*, *26*(12), 1367–1372. https://doi.org/10.1038/nbt.1511

Das, P. P., Bagijn, M. P., Goldstein, L. D., Woolford, J. R., Lehrbach, N. J., Sapetschnig, A., … Miska, E. A. (2008). Piwi and piRNAs Act Upstream of an Endogenous siRNA Pathway to Suppress Tc3 Transposon Mobility in the Caenorhabditis elegans Germline. *Molecular Cell*. https://doi.org/10.1016/j.molcel.2008.06.003

de Albuquerque, B. F. M., Luteijn, M. J., Cordeiro Rodrigues, R. J., van Bergeijk, P., Waaijers, S., Kaaij, L. J. T., … Ketting, R. F. (2014). PID-1 is a novel factor that operates during 21U-RNA biogenesis in Caenorhabditis elegans. *Genes and Development*. https://doi.org/10.1101/gad.238220.114

de Albuquerque, B. F. M., Placentino, M., & Ketting, R. F. (2015). Maternal piRNAs Are Essential for Germline Development following De Novo Establishment of Endo-siRNAs in Caenorhabditis elegans. *Developmental Cell*. https://doi.org/10.1016/j.devcel.2015.07.010

Forster, T. (1946). Energiewanderung und Fluoreszenz. *Die Naturwissenschaften*, *33*(6), 166–175. https://doi.org/10.1007/BF00585226

Friedland, A. E., Tzur, Y. B., Esvelt, K. M., Colaiácovo, M. P., Church, G. M., & Calarco, J. A. (2013). Heritable genome editing in C. elegans via a CRISPR-Cas9 system. *Nature Methods*, *10*(8), 741–743. https://doi.org/10.1038/nmeth.2532

Frøkjær-Jensen, C., Wayne Davis, M., Hopkins, C. E., Newman, B. J., Thummel, J. M., Olesen, S. P., … Jorgensen, E. M. (2008). Single-copy insertion of transgenes in Caenorhabditis elegans. *Nature Genetics*. https://doi.org/10.1038/ng.248

Ghildiyal, M., & Zamore, P. D. (2009). Small silencing RNAs: An expanding universe. *Nature Reviews Genetics*. https://doi.org/10.1038/nrg2504

Gu, W., Lee, H.-C., Chaves, D., Youngman, E. M., Pazour, G. J., Conte, D., & Mello, C. C. (2012). CapSeq and CIP-TAP identify Pol II start sites and reveal capped small RNAs as C. elegans piRNA precursors. *Cell*, *151*(7), 1488–1500. https://doi.org/10.1016/j.cell.2012.11.023

Gu, W., Shirayama, M., Conte, D., Vasale, J., Batista, P. J., Claycomb, J. M., … Mello, C. C. (2009). Distinct argonaute-mediated 22G-RNA pathways direct genome surveillance in the C. elegans germline. *Molecular Cell*, *36*(2), 231–244. https://doi.org/10.1016/j.molcel.2009.09.020

Hirota, K., Shigekawa, C., Araoi, S., Sha, L., Inagawa, T., Kanou, A., … Fukamizu, A. (2017). Simultaneous ablation of prmt-1 and prmt-5 abolishes asymmetric and symmetric arginine dimethylations in Caenorhabditis elegans. *The Journal of Biochemistry*, *161*(6), 521–527. https://doi.org/10.1093/jb/mvw101

Iyer, S., La-Borde, P. J., Payne, K. A. P., Parsons, M. R., Turner, A. J., Isaac, R. E., & Acharya, K. R. (2015). Crystal structure of X-prolyl aminopeptidase from *Caenorhabditis elegans* : A cytosolic enzyme with a di-nuclear active site. *FEBS Open Bio*, *5*(1), 292–302. https://doi.org/10.1016/j.fob.2015.03.013

Kamminga, L. M., van Wolfswinkel, J. C., Luteijn, M. J., Kaaij, L. J. T., Bagijn, M. P., Sapetschnig, A., … Ketting, R. F. (2012). Differential Impact of the HEN1 Homolog HENN-1 on 21U and 26G RNAs in the Germline of Caenorhabditis elegans. *PLoS Genetics*, *8*(7), e1002702. https://doi.org/10.1371/journal.pgen.1002702

Kappei, D., Butter, F., Benda, C., Scheibe, M., Draškovič, I., Stevense, M., … Buchholz, F. (2013). HOT1 is a mammalian direct telomere repeat-binding protein contributing to telomerase recruitment. *The EMBO Journal*, *32*(12), 1681–1701. https://doi.org/10.1038/emboj.2013.105

Kawasaki, I., Shim, Y. H., Kirchner, J., Kaminker, J., Wood, W. B., & Strome, S. (1998). PGL-1, a predicted RNA-binding component of germ granules, is essential for fertility in C. elegans. *Cell*, *94*(5), 635–645. https://doi.org/10.1016/S0092-8674(00)81605-0

Ketting, R. F. (2011). The Many Faces of RNAi. *Developmental Cell*. https://doi.org/10.1016/j.devcel.2011.01.012

Ketting, R. F., Haverkamp, T. H. A., Van Luenen, H. G. A. M., & Plasterk, R. H. A. (1999). mut-7 of C. elegans, required for transposon silencing and RNA interference, is a homolog of werner syndrome helicase and RNaseD. *Cell*.

https://doi.org/10.1016/S0092-8674(00)81645-1

Langmead, B., Trapnell, C., Pop, M., & Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology*, *10*(3), R25. https://doi.org/10.1186/gb-2009-10-3-r25

Laurent, V., Brooks, D. R., Coates, D., & Isaac, R. E. (2001). Functional expression and characterization of the cytoplasmic aminopeptidase P of *Caenorhabditis elegans*. *European Journal of Biochemistry*, *268*(20), 5430–5438. https://doi.org/10.1046/j.0014-2956.2001.02483.x

Lawrence, M., & Morgan, M. (2014). Scalable Genomics with R and Bioconductor. *Statistical Science : A Review Journal of the Institute of Mathematical Statistics*, *29*(2), 214–226. https://doi.org/10.1214/14-STS476

Lee, H. C., Gu, W., Shirayama, M., Youngman, E., Conte, D., & Mello, C. C. (2012). C. elegans piRNAs mediate the genome-wide surveillance of germline transcripts. *Cell*. https://doi.org/10.1016/j.cell.2012.06.016

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., … 1000 Genome Project Data Processing Subgroup. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Luteijn, M. J., Van Bergeijk, P., Kaaij, L. J. T., Almeida, M. V., Roovers, E. F., Berezikov, E., & Ketting, R. F. (2012). Extremely stable Piwi-induced gene silencing in Caenorhabditis elegans. *EMBO Journal*. https://doi.org/10.1038/emboj.2012.213

