Regulatory small RNA pathways in the nematode *C. elegans* and the ant *T. rugatulus*

Dissertation
Zur Erlangung des Grades
Doktor der Naturwissenschaften

Am Fachbereich Biologie
Der Johannes Gutenberg-Universität Mainz

Ann-Sophie Seistrup
geboren am [redacted] in [redacted]

Mainz, 2025

Dekan: Prof. Dr. Eckhard Thines

1. Berichterstatter: Prof. Dr. René F. Ketting
2. Berichterstatter: Prof. Dr. Kathi Zarnack

Tag der mündlichen Prüfung: 21.07.2025

Page left intentionally blank

Page left intentionally blank

Abstract

Epigenetic regulation via RNA interference (RNAi) is an important method for altering gene expression and for silencing transposons; transposable genetic elements which can self-insert into the genome at various locations, thus wreaking havoc on the genome when not kept in check. RNAi functions via proteins known as Argonautes which associate with small RNAs, using these to direct them to their targets via sequence complementarity. This thesis concerns different pathways related to RNAi in the nematode *Caenorhabditis elegans*, in which RNAi has long been studied, as well as in the ant, *Temnothorax rugatulus*, in which RNAi has not priorly been studied.

In Chapter I, we show how the worm-specific Argonaute WAGO-3 is paternally inherited via the sperm in *C. elegans*, where it localizes to the newly found PEI-granule, whose phase-separation is controlled by the proteins PEI-1 and PEI-2. We show evidence to support the segregation into budding spermatids of PEI-1/2 and WAGO-3 via anchoring to fibrous body–membranous organelles (FB-MOs). We also show that WAGO-3 is important for paternal inheritance of transgenerational epigenetic memory (RNAe).

In Chapter II, we show that *T. rugatulus* expresses at least two different classes of small RNAs; miRNAs and piRNAs. We show that miRNAs are related to caste-differentiation and that the piRNA pathway, relevant for transposon silencing, is active even in rudimentary ovaries incapable of reproducing.

Finally, in Chapter III, we provide new insights into the functionalities of the three germline-expressed, worm-specific Argonautes WAGO-1, WAGO-3, and WAGO-4 in *C. elegans*. We show evidence that WAGO-1 and WAGO-4 influence the paternal ALG-3/4 pathway, possibly via interaction with the nuclear Argonaute WAGO-9/HRDE-1. WAGO-3, on the other hand, we find to influence the maternal ERGO-1 pathway, most likely also via interaction with one or more nuclear Argonautes. We find that Argonaute regulation depends on the lifestage of *C.* elegans, and we lastly show that mRNA misregulation caused by loss of WAGO-4 is remembered transgenerationally.

Page left intentionally blank

Zusammenfassung

Epigenetisches Regulierung durch RNA Interferenz (RNAi) ist eine wichtige Methode um Genexpressionzu beeinflussen und um Transposons zu unterdrücken. Transposons sind genetische Elemente die sich selbst an verschiedene Orte des Genoms einfügen können, und dabei das Genom zerstören, wenn sie nicht kontrolliert werden. RNAi funktioniert durch Proteine, sogenannte Argonauten, der mit Small RNAs interagieren und diesen benutzten, um ihre Ziele durch Sequenzkomplementarität zu finden. Dieser Dissertation betrifft verschiedene Pathways bezogen auf RNAi in der Nematode *Caenorhabditis elegans*, bei dem RNAi schon lange studiert wurde, sowie in die Ameise *Temnothorax rugatulus*, bei der RNAi nicht zuvorstudiert wurde.

In Kapitel I zeigen wir wie der wurmspezifischen Argonaut WAGO-3 paternal durch die Spermien in *C. elegans* geerbt wird. Hier lokalisiert es sich auf dem neu entdeckten PEI-Granulum, dessen Phasentrennung durch die Proteine PEI-1 und PEI-2 kontrolliert wird. Wir zeigen Beweise, die die Segregation von PEI-1/2 und WAGO-3 in budding Spermatiden durch Verankerung an Fibrous Body-Membranous Organellen (FB-MOs) unterstützen. Wir zeigen, dass WAGO-3 wichtig für paternaler Erbe vom transgenerationelles epigenetisches Gedächtnis ist.

In Kapitel II zeigen wir, dass *T. rugatulus* mindesten zwei verschiedenen Klassen von Small RNAs exprimieren; miRNAs und piRNAs. Wir zeigen, dass miRNAs mit der Kastendifferenzierung zusammenhängen, und dass das piRNA Pathway, die für Transposon-unterdrückung notwendig ist, sogar in rudimentären Eierstöcken, die sich nicht reproduzieren können, aktiv ist.

Schließlich geben wir in Kapitel III neue Einblicke in die Funktionalitäten der drei in der Keimbahn exprimierten wurmspezifischen Argonauten WAGO-1, WAGO-3 und WAGO-4 in *C. elegans*. Wir zeigen Evidenz dafür, dass WAGO-1 und WAGO-4 mit der paternalen ALG-3/4 Pathway interagieren, möglicherweise durch Interaktion mit der nuklearen Argonaut WAGO-9/HRDE-1. Anderseits finden wir, dass WAGO-3 mit der maternalen ERGO-1 Pathway interagiert, höchstwahrscheinlich auch durch Interaktion mit einen oder mehreren nuklearen Argonauten. Wir stellen fest, dass Argonaut-regulierung vom Lebensstadium von *C. elegans* abhängig ist, und zeigen schließlich, dass Fehlregulierung von mRNAs, die durch Verlust von WAGO-4 verursacht ist, transgenerationell erinnert wird.

Page left intentionally blank

Author's note

This thesis contains two published papers (chapters I and II) and some as of yet unpublished data (chapter III), whose generation has been a collaborative effort between me and several other scientists. The term "we" will be generally used throughout the thesis with no distinction of who carried out any specific task.

The chapters are listed chronologically in terms of publication, with the unpublished data coming last.

For individual contributions to the two published works, I will refer to the papers themselves. As for my own contribution to the three chapters, this is outlined below.

Chapter I – Membrane-associated cytoplasmic granules carrying the Argonaute protein WAGO-3 enable paternal epigenetic inheritance in *Caenorhabditis elegans*

This paper was first published in Nature Cell Biology in February 2022, and the full list of authors, in appropriate order, is:

> Jan Schreier, Sabrina Dietz, Mandy Boermel, Viola Oorschot, **Ann-Sophie Seistrup**, Antonio M. de Jesus Domingues, Alfred W. Bronkhorst, Dieu An H. Nguyen, Stephanie Phillis, Elizabeth J. Gleason, Steven W. L'Hernault, Carolyn M. Phillips, Falk Butter, René F. Ketting

My own contribution to the paper was the computational analysis of RNA Immunoprecipitation Sequencing (RIP-Seq) data. I have generated **Extended Figure 2** and **Extended Figure 3**.

Supervisor confirmation: _____

Chapter II – Age- and caste-independent piRNAs in the germline and miRNA profiles linked to caste and fecundity in the ant *Temnothorax rugatulus*

This paper was first published Molecular Ecology in November 2023, and the full list of authors, in appropriate order, is:

> **Ann-Sophie Seistrup**, Marina Choppin, Shamitha Govind, Barbara Feldmeyer, Marion Kever, Emil Karaulanov, Alice Séguret, Sivarajan Karunanithi, Miguel V. Almeida, René F. Ketting, Susanne Foitzik

I carried out all sequencing data analyses in collaboration with coauthors Emil Karaulanov and Sivarajan Karunanithi. I have made all figures apart from Figure 2 and have written the text with the aid of the coauthors.

Supervisor confirmation: _____

Chapter III – Deletion of worm-specific Argonautes expressed in the germline of *Caenorhabditis elegans* causes unpredicted changes in mRNA levels that can be remembered upon Argonaute reintroduction

For this chapter, I have carried out all computational analyses with the aid of the Bioinformatics Core Facilities at the Institute of Molecular Biology, Mainz.

The experimental procedures were largely carried out by me, with the following exceptions:

i) Library preparation and sequencing of all sequencing data has been carried out by the Genomics Core Facilities at the Institute of Molecular Biology, Mainz.

ii) Immunoprecipitation-mass spectrometry of *gfp::3xflag::wago-3(xf119)* was performed by [redacted]

iii) All other mass spectrometry data has been generated by [redacted]

iv) I have been aided in outcrossing, genotyping, and harvesting of worms by [redacted]

v) **Figure 4I** has been generated by [redacted]

It should also be mentioned that the project was aided by the Media Lab at the Institute of Molecular Biology, Mainz, who cast all growth plates.

Table of Contents

List of Abbreviations

Aub	Aubergine (PIWI Argonaute in *Drosophila*)
dsRNA	double-stranded RNA
endo-siRNA	endogenous small interfering RNA
exo-siRNA	exogenous small interfering RNA
FB-MO	fibrous body-membranous organelle
m^7G	7-methylguanosine
miRNA	microRNA
mRNA	messenger RNA
mRNA-Seq	poly(A)-selected RNA sequencing
ncRNA	non-coding RNA
NGM	nematode growth medium
NGS	next-generation sequencing
NPC	nuclear pore complex
nt	nucleotides
PCR	polymerase chain reaction
PGC	primordial germ cell
piRNA	PIWI-interacting RNA
P-bodies	processing bodies
PIWI	P-element-induced wimpy testis
PTM	post-transcriptional modifications
RBP	riboprotein
RIP-Seq	RNA immunoprecipitation sequencing
RISC	RNA-induced silencing complex
RNAe	RNA-induced epigenetic silencing
RNAi	RNA interference
sRNA-Seq	small RNA sequencing
TEI	transgenerational epigenetic inheritance
TMG	2,2,7-trimethyl-guanosine
UTR	untranslated region
WAGO	worm-specific Argonaute
Zuc	Zucchini (endonuclease in *Drosophila*)

Page left intentionally blank

General introduction

Transgenerational epigenetic inheritance

When the central dogma of biology was first coined by Francis Crick in the 1950s, it was generally thought that each gene would encode a protein and that all other DNA in an organism was useless, commonly referred to as junk DNA. We have since learned that this is not true. While the central dogma of biology adequately describes classical genetics – DNA is transcribed into messenger RNA (mRNA) which is later translated into a protein which can exert one or multiple functions – it fails to take into account the discovery of epigenetics.

In short, epigenetics concerns all types of genetic regulation which doesn't alter the sequence of the DNA, and, like the DNA sequence itself, these features can often be inherited to an offspring (Fitz-James & Cavalli, 2022). Furthermore, epigenetic changes caused by environmental factors can persist for several generations after exposure, a phenomenon termed transgenerational epigenetic inheritance (TEI) (Fitz-James & Cavalli, 2022). This is opposed to intergenerational epigenetic inheritance, which only lasts for one generation (Fitz-James & Cavalli, 2022).

Several types of epigenetic regulation exist. For one, DNA is wound around proteins called histones which aid in organizing the genome (Wolffe & Guschin, 2000). Several different types of post-translational modifications (PTMs) of the histones can cause the DNA to be wound tighter (into heterochromatin) or looser (into euchromatin), either disallowing or allowing transcription (Wolffe & Guschin, 2000). Because new DNA is synthesized using existing DNA as a template during replication, Readers and Writers recognizing the established histone PTMs can ensure that the histones associating with the newly synthesized DNA are modified identically (Fitz-James & Cavalli, 2022). Even in the case of mammals, where germ cells are often reprogrammed, secondary signals such as non-coding RNAs (ncRNAs), genome organization, or changes to intercellular biophysical properties, e.g. concentration changes affecting the organization and recruitment of biomolecular condensates, can drive maintenance and reestablishment of epigenetic markers (Fitz-James & Cavalli, 2022).

The DNA itself can also be methylated, which can both alter the protein binding capacity of that region (Tate & Bird, 1993), recruit histone writers (Nan et al., 1998), or drive heterochromatin formation directly (Choy et al., 2010). DNA methylation can also be reestablished during replication and inherited, although the transgenerational efficacy of this, evident though in plants, has been debated in mammals (Fitz-James & Cavalli, 2022).

Once a mRNA has been transcribed, this can also be regulated epigenetically at several levels. It's transport to the nucleus may be impaired (Suter, 2018), it may be chemically altered (Peixoto et al., 2020) or withheld in cellular locations void of ribosomes, the protein-RNA complexes (riboproteins, RBPs) responsible for translating RNA into protein (Suter, 2018), or it might be degraded altogether (Suter, 2018).

The protein itself may also be regulated via folding, cellular localization, or PTMs (Peixoto et al., 2020). While mRNA or protein level regulation can be classified as epigenetic depending on the slightly varied definition of the term, these molecules are short-lived and therefore not inherited, which is why TEI usually refers to modifications in or around the DNA, with one, important exception; RNA interference (RNAi).

RNAi was first discovered in 1998 (Fire et al., 1998), and it describes the observation that small RNAs (RNAs with lengths of 20-30 nucleotides (nt) (Ha & Kim, 2014)[1]) can interact with a family of proteins known as Argonautes and, via sequence complementarity, guide these proteins to mRNA or even to DNA where it can exert a variety of functions, generally leading to genetic repression (Ketting, 2011).

This thesis aims to help shed light on several RNAi pathways, both specific pathways in the established model organism *Caenorhabditis elegans*, and general small RNA functionalities of the ant species *Temnothorax rugatulus*, in which RNAi-like pathways have not previously been studied. To provide proper insight into these topics, this general introduction will delve into a more specific description of RNAi in general and its specifics in *C. elegans* and insects, respectively, before a more detailed description of the two model organisms, including their biology of inheritance, will be presented.

RNA interference

While the general mechanism of RNAi involves an Argonaute protein and a small RNA to guide binding, several different types of small RNAs exist, and they all bind distinct Argonaute proteins (Ketting, 2011). Different classes of small RNAs include micro RNAs (miRNAs), P-element-induced wimpy testis (PIWI)-interacting RNAs (piRNAs), and endogenous small interfering RNAs (endo-siRNAs), among others, and these all differ in their origin, biogenesis, and structure (Ketting, 2011).

This section will focus on the general characteristics of miRNAs and piRNAs before moving on to the general characteristics of the Argonaute proteins which bind them. While several differences exist between species, these will only briefly be touched upon in this section. Characteristics relevant for our two model organisms, *T. rugatulus*, and *C. elegans*, will be discussed in the following sections on RNAi in social insects and nematodes, respectively.

MICRORNAS

miRNAs are generally 21-25 nt in length, differing slightly between organisms, and they are produced from distinct loci in the genome, generally expressed somatically (Ha & Kim, 2014). miRNAs belong to the family of small RNAs defined by their necessity of the RNAse Dicer for their biogenesis, although they can also be

[1] The lengths defining small RNAs vary greatly between publications. Some classes of small RNAs may be short than 20 nt or longer than 30 nt; this range is simply indicated as a rough estimate of highest abundance of small RNAs.

produced via non-canonical pathways, including one which is Dicer-independent (Ha & Kim, 2014), although this will not be described further here. Endo-siRNAs are another example of small RNAs in the Dicer-dependant pathway; a pathway generally requiring the presence of double-stranded RNA (dsRNA) (Ketting, 2011). Double-stranded RNA can arise from different sources (Ketting, 2011), but for miRNAs, it is a general feature of the sequence of the pri-miRNA, which contains self-complementarity and can therefore form a hairpin structure (Bartel, 2009).

The function of miRNAs is to regulate gene expression, and they have been indicated in many diverse functions in different organisms, including cell differentiation, stress responses, and regulation of DNA damage pathways (Kaufman & Miska, 2010). Generally, this is obtained via post-transcriptional gene silencing, either inhibiting translation or degrading the mRNA altogether (Ha & Kim, 2014). The Argonautes which bind miRNAs can cleave mRNAs when the miRNA and mRNA bind with high sequence complementarity, although it is more common in animals that repression happens via mRNA destabilisation or inhibition of translation (Bartel, 2009). Some Argonautes binding other families of small RNAs are fully dependent on non-cleavage mechanisms, as they are not catalytically active (Ketting, 2011).

The biogenesis of miRNAs starts with RNA polymerase II, the same polymerase responsible for transcribing mRNAs. As such, pri-miRNAs share the features of mRNAs, having a 7-methylguanosine (m^7G) cap structure and a poly(A) tail (Ha & Kim, 2014). While miRNA genes can exist in many different genetic contexts, both in the introns and exons of other genes, it is not uncommon for many to exists close together in proximity (Ha & Kim, 2014). This close proximity has functional relevance, with the complex recognising a miRNA hairpin for processing, the microprocessor, being more likely to recognise a suboptimal hairpin if it is next to a more optimal hairpin. It has been suggested that the microprocessor may recruit more microprocessors to the same area (Fang & Bartel, 2020).

On its journey from pri-miRNA to mature miRNA, the miRNA must undergo two processing steps, both carried out by RNAse type III endonucleases (Ha & Kim, 2014). The first processing step occurs in the nucleus, and the responsible RNAse is called Drosha, which is aided by the essential cofactor Pasha (DGCR8 in humans and mice) (Ha & Kim, 2014). Before processing by Drosha, pri-miRNA typically consist of a cap, a 5'-single stranded RNA moiety, a stem of 33-35 nt and a terminal loop, as well as a 3'-single stranded moiety and the poly(A) tail. After processing, the RNA, now termed pre-miRNA, consists solely of the stem-loop with a 3'-overhang of 2 nt (Ha & Kim, 2014).

After processing by Drosha, the pre-miRNA is transported out of the nucleus to the cytosol, where the 2 nt 3'-overhang characteristic of Drosha-processing is recognized by the RNAse Dicer. Dicer cleaves the stem loop, completing the processing of the mature miRNA which can then be taken up by an Argonaute protein (Ha & Kim, 2014). Because the stem loop contains two strands with partial-to-complete complementarity,

two miRNAs are formed. These are usually denoted with the suffixes '-5p' or '-3p', depending on whether they were derived from the 5'-end or the 3'-end of the pre-miRNA. Typically, however, one of the two is much more prevalent and more biologically active (Ha & Kim, 2014).

In some organisms, uptake of the miRNA by the Argonaute is highly specific, with some Argonautes binding distinct subsets of miRNAs and other Argonautes binding other subsets. In yet other organisms, Argonautes show no such preferences (Ha & Kim, 2014).

When the mature miRNA has been formed, it associates with the Argonaute and other accessory proteins in order to form what is known as the RNA-induced silencing complex (RISC). As described, and central to the concept of RNAi, the miRNA guides the RISC to its targets. However, for miRNAs, targets are not defined by complementarity to the whole sequence (Bartel, 2009). While animal miRNAs are usually ~22 nt in length, the nucleotides at positions 2-7 show much higher importance for target recognition, and this sequence is called the seed sequence (Bartel, 2009). Although perfect complementarity is not required within the seed sequence, seed matching is the most reliable target predictor, with inclusion of some other position for secondary scoring common to many available miRNA target prediction algorithms. Indeed, predicting the targets of a miRNA via *in silico* methods is a quite cumbersome task, due to the many allowances of infidelity within the recognition sites (Bartel, 2009).

This is slightly different in plants, however. While all the mechanisms described above are common to miRNAs in many different animals, several differences exist in plants. For one, while Drosha and Pasha are highly conserved in animals, they do not exist in plants, which rely only on Dicer for miRNA processing (Ha & Kim, 2014). Furthermore, while animal-like repression mechanisms had been found to operate in plants (Fabian et al., 2010), using full sequence complementarity as a targets predictor is highly reliable in plants (Bartel, 2009).

Generally, miRNAs are also cooperative, and proper mRNA repression often requires binding of several miRNAs, either the same or different ones, within the 3'-UTR of a target gene (Fabian et al., 2010).

A final remark to be made for the conservation of miRNAs as a way of controlling gene expression, is the fact that not only miRNAs and their accessory proteins are highly conserved, so are the mRNAs which they regulate (Bartel, 2009). miRNA-RISCs generally bind the mRNAs which they repress in their 3'-untranslated regions (UTRs), and conservation of these regions have been observed (Bartel, 2009). Simultaneously, it has been shown that non-conservation of 3'-UTRs generally occurs in genes whose expression differ from their apparent cognate miRNA, i.e. when miRNAs and mRNAs are expressed in the same tissue, they tend to co-evolve, and when they are expressed in different tissues, they more frequently do not (Bartel, 2009).

PIWI-INTERACTING RNAS

While miRNAs have been shown to function in silencing of protein coding genes and thus impact protein homeostasis, the function of piRNAs is largely focused on silencing of transposons.

Transposons, also known as transposable elements or, more colloquially, as jumping genes, are genetic elements that possess the ability to self-insert into various location in the genome, effectively 'jumping around' in the DNA (Kumar, 2020). Many subtypes of transposons exist, although they are often classified into two major groups: type I retrotransposons and type II DNA transposons, defined by the way in which they self-insert. The former employ a reverse transcriptase to transpose via an RNA intermediate, while the latter make use of DNA as an intermediate (Kumar, 2020). Although the action of transposons can be lethal when the insertion happens into vital genes – and, indeed, transposon activity has been reported in human cancers highlighting their detrimental effects even in somatic tissues (Kumar, 2020) – DNA sequences derived from transposon activity are commonly found in all genomes. While the *C. elegans* genome, for instance, is only predicted to have 12% DNA derived from transposons (Bessereau, 2006), 46% of the human genome is estimated to have been derived from transposons (Kumar, 2020), although this is not to say that such a high amount of either genome contains active transposons in their current state.

Due to the possibly detrimental effects that transposon activity could have in the germline, it is unsurprising that piRNAs have been shown to be highly active in the germline and during meiosis (Ketting, 2011). In fact, while it is speculated that piRNAs were once present in both somatic cells and the germline, many animals have lost them from somatic tissue (Czech et al., 2018). However, while inhibiting the ability of transposons to wreak havoc on the cells responsible for reproduction is certainly important, this is not the only reported function of piRNAs in the germline (Ketting, 2011).

For one, the piRNA pathway has been shown to be involved in reestablishment of DNA methylation in primordial germ cells (PGCs), proving the function of RNAi as a trigger for maintenance of epigenetically inherited markers (Ketting, 2011). The piRNA pathway has also been shown to be involved in *de novo* DNA methylation, showing that piRNAs can work in the nucleus to both maintain epigenetic markers or to regulate and alter them in a context-dependent manner (Czech & Hannon, 2016). Changes to the piRNA machinery has also been shown in different organisms to impact meiosis, oocyte maturation, or spermatogenesis, although it cannot be ruled out that these effects are secondary causes of the role piRNAs play in transposon silencing (Ketting, 2011).

Like miRNAs, piRNAs are transcribed by RNA polymerase II, and they are present in distinct genomic loci, commonly referred to as piRNA clusters (Czech & Hannon, 2016). Unlike miRNAs, however, piRNA precursors are single-stranded and their formation is not dependent on Dicer (Czech & Hannon, 2016). piRNA clusters are mostly unistranded, meaning that they produce a piRNA precursor from only one of the

two DNA strand. Transcribed by RNA polymerase II, these precursors are canonical transcripts with 5'- m^7G caps and poly(A) tail. In the germline of *Drosophila*, however, most clusters are dual-stranded, producing piRNA precursors from both strands. While also transcribed by RNA polymerase II, these transcripts are noncanonical and lack both a 5'-cap structure and a poly(A) tail (Czech et al., 2018).

Because most piRNA precursors resemble mRNA transcripts, it is unclear exactly how they are recognized for processing, although several factors involved in the process have been found (Czech & Hannon, 2016). The production of mature piRNAs can furthermore be divided into two distinct pathways. The primary pathway involves RNA polymerase II and takes place in the nucleus, as explained above (**Figure 1A)**. The secondary pathway, known as the ping-pong cycle, is an amplification step which increases the accumulation of piRNAs targeting active transposons and it takes place in specialized structures in the cytoplasm (**Figure 1C**) (Czech & Hannon, 2016).

There is no unified naming convention for the proteins involved in either of the piRNA biogenesis pathways, so for simplicity, they will be described below using *Drosophila* terminology, as also summarized in **Figure 1**. While many accessory proteins are known for the mechanisms, we will focus only on the main ones. As is true for all mechanisms described in this thesis, differences exist between organisms and many both known and unknown factors help facilitate binding, recognition, and function.

Processing of piRNA precursors happens in subcellular structures in the cytoplasm. Such structures are typically denser than their surroundings, rich in RNA, and capable of including or excluding certain factors (Fitz-James & Cavalli, 2022). They are also common for different types of RNAi, with even miRNAs having been found in processing bodies (P-bodies), although they are thought to be only stored in these rather than exacting any silencing process (Ketting, 2011).

In the case of *Drosophila* piRNA processing, the subcellular structures in question are nuage in the germ cells and Yb-bodies in follicle cells (Czech & Hannon, 2016). The piRNA precursor, which may be a readthrough containing several piRNAs, is recognized in the nuage or Yb-body by a machinery including the endonuclease Zucchini (Zuc). Zuc processes the 5'-end of the piRNA intermediate, usually cleaving right before an uracil, causing piRNAs to have a strong bias for U at the 5'-end (**Figure 1A**) (Czech & Hannon, 2016).

Some, but not all, piRNA transcripts require subsequent trimming of the 3'-end by the exonuclease Trimmer (**Figure 1A**). The exception to this need for trimming happens in the case where Zuc binds and cleaves the piRNA precursor in such a way that it simultaneously generates the 3'-end of one piRNA and the 5'-end of another. Such cleavage products are referred to as phased piRNAs (Czech et al., 2018).

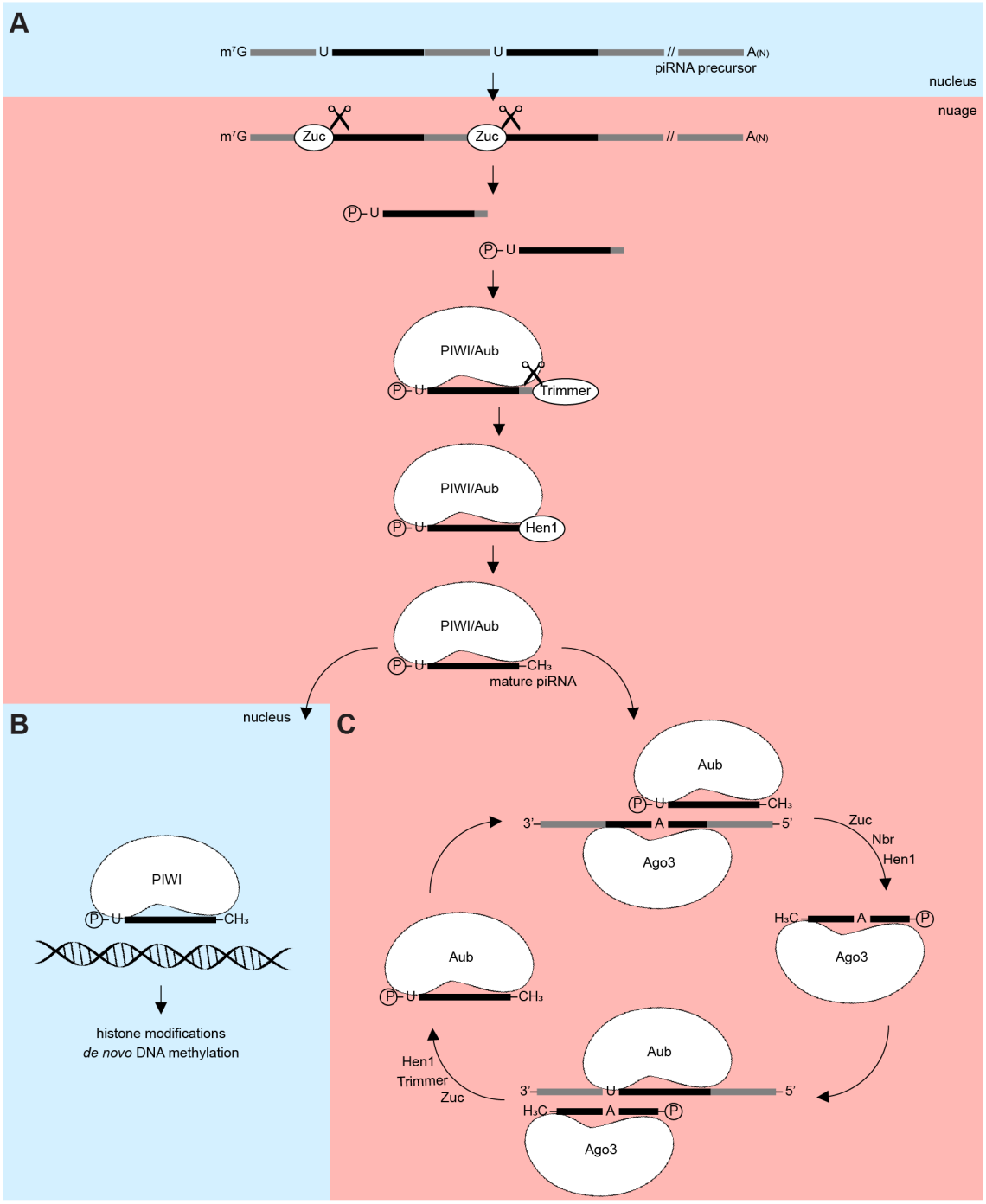


Figure 1: Primary and secondary piRNA production in* D. melanogaster *germ cells. A *Schematic depicting the key proteins involved in primary piRNA production, starting with an RNA polymerase II-generated transcript and ending with a mature piRNA bound to either of the Argonaute proteins PIWI or Aubergine (Aub). Zuc = Zucchini.* ***B*** *Function of the nuclear Argonaute protein PIWI.* ***C*** *Schematic depicting the key elements of the piRNA amplification cycle known as the ping-pong cycle in* Drosophila melanogaster*. This is located in nuage and involves the two Argonaute proteins Aubergine (Aub) and Ago3. Nbr = Nibbler.*

Common to both phased and non-phased piRNAs is the final methylation of the 3'-end by the enzyme Hen1, which takes place after binding to an Argonaute. In *Drosophila*, primary piRNAs – i.e. piRNAs generated by the process just described – bind one of two Argonautes, PIWI or Aubergine (Aub). PIWI is a nuclear Argonaute, meaning that it must return to the nucleus after piRNA loading where it can then guide histone modifications or *de novo* DNA methylation (**Figure 1B**). Aub, on the other hand, is cytoplasmic and remains in the nuage or Yb-body, where it can silence active transposons and where it partakes in the ping-pong cycle (Czech & Hannon, 2016).

Once Aub is loaded with a piRNA, it can bind a transposon transcript via sequence complementarity. Rather than degrading the transposon directly, this dsRNA complex is recognized by a machinery including another Argonaute, Ago3. Zuc then cleaves the transposon, forming a new piRNA intermediate already bound by Ago3. The exonuclease Nibbler can then trim the 3'-end of the Ago3-bound RNA after which Hen1 can methylate it, thus forming a mature piRNA (**Figure 1C**) (Czech & Hannon, 2016). Loaded Ago3 is then free to recognize a new transposon transcript which will recruit proteins including Aub to the site. The action of Zuc, Trimmer, and Hen1 will complete the cycle, forming a new mature piRNA loaded by Aub (**Figure 1C**) (Czech & Hannon, 2016).

Because piRNAs loaded by Aub and Ago3 are shifted 10 nt in relation to one another, Ago3-bound piRNAs do not have the 5'-U bias as seen for Aub- and PIWI-bound piRNAs. Instead, Ago-3 piRNAs are characterized by a strong bias for adenine at position 10 (Czech & Hannon, 2016). While the 5'-U and 10A biases observed by ping-pong pairs are a feature of their biogenesis, it is also reflected in the binding pockets of their respective Argonautes, with both *in silico* analyses and structural data supporting the preferential binding of one or the other in different Argonautes of the same species (Czech & Hannon, 2016).

ARGONAUTES

Argonaute proteins can generally be divided into three classes; the AGO class, which interacts with miRNAs and endo-siRNAs, the PIWI class, which interacts with piRNA, and the WAGO class, which is worm-specific and will be discussed further in the section about RNAi in nematodes (Ha & Kim, 2014). A fourth class, less commonly discussed and likewise not further described here, is the *Trypanosoma* AGO class found in certain protozoa (Wu et al., 2020). Also not discussed here are the three classes of prokaryotic Argonautes, only one of which fully resembles the eukaryotic ones in structure, although conservation of general architecture and function is generally high between all Argonautes (Wu et al., 2020).

The following description of Argonautes refers specifically to the case in eukaryotes.

Argonaute classification is based on sequence homology, yet all classes share the same general structure. This consists of an N-terminal region and the three globular domains, PAZ, MID, and PIWI, always ordered

thusly from N-terminus to C-terminus (Wu et al., 2020). The Argonaute folds into a bilobed structure, with one lobe comprising the N-terminal region and the PAZ domain and the other comprising the MID and PIWI domains. Two linkers exist within this structure, one on either side of the PAZ domain. These are essential for the structure and function of the protein, as they can drive conformation changes upon small RNA binding (Wu et al., 2020).

The MID domain is essential for small RNA binding as it contains a binding pocket for its 5'-end. This interaction is driven by a highly conserved tyrosine residue (Wu et al., 2020). It is also the structure of the MID domain which dictate the preference for 5'-U or 10A in PIWI class Argonautes (Czech & Hannon, 2016). Less critical than the MID domain interaction, the PAZ domain anchors the 3'-end of the small RNA. The N-terminal region plays a role in unwinding and loading of the small RNA (Wu et al., 2020). For Argonautes which are catalytically active, as mostly seen in plants, this cleavage activity resides in the PIWI domain (Wu et al., 2020).

Argonautes and small RNAs are at the core of the RISC, but both the function hereof in silencing and the loading of small RNAs into Argonautes is highly dependent on accessory factors, with small RNA biogenesis and their loading into Argonautes often being linked processes (Wu et al., 2020).

The conformation of the Argonaute itself is also very important. In *Drosophila*, two heat-shock proteins, Hsp70 and Hsp90, help change the conformation of unloaded Ago2 from a closed state to one that is open and can accommodate a small RNA into its binding pocket (Tsuboyama et al., 2018). In the plant *Arabidopsis thaliana*, conformational changes of a N-terminal moiety termed the N-coil driven by small RNA loading determines whether the protein will be marked for degradation (Bressendorff et al., 2023).

The Argonaute proteins are also highly prone to PTMs. They contain several sites for phosphorylation, hydroxylation, and ribosylation, which play a role in stability, localisation, and small RNA binding affinity (Ha & Kim, 2014; Wu et al., 2020). Both phosphorylation and hydroxylation of human AGO2 at certain sites has for instance been shown to regulate affinity for P-bodies, effectively changing the propensity for silencing via localisation changes (Ha & Kim, 2014). Argonautes are also subject to ubiquitylation, and their abundance is regulated via the ubiquitin–proteasome degradation system (Wu et al., 2020).

As discussed, some Argonautes bind small RNAs with little distinction apart from their general class, whereas others bind subsets stemming from distinct loci or sharing distinct features (Czech & Hannon, 2016; Ha & Kim, 2014). While not all is known about the mechanisms that drive and regulate proper RNAi function, it should be clear that this a highly complex system influenced by Argonaute structure and PTMs, small RNA structure and biogenesis, accessory proteins, and cellular localisation of all players in the system.

Social insects and RNAi

When it comes to RNAi in insects, the fruit fly *Drosophila melanogaster* has been the most studied, with many of the mechanisms mentioned in the previous section having been described in this organism. In fact, piRNAs were first discovered in *D. melanogaster* (Czech & Hannon, 2016), as was the ping-pong amplification loop (Brennecke et al., 2007; Gunawardane et al., 2007).

D. melanogaster has five genes encoding Argonaute proteins, although the number of Argonautes varies greatly between different insect species (Zhu & Palli, 2020). The core RNAi machinery genes are highly conserved in insects (Palli, 2023), but they are also found to be highly subjected to duplications, with a type of cockroach (*Blattella germanica*) thus far being the only exception to having duplications in RNAi pathway genes (Zhu & Palli, 2020). The high duplication rate of RNAi-related genes is generally thought to be advantageous, increasing RNAi efficiency via newly acquired functionalities or sub-functionalities (Zhu & Palli, 2020). In insects, the Argonautes themselves have also been found to undergo frequent evolutionary expansions resulting in their functional divergence (Zhu & Palli, 2020).

After its discovery in *D. melanogaster*, RNAi was soon found in many other insects including several types of mosquitoes, beetles, a moth, the German cockroach, the migratory locust, and the brown planthopper (Palli, 2023). The first social insect – i.e. an insect living in a colony and dividing labour between different castes – in which RNAi was described, was the honey bee *Apis mellifera* (Palli, 2023). The pathways involving miRNAs, endo-siRNAs and piRNAs have all been found in insects, and they have been found to play crucial roles in protection against viruses and transposons as well as in regulation of gene expression, as in other organisms (Zhu & Palli, 2020).

Thus far, the description of RNAi outlined here has been centred around endogenous pathways. However, exogenous small interfering RNAs (exo-siRNAs) also exists, and these have been a large focus in insect RNAi research due to the function these can play in pest control (Palli, 2023). Generally, exogenous RNAi pathways are a response to protect an organism when non-self RNA enters the cell, which for instance happens during viral infections. These pathways can be triggered via dsRNA feeding. Nonetheless, the efficacy of this varies largely between insect species due to large variations in proteins involved in uptake and transport as well as the frequent breakdown of dsRNA via nucleases (Palli, 2023). Injection rather than feeding has been shown to improve the RNAi response to dsRNA in some insects, and research is being conducted to improve the efficacy via feeding, as low and variable efficiency have meant that only few currently available pesticides are based on RNAi (Palli, 2023). Pest control, however, is not a subject of interest in this thesis, nor are exogenous RNAi pathways, and neither will be discussed further. Rather, this work focuses on elucidating various aspects of endogenous RNAi pathways.

A. mellifera, as the first and most studied social insect model organism for RNAi, has been shown to have functional RNAi pathways in both embryos, larvae, pupae, and adults (Brutscher & Flenniken, 2015). Small RNAs, both miRNAs and piRNAs, have been shown to be differentially expressed in the different reproductive tissue of *A. mellifera* (ovaries, spermatheca, semen, fertilised and unfertilised eggs, and testes), with fewer piRNAs in semen than in eggs or ovaries (Watson et al., 2022). Furthermore, *A. mellifera* sex is determined in part by certain genomic loci, which are partially targeted by piRNAs (Watson et al., 2022). In another social insect, the jewel wasp *Nasonia vitripennis*, small RNAs found in the eggs are crucial for sex determination (Verhulst et al., 2010). These findings suggest that RNAi plays a role in sex determination in social insects. The differences in miRNA and piRNA populations in male versus female sex organs and reproductive tissues also implicate RNAi in the parent-of-origin effect, where epigenetic factors are inherited differentially between paternal and maternal lineages (Watson et al., 2022).

Also in *A. mellifera*, miRNAs have been shown to be differentially expressed during ovarian activation in queens, and they have been suggested to have a function in oviposition, i.e. the physical process of laying eggs (Chen et al., 2017). miRNAs have also been shown to have a function in ovarian activation of *A. mellifera* workers, a process only happening in the absence of a queen (Macedo et al., 2016).

Apart from their role in sex determination and sex-specific inheritance, miRNAs have also been shown to differ between different castes of the same social insect in both termites and ants (R. Bonasio et al., 2010; Matsunami et al., 2019). In termites, this has also been shown for piRNAs (Elsner et al., 2018; Post et al., 2023).

Because piRNAs are more abundant and transposons, as a result, less expressed in reproductive termites who can live to be many times the age of non-reproductive termites, the piRNA pathway has been proposed to be a key factor in regulating caste-specific longevity (Elsner et al., 2018).

Nematodes and RNAi

Nematodes, the phylum of worms encompassing roundworms and eelworms, contain machineries pertaining to the miRNA, piRNA, and siRNA (both exo-siRNA and endo-siRNA) pathways. More interestingly, the last category of small RNAs thus far not described here has only been observed in nematodes. These small RNAs are produced by secondary pathway, amplifying the response of piRNAs and siRNAs and they bind the worm-specific clade of Argonautes, WAGOs (Hoogstrate et al., 2014). WAGO proteins and WAGO-class small RNAs have been found in many different species of nematodes. Simultaneously, nematodes don't possess the ability to amplify piRNAs via the ping-pong cycle, instead relying only on the WAGO pathway for amplification (Hoogstrate et al., 2014).

C. elegans, which was the first organism in which RNAi was discovered (Fire et al., 1998), is the most extensively studied of all the nematodes. While differences have been reported in the particulars of the nematode pathways (Hoogstrate et al., 2014), *C. elegans* is most commonly used as the reference for nematode RNAi. It is also one of the model organisms used in this thesis. For these reasons, this section will focus solely on *C. elegans* RNAi pathways.

While all branches of RNAi present in *C. elegans* have been studied, miRNAs and exo-siRNAs will not be described further here. Nonetheless, it is worth noting that the structure and function of *C. elegans* miRNAs is similar to the one described for other animals and it binds one of the Argonautes ALG-1, ALG-2, or ALG-5 in *C. elegans* (**Figure 2** and **Figure 3**) (Ambros & Ruvkun, 2018). Also worth mentioning, is that although miRNAs are not generally interconnected with other RNAi pathways, it had been reported that at least one miRNA possesses the ability to activate the secondary small RNA WAGO pathway (Corrêa et al., 2010).

The number of Argonaute genes differ greatly between species. Humans, for instance, encode eight functional Argonaute proteins and fission yeast encode only one. The *C. elegans* genome codes for 27 Argonaute genes, which is the most of all organisms thus far described, and 19 of these have been demonstrated to be expressed (**Figure 2**) (Seroussi et al., 2023). Most of these, as indicated in **Figure 2**, are of the WAGO family and, although much research has gone into these, the full extent of how much their pathways interconnect is not yet fully understood (Liu et al., 2023; Seroussi et al., 2023).

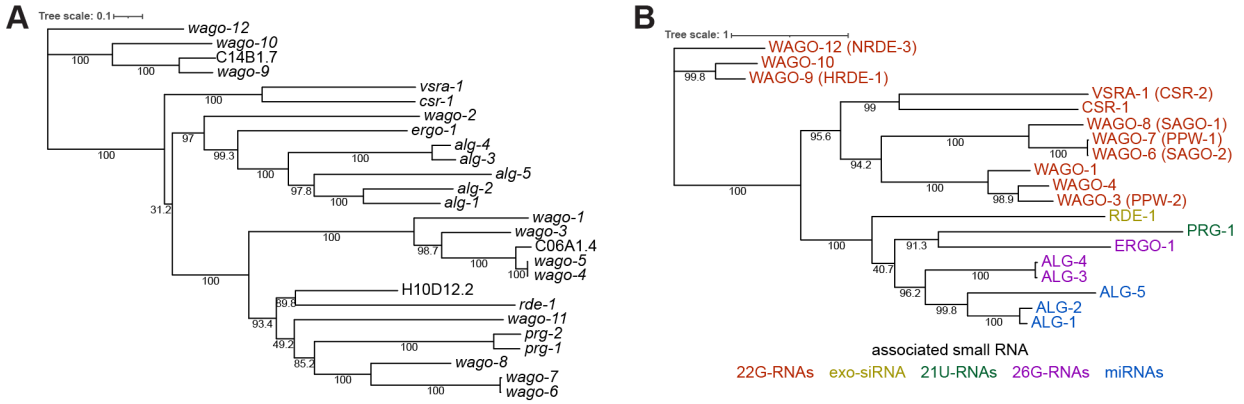
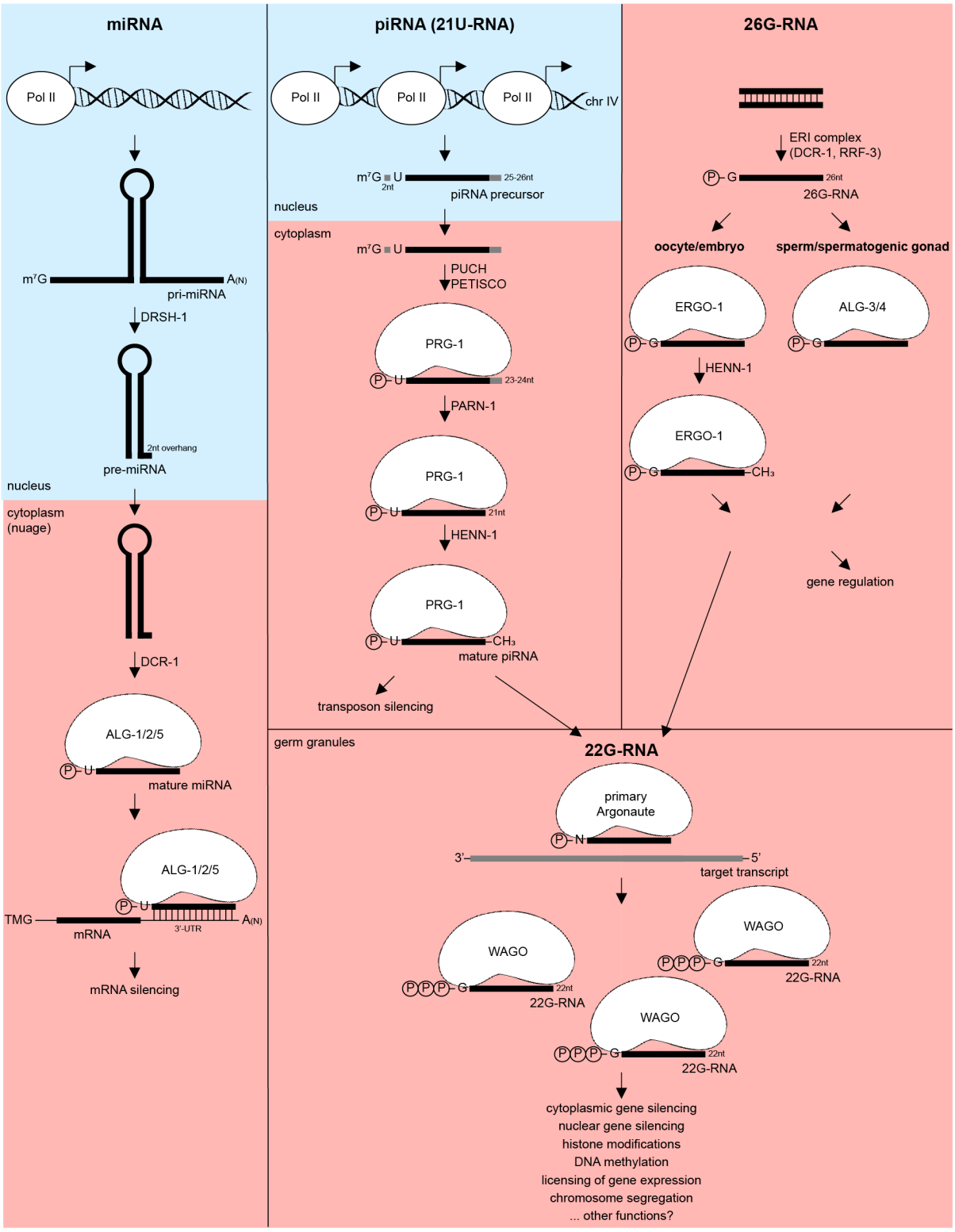


Figure 2: Phylogenetics of the* C. elegans *Argonaute proteins. ***A*** *Phylogenetic tree constructed from the full DNA sequences including UTRs of all* C. elegans *Argonaute genes.* ***B*** *Phylogenetic tree constructed from the protein sequences of the 19 Argonaute proteins with proven expression. Colours represent which class of small RNAs each Argonaute binds. Alternative names given in parentheses.*
Multiple sequence alignments conducted using Clustal Omega (Larkin et al., 2007; Madeira et al., 2024), trees constructed with IQ-TREE (Trifinopoulos et al., 2016) and visualised using (Letunic & Bork, 2024).

The *C. elegans* small RNA pathways are summarized in **Figure 3**. The primary and secondary small RNAs are described in more detail below. Apart from the small RNA pathways described here, it has been proposed

that *C. elegans* might express other classes of less abundant RNAs, although the role of such low-abundance RNAs is unclear (Knittel et al., 2024).



Legend on next page

***Figure 3: small RNA pathways of* C. elegans.** Pri-miRNA are transcribed by DNA polymerase II (Pol II) and form a hairpin structure. They are processed by the nucleases DRSH-1 (Drosha homologue) and DCR-1 (Dicer homologue) before they are loaded into one of the Argonautes ALG-1, -2, or -5. miRNAs silence mRNAs via binding in the 3'-UTR. Precursors to piRNAs (21U-RNAs) are also transcribed by Pol II and require processing of the complexes PUCH and PETISCO as well as the proteins PARN-1 and HENN-1 (Hen1 homologue) for maturation. They bind the Argonaute PRG-1. 26G-RNAs are transcribed from dsRNA via the ERI complex and bind one of the Argonautes ERGO-1, ALG-3, or ALG-4. 22G-RNAs are formed from primary Argonautes binding either 21U-RNAs or 26G-RNAs and they associate with the WAGO class of Argonautes.

Primary small RNA pathways in *C. elegans*; 21U-RNAs and 26G-RNAs

Two types of primary small RNAs exist in *C. elegans*. The first are piRNAs, although these are generally referred to by the name 21U-RNAs in *C. elegans* due to their strong bias for uracil at the 5'-position and their distinct length of 21 nt (Almeida, Andrade-Navarro, et al., 2019). While 21U-RNAs, like the piRNAs of other species, are transcribed by RNA polymerase II and the 21U-RNA-precursors have standard m^7G-caps, the cap structure has been proposed as a mechanism for distinguishing between 21U-RNAs and mRNAs in *C elegans* (Podvalnaya et al., 2023), a feat which remains elusive in other organisms. This is due to the fact that most *C. elegans* mRNAs undergo trans-splicing, where a common sequence known as a splice leader replaces the 5'-end during splicing. As a result of this process, most *C. elegans* mRNAs have 2,2,7-trimethyl-guanosine (TMG) caps (Blumenthal, 2012).

Most 21U-RNAs are transcribed from a locus on chromosome IV with the exception of a less abundant type of 21U-RNAs – termed type II 21U-RNAs – which are derived from capped small RNAs in promoter regions but are also transcribed by RNA polymerase II (Almeida, Andrade-Navarro, et al., 2019). Notably, the *C. elegans* genome does not encode a homologue of Zuc, the endonuclease responsible for processing the 5'-ends of piRNAs (Almeida, Andrade-Navarro, et al., 2019). Instead, two coupled complexes have been identified as essential for recognising and processing the 5'-ends; PUCH (Podvalnaya et al., 2023) and PETISCO (Cordeiro Rodrigues et al., 2019). After cap removal and 5'-end trimming by these complexes, 21U-RNAs can be loaded by the Argonaute PRG-1 and have their 3'-ends processed in a process akin that described for *D. melanogaster*. Whereas *D. melanogaster* relies on the exonuclease Trimmer (or Nibbler, in the case of secondary piRNAs) and the RNA methyltransferase Hen1, *C. elegans* makes use of the homologous proteins PARN-1 and HENN-1 to carry out the functions of 3'-end trimming and methylation (Almeida, Andrade-Navarro, et al., 2019).

As described, there is no ping-pong loop for 21U-RNAs, but they can trigger production of secondary small RNAs via the mechanism described in the next subsection. Conversely, direct targeting of mRNAs by PRG-1-21U-RNAs broadly follows miRNA-recognition rules, where matching to the seed sequence at position 2-8 shows more importance than matching at any other position (Almeida, Andrade-Navarro, et al., 2019).

The other type of primary small RNAs found in *C. elegans* are endo-siRNAs known as 26G-RNAs because they are commonly around 26 nt long and have a 5'-bias for guanine. 26G-RNAs are not produced by RNA polymerase II, but instead by an RNA-dependent RNA polymerase (RdRP) called RRF-3 (Almeida, Andrade-Navarro, et al., 2019).

RRF-3 is a part of the ERI complex which also includes the Dicer homologue DCR-1. The ERI complex recognizes dsRNA and use the nuclease activity of DCR-1, the polymerase activity of RRF-3 as well as other accessory proteins to form a mature 26G-RNA (Almeida, Andrade-Navarro, et al., 2019). The 26G-RNAs can then be taken up by one of two Argonaute proteins in a tissue-specific manner. In oocytes and embryos, ERGO-1 is responsible for 26G-RNA uptake; in the spermatogenic germline, either of the Argonaute proteins ALG-3 or ALG-4 will take up the 26G-RNA. ALG-3 and ALG-4 are not only highly homologous, but they also appear to be redundant, usually referred to simply as ALG-3/4 (Almeida, Andrade-Navarro, et al., 2019).

ALG-3/4-bound 26G-RNAs do not require any further processing, but ERGO-1-bound 26G RNAs are methylated at the 3'-end by HENN-1, much like 21U-RNAs (Almeida, Andrade-Navarro, et al., 2019). The two 26G-RNA pathways furthermore diverge in the way in which they bind their target mRNAs. Unlike 21U-RNAs, 26G-RNAs bind their targets via full sequence complementarity, which enables an easier genome-wide search for the targets which they bind. ALG-3/4-bound 26G-RNAs have been shown to be enriched at either 5'- or 3'-end of their target mRNAs (Conine et al., 2010; Seroussi et al., 2023), but specifically mRNAs against which 26G-RNAs were produced at the 5'-end showed high upregulation upon 26G-RNA accumulation (Conine et al., 2010). Simultaneously, distinct subsets of ALG-3/4-bound 26G-RNAs have been shown to either cause upregulation of downregulation of target mRNAs, dependent on the secondary Argonaute with which they associate (Conine et al., 2013). ERGO-1-bound 26G-RNAs, on the other hand, specifically do not map to the 5'-end of their target genes (Vasale et al., 2010) and they are generally responsible for downregulation of pseudogenes, repeat sequences, long non-coding RNAs and gene duplications (Almeida, Andrade-Navarro, et al., 2019).

Consistent with their appearance in either male or female germlines, loss of ERGO-1 or ALG-3/4 have been associated with defects in spermatogenesis or oogenesis and male- or female-specific inheritance (Almeida, Andrade-Navarro, et al., 2019).

WORMS-SPECIFIC ARGONAUTES AND 22G-RNAS

The secondary small RNAs produced from 21U-RNAs or 26G-RNAs are around 22 nt long and have a 5'-bias for guanine and are thusly called 22G-RNAs. They are transcribed by RdRPs in the cytoplasm and require no further 5'-end processing. For this reason, they can be distinguished from 21U-RNAs and 26G-RNAs in that

they carry a triphosphate at the 5'-end, whereas the primary small RNAs carry monophosphates as a result of the cleavages they must partake in (Ketting, 2011).