Malone, C. D., & Hannon, G. J. (2009). Small RNAs as Guardians of the Genome. *Cell*. https://doi.org/10.1016/j.cell.2009.01.045

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.Journal*, *17*(1), 10. https://doi.org/10.14806/ej.17.1.200

Mouridi, S. El, Lecroisey, C., Tardy, P., Mercier, M., Leclercq-blondel, A., Zariohi, N., … Neuromyogène, I. (2017). Reliable CRISPR / Cas9 Genome Engineering in Caenorhabditis elegans Using a Single Ef fi cient sgRNA and an Easily Recognizable Phenotype. *G3 Genes Genomes Genetics*, *7*(May), 1429–1437. https://doi.org/10.1534/g3.117.040824/-/DC1.1

Paix, A., Wang, Y., Smith, H. E., Lee, C. Y. S., Calidas, D., Lu, T., … Seydoux, G. (2014). Scalable and versatile genome editing using linear DNAs with microhomology to Cas9 sites in Caenorhabditis elegans. *Genetics*, *198*(4), 1347–1356. https://doi.org/10.1534/genetics.114.170423

Pek, J. W., Anand, A., & Kai, T. (2012). Tudor domain proteins in development. *Development*. https://doi.org/10.1242/dev.073304

Phillips, C. M., Brown, K. C., Montgomery, B. E., Ruvkun, G., & Montgomery, T. A. (2015). PiRNAs and piRNA-Dependent siRNAs Protect Conserved and Essential C. elegans Genes from Misrouting into the RNAi Pathway. *Developmental Cell*. https://doi.org/10.1016/j.devcel.2015.07.009

Phillips, C. M., Montgomery, B. E., Breen, P. C., Roovers, E. F., Rim, Y. S., Ohsumi, T. K., … Montgomery, T. A. (2014). MUT-14 and SMUT-1 DEAD box RNA helicases have overlapping roles in germline RNAi and endogenous siRNA formation. *Current Biology*. https://doi.org/10.1016/j.cub.2014.02.060

Phillips, C. M., Montgomery, T. A., Breen, P. C., & Ruvkun, G. (2012). MUT-16 promotes formation of perinuclear Mutator foci required for RNA silencing in the C. elegans germline. *Genes and Development*. https://doi.org/10.1101/gad.193904.112

Quinlan, A. R., & Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics (Oxford, England)*, *26*(6), 841–842. https://doi.org/10.1093/bioinformatics/btq033

Rappsilber, J., Mann, M., & Ishihama, Y. (2007). Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nature Protocols*, *2*(8), 1896–1906. https://doi.org/10.1038/nprot.2007.261

Robert, X., & Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Research*, *42*(W1), W320–W324. https://doi.org/10.1093/nar/gku316

Sapetschnig, A., Sarkies, P., Lehrbach, N. J., & Miska, E. A. (2015). Tertiary siRNAs Mediate Paramutation in C. elegans. *PLOS Genetics*, *11*(3), e1005078. https://doi.org/10.1371/journal.pgen.1005078

Shevchenko, A., Tomas, H., Havli\[sbreve], J., Olsen, J. V, & Mann, M. (2007). In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nature Protocols*, *1*(6), 2856–2860. https://doi.org/10.1038/nprot.2006.468

Shirayama, M., Seth, M., Lee, H. C., Gu, W., Ishidate, T., Conte, D., & Mello, C. C. (2012). PiRNAs initiate an epigenetic memory of nonself RNA in the C. elegans germline. *Cell*. https://doi.org/10.1016/j.cell.2012.06.015

Simon, M., Sarkies, P., Ikegami, K., Doebley, A.-L., Goldstein, L. D., Mitchell, J., … Ahmed, S. (2014). Reduced insulin/IGF-1 signaling restores germ cell immortality to Caenorhabditis elegans Piwi mutants. *Cell Reports*, *7*(3), 762–773. https://doi.org/10.1016/j.celrep.2014.03.056

Siomi, M. C., Mannen, T., & Siomi, H. (2010). How does the royal family of Tudor rule the PIWI-interacting RNA pathway? *Genes & Development*, *24*(7), 636–646. https://doi.org/10.1101/gad.1899210

Spracklin, G., Fields, B., Wan, G., Becker, D., Wallig, A., Shukla, A., & Kennedy, S. (2017). The RNAi Inheritance Machinery of Caenorhabditis elegans. *Genetics*, *206*(3), 1403–1416. https://doi.org/10.1534/genetics.116.198812

Tang, W., Seth, M., Tu, S., Shen, E.-Z., Li, Q., Shirayama, M., … Mello, C. C. (2018). A Sex Chromosome piRNA Promotes Robust Dosage Compensation and Sex Determination in C. elegans. *Developmental Cell*, *44*(6), 762–770.e3. https://doi.org/10.1016/j.devcel.2018.01.025

Wang, G., & Reinke, V. (2008). A C. elegans Piwi, PRG-1, Regulates 21U-RNAs during Spermatogenesis. *Current Biology*. https://doi.org/10.1016/j.cub.2008.05.009

Ward, J. D. (2015). Rapid and precise engineering of the Caenorhabditis elegans genome with lethal mutation co-conversion and inactivation of NHEJ repair. *Genetics*, *199*(2), 363–377. https://doi.org/10.1534/genetics.114.172361

Xiao, Y., Bedet, C., Robert, V. J. P., Simonet, T., Dunkelbarger, S., Rakotomalala, C., … Palladino, F. (2011). Caenorhabditis elegans chromatin-associated proteins SET-2 and ASH-2 are differentially required for histone H3 Lys 4 methylation in embryos and adult germ cells. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(20), 8305–8310. https://doi.org/10.1073/pnas.1019290108

Zhang, C., Montgomery, T. A., Gabel, H. W., Fischer, S. E. J., Phillips, C. M., Fahlgren, N., … Ruvkun, G. (2011). mut-16 and other mutator class genes modulate 22G and 26G siRNA pathways in Caenorhabditis elegans. *Proceedings of the National Academy of Sciences*. https://doi.org/10.1073/pnas.1018695108

Zhou, X., Xu, F., Mao, H., Ji, J., Yin, M., Feng, X., & Guang, S. (2014). Nuclear RNAi Contributes to the Silencing of Off-Target Genes and Repetitive Sequences in Caenorhabditis elegans. *Genetics*, *197*(1), 121–132. https://doi.org/10.1534/GENETICS.113.159780

Chapter 5

Discussion

Caenorhabditis elegans: a model for investigating the complexity of small RNA pathways

Small RNA pathways are present in virtually all organisms, although they are extremely diversified. Despite a lot of variability, these pathways have the main function of recognizing a specific target RNA in order to silence it, and are collectively grouped as RNA interference. This target RNA molecule can be of viral or, more generally, exogenous origin. In this scenario, RNAi exerts an antiviral, defensive function against invasive agents, similarly to the adaptive immune system. This is indeed the most ancient function of RNAi, which later evolved in more diversified pathways; most of them still retain their original defensive function (Castel & Martienssen, 2013; Ishizu et al., 2012; Ketting, 2011; Malone & Hannon, 2009; Olina et al., 2018; Ozata et al., 2018). Nevertheless, it is becoming more and more evident, that RNAi has also a regulatory function towards endogenous genes.