Two RdRPs are responsible for 22G-RNA biogenesis in response to 21U-RNAs: RRF-1 and EGO-1. These two appear to be redundant in this role, but only the former, RRF-1, has been implicated in 22G-RNA biogenesis driven by 26G-RNAs (Almeida, Andrade-Navarro, et al., 2019). 22G-RNAs are transcribed from RNAs brought into distinct subcellular locations known as mutator foci (Phillips et al., 2012), presumably directed there by their association with a primary small RNA RISC. Apart from the RdRPs, several proteins of the mutator complex exist within the mutator foci and play a role in directing 22G-RNA biogenesis and Argonaute loading (Almeida, Andrade-Navarro, et al., 2019; Phillips et al., 2012). At least for the secondary 22G-RNA pathway initiated by exo-siRNA, this involves the endonuclease RDE-8, which cleaves the dsRNA formed by the small RNA and the target (Tsai et al., 2015) and subsequent addition of multiple non-templates uridine-guanosine dinucleotides at the 3'-end of the target RNA, known as an poly(UG), or pUG, tail, by the ribonucleotidyltransferase MUT-2/RDE-3 (Shukla et al., 2020). The pUG tail is recognized by RdRPs, and its addition is necessary for proper initiation of the 22G-RNA response, possibly in response to all primary triggers (Shukla et al., 2020).

After transcription, a 22G-RNA can associate with an Argonaute of the WAGO class. As described, *C. elegans* has many functional WAGOs (**Figure 2**), some of them nuclear and some of them cytoplasmic with no distinction of 21U-RNA-derived 22G-RNAs or 26G-RNA-derived 22G-RNAs doing only one or the other (Almeida, Andrade-Navarro, et al., 2019). Central to all of the WAGOs, however, is that they are not catalytically active and most likely cannot cleave their targets (Ketting, 2011). This thesis is centred around the cytoplasmic branch of Argonaute proteins containing WAGO-1, -3, and -4, which are all present in the germline and will be described in the following subsection. One other interesting WAGO to point out, however, is CSR-1.

While most WAGOs are thought to have a role in silencing of genes and/or transposons, another function entirely has been proposed for CSR-1. CSR-1 is present in the germline and associates with 22G-RNAs targeting germline-expressed genes, but it is independent of mutator complex proteins (Almeida, Andrade-Navarro, et al., 2019). It is also not associated with gene repression – rather, CSR-1 targets expressed genes and has been suggested to directly activate their expression (Seth et al., 2013; Wedeles et al., 2013). Furthermore, while the bulk of CSR-1 is found in the cytoplasm, this Argonaute has also been shown to interact directly with chromatin and it has been indicated as having a function in chromosome segregation (Claycomb et al., 2009). The proper loading, and therefore the proper function, of CSR-1 is dependent on the nucleotidyltransferase CDE-1 which polyuridylates the 3'-ends of CSR-1 22G-RNAs, destabilizing the 22G-RNA and thus keeping CSR-1 from being overly active (van Wolfswinkel et al., 2009). The activating role of

CSR-1 is specific to the germline and has been proposed to only be related to one of two isoforms of the protein, CSR-1b (Charlesworth et al., 2021). Whereas the shorter isoform CSR-1b is thought to be involved in chromosome organisation and proper expression of germline genes, the full isoform CSR-1a may silence genes within immune response pathways and it is present both in somatic tissues and in the germline during spermatogenesis (Charlesworth et al., 2021). The longer isoform has been shown to integrate into the 26G-RNA pathway (Charlesworth et al., 2021), with ALG-3/4 having been shown to upregulate target expression via CSR-1 (Conine et al., 2013). However, whether the two isoforms of CSR-1 indeed have such different functions is unclear. For instance, fine-tuning of the expression of a single microtubule depolymerase via CSR-1 during embryogenesis has been shown to affect chromosome segregation (Gerson-Gurwitz et al., 2016), suggesting that the role of CSR-1 in this process may not be direct, or at least not entirely direct. It also suggests that the two proposed role of CSR-1a and CSR-1b, chromosome segregation and gene silencing, may be directly linked, and thus cannot be contributed to one isoform over the other. Furthermore, CSR-1 has been shown to be involved in maternal mRNA turnover (Quarato et al., 2021), and failure to clear maternal transcripts in the zygote may cause a plethora of phenotypes, making it difficult to fully discern the direct functionalities of CSR-1.

One other Argonaute protein, VSRA-1, the closest homologue to CSR-1 (**Figure 2**), has been proposed to have similar functions to CSR-1 (Tabara et al., 2023), but no other WAGO has been implicated in positive regulatory roles or in chromosome organization.

One final note about WAGOs and 22G-RNAs in general, is that the 22G-RNA response can become independent of the initial trigger, i.e. functional in the absence of 21U-RNAs or 26G-RNAs. This phenomenon is termed RNA-induced epigenetic silencing (RNAe). As 22G-RNAs are transcribed from Argonaute-bound transcripts, this suggests that WAGOs may be able to trigger themselves or that certain pathways are in place, where one WAGO will trigger another. The full extent to which WAGOs interact with one another, however, is not currently known.

GERMLINE-EXPRESSED CYTOPLASMIC WAGOS AND THEIR ROLE IN INHERITANCE

This thesis is focused on a germline-expressed branch of cytoplasmic WAGOs; the one made up of WAGO-1, -3, and -4 (**Figure 2**). Other Argonaute proteins, both primary and secondary, are expressed in the germline of *C. elegans*, but this section will focus on what is known about WAGO-1, -3, and -4 and their localisation into perinuclear foci known as germ granules in the germline of *C. elegans* (Phillips & Updike, 2022).

Germ granules, like structures such as nuage or Yb-bodies, are RNA-dense subcellular locations that work as hotspots for processes such as those involved in RNAi, with most of the proteins thus far found in *C.* elegans germ granules having RNA-binding domains. The germline WAGOs are expressed in so-called P-granules,

named such after the necessity of the scaffolding proteins PGL-1 and PGL-3 to form. P-granules are phase-separated structures whose formation rely on PGL-1 and PGL-3, although the granules remain intact when PGL-1 and PGL-3 are cleared from them, which for instance happens during spermatogenesis. Apart from proteins involved in RNAi, P-granules also house a lot of nuclear pore complexes (NPCs), allowing transport between the nucleus and the granules (Phillips & Updike, 2022).

In the early embryos, all germ granule components exist within a single mixed granule. However, later in development as well as in the adult germline, several separate germ granules take form (**Figure 4C**). At this point, WAGO-1 and WAGO-3 are still found within the structure known as the P-granule, but WAGO-4 is in Z-granules, named after the Z-granule protein ZNFX-1 and always found directly next to P-granules (Wan et al., 2018). Other granules include the mutator foci and the SIMR-foci, although it has been proposed that more may exist and the extent to which they may share proteins and RNAs between them is not well understood (Phillips & Updike, 2022).

Mutator foci, as mentioned, are relevant to the biogenesis of 22G-RNAs, so at least some exchange must happen between these foci and the P/Z-granules in which the WAGOs exact their function. SIMR-foci are the least researched of the known germ granules, but it is thought to play a role in establishing a connection between the 21U-RNA pathway or the exo-siRNA pathway and the 22G-RNA pathway (Phillips & Updike, 2022).

Consistent with the fact that PGL-1 and PGL-3 are not necessary for the stability of P-granules, they are also not necessary for Argonaute recruitment, at least in the case of WAGO-1, whose ability to silence mRNA remains in the absence of PGL-1 and PGL-3 (Aoki et al., 2021). Loss of WAGO-1 also does not affect P-granule structure, but it does significantly reduce germline 22G-RNA levels (Aoki et al., 2021; Gu et al., 2009). Deletion of all three WAGOs, WAGO-1, -3, and -4, reduces germline 22G-RNA levels to near zero, but where neither loss of WAGO-3 or WAGO-4 alone are resistant to exogenous germline RNAi, the double mutant lacking both WAGO-3 and WAGO-4 is (Gu et al., 2009). This indicates at least some redundancy between WAGO-3 and WAGO-4 in responding to exogenous RNAi and a role for both of these proteins in maintaining a small pool of germline 22G-RNAs, the majority of which are stabilized by WAGO-1.

WAGO-1 and WAGO-3 target the majority of transposons (Gu et al., 2009; Liu et al., 2023; Vastenhouw et al., 2003) and both target pseudogenes (Gu et al., 2009; Seroussi et al., 2023). WAGO-4 may also target pseudogenes (Liu et al., 2023), but it has mainly been implicated in exogenous RNAi and in silencing germline-expressed mRNAs, the latter possibly triggered by the former (Xu et al., 2018). All WAGOs, not just WAGO-1, -3, and -4, are predicted to silence protein coding genes (Liu et al., 2023; Seroussi et al., 2023). Interestingly, WAGO-4 targets the same subset of genes as CSR-1, although the two have opposing actions (Xu et al., 2018). Both CSR-1-bound and WAGO-4-bound 22G-RNAs are polyuridylated at the 3'-end by CDE-

1, but the action of CDE-1 has been proposed to play a role in regulating WAGO-4 versus CSR-1 binding, with CDE-1 promoting stability of 22G-RNA-WAGO-4 binding and destabilizing binding of 22G-RNAs to CSR-1 (Xu et al., 2018). Another study, however, suggests that CSR-1-bound 22G-RNAs are more often polyuridylated than WAGO-4-bound 22G-RNAs and that even 22G-RNAs bound by WAGO-3 or WAGO-9 may be polyuridylated (Liu et al., 2023). The exact function of 22G-polyuridylation therefore remains unknown, as does the mechanism that distinguishes 22G-RNA binding between Argonautes. It is known, however, that certain manipulations of an Argonaute or of their accessory proteins can cause an Argonaute to change the 22G-RNAs that it binds. WAGO-1 and -3, for instance, are both processed at the N-terminus by the dipeptidyl peptidase DPF-3 (Gudipati et al., 2021). Loss of DPF-3 causes changes to the 22G-RNAs taken up by WAGO-1, making it preferentially bind 22G-RNAs targeting genes usually targeted by CSR-1 and WAGO-4 (Gudipati et al., 2021).

WAGO-1, -3, and -4 are all cytoplasmic and likely have no catalytic function; hence their method of silencing is unknown. Comparisons of protein and mRNA levels of a WAGO-1 target suggests that WAGO-1 functions in mRNA turnover (Aoki et al., 2021), and the fact that many WAGO-1-bound 22G-RNAs exist within predicted introns suggest that WAGO-1 exerts its function on non-spliced mRNA (Gu et al., 2009). It is both possible that all three WAGOs have a function in mRNA turnover prior to splicing, that they act in three different steps, or that they can all act in different steps in a context-dependent manner. Nonetheless, silencing via the 22G-RNA pathway in general has been shown to implore introns to distinguish between self and non-self transcripts (Makeyeva et al., 2021), suggesting that all WAGOs may act upstream of splicing.

Data also suggests that at least one WAGO-RNAi pathway must be related to the nonsense-mediated decay pathway (Xu et al., 2018), and that WAGO-4 is upregulated upon apoptosis, at least in some instances (Sendoel et al., 2019).

Apart from WAGO-1, -3, and -4, the primary Argonaute proteins as well as CSR-1 exist in germ granules, and the relationship between levels of CSR-1 versus non-CSR-1 WAGO 22G-RNAs is thought to be crucial for germline gene expression (Almeida, Andrade-Navarro, et al., 2019; Phillips & Updike, 2022). This may also be related to whether the germline is spermatogenic or oogenic, since the Argonaute proteins ALG-3/4 – the Argonautes binding 26G-RNAs during spermatogenesis – have been shown to interact with the CSR-1 pathway, as mentioned above (Conine et al., 2013). Interestingly, ALG-3/4 do *not* persist in mature sperm, but are rather deposited in the residual bodies (Conine et al., 2010). One study found WAGO-1 in mature sperm and therefore proposed WAGO-1 as a way of mediating paternal inheritance of the ALG-3/4 response (Conine et al., 2010), However, in **Chapter I** here, we will show that N-terminally tagged WAGO-1 does not localise to mature sperm, but instead is deposited in the residual bodies of budding spermatids. In **Chapter I**, we will also show that WAGO-3 locates to mature sperm and plays a role in paternal inheritance (Schreier

et al., 2022). Loss of DPF-3, which is necessary for WAGO-1 and -3 function, also causes spermatogenesis defects, further implicating one or both of these WAGOs in male fertility and inheritance (Gudipati et al., 2021), with our findings in **Chapter I** confirming that WAGO-3 acts in this role. Contrarily, WAGO-4 is not significantly expressed in the male germline but present in the hermaphroditic germline and oocytes and its expression pattern suggests a capacity of WAGO-4 for transmitting an RNAi phenotype to its offspring via the maternal side (Xu et al., 2018). Nonetheless, WAGO-4 is present in the hermaphrodite during spermatogenesis (Seroussi et al., 2023), which would allow it to have a function during this process as well.

Model organisms

As has been established, this thesis makes use of two different model organisms, *T. rugatulus* and *C. elegans*. This section will focus on describing the basics of their biology and reproduction, as RNAi and reproductive efficacy are linked, as described. A few select notes on the correlation between biology and RNAi machineries will also be presented.

THE ANT *TEMNOTHORAX RUGATULUS*

T. rugatulus is a rock ant common to Northern America, which lives in colonies and make their nests by building walls within narrow gaps in rocks (DiRienzo & Dornhaus, 2017).

Social insects such as ants generally function within a caste-system where females can either be queens or workers (Favreau et al., 2018). Though smaller and less common, ants can also be male. The sex of ants, and of most insects within the order of Hymenoptera, is most commonly determined by a haploid-diploid system where the offspring of a queen will be determined by whether or not she has bred (Heimpel & de Boer, 2008).

The haploid-diploid system of sex determination works in that fertile, reproductive females – in social insects, this mostly refers to females of the queen caste – produce eggs that do not need fertilization. Having undergone meiosis, unfertilized eggs carry only one set of chromosomes and are thus haploid. These invariably develop into males. If the female chooses to breed, however, a second set of chromosomes will be supplied by the male, and the fertilized egg will therefore be diploid and develop into a female (Heimpel & de Boer, 2008). Knockdown of certain sex determination genes in female Hymenoptera, however, can turn diploid offspring male (Heimpel & de Boer, 2008).

Female ants are genetically similar though queens and workers differ in size and reproductive capability (Favreau et al., 2018; Rüppell et al., 2001). This difference is determined by environmental factors during larval development. Whichever factors trigger caste differentiation, it is non-reversible and triggers many changes in gene expression. Certain chromosomes, referred to as social chromosomes, have been found to

contain genes with highly caste-specific expression whereas pleiotropic genes, i.e. genes influencing more than one phenotypic trait, are less likely to differ in expression between castes (Favreau et al., 2018).

T. rugatulus live in colonies with either one or multiple queens which can usually be divided into two morphs; macrogynes and microgynes. While colonies can contain both morphs, macrogynes are more prone to founding new colonies whereas microgynes more commonly stay in their natal nests (Rüppell et al., 2001; Rüppell et al., 2003). Queens are generally larger than workers, but macrogynes are also larger than microgynes. This difference in size affects reproduction, as larger ants have been shown to have a greater reproductive capacity (Rüppell et al., 2001).

Like queens can be divided into different categories, so can workers. Different types of worker differentiation have been reported in *T. rugatulus*. One example is the division into active or inactive workers. This distinction has also been reported in a variety of other social insect with up to 50% of workers being inactive. The function of inactive workers, which may seem disadvantageous, has been proposed to be quick adaptability. Inactive workers require less energy, but they can switch to being active when the active workers are removed from the colony, which is much faster and more efficient than having to rely on reproduction to replenish the work force (Charbonneau et al., 2017). During colony emigration, *T. rugatulus* workers have further been divided into four sub-categories, two active and two inactive, depending on their activity level and mobility between colonies, a division promoting efficient spread of information (Valentini et al., 2020).

Furthermore, different *T. rugatulus* workers may carry out different tasks. While workers can switch task, this has been shown to be highly inefficient, with workers doing task-switching requiring more time between tasks than workers performing the same task repeatedly. This has been proposed to be a main mechanism for behavioural specialization (Leighton et al., 2017). However, workers responsible for either brood care or foraging have also been shown to elicit phenotypic difference, with the former being younger and capable of activating their ovaries in the absence of a queen (Kohlmeier et al., 2023).

Some species of ants completely lack the queen caste and may have both reproductive and non-reproductive workers, the former referred to as gamergates, whereas other ants, such as *T. rugatulus*, have workers that are non-reproductive but can turn reproductive if the queen or queens are removed from the colony (Favreau et al., 2018). The switch of *T. rugatulus* workers from non-reproductive to reproductive has been shown to involve histone acetylation (Choppin et al., 2021) and an increase in DNA damage response pathway genes (Negroni et al., 2020). Interestingly, queens with experimentally induced increases in fecundity – i.e. the propensity for reproduction – also showed an increase in innate reproduction and expression of body maintenance mechanisms (Negroni et al., 2021), suggesting that fertility and the ability

to self-maintain (e.g. activating the DNA damage response pathways) are interlinked both in queens and workers.

Consistent with this idea is the fact that queens generally live much longer than workers, with queens capable of living for years as opposed to the few-months-long lifespan of workers. Males live the shortest, usually only for a few weeks (Favreau et al., 2018). In *T. rugatulus*, ageing queens have been shown to alter their gene expression profile and a link has been found between longevity and the investment into immune responses (Negroni et al., 2019). Similarly, fertility of workers affects the ability to respond to immunity (Negroni et al., 2020). In ants, lifespan has been shown to be connected to expression of genes involved in DNA repair and bacterial symbionts, microbiomes, and viruses have all been proposed to play a role in longevity and caste differences as well (Favreau et al., 2018).

While it is clear that longevity, fertility, and caste are all determined by epigenetics – environmental factors, gene expression, and histone acetylation – it has thus far not been shown whether this is connected to small RNAs in *T. rugatulus*. The current knowledge of the many roles of RNAi and the evidence that miRNAs do differ between castes in two other species of ants (R. Bonasio et al., 2010), however, does provide precedence for the validity of investigating such a hypothesis.

THE NEMATODE *CAENORHABDITIS ELEGANS*

C. elegans is a roundworm commonly found in soil or rotting food which feeds on bacteria. It is transparent and the adult worm mainly consists of a germline with two gonadal arms (depicted in **Figure 4A**) and intestines. *C. elegans* is hermaphroditic, meaning that the same worm produces both sperm and oocytes and it can self-fertilize. In *C. elegans*, spermatogenesis and oogenesis happens in two consecutive steps, with the sperm first being produced and then stored in a compartment called the spermatheca, before the gonad switches and starts producing oocytes, which are fertilized as they pass through the spermatheca on their way from gonad to uterus (**Figure 4A** and **C**) (Corsi et al., 2015).

C. elegans hatch from eggs lain by hermaphroditic parents after developing outside of the uterus at which point the newly hatched worm is characterized as being in its first larval stage, L1. In the fourth and final larval stage, L4, the worm will start producing sperm, with which it will cease in adulthood. Adult worms are generally divided into two categories: young adults, who have made the shift from spermatogenesis to oogenesis but do not yet carry fertilized eggs in their ovaries, and gravid adults, which encompass all gravid worms, regardless of whether they have started laying eggs (Corsi et al., 2015).

Although the vast majority of *C. elegans* worms are hermaphroditic, a small portion (<0.2%, (Corsi et al., 2015)) of males occur in a natural population. The sex is determined via the XO system, in which females –

or, in the case of *C. elegans*, hermaphrodites – carry two sex chromosomes (termed X) and males carry only one copy (Portman, 2007). The male *C. elegans* is depicted in **Figure 4B**, and it has only one gonadal arm.

Nematode sex is determined by the number of sex chromosomes relative to autosomes and heavily relies on dosage compensation. While this can be accomplished in several ways, nematodes perform dosage compensation by reducing transcription from both X chromosomes in the hermaphrodite (Meyer, 2022). The *C. elegans* sex determination pathway ends with the transcription factor *tra-1*, which regulates the kinase and master sex-determination switch gene *xol-1*, which is responsible both for sex determination and for dosage compensation (Hargitai et al., 2009). XOL-1 works both by removing repressors of the pathway leading to repression of male sex-determining genes and by triggering the binding of a dosage compensation complex to the X chromosomes. Mutations in XOL-1 or in proteins downstream of the sex determination pathway can lead to male worms becoming feminized or female worms becoming masculinized (Meyer, 2022). Even with the adult male germline resembling that of the hermaphroditic L4 larvae, several genes are differentially expressed between the sexes, and several sex-specific neurons are necessary to drive behaviours such as egg-laying in the hermaphrodite or mating in the male (Portman, 2007).

Interestingly, while breeding provides the opportunity for renewing the gene pool and is generally seen as advantageous in animals such as ants, *C. elegans* mating comes with a loss for both partners. Breeding increases brood size but decreases the health of both male and hermaphrodite with hermaphrodites becoming smaller and both sexes having reduced lifespans (Shi & Murphy, 2021).

The majority of *C. elegans* embryogenesis happens outside the uterus, with eggs being laid around the 28-cell stage (Corsi et al., 2015). At around this stage, PGC progenitor cells will have formed, and it is around the nucleus of these PGC precursors that the germ granules discussed previously will begin to demix and locate in their characteristic perinuclear manner (**Figure 4C**) (Phillips & Updike, 2022). In *C. elegans*, the PGCs are called Z2 and Z3, and they belong to a set of four progenitor cells, Z1-4, that form the gonad, with Z1 and Z4 forming the somatic moiety thereof. Hermaphrodite and male gonads look the same in L1 larvae, but development happens symmetrically in the hermaphrodite and asymmetrically in the male, which causes the symmetry of the two hermaphroditic gonadal arms in the adult and the presence of only one gonadal arm in the male (Kimble & Hirsh, 1979).

Both male and hermaphroditic gonads display a distal-proximal axis with the maturation of gametes occurring from the distal end, i.e. the one farthest from the vulva or cloaca, towards the proximal end (Kimble & Hirsh, 1979). Also at the distal end is the distal tip cell, which has been derived from Z1/Z4 and whose migration during larval development help shape the gonadal arm (Cecchetelli & Cram, 2017). Next to the distal tip cell is the mitotic region followed by the different steps of meiosis which all can be observed at distinct locations in the *C. elegans* germline (**Figure 4C**), which makes it a good organism for studying meiotic

progression (Corsi et al., 2015). Since the germline RNAi machineries are in place throughout the germline, it also enables the study of these in relation to meiosis, and the successive production of sperm and oocytes makes it an interesting organism for studying sex-dependent inheritance of factors involved in RNAi.

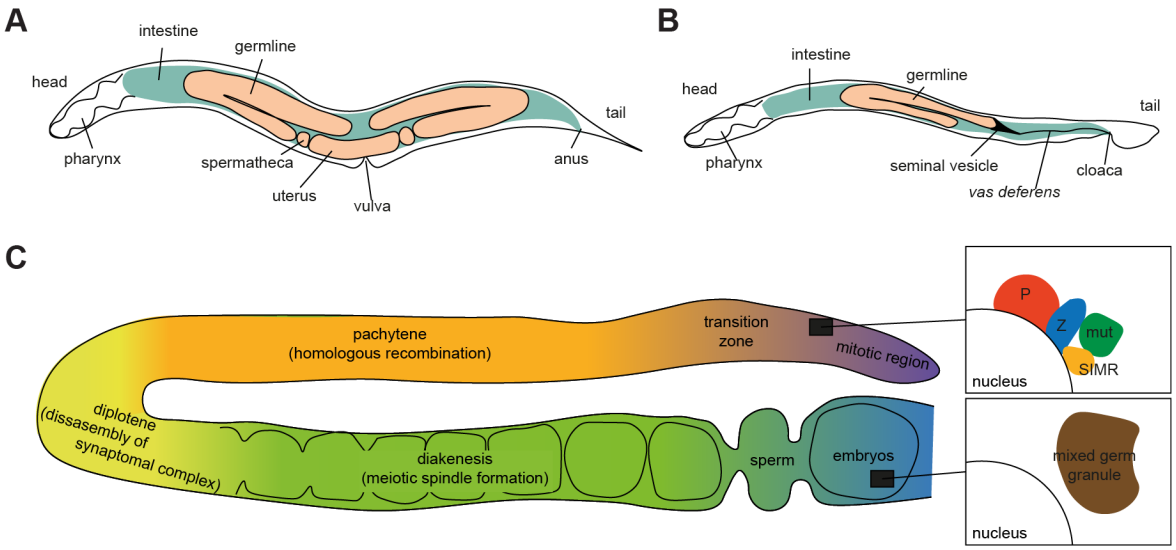


***Figure 4**: **Illustrations of** C. elegans. **A** Structure of a gravid adult hermaphrodite with the germline indicated. **B** Structure of an adult male with the germline indicated. **C** Detailed depiction of the adult hermaphroditic germline and meiotic progression herein. Germ granule structures indicated in zoom-ins.*

Chapter I

Membrane-associated cytoplasmic granules carrying the Argonaute protein WAGO-3 enable paternal epigenetic inheritance in *Caenorhabditis elegans*

Page left intentionally blank

nature cell biology

ARTICLES

https://doi.org/10.1038/s41556-021-00827-2

Membrane-associated cytoplasmic granules carrying the Argonaute protein WAGO-3 enable paternal epigenetic inheritance in *Caenorhabditis elegans*

Jan Schreier[1,2], Sabrina Dietz[2,3], Mandy Boermel[4], Viola Oorschot[4], Ann-Sophie Seistrup[1,2], Antonio M. de Jesus Domingues[1,8], Alfred W. Bronkhorst[1], Dieu An H. Nguyen[5], Stephanie Phillis[6], Elizabeth J. Gleason[6], Steven W. L'Hernault[6], Carolyn M. Phillips[5], Falk Butter[3] and René F. Ketting[1,7] ✉

Epigenetic inheritance describes the transmission of gene regulatory information across generations without altering DNA sequences, enabling offspring to adapt to environmental conditions. Small RNAs have been implicated in this, through both the oocyte and the sperm. However, as much of the cellular content is extruded during spermatogenesis, it is unclear whether cytoplasmic small RNAs can contribute to epigenetic inheritance through sperm. Here we identify a sperm-specific germ granule, termed the paternal epigenetic inheritance (PEI) granule, that mediates paternal epigenetic inheritance by retaining the cytoplasmic Argonaute protein WAGO-3 during spermatogenesis in *Caenorhabditis elegans*. We identify the PEI granule proteins PEI-1 and PEI-2, which have distinct functions in this process: granule formation, Argonaute selectivity and subcellular localization. We show that PEI granule segregation is coupled to the transport of sperm-specific secretory vesicles through PEI-2 in an *S*-palmitoylation-dependent manner. PEI-like proteins are found in humans, suggesting that the identified mechanism may be conserved.

Small RNAs, most notably short-interfering RNAs (siRNAs) and Piwi-interacting RNAs (piRNAs)[1,2], have been implicated in epigenetic inheritance. These molecules act as sequence-specific guides for Argonaute proteins, which in turn regulate gene expression[3–5]. In the nematode *C. elegans*, the 22G-RNAs are siRNAs with an established role in epigenetic inheritance[6–11]. These are generated in a process driven by the *Mutator* proteins MUT-16 and MUT-7 (refs. [12–15]), and are bound by members of the worm-specific Argonaute (WAGO) family, such as the cytoplasmic WAGO-4 and the nuclear HRDE-1 proteins[14–17], both of which have been implicated in maternal epigenetic inheritance[7,10,18]. The Piwi protein PRG-1 is inherited through the oocyte, and maternal piRNAs can initiate WAGO-dependent silencing that can be inherited PRG-1-independently for many generations in a process known as RNA-induced epigenetic silencing (RNAe)[19–21].

Interestingly, a second, *Mutator*-independent class of 22G-RNAs also exists. These are bound by the Argonaute protein CSR-1, and are mostly derived from genes that are expressed and required in the germline[22]. CSR-1 22G-RNAs and *Mutator* 22G-RNAs should not become mixed so that important genes are not inappropriately silenced, and epigenetic inheritance has an important role in this. When embryos have a functional *Mutator* system, but their parents lacked both *Mutator* 22G-RNAs and piRNAs, the *Mutator* system starts to produce 22G-RNAs that are normally restricted to CSR-1 (refs. [8,9]). As a result, WAGO proteins such as HRDE-1 are loaded with CSR-1-type 22G-RNAs, leading to the silencing of CSR-1-target genes, which in turn results in sterility[8,9]. This phenotype, known as *Mutator*-induced sterility (Mis), effectively reveals a prominent self-targeting potential of the *Mutator* 22G-RNAs, and shows that parental 22G-RNAs and piRNAs have an essential role in suppressing this dangerous autoimmune-like property of the *Mutator* system.

Condensates are membraneless compartments that are important for subcellular organization[23,24]. In *C. elegans*, distinct condensates have been implicated in small RNA pathways and epigenetic inheritance—P granules, Z granules and *Mutator* foci, defined by PGL-1, ZNFX-1 and MUT-16, respectively[7,15,23]. Interestingly, so far, evidence of such a germ granule in mature sperm is lacking[25] and, in fact, P granules disappear during spermatogenesis[23]. This is possibly linked to the massive cytoplasmic reduction during spermatogenesis. In *C. elegans*, during meiosis II, a residual body is formed into which, for example, endoplasmic reticulum, Golgi, free ribosomes[26] and the Argonaute protein ALG-3 are discarded[27]. Thus, it is questionable whether paternal epigenetic inheritance[28–30] can be mediated through the cytoplasm.

Results

WAGO-3 is required for inheritance. Mutants that are defective for epigenetic inheritance, such as *hrde-1*, often show a mortal germline phenotype[2], implying that fertility decreases over subsequent generations. We found that *wago-3* mutants display a

[1]Biology of Non-coding RNA group, Institute of Molecular Biology, Mainz, Germany. [2]International PhD Programme on Gene Regulation, Epigenetics and Genome Stability, Institute of Molecular Biology, Mainz, Germany. [3]Quantitative Proteomics Group, Institute of Molecular Biology, Mainz, Germany. [4]Electron Microscopy Core Facility, European Molecular Biology Laboratory, Heidelberg, Germany. [5]Department of Biological Sciences, University of Southern California, Los Angeles, CA, USA. [6]Department of Biology, Emory University, Atlanta, GA, USA. [7]Institute of Developmental Biology and Neurobiology, Johannes Gutenberg University, Mainz, Germany. [8]Present address: Dewpoint Therapeutics GmbH, Dresden, Germany. ✉e-mail: r.ketting@imb-mainz.de

mortal germline phenotype (Extended Data Fig. 1a), suggesting a role for WAGO-3 in epigenetic inheritance. To test this, we analysed whether WAGO-3 affects the heritability of a germline-specific *mCherry::H2B* transgene silenced by RNAe[19–21,31]. The RNAe status implies that the silencing was induced by PRG-1, but afterwards maintained in a *prg-1*-mutant background[31–33]. We therefore assessed the RNAe-associated silencing in *prg-1*-mutant strains. We found that *wago-3;prg-1* double mutants, but not *prg-1* single mutants, stochastically lost silencing (Extended Data Fig. 1b,c). As expected[32,33], *mut-7;prg-1* double mutants directly lost all silencing (Extended Data Fig. 1b,c). These data show that WAGO-3 has a role specifically in the inheritance of RNAe-related silencing.

WAGO-3 associates with paternal 22G-RNAs. To enable immunoprecipitation (IP) experiments, we endogenously tagged WAGO-3 with GFP::3×FLAG at its N terminus. We note that this may affect the function of WAGO-3, as its N terminus is processed[34]. However, disruption of N-terminal WAGO-3 processing does not result in phenotypes in WAGO-1-wild-type animals[34], as is the case in our experiments. We used this tag to IP WAGO-3 from adult hermaphrodites, containing mature sperm and oogenic gonads, and adult males, and sequenced WAGO-3-bound 22G-RNAs (Extended Data Fig. 2a–d) to identify targets of WAGO-3 (Extended Data Fig. 3a–c). Many of these are protein-coding transcripts that are known *Mutator* targets[15], while their overlap with CSR-1 targets is small[22] (Extended Data Fig. 3d). In both sexes, many transposable elements were targeted by WAGO-3 (Extended Data Fig. 3c,e), including Tc1, consistent with the role of WAGO-3 in Tc1 silencing[35,36]. Interestingly, Tc3-derived 22G-RNAs were consistently depleted from WAGO-3 in males (Extended Data Fig. 3c), suggesting sex-specific regulation of this element. Finally, we found a substantial overlap between WAGO-3 targets and previously determined sperm-derived 22G-RNA targets[37] (Extended Data Fig. 3f), suggesting that WAGO-3 is present in sperm.

WAGO-3 is guided into sperm by PEI granules. Confocal microscopy analysis of GFP::3×FLAG::WAGO-3 (hereafter, WAGO-3) revealed expression throughout the germline at all stages, with localization to P granules in mitotic, meiotic and primordial germ cells (Extended Data Fig. 4a,b). Notably, we found strong WAGO-3 signals within the sperm-containing spermatheca (Extended Data Fig. 4a). The presence of WAGO-3 in sperm was confirmed by analysing isolated, male-derived germ cells at different stages of spermatogenesis, which also revealed a punctate subcellular localization (Fig. 1a and Extended Data Fig. 4c). We next performed IP analysis of WAGO-3 from late fourth larval stage (L4) hermaphrodites, a stage during which spermatogenesis is ongoing, followed by label-free quantitative mass spectrometry (IP–MS/MS). Besides known P granule components such as DEPS-1, PRG-1 and WAGO-1 (ref. [38]; Fig. 1b), we identified F27C8.5 (PEI-1). We confirmed the interaction (Extended Data Fig. 4d), and found that PEI-1::mTagRFP-T (hereafter, PEI-1) was exclusively expressed during the later stages of spermatogenesis, both in L4 hermaphrodites (Extended Data Fig. 4e) and in males (Extended Data Fig. 4f).

We analysed late-L4-stage hermaphrodites to examine PEI-1 and WAGO-3 expression in relation to P granules. In naive germ cells, which are PEI-1 negative, WAGO-3 localized to peri-nuclear P granules, marked by PGL-1 (Fig. 1c,d; region of interest 1 (ROI1)). Starting at the primary spermatocyte stage, WAGO-3 accumulated in non-peri-nuclear, cytoplasmic foci and, as previously described[23], P granules began to disappear (Fig. 1c,d; ROI2–3). Moreover, PEI-1 started to be expressed and co-localized with WAGO-3 (Fig. 1e,f; ROI2). PEI-1 did not co-localize with the P granule marker DEPS-1 (Extended Data Fig. 4g,k).

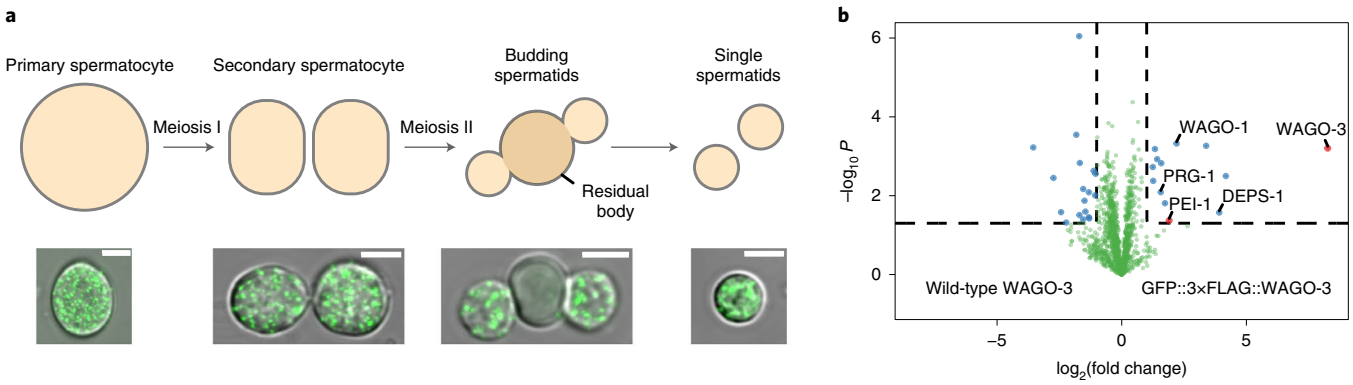
WAGO-3 still co-localized with PGL-1 in *pei-1* mutants, but was absent from spermatozoa and instead was found in the residual body (Fig. 1g–i). Thereby, WAGO-3 followed the same fate as the Argonaute proteins WAGO-1, ALG-3 and CSR-1 in wild-type animals (Extended Data Fig. 4h–j,l–n).

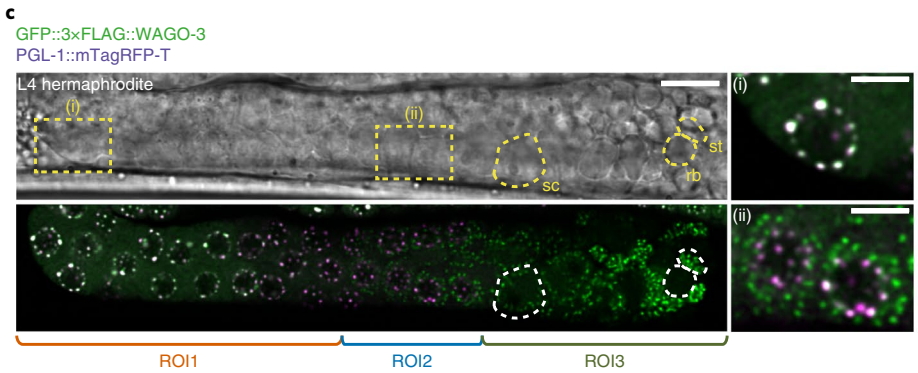
Finally, we tested whether PEI-1 foci depend on known germ granules by removing MUT-16 (*Mutator* foci) or DEPS-1 (P granules)[15,39], or on its resident Argonaute protein WAGO-3. None of these proteins were required for PEI-1 foci (Extended Data Fig. 5a).

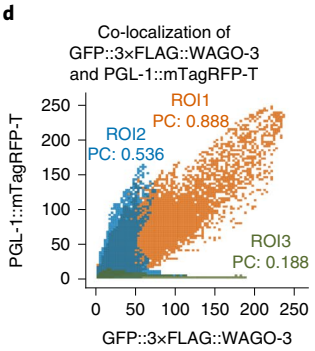
We conclude that PEI-1 defines a spermatogenesis-specific germ granule—the PEI granule—that recruits WAGO-3 and enables its segregation into mature sperm.

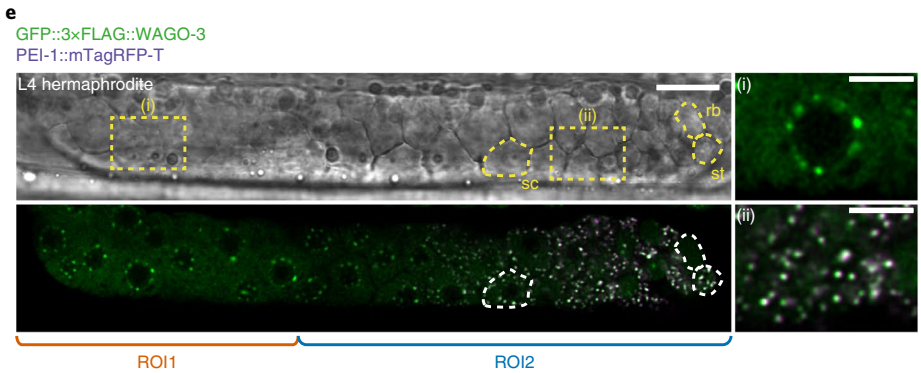
Paternal epigenetic inheritance requires WAGO-3 and PEI-1. The Mis phenotype[8,9] (described above) enables us to examine the relevance and mechanisms of epigenetic inheritance. The precise set-up that we used in the experiment is shown in Fig. 2a. Note that the *mut-7* and *mut-16* mutants can be used interchangeably in both sexes, as they both result in *Mutator* system dysfunction[8,9,14]. This set-up generates embryos that can make *Mutator* 22G-RNAs. Depending on the specific cross, the mother, the father or neither parent can make *Mutator* 22G-RNAs. All strains carry a *prg-1* mutation to remove the partially redundant activity of inherited piRNAs in this system[8,9]. Using this set-up, we found that maternal or paternal 22G-RNAs were sufficient to prevent the Mis phenotype (Fig. 2b,c (top three bars)), enabling us to dissect male- and

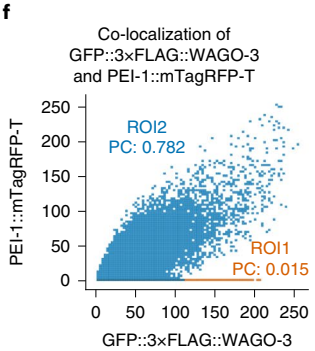
Fig. 1 | WAGO-3 is guided into sperm by PEI-1. a, Schematic of spermatogenesis in *C. elegans*. Corresponding confocal images of male-derived cells expressing GFP::3xFLAG::WAGO-3 are shown below. Scale bars, 4 µm. **b**, Label-free proteomic quantification of quadruplicate GFP::3xFLAG::WAGO-3 IP experiments from late-L4 stage hermaphrodite extracts. The *x* axis indicates the mean fold enrichment of individual proteins in the control versus the genome-edited strain. The *y* axis represents the $-\log_{10}P$ of observed enrichments. The dashed lines show thresholds at $P = 0.05$ and twofold enrichment. The blue and green data points represent above and below the threshold, respectively. WAGO-3 and PEI-1 are highlighted by red data points. **c**, Confocal micrograph showing spermatogenesis of late-L4 stage hermaphrodites expressing the indicated proteins. PGL-1::mTagRFP-T was used as a P granule marker. Germ cell development progresses from left to right. The areas indicated by dashed yellow boxes (i and ii) are magnified on the right. ROI1: PGL-1::mTagRFP-T and GFP::3xFLAG::WAGO-3 co-localize; ROI2: GFP::3xFLAG::WAGO-3 leaves P granules; ROI3: PGL-1::mTagRFP-T signal is not detectable anymore. rb, residual body; sc, spermatocyte; st, spermatid. Scale bars, 10 µm (proximal gonad) and 4 µm (magnified images). **d**, Co-localization analysis between GFP::3xFLAG::WAGO-3 and PGL-1::mTagRFP-T based on the image shown in **c**. Signals from ROI1–ROI3 are plotted in orange, blue and green, respectively. The *x* and *y* axes indicate fluorescence intensity. PC, Pearson's correlation coefficient. **e**, Confocal micrograph as in **c** for PEI-1::mTagRFP-T instead of PGL-1::mTagRFP-T. ROI1: no PEI-1::mTagRFP-T expression. ROI2: PEI-1::mTagRFP-T is expressed. The areas indicated by dashed yellow boxes (i and ii) are magnified on the right. Scale bars, 10 µm (proximal gonad) and 4 µm (magnified images). **f**, Co-localization analysis as in **d** for PEI-1::mTagRFP-T instead of PGL-1::mTagRFP-T. **g**, Confocal maximum intensity projections of male-derived budding spermatids expressing GFP::3xFLAG::WAGO-3 in the presence and absence of PEI-1. Scale bars, 4 µm. **h**, Confocal micrograph showing spermatogenesis of late-L4 stage hermaphrodites expressing GFP::3xFLAG::WAGO-3 and PGL-1::mTagRFP-T in the *pei-1(ok1050)* mutant background. Germ cell development progresses from left to right. The areas indicated by dashed yellow boxes (i and ii) are magnified on the right. ROI1: PGL-1::mTagRFP-T is expressed; ROI2: no PGL-1::mTagRFP-T expression is detectable. Scale bars, 10 µm (proximal gonad) and 4 µm (magnified images). **i**, Co-localization analysis as in **d** in the *pei-1(ok1050)* mutant background based on the image shown in **h**. The images in **c**, **e**, **g** and **h** represent three biologically independent experiments. Source data are available online.

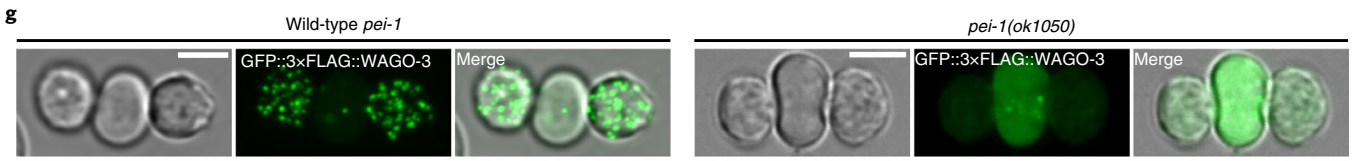

a
Primary spermatocyte
Secondary spermatocyte
Budding spermatids
Single spermatids
Meiosis I
Meiosis II
Residual body
b
WAGO-1
WAGO-3
PRG-1
PEI-1
DEPS-1
Wild-type WAGO-3
GFP::3×FLAG::WAGO-3
log2(fold change)
−log10 P

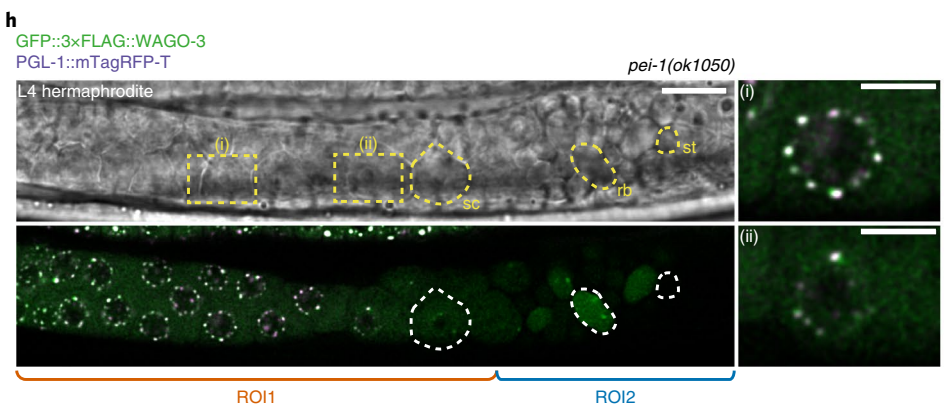

c
GFP::3×FLAG::WAGO-3
PGL-1::mTagRFP-T
L4 hermaphrodite
ROI1
ROI2
ROI3

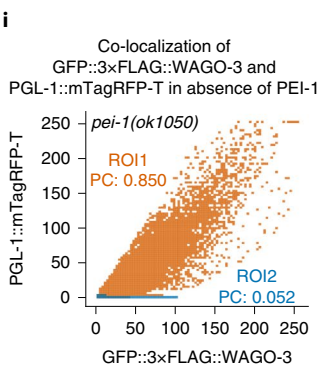

d
Co-localization of GFP::3×FLAG::WAGO-3 and PGL-1::mTagRFP-T
ROI1 PC: 0.888
ROI2 PC: 0.536
ROI3 PC: 0.188
PGL-1::mTagRFP-T
GFP::3×FLAG::WAGO-3


e
GFP::3×FLAG::WAGO-3
PEI-1::mTagRFP-T
L4 hermaphrodite
ROI1
ROI2


f
Co-localization of GFP::3×FLAG::WAGO-3 and PEI-1::mTagRFP-T
ROI2 PC: 0.782
ROI1 PC: 0.015
PEI-1::mTagRFP-T
GFP::3×FLAG::WAGO-3


g
Wild-type pei-1
pei-1(ok1050)
GFP::3×FLAG::WAGO-3
Merge


h
GFP::3×FLAG::WAGO-3
PGL-1::mTagRFP-T
pei-1(ok1050)
L4 hermaphrodite
ROI1
ROI2


i
Co-localization of GFP::3×FLAG::WAGO-3 and PGL-1::mTagRFP-T in absence of PEI-1
pei-1(ok1050)
ROI1 PC: 0.850
ROI2 PC: 0.052
PGL-1::mTagRFP-T
GFP::3×FLAG::WAGO-3

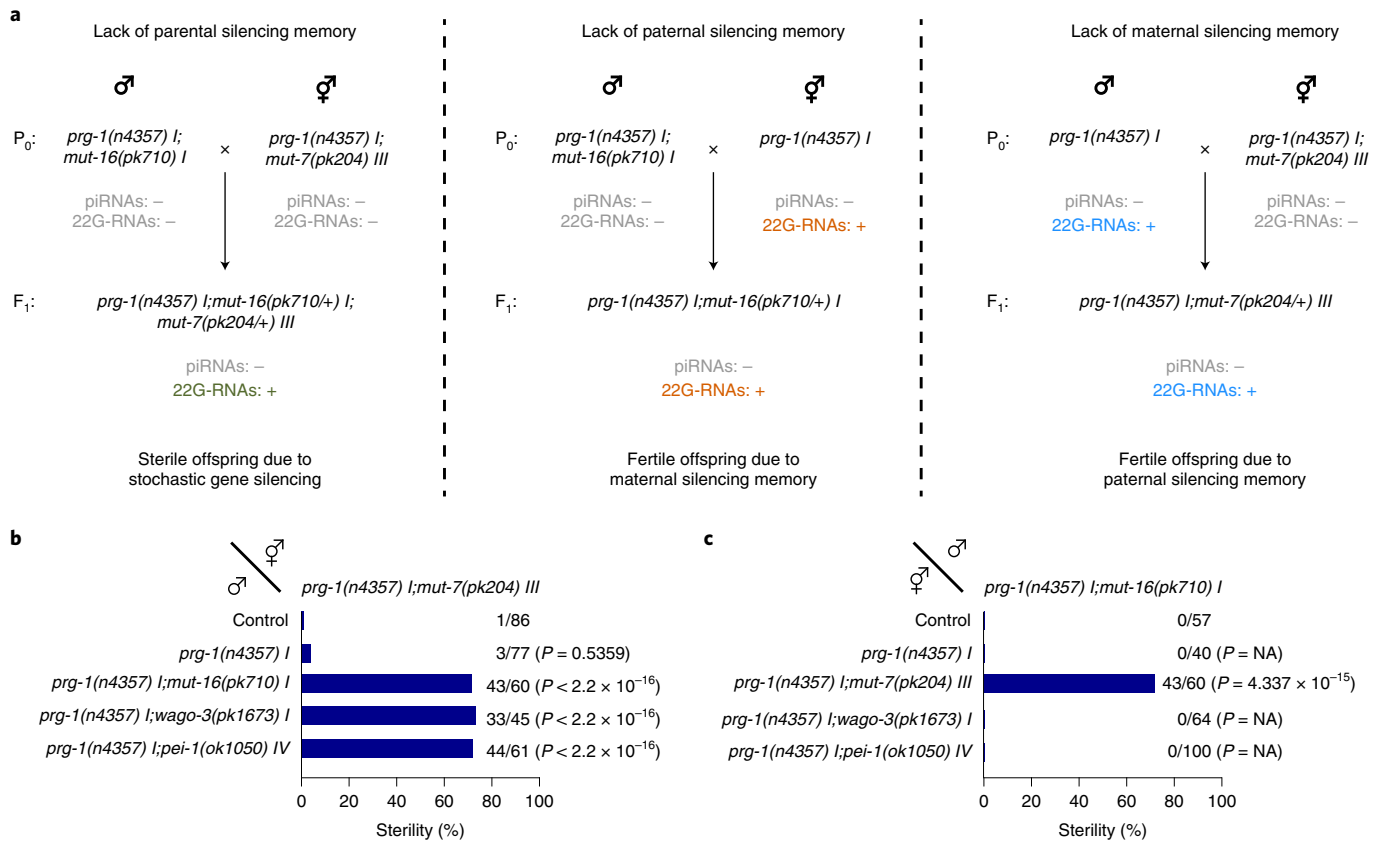


Fig. 2 | WAGO-3 and PEI-1 are required for paternal epigenetic inheritance. a, Schematic of the crosses that were used to examine the specific effects of maternal and paternal epigenetic inheritance on the Mis phenotype. The *mut-7(pk204)* and *mut-16(pk710)* alleles both cause global depletion of *Mutator* 22G-RNAs, and can be used interchangeably. Small RNA colour code: grey, absent without parental influence; green, present; without parental influence; orange, present with maternal influence; blue, present with paternal influence. **b**,**c**, The percentage of fertile F_1 animals generated by crosses between males and hermaphrodites of the indicated genotypes. Fertility implies the presence of paternal (**b**) or maternal (**c**) epigenetic inheritance. Sterility implies no epigenetic inheritance. Statistical significance was tested using a Pearson's χ^2 test with Yates continuity correction. NA, not available.

female-specific contributions to the Mis phenotype. We next examined the roles of WAGO-3 and PEI-1 in this process, and found that both PEI-1 and WAGO-3 are specifically required in the male (Fig. 2b), but not in the female (Fig. 2c). We conclude that PEI-1 and WAGO-3 have critical roles in paternal epigenetic inheritance.

PEI-1 recruits WAGO-3 to PEI granules through its IDR. The N-terminal region of PEI-1 is predicted to adopt a BTB fold followed by a BACK domain[40,41], whereas the C-terminal part of PEI-1 is predicted to be an intrinsically disordered region (IDR) (Fig. 3a). Following these predictions, we edited the endogenous PEI-1::mTagRFP-T locus to generate five different PEI-1 variants (Fig. 3b), and analysed their effects on PEI-1 and WAGO-3 expression in primary spermatocytes (Fig. 3c–k) and budding spermatids (Fig. 4a–n). As a control, free GFP was expressed from the *wago-3* locus (Figs. 3b,l,m and 4h). This revealed that the BTB and BACK domains primarily affected granule number and intensity (Figs. 3j and 4i,j,n), but had a very small effect on co-localization between PEI-1 and WAGO-3 (Figs. 3k and 4k). By contrast, deletion of the PEI-1 IDR resulted in WAGO-3 that mostly localized to the residual body (Fig. 4e,f,l,m). WAGO-3 signal was diffuse, as quantified by a loss of high-intensity pixels (Fig. 4i,j). The PEI-1 signal itself was weaker, but remained in foci that segregated into the spermatids (Fig. 4e,f). Deletion of both the BACK and IDR domains did not further affect WAGO-3, but did result in a diffuse PEI-1 signal that accumulated in the residual body, together with WAGO-3 (Fig. 4g,k).

Residual PEI-1 signal was always detected in spermatozoa within the spermatheca, even when both the BACK and IDR domains were deleted (Extended Data Fig. 5b–g). We observed the same for free GFP (Extended Data Fig. 5h). However, WAGO-3 is undetectable in mature sperm when PEI-1 misses its IDR (Extended Data Fig. 5f,g), indicating that WAGO-3 cannot be stably maintained in sperm without PEI-1 interaction.

We conclude that the IDR of PEI-1 is essential to recruit and stabilize WAGO-3, whereas the BTB, BACK and IDR domains have important roles in forming and stabilizing PEI-1 foci during spermatogenesis.