RNAi, during evolution, lost or gained factors involved in the silencing pathways; for example, *C. elegans* still has RdRPs, which are instead lost in many other eukaryotes. Yet, the RNAi pathways can vary substantially also among closely related species, such as within the nematode phylum, indicating a fast rate of evolution, possibly to adapt to the threats that each species is exposed to (Almeida et al., 2019). *C. elegans* represents a fast and easy-to-handle model organism, to investigate the details of different biological processes.

We have shown, using the nematode *C. elegans*, that maternally deposited small RNAs are required to instruct the RNAi machinery in the next generation for proper gene silencing and also to ensure proper gonad development and fertility (chapters 2 and 3) (de Albuquerque et al., 2015). We also identified three novel proteins, PID-2, PID-4 and PID-5, that form a germline specific complex required to establish *de novo* silencing of a 21U RNA target transgene. Furthermore, this novel complex is also required for germline immortality, and PID-2 has a role in enhancing the production of 22G RNAs on target mRNAs, downstream of PRG-1 target recognition mediated by 21U RNAs (chapter 4). In this chapter, I will discuss the broader implications of the experimental work described in this thesis.

Maternally inherited small RNAs are required to direct gene expression in the next generation and for fertility

Maternally deposited RNAs and proteins are essential to guarantee embryonic development and instruct the silencing machinery for proper gene expression and silencing (Despic & Neugebauer, 2018; Skvortsova et al., 2018). In *C. elegans*, 22G RNAs deposited from the maternal and paternal gametes are necessary to ensure gonad development and fertility in the next generation. Such effects become apparent in absence of parental 21U RNAs, that are the initial input for gene silencing. On the other hand, a subset of 22G RNAs is essential to guarantee the expression of their target mRNAs, by routing them to the CSR-1 pathway. Therefore, inheritance of 22G RNAs is fundamental to ensure the balance between the silencing pathway, executed by PRG-1 and downstream WAGO proteins, and its counterpart, mediated by CSR-1 (de Albuquerque et al., 2015; Phillips et al., 2015). It is clear that the 21U RNA pathway does not only target exogenous mRNAs for silencing, but it is also involved in regulating endogenous gene expression (Bagijn et al., 2012; Batista et al., 2008; Das et al., 2008; Kasper et al., 2014; Shen et al., 2018; Tang et al., 2018). PRG-1 is only required to recognize a target mRNA and to initiate the silencing, but the silencing reaction is then amplified and exerted by downstream WAGO proteins and their 22G RNA cofactors. The 21U RNA pathway therefore becomes essential for fertility and proper gene expression pattern, only in absence of parental 22G RNAs and RNAe memory, as the RNAi machinery lacks the information to properly direct gene silencing (de Albuquerque et al., 2015; Phillips et al., 2015). Interestingly, the sterility that is thus induced can be rescued by the absence of HRDE-1, as RNAe cannot be established in this setup. These findings indicate that the absence of silencing mechanisms is somehow less detrimental than incorrect gene silencing. Nonetheless, both *prg-1* and *hrde-1* mutants have a quite severe mortal germline phenotype (Mrt) (Buckley et al., 2012; Simon et al., 2014), as they lose germ cells within few generations. The cause of the Mrt phenotype in these mutants is not known yet, and whether this relates to the defects described in chapter 2 or not remains to be addressed. On the other hand, mutator mutants, such as *mut-7* and *mut-16*, do not have a Mrt phenotype, indicating that loss of 22G RNAs production as well as gradual loss of RNAe from endogenous loci alone does not trigger the Mrt phenotype.

It may be worthwhile to consider also the contribution of the 26G RNA pathway, as these small RNAs can also direct silencing as well as transmit information about expressed genes via the male gametes (Conine et al., 2010; Han et al., 2009). Furthermore, different proteins are involved in multiple small RNA pathways (Almeida et al., 2019; Billi et al., 2014), so possibly in absence of a functional RNAe pathway or of input from the 21U RNA pathway, these proteins are re-wired to an alternative silencing pathway, which could compensate to some extent for the absence of the others.

Similarly to *C. elegans*, initiation and maintenance of gene silencing may well be distinct also in other organisms. However, in many organisms piRNAs take care of both processes, making it very difficult to tease apart these different aspects. Nevertheless, co-factors may exist that are specialized in one or the other function. These have thus far not been identified.

RNAi-like pathways act differentially during development and their functionality relies on maternal contribution

We have shown in chapter 3 that 21U RNAs, inherited from the maternal lineage, are necessary and sufficient to establish *de novo* silencing of a 21U RNA target transgene (de Albuquerque et al., 2014). Possibly, 21U RNAs need to act early during embryonic development to establish target silencing. Outside of this time window, target silencing may no longer be established, such as after the onset of embryonic transcription (Mello et al., 1992; Seydoux & Dunn, 1997), when the offspring can synthetize their own 21U RNAs. Interestingly, also in other organisms, such as *Drosophila*, a similar observation indicates that Piwi-mediated gene silencing has to be established in primordial germ cells (Akkouche et al., 2017; Siddiqui et al., 2012). Nonetheless, we have observed that the silencing triggered by maternally deposited 21U RNAs only is less robust compared to the situation when also zygotic 21U RNAs can be produced. This suggests that besides maternal 21U RNAs, also zygotic 21U RNAs do have an effect on silencing, even though zygotic 21U RNAs alone are not sufficient to initiate the silencing. These results seem to indicate that 21U RNAs induce silencing in a time window when both maternal and zygotic 21U RNAs are present.

Given these results, and that the silencing induced by 21U RNAs can often be maintained in their absence, it may be possible that the 21U RNAs that are made during adulthood in *C. elegans* are purely meant for loading into the embryo, and not to induce silencing in the adult gonad. With this in mind, it is interesting to note the differential effects that have been described upon loss of HENN-1 activity in the embryo versus the adult (Billi et al., 2012; Kamminga et al., 2012; Montgomery et al., 2012). In the embryo, loss of HENN-1 led to a reduction of 21U RNA levels, accompanied by increased 3' end uridylation (Billi et al., 2012; Montgomery et al., 2012), both aspects expected to occur in *henn-1* mutants (Kamminga et al., 2010). In the adult, however, loss of HENN-1 did not trigger these effects on 21U RNAs, even though they were detected on another small RNA species that is modified by HENN-1, 26G RNAs (Kamminga et al., 2012). Given the idea that 3' end uridylation, followed by destabilization, is triggered by target recognition (Ameres et al., 2010), this may indirectly reveal that 21U RNAs in the adult gonad are not very actively binding to target RNAs, and hence may not be involved in driving silencing.