Characteristics of PEI granules. WAGO-3 was found to be highly sensitive to 1,6-hexanediol, a compound that is often used to probe condensates[42], as no foci remained in the presence of only 1.25% (Extended Data Fig. 6a,b). By contrast, PEI-1 foci were more resistant, especially in budding spermatids in which even 5% 1,6-hexanediol treatment did not cause complete disassembly of PEI granules (Extended Data Fig. 6b). This resistance to 1,6-hexanediol possibly derives from additional interactions between the folded BTB and BACK domains, which have been shown to drive oligomerization[41,43].

We also compared WAGO-3 mobility between PEI granules and P granules by measuring fluorescence recovery after photobleaching (FRAP) (Extended Data Fig. 6c,d). Proteins localizing to the liquid phase of P granules have been reported to exhibit high recovery rates[44,45]. Consistently, we found that WAGO-3 showed relatively

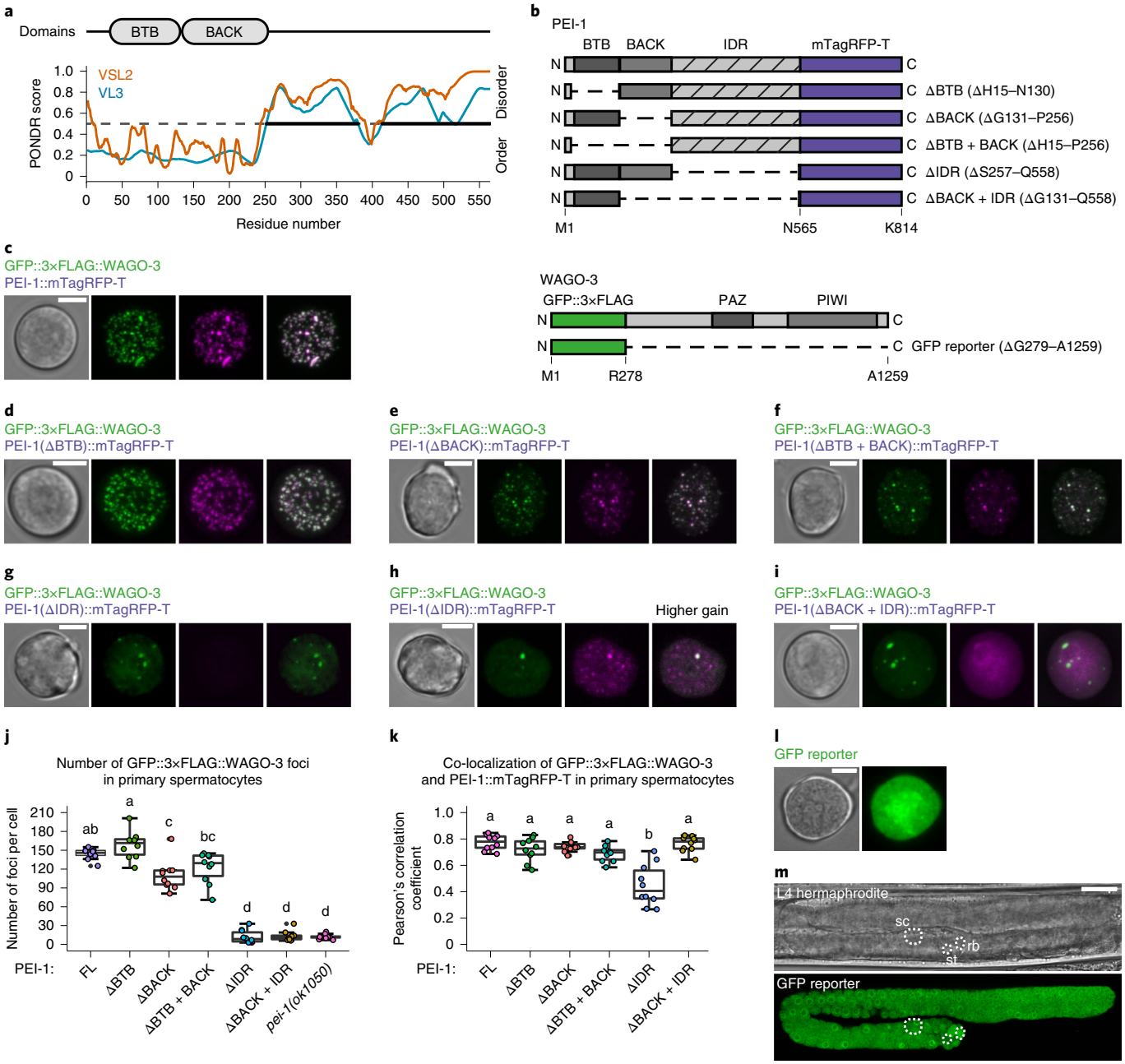


Fig. 3 | Granule formation and WAGO-3 interaction are mediated through different PEI-1 domains. a, PEI-1 domain composition. Prediction of naturally disordered regions was performed using the PONDR VSL2 and PONDR VL3 algorithms. **b**, Schematic of truncated PEI-1::mTagRFP-T and GFP::3xFLAG::WAGO-3 proteins generated using CRISPR–Cas9-mediated genome editing. **c**–**i**,**l**, Confocal maximum intensity projections of isolated, male-derived primary spermatocytes expressing GFP::3xFLAG::WAGO-3 (**c**–**i**) PEI-1::mTagRFP-T (**c**), PEI-1(ΔBTB)::mTagRFP-T (**d**), PEI-1(ΔBACK)::mTagRFP-T (**e**), PEI-1(ΔBTB + BACK)::mTagRFP-T (**f**), PEI-1(ΔIDR)::mTagRFP-T (**g**,**h**), PEI-1(ΔBACK + IDR)::mTagRFP-T (**i**) and a GFP reporter (**l**). For **c**–**i**, GFP::3xFLAG::WAGO-3 appears in green and PEI-1::mTagRFP-T variants appear in magenta. The images represent two biologically independent experiments. The image in **h** was acquired with a higher gain compared with the image in **g** to visualize the remaining PEI-1::mTagRFP-T signal within the spermatocyte. **l**, Free GFP was expressed from the *wago-3* locus. **j**,**k**, Quantification of the GFP::3xFLAG::WAGO-3 foci number (**j**), and co-localization of GFP::3xFLAG::WAGO-3 and PEI-1::mTagRFP-T (**k**) in isolated, male-derived primary spermatocytes expressing the indicated PEI-1::mTagRFP-T variants. FL, full length. $n = 10$ cells pooled from two independent experiments for each condition. Statistically significant differences were determined using one-way analysis of variance (ANOVA) ($P \leq 0.001$) followed by Tukey's honest significant difference post hoc test ($P \leq 0.05$). Different letters represent significant differences. The exact P values are provided as source data. The box plots show the median (centre line), 25th or 75th percentiles (box edges), and the whiskers indicate the median ± 1.5 × interquartile range. Note that the full-length data in **j** and **k** are the same as those shown in Extended Data Fig. 4c (primary spermatocyte) and Extended Data Fig. 10f,g, respectively. For **c**–**i** and **l**, scale bars, 4 µm. **m**, Confocal micrograph of an L4-stage hermaphrodite expressing free GFP from the endogenous *wago-3* locus. The image represents two biologically independent experiments. Scale bar, 20 µm. Source data are available online.

rapid FRAP in P granules ($t_{1/2}$ = 4.9 s). In PEI granules of budding spermatids, WAGO-3 exhibited much slower exchange dynamics ($t_{1/2}$ = 42.2 s). Moreover, we found that the mobile fraction of WAGO-3 was reduced in PEI granules compared with in P granules.

The prevalence of certain amino acids has been shown to modulate the material properties of condensates[46]. In particular, glycine residues maintain liquidity, whereas serine and glutamine residues promote hardening. We analysed the amino acid composition of the PEI-1 IDR, and compared it with the IDRs of PGL-1 and PGL-3, which are both known to localize to the liquid phase of P granules[47–49], and MEG-3 and MEG-4, which are both reported to form gel-like assemblies[45] (Extended Data Fig. 7a–h). This revealed that the PGL-1 and PGL-3 IDRs were strongly enriched for glycine, due to a glycine-rich C-terminal domain (Extended Data Fig. 7b,e,f), whereas such enrichment was absent from PEI-1, MEG-3 and MEG-4 (Extended Data Fig. 7c,g,h). Introduction of this PGL-1-derived glycine-rich stretch into PEI-1::mTagRFP-T rendered the glycine enrichment of the PEI-1 IDR similar to that of PGL-1 (Extended Data Fig. 7d), and enhanced WAGO-3 recovery (Extended Data Fig. 7i–k), indicating that the PEI-1 IDR composition affects WAGO-3 mobility. Note that the mTagRFP-T tag reduced WAGO-3 recovery compared with untagged PEI-1 (compare Extended Data Fig. 6c with Extended Data Fig. 7k), indicating that the tags that we introduced affected PEI granule properties.

Finally, fusion and fission of individual foci is a strong indication for liquid-like properties of condensates. We therefore performed live imaging of PEI granules by monitoring WAGO-3. However, we found that they were rather static; we did not observe any major movements, and we therefore observed no fission or fusion events in a period of 1 h (Extended Data Fig. 6e and Supplementary Video 1).

PEI granules associate with membranous organelles. In our deletion analysis of PEI-1, we generated deletions that removed all of PEI-1, or all but a few amino acids at the N and C termini (Fig. 5a). Although the full deletion produced a diffuse signal that segregated significantly into the residual body, the remaining PEI-1 peptides guided the mTagRFP-T signal to discrete structures that were maintained in spermatids (Fig. 5b,c and Extended Data Fig. 8a–c), and that were distinct from PEI granules (Extended Data Fig. 8d,e). Besides the nucleus, only two organelles are sorted into spermatids—mitochondria and FB-MOs[26]. The latter are sperm-specific membranous organelles that help to sort major sperm protein (MSP) and other proteins into spermatids. They consist of a membranous part (MO) and a fibrous body (FB) made of MSP. The PEI-1-marked structures did not overlap with mitochondria (Extended Data Fig. 8f,g), and their numbers approximately match that of FB-MOs (Fig. 5d), suggesting that the large PEI-1 deletion possibly marks FB-MOs. Furthermore, the sorting of PEI granules depends on the myosin VI motor protein SPE-15 (Fig. 5e), a protein that is known to drive FB-MO, but also mitochondria localization in sperm[50].

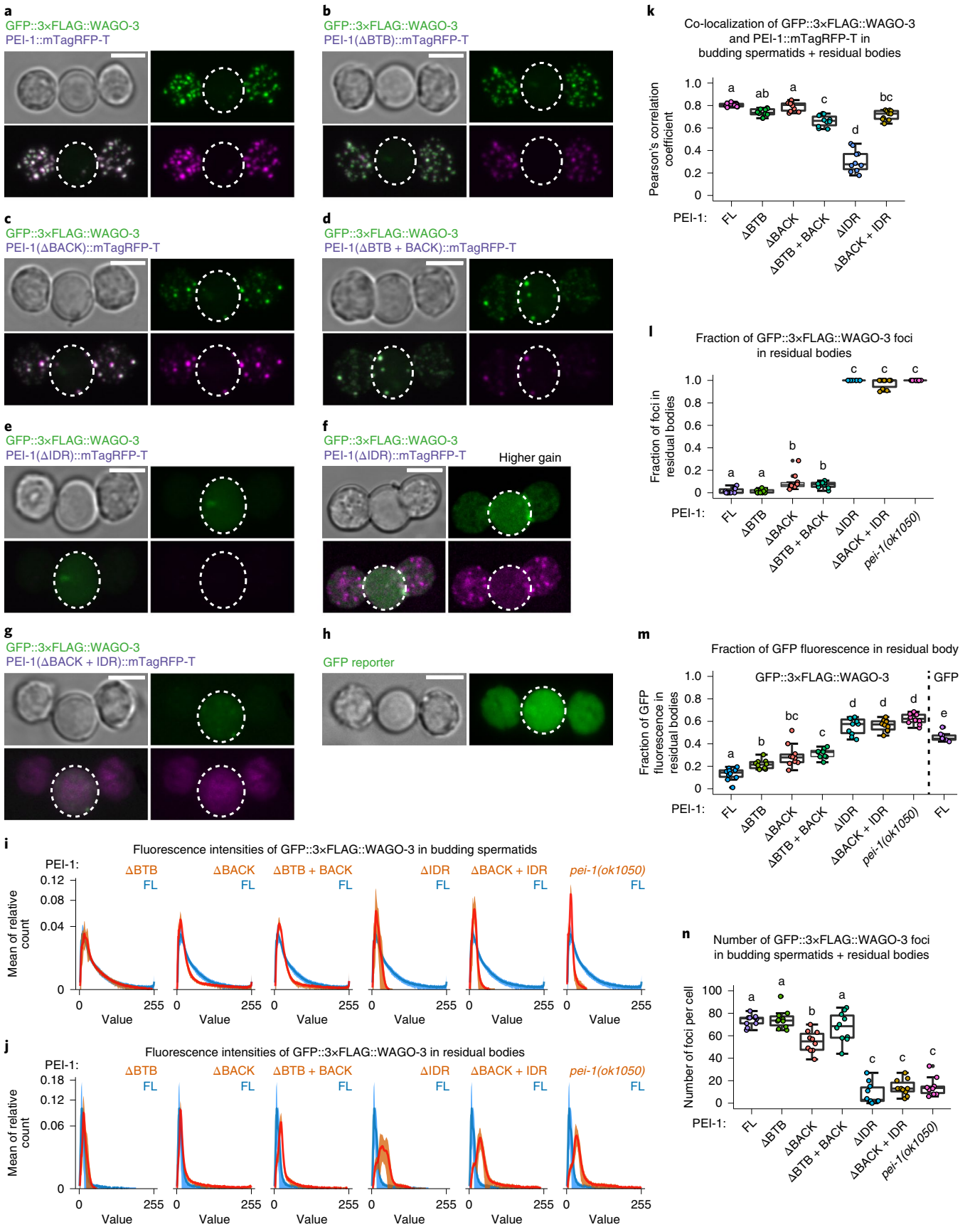
To resolve PEI granule localization at a high resolution, we used correlative light and electron microscopy (CLEM) (Fig. 6a–e and Extended Data Fig. 9a–c; see the BigDataViewer at https://doi.org/10.17632/dgb8d7h2hz.1, further details about the BigData Viewer are provided in the Supplementary Information). In early primary spermatocytes, when the MOs are just starting to form from the Golgi, and no FBs are associated yet[26], PEI granules were found close to and overlapping with MOs (Fig. 6a). This situation remained in the later stages of spermatogenesis: when the FB showed its typical fibrous structure and was enwrapped by the MO (Fig. 6b,c); when the MO started to retract from the FB (Fig. 6d); and when FBs were fully released from the MOs (Fig. 6e). From a total of 10 precise CLEM overlays, we found that 18 out of 18 and 17 out of 17 foci in spermatocytes and spermatids, respectively, were positioned immediately next to, or overlapping with, an MO. In both stages, only three of these foci were found to also contact a mitochondrion. We conclude that PEI granules associate with MOs.

PEI granule segregation affects epigenetic inheritance. We identified R09A1.2, or PEI-2, as a PEI-1-interacting protein (Fig. 7a and Extended Data Fig. 10a) and revealed spermatogenesis-specific expression of PEI-2 (Fig. 7b,c), with strong co-localization to PEI-1 (Fig. 7d–f). Similar to PEI-1, PEI-2 has a BTB and BACK domain (Fig. 7g), and through co-IP experiments in a heterologous cell culture system, we showed that the BTB and BACK domains of PEI-2 interact with PEI-1 (Extended Data Fig. 10b).

In *pei-2* mutants, FB-MOs, as visualized by the ΔH15–Q558 PEI-1 deletion (Fig. 5a,b), segregated normally into spermatids (Extended Data Fig. 10c,d). Moreover, PEI granules still formed (Extended Data Fig. 10e) and still recruited WAGO-3 (Fig. 7h,i and Extended Data Fig. 10e–g). However, in budding spermatids, these PEI granules did not properly segregate and were often lost in the residual bodies (Fig. 7h–l), similar to *spe-15* mutants (Fig. 5e). Mutants lacking SPE-15 are sterile, making it impossible to test the relevance of PEI-granule segregation in paternal epigenetic inheritance using the Mis phenotype. However, *pei-2* mutants are fertile, enabling us to reveal that effective PEI-granule segregation is also required for epigenetic inheritance by sperm (Fig. 7m).

Segregation of PEI granules requires *S*-palmitoylation. *S*-palmitoylation can guide proteins to membranes and typically occurs on Golgi-related membranes[51]. The palmitoyltransferase

Fig. 4 | The IDR of PEI-1 is required for WAGO-3 segregation into spermatids. a–h, Confocal maximum intensity projections of isolated, male-derived budding spermatids expressing GFP::3xFLAG::WAGO-3 (**a–h**) PEI-1::mTagRFP-T (**a**), PEI-1(ΔBTB)::mTagRFP-T (**b**), PEI-1(ΔBACK)::mTagRFP-T (**c**), PEI-1(ΔBTB + BACK)::mTagRFP-T (**d**), PEI-1(ΔIDR)::mTagRFP-T (**e** and **f**), PEI-1(ΔBACK + IDR)::mTagRFP-T (**g**) and a GFP reporter (**h**). For **a–g**, GFP::3xFLAG::WAGO-3 appears in green and PEI-1::mTagRFP-T variants appear in magenta. Residual bodies are indicated by dashed circles. For **a–h**, scale bars, 4 µm. The image in **f** was acquired with higher gain compared with the image in **e** to visualize the remaining PEI-1::mTagRFP-T signal within the budding spermatids. **h**, Free GFP was expressed from the *wago-3* locus. **i,j**, The fluorescence intensity (*x* axis) versus the mean of relative pixel count (*y* axis) of GFP::3xFLAG::WAGO-3 signal in budding spermatids (**i**) and residual bodies (**j**) expressing the indicated PEI-1::mTagRFP-T variants. $n = 10$ cells pooled from two independent experiments for each condition. The relative pixel count is the number of pixels with a given intensity within a selected region, divided by the total number of pixels in that region. In each plot, the curve derived from full-length PEI-1::mTagRFP-T is also shown in blue. The width of the curves reflects the s.d. of the mean. **k–n**, Quantification of the co-localization of GFP::3xFLAG::WAGO-3 and PEI-1::mTagRFP-T in budding spermatids + residual bodies (**k**), the fraction of GFP::3xFLAG::WAGO-3 foci in residual bodies (**l**), total GFP::3xFLAG::WAGO-3 signal in residual bodies (**m**) and GFP::3xFLAG::WAGO-3 foci number in budding spermatids + residual bodies (**n**) of male-derived cells expressing indicated PEI-1::mTagRFP-T variants. $n = 10$ cells pooled from two independent experiments for each condition. Statistically significant differences were determined using one-way ANOVA ($P \leq 0.001$) followed by Tukey's honest significant difference post hoc test ($P \leq 0.05$). Different letters represent significant differences. The exact *P* values are provided as source data. The box plots show the median (centre line), 25th or 75th percentiles (box edges), and the whiskers indicate the median ± 1.5 × interquartile range. Note that the full-length data in **k**, **l**, **m** and **n** are the same as the wild-type data shown in Fig. 7i, Fig. 7k, Fig. 7l,j and Extended Data Fig. 4c (budding spermatid), respectively. Representative images from two biologically independent experiments are shown in **a–h**. Source data are available online.


a
GFP::3×FLAG::WAGO-3
PEI-1::mTagRFP-T
b
GFP::3×FLAG::WAGO-3
PEI-1(ΔBTB)::mTagRFP-T
c
GFP::3×FLAG::WAGO-3
PEI-1(ΔBACK)::mTagRFP-T
d
GFP::3×FLAG::WAGO-3
PEI-1(ΔBTB + BACK)::mTagRFP-T
e
GFP::3×FLAG::WAGO-3
PEI-1(ΔIDR)::mTagRFP-T
f
GFP::3×FLAG::WAGO-3
PEI-1(ΔIDR)::mTagRFP-T
Higher gain
g
GFP::3×FLAG::WAGO-3
PEI-1(ΔBACK + IDR)::mTagRFP-T
h
GFP reporter
k
Co-localization of GFP::3×FLAG::WAGO-3 and PEI-1::mTagRFP-T in budding spermatids + residual bodies
Pearson's correlation coefficient
PEI-1: FL ΔBTB ΔBACK ΔBTB + BACK ΔIDR ΔBACK + IDR
l
Fraction of GFP::3×FLAG::WAGO-3 foci in residual bodies
Fraction of foci in residual bodies
PEI-1: FL ΔBTB ΔBACK ΔBTB + BACK ΔIDR ΔBACK + IDR pei-1(ok1050)
m
Fraction of GFP fluorescence in residual body
GFP::3×FLAG::WAGO-3
GFP
Fraction of GFP fluorescence in residual bodies
PEI-1: FL ΔBTB ΔBACK ΔBTB + BACK ΔIDR ΔBACK + IDR pei-1(ok1050) FL
i
Fluorescence intensities of GFP::3×FLAG::WAGO-3 in budding spermatids
PEI-1: ΔBTB FL ΔBACK FL ΔBTB + BACK FL ΔIDR FL ΔBACK + IDR FL pei-1(ok1050) FL
Mean of relative count
Value
j
Fluorescence intensities of GFP::3×FLAG::WAGO-3 in residual bodies
PEI-1: ΔBTB FL ΔBACK FL ΔBTB + BACK FL ΔIDR FL ΔBACK + IDR FL pei-1(ok1050) FL
Mean of relative count
Value
n
Number of GFP::3×FLAG::WAGO-3 foci in budding spermatids + residual bodies
Number of foci per cell
PEI-1: FL ΔBTB ΔBACK ΔBTB + BACK ΔIDR ΔBACK + IDR pei-1(ok1050)

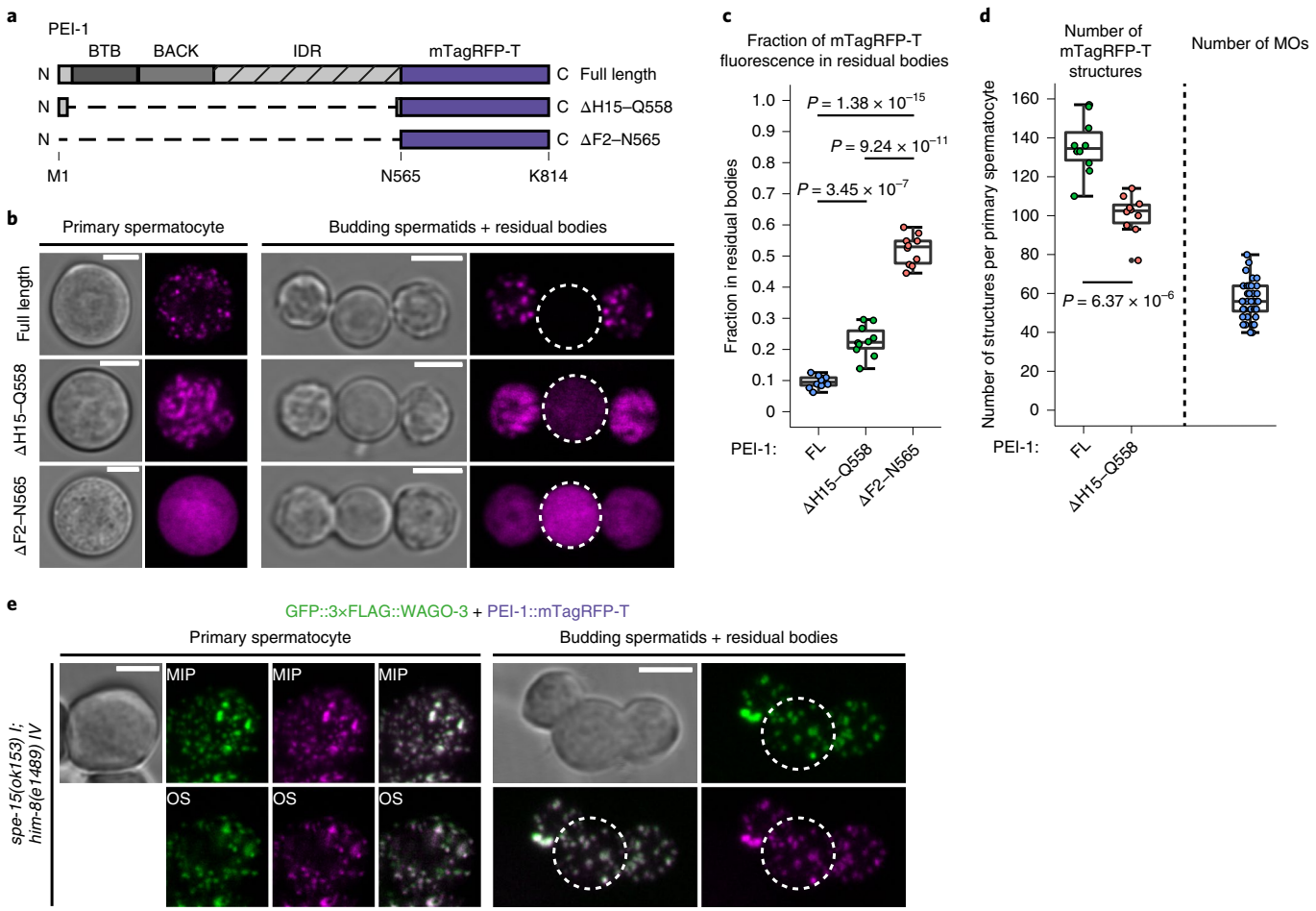


Fig. 5 | PEI granule segregation is dependent on SPE-15. a, Schematic of two mTagRFP-T proteins generated using CRISPR–Cas9-mediated genome editing of the *pei-1* locus. The ΔH15–Q558 deletion leaves 14 and 7 amino acids of the PEI-1 N and C terminus, respectively. The ΔF2–N565 deletion removes all PEI-1-specific amino acids. **b**, Confocal micrographs of isolated, male-derived spermatocytes (left) and budding spermatids (right) each expressing one of the proteins shown in **a**. Residual bodies are indicated by the dashed circles. The images represent two biologically independent experiments. Scale bars, 4 µm. **c**, The fraction of total mTagRFP-T signal within the residual bodies of male-derived budding spermatids expressing indicated proteins. $n = 10$ cells, which were pooled from multiple experiments for each condition. Note that the ΔH15–Q558 data are the same as the those displayed in Extended Data Fig. 10c (wild type). **d**, Quantification of mTagRFP-T structures in isolated, male-derived primary spermatocytes expressing the indicated proteins. $n = 10$ cells pooled from two independent experiments for each condition. Right, extrapolated number of FB-MOs in primary spermatocytes based on LysoSensor Blue DND-192 staining in spermatids. $n = 36$ cells pooled from two independent experiments. For **c** and **d**, statistically significant differences were determined using one-way ANOVA ($P \leq 0.001$) followed by Tukey's honest significant difference post hoc test ($P \leq 0.05$). The box plots show the median (centre line), 25th or 75th percentiles (box edges), and the whiskers indicate the median ± 1.5 × interquartile range. Representative images are shown in **b**. **e**, Confocal maximum intensity projections and optical sections of isolated, male-derived spermatocytes and budding spermatids expressing GFP::3xFLAG::WAGO-3 and PEI-1::mTagRFP-T in the *spe-15(ok153)* mutant background. The strain contained a *him* mutation to increase the frequency of males in the cultures. Residual bodies are indicated by the dashed circles. The images represent two biologically independent experiments. MIP, maximum intensity projection; OS, optical section. Scale bars, 4 µm. Source data are available online.

SPE-10 localizes to MOs, and is required for their stable interaction with FBs[52]. We found that PEI granules were severely defective in *spe-10* mutants. Large and irregularly shaped PEI granules formed along the cell periphery in *spe-10* spermatocytes (Fig. 8a). Similar to the much smaller wild-type PEI granules, these structures were static and did not show signs of fusion or fission (Extended Data Movie 2). At the later stages, large PEI-1 aggregates were detected in the residual body (Fig. 8a). WAGO-3 and PEI-1 still co-localized in the absence of SPE-10 (Fig. 8a), indicating that *S*-palmitoylation affects the subcellular localization of PEI granules, but not their recruitment of WAGO-3.

Western blotting showed a clear doublet band for PEI-2 (Fig. 8b), compatible with palmitoylation[53,54]. Interestingly, the top band of PEI-2 disappeared in *spe-10* mutants (Fig. 8b), suggesting PEI-2 is a substrate of the SPE-10 enzyme. Our western blot analysis also revealed evidence for PEI-1 modification, although this appears more as a smear than as a discrete band (Fig. 8c). PEI-1 modification did not depend on SPE-10 (Fig. 8c), suggesting that PEI-1 may carry a different kind of modification, or that another palmitoyltransferase may act on PEI-1. Strikingly, PEI-1 and PEI-2 affected each other's modification status—although PEI-2 was required for PEI-1 modification (Fig. 8c), PEI-1 inhibited PEI-2 modification (Fig. 8b). In a heterologous expression system, PEI-1 and PEI-2 also appeared as doublets, and showed decreased stability after palmitoylation inhibition (Extended Data Fig. 10h,i), similar to what has been reported for the palmitoylated protein PD-L1 (ref. [55]). We conclude that PEI-2 is a potential direct substrate of SPE-10, and that PEI-1 can also be modified, but only in the presence of PEI-2.

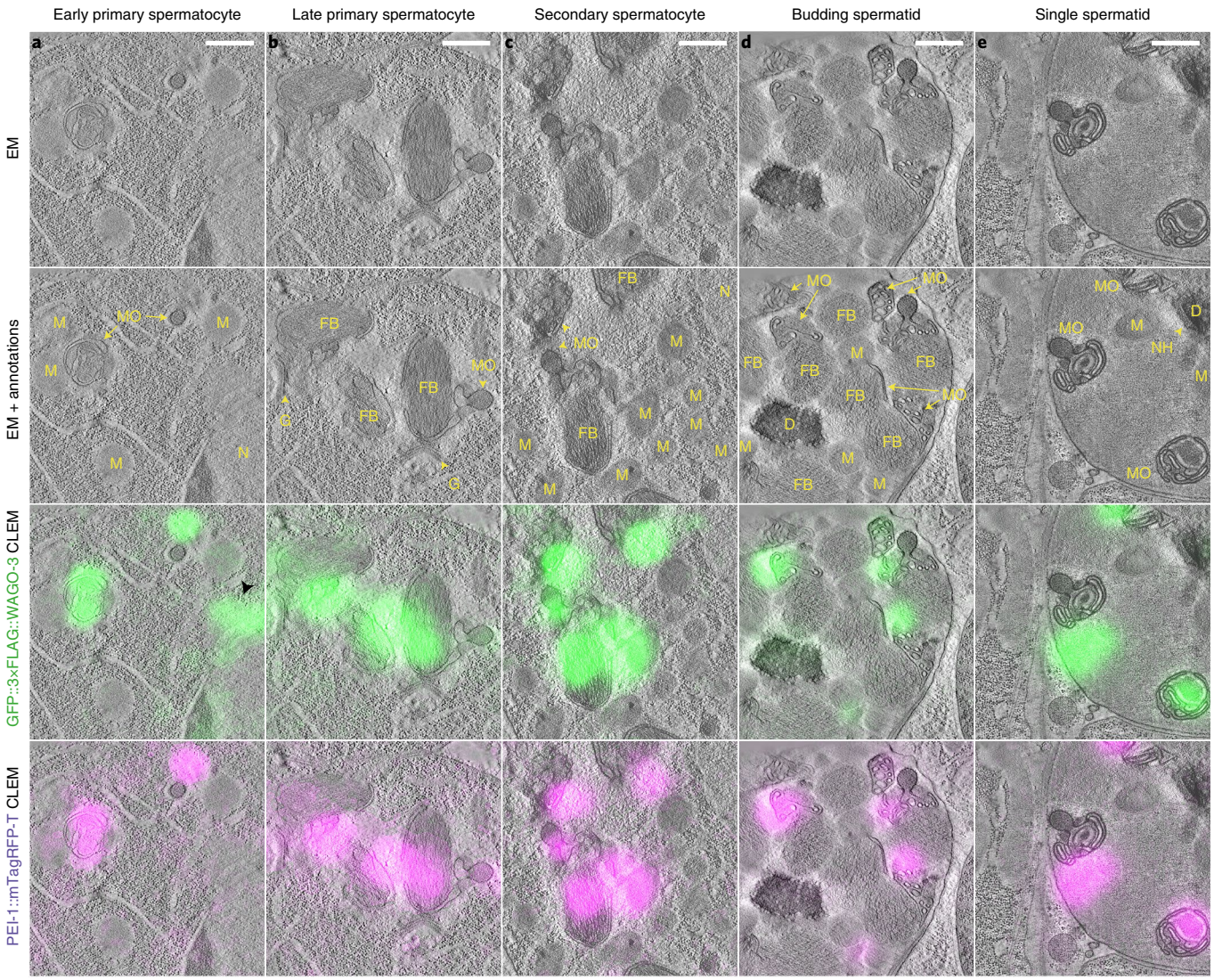


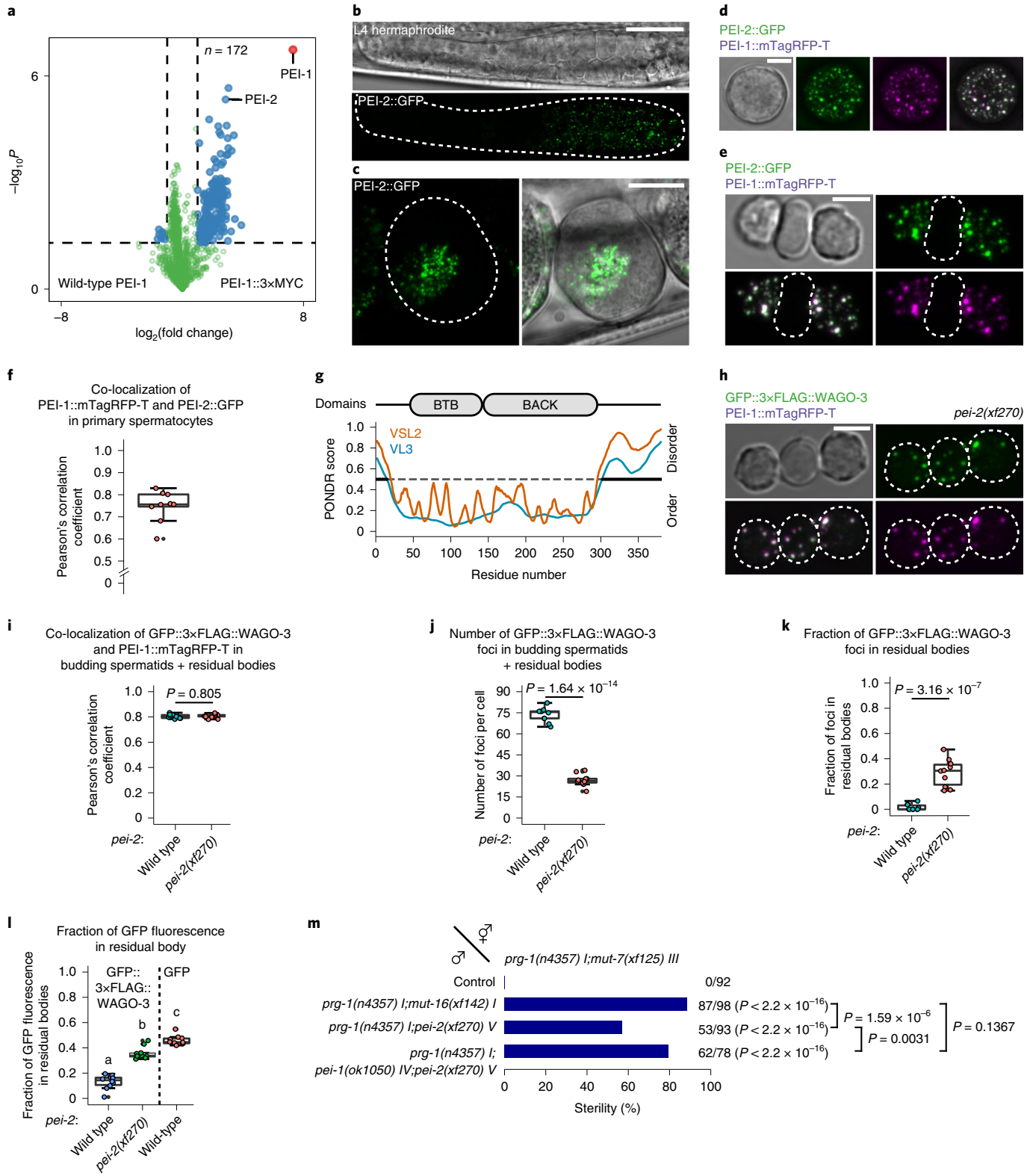
Fig. 6 | PEI granules are associated with membranous organelles. a–e, Representative dual-colour CLEM images acquired in germ cells of spermatogenesis (early primary spermatocyte (**a**), late primary spermatocyte (**b**) secondary spermatocyte (**c**), budding spermatid (**d**) and single spermatid (**e**)) of high-pressure frozen adult males expressing GFP::3xFLAG::WAGO-3 and PEI-1::mTagRFP-T. The four rows show electron microscopy (EM)-only images (row 1), EM with annotations (row 2), EM with GFP::3xFLAG::WAGO-3 fluorescence (row 3) and EM with PEI-1::mTagRFP-T fluorescence (row 4). The GFP::3xFLAG::WAGO-3-only positive focus in row 3 of **a** (black arrow head) is adjacent to the nucleus, and most likely represents a P granule. The images represent two biologically independent experiments. D, DNA; FB, fibrous body; G, Golgi complex; M, mitochondrion; N, nucleus; NH, peri-nuclear halo. For **a–e**, scale bars, 500 nm. Source data are available online.

Discussion

We identified a previously undescribed, sperm-specific compartment—the PEI granule—and define its role in sperm-borne cytoplasmic inheritance of a specific Argonaute protein, WAGO-3 (Fig. 8d). PEI granules are made by PEI-1 and PEI-2 proteins, which contain a BTB–BACK domain, followed by an IDR. Interestingly, BTB and BACK domains can mediate homo- and heteromeric oligomerization[40,43], providing multivalency, a property that is known to drive phase separation[24]. As has been found in other condensates[56], our data are consistent with the idea that BTB–BACK domain interactions help to stabilize IDR–IDR interactions, together driving PEI granule formation. As such, we consider PEI-1 and PEI-2 as scaffold proteins[24] of PEI granules. Note that the material state of the PEI granules has not been clarified. To examine whether PEI granules have a liquid character, are more gel-like or represent some other form of complex, experiments with purified proteins will be required.

The PEI-1 IDR also recruits WAGO-3, which we propose is a client[24] of PEI granules. Even though WAGO-1 and CSR-1 have been proposed to be also present in sperm[27,57], we found that N-terminally tagged versions of these proteins expressed from their endogenous loci accumulated in the residual bodies. We speculate that the PEI-1 IDR may create a condensate that is selective for some feature of the WAGO-3 protein, or possibly non-permissive for characteristics of depleted proteins, such as WAGO-1 and CSR-1. Such features could be sequence intrinsic, but could also relate to post-translational modifications. Further experiments will be required to test these ideas.

PEI granules interact with MOs through *S*-palmitoylation and PEI-2 is probably palmitoylated by SPE-10, providing an example of how acylation of germ granule components may promote their membrane-affinity. *S*-palmitoylation is reversible[51], raising the possibility that PEI granules as a whole may be released from the MOs after fertilization by depalmitoylation. However, other processes that may affect PEI granules and their cargo in the zygote



are the effect of dilution, which could directly affect PEI granule stability, and potential post-translational modification of PEI granules by maternal factors. We note that environmental cues, such as temperature, could also affect PEI granule behaviour and therefore paternal epigenetic inheritance.

When domain organization is considered, PEI-1-related proteins can be easily identified within nematodes (Fig. 8e) and in humans (such as BTBD7; Fig. 8f). Interestingly BTBD7 carries a predicted myristylation site close to its N terminus, suggesting that it might be membrane-associated like PEI-2. The other human proteins shown in Fig. 8f are also known to be expressed in the testis and, for two of these, functions during spermatogenesis have been described previously[58,59]. BTBD18 forms nuclear foci[58], and GMCL1 interacts with IDRs found in primate-specific GAGE proteins, and affects

Fig. 7 | PEI-2 affects PEI granule segregation. a, Label-free proteomic quantification of PEI-1::3xMYC IP experiments from late-L4 stage hermaphrodite extracts. See Fig. 1b for further information. **b**, Confocal micrograph of an L4 hermaphrodite expressing PEI-2::GFP. The dashed line indicates the gonad. Scale bar, 20 µm. **c**, Confocal maximum intensity projection of spermatozoa within the spermatheca of an adult hermaphrodite expressing PEI-2::GFP. The dashed line encloses the spermatheca. Scale bar, 10 µm. **d**,**e**, Confocal maximum intensity projection of an isolated male-derived primary spermatocyte (**d**) and budding spermatids (**e**) expressing PEI-1::mTagRFP-T and PEI-2::GFP. The dashed lines in **e** indicate the residual body. Scale bars, 4 µm. **f**, Co-localization analysis between PEI-1::mTagRFP-T and PEI-2::GFP in isolated male-derived primary spermatocytes. $n = 9$ cells pooled from two independent experiments. **g**, PEI-2 domain organization. Features are as described in Fig. 3a. **h**, Confocal maximum intensity projections of isolated male-derived budding spermatids expressing GFP::3xFLAG::WAGO-3 and PEI-1::mTagRFP-T in the *pei-2(xf270)*-mutant background. The dashed circles indicate cells. Scale bar, 4 µm. **i**–**l**, Quantification of GFP::3xFLAG::WAGO-3 and PEI-1::mTagRFP-T expression in wild-type and *pei-2(xf270)*-mutant, male-derived budding spermatids. $n = 10$ cells pooled from two independent experiments for each condition. **i**, Co-localization of GFP::3xFLAG::WAGO-3 and PEI-1::mTagRFP-T. **j**, The total number of GFP::3xFLAG::WAGO-3 foci. **k**, The fraction of GFP::3xFLAG::WAGO-3 foci within the residual body. **l**, The fraction of total GFP::3xFLAG::WAGO-3 signal within the residual body. The wild-type (**i**–**l**) and free GFP (**l**) data are the same as those shown as full-length data in Fig. 4k–n. Statistically significant differences were determined using one-way ANOVA ($P \leq 0.001$) followed by Tukey's honest significant difference post hoc test ($P \leq 0.05$). Different letters in **l** represent significant differences. **m**, The percentage of fertile F_1 animals generated by crosses between males and hermaphrodites with the indicated genotypes. See Fig. 2 for further information. Statistical significance was tested using a Pearson's χ^2 test with Yates continuity correction. For **f** and **i**–**l**, The box plots show the median (centre line), 25th or 75th percentiles (box edges), and the whiskers indicate the median ± 1.5 × interquartile range. The images in **b**–**e** and **h** represent two biologically independent experiments. Exact *P* values (**l**) and source data are available online.

their localization[60]. Thus, the mechanism that we revealed may be broadly conserved.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41556-021-00827-2.

Received: 17 August 2021; Accepted: 3 December 2021;
Published online: 7 February 2022

References

1. Bošković, A. & Rando, O. J. Transgenerational epigenetic inheritance. *Annu. Rev. Genet.* **52**, 21–41 (2018).
2. Perez, M. F. & Lehner, B. Intergenerational and transgenerational epigenetic inheritance in animals. *Nat. Cell Biol.* **21**, 143–151 (2019).
3. Hutvagner, G. & Simard, M. J. Argonaute proteins: key players in RNA silencing. *Nat. Rev. Mol. Cell Biol.* **9**, 22–32 (2008).
4. Peters, L. & Meister, G. Argonaute proteins: mediators of RNA silencing. *Mol. Cell* **26**, 611–623 (2007).
5. Castel, S. E. & Martienssen, R. A. RNA interference in the nucleus: Roles for small RNAs in transcription, epigenetics and beyond. *Nat. Rev. Genet.* **14**, 100–112 (2013).
6. Xu, F., Guang, S. & Feng, X. Distinct nuclear and cytoplasmic machineries cooperatively promote the inheritance of RNAi in *Caenorhabditis elegans*. *Biol. Cell* **110**, 217–224 (2018).
7. Wan, G. et al. Spatiotemporal regulation of liquid-like condensates in epigenetic inheritance. *Nature* **557**, 679–683 (2018).
8. de Albuquerque, B. F. M., Placentino, M. & Ketting, R. F. Maternal piRNAs are essential for germline development following de novo establishment of endo-siRNAs in *Caenorhabditis elegans*. *Dev. Cell* **34**, 448–456 (2015).
9. Phillips, C. M., Brown, K. C., Montgomery, B. E., Ruvkun, G. & Montgomery, T. A. PiRNAs and piRNA-dependent siRNAs protect conserved and essential *C. elegans* genes from misrouting into the RNAi pathway. *Dev. Cell* **34**, 457–465 (2015).
10. Buckley, B. A. et al. A nuclear Argonaute promotes multigenerational epigenetic inheritance and germline immortality. *Nature* **489**, 447–451 (2012).
11. Mao, H. et al. The Nrde pathway mediates small-RNA-directed histone H3 lysine 27 trimethylation in *Caenorhabditis elegans*. *Curr. Biol.* **25**, 2398–2403 (2015).
12. Gent, J. I. et al. Distinct phases of siRNA synthesis in an endogenous RNAi pathway in *C. elegans* soma. *Mol. Cell* **37**, 679–689 (2010).
13. Ketting, R. F., Haverkamp, T. H. A., Van Luenen, H. G. A. M. & Plasterk, R. H. A. *mut-7* of C. elegans, required for transposon silencing and RNA interference, is a homolog of Werner syndrome helicase and RNaseD. *Cell* **99**, 133–141 (1999).
14. Zhang, C. et al. *mut-16* and other mutator class genes modulate 22G and 26G siRNA pathways in *Caenorhabditis elegans*. *Proc. Natl Acad. Sci. USA* **108**, 1201–1208 (2011).
15. Phillips, C. M., Montgomery, T. A., Breen, P. C. & Ruvkun, G. MUT-16 promotes formation of perinuclear Mutator foci required for RNA silencing in the *C. elegans* germline. *Genes Dev.* **26**, 1433–1444 (2012).
16. Gu, W. et al. Distinct Argonaute-mediated 22G-RNA pathways direct genome surveillance in the *C. elegans* Germline. *Mol. Cell* **36**, 231–244 (2009).
17. Yigit, E. et al. Analysis of the *C. elegans* argonaute family reveals that distinct Argonautes act sequentially during RNAi. *Cell* **127**, 747–757 (2006).
18. Xu, F. et al. A cytoplasmic Argonaute protein promotes the inheritance of RNAi. *Cell Rep.* **23**, 2482–2494 (2018).
19. Ashe, A. et al. PiRNAs can trigger a multigenerational epigenetic memory in the germline of *C. elegans*. *Cell* **150**, 88–99 (2012).
20. Shirayama, M. et al. PiRNAs initiate an epigenetic memory of nonself RNA in the *C. elegans* germline. *Cell* **150**, 65–77 (2012).
21. Luteijn, M. J. et al. Extremely stable Piwi-induced gene silencing in *Caenorhabditis elegans*. *EMBO J.* **31**, 3422–3430 (2012).
22. Claycomb, J. M. et al. The Argonaute CSR-1 and Its 22G-RNA cofactors are required for holocentric chromosome segregation. *Cell* **139**, 123–134 (2009).
23. Updike, D. & Strome, S. P granule assembly and function in *Caenorhabditis elegans* germ cells. *J. Androl.* **31**, 53–60 (2010).
24. Banani, S. F., Lee, H. O., Hyman, A. A. & Rosen, M. K. Biomolecular condensates: organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.* **18**, 285–298 (2017).
25. Voronina, E., Seydoux, G., Sassone-Corsi, P. & Nagamori, I. RNA granules in germ cells. *Cold Spring Harb. Perspect. Biol.* **3**, a002774 (2011).
26. Ellis, R. E. & Stanfield, G. M. The regulation of spermatogenesis and sperm function in nematodes. *Semin. Cell Dev. Biol.* **29**, 17–30 (2014).
27. Conine, C. C. et al. Argonautes ALG-3 and ALG-4 are required for spermatogenesis-specific 26G-RNAs and thermotolerant sperm in *Caenorhabditis elegans*. *Proc. Natl Acad. Sci. USA* **107**, 3588–3593 (2010).
28. Grishok, A., Tabara, H. & Mello, C. C. Genetic requirements for inheritance of RNAi in *C. elegans*. *Science* **287**, 2494–2497 (2000).
29. Alcazar, R. M., Lin, R. & Fire, A. Z. Transmission dynamics of heritable silencing induced by double-stranded RNA in *Caenorhabditis elegans*. *Genetics* **180**, 1275–1288 (2008).
30. Lev, I. et al. Germ granules govern small RNA inheritance. *Curr. Biol.* **29**, 2880–2891.e4 (2019).
31. Ozata, D. M., Gainetdinov, I., Zoch, A., O'Carroll, D. & Zamore, P. D. PIWI-interacting RNAs: small RNAs with big functions. *Nat. Rev. Genet.* **20**, 89–108 (2019).
32. Bagijn, M. P. et al. Function, targets, and evolution of *Caenorhabditis elegans* piRNAs. *Science* **337**, 574–578 (2012).
33. Lee, H. C. et al. *C. elegans* piRNAs mediate the genome-wide surveillance of germline transcripts. *Cell* **150**, 78–87 (2012).
34. Gudipati, R. K. et al. Protease-mediated processing of Argonaute proteins controls small RNA association. *Mol. Cell* https://doi.org/10.1016/j.molcel.2021.03.029 (2021).
35. Robert, V. P. V., Sijen, T., van Wolfswinkel, J. & Plasterk, R. H. A. Chromatin and RNAi factors protect the *C. elegans* germline against repetitive sequences. *Genes Dev.* **19**, 782–787 (2005).
36. Vastenhouw, N. L. et al. A genome-wide screen identifies 27 genes involved in transposon silencing in *C. elegans*. *Curr. Biol.* **13**, 1311–1316 (2003).
37. Stoeckius, M., Grün, D. & Rajewsky, N. Paternal RNA contributions in the *Caenorhabditis elegans* zygote. *EMBO J.* **33**, 1740–1750 (2014).
38. Barucci, G. et al. Small RNA-mediated transgenerational silencing of histone genes impairs fertility in piRNA mutants. *Nat. Cell Biol.* **22**, 235–245 (2020).

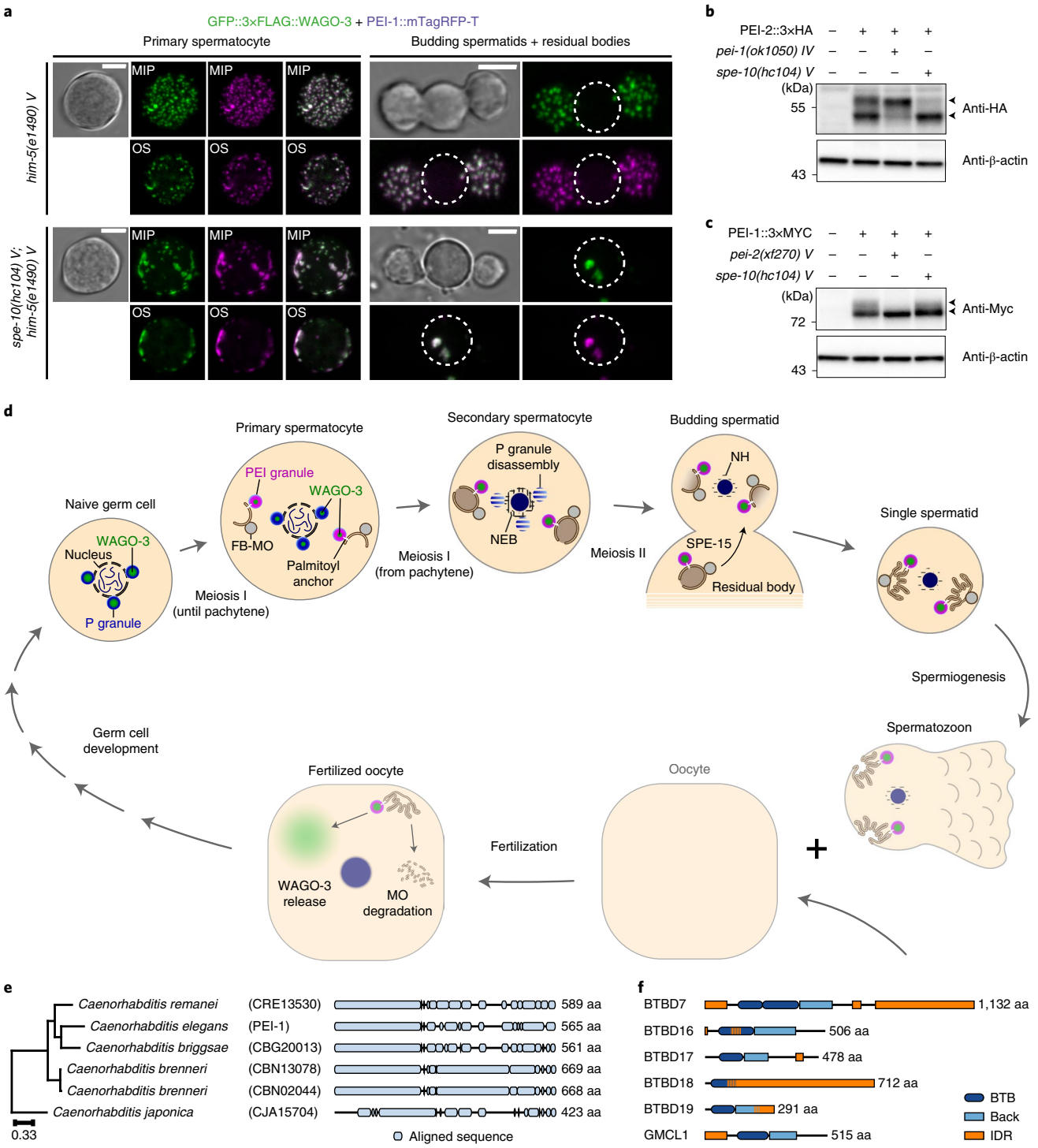


Fig. 8 | PEI granule formation and segregation depends on *S*-palmitoylations. a, Confocal maximum intensity projections and optical sections of isolated male-derived spermatocytes and budding spermatids expressing GFP::3xFLAG::WAGO-3 and PEI-1::mTagRFP-T in the indicated mutants. Strains contained *him* mutations to increase the frequency of males in the cultures. The dashed circles indicate residual bodies. The images represent two biologically independent experiments. Scale bars, 4 µm. **b**,**c**, Whole-worm extracts of late-L4 stage hermaphrodites were separated by SDS–PAGE, followed by western blot transfer and chemiluminescence detection of PEI-2::3xHA (**b**) and PEI-1::3xMYC (**c**) in the indicated mutants. The doublet signals of both proteins are indicated by arrowheads. β-Actin was used as a loading control. Unprocessed original scans of blots are provided as source data. The experiment in **b** and **c** was performed once. **d**, We provide data for the top half of the model: WAGO-3 starts in P granules, gradually moves to PEI granules that associate with FB-MOs through *S*-palmitoylation, which ensures spermatid localization through SPE-15 dependent transport. The bottom half is hypothetical, and is depicted in reduced opacity. We speculate that PEI granules release their content into the oocyte, helping to establish/maintain silencing of specific targets (Extended Data Fig. 1). The schematic is not to scale. **e**, Phylogenetic analysis showing PEI-1 conservation within the *Caenorhabditis* genus. The phylogenetic tree was generated using EggNOG (v.4.5.1). PEI-1 was defined as a query and compared with all eukaryote entries. aa, amino acids. **f**, Protein length and domain composition of six human BTB-domain-containing proteins that resemble the PEI-1 protein composition.

39. Spike, C. A., Bader, J., Reinke, V. & Strome, S. DEPS-1 promotes P-granule assembly and RNA interference in *C. elegans* germ cells. *Development* **135**, 983–993 (2008).
40. Collins, T., Stone, J. R. & Williams, A. J. All in the family: the BTB/POZ, KRAB, and SCAN domains. *Mol. Cell. Biol.* **21**, 3609–3615 (2001).
41. Stogios, P. J. & Privé, G. G. The BACK domain in BTB-kelch proteins. *Trends Biochem. Sci.* **29**, 634–637 (2004).
42. Kroschwald, S., Maharana, S. & Simon, A. Hexanediol: a chemical probe to investigate the material properties of membrane-less compartments. *Matters* https://doi.org/10.19185/matters.201702000010 (2017).
43. Marzahn, M. R. et al. Higher-order oligomerization promotes localization of SPOP to liquid nuclear speckles. *EMBO J.* **35**, 1254–1275 (2016).
44. Brangwynne, C. P. et al. Germline P granules are liquid droplets that localize by controlled dissolution/condensation. *Science* **324**, 1729–1732 (2009).
45. Putnam, A., Cassani, M., Smith, J. & Seydoux, G. A gel phase promotes condensation of liquid P granules in *Caenorhabditis elegans* embryos. *Nat. Struct. Mol. Biol.* **26**, 220–226 (2019).
46. Wang, J. et al. A molecular grammar governing the driving forces for phase separation of prion-like RNA binding proteins. *Cell* **174**, 688–699 (2018).
47. Hanazawa, M., Yonetani, M. & Sugimoto, A. PGL proteins self associate and bind RNPs to mediate germ granule assembly in *C. elegans*. *J. Cell Biol.* **192**, 929–937 (2011).
48. Kawasaki, I. et al. The PGL family proteins associate with germ granules and function redundantly in *Caenorhabditis elegans* germline development. *Genetics* **167**, 645–661 (2004).
49. Kawasaki, I. et al. PGL-1, a predicted RNA-binding component of germ granules, is essential for fertility in *C. elegans*. *Cell* **94**, 635–645 (1998).
50. Kelleher, J. F. et al. Myosin VI is required for asymmetric segregation of cellular components during *C. elegans* spermatogenesis. *Curr. Biol.* **10**, 1489–1496 (2000).
51. Tabaczar, S., Czogalla, A., Podkalicka, J., Biernatowska, A. & Sikorski, A. F. Protein palmitoylation: palmitoyltransferases and their specificity. *Exp. Biol. Med.* **242**, 1150–1157 (2017).
52. Gleason, E. J., Lindsey, W. C., Kroft, T. L., Singson, A. W. & L'Hernault, S. W. Spe-10 encodes a DHHC-CRD zinc-finger membrane protein required for endoplasmic reticulum/golgi membrane morphogenesis during *Caenorhabditis elegans* spermatogenesis. *Genetics* **172**, 145–158 (2006).
53. Gonzalo, S. & Linder, M. E. SNAP-25 palmitoylation and plasma membrane targeting require a functional secretory pathway. *Mol. Biol. Cell* **9**, 585–597 (1998).
54. Fukata, Y., Bredt, D. S. & Fukata, M. in *The Dynamic Synapse: Molecular Methods in Ionotropic Receptor Biology* (eds Kittler, J. T. & Moss, S. J.) Ch. 5 (CRC Press/Taylor & Francis, 2006).
55. Yao, H. et al. Inhibiting PD-L1 palmitoylation enhances T-cell immune responses against tumours. *Nat. Biomed. Eng.* **3**, 306–317 (2019).
56. Wu, H. & Fuxreiter, M. The structure and dynamics of higher-order assemblies: amyloids, signalosomes, and granules. *Cell* **165**, 1055–1066 (2016).
57. Conine, C. C. et al. Argonautes promote male fertility and provide a paternal memory of germline gene expression in *C. elegans*. *Cell* **155**, 1532–1544 (2013).
58. Zhou, L. et al. BTBD18 regulates a subset of piRNA-generating loci through transcription elongation in mice. *Dev. Cell* **40**, 453–466 (2017).
59. Kleiman, S. E. et al. Reduced human germ cell-less (HGCL) expression in azoospermic men with severe germinal cell impairment. *J. Androl.* **24**, 670–675 (2003).
60. Gjerstorff, M. F. et al. GAGE cancer-germline antigens are recruited to the nuclear envelope by germ cell-less (GCL). *PLoS ONE* **7**, e45819 (2012).

Methods

***C. elegans* culture and strains.** Unless otherwise stated, all worm strains were cultured according to standard laboratory conditions at 20 °C on nematode growth medium (NGM) plates seeded with *Escherichia coli* OP50 (ref. [61]). Animals for IP–MS/MS experiments were grown on egg plates (diameter, 90 mm)[62] for one generation, synchronized by bleaching, and then grown on standard NGM plates (diameter, 90 mm) for one generation before collection. Egg plates were generated by thoroughly mixing egg yolk with 50 ml LB medium/egg. After incubation at 65 °C for 2–3 h, the mixture was allowed to cool to room temperature before adding 10 ml OP50 culture per egg. About 10 ml was put on top of standard NGM plates (diameter, 90 mm) and incubated at room temperature. The next day, excess liquid was decanted and egg plates were incubated at room temperature for another 2 d. All of the strains are in the N2 Bristol background. A list of the strains used in this study is provided in Supplementary Table 1.

Mortal germline assay. All of the mutant strains were out-crossed four times with wild-type animals before starting the experiment, to clean the genetic background from potential mutations that occurred during culturing. For each strain, 90 L2 or L3 animals were distributed to 15 NGM plates (diameter, 90 mm), resulting in six larvae per plate. Animals were grown at 25 °C. Worms were picked onto fresh plates every second generation. The experiment was stopped after 17 generations.

RNAe reporter reactivation. A *mCherry::H2B* (RNAe) transgene (*mjSi22*) from the *prg-1(n4357)*-mutant background was crossed into *wago-3(pk1673)*- and *mut-7(xf125)*-mutant backgrounds to generate *prg-1(n4357);wago-3(pk1673/+)* and *prg-1(n4357);mut-7(xf125/+)* animals that also were homozygous for the integrated *mjSi22* transgene. From the offspring of these animals, we identified homozygous mutant and homozygous wild-type animals using PCR. From these, ten offspring were singled, and 50 of their offspring were analysed by microscopy for *mjSi22* expression. If not all animals were expressing the *mjSi22* transgene, again 10 non-expressing animals were singled, of which the offspring were scored by microscopy.

***Mutator*-induced sterility crosses.** All strains were confirmed and out-crossed twice before setting up crosses. Note that out-crossing ensured comparable results as an enhanced Mis phenotype was observed when using non-out-crossed animals. The transgenic allele *otIs45[unc-119p::gfp] V* was always present in paternal strains and served as mating control to avoid picking self-fertilized offspring. Only L2-stage F_1 animals were picked onto individual plates to avoid any biased selection. After 3 d, male or dead F_1 animals were excluded from the analysis. The fertility of F_1 animals was determined by the presence of F_2 animals after another 2–4 d.

CRISPR–Cas9-mediated genome editing. All protospacer sequences were chosen using CRISPOR (http://crispor.tefor.net)[63] and, unless otherwise stated, cloned in either pRK2411 (plasmid expressing *Cas9* + sgRNA(F+E)[64]; derived from pDD162) or pRK2412 (plasmid expressing sgRNA(F+E)[64] with *Cas9* deleted; derived from pRK2411) by site-directed, ligase-independent mutagenesis (SLIM)[65,66]. pDD162 (*Peft-3::Cas9* + empty sgRNA) was a gift from B. Goldstein (Addgene plasmid, 47549)[67]. SLIM reactions were transformed in subcloning efficiency DH5α competent cells (18265017, Invitrogen) and plated on LB agar plates supplemented with 100 μg ml^{-1} ampicillin. A list of all of the protospacer sequences is provided in Supplementary Table 2.

Insertions of a *gfp::3flag* sequence were based on plasmid DNA donor templates containing a self-excising drug selection cassette (SEC), which were designed and cloned as previously described[68]. pDD282 was a gift from B. Goldstein (Addgene plasmid, 66823)[68]. pJW1259 was used as *Cas9* plasmid and was a gift from J. Ward (Addgene plasmid, 61251)[69]. pGH8, pCFJ90 and pCFJ104 served as co-injection markers and were gifts from E. Jorgensen (Addgene plasmid, 19359; Addgene plasmid, 19327; Addgene plasmid, 19328)[70]. All plasmids were purified from 4 ml bacterial culture using either NucleoSpin Plasmid (740588.50, Macherey-Nagel) or the PureLink HiPure Plasmid Miniprep Kit (K210011, Invitrogen), eluted in sterile water and confirmed by enzymatic digestion and sequencing.

PCR products served as linear, double-stranded DNA donor templates for the insertion of *mTagRfp-t* and *gfp* sequences. The *mTagRfp-t* coding sequence, including three introns and flanking homology regions, was amplified from pDD286, which was a gift from B. Goldstein (Addgene plasmid, 70684). The *gfp* coding sequence, including three introns and flanking homology regions, was amplified from pDD282. All PCR products were purified using the QIAquick PCR Purification Kit (28106, Qiagen), eluted in sterile water and confirmed by agarose gel electrophoresis. For all epitope tag insertions, co-conversions and precise deletions, we ordered 4 nmol Ultramer DNA oligodeoxynucleotides from Integrated DNA Technologies, which serves as linear, single-stranded DNA (ssODN) donor templates. All Ultramer DNA oligodeoxynucleotides were resuspended in sterile water. All linear DNA donor templates contained ~35 bp homology regions[71,72]. A list of all of the DNA donor templates is provided in Supplementary Table 3.

To generate the *mut-16-*, *mut-7-* and *pei-2*-deletion alleles, we injected animals with 50 ng μl^{-1} pJW1259, 30 ng μl^{-1} of each sgRNA(F+E), 10 ng μl^{-1} pGH8, 5 ng μl^{-1} pCFJ104 and 2.5 ng μl^{-1} pCFJ90. F_1 animals expressing all three co-injection markers were selected for subsequent screening of deletion alleles. To insert *gfp::3×Flag* sequences, injection mixes included 50 ng μl^{-1} pJW1259, 50 ng μl^{-1} sgRNA(F+E), 10 ng μl^{-1} SEC donor plasmid, 10 ng μl^{-1} pGH8, 5 ng μl^{-1} pCFJ104 and 2.5 ng μl^{-1} pCFJ90. Screening of F_1 animals was performed as previously described[68]. All other CRISPR–Cas9-mediated genome editing was performed using either *dpy-10(cn64)* or *unc-58(e665)* co-conversion strategies[73]. To insert epitope tag or protospacer sequences, we injected 50 ng μl^{-1} Cas9 + sgRNA(F+E) (co-conversion), 50 ng μl^{-1} sgRNA(F+E) (gene of interest), 750 nM ssODN donor 1 (co-conversion) and 750 nM ssODN donor 2 (gene of interest). To insert a *mTagRfp-t* sequence in *pei-1* and *pgl-1*, we first transplanted the protospacer sequence used for the *dpy-10* co-conversion directly upstream of the respective stop codon to generate *d10*-entry strains[74]. These strains served as reference strains for the insertion of a *mTagRfp-t* sequence by injecting 50 ng μl^{-1} Cas9 + sgRNA(F+E) (*dpy-10* co-conversion), 1,000 nM ssODN donor (*dpy-10* co-conversion) and 300 ng μl^{-1} linear, double-stranded DNA donor. Similarly, a *gfp* sequence was inserted in *pei-2*. Precise deletions in *pei-1* and *wago-3* were generated by injecting 50 ng μl^{-1} Cas9 + sgRNA(F+E) (co-conversion), 50 ng μl^{-1} of each sgRNA(F+E) (gene of interest), 750 nM ssODN donor 1 (co-conversion) and 750 nM ssODN donor 2 (gene of interest). Unless otherwise stated, DNA injection mixes were injected in both gonad arms of five to 20 young adult N2 hermaphrodites maintained at 20 °C. Selected F_1 progeny were screened for insertion or deletion by PCR. Successful editing events were confirmed by Sanger sequencing. All generated mutant strains were out-crossed at least twice before any further cross or analysis.

Immunoprecipitation experiments. Unless otherwise stated, synchronized animals were cultured at 20 °C until late-L4 stage, collected with M9 buffer and frozen on dry ice in sterile water and 200 μl aliquots. Aliquots were thawed on ice, mixed with same volume of 2× lysis buffer (50 mM Tris HCl pH 7.5, 300 mM NaCl, 3 mM $MgCl_2$, 2 mM dithiothreitol (DTT), 0.2% Triton X-100, cOmplete Mini EDTA-free Protease Inhibitor Cocktail (11836170001, Roche)) and sonicated using a Bioruptor Plus device (B01020001, Diagenode) (4 °C, 10 cycles, 30 s on and 30 s off). After centrifugation for 10 min at 4 °C and 21,000*g*, supernatants were carefully transferred into a fresh tube without taking any material from the pellet or lipid phase. Pellet fractions were washed three times in 1× lysis buffer and resuspended in 1× Novex NuPAGE LDS sample buffer (NP0007, Invitrogen) supplemented with 100 mM DTT. Total protein concentrations of soluble worm extracts were determined using the Pierce BCA Protein-Assay (23225, Thermo Fisher Scientific) and an Infinite M200 Pro plate reader (Tecan). Extracts were diluted with 1× lysis buffer to reach 550 μl and a total protein concentration of 3 μg μl^{-1}. For each sample, 50 μl of this extract was added to 50 μl 1× Novex NuPAGE LDS sample buffer supplemented with 100 mM DTT and served as input control sample. For each IP experiment, 30 μl Novex DYNAL Dynabeads Protein G (10004D, Invitrogen) were washed three times with 500 μl 1× wash buffer (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), combined with the remaining 500 μl extract and incubated with rotation for 1 h at 4 °C. In the meantime, 8 μg antibody (monoclonal anti-Flag M2, F3165, Sigma-Aldrich; anti-Myc (9B11) mouse monoclonal antibodies, 2276, Cell Signaling Technology; monoclonal anti-HA (12CA5) mouse antibody, in-house production) was conjugated to another 30 μl Novex DYNAL Dynabeads Protein G according to the manufacturer's instructions. Extracts were separated from non-conjugated Dynabeads, combined with antibody-conjugated Dynabeads and incubated with rotation for 2 h at 4 °C. Following three washes with 500 μl 1× wash buffer, antibody-conjugated Dynabeads were resuspended in 25 μl 1.2× Novex NuPAGE LDS sample buffer supplemented with 120 mM DTT.

For RNA immunoprecipitation experiments, IPs were performed as described above with the following modifications: (1) adult animals were collected; (2) soluble worm extracts were diluted to 650 μl and a total protein concentration of 7 μg μl^{-1}, of which 150 μl served as input sample for later RNA extraction; and (3) antibody-conjugated Dynabeads were resuspended in 50 μl nuclease-free water.

For RIP experiments on males, synchronized *wago-3(xf119) I;him-5(1490) V* animals were cultured at 20 °C until adulthood. Adults were collected in M9 buffer and filtered through a 35 μm mesh using CellMicroSieves (35 μm pore size; N35R, BioDesign)[75]. Animals at the bottom of the mesh (>98% males) were collected and frozen on dry ice in sterile water and 200 μl aliquots.

Immunoprecipitation experiments associated with MS and small RNA sequencing were performed in quadruplicates and triplicates, respectively.

Western blot. Equal amounts of input samples (2%) and IP samples (10%) were adjusted to same volume with 1× Novex NuPAGE LDS sample buffer (NP0007, Invitrogen) supplemented with 100 mM DTT and incubated for 10 min at 95 °C. Together with PageRuler Prestained Protein Ladder (10–180 kDa, 26616, Thermo Fisher Scientific) or Color Prestained Protein Standard (10–250 kDa, P7719S, New England BioLabs), samples were separated on a Novex NuPAGE 4–12% Bis-Tris Mini Protein Gel (NP0323, Invitrogen) in 1× Novex NuPAGE MOPS SDS Running Buffer (NP0001, Invitrogen) at 50 mA. Proteins were next transferred on an Immobilon-P Membrane (PVDF, 0.45 μm, IPVH00010, Merck

Millipore) for 16 h at 15 V using a Mini Trans-Blot Cell (1703930, Bio-Rad) and 1× NuPAGE Transfer Buffer (NP0006, Invitrogen) supplemented with 20% methanol. After incubation in 1× PBS supplemented with 5% skimmed milk and 0.05% Tween-20 for 1 h, the PVDF membrane was cut according to the molecular mass of the proteins of interest. Each part was incubated in 1× PBS supplemented with 0.5% skimmed milk, 0.05% Tween-20 and the primary antibodies (1:5,000 monoclonal anti-Flag M2, F3165, Sigma-Aldrich; 1:1,000 anti-Myc-Tag (9B11) mouse monoclonal antibodies, 2276, Cell Signaling Technology; 1:5,000 anti-histone H3, H0164, Sigma-Aldrich; 1:1,000 polyclonal anti-Flag antibodies produced in rabbit, F7425, 086M4803V, Sigma-Aldrich; 1:1,000 Myc-Tag (71D10) Rabbit monoclonal antibodies, 2278, Cell Signaling Technology; 1:1,000 polyclonal anti-HA-Tag antibodies produced in rabbit, SAB4300603, 5117T501, Sigma-Aldrich; 1:1,000 β-actin (D6A8) rabbit monoclonal antibodies, 8457, Cell Signaling Technology; 1:2,500 monoclonal anti-α-tubulin antibodies produced in mouse, B-5-1-2, T6074, Sigma-Aldrich; 1:1,000 monoclonal anti-GFP antibodies (B-2), Santa Cruz, sc-9996, K1115; 1:1,000 polyclonal anti-RFP antibody pre-adsorbed, produced in rabbit, 33754, 600-4001-379, Rockland) for 1 h, followed by three washes with 1× PBS supplemented with 0.05% Tween-20 (hereafter, 0.05% PBS-T) for 10 min each, 1 h incubation in 0.05% PBS-T supplemented with the secondary antibody (1:10,000 anti-mouse IgG, HRP-linked antibodies, 7076, Cell Signaling Technology; 1:10,000 anti-rabbit IgG, HRP-linked antibodies, 7074, Cell Signaling Technology) and three final washes with 0.05% PBS-T for 10 min each. Chemiluminescence detection was performed using the Amersham ECL Select Western Blotting Detection Reagent (RPN2235, GE Healthcare) and a ChemiDoc XRS+ System (1708265, Bio-Rad). The samples shown in Fig. 8b–c were separated on a Novex NuPAGE 10% Bis-Tris Mini Protein Gel (NP0301, Invitrogen).

MS and proteome comparison. IP was performed in quadruplicate. After resuspending the precipitates in Novex NuPAGE LDS sample buffer (NP0007, Invitrogen), samples were incubated at 70 °C for 10 min and separated on a Novex NuPAGE 4–12% Bis-Tris Mini Protein Gel (NP0321, Invitrogen) in 1× Novex NuPAGE MOPS SDS Running Buffer (NP0001, Invitrogen) at 180 V for 10 min. After separation the samples were processed by in-gel digest as previously described[76,77]. After protein digest, the peptides were desalted using a C18 StageTip[78]. For measurement, the digested peptides were separated on a 25 cm reverse-phase capillary (inner diameter, 75 µm) packed with Reprosil C18 material (Dr. Maisch). Elution was carried out along a 2 h gradient of 2–40% of a mixture of 80% acetonitrile/0.5% formic acid using the EASY-nLC 1000 system (LC120, Thermo Fisher Scientific). A Q Exactive Plus mass spectrometer (Thermo Fisher Scientific) operated with a Top10 data-dependent MS/MS acquisition method per full scan was used for measurement[79]. Processing of the obtained results was performed with the MaxQuant software (v.1.5.2.8) against the Wormbase protein database (version WS263) for quantification[80]. The processed data were visualized with R and R-Studio using in-house scripts.

RNA extraction, library preparation and sequencing. RNA of input and GFP::3×Flag::WAGO-3 IP samples was extracted using TRIzol LS Reagenz (10296010, Invitrogen) according to the manufacturer's instructions, and resuspended in nuclease-free water. RNA quality and quantity was assessed using the Bioanalyzer RNA 6000 Nano Kit (5067-1511, Agilent Technologies) and the Qubit RNA BR Assay Kit (Q10210, Invitrogen), respectively.

RNA 5′ pyrophosphohydrolase (RppH) (M0356S, New England BioLabs) treatment was performed with a starting amount of 690 ng RNA. After purification, samples were quantified using the Qubit RNA HS Assay Kit (Q32852, Invitrogen). Next-generation sequencing library preparation was performed using the NEXTFLEX Small RNA-Seq Kit v3 (Bioo Scientific) following step A to step G of the manufacturer's standard protocol (v.16.06). Libraries were prepared with a starting amount ranging between 426 ng and 896 ng and amplified in 16 PCR cycles. Amplified libraries were purified by running an 8% TBE gel and size-selected for 15–50 bp. Libraries were profiled in a High Sensitivity DNA Chip on a 2100 Bioanalyzer Instrument (Agilent Technologies) and quantified using the Qubit dsDNA HS Assay Kit (Q32851, Invitrogen), in a Qubit 2.0 Fluorometer (Invitrogen). All of the samples were pooled at an equimolar ratio and sequenced on one NextSeq 500/550 flow cell, single read for 1×84 cycles plus seven cycles for the index read.

Read processing and mapping. Raw sequenced reads from high-quality libraries, as assessed by FastQC, were processed using Cutadapt[81] for adapter removal (-a TGGAATTCTCGGGTGCCAAGG -O 5 -m 26 -M 48) and low-quality reads were filtered out using the FASTX-Toolkit (fastq_quality_filter, -q 20 -p 100 -Q 33). Unique molecule identifiers were used to remove PCR duplicates using a custom script and were subsequently removed using seqtk (trimfq-l 4 – b 4). Finally, reads shorter than 15 nucleotides were removed using seqtk (seq -L 15).

A custom GTF file was created by adding transposons retrieved from Wormbase (PRJNA13758.WS264) to the Ensembl reference WBcel235.84 and reads were aligned using bowtie (v.1.2.2)[82] (--phred33-quals --tryhard --best --strata --chunkmbs 256 -v 2 -M 1).

Small RNA classification and target identification. All mapped reads were divided into sense and antisense reads using BEDTools intersect[83], and reads of differing lengths and 5′ nucleotides were counted using a custom Python script.

Small RNA classes were then extracted from the mapped reads with the different classes defined as: 21U-RNAs, 21-nucleotide-long mapped reads that map sense to annotated piRNA loci; 22G-RNAs, mapped reads of lengths 20–23 nucleotides with a G at the 5′ position that map antisense to protein-coding genes/ncRNAs/pseudogenes; 26G-RNAs, mapped reads 26 nucleotides long that map antisense to annotated protein-coding genes/ncRNAs/pseudogenes; miRNAs are 20–24 nucleotide reads that map sense to annotated miRNA loci; finally all mapped reads longer than 26 nucleotides were classed in a separate group. ncRNAs were defined to include not only annotated ncRNAs but also RNAs annotated as lincRNAs, snRNAs, snoRNAs, tRNAs and rRNAs. Read filtering was performed using a Python script based on pysam (v.0.15)[84] in combination with BEDTools intersect to extract miRNA and piRNA information[83].

All mapped reads from sequences of 20–23 nucleotides in length, regardless of 5′-nucleotide and mapping direction, were counted using htseq-count (v.0.11.1)[85] (-s no -m intersection-nonempty). Reads per kb of transcript per million mapped reads (RPKM) values were calculated for 22G-RNAs mapping to protein-coding genes/ncRNA/pseduogenes relative to all mapped reads. Targets were defined as genes that were positive in at least two out of three replicates, with positive meaning that the 22G-RNA RPKM was (1) above 4 in the IP; (2) at least twice as high in the IP relative to the input; and (3) non-zero in the input.

For 22G-RNAs mapping to transposons, RPKM values were calculated relative to the predicted number and length of insertions in the custom annotation file and positives were defined using only requirements 2 and 3 above with no minimal RPKM requirement in the IP.

Protein-coding target genes of WAGO-3 were compared to (1) protein-coding target genes of CSR-1 (ref. [22]); (2) protein-coding target genes of siRNAs downregulated in *mut-16*-mutant animals[86]; and (3) protein-coding target genes of sperm-derived 22G-RNAs[37]. To determine germline expression, protein-coding target genes of WAGO-3 were compared to lists of genes expressed in the *C. elegans* germline of either *fem-3*- or *fog-2*-mutant animals[87].

22G-RNA coverage on protein-coding genes. Coverage of 22G-RNA along targeted protein-coding genes was visualized by (1) creating bigwig tracks normalized to mapped non-structural reads (rRNA/tRNA/snoRNA/snRNA)×1 million (RPM) using a combination of BEDTools (genomeCoverageBed -bg -scale -split)[83] followed by bedGraphToBigWig; (2) $\log_2$[IP/input]-normalized tracks were created using deepTools (v.3.4.1)[88] (bigwigCompare --binSize 10 --ratio log2); (3) coverage for each gene was determine with deepTools (computeMatrix scale-regions --metagene --missingDataAsZero -b 250 -a 250 --regionBodyLength 2000 --binSize 50 --averageTypeBins median) with the male samples only being compared to targets found in males and hermaphrodite samples only being compared to targets found in hermaphrodites; and plots were generated with plotProfile (--plotType se --averageType mean --perGroup) to scale and visualize 22G-RNA abundance along targeted genes.

Reads mapping to intronic, exonic or untranslated regions were counted using a custom Python script with reads mapping at exon–intron junctions counted as 0.5 intronic and 0.5 exonic regardless of the spanned region.

Microscopy. For L4 larvae, adults and males, 20–30 animals were washed in a drop of 100 µl M9 buffer and subsequently transferred to a drop of 50 µl M9 buffer supplemented with 40 mM sodium azide on a coverslip. After 15–30 min, excess buffer was removed and a glass slide containing a freshly made agarose pad (2% (w/v) in water) was placed on top of the coverslip. For imaging embryos, adult hermaphrodites were washed and dissected in M9 buffer before mounting. To image sperm, L4 males were singled from hermaphrodites, grown over night, washed and dissected in SMG buffer (50 mM HEPES pH 7.5, 50 mM NaCl, 25 mM KCl, 5 mM $CaCl_2$, 1 mM $MgSO_4$, 10 mM glucose) by cutting near the vas deferens. Animals and sperm were immediately imaged using a TCS SP5 Leica confocal microscope equipped with the HCX PL APO ×63 water-immersion objective (NA, 1.2) or the HCX PL APO CS ×40 oil-immersion objective (NA, 1.3). Fluorescence emission was detected by either photomultiplier tubes or hybrid detectors. Depending on the experiment, SMG buffer was supplemented with 1:2,000 Hoechst 33342 (H1399, Invitrogen), 200 nM MitoTracker Green FM (M7514, Invitrogen) or 1,6-hexanediol (240117, Sigma-Aldrich), and sperm were imaged after 30 min incubation. To score the expression of a germline-specific *mCherry::H2B* transgene, we used a Leica DM6000 B research microscope equipped with a HC PL Fluotar ×20 dry objective (NA, 0.5). Images were processed using ImageJ v.1.53i. The following figures were deconvolved using the Huygens Remote Manager v3.6: Figs. 1c,e,h and 7b and Extended Data Figs. 4a,b,e–j, 5a and 8c.

Time series of spermatocytes expressing GFP::3×Flag::WAGO-3 were acquired using a fluorescence spinning-disk confocal microscope from Visitron Systems (VisiSope 5Elements) based on a Nikon Ti-2E stand and a spinning disk from Yokogawa (CSU-W, 50 µm pinhole) controlled by the VisiView software. The microscope was equipped with a ×60 plan apochromatic water-immersion

objective (CFI Plan Apo VC; NA,1.2), a twofold magnification lens in front of the sCMOS camera (BSI, Photometrics), and a stage-top incubation chamber for live imaging (20 °C, ambient CO_2). The sample was excited by an argon laser at λ_{ex} = 488 nm (200 mW, power set to 20%) and the emission was detected in a range of λ_{em} = 500–550 nm (ET525/50m, Chroma).

Image quantification. Co-localization analyses of confocal micrographs of *C. elegans* gonads were performed using ImageJ v.1.53i and the JACoP plugin. Foci quantification and co-localization analyses of confocal *z* stacks of isolated, male-derived germ cells were performed using IMARIS 9.7.2. The distribution of fluorescence intensities between budding spermatids and residual bodies, as well as quantification of pixel counts per fluorescence intensity (grey value) were determined using ImageJ v.1.53i. The relative pixel count is the number of pixels with a given fluorescence intensity within a selected region of interest (either spermatid or residual body), divided by the total number of pixels in that region. Every other data distribution is represented as a box plot, with the whiskers defining the median ± 1.5 × interquartile range, a rectangle marking the first and third quartile, and the centre line showing the median.

FRAP. FRAP measurements were performed on a TCS SP5 Leica confocal microscope equipped with a FRAP-booster and a HCX PL APO ×63 water-immersion objective (NA, 1.2). An entire granule was bleached in a fixed region of interest (diameter, 0.9 µm), while two additional control ROIs of same size were used to detect fluorescence emission of an unbleached granule and background signal, respectively. Five prebleach frames were recorded (5 × 0.374 s per frame), followed by two bleach frames (2 × 0.374 s per frame), and 3 sets of post-bleach frames (10 × 0.5 s per frame, 10 × 5 s per frame, 15 × 10 s per frame). Data analysis, including full-scale normalization and curve fitting using a double term exponential equation, was performed using EasyFRAP-web[89].

MO counting. The LysoSensor Blue DND-192 stained MOs[90] within living spermatids were viewed using an Olympus BX60 with a ×100/1.35 NA oil-immersion objective lens. Epi-fluorescence of stained MOs was imaged using a DAPI filter pack and captured using a SensiCam digital camera (Cooke) controlled by SlideBook software (Intelligent Imaging Innovations). SlideBook software was used to collect *z*-axis stacks of 11–30, and 12-bit images were captured every 0.44–0.88 µm. The majority of images within the *z* stacks was collected approximately every 0.65 µm. A nearest-neighbour deconvolution algorithm within the SlideBook software was applied to the images. Images were then converted to *z*-axis projections, again using the SlideBook software. The diameters of individual spermatids were measured within SlideBook using the software's ruler function over images captured with a DIC filter. The manipulated images were exported from SlideBook as 16-bit tif images. The 16-bit tif images exported from SlideBook were reopened in ImageJ (NIH). The cell counter plugin within ImageJ was used to assist in counting MOs. Images were compiled using PhotoShop CS3 (Adobe).

BmN4 cell culture and transfection. BmN4 cells were cultured at 27 °C in IPL-41 (Gibco) medium supplemented with 10% FBS (Gibco) and 1% penicillin–streptomycin. BmN4 cells (6.0×10^5) were seeded in a single well of a 6-well plate and, the next day, were transfected with 2 µg of plasmid DNA using X-tremeGene HP (Roche). 2-BP was added 3 h after transfection and after 24 h cells were collected and lysed in IP-150 lysis buffer (30 mM HEPES (pH 7.4), 150 mM potassium acetate, 5 mM magnesium acetate, 5 mM DTT, 0.1% tergitol-type NP-40, 2 mg ml^{-1} pepstatin, 2 mg ml^{-1} leupeptin and 0.5% aprotinin).

For the IP experiments, 4.0×10^6 BmN4 cells were seeded in a 10 cm dish and transfected with 10 µg of plasmid DNA. Cells were lysed in IP-150 lysis buffer (supplemented with 0.5% Triton X-100) and IP experiments were performed using ChromoTek RFP-TRAP beads. After IP, beads were washed five times in IP-150 lysis buffer (supplemented with 0.5% Triton X-100).

CLEM analysis. For CLEM analyses, *C. elegans* males (*wago-3(xf119) I; pei-1(xf193) IV; him-5(e1490) V*) were selected using a stereomicroscope. Using a platinum wire, around 50 worms were transferred into the 100-µm-deep cavity of an A-type carrier (Wohlwend, imbibed with 1-hexadecene) filled with thick *E. coli* paste (serving as a cryo-protectant; thick OP50 *E.coli* suspension in M9 worm buffer, 20% BSA, 150 mM NaCl), closed with the flat side of a B-type carrier and subsequently high-pressure frozen (HPM010, AbraFluid). All samples were further processed by freeze-substitution in a temperature-controlling device (EM-AFS2, Leica Microsystems). Freeze-substitution was carried out at −90 °C for 72 h with 0.1% (w/v) uranyl acetate in glass distilled acetone (EMS). The temperature was then raised to −45 °C (3.5 °C h^{-1}), and samples were further incubated for 5 h. After rinsing in acetone, the samples were infiltrated with increasing concentrations (10%, 25%, 50% and 75%; 6 h each) of Lowicryl HM20 resin (EMS) in acetone, while the temperature was further raised to −25 °C. Lowicryl (100%) was exchanged three times in 10 h steps, and samples were ultraviolet-light polymerized at −25 °C for 48 h, after which the temperature was raised to 20 °C (5 °C h^{-1}), and ultraviolet-light polymerization continued for 6 h. Longitudinal sections (thickness, 300 nm) were cut using an ultra-microtome (UC7, Leica) and a diamond knife (ultra semi, DiATOME). Targeting of areas containing spermatocytes and spermatids was performed using toluidine blue staining. Sections of interest were picked up onto carbon-coated formvar-slot grids. Ultramicrotomy and acquisition of the in-resin retained fluorescence within the sections were best performed on the same day to avoid bleaching of the fluorescence. The fluorescence microscopy imaging of the sections (stained also with HOECHST) was performed as previously described[91] using a wide-field fluorescence microscope (Olympus IX81) equipped with an Olympus PlanApo ×100/1.40 NA oil-immersion objective. After post-staining, tilt series of the area of interest (1° increments, −60° to 60°) were acquired using a FEI TECNAI F30 TEM operated at 300 kV and a fast Gatan OneView 4K camera. Tomograms were reconstructed at a final voxel size of 1.56 nm using gold fiducials and weighted-back projection algorithms of the software package IMOD[92]. Correlation between light and electron micrographs was carried out using the plugin ec-CLEM[93] of the software platform Icy[94]. The coordinates of pairs of corresponding features in the two imaging modalities (50 nm Tetraspecks, HOECHST-stained condensed chromosomes, auto-fluorescent uranyl-acetate-stained *E. coli*) were used to calculate a linear transformation, which enabled mapping of the coordinates of the fluorescent spots of interest to overlay them onto the electron micrograph. Electron tomograms were displayed and analysed using the IMOD software package[92].

Statistics and reproducibility. Statistical analyses were performed using R-based packages. The log-rank test was used for the mortal germline assay. For multiple group comparison, either one-way ANOVA followed by Tukey honest significant difference post hoc test, Pearson's χ^2 test with Yates continuity correction or Fisher's exact test was used. $P < 0.05$ was considered to be significant. Pearson's correlation analyses were performed to determine the relationship between two different factors. No statistical method was used to predetermine sample size. No data were excluded from the analyses. Fluorescence microscopy, CLEM, time-lapse microscopy, 1,6-hexanediol assays as well as *C. elegans* experiments comprising *mCherry::H2B*(RNAe) reactivation and *Mutator*-induced sterility were performed twice. Mortal germline assay (Extended Data Fig. 1a) and western blots (Fig. 8a,b and Extended Data Fig. 10a,b,h,i) were performed once unless specified in the legends. Immunoprecipitation experiments associated with MS and small RNA sequencing were performed in biological quadruplicates and triplicates, respectively. Quantification and statistical analyses of microscopy images were derived from the number of cells analysed across two independent experiments as indicated in the figure legends, with the presented data being derived from one representative independent experiment.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The accession number for the small-RNA-seq data generated in this study is PRJNA629991. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE[95] partner repository under dataset identifier PXD019099. The BigDataViewer supporting the CLEM analyses is available at Mendeley Data (V1; https://doi.org/10.17632/dgb8d7h2hz.1; https://data.mendeley.com/datasets/dgb8d7h2hz/1). All other data supporting the findings of this study are available from the corresponding author on reasonable request. Source data are provided with this paper.

Code availability

All code developed for this analysis is available via https://github.com/Tunphie/SequencingTools/blob/main/smRNA_TypeCounter.py and https://github.com/Tunphie/SequencingTools/blob/main/CoverageOnProteinCodingGenes.py.

References

61. Brenner, S. The genetics of *Caenorhabditis elegans*. *Genetics* **77**, 71–94 (1974).
62. Schweinsberg, P. J. & Grant, B. D. *C. elegans* gene transformation by microparticle bombardment. *WormBook* https://doi.org/10.1895/wormbook.1.166.1 (2013).
63. Haeussler, M. et al. Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. *Genome Biol.* **17**, 148 (2016).
64. Chen, B. et al. Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell* **155**, 1479–1491 (2013).
65. Chiu, J., March, P. E., Lee, R. & Tillett, D. Site-directed, ligase-independent mutagenesis (SLIM): a single-tube methodology approaching 100% efficiency in 4h. *Nucleic Acids Res.* **32**, e174 (2004).
66. Chiu, J., Tillett, D., Dawes, I. W. & March, P. E. Site-directed, ligase-independent mutagenesis (SLIM) for highly efficient mutagenesis of plasmids greater than 8kb. *J. Microbiol. Methods* **73**, 195–198 (2008).
67. Dickinson, D. J., Ward, J. D., Reiner, D. J. & Goldstein, B. Engineering the *Caenorhabditis elegans* genome using Cas9-triggered homologous recombination. *Nat. Methods* **10**, 1028–1034 (2013).

68. Dickinson, D. J., Pani, A. M., Heppert, J. K., Higgins, C. D. & Goldstein, B. Streamlined genome engineering with a self-excising drug selection cassette. *Genetics* **200**, 1035–1049 (2015).
69. Ward, J. D. Rapid and precise engineering of the caenorhabditis elegans genome with lethal mutation co-conversion and inactivation of NHEJ repair. *Genetics* **199**, 363–377 (2014).
70. Frøkjær-Jensen, C. et al. Single-copy insertion of transgenes in *Caenorhabditis elegans*. *Nat. Genet.* **40**, 1375–1383 (2008).
71. Paix, A. et al. Scalable and versatile genome editing using linear DNAs with microhomology to Cas9 sites in *Caenorhabditis elegans*. *Genetics* **198**, 1347–1356 (2014).
72. Paix, A., Schmidt, H. & Seydoux, G. Cas9-assisted recombineering in *C. elegans*: genome editing using in vivo assembly of linear DNAs. *Nucleic Acids Res.* **44**, e128 (2016).
73. Arribere, J. A. et al. Efficient marker-free recovery of custom genetic modifications with CRISPR/Cas9 in *Caenorhabditis elegans*. *Genetics* **198**, 837–846 (2014).
74. El Mouridi, S. et al. Reliable CRISPR/Cas9 genome engineering in *Caenorhabditis elegans* using a single efficient sgRNA and an easily recognizable phenotype. *G3* **7**, 1429–1437 (2017).
75. Klass, M. R. & Hirsh, D. Sperm isolation and biochemical analysis of the major sperm protein from *Caenorhabditis elegans*. *Dev. Biol.* **84**, 299–312 (1981).
76. Shevchenko, A., Tomas, H., Havliš, J., Olsen, J. V. & Mann, M. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat. Protoc.* **1**, 2856–2860 (2007).
77. Kappei, D. et al. HOT1 is a mammalian direct telomere repeat-binding protein contributing to telomerase recruitment. *EMBO J.* **32**, 1681–1701 (2013).
78. Rappsilber, J., Mann, M. & Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* **2**, 1896–1906 (2007).
79. Bluhm, A., Casas-Vila, N., Scheibe, M. & Butter, F. Reader interactome of epigenetic histone marks in birds. *Proteomics* **16**, 427–436 (2016).
80. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367–1372 (2008).
81. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* **17**, 10–12 (2011).
82. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
83. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
84. Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
85. Anders, S., Pyl, P. T. & Huber, W. HTSeq-A Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).
86. Phillips, C. M. et al. MUT-14 and SMUT-1 DEAD box RNA helicases have overlapping roles in germline RNAi and endogenous siRNA formation. *Curr. Biol.* **24**, 839–844 (2014).
87. Ortiz, M. A., Noble, D., Sorokin, E. P. & Kimble, J. A new dataset of spermatogenic vs. oogenic transcriptomes in the nematode *Caenorhabditis elegans*. *G3* **4**, 1765–1772 (2014).
88. Ramírez, F., Dündar, F., Diehl, S., Grüning, B. A. & Manke, T. DeepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res.* **42**, 187–191 (2014).
89. Koulouras, G. et al. EasyFRAP-web: a web-based tool for the analysis of fluorescence recovery after photobleaching data. *Nucleic Acids Res.* **46**, W467–W472 (2018).
90. Gleason, E. J. et al. Developmental genetics of secretory vesicle acidification during *Caenorhabditis elegans* spermatogenesis. *Genetics* **191**, 477–491 (2012).
91. Kukulski, W. et al. Correlated fluorescence and 3D electron microscopy with high sensitivity and spatial precision. *J. Cell Biol.* **192**, 111–119 (2011).
92. Kremer, J. R., Mastronarde, D. N. & McIntosh, J. R. Computer visualization of three-dimensional image data using IMOD. *J. Struct. Biol.* **116**, 71–76 (1996).
93. Paul-Gilloteaux, P. et al. eC-CLEM: flexible multidimensional registration software for correlative microscopies. *Nat. Methods* **14**, 102–103 (2017).
94. de Chaumont, F. et al. Icy: an open bioimage informatics platform for extended reproducible research. *Nat. Methods* **9**, 690–696 (2012).
95. Perez-Riverol, Y. et al. The PRIDE database and related tools and resources in 2019: Improving support for quantification data. *Nucleic Acids Res.* **47**, D442–D450 (2019).

Acknowledgements

We thank the members of the Ketting laboratory for discussions; H. Grosshans for reading the manuscript; M. Dörr and S. Hellmann for technical support; C. Werner and A. Dold of the IMB Genomics Core Facility for small RNA library preparation; the staff at the IMB Media Laboratory, Microscopy, Proteomic and Genomic Core Facilities for consumables, equipment and experimental support; and S. Köhler and M. Schorb for their help in CLEM sample preparation and visualization. Some of the strains were provided by the *Caenorhabditis* Genetics Center (CGC), funded by NIH Office of Research Infrastructure Programs (P40 OD010440). We acknowledge the GenEvo RTG funded by the Deutsche Forschungsgemeinschaft (DFG; 407023052/GRK2526/1), which enabled the conception of this project. E.J.G., S.P. and S.W.L. were supported for this work by Emory College of Arts and Sciences. This work was supported by grants of the DFG KE 1888/1-1, KE1888/1-2 and KE 1888/6-1 (to R.F.K.) and the National Institute of Health R35 GM119656 (to C.M.P.), and T32 GM118289 (to D.A.H.N.).

Author contributions

J.S. and R.F.K. conceived the study and designed experiments. J.S. executed experiments and performed data analysis. S.D. and F.B. performed MS analysis. A.M.d.J.D. and A.-S.S. performed small-RNA-seq analysis. M.B. and V.O. performed the CLEM experiments. A.W.B. performed PEI-1/2 studies in cell culture. D.A.H.N. and C.M.P. provided unpublished strains. E.J.G., S.P. and S.W.L. shared unpublished data on MO counts. R.F.K. supervised the project. J.S. and R.F.K. wrote the manuscript with input from all authors.

Competing interests

The authors declare no competing interests.

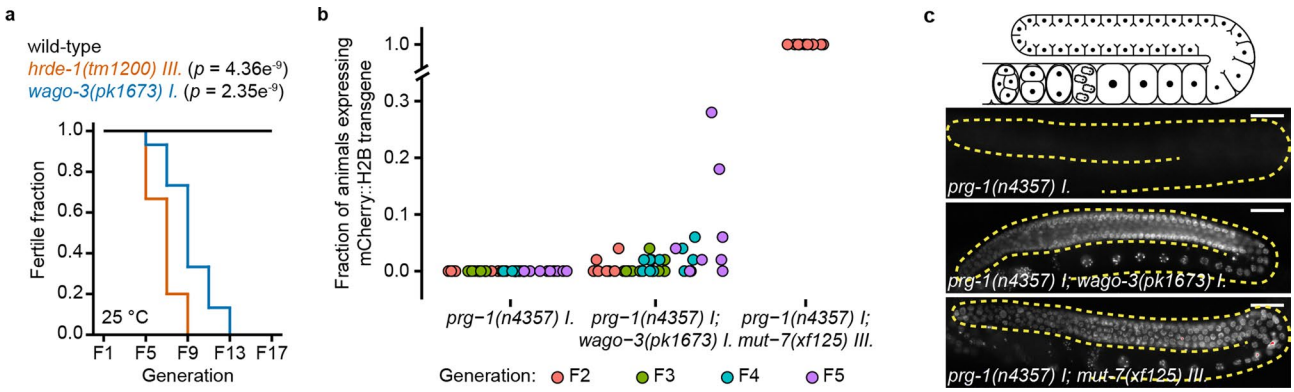
Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41556-021-00827-2.

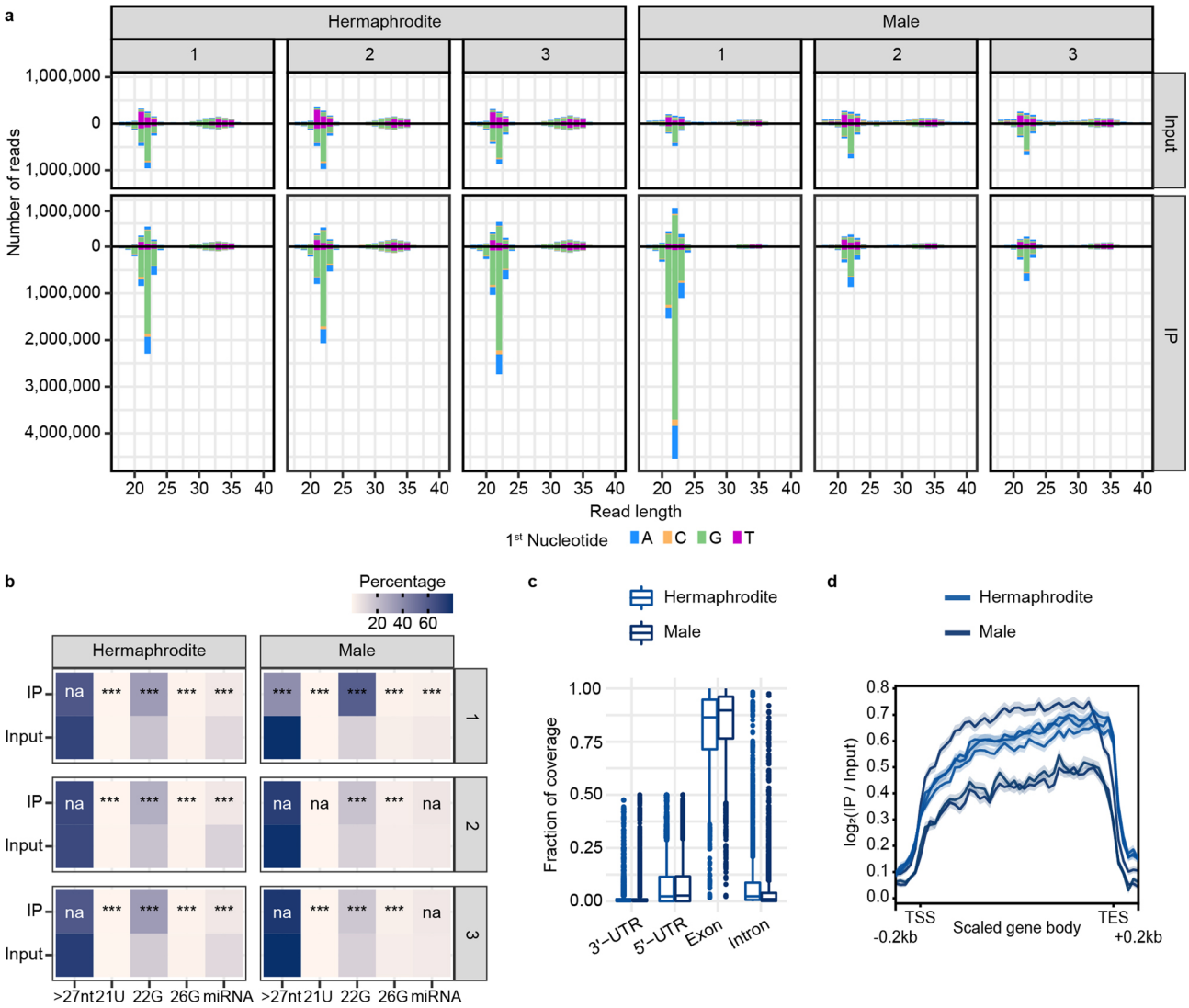
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41556-021-00827-2.

Correspondence and requests for materials should be addressed to René F. Ketting.

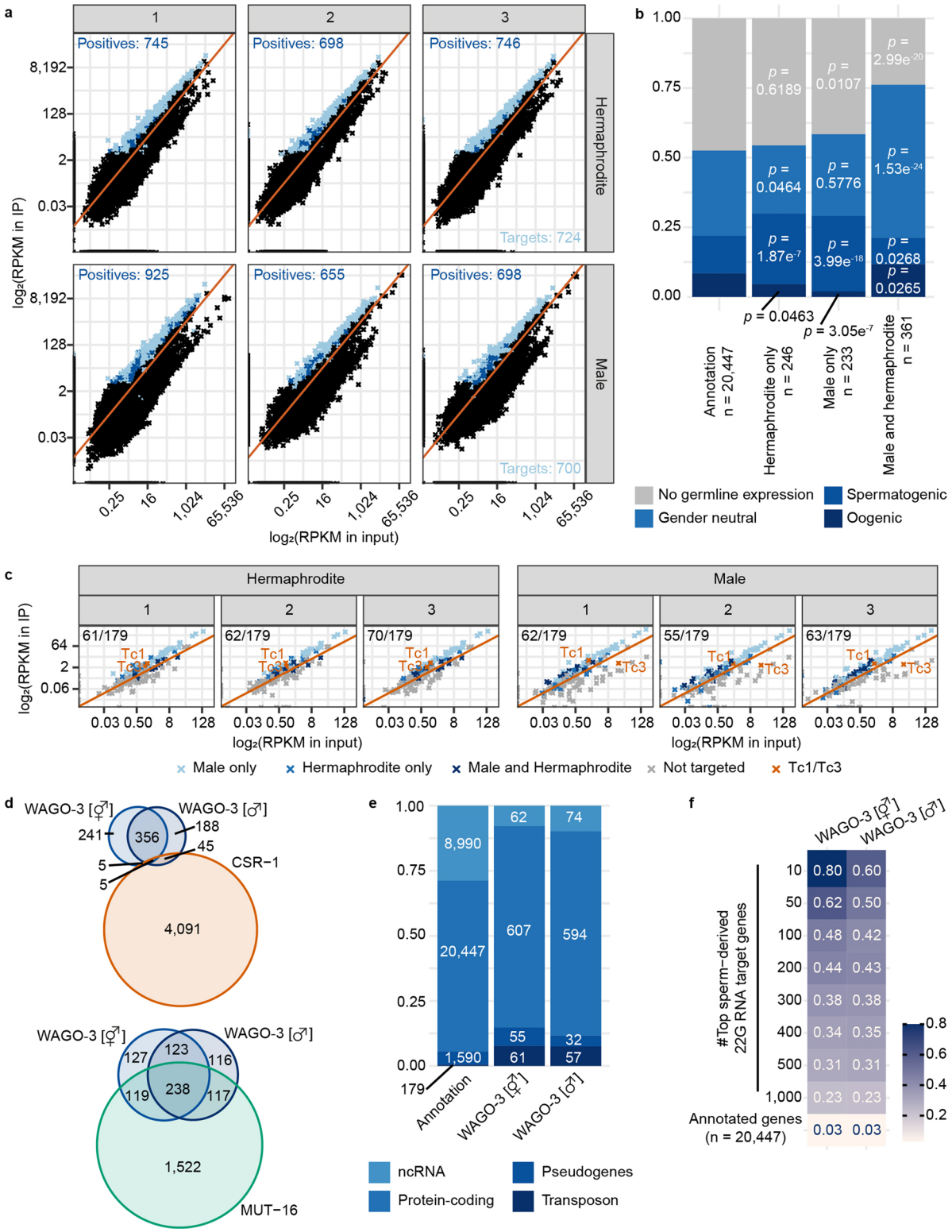
Peer review information *Nature Cell Biology* thanks Ben Lehner, Oded Rechavi and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.



Extended Data Fig. 1 | WAGO-3 is required for germline immortality and transgenerational maintenance of RNAe. a, Mortal germline assay representing loss of fertility of strains with indicated genotype at 25 °C. Statistical significance was tested with a log-rank-test (n = 15 populations per strain assayed in a single experiment). **b**, Diagram displaying mCherry::H2B(RNAe) reactivation in *prg-1(n4357), prg-1(n4357);mut-7(xf125)* and *prg-1(n4357);wago-3(pk1673)* mutant generations. F2-F5: second-fifth homozygous generation. For each generation, reactivation in 10 populations of 50 animals each was scored. Each plotted point represents the fraction of 50 animals that express the mCherry::H2B transgene. Since no *prg-1(n4357)* single mutant animal was found to reactivate mCherry::H2B expression, the value of this group is deterministically zero due to lack of variability/statistical noise. Thus, any positive number of animals that expresses the mCherry::H2B transgene in either the *prg-1(n4357);mut-7(xf125)* or *prg-1(n4357);wago-3(pk1673)* group causes a significant difference from the *prg-1(n4357)* group. **c**, Micrographs of three example animals with the mCherry::H2B transgene in either RNAe (*prg-1(n4357)*) or activated (*prg-1(n4357);wago-3(pk1673)* and *prg-1(n4357);mut-7(xf125)*) status. Top panel shows schematic representation of an adult hermaphroditic gonad. Activity status of the transgene was homogeneous in F2 homozygous *prg-1(n4357)* and *prg-1(n4357);mut-7(xf125)* mutants. Images represent two biologically independent experiments. Scale bars: 30 µm. Source data are provided.