We have shown that different populations isolated from the same offspring, receiving maternal 21U RNAs only, have a high variability among their transcriptomes, despite being genetically identical. Possibly, this variability reflects the composition of the initial 21U RNA pool inherited, or stochastic fluctuations in reaching thresholds required for establishing silencing. Given that homozygous mutants, at some point in their generation, have experienced exposure to maternal 21U RNAs, this finding has implications for studies using mutants from this pathway. Our results imply that different individuals, used to establish a new mutant line, will have different mRNA expression profiles, possibly affecting the outcomes of downstream studies. Multiple interpedently established lines should therefore better be used.

In order to further dissect the 21U RNA pathway, we will need to take into account a more complicated relationship with other small RNA pathways as well as the developmental stage studied. Indeed, different small RNA pathways may act differently at different time points, or they could only exert their function during one particular developmental stage, such as in early embryonic development when they are maternally deposited, or only at a later step of the life cycle.

Such temporal specificity is a conserved feature among organisms. For example, in mice, different PIWI proteins are expressed at different stages of spermatogenesis (Chuma & Nakano, 2012; Rojas-Ríos & Simonelig, 2018); therefore, they can only exert their functions within those specific time windows. Furthermore, also the composition of the piRNA populations varies during development. The PIWI pathway could hence function differently at different developmental stages, adapting to the specific needs of the developmental stage. Thus far, such a developmental layer is often absent from piRNA-related studies, and will need to be included to fully understand how these intriguing molecules work.

Why do maternal 21U RNAs have such a strong impact?

In many organisms, inheritance of proteins and nucleic acids from the parents is essential for the next generation. In *C. elegans*, RNAe and CSR-1 pathways represent a memory system of formerly silenced and expressed genes, respectively, fundamental to direct the silencing machinery transgenerationally. We have shown that maternally provided 21U RNAs are necessary and sufficient for instructing the RNAi machinery in the next generation; they are the prerequisite for *de novo* target silencing, independently of the integrity of the silencing pathways in the offspring (see chapters 2 and 3) (de Albuquerque et al., 2014, 2015; Phillips et al., 2015). Why do maternal 21U RNAs influence so strongly the silencing in the next generation? It has been shown that environmental stimuli can influence gene expression as well as the small RNA pathways, and this information can be transmitted transgenerationally (Klosin et al., 2017; Rechavi et al., 2014). Lack of food sources induces a developmental arrest at the first larval stage in *C. elegans*, that can be reversed upon restoration of favourable environmental conditions. If the parents have been starved, they encounter changes in gene expression, resulting from changes in 22G RNAs production. Such changes affect specifically genes involved in nutrition and can still be detected not only in the offspring, but also after three generations, although these generations have not been food-deprived (Rechavi et al., 2014). Another form of stress for *C. elegans* is represented by high temperature. Exposure of animals at 25 °C can also affect gene expression, in particular it increases the expression of a multi-copy transgene, and such effect can be detected up to 14 generations after the exposure to high temperature (Klosin et al., 2017). Possibly, maternal 21U RNAs are involved in such heritability, in

order to ensure a prompt adaptation of the offspring to environmental conditions, as instructed by their parents. Considering that *C. elegans* has a very fast life cycle, such adaptation should occur extremely rapidly.

PID-2 enhances the production of 22G RNAs to ensure the silencing function of maternally provided 21U RNAs

We showed that maternal 21U RNA activity can induce RNAe on a 21U sensor, in absence of a zygotic 21U RNA functional pathway, and this requires a protein that we newly identified: PID-2. The precise molecular function of PID-2 is not clear yet, but from our data, it appears that PID-2 is required to mediate inheritance of small RNA populations and/or to boost somehow the 22G RNA response in primordial germ cells or in the early gonad.

Of note, expression of the 21U sensor in *pid-2* mutants is lower than in mutants lacking either 21U or 22G RNAs (e.g. *prg-1* or *mut-7* mutants, respectively). Similar mild expression of such a sensor has been seen in *henn-1* mutants, in which 21U RNA maturation is compromised. Despite the lower expression, in *pid-2* mutants, we have never observed individuals in which the mild signal was lost and the 21U sensor was thereafter silenced. This is instead readily observed in *henn-1* mutants and led to the discovery of RNAe. Therefore, PID-2 activity, and its enhancing effects on 22G RNA levels, may be a pre-requisite for the establishment of RNAe. Possibly, only if a critical threshold of 22G RNA levels is reached, RNAe might be established. In absence of PID-2, such levels may simply not be achieved, as our small RNA sequencing data from *pid-2* mutants suggest.

This finding, together with the fact that 21U RNAs themselves are not affected in *pid-2* mutants, labels PID-2 as a link between PRG-1 and secondary 22G RNA binding Argonaute proteins, such as HRDE-1. However, even though RNAe maintenance is not visibly affected in *pid-2* mutants, the levels of 22G RNA populations produced on 21U RNA targets, that are already stably silenced via RNAe, independently of PRG-1, are reduced in *pid-2* mutants. This shows that PID-2 is not specifically coupled to 22G RNA production in response to PRG-1. However, we could still not pinpoint the molecular function(s) of PID-2. It is exciting to speculate that PID-2 recruits an RdRP to the target mRNA, or somehow interacts with WAGO or mutator proteins to enhance the production,

or stabilization of 22G RNAs. We were not able to identify such interactions through IP/MS experiments, possibly because they are very dynamic and transient, or possibly because they do not exist and PID-2 influences the synthesis of 22G RNAs via a different mechanism.

PID-2 colocalizes with PGL-1, a P granule marker, both in the germline as well as in embryos, so possibly it could have a function within the P granule, where many loaded Argonaute proteins reside and the target silencing takes place. PID-2 could help secondary Argonaute proteins to amplify the silencing, or mediate the connection between P granules and mutator foci, where 22G RNAs are generated, thereby acting as an enhancer of 22G RNAs production.

Interestingly, when scoring for the expression of the 21U sensor in the offspring of *pid-2* mutant hermaphrodites, every now and then we observed animals in which the 21U sensor was expressed in one gonadal arm, but silenced in the other one (not shown). Given that the two different gonad arms are derived from the two primordial germ cells, Z2 and Z3, this could indicate that the silencing is established in the Z2/Z3 cells. At this stage, maternal factor are still present, but zygotic transcription also starts, consistent with the idea that 21U RNA-driven silencing is established using both maternal and zygotic RNAs. Indeed, PID-2 itself, plus its binding partners PID-4 and PID-5, are found in Z2/Z3 cells as well.