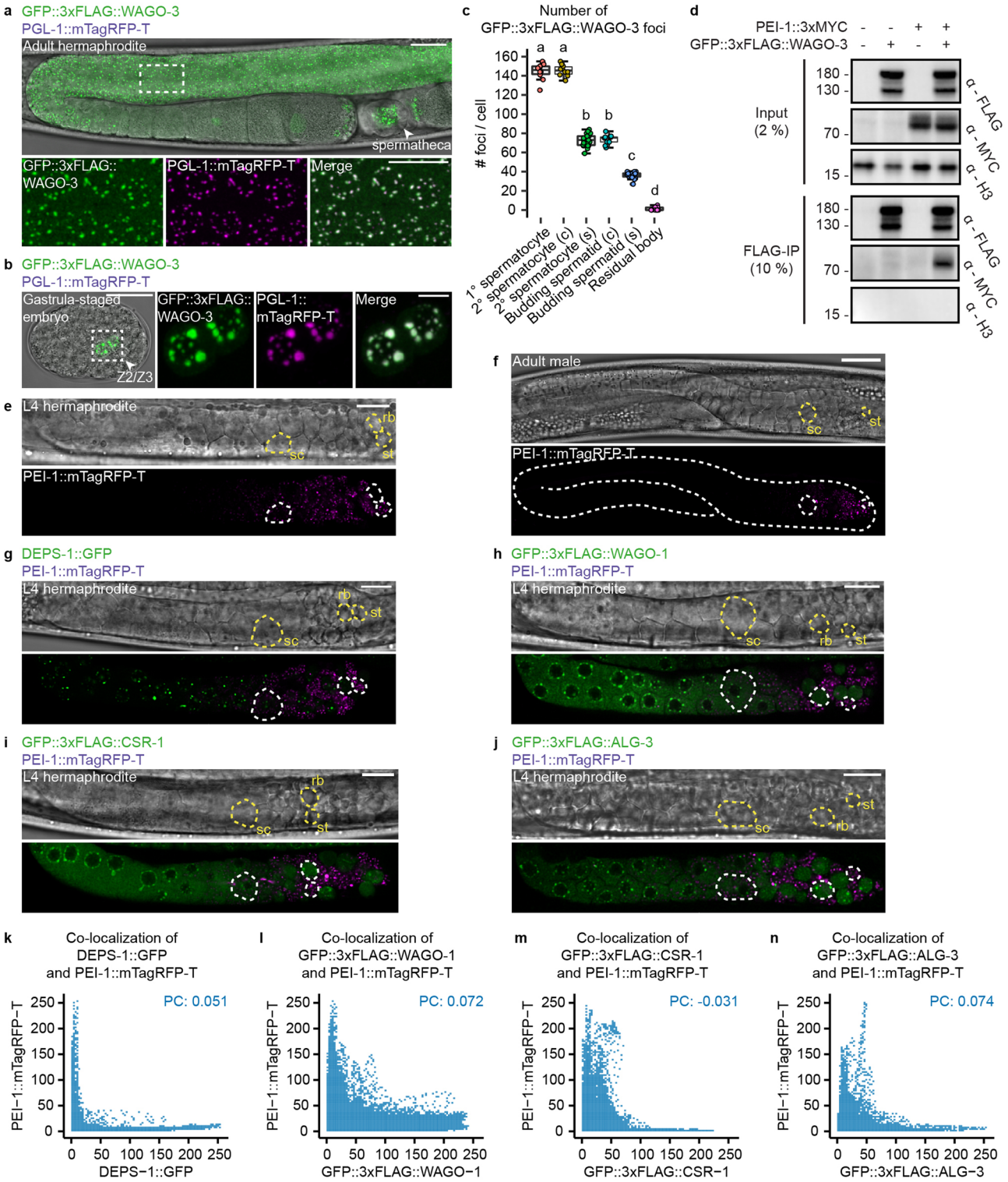


Extended Data Fig. 2 | WAGO-3 is associated with 22G RNAs. a, Read length distribution and first nucleotide bias of indicated small RNA libraries. Both hermaphrodite and male libraries were prepared and sequenced in biological triplicates. Each panel represents a replicate, top and bottom bars represent sense and anti-sense small RNAs mapping to annotated loci. **b**, Heat map showing the enrichment or depletion of small RNA species of the libraries shown in **a**. >27 nt are non-specific RNA fragments, also including fragments of structural RNAs, such as rRNA and tRNA. 21U, 22G, 26 G and miRNA are known Argonaute-associated small RNA species. Statistically significant differences between small RNA types were determined by Fisher's exact tests (***: $p \leq 0.001$, ns: $p > 0.05$). The exact *P* values are provided as source data. **c**, Box plots displaying the distribution of WAGO-3 associated 22G RNAs mapping to protein-coding genes across the indicated gene segments. Each box plot represents average data from three biological replicates. Boxplot centre and box edges indicate median and 25th or 75th percentiles, respectively, while whiskers indicate the median ± 1.5 x interquartile range. **d**, Metagene analysis of 22G RNA reads mapping to protein-coding WAGO-3 target genes. TSS – transcription start site, TES – transcription end site. Shading around each line represents the standard error of the median of each bin. Source data are provided.



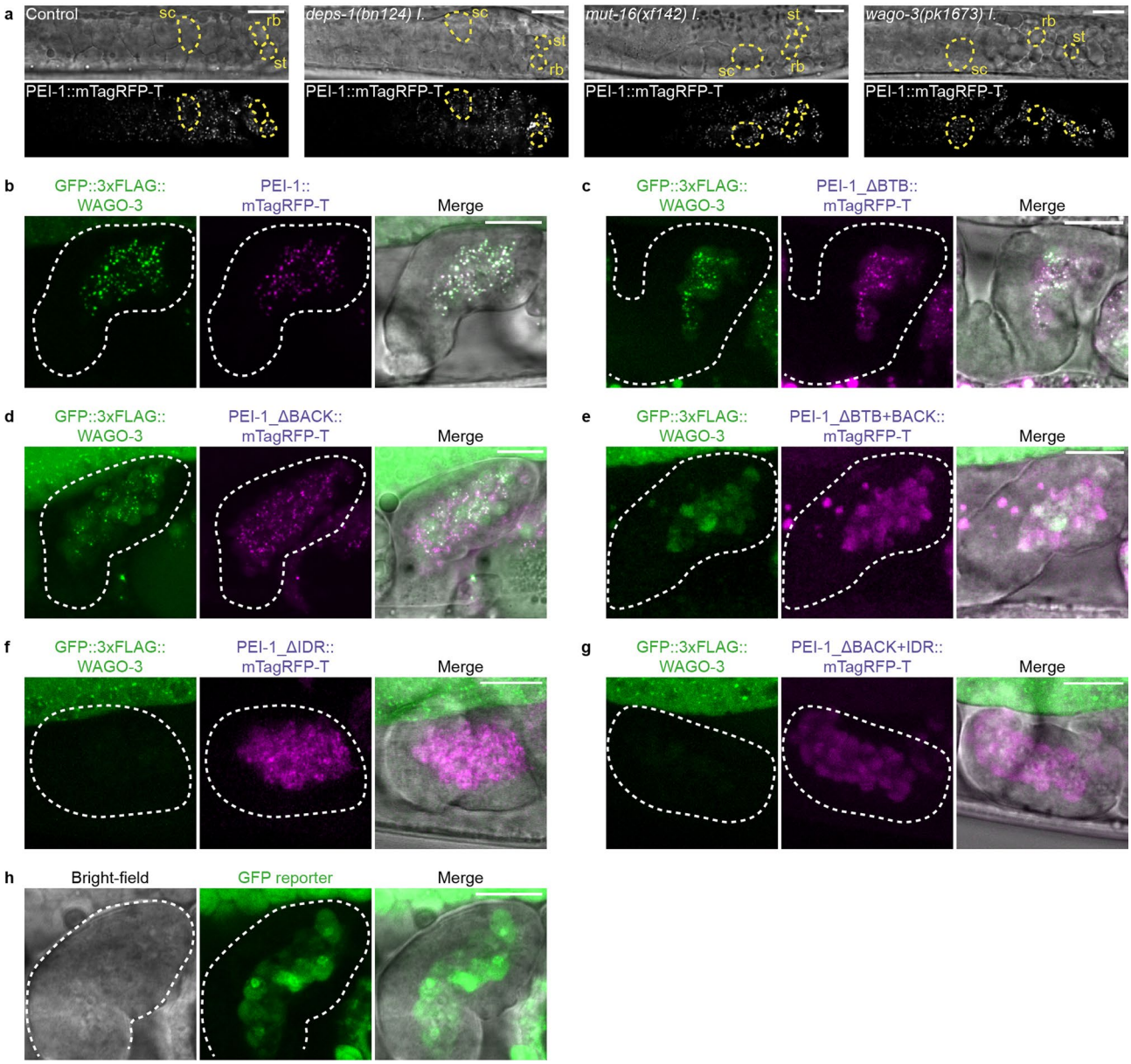
Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | WAGO-3 targets overlap with *Mutator* targets and sperm-derived 22 G RNAs. a, Scatter plots displaying the RPKM of 22 G RNAs mapping to individual genes in immunoprecipitation (IP) (*Y*-axis) versus input (*X*-axis) samples for all six sequenced libraries. Transposons are not included in these plots. Light blue: significant enrichment of genes in at least two replicates. Dark blue: significant enrichment of genes in only one replicate. **b**, Germline expression status of protein-coding WAGO-3 target genes. Left column shows the distribution of expression profiles of all annotated protein-coding genes. Columns 2 and 3 show the same, but for hermaphrodite and male-specific WAGO-3 targets, respectively. Column 4 shows the same for targets found in both sexes. Statistically significant differences with respect to complete gene annotations were determined by Chi-square tests. **c**, Scatter plots displaying the RPKM of 22G RNAs mapping to transposons in the six individual experiments in input (*X*-axes) and IP samples (*Y*-axes). Red line represents the diagonal. **d**, Venn diagrams displaying the overlaps of WAGO-3 targets (protein-coding) called in hermaphrodites and males with previously determined CSR-1 and MUT-16 targets. **e**, Stacked bar plot showing number and types of WAGO-3 targets. **f**, Heat map showing the overlap of WAGO-3 targets (protein-coding) with previously determined sperm-derived 22G RNA targets (protein-coding), which were binned into indicated groups. Values inside the boxes indicate fraction of overlap. Source data are provided.

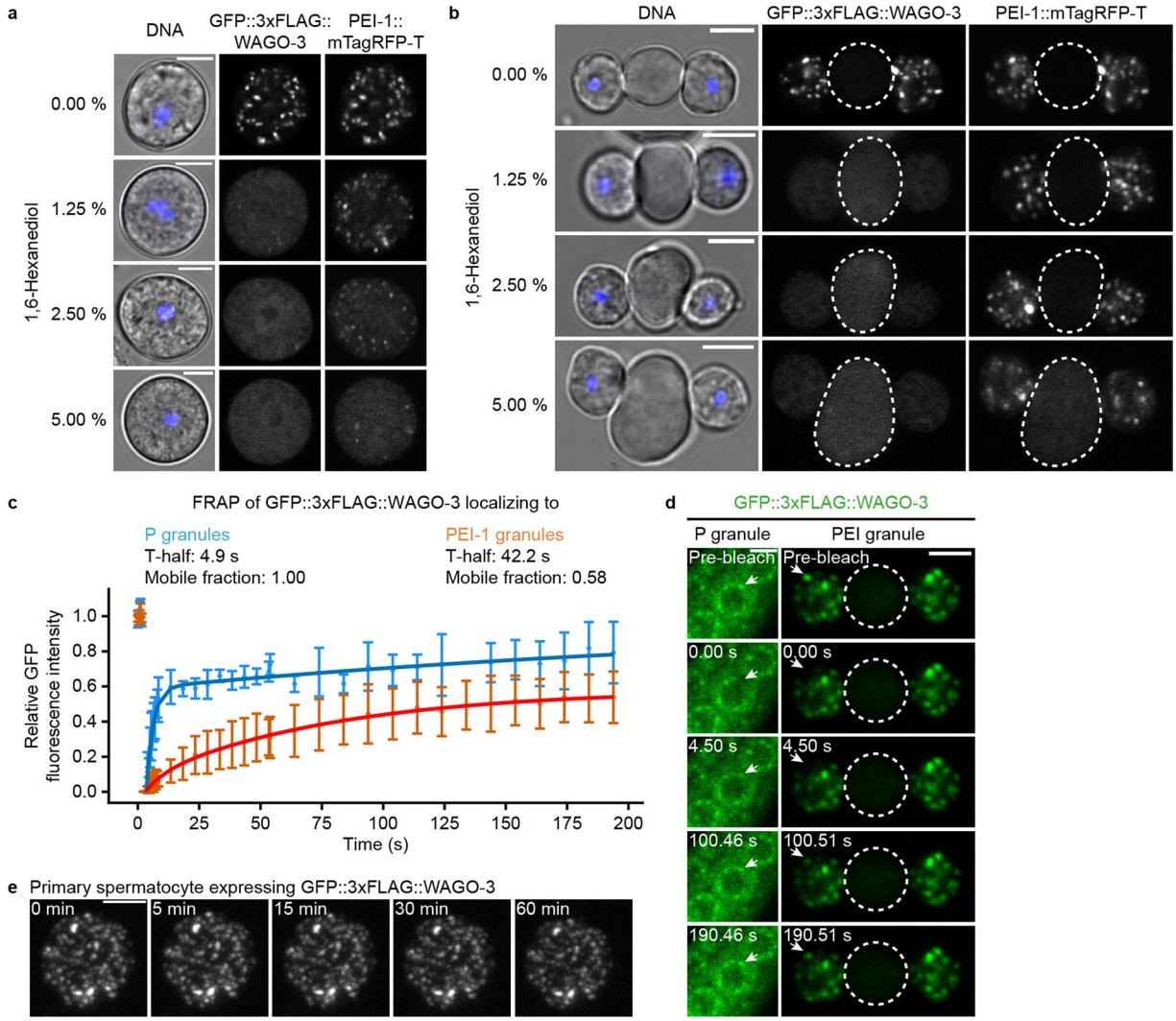


Extended Data Fig. 4 | See next page for caption.

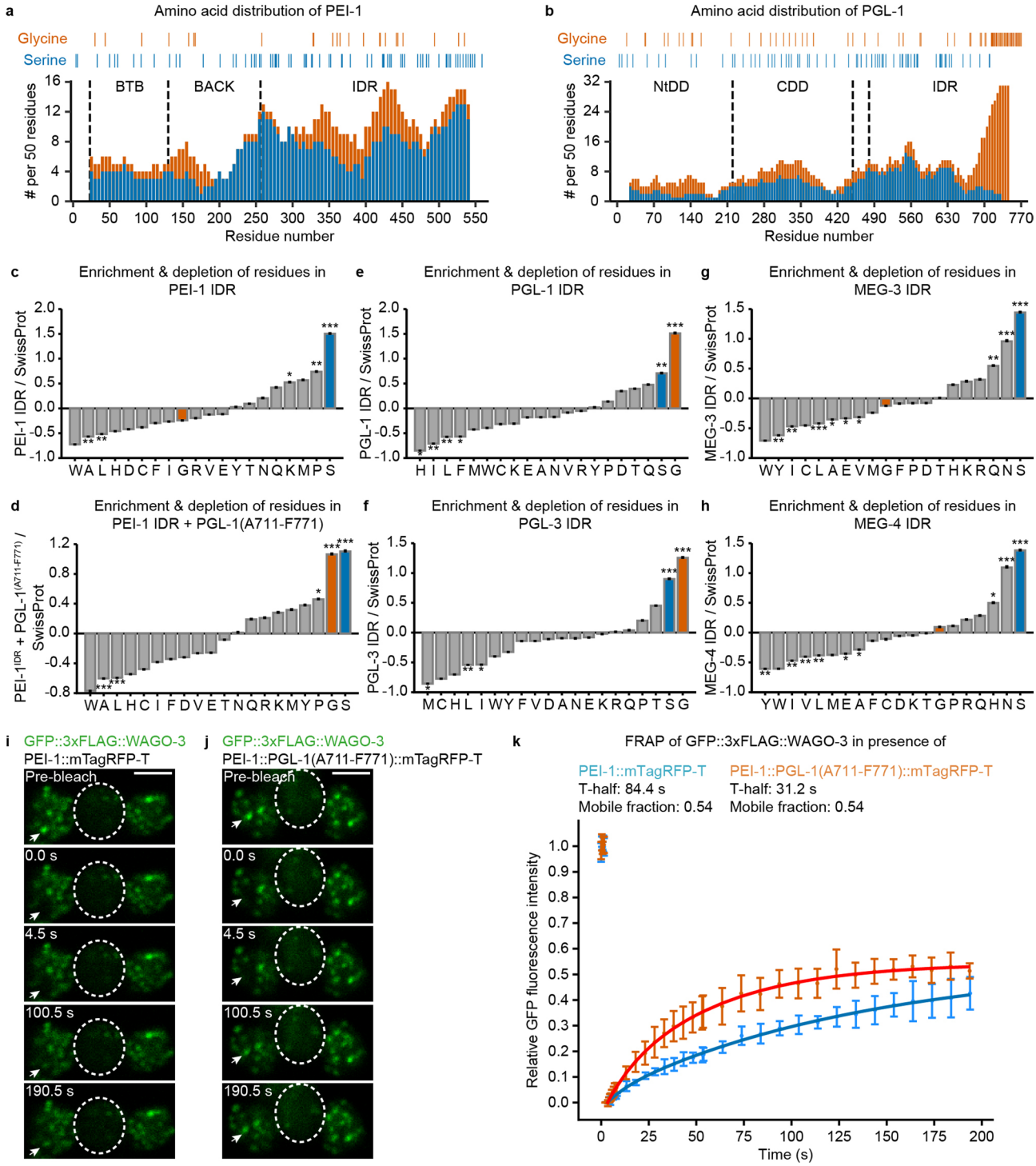
Extended Data Fig. 4 | PEI granules specifically recruit WAGO-3. a,b, Confocal maximum intensity projections of an adult hermaphrodite (**a**) and gastrula-staged embryo (**b**) expressing indicated proteins. Zooms show perinuclear co-localization of GFP::3xFLAG::WAGO-3 and PGL-1::mTagRFP-T in meiotic (**a**) and primordial (**b**) germ cells. Z2/Z3 are the primordial germ cells of *C. elegans*. Scale bars: 20 µm (**a**, adult), 20 µm (**b**, embryo), 10 µm (**a**, zoom), 4 µm (**b**, zoom). **c**, Quantification of GFP::3xFLAG::WAGO-3 foci number within indicated, male-derived germ cells (n = 10 cells pooled from two independent experiments, for each condition). Statistically significant differences were determined by one-way ANOVA ($p \leq 0.001$) followed by Tukey's honestly significant difference post hoc test ($p \leq 0.05$). Different letters represent significant differences. Note that the values for primary spermatocytes and budding spermatids (c) are the same as those shown in Fig. 3j (FL) and Extended Data Fig. 10f (wild-type), and Fig. 4n (FL) and Fig. 7j (wild-type), respectively. Secondary spermatocyte and budding spermatid stages are separated into 'c' and 's'. c: coupled; due to incomplete cytokinesis, leaving the two sister cells coupled and both cells were analyzed; s: separate, each of the coupled cells in 'c' was analyzed individually. Boxplot centre and box edges indicate median and 25th or 75th percentiles, respectively, while whiskers indicate the median ± 1.5 x interquartile range. **d**, Co-immunoprecipitation experiments using whole-worm extracts of late-L4 stage hermaphrodites analyzed by Western blotting. Sample processing control was run on a different gel. Data represent two biologically independent experiments. **e-j**, Confocal micrographs of an adult male (**f**) and late-L4 stage hermaphrodites (**e**,**g**-**j**) expressing indicated proteins. sc – spermatocyte, st – spermatid, rb – residual body. Scale bars: 10 µm (**e**-**j**). **a**,**b**,**e**-**j**, Images represent two biologically independent experiments. **k-n**, Co-localization analyses between PEI-1::mTagRFP-T and DEPS-1::GFP (**k**), GFP::3xFLAG::WAGO-1 (**l**), GFP::3xFLAG::CSR-1 (**m**) and GFP::3xFLAG::ALG-3 (**n**) based on the images shown in **g**-**j**, respectively (n = 10 worms for each condition). *X* and *Y* axes indicate fluorescence intensity. PC: Pearson's correlation coefficient. Exact *P* values (**c**), unprocessed original scans of blots and the source data for all graphical representations are provided.



Extended Data Fig. 5 | Presence of WAGO-3 in spermatozoa is dependent on the IDR of PEI-1. a, Confocal micrographs showing spermatogenesis of late-L4 stage hermaphrodites expressing PEI 1::mTagRFP-T in indicated mutants. sc – spermatocyte, rb – residual body, st – spermatid. Images represent two biologically independent experiments. Scale bars: 10 μm. **b**-**h**, Confocal maximum intensity projections of hermaphrodite-derived spermatozoa within the spermatheca expressing indicated proteins. In all panels, except **c**, a piece of a gonad arm expressing GFP::3xFLAG::WAGO-3 is visible in the top part of the image. Dashed lines indicate spermatheca. Images represent two biologically independent experiments. Scale bars: 10 μm.

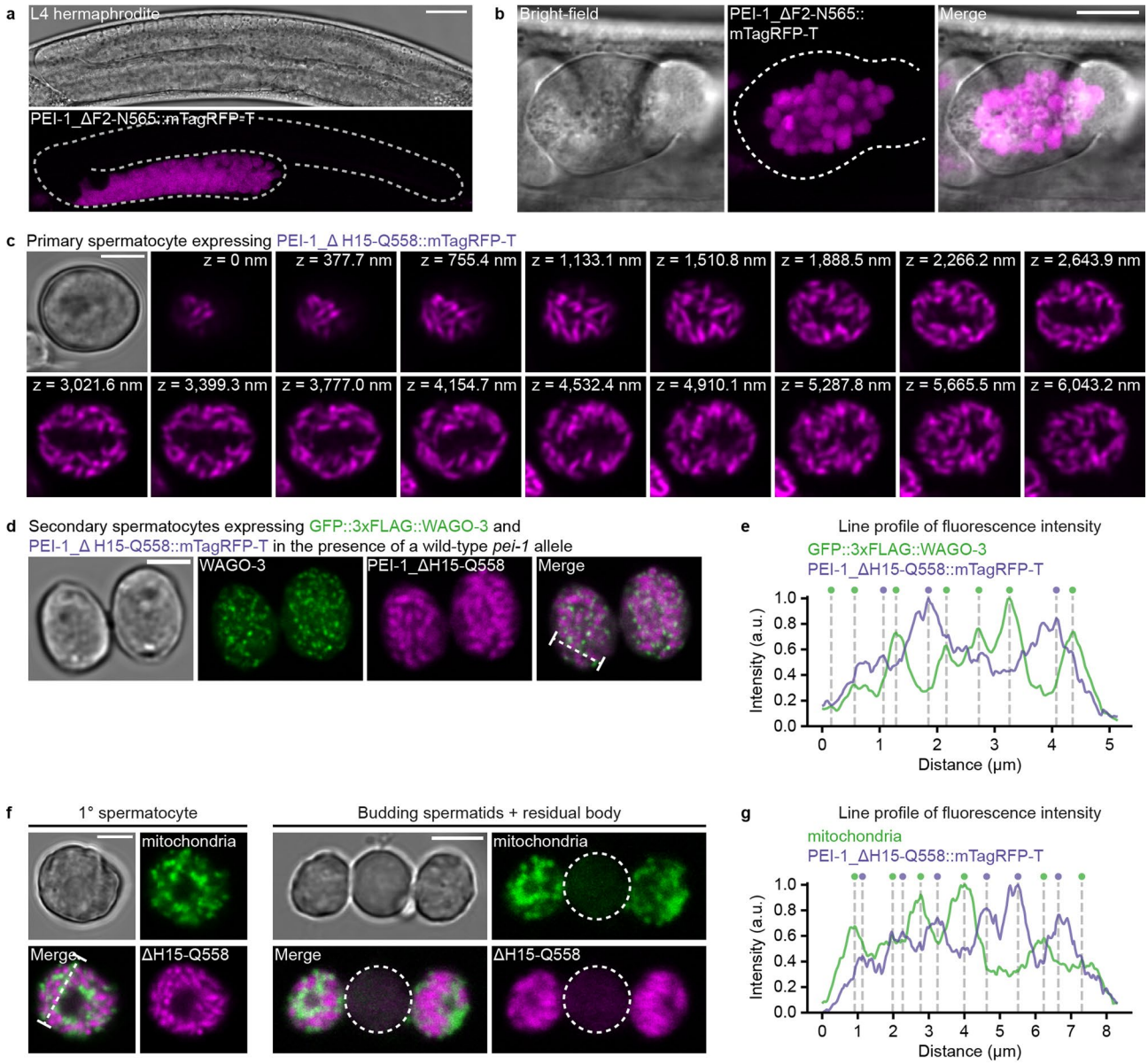


Extended Data Fig. 6 | PEI granules are static condensates with liquid-like properties. a,b, Confocal micrographs of isolated, male-derived spermatocytes (**a**) and budding spermatids (**b**) expressing GFP::3xFLAG::WAGO-3 and PEI-1::mTagRFP-T. Images were taken after a 30 minute treatment with 1,6-hexanediol. Hoechst33342 was used to stain DNA. Residual bodies are marked by a dashed circle. Images represent two biologically independent experiments. Scale bars: 4 µm. **c**, FRAP recovery curve of GFP::3xFLAG::WAGO-3 localizing to either P granules in L4 gonads or PEI granules in male-derived spermatids. Normalized data is presented as mean+/− SD and was fitted to a double exponential curve (n = 4 granules pooled from one independent experiment). **d**, Time sequence showing fluorescence recovery after photobleaching (FRAP) of GFP::3xFLAG::WAGO-3 localizing to either P granules in L4 gonads or PEI granules in male-derived spermatids. Residual bodies are marked by a dashed circle. Images represent two biologically independent experiments. Scale bars: 4 µm. **e**, Time sequence of GFP::3xFLAG::WAGO-3, taken from Extended Data Movie 1. Images are confocal maximum intensity projections of an isolated, male-derived spermatocyte. Images represent two biologically independent experiments. Scale bar: 4 µm. Source data are provided.

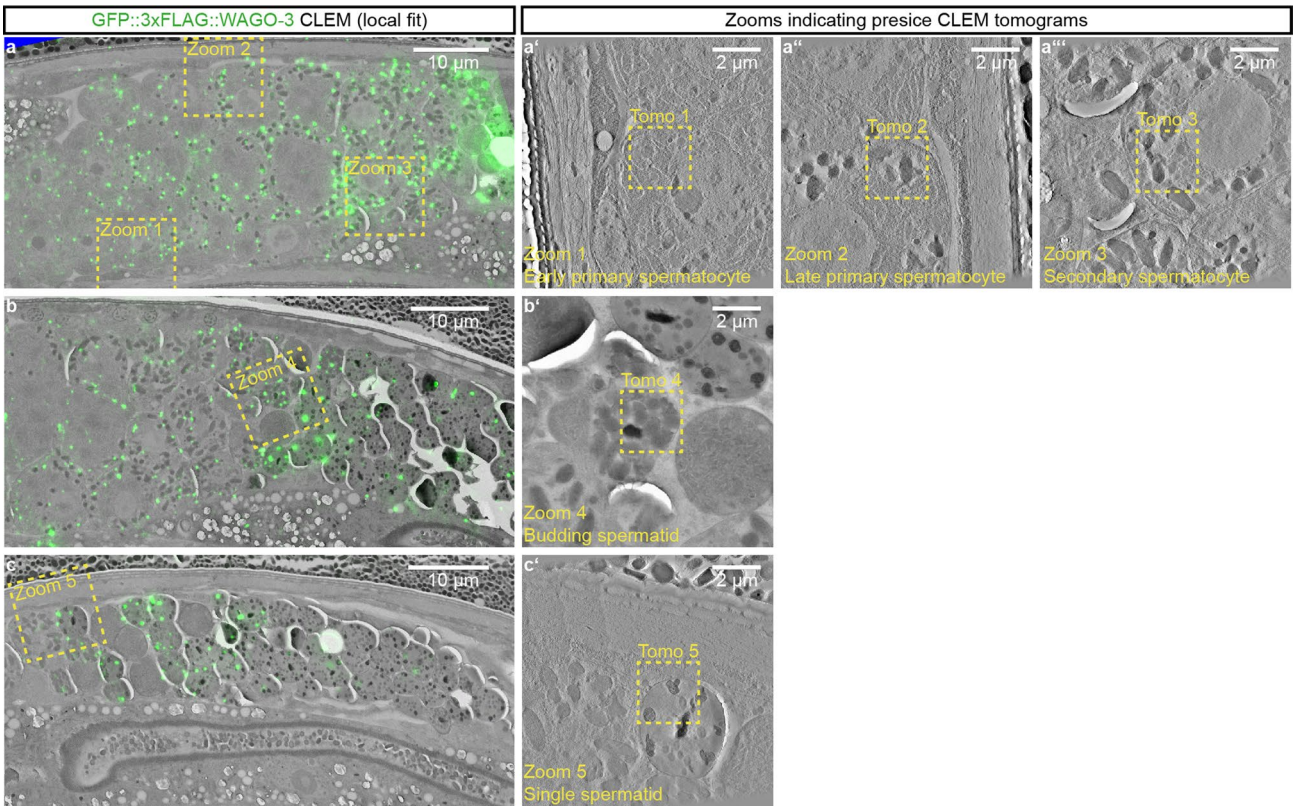


Extended Data Fig. 7 | See next page for caption.

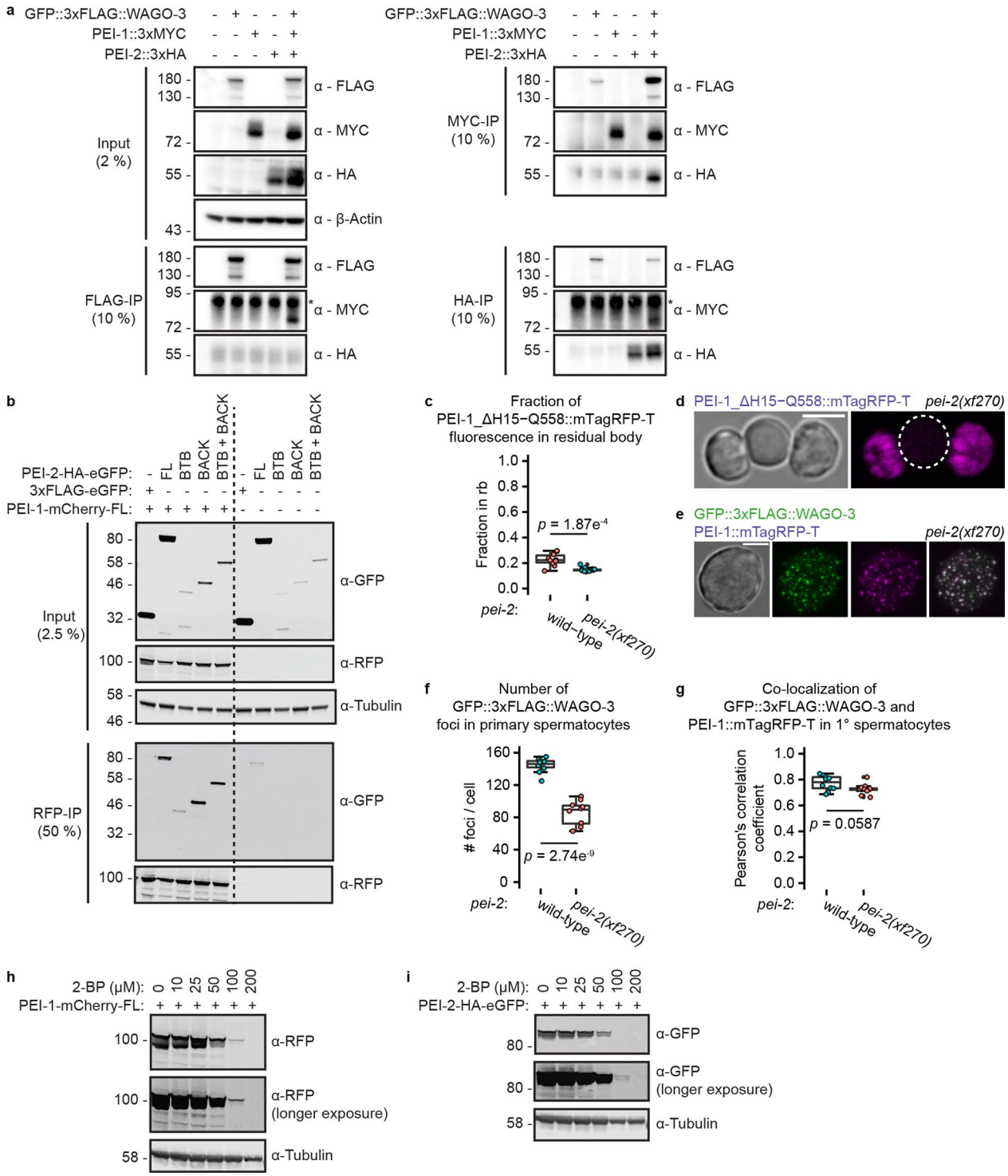
Extended Data Fig. 7 | The amino acid composition of the PEI-1 IDR affects exchange dynamics of WAGO-3. a,b, Occurrence of glycine and serine residues in PEI-1 (**a**) and PGL-1 (**b**) was counted in amino acid 50-mers, starting at position one, shifting 5 residues at a time, and displayed as stacked columns. Indicated residue positions in the diagrams are the mid-point of the 50-mer. *Y*-axes display number of relevant residues in amino acid 50-mers. *X*-axes indicate the position along the respective proteins. The various domains are indicated by vertical, dashed lines. NtDD and CDD indicate the N-terminal and C-terminal dimerization domains of PGL-1, respectively. The exact positions of glycine and serine residues for each protein are indicated above the stacked bar diagrams. IDR – intrinsically disordered region. **c-h**, Amino acid composition profiles of the intrinsically disordered region of the indicated proteins. Bars representing serine and glycine residues are highlighted in blue and orange, respectively. Panel **d** reflects a fusion between the PEI-1 IDR and the very C-terminal end of PGL-1(A711-F771). The profiles were generated using Composition profiler. Sequences were analyzed against the SwissProt database using 10,000 bootstrap iterations. Statistical significance was tested using the two sample t test (***: $p \leq 0.001$, **: $p \leq 0.01$, *: $p \leq 0.05$, ns: $p > 0.05$). The exact *P* values are provided as source data. **i-j**, Time lapse images showing fluorescence recovery after photobleaching (FRAP) of GFP::3xFLAG::WAGO-3 localizing to PEI granules via PEI-1::mTagRFP-T (**i**) or via PEI-1::PGL-1(A711-F771)::mTagRFP-T (**j**) in isolated, male-derived budding spermatids. Residual bodies are marked by a dashed circle. Images represent two biologically independent experiments. Scale bars: 4 µm. **k**, FRAP recovery curves of GFP::3xFLAG::WAGO-3 localizing to PEI granules, containing indicated PEI-1 proteins, in male-derived budding spermatids. Normalized data is presented as mean +/− SD and was fitted to a double exponential curve (n = 5 granules pooled from one independent experiment). Source data are provided.



Extended Data Fig. 8 | PEI-1 peptides at the N- and C-termini localize to asymmetrically segregated structures of defined shape. a, Confocal micrograph of an L4 hermaphrodite expressing free mTagRFP-T from the *pei-1* locus. The dashed line indicates the outline of the gonad. Scale bar: 20 μm. **b**, Confocal maximum intensity projection of spermatozoa within the spermatheca of an adult hermaphrodite expressing free mTagRFP-T from the *pei-1* locus. The dashed line indicates the outline of the spermatheca. Scale bar: 10 μm. **c**, Confocal Z-stack of an isolated, male-derived spermatocyte expressing PEI-1_ΔH15-Q558::mTagRFP-T (see Fig. 4a, middle construct). Z-size: 125.9 nm. Scale bar: 4 μm. **d**, Confocal micrograph of two isolated, male-derived secondary spermatocytes expressing GFP::3xFLAG::WAGO-3 and PEI-1_ΔH15-Q558::mTagRFP-T. The *pei-1* locus was heterozygous: *pei-1_ΔH15-Q558::mTagRfp-t/pei-1(+)*. This allowed the formation and visualization of both PEI granules and PEI-1_ΔH15-Q558-specific structures within the same animal. Scale bar: 4 μm. **e**, Line profiles displaying relative fluorescence intensity for PEI-1_ΔH15-Q558::mTagRFP-T signals versus GFP::3xFLAG::WAGO-3 signals over the indicated, dashed line shown in panel **d**. Vertical lines and colored circles indicate fluorescence peaks. a.u. – arbitrary unit. **f**, Confocal micrograph of an isolated, male-derived spermatocyte (left) and budding spermatids (right) showing the subcellular distribution of mitochondria and PEI-1_ΔH15-Q558-specific structures. MitoTracker Green FM was used to stain mitochondria. Residual bodies are marked by dashed circles. Scale bar: 4 μm. **g**, Line profiles displaying relative fluorescence intensity for PEI-1_ΔH15-Q558::mTagRFP-T signals versus mitochondria signals over the indicated, dashed line shown in panel **f**. Vertical lines and colored circles indicate fluorescence peaks. a.u. – arbitrary unit. All images represent two biologically independent experiments. Source data are provided.



Extended Data Fig. 9 | PEI granules are associated with membranous organelles throughout spermatogenesis. a-c, Overview GFP::3xFLAG::WAGO-3 CLEM montages acquired from three high-pressure frozen adult males expressing GFP::3xFLAG::WAGO-3 and PEI-1::mTagRFP-T. The depicted animals were used to collect the high-resolution CLEM images shown in Fig. 6a-e. In all three panels, germ cell development progresses from left to right. The GFP::3xFLAG::WAGO-3 signal is fitted locally, implying that fluorescence signal was fitted using landmarks (tetraspecks and Hoechst staining) spread over the entire region of interest, spanning a larger field of view. The panels **a'-a'''**, **b'** and **c'** depict zoom-ins of the indicated areas in panels **a**-**c**. The indicated 'Tomo' regions within these zoomed-in panels are shown in greater detail in Fig. 6a-e. Precise CLEM overlays at specific ROIs (Tomos 1 to 5) were done using landmarks more locally and close to each ROI, covering a smaller field of view. The EM grids shown in panels **a**-**c** can be navigated using Fiji software and data deposited to Mendeley Data (https://data.mendeley.com/datasets/dgb8d7h2hz/1), following the instruction listed in Supplementary Information. All images represent two biologically independent experiments.



Extended Data Fig. 10 | See next page for caption.

Extended Data Fig. 10 | PEI-2 interacts with PEI-1. a, Co-immunoprecipitation experiments using whole-worm extracts of late-L4 stage hermaphrodites analyzed by Western blotting. Labels above the blots indicate the presence (+) or absence (-) of the respective tags. Asterisks indicate non-specific signals. **b**, Pull-down experiments on extracts of transfected BmN4 cells expressing the indicated PEI-1 and PEI-2 variants. Full-length (FL) PEI-1-mCherry was pulled down, followed by detection of the various PEI-2 fragments. Expression of free 3xFLAG-eGFP served as negative control. **c**, Fraction of total mTagRFP-T signal within the residual body of male-derived budding spermatids expressing PEI-1_ΔH15-Q558::mTagRFP-T in wild-type or *pei-2(xf270)* mutant background (n = 10 cells pooled from two independent experiments, for each condition). Note that the wild-type data is the same as shown in Fig. 5c (ΔH15-Q558). **d**, Confocal maximum intensity projection of isolated, male-derived budding spermatids expressing PEI-1_ΔH15-Q558::mTagRFP-T in *pei-2(xf270)* mutant background. The residual body is marked by a dashed circle. Scale bar: 4 µm. **e**, Confocal maximum intensity projection of an isolated, male-derived primary spermatocyte expressing GFP::3xFLAG::WAGO-3 and PEI-1::mTagRFP-T in *pei-2(xf270)* mutant background. Scale bar: 4µm. **d**,**e**, Images represent two biologically independent experiments. **f**-**g**, Co-localization analysis between GFP::3xFLAG::WAGO-3 and PEI-1::mTagRFP-T (**g**), and quantification of GFP::3xFLAG::WAGO-3 foci number (**f**) in wild-type and *pei-2(xf270)* mutant, male-derived primary spermatocytes (n = 10 cells pooled from two independent experiments, for each condition). **c**,**f**,**g**, Statistically significant differences were determined by one-way ANOVA ($p \leq 0.001$). Boxplot centre and box edges indicate median and 25th or 75th percentiles, respectively, while whiskers indicate the median ± 1.5 x interquartile range. Note that the wild-type data in **f** and **g** are the same as those shown in Fig. 3j (FL) and Extended Data Fig. 4c (primary spermatocyte), and Fig. 3k (FL), respectively. **h**-**i**, Transfected BmN4 cells were treated with the palmitoylation inhibitor 2-BP at indicated concentrations, followed by Western Blot detection of PEI-1-mCherry (**h**) and PEI-2-HA-eGFP (**i**). α-tubulin served as loading control. **a**,**b**,**h**,**i**, The experiment has been performed once. Unprocessed original scans of blots are provided in source data.

Page left intentionally blank

Chapter II

Age- and caste-independent piRNAs in the germline and miRNA profiles linked to caste and fecundity in the ant *Temnothorax rugatulus*

Page left intentionally blank

Received: 22 May 2023 | Accepted: 26 September 2023

DOI: 10.1111/mec.17162

ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

Age- and caste-independent piRNAs in the germline and miRNA profiles linked to caste and fecundity in the ant *Temnothorax rugatulus*

Ann-Sophie Seistrup[1,2] | Marina Choppin[3] | Shamitha Govind[1,2] | Barbara Feldmeyer[4] | Marion Kever[3] | Emil Karaulanov[1] | Alice Séguret[5] | Sivarajan Karunanithi[1] | Miguel V. Almeida[1,6] | René F. Ketting[1,7] | Susanne Foitzik[3]

[1]Institute of Molecular Biology, Mainz, Germany

[2]International PhD Programme on Gene Regulation, Epigenetics & Genome Stability, Mainz, Germany

[3]Institute of Organismic and Molecular Evolution, Johannes Gutenberg University Mainz, Mainz, Germany

[4]Senckenberg Biodiversity and Climate Research Centre (SBiK-F), Molecular Ecology, Frankfurt, Germany

[5]School of Biological Sciences, University of Bristol, Bristol, UK

[6]Department of Biochemistry, University of Cambridge, Cambridge, UK

[7]Institute of Developmental Biology and Neurobiology, Johannes Gutenberg University Mainz, Mainz, Germany

Correspondence
René F. Ketting, Institute of Molecular Biology, Mainz, Germany.
Email: r.ketting@imb-mainz.de

Susanne Foitzik, Institute of Organismic and Molecular Evolution, Johannes Gutenberg University Mainz, Germany
Email: foitzik@uni-mainz.de

Funding information
Deutsche Forschungsgemeinschaft, Grant/Award Number: GRK2526/1 - Projectnr. 407023052

Handling Editor: Jacob A. Russell

Abstract

Social insects are models for studies of phenotypic plasticity. Ant queens and workers vary in fecundity and lifespan, which are enhanced and extended in queens. Yet, the regulatory mechanisms underlying this variation are not well understood. Ant queens live and reproduce for years, so that they need to protect their germline from transposable element (TE) activity, which may be redundant in short-lived, often sterile workers. We analysed the expression of two protective classes of small RNAs, microRNAs (miRNAs) and Piwi-interacting RNAs (piRNAs), in various tissues, castes and age classes of the ant *Temnothorax rugatulus*. In queens, piRNAs were highly abundant in ovaries with TEs being their clear targets, with reduced but still detectable piRNA-specific ping-pong signatures in thorax and brains. piRNA pathway activity varied little with age in queens. Moreover, the reduced ovaries of workers also exhibited similar piRNA activity and this not only in young, fertile workers, but also in older foragers with regressed ovaries. Therefore, these ants protect their germline through piRNA activity, regardless of ovarian development, age or caste, even in sterile workers often considered the soma of the *superorganism*. Our tissue-specific miRNA analysis detected the expression of 304 miRNAs, of which 105 were expressed in all tissues, 10 enriched in the brain, three in the thorax, whereas 83 were ovarian-specific. We identified ovarian miRNAs whose expression was related to caste, fecundity and age, and which likely regulate group-specific gene expression. sRNA shifts in young- to middle-aged queens were minor, suggesting delayed senescence in this reproductive caste.

KEYWORDS
ageing, castes, gene regulation, miRNAs, social insects, transposable elements

1 | INTRODUCTION

The origin of eusociality is considered one of the major transitions in evolution (Szathmáry & Smith, 1995). Insect societies show a high level of complexity; their colonies have been described as superorganisms (Boomsma & Gawne, 2018; Wheeler, 1911), in which the queen functions as the germline and the workers represent the soma. In ant colonies, for example, queens and workers are closely adapted to their specific roles. Ant queens focus on reproductive activity and are exceptionally long-lived, particularly in single queen societies (Keller & Genoud, 1997), wherein some queen ants can live for more than 20 years. In contrast, the typically sterile workers show high behavioural complexity performing all other tasks in the colony, including foraging and brood care. Relative to the queen, the average worker lifespan is considerably shorter, with most workers living less than a year. The different castes of social insects have been long regarded as models of phenotypic plasticity, as they are usually not genetically determined but arise due to differential transcriptional activity in the larval phase (Corona et al., 2016). While a strong investment in reproduction shortens the lifespan in many organisms (Flatt, 2011; Maklakov & Chapman, 2019), insect queens can maintain—especially in comparison to workers—a high fertility and long lifespan. Indeed, fecundity has been shown to be positively linked to a long lifespan in ants, even within a caste (Heinze & Schrempf, 2012; Negroni et al., 2020).

Our previous studies investigated the molecular regulation of ageing and fecundity in the small Myrmicine ant *Temnothorax rugatulus*, in which queens can reach lifespans of two decades. We saw that an increase in queen fecundity results in an upregulation of longevity pathways such as DNA repair and autophagy (Negroni et al., 2021). Moreover, workers that developed their ovaries upon queen removal showed an extended lifespan and transcriptional shifts in TOR and insulin-like/IGF-1 signalling (Choppin et al., 2021; Negroni et al., 2020, 2021). Gene expression shifts with age and fecundity in ant queens as well, with middle-aged queens investing more in antioxidants, but less in immunity and starvation resistance, compared to less fertile, young founding queens (Negroni et al., 2019). Although we have gained insights into the molecular pathways of ageing and fecundity in this ant and its different castes, we know little about the regulatory mechanisms underlying these shifts. All that has been shown is that histone acetyltransferase activity plays a role in the molecular regulation of ovarian development in workers in the absence of the queen (Choppin et al., 2021). We also do not know how the transcriptional differences between the castes, which lead to these divergent phenotypes, arise and are maintained.

In many social hymenopterans, workers can develop their ovaries and lay male-destined haploid eggs, a phenomenon known as arrhenotokous reproduction (Bourke, 1988). This occurs in response to specific social conditions, specifically when the queen dies or is lost. Indeed, the development and expression of reproductive traits in worker ants are socially regulated, which ensures the maintenance of the reproductive division of labour within the colony. In the ant *T. rugatulus*, workers start fighting over dominance within hours and activate their ovaries within weeks of queen loss. These workers do not only exhibit activated ovaries but also an extended lifespan and altered regulation of mTOR or insulin-like/IGF-1 signalling (Choppin et al., 2021; Negroni et al., 2020, 2021).

In this study, we aimed to describe small regulatory RNAs (sRNAs) in the ant *T. rugatulus*, and to investigate their relationship to ovarian activity, caste and age in female ants. As will be described below, different classes of sRNAs exist, most notably microRNAs (miRNAs) and Piwi-interacting RNAs (piRNAs), and their activities have been linked to germ cells and ageing. For example, the expression of miRNAs changes during ageing in many organisms associated with transcriptional shifts of genes involved in senescence and age-related diseases (Kinser & Pincus, 2020). In the honeybee *Apis mellifera*, specific miRNAs are associated with age-dependent behavioural changes (Behura & Whitfield, 2010). PiRNAs represent a class of sRNAs that is particularly active in the germline (Ketting & Cochella, 2021; Ozata et al., 2019), where they regulate the activity of transposable elements (TEs). There is also growing evidence that TE activity is a consistent molecular hallmark of ageing (Sturm et al., 2017). In the termite *M. bellicosus*, TE-associated genes are upregulated and genes from piRNA pathways are downregulated in old workers with short residual lifespans, compared to long-lived reproductives (Elsner et al., 2018). Moreover, TE expression in the fat body of the termite *Macrotermes natalensis* does not rise with increasing queen age, likely due to the protective upregulation of piRNA pathways in this reproductive caste with delayed ageing (Post et al., 2023). Similarly, in *A. mellifera*, the piRNA repertoire and expression levels are greater in reproductive individuals than in sterile workers, with the ovaries, representing the female germline, being the tissue of highest expression (Wang et al., 2017). Haploid males, in which deleterious effects of TEs dominate, exhibit the highest piRNA levels, but surprisingly also the highest TE expression. The honeybee is a TE-deficient species with TEs making up only 3% of the genome (Elsik et al., 2014; Kapheim et al., 2020). Yet, an active piRNA system is still required to protect the vulnerable haploid male genome and the reproductive castes from TE-induced genomic instability (Wang et al., 2017). Ant genomes often contain higher proportions of TEs (Gilbert et al., 2021). For example, recently published ant genomes reported values around 20%–32% (Bohn et al., 2021; Faulk, 2023; Nouhaud et al., 2022) and we estimate that 33.3% of the *T. rugatulus* genome is covered by repetitive elements.

Different classes of sRNAs exist, which all form effector complexes with proteins of the Argonaute family and typically repress gene activity (Ender & Meister, 2010). Two of the major sRNA pathways are the miRNA pathway (reviewed in Bartel, 2009) and the piRNA pathway (reviewed in Ozata et al., 2019) and we focus here on the expression of these two classes of sRNAs. miRNAs are transcribed from distinct loci and the resulting transcripts form hairpin structures, which are processed by specific nucleases to produce mature miRNAs (Bartel, 2009). After binding their respective Argonaute proteins, miRNAs direct the binding of this protein to specific mRNAs, based on base-pairing between the miRNA and the mRNA, and repress their activity (Filipowicz et al., 2008). As

miRNAs bind their target mRNAs through only partial complementary base-pairing, it can be difficult to predict which genes a certain miRNA targets, and many tools exist to tackle this (Bartel, 2009; Lewis et al., 2003). The targets of miRNAs are typically protein-coding genes that have regulatory roles themselves (Bartel, 2009).

Much like miRNAs, piRNAs are transcribed from distinct loci in the genome. However, unlike miRNAs, which can be found in any tissue, piRNAs are often restricted to the germline. This is at least true for model organisms like *Drosophila*, *Caenorhabditis elegans* and mice. However, in various arthropods, somatic piRNAs have been detected (Lewis et al., 2018) suggesting that the ancestral state of piRNA expression may have been more widespread. The processing of piRNA precursor transcripts into mature piRNAs does not proceed via a hairpin intermediate, but utilizes a different, complex molecular machinery as reviewed in Zhang et al. (2022). Mature piRNAs are bound by a specific subfamily of Argonaute proteins: Piwi proteins. Piwi–piRNA complexes are known to elicit strong silencing effects, both at the transcriptional and post-transcriptional levels (Ozata et al., 2019). When a Piwi–piRNA binds a target, it may turn the targeted mRNA into a new piRNA in a process known as the ping-pong cycle (Luo & Lu, 2017; Ozata et al., 2019), which results in sense and antisense piRNAs that overlap precisely 10 nucleotides at their 5′ ends. Finally, piRNAs have a strong bias towards uracil at the 5′-position, matched by an adenine at position 10, due to the described ping-pong cycle. A deeply conserved role of piRNAs is the silencing of TEs and the protection from non-self DNA (Madhani, 2013). As a consequence, piRNA sequences are typically not conserved, even between closely related species, as they co-evolve with the TEs they control.

Here, we characterize the sRNA classes in the ant *T. rugatulus* with a focus on miRNAs and piRNAs and by using high-throughput sequencing of various castes, age classes and tissues, and determine whether sRNA profiles are tissue and/or caste specific. We also explored whether they change with age or reproductive status of the queen and worker castes.

2 | MATERIALS AND METHODS

2.1 | Ant collection and maintenance

Temnothorax rugatulus ants are widespread in western North America and live in coniferous forests, under rocks or in crevices. In August 2015 and 2018, we collected several hundred colonies in the Chiricahua Mountains (see Choppin et al., 2021; Negroni et al., 2019 for GPS information on exact locations). Collection permits for the Coronado National Forest were obtained through the Southwestern Research Station of the Museum of Natural History in Portal, Arizona. The ant colonies were transported to our laboratory in Mainz, Germany, where they were relocated to individual boxes with plastered flours containing nests made of plastic inserts between two glass plates covered with red foils to block the light. The colonies were kept in climate chambers at 18°C and 70% humidity in a 12:12 light–dark cycle. They were fed with honey and crickets and received ad libitum water.

2.2 | Analysis of miRNA and piRNA expression depending on queen age and tissue

For our first aim of investigating differences between tissues and age of queens, we sequenced RNAs from young and old queens. Seven young queens were collected during our 2018 trip as founding queens either alone or with their first workers, invariably less than six. These queens likely had established their colony in July or August of that year, so they were less than 6 months old. Seven older queens were collected in Arizona in August 2015 in large single-queen colonies, indicating that, at the time of collection, they were likely several years old. We kept them for another 3 ½ years so that at the time of sampling, we estimate their age to be between 6 and12 years. *Temnothorax* ant queens can live for up to 20 years (Plateaux, 1986), so our older queens were probably middle-aged when sampled. At the time of sampling, colonies of these young and older queens comprised between 60 and 230 workers. In November 2018, the queens were provided with fresh boxes, nests and food and housed in a climate chamber at 25°C and 70% humidity on a 12:12 light–dark cycle. All queens came from single-queen colonies.