PID-2 has a perinuclear localization around germ cells nuclei, specifically during pachytene stage. Why is the expression of PID-2 so specific for this meiotic stage? At the pachytene stage, the synaptonemal complex covers the interface between homologous chromosomes, previously paired during leptotene/zygotene, and at this stage of meiosis, crossovers are formed (Woglar & Jantsch, 2014), so possibly PID-2 could have some relevance in this process. We know that unpaired chromatin during meiosis is recognized and targeted for silencing (Hammond, 2017), such as the X chromosome in males, and transgenes, if present. In *C. elegans*, the Argonaute protein CSR-1 is involved in chromosome segregation and silencing of unpaired DNA (Claycomb et al., 2009; Wedeles et al., 2013). It is possible that PID-2 in some ways has a more general role in production of 22G RNAs, thereby affecting also the CSR-1 pathway, to ensure proper gene expression and chromosome segregation.

PID-2 interacts with two novel proteins, PID-4 and PID-5, to promote silencing

We identified two novel proteins that bind PID-2: PID-4 and PID-5. This novel complex has a role in boosting the production of 22G RNAs, which are not only required to establish *de novo* silencing, but also for maintenance of RNAe memory, as well as for ensuring correct expression of endogenous genes. PID-2/-4/-5 may play a role in helping to establish the correct pattern of gene expression and silencing in the embryo, using the small RNA populations that are parentally inherited. We know that PID-2 acts downstream of PRG-1, so possibly this novel complex could interact with downstream factors to ensure appropriate gene silencing, such as an RdRP or WAGO proteins. An interesting hypothesis is that PID-5 could use its Xaa-Pro aminopeptidase domain to recruit WAGO proteins, such as WAGO-3/PPW-2, WAGO-4 and WAGO-10, which have a proline at position 2 at their N-terminal sequence and are involved in secondary 22G RNA-mediated response and loading of 22G RNAs into *C. elegans* embryos. Intriguingly, in PID-2, PID-4 and PID-5 IP/MS experiments, WAGO-3 and WAGO-4 consistently tend to be slightly enriched, even though these enrichments were never close to being significant. Maybe the complex PID-2/-4/-5 merges the information about gene silencing inherited both from the maternal lineage, transmitted via WAGO-4 (Xu et al., 2018), and from the paternal lineage, via WAGO-3 (Schreier et al., unpublished), to then ensure appropriate gene silencing in the embryo. Also, CSR-1 transmits information about expressed genes via the sperm (Conine et al., 2013) and they could also funnel into PID-2/-4/-5 complex, which would then function as a hub for instructing the gene expression/silencing machinery.

PID-4 and PID-5, in analogy to other Tudor domain proteins (Pek et al., 2012), possibly use their Tudor domains to act as a platform to recruit and assemble a multiprotein complex, which is seemingly involved in establishing *de novo* gene silencing as well as in enhancing the production of 22G RNAs. In other organisms, Tudor proteins contain multiple Tudor domains, which are required to recruit multiple effector proteins (Pek et al., 2012). In *C. elegans* instead, Tudor domains come mostly alone. This possibly relates to the fact that the silencing reaction mediated by 22G RNAs and WAGO proteins is strongly amplified, starting from few 21U or 26G RNA input molecules. Therefore, a single Tudor domain could be sufficient to recruit one of the effector proteins, such as a WAGO protein or an RdRP, which then would trigger a massive downstream amplification reaction.

Furthermore, the perinuclear granules, such as P granules and mutator foci, in which the mRNA surveillance and silencing take place, have a very high concentration of factors involved in these processes. Hence, a single Tudor domain could bring together all the proteins necessary to ensure appropriate target silencing. This possibly relates also to the small size of germ cells and, consequently, of these perinuclear granules. In fact, in other organisms, such as zebrafish and mouse, analogous structures, such as the chromatoid body and the Balbiani body, respectively, are much larger; therefore, they may require proteins with multiple Tudor domains, or numerous Tudor proteins, in order to achieve an adequate size. For example, the Tudor protein Tdrd6 in mouse, and Tdrd6a in zebrafish, have seven Tudor domains and are essential components of the chromatoid body and the Balbiani body, respectively (Hosokawa et al., 2007; Roovers et al., 2018; Vasileva et al., 2009).

We made an intriguing observation regarding *pid-4* and *pid-5*: adjacent to the genomic location of *pid-4*, the *app-1* gene is located. The *app-1* gene encodes the catalytically active Xaa-Pro aminopeptidase (Iyer et al., 2015; Laurent et al., 2001). It is exciting to speculate that *pid-5* is the result of some genomic rearrangement between the *pid-4* and *app-1* genes, leading to the fusion of the aminopeptidase domain of APP-1 to PID-4, generating PID-5. We also did conservation analysis in order to identify *pid-5* orthologs in other nematode species. Interestingly, we found *pid-5* orthologs in *C. remanei*, *C. brenneri* and *C. briggsae*. Furthermore, in the same species, an ortholog of *pid-4* is always found adjacent to *app-1*. Interestingly, in *C. japonica*, which is significantly further away from *C. elegans* than *C. remanei*, *C. brenneri* and *C. briggsae* (Kanzaki et al., 2018), the ortholog of *pid-4* is also located adjacent to *app-1*, but an ortholog of *pid-5* is not present. This indeed suggests that there has been some sort of rearrangement and duplication event that led to the generation of *pid-5* after the last common ancestor of *C. elegans* and *C. japonica*.

The role of PID-2/-4/-5 in maintaining an immortal germline

PID-2 and, to a weaker extent, PID-4 and PID-5 are required to ensure a healthy germline over multiple generations, as mutants for these proteins progressively lose germ cells. This mortal germline phenotype (Mrt) has been described also for many other factors involved in the 21U RNA pathway as well as in RNAe (Buckley et al., 2012; Simon et al., 2014; Spracklin et al., 2017; Xiao et al., 2011). It has been proposed that such Mrt phenotype is

somehow connected to metabolic pathways; specifically, reduction of insulin/insulin-like growth factor-1 (IGF-1) signalling contributes to germ cell immortality. Upon starvation, the transcription factor DAF-16 is activated to promote stress resistance and longevity, and additionally germline immortality is restored in *prg-1* mutants (Simon et al., 2014). Apart from starvation, a different food source can also induce a reversion of the *prg-1* induced sterility at late generations; yet, this phenomenon is quite rare (Heestand et al., 2018). It has been proposed that an alternative small RNA silencing pathway, independent of PRG-1, is triggered by DAF-16, to ensure silencing of repetitive elements, thereby contributing to fertility. This alternative silencing pathway requires factors that are involved in 22G RNAs biogenesis and function, such as MUT-7, RDE-2/MUT-8 and PPW-1/WAGO-7 (Simon et al., 2014).