After 1 month under the same conditions, the queens were dissected in ice-cold PBS under a Leica stereomicroscope. The head with the thorax and the ovaries were removed from each queen, separately frozen in liquid nitrogen and stored at −80°C. We measured ovariole length and counted the number of white eggs (i.e. eggs in development) in the ovaries using the Leica software LAS v4.5. One young queen had died before the dissection date and was therefore not sampled. Differences between old and young queens were assessed using R v.4.3.1. For the number of eggs, we used a generalized linear model (glm(eggs ~ age, family = quasipoisson)), and for the lengths, we used a linear model (lm(length ~ age)).

In August 2020, the brains of the queens were isolated from their heads on ice, placed in 50 μL TRizol™ (Life Technologies), frozen in liquid nitrogen and stored at −80°C. The brain of one young queen was not removed because her head was damaged. The brain was placed in 50 μL TRIzol, frozen in liquid nitrogen and also stored at −80°C. Two weeks later, the RNA was extracted from the brain and thorax samples. For this, 50 μL chloroform was added to the samples, the tubes were swirled and placed on normal ice for 5 min. The samples were then centrifuged at 12000 × *g* and 4°C for 15 min. The resulting upper phases were transferred to new 1.5 mL Eppendorf tubes using gel loading tips. Then, 25 μL of 100% ethanol was added to the upper phases, the tubes were swirled and the samples were transferred to Zymo-Spin IC columns (Direct-zol RNA Microprep Kit from Zymo Research) in collection tubes. The collection tubes containing the columns were centrifuged at 12000 × *g* for 30 s at room temperature, and the columns were then transferred to new collection tubes. 400 μL RNA wash buffer was added to the columns, which were then centrifuged at 12000 × *g* for 30 s at room temperature.

5 μL DNase I was carefully mixed with 35 μL DNA digestion buffer and added to the columns. The samples were incubated for 15 minutes at room temperature. Then, 400 μL Direct-zol PreWash was added to the columns and they were centrifuged at 12000×*g* for 30 s at room temperature. The residues were discarded, and the previous step was repeated. 700 μL of RNA wash buffer was added to the columns, which were then centrifuged at 12000×*g* for 1 min at room temperature. The columns were transferred to RNAse-free tubes and 15 μL of DNAse/RNAse-free water was added directly to the column matrices. Finally, the collection tubes containing the columns were centrifuged at 12000×*g* for 30 s at room temperature to elute the RNA and the columns were discarded. The RNA samples were stored at −80°C until sequencing. In November 2020, 50 μL of ice-cold TRIzol was added to the ovary samples and the RNA was extracted using the same procedure as described above.

2.3 | Analysis of miRNA and piRNA expression in ovarian tissue of queens and workers

For our second aim of comparing queens to workers of different fertility stages, we sequenced RNA from queens and workers either from colonies with a queen (queenright, ql) or from colonies lacking a queen (queenless, ql). *T. rugatulus* female ants display a gradient in terms of reproductive potential and life expectancy. Both traits are most pronounced in the queen; only she can mate, and her ovaries also possess a spermatheca and eight ovarioles. Worker ovaries are smaller and are composed of only two ovarioles. Fertile nurses in ql colonies are the longest lived and most fertile individuals after the queen. They are followed by the nurses from qr colonies and finally by the old foragers from these colonies. This gradient allows us to analyse at which point these female ants stop protecting their germline from TEs or whether germline tissue always expresses Piwi protection.

At the end of 2021, 12 qr and 12 ql colonies with between 20 and 100 workers were provided with fresh boxes, nests and food and housed in a climate chamber at 25°C and 70% humidity on a 12:12 light:dark cycle. We aimed to sample four independent replicates of each of the four groups: queens, qr nurses (qr-n) and foragers (qr-f) and ql nurses (ql-n). After 1 month, in January 2022, we dissected the ovaries of the queen, four nurses and five foragers from 12 qr colonies each, and the ovaries of four nurses from 12 ql colonies. For each queen replicate, the ovaries of three queens from three different colonies were pooled. For each nurse replicate (qr and ql alike), we pooled the ovaries of four workers from three colonies, with no overlap between replicates, for a total of 12 ovaries. For the forager replicates, we pooled the ovaries of five foragers from three colonies, for a total of 15 ovaries each. We have pooled 15 ovaries from foragers to get enough material, as foragers are known to have the most regressed ovaries, as we show below. Dissections were conducted in ice-cold PBS under a stereomicroscope and ovaries were photographed using the Leica system LAS v4.5. We noted the number of eggs in development and grouped the ovaries of queens and workers into five developmental stages (0 = regressed ovaries: thin, clear tubes; 1 = undeveloped ovaries with rounded tubes and a wider end, 2 = slightly developed ovaries containing immature eggs; 3 = developed ovaries with one mature egg, 4 = well-developed ovaries with 2–5 mature eggs, 5 = extremely well-developed ovaries with >5 eggs). Dissections took less than 5 minutes and ovaries were transferred into 100 μL Trizol, which was kept on ice until all ovaries of one pooled sample were added and then stored at −80°C. Differences between groups were assessed using R v.4.3.1. For the number of eggs, we used a generalized linear mixed effects model (glmer(eggs ~ group + (1|Colony), family = poisson). Due to a singularity in the model, we could not include colony ID in the analyses of developmental stages and instead performed a glm (glm(stage ~ group, family = poisson)).

Each sample was ground while frozen and then RNA was extracted with the Direct-zol RNA Microprep kit following the standard instructions. RNA was resolved in 15 μL RNAse/DNAse-free water and stored at −80°C before sequencing.

2.4 | Sequencing

NGS library prep was performed with NEXTflex Small RNA-Seq Kit V3 following Step A to Step G of Bio Scientific's standard protocol (V19.01) using the NEXTFlex 3' SR Adaptor and 5' SR Adaptor

(5'rApp/NNNNTGGAATTCTCGGGTGCCAAGG/3ddC/and 5'GUUCAGAGUUCUACAGUCCGACGAUCNNNN, respectively). Libraries were prepared with a starting amount of 5 ng and amplified in 21 PCR cycles (1st data set, different tissues of old and young queens), or a starting amount of 1.5 ng and amplified in 22 PCR cycles (2nd data set; ovaries of different castes). Note that we do not directly compare these two different data sets, such that this difference in PCR cycle number does not affect our conclusions. Amplified libraries were purified by running an 8% TBE gel and size selected for 15–40 nt. Libraries were profiled in a high sensitivity DNA Chip on a 2100 bioanalyser (Agilent Technologies) and quantified using the Qubit dsDNA HS Assay Kit, in a Qubit 2.0 fluorometer (Life Technologies). All samples were pooled in equimolar ratio and sequenced on 1 high-output NextSeq 500/550 flow cell, SR for 1x 84 cycles plus 7 cycles for the index read.

2.5 | Read preprocessing, mapping and filtering

NGS library quality was assessed using FastQC v0.11.8 (Andrews, 2010) before removing the constant sequence of the 3'-NextFlex adapter using Cutadapt v2.4 (Martin, 2011) (–a TGGAA TTCTCGGGTGCCAAGG -m 26 -M 43). Trimmed libraries were once again assessed for quality using FastQC. Reads were mapped to the draft genome assembly 'trug_v1.0' (BioProject PRJNA750352, in submission) using Bowtie v1.2.2 (Langmead et al., 2009) allowing for one mismatch and reporting one best alignment in cases of multimapping while concomitantly removing the random 2x 4 nt random NextFlex adapter bases (bowtie -p 16 -v 1 -M 1 -y --best

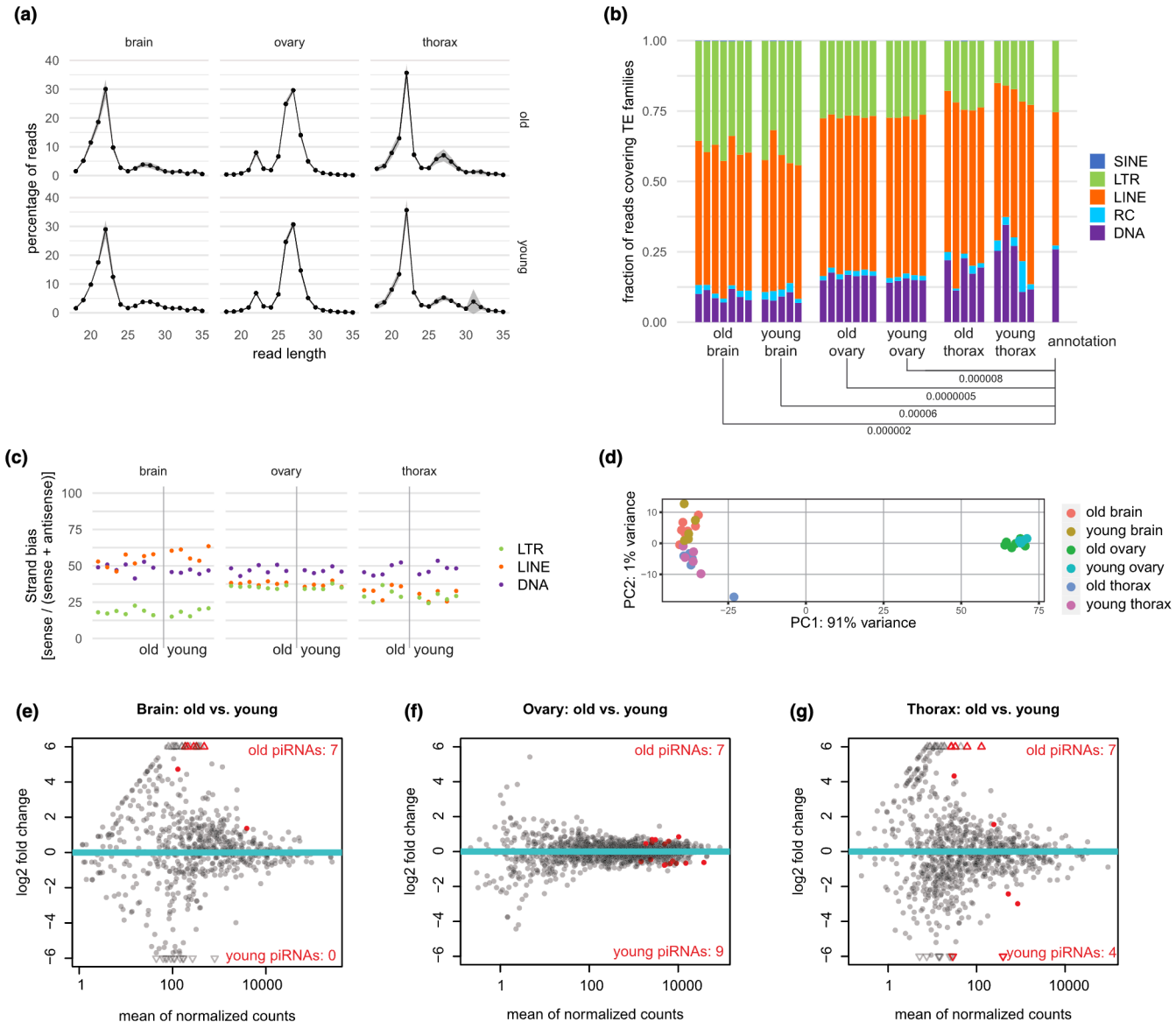


FIGURE 1 piRNA expression in different tissues of young- and middle-aged queens. (a) Length distribution of 18–35 nt reads generated by sRNA-Seq. Shaded areas depict standard deviation, obtained from the five to seven biological replicates. (b) Stacked bar plots showing piRNA-like reads (25–30 nt) covering different TE families. *T*-tests carried out for each TE family to find difference between old and young and between each group and annotation (random expectation). For each pair of comparisons, the lowest Holm-adjusted *p*-value is shown, if this is significant ($p < .05$). All values not shown are listed in the supplementary data. (c) Strand bias (%) of piRNA-like reads derived from LTR, LINE and DNA elements in the three tissues, obtained from young and older queens ($n = 5–7$). (d) Principal component (PC) analysis of piRNA-like reads (25–30 nt) that map to annotated TEs. (e–g) MA plots showing the pairwise differential TE comparisons of piRNA targeting in young versus older queens in the three tissues. Each dot represents one TE type and red indicates significant changes (FDR <0.05). All hits are reported in Table S1.

--strata --trim5 4 --trim3 4). Read lengths were assessed by a custom Python script (https://github.com/Tunphie/SequencingTools/blob/main/summarizeNucleotideByReadLength.py) and plotted using ggplot2 (Wickham, 2016) in R v3.13 (R Core Team, 2013). Based on the length profiles as shown in Figure 1a, mapped reads were subsequently split into different lengths for further analysis of piRNAs (reads of length 25–30 nt) and miRNAs (reads of length 18–24 nt) using samtools v1.9 (Danecek et al., 2021) and GNU awk. These mapped reads were further split into non-overlapping and overlapping reads with miRNA or transposon annotation in sense or antisense orientation using the function intersectBed from bedtools v2.27 (Quinlan & Hall, 2010). This was done in order to determine lengths and 5′-biases using the aforementioned Python script before plotting these with ggplot2 in R.

2.6 | piRNA analyses

Reads in the piRNA-range (25–30 nt) existing outside of our annotated transposons were split into plus and minus strand using

samtools v1.9. BigWig files were then generated for each strand separately using the deepTools v3.5.1 (Ramírez et al., 2016) function bamCoverage. Means of groups were calculated using WiggleTools (Zerbino et al., 2014) and converted back into BigWig files using wigToBigWig from UCSC. The tools computeMatrix and plotProfile from deepTools v3.5.1 were used to generate views of the mapped data coverage over selected loci provided as a BED file.

Reads in the piRNA-range (25–30 nt) overlapping with our annotated transposons (supplementary file TE_Annotation.trug.fa.out.gv; courtesy of Jongepier et al. (2022)) were counted using Subread v2.0.0 (Liao et al., 2013) (featureCounts -s 0 -M -F SAF), thereby summarizing read counts per transposon of the same type ($n = 1903$). For strand-bias analysis, featureCounts was used with parameter '-s 1' for TE-sense and '-s 2' for TE-antisense read counting. The provided annotation contain TEs covering 27.1% of the *T. rugatulus* genome, defined from an initial 33.3% repetitive elements found by use of RepeatMasker (v4.0.8; Smit et al., 2013–2015). Reads in different transposon families were subsequently summarized in R by comparison to the used annotation file. *T*-tests and subsequent Holm-correction for multiple testing were carried out in R. Differential analyses were carried out using DESeq2 v1.30.1 (Love et al., 2014) in R v4.0.3. A false discovery rate (FDR) or 0.05 was used without a fold-change cut-off. PCA plots and MA plots were generated using plotPCA and plotMA from the DESeq2 package.

Samples were randomly downsized to 500 K (old and young queens; thorax, brain and ovaries) or 2 million (ovaries of queens, nurses and foragers) reads per sample for our two data sets, respectively, using first samtools collate and then shuf (−n 500,000/-n 2,000,000). These reads were then used to search for ping-pong signatures using PingPongPro v1.0 (Uhrig & Klein, 2019, −p -v -o). Ping-pong pairs were compared to the TE annotation and divided into signatures existing within TEs and outside of TEs using R. PingPongPro was further run using the TE annotation as input, and mapping relative to TE families was counted using R. Plots were generated using ggplot2 in R. UpSet plots were generated using UpSetR (Conway et al., 2017).

2.7 | miRNA prediction and annotation

We used miRDeep v2.0.1.3 (Friedländer et al., 2012) on reads of length 18–24 nt to predict miRNA de novo. We first used mapper.pl script (−d -e -h -i -j -m -o 30 -v -p -s -t) to process the reads and map them to our partial *T. rugatulus* genome. The outputs from mapper.pl along with our partial genome and miRNA annotations of four insect species (*A. mellifera*, *D. melanogaster*, *T. castaneum* and *B. mori)* from miRBase v22 were used to predict miRNAs using the miRDeep2.pl script (−g 100,000 -P). We quantified the predicted miRNAs using the quantifier.pl (−d) script. Furthermore, we removed predicted miRNAs that did not satisfy the following quality filters adapted from Coenen-Stass et al. (2018). The miRNAs should (i) have a miRDeep2 score above 1, (ii) have a statistically significant miRDeep2 randfold *p*-value ($p < 0.05$), (iii) have a sequence that does not match any tRNAs or rRNAs in RFAM (Griffiths-Jones et al., 2003) and (iv) have at least one read in three or more of our samples from our first data set (old and young queens; thorax, brain and ovaries). The predicted miRNAs were stringently annotated by miRDeep2 looking for exact seed match in the four insect species. Additional low-confidence annotation was made by blasting the sequences of the miRNAs of interest against a broad category of known hexapod miRNAs from miRBase using blast v2.10.1 with relaxed parameters (−evalue 20 -word_size 4 -reward 5 -penalty −4 -gapopen 25 -gapextend 10 -num_threads 10 -max_target_seqs 1 -outfmt 6). The specific software parameters used are mentioned in parentheses following the respective scripts.

2.8 | miRNA analyses

Reads in the miRNA-range (18–24 nt) aligning to our miRNA annotation (mature sequences) were counted using Subread v2.0.0 (featureCounts -s 1 -M -F SAF --minOverlap 18). miRNA expression for a given tissue (brain, ovary and thorax) was defined as non-zero read counts in at least three samples from that tissue regardless of age. This number was chosen arbitrarily, as choosing two or four samples yielded almost the same results (115 or 96 miRNAs, respectively, instead of the 105 we now report). Overlap of these reads were plotted using VennDiagram in R (Chen & Boutros, 2011). Differential analyses were carried out using DESeq2 v1.30.1 in R v4.0.3. A false discovery rate (FDR) of 0.05 was used without a fold-change cut-off. PCA plots and MA plots were generated using plotPCA and plotMA from the DESeq2 package.

We note that some of our reads may be misclassified; either as the wrong sRNA type or as a sRNA when it was indeed the breakdown product of a functional RNA (Ludwig et al., 2017). Some piRNAs might also have been breakdown product of rRNAs (Figure 3d,f,g and Figure S2d). A few non-TE-derived piRNA reads resembled structural reads (Figure S1–S2d). Also some piRNAs may have been misclassified as miRNAs due to the closeness of their length distributions and their similar 5′-bias. We do, however, believe that most piRNAs would be filtered out of our differential miRNA analyses since these are the only reads mapping sense to our miRNA annotation. We also are aware that sRNA populations other than miRNAs and piRNAs exist, and that some of our sRNAs may have been misclassified due to not investigating further sRNA families such as small interfering RNAs (siRNAs).

3 | RESULTS

3.1 | Basic sRNA profiling of three *T. rugatulus* tissues

In order to investigate tissue-specific expression of sRNAs in *T. rugatulus*, we sequenced sRNAs of 15–40 nucleotides (nt) from the thorax, ovaries and brain of young, founding queens below 6 months of age and middle-aged established queens above 3.5 years. When comparing the read length distribution, two distinct populations

were observed; one in the typical miRNA range (18–24 nt) and one in the piRNA range (25–30 nt; Figure 1a). The peak in the piRNA range was more prominent in the ovaries than in the somatic tissues, consistent with findings in other animals of high piRNA pathway activity in the germline. In contrast, the miRNA-range peak was more prominent in the somatic tissues, although a distinct piRNA-range peak could still be observed in these tissues (Figure 1a). To probe deeper into the characteristics of the sRNA molecules, we aligned the reads to a genome assembly from Jongepier et al. (2022). A large fraction of the reads mapped to multiple locations in the genome (Figure S1–S2a), which could mean that these map to repetitive elements such as TEs. Below, we will describe the miRNA- and piRNA-range sRNAs in soma and germline in separate sections. Lastly, comparisons between young- and middle-aged queens will be described.

3.2 | Queen ovaries express TE-targeting piRNAs

After having determined that ovaries have a prominent piRNA-like sRNA population, which peaked at abundance at lengths of ~26–28 nt, we extracted all reads in the range 25–30 nt and compared these to a TE annotation provided by Jongepier et al. (2022), since piRNAs are known to be involved in TE repression, particularly in gametes. We saw that many reads, up to 60%, overlapped with annotated TEs in the queen ovaries (Figure S1–S2b). We further extracted these TE-derived reads and saw that they mapped to both sense and antisense strands of TEs, with some bias for the antisense strands, as has been previously shown for piRNAs in other species (Figure S1–S2c). The reads furthermore had a strong bias for U at the 5′-position, another common feature of piRNAs (Figure S1–S2c). Compared to the genome-wide TE annotation, DNA and rolling circle-derived piRNAs were slightly underrepresented in the piRNA-like reads in the ovaries, whereas LTR- and LINE-derived reads were mildly overrepresented (Figure 1b).

In the absence of available gene annotation, we could not determine where the piRNA-like reads, that were not assigned to TEs, mapped. Nevertheless, the reads mapping outside of TEs also had a 5′-U bias and showed mapping on both genomic strands similar to the TE-derived sRNAs (Figure S1–S2d) with at least some of the sense/antisense reads stemming from the same genomic position (Figure 1e). We noticed that some loci were strongly biased to produce sRNAs from only one of the strands (Figure S1–S2e), suggesting that these may not be true piRNAs. Alternatively, these may represent highly strand-biased populations of piRNAs (Brennecke et al., 2007; Hirano et al., 2014).

Taken together, the above results are consistent with the idea that this 25–30 nt sRNA population corresponds to piRNAs. To further assess this, we probed for so-called ping-pong signatures, which arise when cleaved piRNA target transcripts are themselves turned into piRNAs. Such piRNAs will display a characteristic 10 nt overlap on their 5′ end with the piRNA that induced the cleavage. We first downscaled all data sets of 25–30 nt reads to 500 K reads in order to enable comparison between them, and then searched for possible ping-pong signatures using PingPongPro (Uhrig & Klein, 2019). We found around 4000 high-fidelity ping-pong pairs in the ovaries across the entire genome, around two-thirds of which were found inside annotated TEs (Figure S1–S2f). We then looked for ping-pong pairs specifically in the TEs and noticed that the majority of them were derived from LTR and LINE elements with more than 50% of TE-derived ping-pong signatures stemming from LTR elements (Figure S1–S2g). Given that LTR elements only amount to 25% of the annotated transposons, this represents a significant overrepresentation and suggests that these TE types are most actively targeted by the Piwi pathway. In the absence of functional antibodies, which would allow us to test whether the identified sRNAs are bound to Piwi proteins, we conclude that the reported sRNA species represent piRNAs, and that *T. rugatulus* ovaries express an active piRNA pathway, with LTR retrotransposons as important targets.

3.3 | Thorax and brain have limited piRNA expression in queens

Next, we probed the queens' thorax and brain samples for the existence of piRNAs. In both tissues, we did find a population of small RNAs that displayed similar features as the piRNAs in the ovaries with length peaks at roughly 27 nt (Figure 1a). They showed a mild bias towards antisense polarity with respect to their matching transposons, and they displayed a strong bias for a 5′-U (Figure S1–S2c). Interestingly, the brain had more piRNAs mapping outside of annotated TEs (Figure S1–S2b,d,e), possibly pointing towards a function of piRNAs other than transposon surveillance.

In contrast to ovaries, ping-pong signatures were practically absent in the brain (0–200 signatures per 500 K reads per sample). Some high-fidelity ping-pong signatures, around 500 per 500 K reads (overall average: 459, average in old samples: 572, average in young samples: 346), were found in the thorax (Figure S1–S2f), but their abundance was still roughly 5- to 10-fold lower compared to ovaries. Ping-pong results were not strongly affected by age (Figure S1–S2f).

Then, we checked how the available piRNAs were distributed over the different TE families. In the thorax, the piRNAs were distributed along the annotated transposons roughly according to their representation in the genome, albeit with high variance between samples (Figure 1b). Interestingly, the brain samples showed a consistent overrepresentation of LTR elements relative to DNA elements (Figure 1b). These results could point to differential activities of TE families in these two tissues.

We noticed that the strand bias of LTR and LINE elements differed in the brain relative to the ovaries and thorax (Figure 1c): LTR-derived piRNAs showed a much more extreme anti-sense bias, whereas LINE-derived populations showed a weaker anti-sense bias and even a mild over-representation of sense piRNAs. These findings may also reflect differential TE activities in the brain compared to the thorax and ovaries.

3.4 | piRNA expression shifts only minimally with queen age

We saw no difference in TE family targeting or strand bias between young- and middle-aged queens (Figure 1b,c). In order to further determine whether ageing of *T. rugatulus* queens is associated with differential targeting of specific TEs by piRNAs, we carried out differential analyses of this subset of sRNAs using DESeq2 (Love et al., 2014). piRNAs found to be differentially expressed between young and older queens in different tissues are provided in Table S1.

The principal component analysis revealed that the largest differences in piRNA expression existed between ovarian and somatic tissues, irrespective of age (Figure 1d). Nevertheless, we investigated further the effects of age in all three tissues. In the ovaries, we detected 16 differentially targeted transposons, 7 of which were Penelope elements. The differential abundance was, however, minor (Figure 1f). The piRNA profiles of thorax and brain were noisier, but we were able to define 11 and 7 differentially targeted transposons respectively (Figure 1e,g). In the thorax, much like in the ovaries, these were mainly Penelope elements (6/11), whereas the transposons significantly affected in the brain were mainly targeted by lowly expressed piRNAs (Figure 1e). We conclude that piRNA expression changes only marginally with age in *T. rugatulus* queens.

3.5 | Foragers have less developed ovaries than nurses and queens

The weak differences in piRNA expression between young- and middle-aged queens led us to determine their fecundity, thinking that low fertility and loss of ovarian piRNA expression might be associated. The ovaries of young- and middle-aged queens did not differ in the number of eggs in development (glm (quasipoisson): $df = 1$; $\chi^2 = 1.54$; $p = .21$, Figure 2a), nor in ovariole length (lm(length ~ age), Figure 2b). In contrast, we found strong differences in egg production and ovarian development between the different groups of queens and workers, when we dissected the ovaries of four different groups; queens, nurses and foragers from colonies with a queen (queenright, qr) and nurses from colonies lacking a queen (queenless, ql). Queens had more eggs than all worker types, while foragers had fewer eggs in development than nurses (glmer(family = poisson + (1|Colony)): $df = 3$; $\chi^2 = 427.03$; $p < 2.2e\text{-}16$; queen – ql-nurse: $p < .001$; queen – qr-forager: $p < .001$; queen – qr-nurse: $p < .001$; qr-forager – ql-nurse: $p < .001$; ql-nurse – qr-nurse: $p = .79$; qr-nurse – qr-forager: $p < .001$, Figure 2c). We further defined the developmental stage of the ovaries ranging from 0 (undeveloped; no eggs in development) to 5 meaning (fully developed; more than five eggs in development) (Figure 2d). Here, we could also see that queens had much more developed ovaries than any worker type, while nurses generally had more developed ovaries than foragers (glmer (family = poisson): $df = 3$; $\chi^2 = 72.483$; $p = 1.255e\text{-}15$; queen - ql-nurse: $p < .003$; queen – qr-forager: $p < .001$; queen – qr-nurse $p < .001$; qr-forager – ql-nurse $p < .001$; ql-nurse – qr-nurse $p = .68$; qr-nurse – qr-forager $p < .001$, Figure 2e). We were surprised to observe that there were no significant differences between the ovarian developmental stages of nurses from qr and ql colonies (Figure 2c,e), though workers from ql colonies were shown to undergo ovarian activation (Choppin et al., 2021; Negroni et al., 2020, 2021).

3.6 | piRNAs are expressed in the ovaries of queens and all worker types

After having determined that, in queens, piRNAs act primarily in the ovaries, we sought to investigate the piRNA profiles of ovaries from workers, even if they are only present as almost rudimentary organs. To allow proper comparison, we again sequenced sRNAs from the ovaries of queens, but this time also included the ovaries of foragers and nurses from qr colonies as well as nurses from ql colonies, where activation of the nurses' ovaries is triggered. Analysing these four different groups allowed us to investigate differences in caste, worker task and conditionary changes to nurses. Surprisingly, we obtained near-similar piRNA expression profiles in the ovaries of all four groups (Figure 3a and Figure S2a) and only a slight depletion of ping-pong signatures in worker ovaries (Figure S2b). Multimapping reads were again abundant in all samples (Figure S2c). As previously seen for the different queen tissues (Figure 1b), TE family targeting did not differ between these different groups. Relative to the distribution of all annotated TEs, all samples showed a bias towards LINE elements at the expense of DNA elements (Figure 3b).

We carried out a differential analysis of our defined piRNAs using DESeq2 and observed only minor differences between the different groups, reflected in the clustering in the principal component analysis (Table S2, Figure 3c). Nonetheless, we could define several TEs with differential piRNA expression between groups with a particular upregulation of piRNAs in queens compared to foragers (Figure 3d–i). In general, those TEs that were found to be differentially targeted were those for which piRNAs were generally more abundant (Figure 3d–h), suggesting that, although the general upregulation of piRNAs in active versus inactive ovaries is modest (Figure 3a), the effect likely is biologically relevant.

We found one TE to be highly upregulated in foragers in several analyses (Figure 3d,f,g), namely the Copia element. However, when further investigating the various Copia loci, we noticed that only one peak existed in each locus and that it consisted of only sense reads with no ping-pong signatures. We therefore extracted the sequences in the peaks and saw that these all aligned not only to each other but also to the 3′ end of large subunit rRNA from *Bombus affjnis* (Figure S2d). Given the strong homologies of rRNA sequences between species, we anticipate that this potential Copia-derived piRNA is not a genuine piRNA but rather an rRNA fragment that was relatively higher expressed in the foragers as a result of their slightly lower piRNA expression. This finding indicates that effects on individual TEs need to be treated with care. Therefore, we restricted our analyses to overall patterns and did not investigate individual loci. We conclude that piRNAs are similarly active in the ovaries of

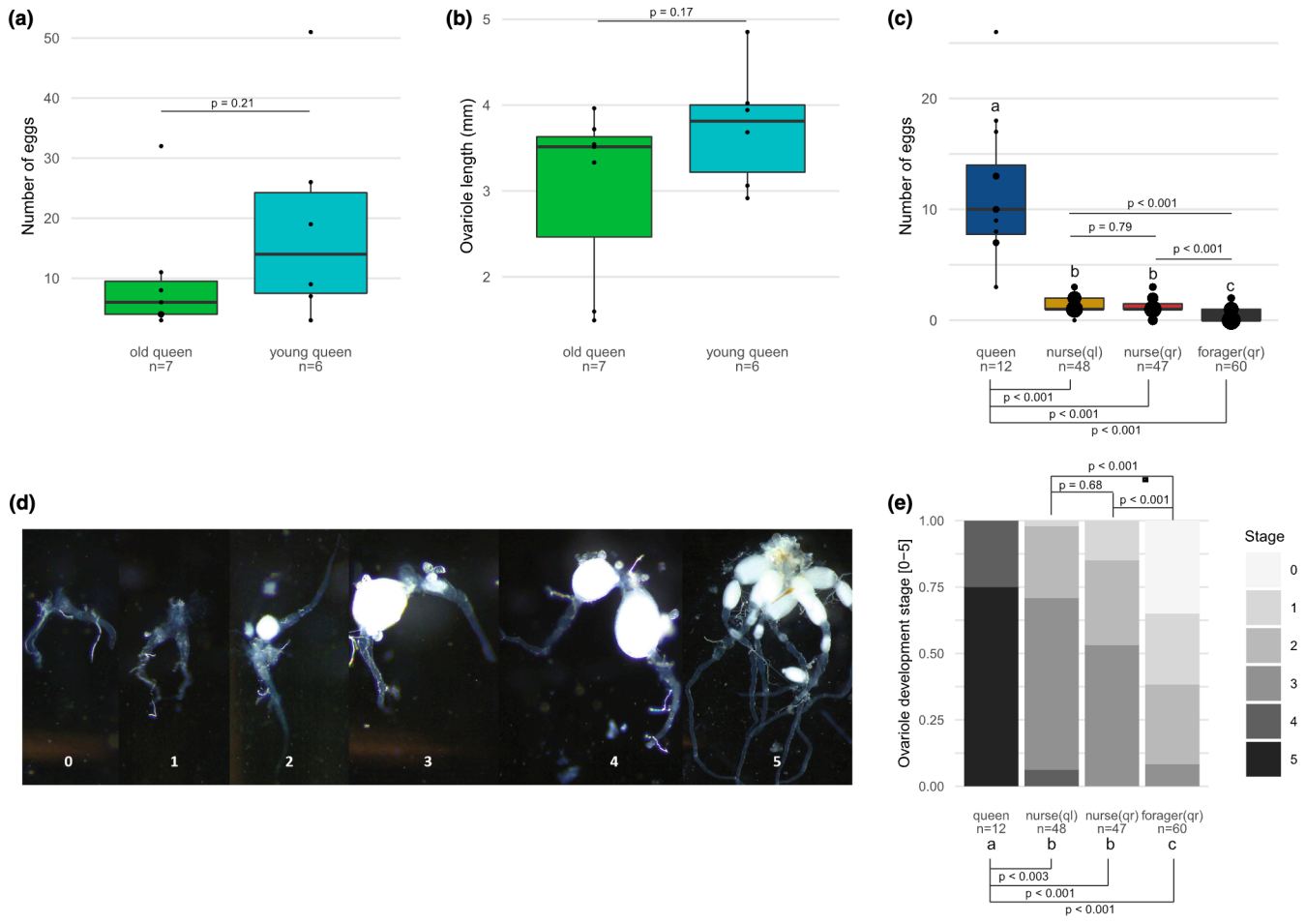


FIGURE 2 Fecundity of *T. rugatulus* queens and workers of different ages and status as determined by ovarian development and egg production. (a) Boxplots of the number of eggs in development found in the ovaries of young, founding queens (<6 months old) and middle-aged established queens (>3.5 years old). *p*-value for general linear model (quasipoisson) indicated. (b) Boxplot of ovariole lengths of young and older queens. Point size is relative to number of replicates. *p*-value for linear model indicated. (c) Boxplots of the number of eggs in development found in queens, nurses and foragers of qr colonies and nurse workers of ql colonies. Point size is relative to number of replicates. *p*-values for generalized linear mixed-effects model (family = poisson+ (1|Colony)) indicated, letters indicate significance. (d) Representative pictures of five defined developmental stages of ovaries. 0 = regressed ovaries: thin, clear tubes; 1 = undeveloped ovaries with rounded tubes and a wider end, 2 = slightly developed ovaries containing immature eggs; 3 = developed ovaries with one mature egg, 4 = well-developed ovaries with two to five mature eggs, 5 = extremely well-developed ovaries with >5 eggs. (e) Boxplot of the distribution of each caste into the five developmental stages as defined in (d). *p*-values for generalized linear mixed-effects model (family = poisson) indicated, letters indicate significance. qr, queenright, ql, ueenless.

all female castes, with only marginal upregulation of already heavily targeted TEs in fully active ovaries. We also infer that piRNAs are unlikely directly involved in the activation of ovaries in nurses upon removal of the queen.

3.7 | *T. rugatulus* miRNAs

We then investigated miRNA-like reads. First, we used miRDeep2 (Friedländer et al., 2012) to carry out a de novo miRNA prediction in our draft genome using miRNAs from four related species (*Bombyx mori*, *Drosophila melanogaster*, *Apis mellifera* and *Tribolium castaneum*) as a reference and the pooled small RNA-seq data of all samples. Thereby, we could detect 1191 possible miRNAs. We then filtered these putative miRNAs using four parameters (see Methods) to ensure high fidelity of our generated miRNA annotation. After filtering, we were left with 372 miRNA loci coding for 304 distinct miRNAs (Table S3). Most of the miRNAs were unique, that is, only found at one locus, and no miRNA was present at more than five loci. Notably, the identified miRNAs included strongly conserved miRNAs such as let-7 and mir-9 known to be ubiquitously expressed in other species.

After having generated our reference, we overlapped this with our mapped reads in the range 18–24 nt from young and older queens. We saw that the reads mainly mapped sense to our defined miRNA loci, as would be expected for miRNAs (Figure 4a). We could further show that *T. rugatulus* miRNAs have a strong 5'-U bias (Figure 4a). A strong bias for 5'-U has also been indicated in other

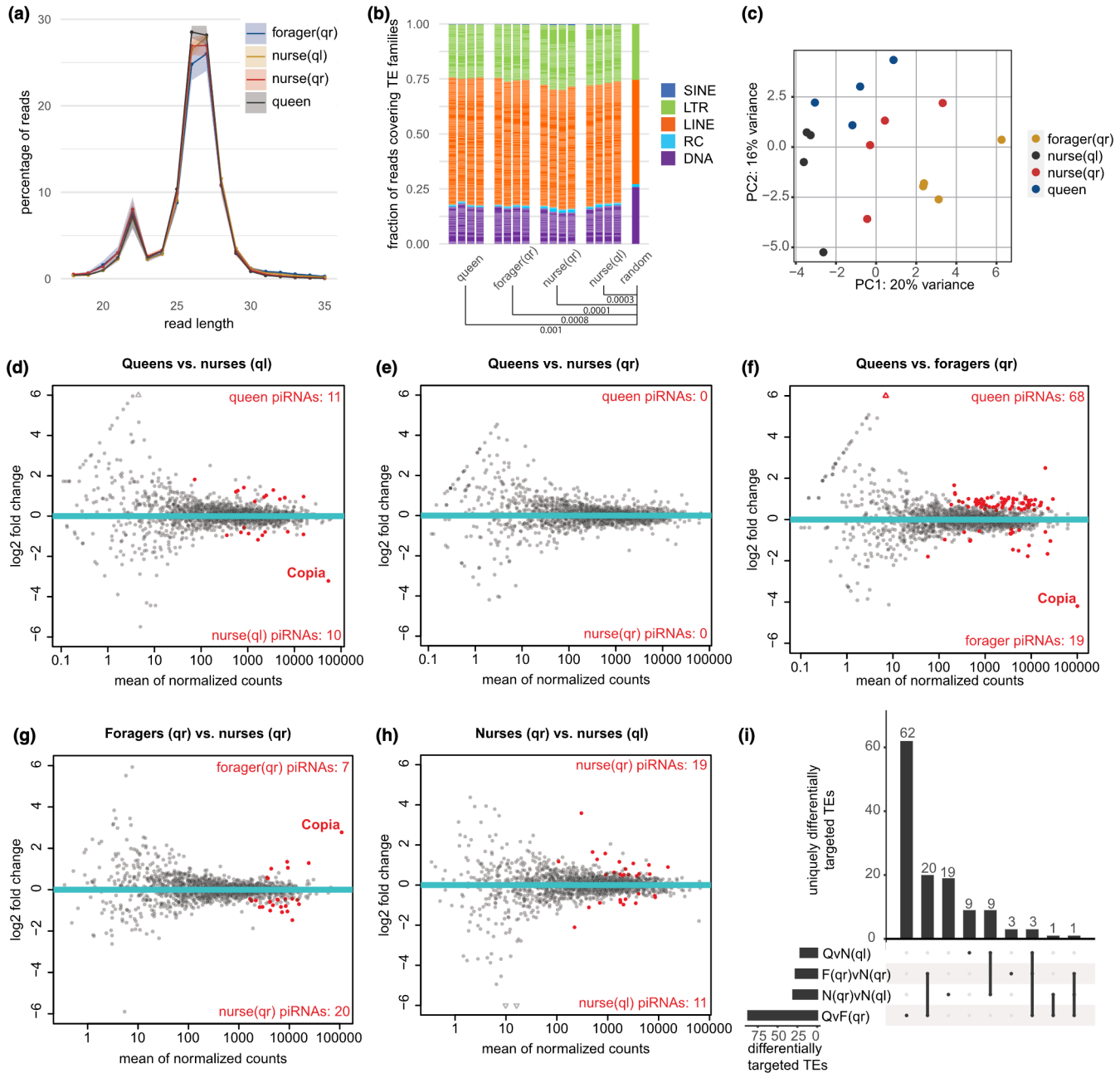


FIGURE 3 Ovarian piRNA expression in different castes. (a) Length distribution of all 18–35 nt ovarian reads generated by sRNA-Seq. Shaded areas depict standard deviation obtained from the four biological replicates. (b) Stacked bar plots showing piRNA-like reads (25–30 nt) attributed to different TE families. *t*-tests carried out for each TE family to find difference between each caste and annotation (random expectation). For each pair of comparisons, the lowest Holm-adjusted *p*-value is shown. All *p*-values are listed in the supplement. (c) Principal component (PC) analysis of piRNA-like reads (25–30 nt) that map to annotated TEs. (d–h) MA plots showing pairwise differential TE comparisons of piRNA targeting in the four castes. Each dot represents one TE type and red indicates significant changes (FDR <0.05). All hits are reported in Table S2. (i) UpSet plot showing overlap of significantly differentially targeted TEs as defined in (d–h). Hits per comparison are shown in the horizontal bar plot, overlaps between comparisons are shown in the vertical bar plot. QvN(ql) = queens versus ql nurses, panel (d). F(qr)vN(qr) = foragers versus qr nurses, panel (g). N(qr)vN(ql) = qr nurses versus ql nurses, panel (h). QvF(qr) = queens versus foragers, panel (f).

species, where this has been seen to aid in AGO recognition and direction into the correct AGO (Seitz et al., 2011). In *D. melanogaster*, for instance, a uridine at the 5′-position will favour the uptake by the Argonaute AGO1 whereas a cytosine at the same position will favour uptake of the miRNA by AGO2 (Czech et al., 2009). In mammals, AGO1 preferentially binds 5′-U and AGO-2 preferentially binds 5′-A (Frank et al., 2010).

Although relative miRNA expression was highest in the brain and thorax, ovaries expressed a more varied selection of miRNAs (Figure 4b). In the ovaries, we could detect the expression of 227

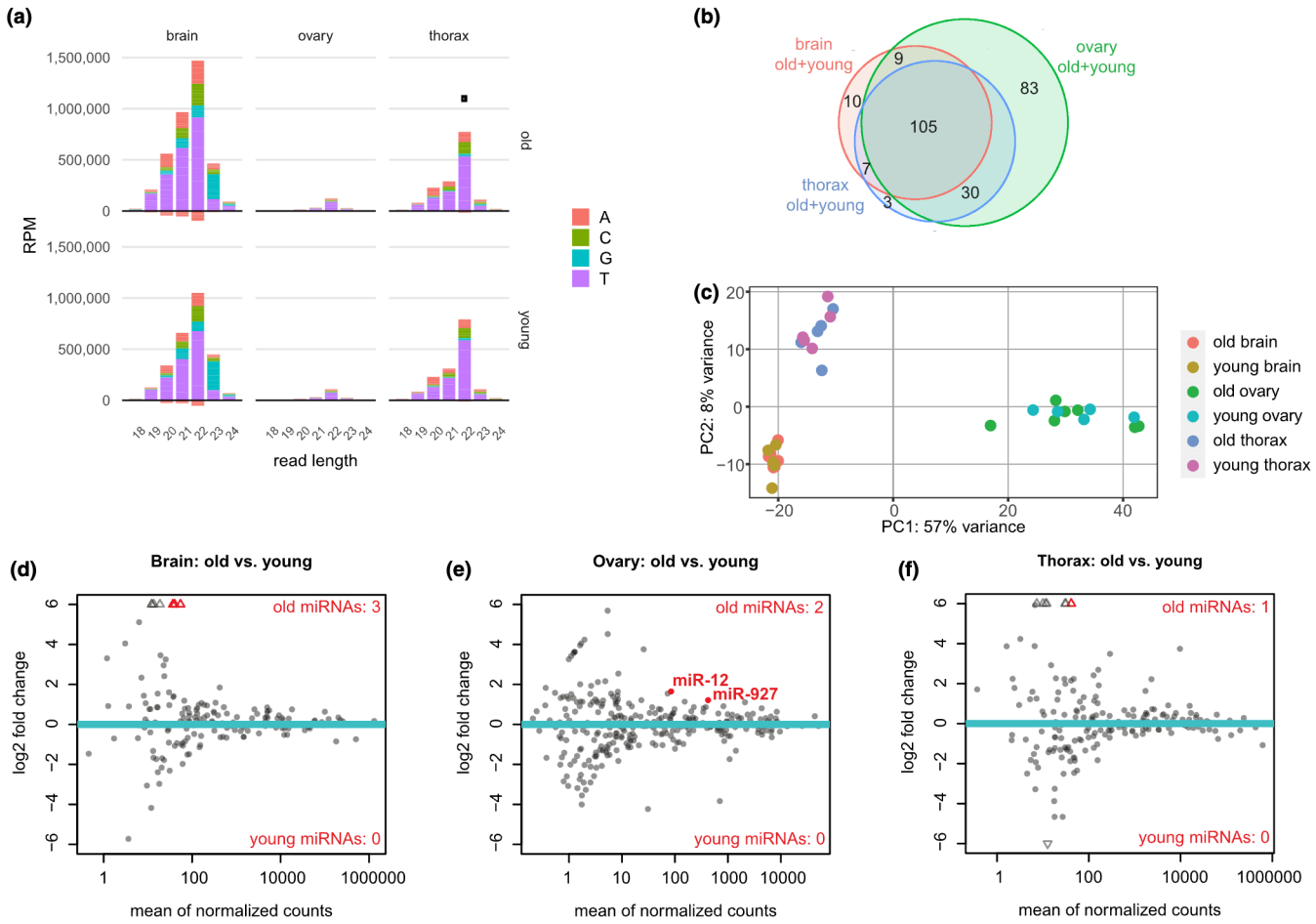


FIGURE 4 miRNA expression in different tissues of young- and middle-aged queens. (a) Length distribution of reads length 18–24 nt mapping sense (up) and antisense (down) to miRNAs. Values represent average of all five to seven replicates, colours represent the starting base. RPM, reads per million. (b) Venn diagram showing the overlap of distinct miRNAs found in at least three samples of brain, ovary and thorax. Old and young queens are pooled together. All miRNAs are listed in Table S4. (c) Principal component analysis of miRNA reads. (d–f) MA plots showing pairwise differential expression comparisons of miRNAs in young versus older queens in the three tissues. Each dot represents one miRNA. Red indicates significant changes (FDR <0.05). All hits are reported in Table S5.

different miRNAs, 83 of which were unique for this tissue. In comparison, only 145 different miRNAs were detected in the thorax and 131 in the brain. We could detect 105 of the expressed miRNAs in all three tissues (Figure 4b). It is worth noting that expression here is defined as detection in at least three replicates of the same tissue, regardless of age, explaining why we find fewer expressed miRNAs (247 as shown in Figure 4b) than found in the genome (304, see above), where detection in three samples of *any* tissue was used as the filtering parameter. Altogether, we conclude that *T. rugatulus* does indeed express miRNAs and that these can be expressed tissue specifically.

3.8 | miRNA expression shifts only slightly with queen age

We next carried out differential analyses in order to detect miRNAs involved in ageing. Similar to the piRNAs, the major difference in miRNA expression was related to tissue type, not age (Figure 4c). In the thorax and brain, we could define one and three miRNAs that were upregulated in the middle-aged queens respectively. All of these were, however, lowly expressed (Figure 4d–f). In the ovaries, we found two miRNAs, miR-12 and miR-927a (both according to *A. mellifera* annotation), to be upregulated in middle-aged queens, although both showed modest change (Figure 4e). In general, miRNAs seemed to be only mildly affected by age in *T. rugatulus* queens.

3.9 | Ovarian miRNA expression is specific to caste and worker task

Finally, we investigated the ovarian miRNAs of queens, nurses from qr and ql colonies and foragers from qr colonies by isolating the reads in the range 18–24 nt and mapping these to our miRNA annotation. Similar to results in queens (Figure 4a), ovarian miRNAs of all groups mapped primarily sense to annotated miRNAs with a 5′-U bias (Figure 5a).

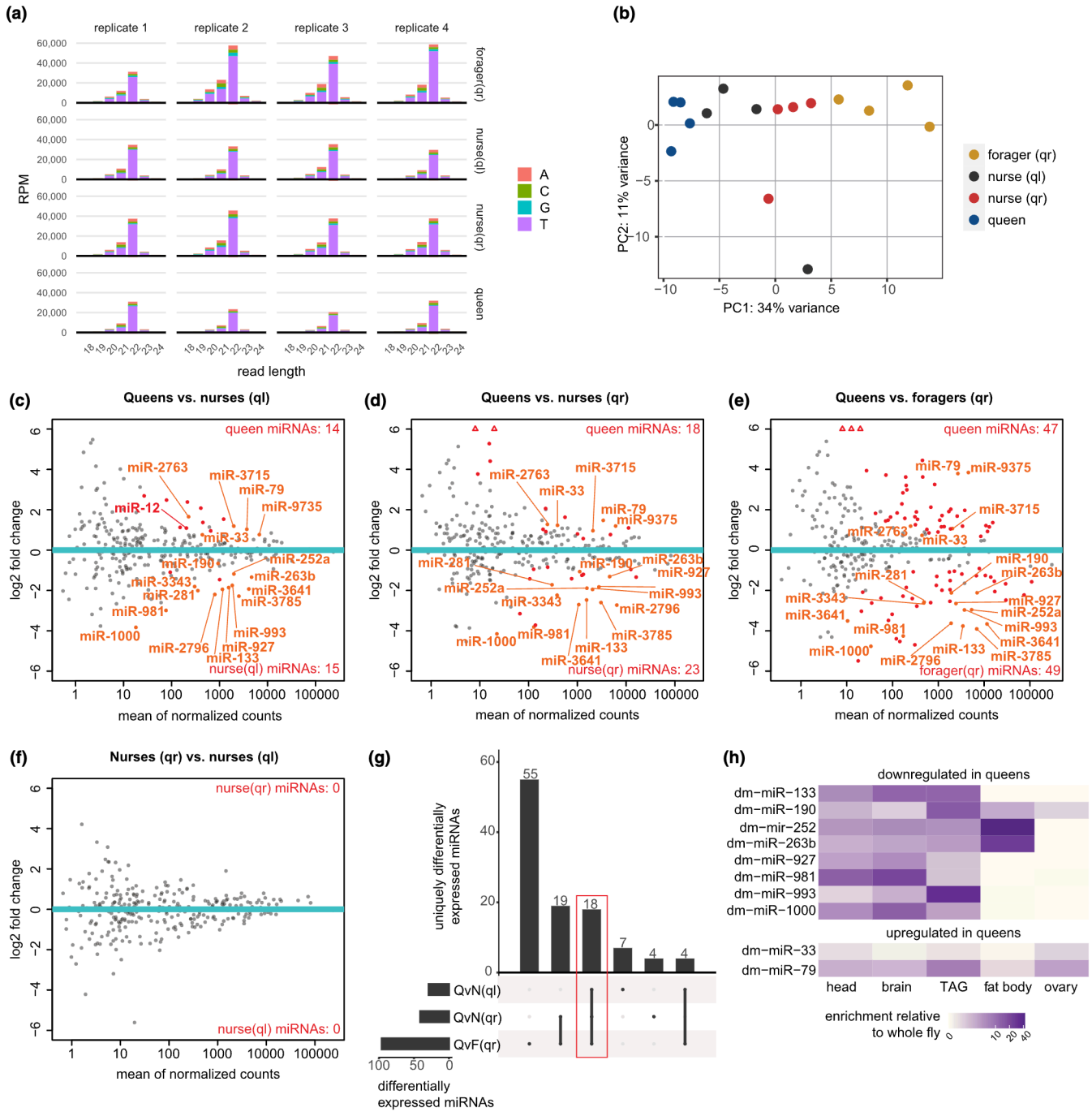


FIGURE 5 Ovarian miRNA expression in different castes. (a) Length distribution of ovarian reads length 18–24 nt mapping to miRNAs on sense (up) and antisense (down) strands. Colours represent the starting base. (b) Principal component (PC) analysis of miRNA reads. (c–f) MA plots showing pairwise differential expression comparisons of miRNAs in different castes. Each dot represents one miRNA. Named miRNAs are those indicated in the orange square in panel (g). Red indicates significant changes (FDR <0.05). All hits are reported in Table S6. (g) UpSet plot showing overlap of significantly differentially expressed miRNAs as defined in (c–f). Hits per comparison are shown in the horizontal bar plot, overlaps between comparisons are shown in the vertical bar plot. QvN(ql) = queens versus ql nurses, panel (c). QvN(qr) = queens versus qr nurses, panel (d). QvF(qr) = queens versus foragers, panel E. (h) Heatmap showing the relative expression of 10 miRNAs in *D. melanogaster* as reported in FlyAtlas2. These 10 miRNAs are those found in all comparisons (c–e) in five tissues (shown in the orange square in panel (g)) for which homologues exist in *D. melanogaster*. TAG, Thoracicoabdominal ganglion.

We next overlapped our mapped miRNAs with the miRNA annotation that we had already generated using our first data set. In stark contrast to piRNAs, which showed little correlation between ovarian expression and group, we observed that miRNA expression and ovarian activation state were highly associated. Principle component analysis revealed gradual shifts from active queen ovaries to activated ovaries of ql nurses, onward to inactive ovaries of qr nurses and the inactive ovaries of foragers (Figure 5b). Despite

this segregation of our samples, we could not identify any specific miRNAs that could be directly involved in the activation of ovaries, as we did not detect any significantly regulated miRNAs between nurses of ql and qr colonies (Figure 5f). However, we could define more than 100 miRNAs that were differentially expressed in different castes across all comparisons (Figure 5c-g, Table S6).

We wanted to investigate a subset of these miRNAs, and therefore decided to search FlyAtlas2 (Krause et al., 2022) for tissue-specific expression in *Drosophila melanogaster* of the 18 miRNAs that we found to be differentially expressed in queens in all three comparisons (queen vs. ql nurses, queen vs. qr nurses and queen vs. foragers). Two of the five miRNAs upregulated in queens and eight of the 13 miRNAs downregulated in queens had homologues with expression data in FlyAtlas2. We asked whether these miRNAs were generally known to be expressed in ovaries. Both of the queen-upregulated miRNAs are present in the *D. melanogaster* ovaries, but most of the queen-downregulated miRNAs did not show ovarian expression in *D. melanogaster* (Figure 5h). We noticed that all of the queen-downregulated miRNAs showed high expression in non-reproductive tissues; head, brain and the thoracicoabdominal ganglion in *D. melanogaster* (Figure 5h). In other tissues such as the fat body, the expression level varied (Figure 5h). This indicates that non-queen ovary samples may carry significant amounts of other tissue(s). This could be due to their small size making them hard to dissect, or to the fact that these ovaries contain a lower fraction of germ cell mass, and hence relatively speaking more somatic tissue, or it could imply that these ovaries simply do express more miRNAs that are usually found in non-ovarian tissues.

We then compared the expression of specific miRNAs to studies on other social insects. We found miR-927a and miR-12 to be upregulated in the ovaries of middle-aged relative to young queens (Figure 4e) and miR-12 in addition upregulated in ovaries of queens relative to ql nurses (Figure 5c). In the honeybee *A. mellifera*, miR-927a and miR-12 were found to be overexpressed in virgin and inactive ovaries (Chen et al., 2017; Macedo et al., 2016). Indeed, we detected quite some overlap in differentially expressed miRNAs between the aforementioned honeybee studies and ours in ant ovaries, including miR-3049, miR-279b, miR-33, miR-190, miR-79 and miR-993. This points to ovarian miRNA expression being linked to ovarian development and female fecundity. Evidence for a conserved function of caste-specific expressed miRNAs comes from miR-133, which was universally downregulated in *T. rugatulus* queen ovaries compared to all worker samples (Figure 5c–e), and showed worker caste-specific expression in the termite *R. speratus* (Matsunami et al., 2018).

4 | DISCUSSION

4.1 | *T. rugatulus* expresses miRNAs and piRNAs

sRNA populations have been described in fungi, plants, invertebrates and mammals (Bartel, 2009; Ozata et al., 2019). In ants, sRNAs have so far only been described in the formicine *Camponotus floridanus* and in the ponerine *Harpegnathos saltator* (Bonasio et al., 2010). Our study on the myrmicine *T. rugatulus* is thus the first on this subfamily of ants. We sampled *T. rugatulus* ants of different ages, castes and worker types and sequenced sRNAs from brain, thorax and ovaries in order to describe the expression landscape of sRNAs across these different groups.

Our analyses confidently show that the ant *T. rugatulus* expresses two distinct sRNA populations coinciding with piRNAs and miRNAs. We also demonstrate that the relative expression of piRNAs in *T. rugatulus* queens is higher in the ovaries than in somatic tissues, which is consistent with the general function of piRNAs in reproductive tissues, where TE surveillance is particularly important (Madhani, 2013) to reduce the risks of mutations in the germline.

4.2 | Only minor changes of piRNA expression exist between different castes of *T. rugatulus* ants

Our results indicate that piRNAs may also play a role in somatic tissues in ants, and particularly in the thorax (Figure 1). This is consistent with a previous report describing piRNAs in somatic tissues in several ancestral arthropod species (Lewis et al., 2018). We found divergent biases and different targets of piRNAs in queen brains, pointing to possibly alternate functions of these sRNAs in this tissue. In the termite *Macrotermes bellicosus*, TEs were found to be expressed in the heads of old workers but not in those of queens, which are much longer lived (Elsner et al., 2018). TEs were also much more active in the fat body of workers of the termite *M. natalensis* than in reproductive individuals (Post et al., 2023), leading to the prediction that pathway targeting TEs might be less active in non-reproductive individuals than in reproductive individuals. We demonstrated that the non-reproductive and regressed ovaries of nurses and foragers from queenright colonies not only have piRNAs at similar levels to those of queens, but also show ping-pong signatures and target TEs, all of which are features of an active TE-silencing piRNA pathway. We find that piRNAs are only slightly downregulated in foragers compared to queens (Figure 3a,f), but not nearly to the extent expected, as ovaries of foragers had regressed to the point where they were only translucent membranes. We therefore conclude that in *T. rugatulus*, reproductive and non-reproductive ovaries alike seem to have the capacity to silence TEs via piRNAs and this pathway is likely not linked to ovarian development. Rather, the high activity of piRNAs is a characteristic of ovarian tissue per se maintained even in old workers, despite them sometimes being regarded as the soma of the superorganism ant colony (Johnson & Linksvayer, 2010).

We detected 7–16 differentially targeted TEs between young- and middle-aged queens (Figure 1e–g). These had either a low sRNA levels overall, as in the brain and thorax, or were strongly expressed but only weakly differentially regulated, as in the ovaries, the tissue that showed the most pronounced differences. The weak differences could indicate that piRNAs and miRNAs—we also found only minor differences in those regulatory RNAs—do

not play a major role in physiological ageing in this species. The minor shifts in sRNAs could also be due to slow ageing in *T. rugatulus* queens. Although *Temnothorax* queens can live up to 20 years (Plateaux, 1986), our study compared queens with an age of half a year and more than 4 years and probably did not include senescent queens at the end of their life. Moreover, the young- and middle-aged queens in our study showed no differences in ovarian development, which distinguished them from those in a previous gene expression study in which the young queens had significantly less developed ovaries and also showed strong transcriptional differences in the fat body (Negroni et al., 2019). In the ant species *Cardiocondyla obscurior*, there were also only minor expression changes between young- and middle-aged queens, which indicated a stable physiology over a long period of time (von Wyschetzki et al., 2015), but dramatic changes at the very end of life indicative of rapid physiological degeneration (Jaimes-Nino et al., 2022). Therefore, it is conceivable that the minimal change in sRNAs with queen age in our study indicates slowed ageing in the young- to middle-aged queens of our model *T. rugatulus*.

4.3 | miRNA expression in relation to ageing and ovarian activation state

We detected 304 miRNAs in *T. rugatulus*, which is about double to triple the number annotated in the genomes of *H. saltator* (159) and in *C. floridanus* (96) (Bonasio et al., 2010). Nevertheless, many of our miRNAs are clear homologues of known miRNAs. Some of them are broadly expressed, but we also detected tissue-specific expression. We found only six miRNAs that were differentially expressed between young- and middle-aged queens in the tissues analysed (Figure 4d–f), and these either differed only slightly in expression or were expressed at very low levels. One miRNA upregulated in the ovaries of older queens, miR-12, was also found to be upregulated in active ovaries of queens relative to nurses from queenless colonies (Figure 5c). Interestingly, in *A. mellifera* (Macedo et al., 2016), miR-12 is predominantly expressed in inactive ovaries of virgin queens, suggesting that miR-12 and possibly also miR-927a have a conserved role linked to ovarian activity, but may have different target genes in *A. mellifera* compared to *T. rugatulus*. Overall, our data suggest that miRNAs are rather stably expressed in queens over many years in *T. rugatulus*, suggesting little changes in molecular physiology with age. In contrast, we discovered very large differences in miRNA populations between the ovaries of queens and workers. Indeed, 96 miRNAs were different in expression between queens and foragers (Figure 5c–e). The latter represent the oldest worker generation (Kohlmeier et al., 2019) with most regressed ovaries. In contrast, only a quarter to half as many miRNAs were differentially between queens and nurses from queenless (25) and queenright (41) colonies. These younger workers mostly exhibit extended ovaries with developing eggs. Also, the principal component analysis based on all miRNAs does not show a dichotomy between queens and workers, but rather a gradual shift with queen samples at one end and forager samples at the other end and nurses in the middle, suggesting a possible correlation between miRNA expression profiles and ovarian activation (Figure 5b). Several miRNAs were differentially expressed in individual comparisons between the reproductive ovaries of queens, and less activated ovaries of nurses or foragers. We found no miRNAs to be differentially expressed between nurses from either queenright or queenless colonies (Figure 5f), which might be due to the fact that the number of eggs and ovarian development between these two groups was insignificant in our study (Figure 2c,e), albeit earlier studies showing that queen removal can cause changes in fat body gene expression linked to reproductive state (Choppin et al., 2021; Negroni et al., 2020, 2021). As more social insects have their miRNA profiles mapped, it will in the future become possible to do a more in-depth comparison and determine conserved versus species-specific caste-associated miRNAs and possibly to elucidate the evolutionary mechanisms behind the roles of specific sRNAs in the complex lifestyles of social insects. Based on our findings, we hypothesise that miRNAs in the ovaries play an important role in regulating fertility in ants and do not solely reflect caste status.

AUTHOR CONTRIBUTIONS

Susanne Foitzik and René Ketting designed the study. Marina Choppin, Barbara Feldmeyer and Susanne Foitzik collected the ant colonies. Marina Choppin, Marion Kever and Susanne Foitzik dissected the ants, measured the ovaries and prepared the samples. The TE annotation was done by Alice Séguret. The RNA was extracted by Marion Kever, Ann-Sophie Seistrup and Shamitha Govind, the latter two also performing all other molecular biology work in the lab. The miRNA annotation and bioinformatic analyses were performed by Emil Karaulanov, Sivarajan Karunanithi, Miguel V. Almeida, Ann-Sophie Seistrup and Shamitha Govind. Data analysis and manuscript preparation were carried out by Ann-Sophie Seistrup, René Ketting and Susanne Foitzik with the support of all co-authors.

ACKNOWLEDGEMENTS

This project was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—GRK2526/1—Projectnr. 407023052 and FO 298/19–2 within the research unit FOR-2281. Support by the IMB Genomics Core Facility and the use of its NextSeq500 (funded by DFG—INST 247/870-1 FUGG) is gratefully acknowledged. Open Access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST STATEMENT

The authors report no conflict of interest.

DATA AVAILABILITY STATEMENT

Raw sequence data were deposited in the NCBI Sequence Read Archive (SRA) under BioProject PRJNA955004. Draft genome was provided by Jongepier et al. (2022), available under BioProject PRJNA750352 and TE annotation can be found in supplementary data (TE_annotation_trug.fa.out.gz) respectively.

ORCID

Barbara Feldmeyer https://orcid.org/0000-0002-0413-7245
Susanne Foitzik https://orcid.org/0000-0001-8161-6306

REFERENCES

Andrews, S. (2010). FASTQC. A quality control tool for high throughput sequence data.

Bartel, D. P. (2009). MicroRNAs: Target recognition and regulatory functions. *Cell*, *136*(2), 215–233. https://doi.org/10.1016/j.cell.2009.01.002

Behura, S. K., & Whitfield, C. W. (2010). Correlated expression patterns of microRNA genes with age-dependent behavioural changes in honeybee. *Insect Molecular Biology*, *19*(4), 431–439. https://doi.org/10.1111/j.1365-2583.2010.01010.x

Bohn, J., Halabian, R., Schrader, L., Shabardina, V., Steffen, R., Suzuki, Y., Ernst, U. R., Gadau, J., & Makałowski, W. (2021). Genome assembly and annotation of the California harvester ant *Pogonomyrmex californicus*. *G3 Genes|Genomes|Genetics*, *11*(1), jkaa019. https://doi.org/10.1093/g3journal/jkaa019

Bonasio, R., Zhang, G., Ye, C., Mutti, N. S., Fang, X., Qin, N., Donahue, G., Yang, P., Li, Q., Li, C., Zhang, P., Huang, Z., Berger, S. L., Reinberg, D., Wang, J., & Liebig, J. (2010). Genomic comparison of the ants *Camponotus floridanus* and *Harpegnathos saltator*. *Science*, *329*(5995), 1068–1071. https://doi.org/10.1126/science.1192428

Boomsma, J. J., & Gawne, R. (2018). Superorganismality and caste differentiation as points of no return: How the major evolutionary transitions were lost in translation. *Biological Reviews*, *93*(1), 28–54. https://doi.org/10.1111/brv.12330

Bourke, A. F. G. (1988). Worker reproduction in the higher eusocial hymenoptera. *The Quarterly Review of Biology*, *63*(3), 291–311. https://doi.org/10.1086/415930

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

Chen, H., & Boutros, P. C. (2011). VennDiagram: A package for the generation of highly-customizable Venn and Euler diagrams in R. *BMC Bioinformatics*, *12*(1), 35. https://doi.org/10.1186/1471-2105-12-35

Chen, X., Ma, C., Chen, C., Lu, Q., Shi, W., Liu, Z., Wang, H., & Guo, H. (2017). Integration of lncRNA–miRNA–mRNA reveals novel insights into oviposition regulation in honey bees. *PeerJ*, *5*, e3881. https://doi.org/10.7717/peerj.3881

Choppin, M., Feldmeyer, B., & Foitzik, S. (2021). Histone acetylation regulates the expression of genes involved in worker reproduction in the ant Temnothorax rugatulus. *BMC Genomics*, *22*, 871. https://doi.org/10.1186/s12864-021-08196-8

Coenen-Stass, A. M. L., Sork, H., Gatto, S., Godfrey, C., Bhomra, A., Krjutškov, K., Hart, J. R., Westholm, J. O., O'Donovan, L., Roos, A., Lochmüller, H., Puri, P. L., El Andaloussi, S., Wood, M. J. A., & Roberts, T. C. (2018). Comprehensive RNA-sequencing analysis in serum and muscle reveals novel small RNA signatures with biomarker potential for DMD. *Molecular Therapy - Nucleic Acids*, *13*, 1–15. https://doi.org/10.1016/j.omtn.2018.08.005

Conway, J. R., Lex, A., & Gehlenborg, N. (2017). UpSetR: An R package for the visualization of intersecting sets and their properties. *Bioinformatics*, *33*(18), 2938–2940. https://doi.org/10.1093/bioinformatics/btx364

Corona, M., Libbrecht, R., & Wheeler, D. E. (2016). Molecular mechanisms of phenotypic plasticity in social insects. *Current Opinion in Insect Science*, *13*, 55–60. https://doi.org/10.1016/j.cois.2015.12.003

Czech, B., Zhou, R., Erlich, Y., Brennecke, J., Binari, R., Villalta, C., Gordon, A., Perrimon, N., & Hannon, G. J. (2009). Hierarchical rules for Argonaute loading in drosophila. *Molecular Cell*, *36*(3), 445–456. https://doi.org/10.1016/j.molcel.2009.09.028

Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M., & Li, H. (2021). Twelve years of SAMtools and BCFtools. *GigaScience*, *10*(2), giab008. https://doi.org/10.1093/gigascience/giab008

Elsik, C. G., Worley, K. C., Bennett, A. K., et al. (2014). Finding the missing honey bee genes: lessons learned from a genome upgrade. *BMC Genomics*, *15*, 86. https://doi.org/10.1186/1471-2164-15-86

Elsner, D., Meusemann, K., & Korb, J. (2018). Longevity and transposon defense, the case of termite reproductives. *Proceedings of the National Academy of Sciences*, *115*(21), 5504–5509. https://doi.org/10.1073/pnas.1804046115

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

Faulk, C. (2023). De novo sequencing, diploid assembly, and annotation of the black carpenter ant, *Camponotus pennsylvanicus*, and its symbionts by one person for $1000, using nanopore sequencing. *Nucleic Acids Research*, *51*(1), 17–28. https://doi.org/10.1093/nar/gkac510

Filipowicz, W., Bhattacharyya, S. N., & Sonenberg, N. (2008). Mechanisms of post-transcriptional regulation by microRNAs: Are the answers in sight? *Nature Reviews. Genetics*, *9*(2), 102–114. https://doi.org/10.1038/nrg2290

Flatt, T. (2011). Survival costs of reproduction in *drosophila*. *Experimental Gerontology*, *46*(5), 369–375. https://doi.org/10.1016/j.exger.2010.10.008

Frank, F., Sonenberg, N., & Nagar, B. (2010). Structural basis for 5′-nucleotide base-specific recognition of guide RNA by human AGO2. *Nature*, *465*(7299), 818–822. https://doi.org/10.1038/nature09039

Friedländer, M. R., Mackowiak, S. D., Li, N., Chen, W., & Rajewsky, N. (2012). miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. *Nucleic Acids Research*, *40*(1), 37–52. https://doi.org/10.1093/nar/gkr688

Gilbert, C., Peccoud, J., & Cordaux, R. (2021). Transposable Elements and the Evolution of Insects. *Annual Review of Entomology*, *66*(1), 355–372.

Griffiths-Jones, S., Bateman, A., Marshall, M., Khanna, A., & Eddy, S. R. (2003). Rfam: An RNA family database. *Nucleic Acids Research*, *31*(1), 439–441. https://doi.org/10.1093/nar/gkg006

Heinze, J., & Schrempf, A. (2012). Terminal investment: Individual reproduction of ant queens increases with age. *PLoS One*, *7*(4), e35201. https://doi.org/10.1371/journal.pone.0035201

Hirano, T., Iwasaki, Y. W., Lin, Z. Y., Imamura, M., Seki, N. M., Sasaki, E., Saito, K., Okano, H., Siomi, M. C., & Siomi, H. (2014). Small RNA profiling and characterization of piRNA clusters in the adult testes of the common marmoset, a model primate. *RNA*, *20*(8), 1223–1237. https://doi.org/10.1261/rna.045310.114

Jaimes-Nino, L. M., Heinze, J., & Oettler, J. (2022). Late-life fitness gains and reproductive death in Cardiocondyla obscurior ants. *eLife*, *11*, e74695. https://doi.org/10.7554/eLife.74695

Johnson, B. R., & Linksvayer, T. A. (2010). Deconstructing the superorganism: Social physiology, Groundplans, and Sociogenomics. *The Quarterly Review of Biology*, *85*(1), 57–79. https://doi.org/10.1086/650290

Jongepier, E., Séguret, A., Labutin, A., Feldmeyer, B., Gstöttl, C., Foitzik, S., Heinze, J., & Bornberg-Bauer, E. (2022). Convergent loss of chemoreceptors across independent origins of slave-making in ants. *Molecular Biology and Evolution*, *39*(1), msab305. https://doi.org/10.1093/molbev/msab305

Kapheim, K. M., Jones, B. M., Søvik, E., Stolle, E., Waterhouse, R. M., Bloch, G., & Ben-Shahar, Y. (2020). Brain microRNAs among social and solitary bees. *Royal Society Open Science*, *7*(7), 200517. https://doi.org/10.1098/rsos.200517

Keller, L., & Genoud, M. (1997). Extraordinary lifespans in ants: A test of evolutionary theories of ageing. *Nature*, *389*(6654), 958–960. https://doi.org/10.1038/40130

Ketting, R. F., & Cochella, L. (2021). Concepts and functions of small RNA pathways in C. Elegans. *Current Topics in Developmental Biology*, *144*, 45–89. https://doi.org/10.1016/bs.ctdb.2020.08.002

Kinser, H. E., & Pincus, Z. (2020). MicroRNAs as modulators of longevity and the aging process. *Human Genetics*, *139*(3), 291–308. https://doi.org/10.1007/s00439-019-02046-0

Kohlmeier, P., Alleman, A. R., Libbrecht, R., Foitzik, S., & Feldmeyer, B. (2019). Gene expression is more strongly associated with behavioural specialization than with age or fertility in ant workers. *Molecular Ecology*, *28*(3), 658–670. https://doi.org/10.1111/mec.14971

Krause, S. A., Overend, G., Dow, J. A. T., & Leader, D. P. (2022). FlyAtlas 2 in 2022: Enhancements to the *Drosophila melanogaster* expression atlas. *Nucleic Acids Research*, *50*(D1), D1010–D1015. https://doi.org/10.1093/nar/gkab971

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

Lewis, B. P., Shih, I. H., Jones-Rhoades, M. W., Bartel, D. P., & Burge, C. B. (2003). Prediction of mammalian microRNA targets. *Cell*, *115*(7), 787–798. https://doi.org/10.1016/s0092-8674(03)01018-3

Lewis, S. H., Quarles, K. A., Yang, Y., Tanguy, M., Frézal, L., Smith, S. A., Shrma, P. P., Cordaux, R., Gilbert, C., Giraud, I., Collins, D. H., Zamore, P. D., Miska, E. A., Sarkies, P., & 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

Liao, Y., Smyth, G. K., & Shi, W. (2013). The subread aligner: Fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Research*, *41*(10), e108. https://doi.org/10.1093/nar/gkt214

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

Ludwig, N., Becker, M., Schumann, T., Speer, T., Fehlmann, T., Keller, A., & Meese, E. (2017). Bias in recent miRBase annotations potentially associated with RNA quality issues. *Scientific Reports*, *7*(1), 5162. https://doi.org/10.1038/s41598-017-05070-0

Luo, S., & Lu, J. (2017). Silencing of transposable elements by piRNAs in *drosophila*: An evolutionary perspective. *Genomics, Proteomics & Bioinformatics*, *15*, 164–176. https://doi.org/10.1016/j.gpb.2017.01.006

Macedo, L. M., Nunes, F. M., Freitas, F. C., Pires, C. V., Tanaka, E. D., Martins, J. R., Plulachs, M. D., Cristino, A. S., Pinheiro, D. G., & Simões, Z. L. (2016). MicroRNA signatures characterizing caste-independent ovarian activity in queen and worker honeybees (*Apis mellifera* L.). *Insect Molecular Biology*, *25*(3), 216–226. https://doi.org/10.1111/imb.12214

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

Maklakov, A. A., & Chapman, T. (2019). Evolution of ageing as a tangle of trade-offs: Energy versus function. *Proceedings of the Royal Society B: Biological Sciences*, *286*(1911), 20191604. https://doi.org/10.1098/rspb.2019.1604

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet Journal*, *17*(1), 3. doi:10.14806/ej.17.1.200

Matsunami, M., Nozawa, M., Suzuki, R., Toga, K., Masuoka, Y., Yamaguchi, K., Maekawa, K., Shigenobu, S., & Miura, T. (2018). Caste-specific microRNA expression in termites: Insights into soldier differentiation. *Insect Molecular Biology*, *28*(1), 86–98. https://doi.org/10.1111/imb.12530

Negroni, M. A., Foitzik, S., & Feldmeyer, B. (2019). Long-lived *Temnothorax* ant queens switch from investment in immunity to antioxidant production with age. *Scientific Reports*, *9*(1), 7270. https://doi.org/10.1038/s41598-019-43796-1

Negroni, M. A., Macit, M. N., Stoldt, M., Feldmeyer, B., & Foitzik, S. (2021). Molecular regulation of lifespan extension in fertile ant workers. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *376*(1823), 20190736. https://doi.org/10.1098/rstb.2019.0736

Negroni, M. A., Segers, F. H. I. D., Vogelweith, F., & Foitzik, S. (2020). Immune challenge reduces gut microbial diversity and triggers fertility-dependent gene expression changes in a social insect. *BMC Genomics*, *21*(1), 816. https://doi.org/10.1186/s12864-020-07191-9

Nouhaud, P., Beresford, J., & Kulmuni, J. (2022). Assembly of a hybrid *Formica aquilonia* × *F. Polyctena* ant genome from a haploid male. *Journal of Heredity*, *113*(3), 353–359. https://doi.org/10.1093/jhered/esac019

Ozata, D. M., Gainetdinov, I., Zoch, A., O'Carroll, D., & Zamore, P. D. (2019). PIWI-interacting RNAs: Small RNAs with big functions. *Nature Reviews. Genetics*, *20*(2), 89–108. https://doi.org/10.1038/s41576-018-0073-3

Plateaux, L. (1986). Comparaison des cycles saisonniers, des durées des sociétés et des productions des trois espèces de fourmis Leptothorax (Myrafant) du groupe nylanderi. *Actes Des Colloques Insectes Sociaux*, *3*, 221–234.

Post, F., Bornberg-Bauer, E., Vasseur-Cognet, M., & Harrison, M. C. (2023). More effective transposon regulation in fertile, long-lived termite queens than in sterile workers. *Molecular Ecology*, *32*(2), 369–380. https://doi.org/10.1111/mec.16753

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

R Core Team. (2013). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. http://www.R-project.org/

Ramírez, F., Ryan, D. P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A. S., Heyne, S., Dündar, F., & 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

Seitz, H., Tushir, J. S., & Zamore, P. D. (2011). A 5′-uridine amplifies miRNA/miRNA* asymmetry in *drosophila* by promoting RNA-induced silencing complex formation. *Silence*, *2*, 4. https://doi.org/10.1186/1758-907x-2-4

Smit, A. F. A., Hubley, R., & Green, P. (2013–2015). RepeatMasker Open-4.0. http://www.repeatmasker.org

Sturm, Á., Perczel, A., Ivics, Z., & Vellai, T. (2017). The Piwi-piRNA pathway: Road to immortality. *Aging Cell*, *16*(5), 906–911. https://doi.org/10.1111/acel.12630

Szathmáry, E., & Smith, J. M. (1995). The major evolutionary transitions. *Nature*, *374*(6519), 227–232. https://doi.org/10.1038/374227a0

Uhrig, S., & Klein, H. (2019). PingPongPro: A tool for the detection of piRNA-mediated transposon-silencing in small RNA-Seq data. *Bioinformatics*, *35*(2), 335–336. https://doi.org/10.1093/bioinformatics/bty578

von Wyschetzki, K., Rueppell, O., Oettler, J., & Heinze, J. (2015). Transcriptomic Signatures Mirror the Lack of the Fecundity/Longevity Trade-Off in Ant Queens. *Molecular Biology and Evolution*, *32*(12), 3173–3185. https://doi.org/10.1093/molbev/msv186

Wang, W., Ashby, R., Ying, H., Maleszka, R., & Forêt, S. (2017). Contrasting sex-and caste-dependent piRNA profiles in the transposon depleted Haplodiploid honeybee *Apis mellifera*. *Genome Biology and Evolution*, *9*(5), 1341–1356. https://doi.org/10.1093/gbe/evx087

Wheeler, W. M. (1911). The ant-colony as an organism. *Journal of Morphology*, *22*(2), 307–325. https://doi.org/10.1002/jmor.1050220206

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag.

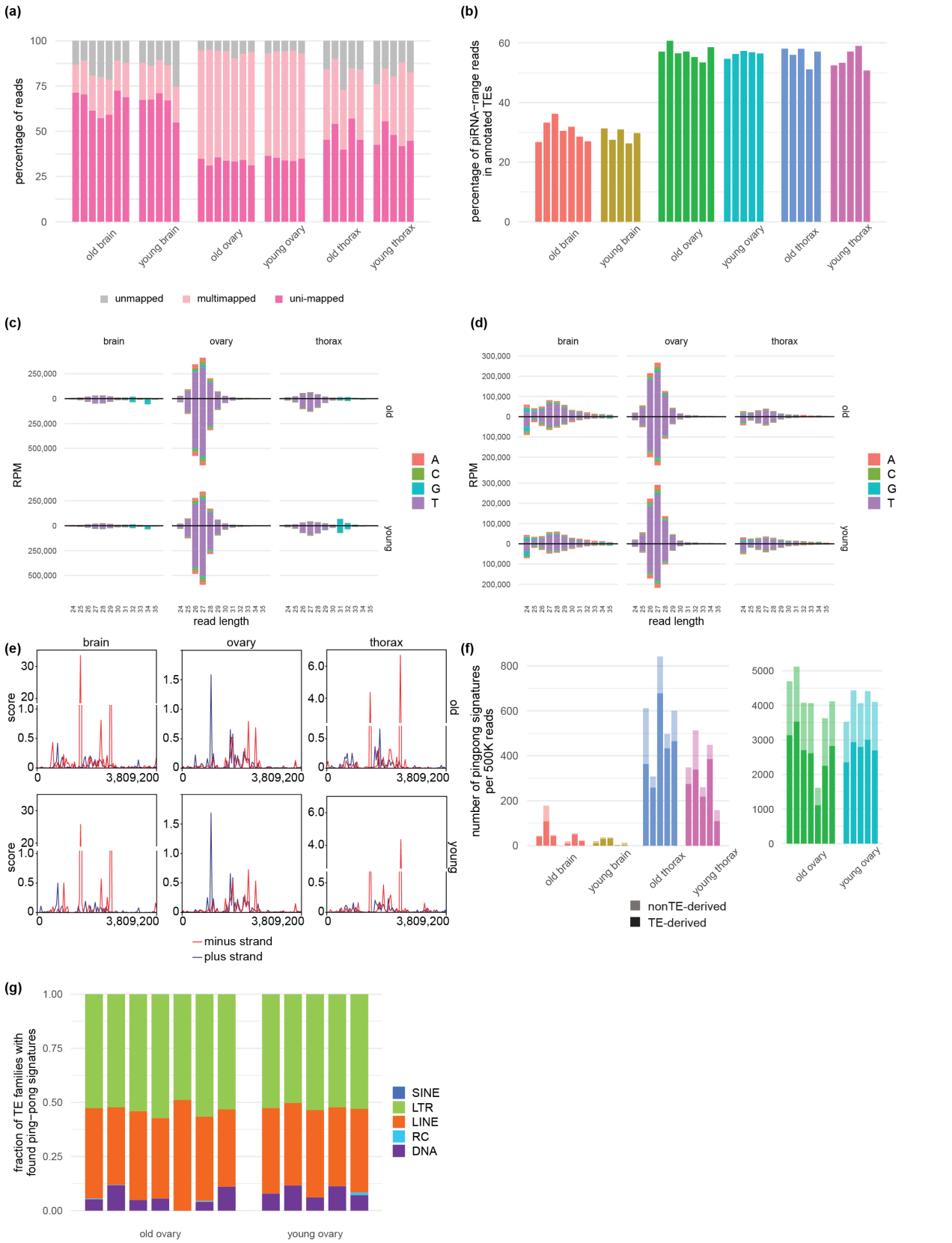
Zerbino, D. R., Johnson, N., Juettemann, T., Wilder, S. P., & Flicek, P. (2014). WiggleTools: Parallel processing of large collections of genome-wide datasets for visualization and statistical analysis. *Bioinformatics*, *30*(7), 1008–1009. https://doi.org/10.1093/bioinformatics/btt737

Zhang, J., Chen, S., & Liu, K. (2022). Structural insights into piRNA biogenesis. *Biochim Biophys Acta Gene Regul Mech*, *1865*(2), 194799. https://doi.org/10.1016/j.bbagrm.2022.194799

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Seistrup, A.-S., Choppin, M., Govind, S., Feldmeyer, B., Kever, M., Karaulanov, E., Séguret, A., Karunanithi, S., Almeida, M. V., Ketting, R. F., & Foitzik, S. (2023). Age- and caste-independent piRNAs in the germline and miRNA profiles linked to caste and fecundity in the ant *Temnothorax rugatulus*. *Molecular Ecology*, *32*, 6027–6043. https://doi.org/10.1111/mec.17162



Legend on next page

FIGURE S1 piRNA characteristics of sRNA reads in different tissues of queens. (a) Stacked bar plot showing the percentage of all 18-40nt reads from each sample that either mapped uniquely (uni-mapped), at multiple locations (multi-mapped), or not at all (unmapped) to our draft genome. (b) Bar plot showing the percentage of piRNA-range reads (25-30nt) found within annotated TEs. (c) Length distribution of 24-35nt reads mapping sense (up) and antisense (down) to TEs. Colours represent the starting base. Reads were RPM normalized. (d) Length distribution of 24-35nt reads that do not map to annotated TEs. Reads are separated into plus strand (up) and minus strand (down). Colours represent the starting base. Reads were RPM normalized. (e) Coverage plot of piRNA-range reads mapping outside of TEs across the longest contig tig0031 which is 3.8Mbp in length, as indicated on the x-axis. Y-axis depicts the average read coverage score of the 5-7 replicates on plus (blue) and minus (red) strands as calculated in bins of 10bp. (f) Bar plot showing the normalised number of ping-pong signatures using piRNA-range reads found in each sample either within or outside of annotated TEs, as indicated below. Each replicate is depicted by a separate bar and replicates of the same age and tissue are shown in the same colour. T-tests old versus young: brain p = 0.23; thorax p = 0.07; ovary p = 0.66. (g) Stacked bar plot showing the fraction of TE-derived ping-pong pairs matching to each of the indicated TE families. T-tests old versus young: SINE p = 1; LTR p = 1; LINE p = 1; RC p = 1; DNA p = 1. T-tests old and young versus annotation: SINE p = 0.0; LTR p = 7.1e-13; LINE p = 4.3e-5; RC p = 3.8e-7; DNA p = 6.0e-9.

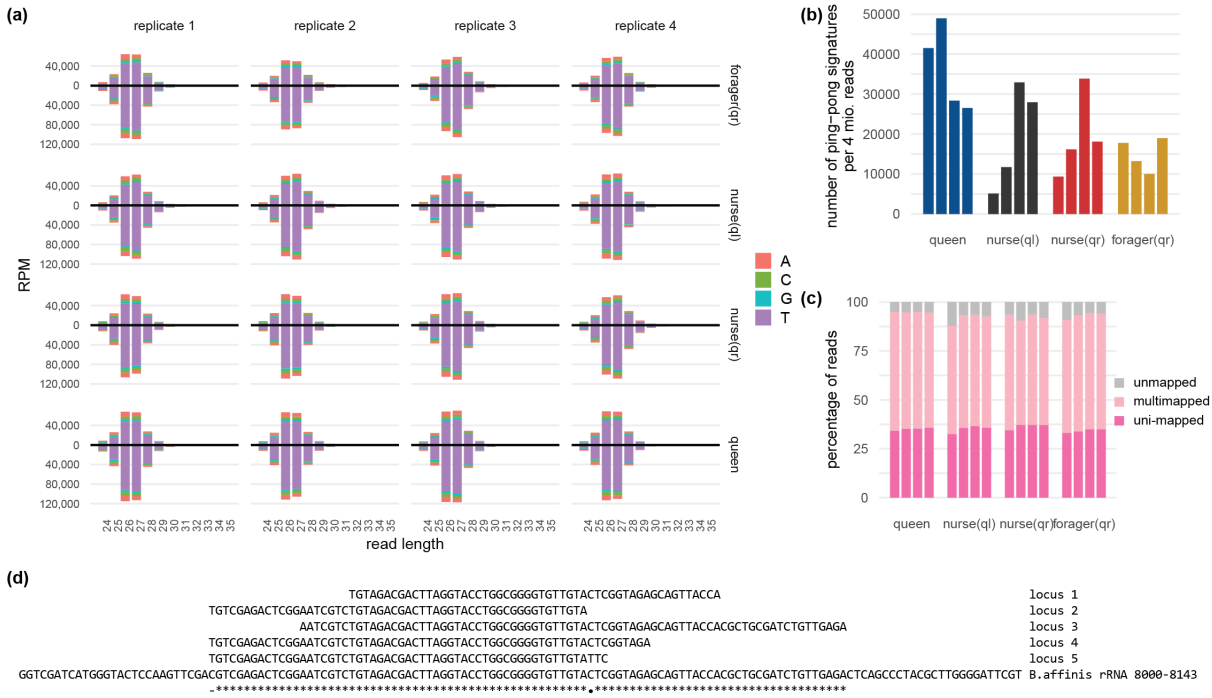


FIGURE S2 Ovarian piRNA characteristics of sRNA reads in different castes. (a) Length distribution of ovarian reads length 24-35nt mapping sense (up) and antisense (down) to TEs. Colours represent the starting base. (b) Bar plot showing the normalized number of ping-pong signatures found in each sample using piRNA-range reads (25-30nt). Replicates of the same caste are shown in the same colour. (c) Stacked bar plot showing the percentage of reads from each sample that either mapped uniquely (uni-mapped), at multiple locations (multimapped), or not at all (unmapped) to our draft genome. (d) Alignment of the sequences where the highest five peaks in Copia loci were found and the *B. affinis* large rRNA C-terminus, position 8000-8143. * 100% similar; • > 80% similar; - similar for loci, different from rRNA.

Chapter III

Deletion of worm-specific Argonautes expressed in the germline of *Caenorhabditis elegans* causes unpredicted changes in mRNA levels that can be remembered upon Argonaute reintroduction

Page left intentionally blank

Aims

This chapter is focused on the three germline expressed WAGOs, WAGO-1, -3, and -4 in *C. elegans*. We wanted to understand the specific role of each of these WAGOs and whether they show mutual redundancy. We wanted to investigate this in different lifestages as a proxy for male versus female inheritance, although we note that investigation of obligate males and females and/or of sperm and oocytes would be necessary to properly ascertain paternal versus maternal inheritance phenomena.

We chose three different lifestages to look at; L4 larvae, where the *C. elegans* germline is spermatogenic, gravid adults, where the germline is oogenic, and embryos, where the two are brought together. We note that harvesting of gravid adults will cause our samples to contain embryo material, yet the lifestage was chosen to maximize the temporal distance between the spermatogenic and oogenic worms. As will be shown in the following section, differences between embryos and gravid adults were large, wherefore we do not believe the presence of embryos in the gravid adults to cause major error in our findings.

We primarily used Next-Generation Sequencing (NGS) data to answer our questions; these being summarized in **Figure 1**. We used three different kinds of NGS data in two different types of mutants. In an endogenously tagged mutant, we used RNA immunoprecipitation sequencing (RIP-Seq) to determine 22G-RNAs bound by WAGO-3. In single deletion mutants, we used small RNA-Seq (sRNA-Seq) and poly(A)-selected RNA sequencing (mRNA-Seq) to determine how loss of WAGO-1, -3, or -4 changes the RNA landscape. This is all summarized in **Figure 1**, which details the separate subsections of the following section.

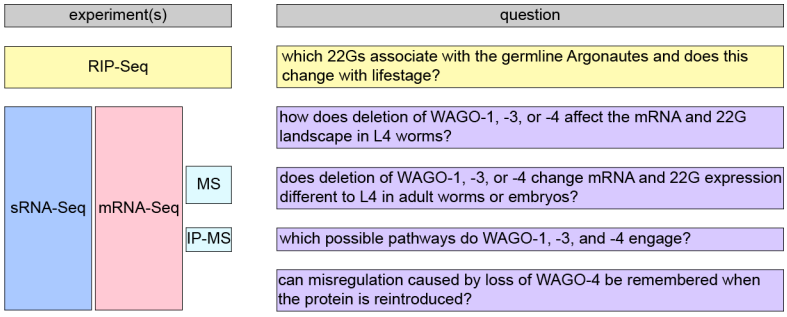


Figure 1: Workflow of the chapter. *The questions posed represent subsections of the results section, in appropriate order.*

Results

Interaction of WAGOs and 22G-RNAs at different lifestages

WAGO-3 binds slightly different subsets of 22G-RNAs in L4 larvae and adult worms

Having established in **Chapter I** that GFP::3xFLAG::WAGO-3 in *C. elegans* associates with slightly different subsets of 22G-RNAs in males and hermaphrodites, we sought to investigate whether the same was true for hermaphrodites at a different lifestages. For this purpose, we carried out RIP-Seq of *gfp:3xflag::wago-3(xf119)* in L4 worms in order to compare WAGO-3 RIP-Seq data from adults (male and hermaphrodite) and L4 larvae. We saw a clear enrichment of 22G-RNAs in our IP relative to our input samples (**Figure 2A**), and we were able to determine several target genes using the same parameters as in **Chapter I** (**Figure 2B** and **C**).

As is the case for adult hermaphrodites and adult males, 22G-RNAs mapping to transposons were overrepresented in the WAGO-3 IPs from L4 hermaphrodites (**Figure 2D**), with highly active transposons such as Tc1 and Tc3 enriched in all replicates (**Figure 2C**). WAGO-3 also appears to distinctively *not* bind 22G-RNAs against non-coding RNAs in all cases (**Figure 2D**). Overall, we detected many more targets from the L4 IPs than from the other two stages. This could be a true effect, supporting the idea that WAGO-3 is more important in the spermatogenic gonad than in the adult gonad, however, due to the differences in our libraries (**Figure 2A** and **Chapter I**, **Extended Figure 2D**), we cannot definitively make such conclusions.

The overlap between targets found in each case was large with some targets unique for the specific sex and lifestage, partially due to the much larger set of targets found in the L4 hermaphrodite (1475 targets against 785 in the adult hermaphrodite and 757 in the adult male) (**Figure 2E**). Furthermore, when comparing our findings to WAGO-3 targets as defined by Seroussi and coworkers (Seroussi et al., 2023) in the L4-to-YA transition of hermaphrodites, we find that the overlap, although large, is not perfect (**Figure 2F**), suggesting that WAGO-3 may change targets during development and in a context-dependent manner, i.e. differently in male and hermaphroditic worms. In addition, we cannot rule out that technical differences, differences in strain history, or differential tagging of WAGO-3 contribute to the relatively large number of non-shared targets.

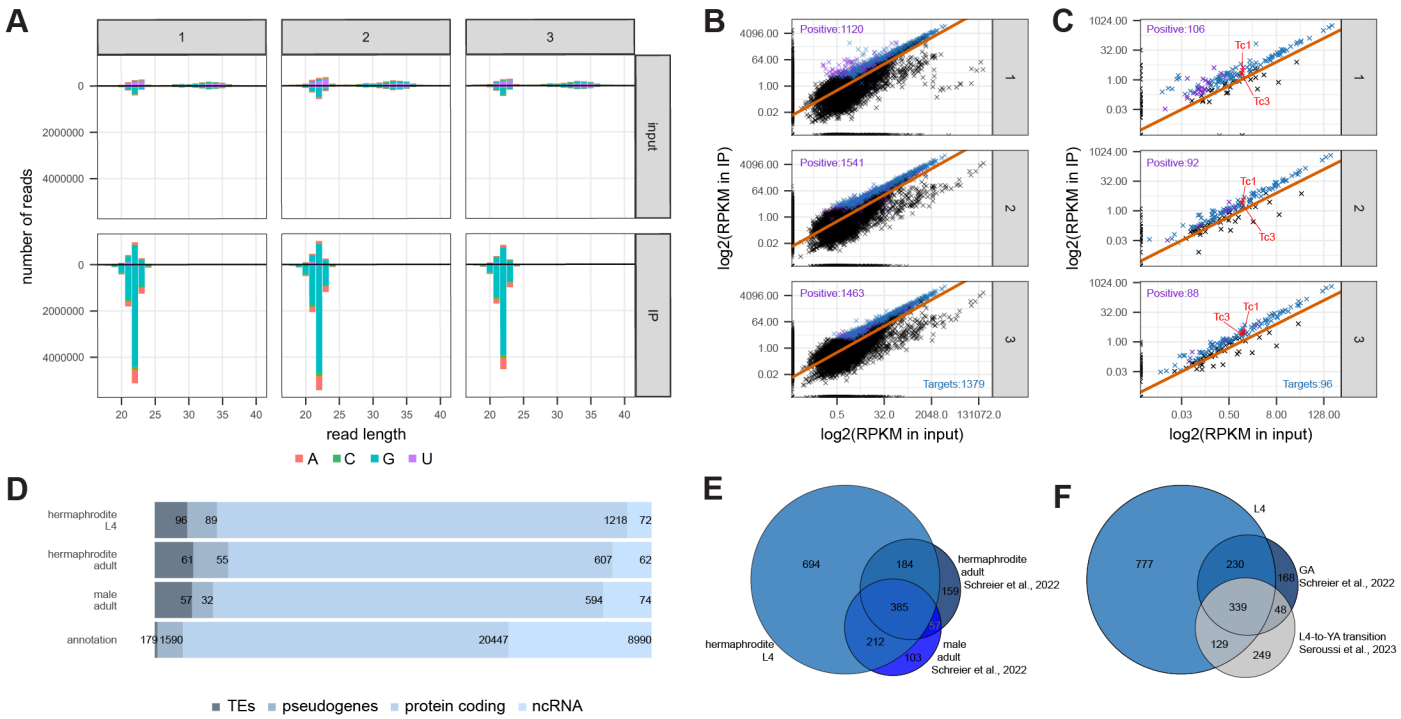


Figure 2: RIP-Seq of* gfp::3xflag::wago-3(xf119) *L4 worms. A *Distribution of reads found in input and IP samples and mapping sense (up) or antisense (down) to their corresponding genetic elements. Colours represent the 5'-nucleotide.* ***B*** *Scatterplot showing RPKM in input versus RPKM in IP in each of the three replicates. Colours represent positive hits (RPKM above 4 in IP and at least two-fold enriched over input and non-zero RPKM in input) and targets (positive in two or more replicates), respectively. Transposons omitted.* ***C*** *Scatterplot showing RPKM of transposons in input versus RPKM of transposons in IP in each of the three replicates. Colours represent positive hits (RPKM at least two-fold enriched in IP over input and non-zero RPKM in input) and targets (positive in two or more replicates), respectively.* ***D*** *Bar plot showing the biotypes of WAGO-3 targets in L4 larvae as well as those defined in Chapter I and in the full annotation.* ***E*** *Venn diagram showing the overlap between WAGO-3 targets in L4 larvae and adult hermaphrodites as well as in male.* ***F*** *Venn diagram showing the overlap between hermaphrodite WAGO-3 targets defined by us and by (Seroussi et al., 2023).*

DISTINCT SUBSETS OF SMALL RNAS ASSOCIATE WITH DIFFERENT WAGOS

Publicly available RNA-Seq data shows that WAGO-3 expression starts before L4 stage (**Figure 3A**), although we have not sought to further investigate the distinctions between lifestages. Similarly, expression of WAGO-1 and WAGO-4 starts early with WAGO-1 being highly expressed in early embryos as well (**Figure 3A**). All three of the WAGOs are found in male worms (**Figure 3A**).

While we have not conducted RIP-Seq experiments for either WAGO-1 nor WAGO-4, other groups have. In 2009, Gu and coworkers (Gu et al., 2009) defined a small number of WAGO-1 targets by conducting RIP-Seq in gravid adult hermaphrodites kept at 24°C, and 14 years later, Seroussi and coworkers (Seroussi et al., 2023) defined a larger number of targets in hermaphrodites at the L4-to-YA transition kept at 20°C. Even

with the large difference in number of found targets (97 in 2009 versus 2,011 in 2023), the latter did not fully encompass the former (**Figure 3B**), once again suggesting that targets may vary at different lifestages whilst simultaneously asking the question of whether 22G-RNA loading may also be affected by temperature differences. However, it should be noted that two different tagged mutants were used by the two groups. Furthermore, the difference in the abundance of found genes, although the parameters for the analyses carried our were near identical, highlights the variance of results obtained from the same experiment carried out at different times and/or by different hands.

Seroussi and coworkers (Seroussi et al., 2023) furthermore carried out RIP-Seq on tagged WAGO-4 in hermaphrodites at the L4-to-YA transition. In 2018, Xu and coworkers (Xu et al., 2018) had defined WAGO-4 targets in gravid adult hermaphrodites, and once again, the overlap between targets found were large but could suggest some differences in how WAGO-4 is loaded at different lifestages, although the number of hits found by both groups once again differed greatly (**Figure 3C**).

Altogether, currently available RIP-Seq data can neither confirm nor refute differences in binding capacities of WAGO-1, -3, or -4 at different lifestages. While there is some indication that overlaps may be large but varied, experiments at different lifestages would have to be carried out in tandem in order to fully test this hypothesis. What we can conclude, however, is that when we compare all targets of WAGO-1, -3 and -4 (**Figure 3D**), it becomes clear that, while some overlaps exists, each WAGO associates with 22Gs targeting distinct subset of genes.

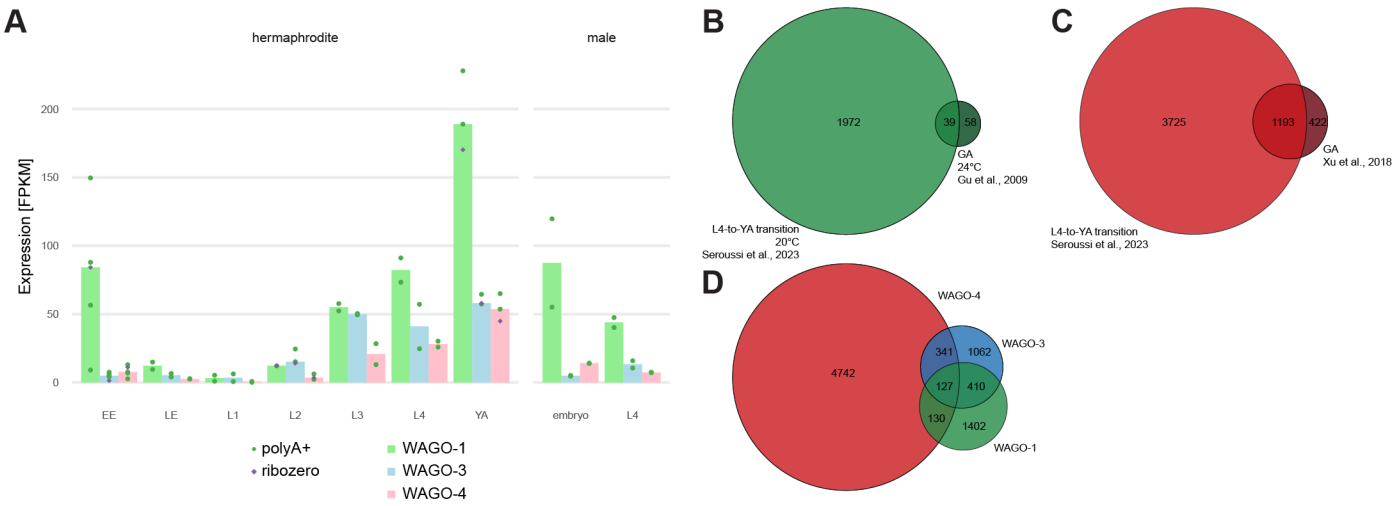


Figure 3: Expression and targets of WAGO-1, -3, and 4. ***A*** *Expression of WAGO-1, -3 and -4 during development as based on publicly available RNA-Seq data in hermaphrodites and males. EE = early embryo, LE = late embryo, YA = young adult.* ***B*** *Venn diagram showing the targets of WAGO-1 as defined by RIP-Seq.* ***C*** *Venn diagram showing the targets of WAGO-4 as defined by RIP-Seq.* ***D*** *Venn diagram showing the overlap of targets of WAGO-1, -3, and -4. Targets of either WAGO is the combination of all RIP-Seq experiments carried out in hermaphrodites in any lifestage.*

Effects of WAGO-1, -3, or -4 deletion on mRNA and 22G-RNA expression in L4 larvase

After having noted that WAGO-1, -3, and -4 interact with 22G-RNAs targeting different subsets of genes, we hypothesized that loss of either one of the WAGOs may cause a decrease of its specific 22Gs which could lead to an increase in mRNA levels of those same target genes, given that these can no longer be silenced. In order to investigate this hypothesis, we carried out sRNA-Seq and mRNA-Seq of hermaphrodites with no functional WAGO-1, -3, or -4, respectively.

Since all further experiments were conducted in *C. elegans* hermaphrodites, this distinction will no longer be mentioned, although we note that differences likely exist between hermaphrodites and males. Rather than testing the differences between males and hermaphrodites, we sought to investigate the differences between lifestages of hermaphrodites. To that end, we synchronized *wago-1(ok1074)*, *wago-3(pk1673)*, and *wago-4(tm1019)* worms and harvested them at L4 stage and at gravid adult stage. We furthermore harvested mixed-stage embryos via bleaching from plates containing gravid adults that had already started laying eggs. We wanted to minimize secondary effects, and for that purpose we outcrossed all worms to a wildtype strain prior to harvest. We outcrossed the worms twice; once from either maternal or paternal side and confirmed them as homozygous for the deletion via PCR. The harvested worms were in their fifth generation of being homozygous.

While this section focuses on the effects in L4 worms, the general features of all our libraries will first be presented.

Deletion of WAGO-1, -3, and -4 could be confirmed via mRNA-Seq

Due to technical issues, we sequenced our samples in two batches, as indicated in **Figure 4A-D**. The largest difference between samples, both for sRNA-Seq and mRNA-Seq, existed between samples from different lifestages rather than samples from different mutants (**Figure 4A-D**).

In order to confirm that we were indeed sequencing *wago-1(ok1074)*, *wago-3(pk1673)*, and *wago-4(tm1019)* deletion mutants, we generated coverage plots for the mRNA-Seq data along each of the affected genes (**Figure 4E-G**). While we could confirm that we had zero reads mapping to the deleted regions of *wago-3* and *wago-4* (**Figure 4F** and **G**), we found reads in the area of presumed deletion of *wago-1* in *wago-1(ok1074)* worms (**Figure 4E**). Since our PCR analysis had confirmed that we saw no wildtype bands in *wago-1(ok1074)*, we theorized that the mRNA-Seq reads mapping to the deleted region were a cause of a translocation of the deleted region. To further investigate this, we created primers internally in the deleted region of *wago-1* (**Figure 3H**). PCR analysis confirmed that the deleted region of *wago-1* was present in *wago-*

1(ok1074), but that the region spanning the deletion was not (**Figure 3I**). We conclude that even though wago-1(ok1074) is expressed at the RNA level, this cannot not result in fully functional WAGO-1 protein. However, we cannot exclude that this allele may produce some protein with partial functionality.

Deletion of WAGO-1, -3, or -4 does not change global 22G-RNA levels

We next turned our focus to the sRNA-Seq data. We found that all samples had a peak of reads at the length 20-23nt that predominantly had a G at the 5'-position (**Figure 5A**), consistent with the fact that 22G-RNAs should be the most abundant small RNA moiety. Apart from 22G-RNAs, we found that all samples had some expression of miRNAs as well as expression of RNAs longer than 27nt, the latter being more predominant in the embryos, but we saw no difference in small RNA types present in samples of wildtype or mutant worms (**Figure 5**).

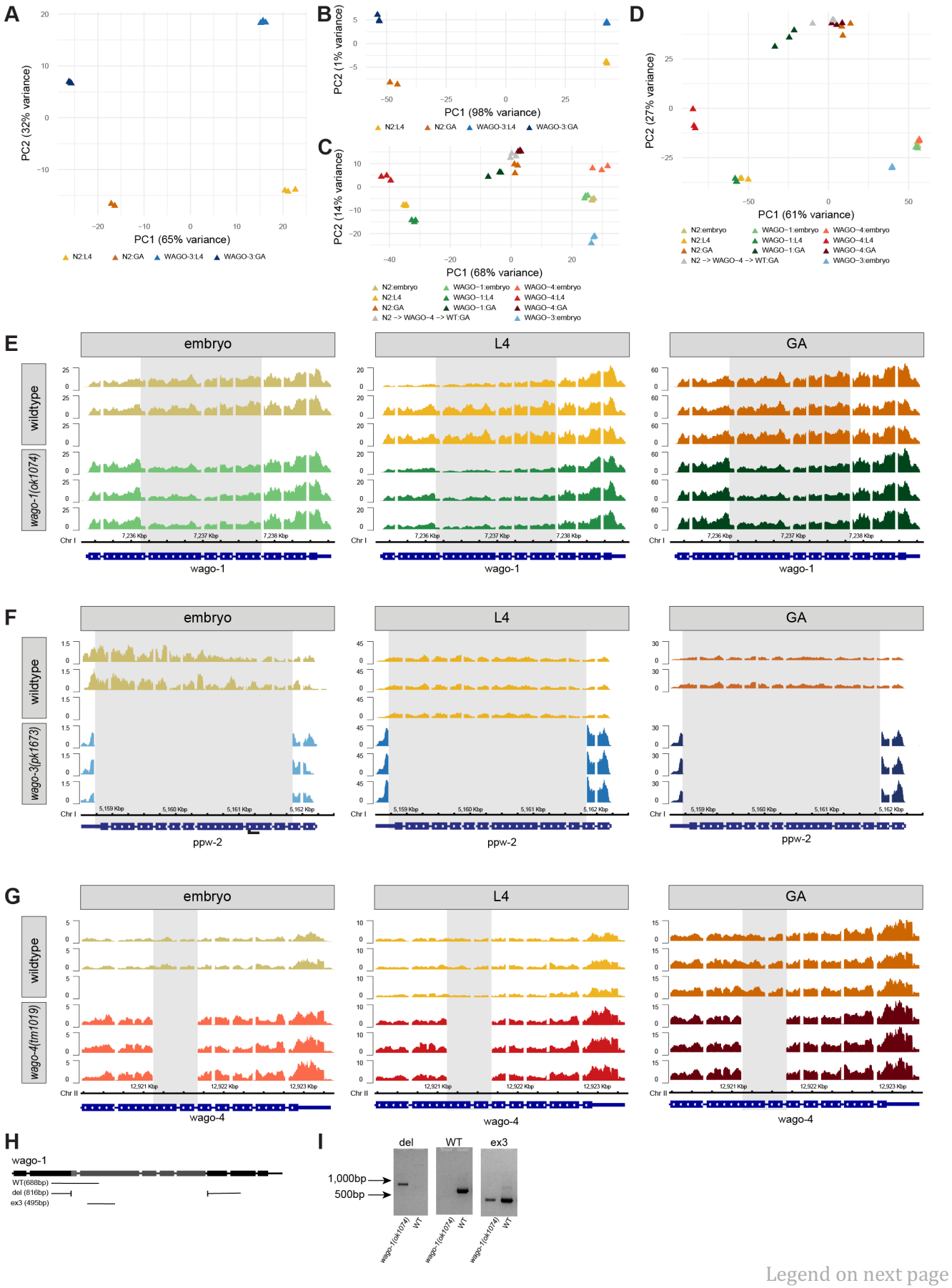
Since all WAGOs are known to associate with 22G-RNAs, we decided to limit our analyses to changes in this subset of small RNAs. We therefore isolated reads in the length 20-23nt with no requirement of 5'-nucleotide, and these reads will be referred to simply as 22G-RNAs going forward.