DAF-2 encodes the insulin/IGF-1 receptor homolog, which is activated upon food availability and suppresses the downstream transcription factor DAF-16. Intriguingly, in favourable growth conditions, DAF-2 activates the downstream RAS-ERK signalling pathway, thereby stimulating progression through prophase of meiosis I, and consequently oogenesis. On the other hand, in absence of nutrients, the IGF-1 signalling is not activated, and the production of oocytes is arrested. In fact, the presence of food is sensed as a positive environment, that would ensure the survival of the embryos. DAF-2 stimulates oogenesis progression by acting specifically at the pachytene stage (Lopez et al., 2013), which is exactly the stage of major expression of PID-2/-4/-5.

Interestingly, the relationship between insulin signalling and female fertility is well established, not only in worms, but is also conserved from *Drosophila* to mammals. Insulin, in fact, influences different aspects of female fertility and successful reproduction, from oocyte growth and maturation to ovaries formation (Das & Arur, 2017). In *Drosophila*, food availability induces the fly adipose tissue to stimulate the release of small insulin-like peptides from a subset of neurons. Such peptides then activate the insulin signalling, and, specifically in ovaries, they stimulate stem cells proliferation and oocytes maturation via the PI3K/AKT signalling cascade (Colombani et al., 2003; LaFever & Drummond-Barbosa, 2005). In mammals, there is no evidence that insulin directly contributes to oocytes growth and/or development. Nonetheless, diet has been shown to influence oocyte quality and fertility in cows, sheep and pigs (Fouladi-Nashta et al., 2007; Papadopoulos et

al., 2001). In humans, it is well documented that uncontrolled levels of insulin are connected with defects in oocytes development, resulting in defects at birth or pathological conditions during pregnancy. Furthermore, the downstream PI3K/AKT signalling pathway has a role in contributing to oocytes formation and female fertility (Das & Arur, 2017).

To conclude, the PID-2/-4/-5 complex could serve as an information centre where all the data carried by the 22G RNAs, both derived from parental epigenetic memory and from environmental clues, are integrated to ensure coordination between gene silencing/expression and survival of the organisms as well as germline development. Such a function has not been described for a protein complex in any of the small RNA pathways that are studied, and thus make PID-2/-4/-5 an extremely interesting set of proteins for further investigation. It is, however, also clear that the network of information flow in *C. elegans* is very complex, and it will not be easy to disentangle. Starting to decipher the precise molecular functions and/or activities of the proteins we identified will therefore be a more feasible way to move forward in the near future.

References

Akkouche, A., Mugat, B., Barckmann, B., Varela-Chavez, C., Li, B., Raffel, R., … Chambeyron, S. (2017). Piwi Is Required during Drosophila Embryogenesis to License Dual-Strand piRNA Clusters for Transposon Repression in Adult Ovaries. *Molecular Cell*, *66*(3), 411–419.e4. https://doi.org/10.1016/j.molcel.2017.03.017

Almeida, M. V., Andrade-Navarro, M. A., & Ketting, R. F. (2019). Function and Evolution of Nematode RNAi Pathways. *Non-Coding RNA*, *5*(1), 8. https://doi.org/10.3390/ncrna5010008

Ameres, S. L., Horwich, M. D., Hung, J.-H., Xu, J., Ghildiyal, M., Weng, Z., & Zamore, P. D. (2010). Target RNA-directed trimming and tailing of small silencing RNAs. *Science (New York, N.Y.)*, *328*(5985), 1534–1539. https://doi.org/10.1126/science.1187058

Bagijn, M. P., Goldstein, L. D., Sapetschnig, A., Weick, E.-M., Bouasker, S., Lehrbach, N. J., … Miska, E. A. (2012). Function, Targets, and Evolution of Caenorhabditis elegans piRNAs. *Science*.

Batista, P. J., Ruby, J. G., Claycomb, J. M., Chiang, R., Fahlgren, N., Kasschau, K. D., … Mello, C. C. (2008). PRG-1 and 21U-RNAs Interact to Form the piRNA Complex Required for Fertility in C. elegans. *Molecular Cell*. https://doi.org/10.1016/j.molcel.2008.06.002

Billi, A. C., Alessi, A. F., Khivansara, V., Han, T., Freeberg, M., Mitani, S., & Kim, J. K. (2012). The Caenorhabditis elegans HEN1 ortholog, HENN-1, methylates and stabilizes select subclasses of germline small RNAs. *PLoS Genetics*, *8*(4), e1002617. https://doi.org/10.1371/journal.pgen.1002617

Billi, A. C., Fischer, S. E. J., & Kim, J. K. (2014). Endogenous RNAi pathways in C. elegans. *WormBook*, 1–49. https://doi.org/10.1895/wormbook.1.170.1

Buckley, B. A., Burkhart, K. B., Gu, S. G., Spracklin, G., Kershner, A., Fritz, H., … Kennedy, S. (2012). A nuclear Argonaute promotes multigenerational epigenetic inheritance and germline immortality. *Nature*, *489*(7416), 447–451. https://doi.org/10.1038/nature11352

Castel, S. E., & Martienssen, R. A. (2013). RNA interference in the nucleus: roles for small RNAs in transcription, epigenetics and beyond. *Nature Reviews Genetics*, *14*(2), 100–112. https://doi.org/10.1038/nrg3355

Chuma, S., & Nakano, T. (2012). piRNA and spermatogenesis in mice. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *368*(1609), 20110338–20110338. https://doi.org/10.1098/rstb.2011.0338

Claycomb, J. M., Batista, P. J., Pang, K. M., Gu, W., Vasale, J. J., van Wolfswinkel, J. C., … Mello, C. C. (2009). The Argonaute CSR-1 and Its 22G-RNA Cofactors Are Required for Holocentric Chromosome Segregation. *Cell*. https://doi.org/10.1016/j.cell.2009.09.014

Colombani, J., Raisin, S., Pantalacci, S., Radimerski, T., Montagne, J., & Léopold, P. (2003). A nutrient sensor mechanism controls Drosophila growth. *Cell*, *114*(6), 739–749. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/14505573

Conine, C. C., Batista, P. J., Gu, W., Claycomb, J. M., Chaves, D. A., Shirayama, M., & Mello, C. C. (2010). Argonautes ALG-3 and ALG-4 are required for spermatogenesis-specific 26G-RNAs and thermotolerant sperm in Caenorhabditis elegans. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(8), 3588–3593. https://doi.org/10.1073/pnas.0911685107

Conine, C. C., Moresco, J. J., Gu, W., Shirayama, M., Conte, D., Yates, J. R., & Mello, C. C. (2013). Argonautes promote male fertility and provide a paternal memory of germline gene expression in C. Elegans. *Cell*. https://doi.org/10.1016/j.cell.2013.11.032

Das, D., & Arur, S. (2017). Conserved insulin signaling in the regulation of oocyte growth, development, and maturation. *Molecular Reproduction and Development*, *84*(6), 444–459. https://doi.org/10.1002/mrd.22806

Das, P. P., Bagijn, M. P., Goldstein, L. D., Woolford, J. R., Lehrbach, N. J., Sapetschnig, A., … Miska, E. A. (2008). Piwi and piRNAs Act Upstream of an Endogenous siRNA Pathway to Suppress Tc3 Transposon Mobility in the Caenorhabditis elegans Germline. *Molecular Cell*. https://doi.org/10.1016/j.molcel.2008.06.003

de Albuquerque, B. F. M., Luteijn, M. J., Cordeiro Rodrigues, R. J., van Bergeijk, P., Waaijers, S., Kaaij, L. J. T., … Ketting, R. F. (2014). PID-1 is a novel factor that operates during 21U-RNA biogenesis in Caenorhabditis elegans. *Genes and Development*. https://doi.org/10.1101/gad.238220.114

de Albuquerque, B. F. M., Placentino, M., & Ketting, R. F. (2015). Maternal piRNAs Are Essential for Germline Development following De Novo Establishment of Endo-siRNAs in Caenorhabditis elegans. *Developmental Cell*. https://doi.org/10.1016/j.devcel.2015.07.010

Despic, V., & Neugebauer, K. M. (2018). RNA tales - how embryos read and discard messages from mom. *Journal of Cell Science*, *131*(5), jcs201996. https://doi.org/10.1242/jcs.201996

Fouladi-Nashta, A. A., Gutierrez, C. G., Gong, J. G., Garnsworthy, P. C., & Webb, R. (2007). Impact of dietary fatty acids on oocyte quality and development in lactating dairy cows. *Biology of Reproduction*, *77*(1), 9–17. https://doi.org/10.1095/biolreprod.106.058578

Hammond, T. M. (2017). Sixteen Years of Meiotic Silencing by Unpaired DNA. In *Advances in genetics* (Vol. 97, pp. 1–42). https://doi.org/10.1016/bs.adgen.2016.11.001

Han, T., Manoharan, A. P., Harkins, T. T., Bouffard, P., Fitzpatrick, C., Chu, D. S., … Kim, J. K. (2009). 26G endo-siRNAs regulate spermatogenic and zygotic gene expression in Caenorhabditis elegans. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(44), 18674–18679. https://doi.org/10.1073/pnas.0906378106

Heestand, B., Simon, M., Frenk, S., Titov, D., & Ahmed, S. (2018). Transgenerational Sterility of Piwi Mutants Represents a Dynamic Form of Adult Reproductive Diapause. *Cell Reports*, *23*(1), 156–171. https://doi.org/10.1016/j.celrep.2018.03.015

Hosokawa, M., Shoji, M., Kitamura, K., Tanaka, T., Noce, T., Chuma, S., & Nakatsuji, N. (2007). Tudor-related proteins TDRD1/MTR-1, TDRD6 and TDRD7/TRAP: Domain composition, intracellular localization, and function in male germ cells in mice. *Developmental Biology*, *301*(1), 38–52. https://doi.org/10.1016/J.YDBIO.2006.10.046

Ishizu, H., Siomi, H., & Siomi, M. C. (2012). Biology of PIWI-interacting RNAs: new insights into biogenesis and function inside and outside of germlines. *Genes & Development*, *26*(21), 2361–2373. https://doi.org/10.1101/gad.203786.112

Iyer, S., La-Borde, P. J., Payne, K. A. P., Parsons, M. R., Turner, A. J., Isaac, R. E., & Acharya, K. R. (2015). Crystal structure of X-prolyl aminopeptidase from *Caenorhabditis elegans* : A cytosolic enzyme with a di-nuclear active site. *FEBS Open Bio*, *5*(1), 292–302. https://doi.org/10.1016/j.fob.2015.03.013

Kamminga, L. M., Luteijn, M. J., den Broeder, M. J., Redl, S., Kaaij, L. J. T., Roovers, E. F., … Ketting, R. F. (2010). Hen1 is required for oocyte development and piRNA stability in zebrafish. *The EMBO Journal*, *29*(21), 3688–3700. https://doi.org/10.1038/emboj.2010.233

Kamminga, L. M., van Wolfswinkel, J. C., Luteijn, M. J., Kaaij, L. J. T., Bagijn, M. P., Sapetschnig, A., … Ketting, R. F. (2012). Differential Impact of the HEN1 Homolog HENN-1 on 21U and 26G RNAs in the Germline of Caenorhabditis elegans. *PLoS Genetics*, *8*(7), e1002702. https://doi.org/10.1371/journal.pgen.1002702

Kanzaki, N., Tsai, I. J., Tanaka, R., Hunt, V. L., Liu, D., Tsuyama, K., … Kikuchi, T. (2018). Biology and genome of a newly discovered sibling species of Caenorhabditis elegans. *Nature Communications*, *9*(1), 3216. https://doi.org/10.1038/s41467-018-05712-5

Kasper, D. M., Gardner, K. E., & Reinke, V. (2014). Homeland security in the *C. elegans* germ line. *Epigenetics*, *9*(1), 62–74. https://doi.org/10.4161/epi.26647

Ketting, R. F. (2011). The Many Faces of RNAi. *Developmental Cell*. https://doi.org/10.1016/j.devcel.2011.01.012

Klosin, A., Casas, E., Hidalgo-Carcedo, C., Vavouri, T., & Lehner, B. (2017). Transgenerational transmission of environmental information in *C. elegans*. *Science*, *356*(6335), 320–323. https://doi.org/10.1126/science.aah6412

LaFever, L., & Drummond-Barbosa, D. (2005). Direct control of germline stem cell division and cyst growth by neural insulin in Drosophila. *Science (New York, N.Y.)*, *309*(5737), 1071–1073. https://doi.org/10.1126/science.1111410

Laurent, V., Brooks, D. R., Coates, D., & Isaac, R. E. (2001). Functional expression and characterization of the cytoplasmic aminopeptidase P of *Caenorhabditis elegans*. *European Journal of Biochemistry*, *268*(20), 5430–5438. https://doi.org/10.1046/j.0014-2956.2001.02483.x

Lopez, A. L., Chen, J., Joo, H.-J., Drake, M., Shidate, M., Kseib, C., & Arur, S. (2013). DAF-2 and ERK Couple Nutrient Availability to Meiotic Progression during Caenorhabditis elegans Oogenesis. *Developmental Cell*, *27*(2), 227–240. https://doi.org/10.1016/J.DEVCEL.2013.09.008

Malone, C. D., & Hannon, G. J. (2009). Small RNAs as Guardians of the Genome. *Cell*. https://doi.org/10.1016/j.cell.2009.01.045

Mello, C. C., Draper, B. W., Krause, M., Weintraub, H., & Priess, J. R. (1992). The pie-1 and mex-1 genes and maternal control of blastomere identity in early C. elegans embryos. *Cell*, *70*(1), 163–176. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1623520

Montgomery, T. A., Rim, Y.-S., Zhang, C., Dowen, R. H., Phillips, C. M., Fischer, S. E. J., & Ruvkun, G. (2012). PIWI associated siRNAs and piRNAs specifically require the Caenorhabditis elegans HEN1 ortholog henn-1. *PLoS Genetics*, *8*(4), e1002616. https://doi.org/10.1371/journal.pgen.1002616

Olina, A. V., Kulbachinskiy, A. V., Aravin, A. A., & Esyunina, D. M. (2018). Argonaute Proteins and Mechanisms of RNA Interference in Eukaryotes and Prokaryotes. *Biochemistry (Moscow)*, *83*(5), 483–497. https://doi.org/10.1134/S0006297918050024

Ozata, D. M., Gainetdinov, I., Zoch, A., O'Carroll, D., & Zamore, P. D. (2018). PIWI-interacting RNAs: small RNAs with big functions. *Nature Reviews Genetics*, 1. https://doi.org/10.1038/s41576-018-0073-3

Papadopoulos, S., Lonergan, P., Gath, V., Quinn, K. M., Evans, A. C., O'Callaghan, D., & Bolan, M. P. (2001). Effect of diet quantity and urea supplementation on oocyte and embryo quality in sheep. *Theriogenology*, *55*(5), 1059–1069. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11322234

Pek, J. W., Anand, A., & Kai, T. (2012). Tudor domain proteins in development. *Development*. https://doi.org/10.1242/dev.073304

Phillips, C. M., Brown, K. C., Montgomery, B. E., Ruvkun, G., & Montgomery, T. A. (2015). PiRNAs and piRNA-Dependent siRNAs Protect Conserved and Essential C. elegans Genes from Misrouting into the RNAi Pathway. *Developmental Cell*. https://doi.org/10.1016/j.devcel.2015.07.009

Rechavi, O., Houri-Ze'evi, L., Anava, S., Goh, W. S. S., Kerk, S. Y., Hannon, G. J., & Hobert, O. (2014). Starvation-Induced Transgenerational Inheritance of Small RNAs in C. elegans. *Cell*, *158*(2), 277–287. https://doi.org/10.1016/j.cell.2014.06.020

Rojas-Ríos, P., & Simonelig, M. (2018). piRNAs and PIWI proteins: regulators of gene expression in development and stem cells. *Development (Cambridge, England)*, *145*(17), dev161786. https://doi.org/10.1242/dev.161786

Roovers, E. F., Kaaij, L. J. T., Redl, S., Bronkhorst, A. W., Wiebrands, K., de Jesus Domingues, A. M., … Ketting, R. F. (2018). Tdrd6a Regulates the Aggregation of Buc into Functional Subcellular Compartments that Drive Germ Cell Specification. *Developmental Cell*, *46*(3), 285–301.e9. https://doi.org/10.1016/J.DEVCEL.2018.07.009

Seydoux, G., & Dunn, M. A. (1997). Transcriptionally repressed germ cells lack a subpopulation of phosphorylated RNA polymerase II in early embryos of Caenorhabditis elegans and Drosophila melanogaster. *Development (Cambridge, England)*, *124*(11), 2191–2201. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9187145

Shen, E.-Z., Chen, H., Ozturk, A. R., Tu, S., Shirayama, M., Tang, W., … Mello, C. C. (2018). Identification of piRNA Binding Sites Reveals the Argonaute Regulatory Landscape of the C. elegans Germline. *Cell*, *172*(5), 937–951.e18. https://doi.org/10.1016/j.cell.2018.02.002

Siddiqui, N. U., Li, X., Luo, H., Karaiskakis, A., Hou, H., Kislinger, T., … Lipshitz, H. D. (2012). Genome-wide analysis of the maternal-to-zygotic transition in Drosophila primordial germ cells. *Genome Biology*, *13*(2), R11. https://doi.org/10.1186/gb-2012-13-2-r11

Simon, M., Sarkies, P., Ikegami, K., Doebley, A.-L., Goldstein, L. D., Mitchell, J., … Ahmed, S. (2014). Reduced insulin/IGF-1 signaling restores germ cell immortality to Caenorhabditis elegans Piwi mutants. *Cell Reports*, *7*(3), 762–773. https://doi.org/10.1016/j.celrep.2014.03.056

Skvortsova, K., Iovino, N., & Bogdanović, O. (2018). Functions and mechanisms of epigenetic inheritance in animals. *Nature Reviews Molecular Cell Biology*, 1. https://doi.org/10.1038/s41580-018-0074-2

Spracklin, G., Fields, B., Wan, G., Becker, D., Wallig, A., Shukla, A., & Kennedy, S. (2017). The RNAi Inheritance Machinery of Caenorhabditis elegans. *Genetics*, *206*(3), 1403–1416. https://doi.org/10.1534/genetics.116.198812

Tang, W., Seth, M., Tu, S., Shen, E.-Z., Li, Q., Shirayama, M., … Mello, C. C. (2018). A Sex Chromosome piRNA Promotes Robust Dosage Compensation and Sex Determination in C. elegans. *Developmental Cell*, *44*(6), 762–770.e3. https://doi.org/10.1016/j.devcel.2018.01.025

Vasileva, A., Tiedau, D., Firooznia, A., Müller-Reichert, T., & Jessberger, R. (2009). Tdrd6 Is Required for Spermiogenesis, Chromatoid Body Architecture, and Regulation of miRNA Expression. *Current Biology*, *19*(8), 630–639. https://doi.org/10.1016/J.CUB.2009.02.047

Wedeles, C. J., Wu, M. Z., & Claycomb, J. M. (2013). A multitasking Argonaute: exploring the many facets of C. elegans CSR-1. *Chromosome Research*, *21*(6–7), 573–586. https://doi.org/10.1007/s10577-013-9383-7

Woglar, A., & Jantsch, V. (2014). Chromosome movement in meiosis I prophase of Caenorhabditis elegans. *Chromosoma*, *123*(1–2), 15–24. https://doi.org/10.1007/s00412-013-0436-7

Xiao, Y., Bedet, C., Robert, V. J. P., Simonet, T., Dunkelbarger, S., Rakotomalala, C., … Palladino, F. (2011). Caenorhabditis elegans chromatin-associated proteins SET-2 and ASH-2 are differentially required for histone H3 Lys 4 methylation in embryos and adult germ cells. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(20), 8305–8310. https://doi.org/10.1073/pnas.1019290108

Xu, F., Feng, X., Chen, X., Weng, C., Yan, Q., Xu, T., … Guang, S. (2018). A Cytoplasmic Argonaute Protein Promotes the Inheritance of RNAi. *Cell Reports*, *23*(8), 2482–2494. https://doi.org/10.1016/J.CELREP.2018.04.072