Changes in mRNAs and 22G-RNAs upon loss of WAGO-1, -3, or -4 are largely unrelated in L4 larvae

We first sought to investigate the change in RNA landscape in each of our three deletion mutants in L4 larvae, where hermaphroditic *C. elegans* undergoes spermatogenesis. We wanted to investigate which mRNAs were affected, which 22G-RNAs were affected, and how these two correlated.

In *wago-3(pk1673)* as well as in *wago-4(tm1019)*, around 4-500 genes had regulated 22Gs in either direction (**Figure 6A**). *wago-1(ok1074)* showed less of an effect, with only 172 genes with upregulated 22Gs and just below 100 genes with downregulated 22Gs (**Figure 6A**). On mRNA level, however, 460 genes were upregulated in *wago-1(ok1074)* with much fewer (73) downregulated, and only a few more than 100 genes were upregulated in either direction in *wago-3(pk1673)* (**Figure 6B**). The regulation in *wago-4(tm1019)* was much starker than either of these, with nearly 1,000 genes downregulated and more than 2,500 genes upregulated (**Figure 6B**).

Surprisingly, we didn't find any correlation between genes regulated at either 22G-RNA level or at mRNA level, and genes defined as targets via RIP-Seq also appeared to only show regulation at one level but not the other or none at all (**Figure 6C**). The only correlation between the three – 22G-RNAs, mRNA, and RIP-Seq targets – we found, was a subset of 162 genes, that were upregulated in *wago-4(tm1019)* both at 22G-RNA level and at mRNA level. GO-term analysis revealed that some of these genes were related to phosphorylation



Legend on next page

Figure 4: Sequencing of WAGO deletion mutants. A *PCA plot of the 22G-RNAs from our first sequencing batch.* ***B*** *PCA plot of the mRNAs from our first sequencing batch.* ***C*** *PCA plot of the 22G-RNAs from our second sequencing batch.* ***D*** *PCA plot of the mRNAs from second first sequencing batch.* ***E*** *mRNA reads along the* wago-1 *gene in wildtype and* wago-1(ok1074)*. Location of deletion marked as a grey box.* ***F*** *mRNA reads along the* wago-3 *gene in wildtype and* wago-3(pk1673)*. Location of deletion marked as a grey box.* ***G*** *mRNA reads along the* wago-4 *gene in wildtype and* wago-4(tm1019)*. Location of deletion marked as a grey box.* ***H*** *Schematic of the primers used to genotype* wago-1(ok1074)*. Deletion marked in grey.* ***I*** *PCR of wildtype and* wago-1(ok1074) *using the primers presented in panel H.*

or dephosphorylation, specifically of histones (**Figure 6D**) and nine of them were membranal (**Figure 6E**). We also found upregulation of the terms "oogenesis", "gamete generation", and "eggshell formation" (**Figure 6D**), which could indicate early onset of adult-specific genes, with oogenesis possibly commencing prematurely in L4 larvae.

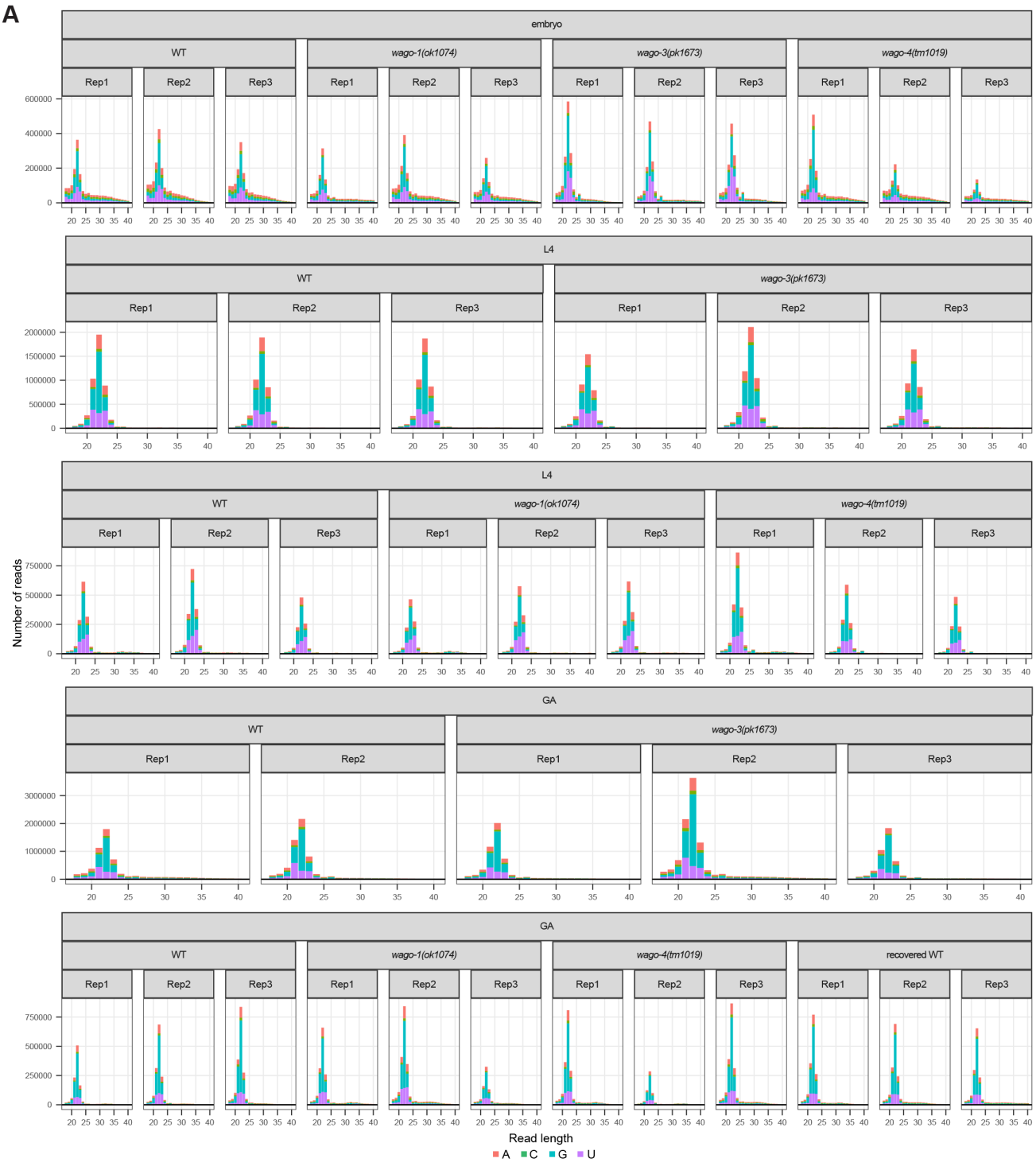
We conclude that loss of each WAGO affect mRNA and 22G-RNA levels to different extents in L4 larvae with WAGO-4 having the largest effect, and that loss or gain of 22G-RNAs does not directly affect corresponding mRNAs.

Loss of one WAGO does not cause loss of its associated 22G-RNAs nor regulation of their corresponding genes in L4 larvae

Having previously established that each of the three WAGOs bind 22G-RNAs against distinct subsets of genes, we reasoned that loss of either one of the WAGOs would cause downregulation of 22G-RNAs against those same genes. Therefore, we next overlay our data with the targets as found by RIP-Seq. What we found, however, was that 22Gs against established WAGO-3 targets were not specifically regulated in either direction in *wago-3(pk1673)*, neither were 22G-RNAs against WAGO-4 targets regulated in *wago-4(tm1019)*, with only a very small – albeit significant – difference in fold changes between targets and non-targets in either case (**Figure 7A** and **B**). Even more interestingly, 22G-RNAs against established WAGO-1 targets were *up*regulated in *wago-1(ok1074)* (**Figure 7A** and **B**). Altogether, this suggests that other Argonaute proteins are capable of binding and stabilizing 22G-RNA targets of WAGO-1, -3, and -4.

On mRNA level, neither WAGO-1 nor WAGO-3 targets were regulated in either direction. In *wago-4(tm1019)*, however, WAGO-4 targets were *less* upregulated than non-targets (**Figure 7C** and **D**).

Since we don't know at what stage either of the WAGOs infer gene regulation, we did consider the possibility that we were incapable of finding RIP-Seq target-specific regulation at mRNA level may have to do with the fact that the WAGOs cause downregulation of translation rather than downregulation of the actual abundance of mRNAs, not that loss of the WAGOs simply didn't cause regulation of their associated targets. In order to investigate this, we carried out mass spectrometry of *wago-3(pk1673)* L4 larvae and



Legend on next page

***Figure 5: sRNA-Seq of WAGO deletion mutants. A** Length distribution profiles of all, unmapped reads. Colour represents 5'-nucleotide. **B** Percentwise distribution of 21U-RNAs, 22G-RNAs, 26G-RNAs, miRNAs and reads longer than 27 nucleotides in unmapped samples.*

compared it to wildtype L4 larvae. We were only able to detect 174 of the 1,475 targets of WAGO-3 at L4 (12%), but these did not show regulation in either direction (**Figure 7E** and **F**).

Although the sample size for our mass spectrometry analysis was too small to conclude with certainty that WAGO-3 targets are not regulated at protein level, we can conclude that there is no evidence to support that loss of WAGO-1, WAGO-3, or WAGO-4 in L4 larvae cause regulation of genes against which 22G-RNAs are bound by that specific WAGO.

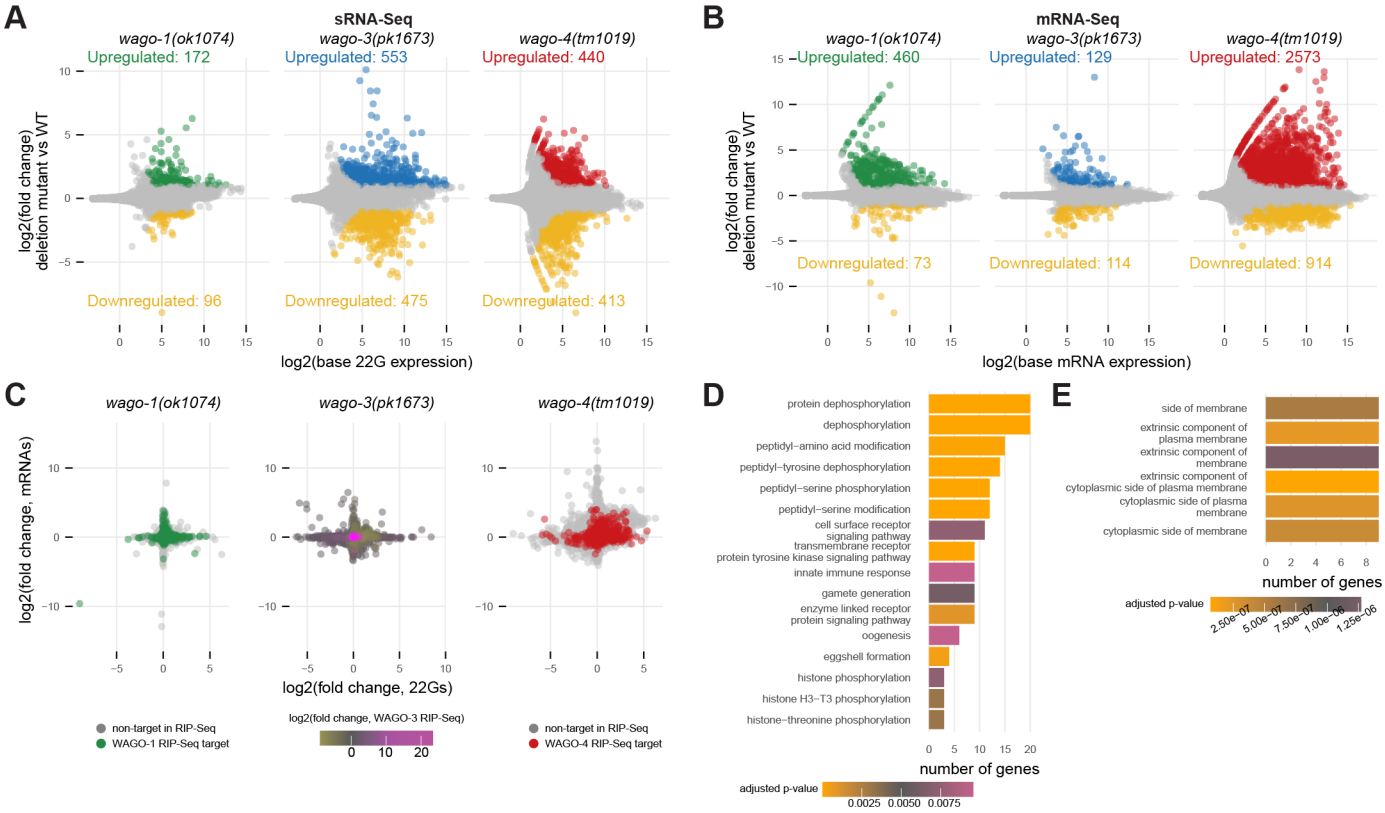


***Figure 6: mRNAs and 22G-RNAs regulated upon WAGO deletion in L4 larvae. A** Scatterplots showing genes with regulated 22G-RNAs. **B** Scatterplots showing genes with regulated mRNAs. **C** Scatterplots comparing fold changes of mRNAs to fold changes of 22G-RNAs. For* wago-1(ok1074) *and* wago-4(tm1019)*, colours represent whether or not genes are defined as targets in RIP-Seq. For* wago-3(pk1673)*, colours represent the fold change of IP versus input of L4* gfp::3xflag::wago-3(xf119)*. **D** Bar plot showing all 'biological process' GO-terms with adjusted p-values below 0.01 found in the 162 genes that were upregulated both at mRNA levels and at 22G-RNA level in* wago-4(tm1019)*. **E** Bar plot showing all 'cellular component' GO-terms with adjusted p-values below 0.01 found in the 162 genes that were upregulated both at mRNA levels and at 22G-RNA level in* wago-4(tm1019)*.*

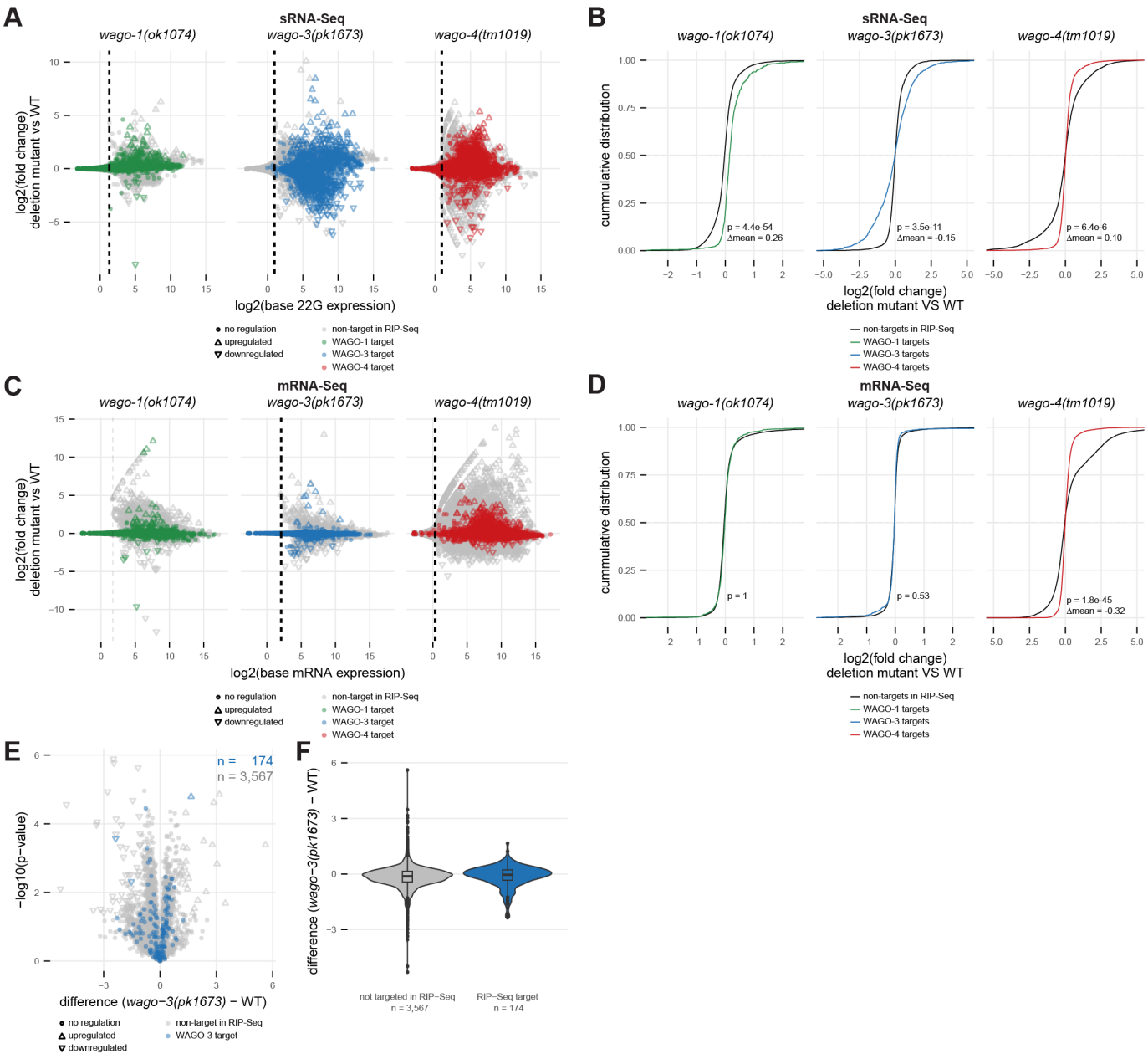


Figure 7: mRNA and 22G-RNA regulation in L4 WAGO deletion mutants compared to known targets. A *Scatterplots showing genes with regulated 22G-RNAs. For* wago-1(ok1074) *and* wago-4(tm1019)*, colours represent whether the gene has been found as a target in any RIP-Seq experiment. For* wago-3(pk1673)*, colours represent whether the gene has been found as a target in our L4 RIP-Seq experiment. Dotted line represents the cutoff used for quantification.* ***B*** *CDF plot quantifying panel A. X-axis has been shortened to show the relevant area and Bonferroni-adjusted p-values from one-way ANOVA indicated along with difference in mean, when* $p \leq 0.01$*.* ***C*** *Scatterplots showing genes with regulated mRNAs. For* wago-1(ok1074) *and* wago-4(tm1019)*, colours represent whether the gene has been found as a target in any RIP-Seq experiment. For*

Legend continues on next page

wago-3(pk1673), *colours represent whether the gene has been found as a target in our L4 RIP-Seq experiment. Dotted line represents the cutoff used for quantification.* ***D*** *CDF plot quantifying panel C. X-axis has been shortened to show the relevant area and Bonferroni-adjusted p-values from one-way ANOVA indicated along with difference in mean, when* $p \leq 0.01$. ***E*** *Scatterplot showing mass spectrometry data of* wago-3(pk1673) *versus wildtype. Genes defined as WAGO-3 targets in L4 larvae are marked in blue.* ***F*** *Violin plot quantifying panel E.*

Loss of WAGO-4 opposes loss of WAGO-1 to regulate adult-type gene expression in spermatogenic *C. elegans* in L4 larvae

Because there is a shift from spermatogenesis to oogenesis between L4 larvae and adult worms and we suspected the different WAGOs to be differentially coupled to these processes, we next sought to investigate how the regulation of 22G-RNAs in L4 larvae matched lifestage specific expression of 22G-RNAs. We were able to detect more than 1,000 genes with 22G-RNAs enriched in either L4 larvae or gravid adult wildtype worms (**Figure 8A**). In L4 larvae, all three deletion mutants regulated 22G-RNAs more prominent in their current lifestage or not enriched in either lifestage in wildtypes (**Figure 8B** and **C**). While both *wago-3(pk1673)* and *wago-4(tm1019)* downregulated L4 enriched 22G-RNAs and upregulated gravid adult enriched 22G-RNAs, this tendency was more prominent in *wago-4(tm1019)* L4 larvae (**Figure 8B** and **C**), which, once again, could suggest that loss of WAGO-4 shifts adult-specific expression to occur earlier.

While all deletion mutants appear to regulate L4 enriched 22G-RNAs in L4 larvae, we did not know whether those 22G-RNAs corresponded to genes also expressed at this lifestage. We did note that some of the L4 enriched 22G-RNAs corresponded to spermatogenesis genes such as spe-9, spe-41, and spe-47, however, this did not answer the question of whether genes for which 22G-RNAs were enriched in L4 larvae or genes for which 22G-RNAs were affected by WAGO deletion were also expressed in L4 larvae. For instance, high expression of 22G-RNAs could be coupled to low mRNA expression or either deletion mutant could be specifically altering 22G-RNA populations against mRNAs generally expressed or not expressed in L4 larvae. In order to investigate mRNA expression of genes corresponding to our different subsets of 22G-RNAs, we downloaded RNA-Seq data from GExplore (Hutter & Suh, 2016). This data contains RNA-Seq from 18 different embryonic time points as well as data from all four larvae stages and young adult worms (Boeck et al., 2016). We normalized each of the genes we found in the database and in our data to its own expression so that we could determine at which lifestage the genes were most prominently expressed. The 22G-RNAs that we defined as gravid adult enriched could indeed correspond to genes expressed primarily in the gravid adult, as their expression peaks in early embryos and, to a lesser extent, in young adults (**Figure 8D**). L4 enriched 22G-RNAs also corresponded to genes with expression in L4 larvae, but the peak of their expression was in the late embryos (**Figure 8D**).

Although the 22G-RNAs downregulated in *wago-1(ok1074)* were not gravid adult enriched (**Figure 8C**), the expression profiles of the corresponding genes did match that of the genes with gravid adult enriched 22G-RNAs, their expression peaking in early embryos and adult worms but relatively low in late embryos and all larval stages (**Figure 8E**). Genes with upregulated 22G-RNAs in *wago-1(ok1074)* had an expression profile opposite of the genes with downregulated 22G-RNAs, i.e. their expression peaked in late embryos and in early larval stages but was lower in early embryos and late larval stages. This was more akin to the genes corresponding to the L4 enriched 22G-RNAs (**Figure 8D** and **E**). In *wago-4(tm1019)*, this was completely opposite, with genes corresponding to downregulated 22G-RNAs peaking in the late embryo/L1 larvae and genes with upregulated 22G-RNAs peaking in early embryos and adult worms (**Figure 8G**). While *wago-1(ok1074)* and *wago-4(tm1019)* displayed opposing effects, *wago-3(pk1673)* was somewhere in-between, with both genes with upregulated 22Gs and genes with downregulated 22Gs having a peak of expression in the late embryo, but also high relative expression in late larval stages and adult worms (**Figure 8F**).

On mRNA level, gene expression of regulated genes in *wago-1(ok1074)* and *wago-3(pk1673)* followed the same pattern, with upregulated genes peaking in expression in the early larval stages and downregulated genes peaking in later larval stages and adult worms as well as in late embryos (**Figure 8H** and **I**). Once again, expression of genes regulated in *wago-4(tm1019)* was opposite of that, with late expressed genes being upregulated and early expressed genes being downregulated (**Figure 8J**). In all mutants, genes regulated in either direction also appeared to have some expression in the late embryo, suggesting that any WAGO deletion may disturb late embryonic genes. For reference, we would like to point out that the genes we found to have highest expression in the mRNA-Seq of our L4 wildtype worms, although highly expressed in L4 stage of the GExplore data, did not actually peak in expression at that stage. Rather, their expression started earlier, somewhat mimicking the expression profile of the genes upregulated in *wago-1(ok1074)* and *wago-3(pk1673)* (**Figure 8K**). ‘

A further consideration of ours was whether the genes regulated in any mutant were germline expressed. In 2014, a study conducted found above 10,000 genes that were enriched in the germline with subsets of these being more expressed in spermatogenic or oogenic gonads of obligate males and females from mutant worms (Ortiz et al., 2014). Overlaying this data with our data, we saw that most of the regulated 22G-RNAs indeed corresponded to germline expressed genes (**Figure 8L**), whereas the proportion of germline expressed genes found in our regulated mRNAs more resembled the proportion of germline expression in the entire genome (**Figure 8M**). Although it is possible that some of the genes not defined as germline expressed by Ortiz and coworkers are present in the germline, this does suggest that some of the mRNA regulation observed is somatic, which could help explain the disparity between 22G-RNA regulation and

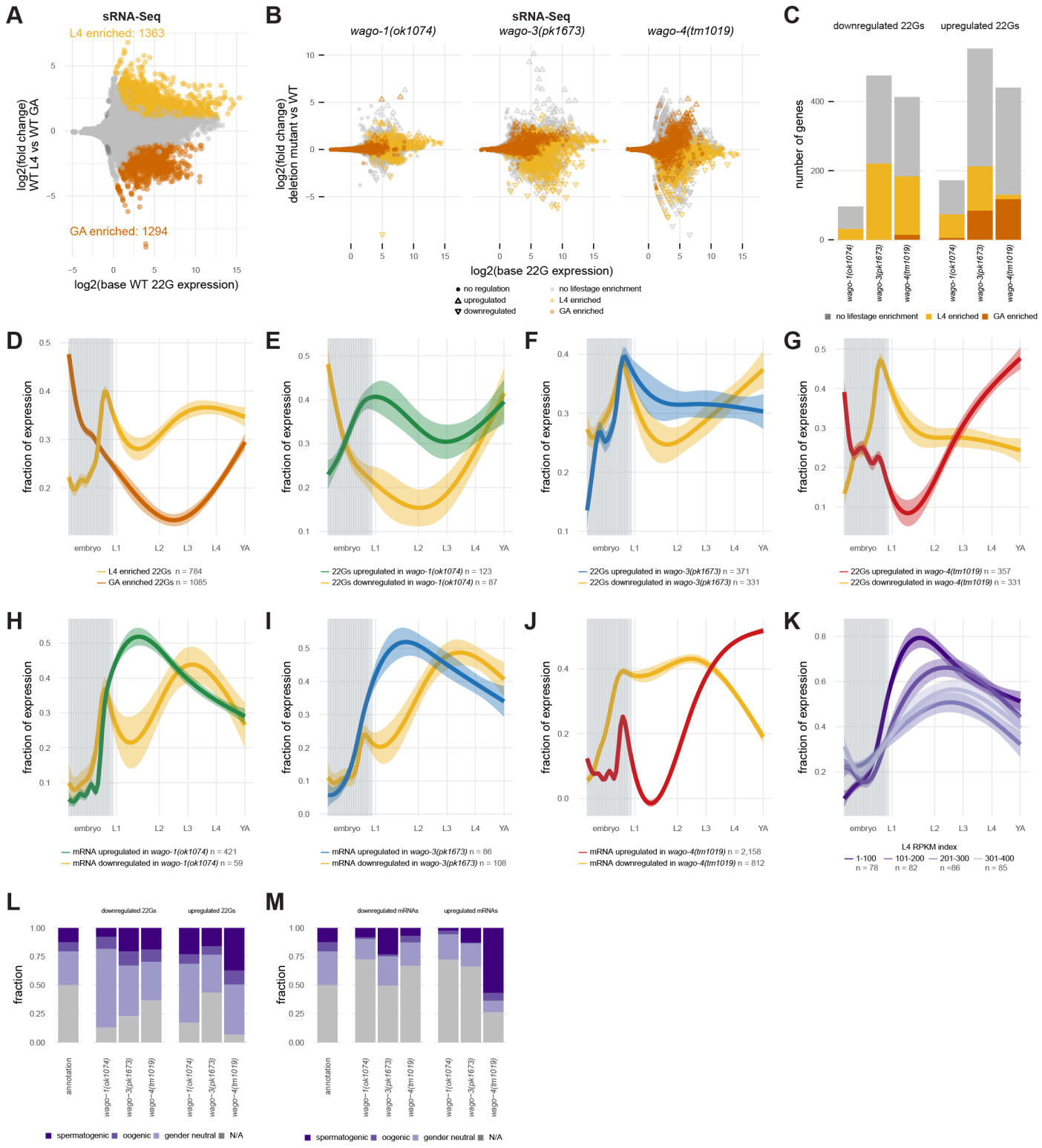


Figure 8: Lifestage specificity of genes regulated in WAGO deletion mutants in L4 larvae. A *Scatterplot comparing 22G-RNAs in wildtype L4 larvae to 22G-RNAs in wildtype gravid adult.* ***B*** *Scatterplot showing genes with regulated 22G-RNAs. Colours represent lifestage enrichment (gravid adult, L4, or none) in wildtype, as defined in panel A.* ***C*** *Bar plot showing the lifestage enrichment of 22G-RNAs significantly regulated in deletion*

Legend continues on next page

mutant versus wildtype. ***D-K*** *Expression profiles along lifestage as downloaded from GExplore (Hutter & Suh, 2016). Profiles made using a generalized additive model and stages beyond embryo given approximate timepoints; L1 = 800 min, L2 = 1520 min, L3 = 2000 min, L4 = 2480 min, YA (young adult) = 2960 min.* ***D*** *Expression profiles of genes with 22G-RNAs enriched in L4 larvae or gravid adults in the wildtype, as defined in panel A.* ***E*** *Expression profiles of genes with 22G-RNAs up- or downregulated in* wago-1(ok1074). ***F*** *Expression profiles of genes with 22G-RNAs up- or downregulated in* wago-3(pk1673). ***G*** *Expression profiles of genes with 22G-RNAs up- or downregulated in* wago-4(tm1019). ***H*** *Expression profiles of genes with mRNAs up- or downregulated in* wago-1(ok1074). ***I*** *Expression profiles of genes with mRNAs up- or downregulated in* wago-3(pk1673). ***J*** *Expression profiles of genes with mRNAs up- or downregulated in* wago-4(tm1019). ***K*** *Expression profiles of the 400 top-most expressed genes in wildtype L4 larvae, divided into hundreds as based on highest expression.* ***L*** *Bar plot showing germline expression of 22G-RNAs regulated in deletion mutants as defined by (Ortiz et al., 2014).* ***M*** *Bar plot showing germline expression of mRNAs regulated in deletion mutants as defined by (Ortiz et al., 2014).*

mRNA regulation in all three of our mutants. Somatic mRNA regulation may be a number of steps downstream of 22G-RNA misregulation in the germline and the effect may therefore be indirect.

Interestingly, we did note that mRNAs upregulated in *wago-4(tm1019)* were predominantly spermatogenic (**Figure 8M**). Since our previous analyses suggested loss of WAGO-4 to induce early onset of adult-type expression, we would have expected oogenic (adult-type) mRNAs to be upregulated and spermatogenic (L4 type) mRNAs to be downregulated. At the same time, a strong shift towards early oogenesis would suggest that *wago-4* deletion mutant worms may have a feminizations phenotype of their L4 gonad and possible less sperm in their spermatheca due to shortened time in the spermatogenic phase, but no such phenotypes have thus far been reported. Since our data is not fully one-sided as to whether or not loss of WAGO-4 drives earlier onset of oogenesis or of adult-specific genes, we cannot draw definitive conclusions in this regard.

Nonetheless, we can conclude that loss of WAGO-1 and loss of WAGO-4 have opposing effects on the RNA landscapes of L4 larvae, with *wago-1(ok1074)* favouring expression of late embryo or early larval genes and 22G-RNAs and disfavouring expression of adult and early embryo genes and 22G-RNAs while *wago-4(tm1019)* favours expression of adult and early embryo genes and 22G-RNAs and disfavours expression of late embryo or early larval genes and 22G-RNAs. The effect of loss of WAGO-3 is an intermediate between the two, more closely resembling WAGO-1 at mRNA level.

Effects of WAGO-1, -3, or -4 deletion on mRNA and 22G-RNA expression in adult worms and embryos

Having seen that loss of either WAGO-1, -3, or -4 had different effects in L4 larvae, we next sought to see how this looked in gravid adult worms and in embryos. We were interested whether we could capture any

correlations between our sRNA-Seq data, our mRNA-Seq data, and known targets defined via RIP-Seq, which we failed to capture in L4 lavae. Furthermore, we wanted to see whether the RNAs regulated in either mutant differed with lifestage.

DELETION OF GERMLINE WAGOS REGULATES NON-TARGET GENE EXPRESSION IN OOGENIC ADULT *C. ELEGANS*

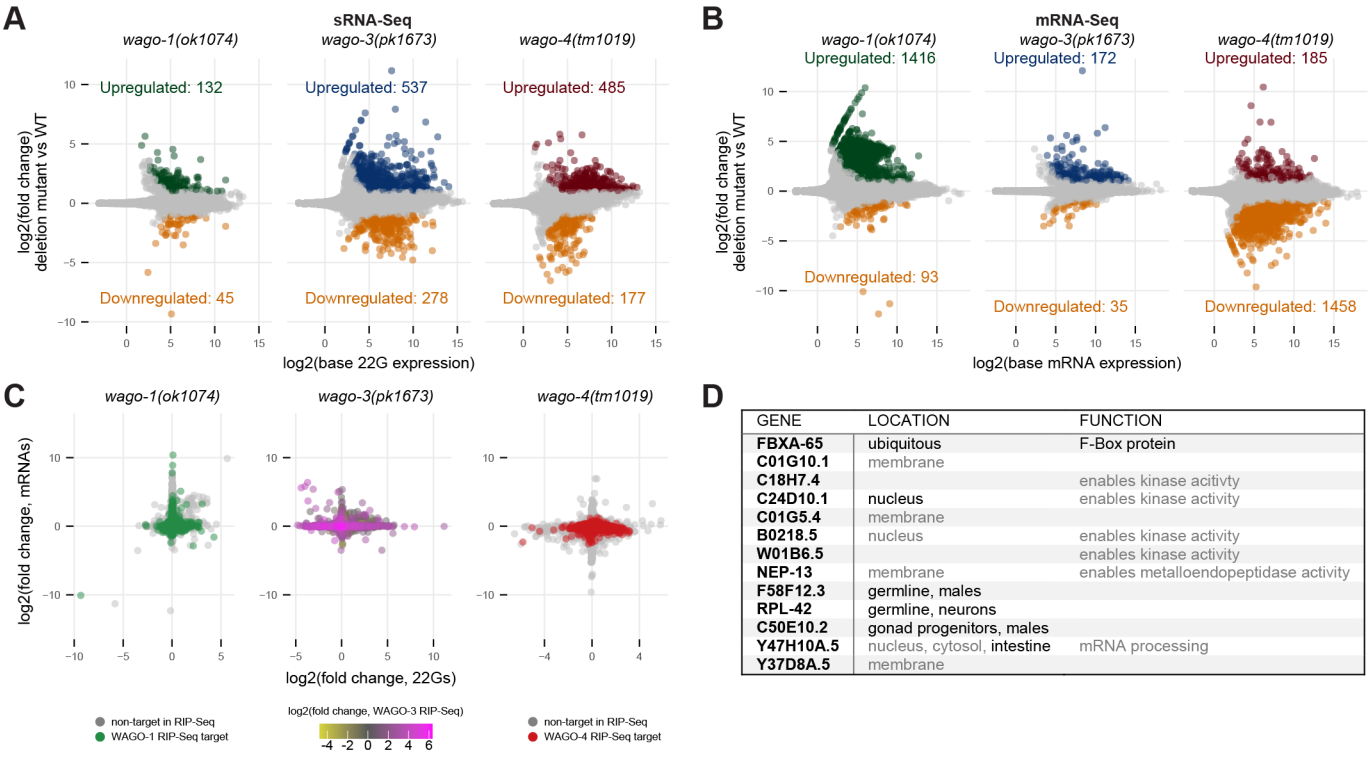
We started by turning our attention to the gravid adult worms. Here, we wanted to see the extent to which 22G-RNAs and mRNAs were regulated and whether this showed any intercorrelation or correlation to defined targets.

The number of genes with regulated 22G-RNAs was similar to the numbers in L4 larvae for all mutants, albeit with some lessening of downregulation in both *wago-3(pk1673)* and *wago-4(tm1019)* (**Figure 9A**). On mRNA level, however, the numbers differed drastically between the two lifestages, with more than 1,400 genes with upregulated 22G-RNAs in *wago-1(ok1074)* in the adult worm (**Figure 9B**) compared to less than 500 in L4 larvae (**Figure 6B**). For *wago-4(tm1019)*, this was reverse; more than 2,500 genes had upregulated 22G-RNAs in L4 larvae (**Figure 6B**) but less than 200 genes had upregulated 22G-RNAs in the adult worm (**Figure 9B**). For *wago-3(pk1673)*, the numbers were comparable between the two lifestages (**Figure 6B** and **Figure 9B**).

Like the case in L4 larvae, there was little to no correlation between changes in 22G-RNA levels and changes in mRNA levels (**Figure 9C**). We did find a small subset of 13 genes that were upregulated at both levels in *wago-1(ok1074)* (**Figure 9C**), and these were generally related to protein modifications, with at least three being germline expressed (**Figure 9D**).

Neither genes with regulated 22G-RNAs nor genes with regulated mRNA levels correlated to genes defined as targets of the respective WAGOs via RIP-Seq, with only very slight differences between targets and non-targets(**Figure 10**). Like in L4 larvae, WAGO-3 targets appeared to be generally regulated in both directions in *wago-3(pk1673)* (**Figure 10A** and **B**) and WAGO-1 target 22G-RNAs were slightly upregulated in *wago-1(ok1074)*. There was a small overall downregulation of WAGO-4 target 22G-RNAs in *wago-4(tm1019)* compared to the wildtype (**Figure 10B**), although it does seem that the downregulated WAGO-4 targets are generally lowly expressed and that a second subset of targets with higher expression are upregulated (**Figure 10A**).

On mRNA level, WAGO-1 targets are slightly downregulated and WAGO-4 targets are slightly upregulated in the respective mutants (**Figure 10C** and **D**). However, since mRNAs are generally upregulated in *wago-1(ok1074)* and mRNAs are generally downregulated in *wago-4(tm1019)*, it seems that the targets are actually not regulated in either mutant. Rather, WAGO-1 target mRNAs are specifically *not* regulated in *wago-1(ok1074)*, and WAGO-4 target mRNAs are specifically *not* regulated in *wago-4(tm1019)*.



GENE	LOCATION	FUNCTION
FBXA-65	ubiquitous	F-Box protein
C01G10.1	membrane	
C18H7.4		enables kinase acitivty
C24D10.1	nucleus	enables kinase activity
C01G5.4	membrane	
B0218.5	nucleus	enables kinase activity
W01B6.5		enables kinase activity
NEP-13	membrane	enables metalloendopeptidase activity
F58F12.3	germline, males	
RPL-42	germline, neurons	
C50E10.2	gonad progenitors, males	
Y47H10A.5	nucleus, cytosol, intestine	mRNA processing
Y37D8A.5	membrane	

Figure 9: mRNAs and 22G-RNAs regulated upon WAGO deletion in gravid adult worms. ***A*** *Scatterplots showing genes with regulated 22G-RNAs.* ***B*** *Scatterplots showing genes with regulated mRNAs.* ***C*** *Scatterplots comparing fold changes of mRNAs to fold changes of 22G-RNAs. For* wago-1(ok1074) *and* wago-4(tm1019)*, colours represent whether or not genes are defined as targets in RIP-Seq. For* wago-3(pk1673)*, colours represent the fold change of IP versus input of gravid adult* gfp::3xflag::wago-3(xf119). ***D*** *Table of genes both upregulated on mRNA and 22G-RNA level in* wago-1(ok1074)*. Predicted locations or functions are shown in grey.*

A final consideration we made, was that perhaps the reason we found little correlation between 22G regulation and mRNA regulation was that we compared within the same lifestage. It could be that 22Gs in L4 worms regulate mRNA expression in adult worms; for instance, they could be there to help dampen the expression of adult genes in the L4 larvae. However, when comparing 22G-RNAs regulated in L4 larve or gravid adult worms in either of the four mutants to mRNAs regulated in gravid adults, there was little difference (**Figure 11**).

Overall, we conclude that there is a difference to the extent to which loss of a certain WAGO affect RNA levels in L4 larvae or gravid adult worms and that no global correlations between changes to 22G-RNAs, changes to mRNAs, and 22G-RNAs bound to tagged WAGOs could be extracted form our data.

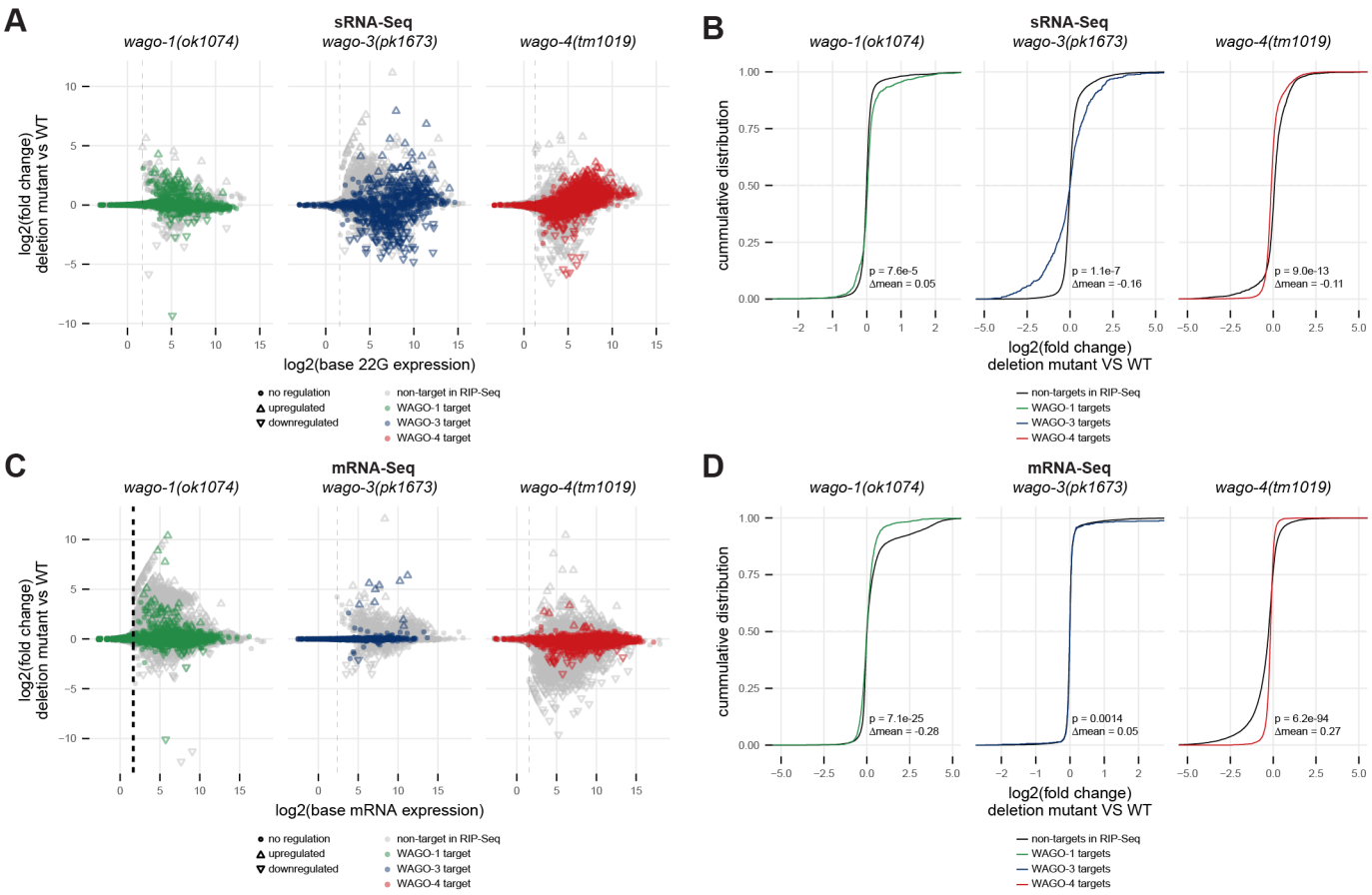


Figure 10: mRNA and 22G-RNA regulation in gravid adult WAGO deletion mutants compared to known targets. A *Scatterplots showing genes with regulated 22G-RNAs. For* wago-1(ok1074) *and* wago-4(tm1019)*, colours represent whether the gene has been found as a target in any RIP-Seq experiment. For* wago-3(pk1673)*, colours represent whether the gene has been found as a target in our gravid adult RIP-Seq experiment. Dotted line represents the cutoff used for quantification.* ***B*** *CDF plot quantifying panel A. X-axis has been shortened to show the relevant area and Bonferroni-adjusted p-values from one-way ANOVA indicated along with difference in mean, when p ≤0.01.* ***C*** *Scatterplots showing genes with regulated mRNAs. For* wago-1(ok1074) *and* wago-4(tm1019)*, colours represent whether the gene has been found as a target in any RIP-Seq experiment. For* wago-3(pk1673)*, colours represent whether the gene has been found as a target in our gravid adult RIP-Seq experiment. Dotted line represents the cutoff used for quantification.* ***D*** *CDF plot quantifying panel C. X-axis has been shortened to show the relevant area and Bonferroni-adjusted p-values from one-way ANOVA indicated along with difference in mean, when p ≤0.01.*

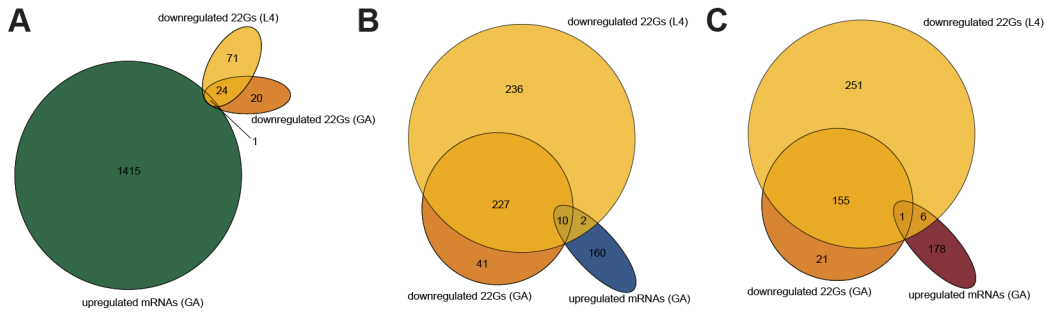


Figure 11: Overlap of downregulated 22G-RNAs and upregulated mRNAs. ***A*** *Venn diagram showing overlap of regulated genes in* wago-1(ok1074)*.* ***B*** *Venn diagram showing overlap of regulated genes in* wago-3(pk1673)*.* ***C*** *Venn diagram showing overlap of regulated genes in* wago-4(tm1019)*.*

WAGO-3 AND -4, BUT NOT WAGO-1, ARE IMPORTANT FOR APPROPRIATE EMBRYONIC GENE EXPRESSION

The last lifestage we investigated was embryos. Once again, we wanted to see the extent to which 22G-RNAs and mRNAs were regulated and whether this showed any intercorrelation or correlation to defined targets.

In embryos, loss of WAGO-1 had almost no effect on either 22G-RNAs or mRNAs (**Figure 12A** and **B**). 22G-RNA regulation in *wago-3(pk1673)* and *wago-4(tm1019)* was also less than in either of the other lifestages (**Figure 12A**), but the number of genes with regulated mRNA levels was much higher than in any other case, with many more mRNAs downregulated; more than 6,000 in *wago-3(pk1673)* and more than 4,000 in *wago-4(tm1019)* (**Figure 12B**).

As is the case for the other lifestages, 22G-RNA regulation and mRNA regulation was only scarcely related (**Figure 12C**).

Also as seen before, genes with regulated 22G-RNAs or mRNA levels didn't correlate with genes defined as targets via RIP-Seq, with only very slight differences between targets and non-targets (**Figure 13**). The only exception to this was the mRNA levels in *wago-4(tm1019)*, where WAGO-4 targets where specifically not regulated, which was a stark contrast to the heavily downregulate non-targets (**Figure 13C** and **D**).

We can thus definitively conclude that loss of WAGO-1, -3, or -4 causes changes to 22G-RNA levels and to mRNA levels that are not directly linked, that these changes are not directly linked to tbe 22G-RNAs natively bound to endogenously tagged versions of those same WAGOs. We can also conclude that these changes are dependent on lifestage, with loss of WAGO-1 having the highest impact in gravid adults, loss of WAGO-3 having the highest impact on the embryo, and loss of WAGO-4 having the highest impact in L4 larvae.

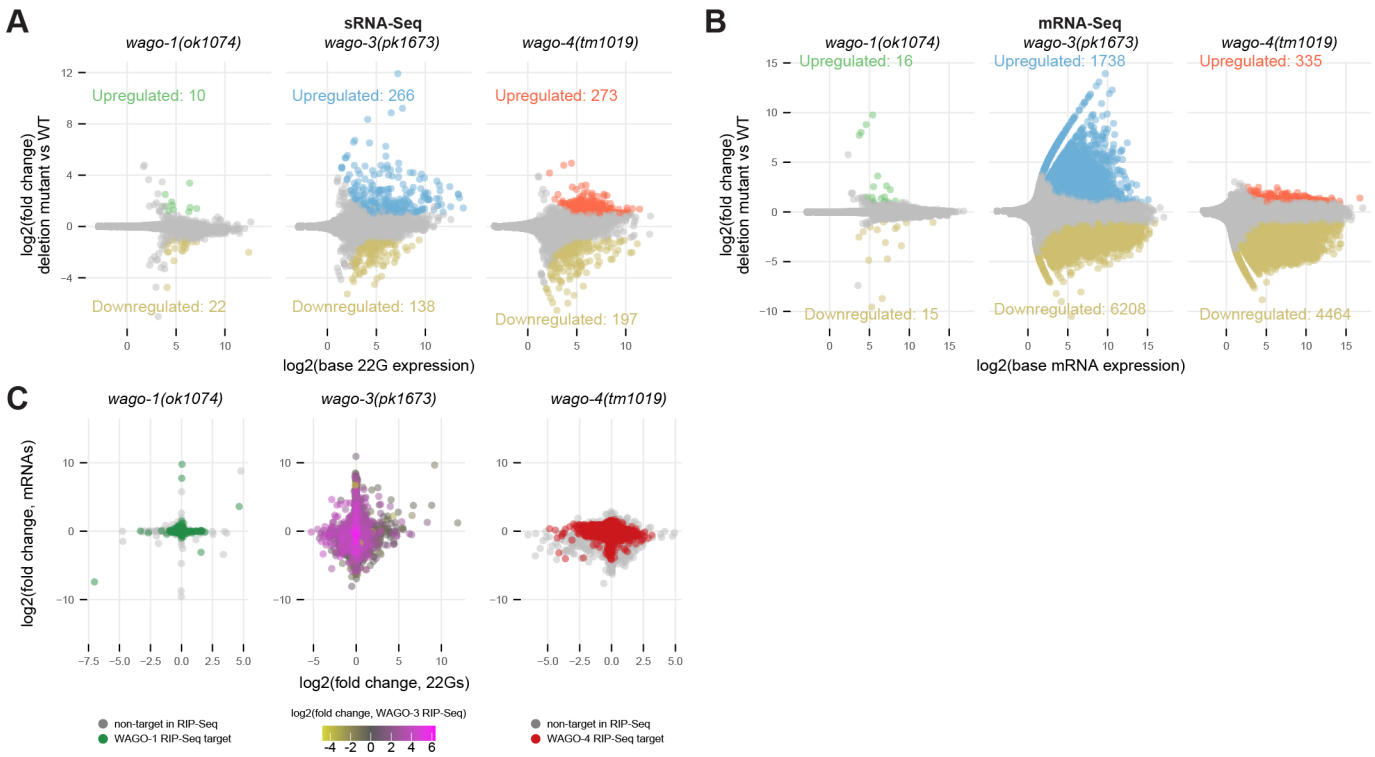


Figure 12: mRNAs and 22G-RNAs regulated upon WAGO deletion in embryos. ***A*** *Scatterplots showing genes with regulated 22G-RNAs.* ***B*** *Scatterplots showing genes with regulated mRNAs.* ***C*** *Scatterplots comparing fold changes of mRNAs to fold changes of 22G-RNAs. For* wago-1(ok1074) *and* wago-4(tm1019)*, colours represent whether or not genes are defined as targets in RIP-Seq. For* wago-3(pk1673)*, colours represent the fold change of IP versus input of gravid adult* gfp::3xflag::wago-3(xf119).

WAGO-1, -3, AND -4 IMPACT DIFFERENT SUBSET OF GENES

As previously stated, it has been shown that WAGO-1, WAGO-3, and WAGO-4 all bind different targets. Here, we have also shown that loss of either of those WAGOs affect at different subset of 22G-RNAs and yet another subset of mRNAs. We wanted to finally show whether there were any correlations between the 22G-RNAs and the mRNAs regulated in either of our three mutants at any lifestage.

We saw that the genes with regulated 22G-RNAs did differ between mutants (**Figure 14A**). Genes with regulated 22G-RNAs were largely specific for both mutant and lifestage, although there were substantial overlaps between different lifestages within the same mutant. There was also some overlap between genes with 22G-RNAs regulated in the same direction in *wago-1(ok1074)* and *wago-4(tm1019)* but at different lifestages. Smaller overlaps existed between genes with 22Gs regulated in either opposing or the same direction in the same lifestage of *wago-3(pk1673)* and *wago-4(tm1019)* (**Figure 14A**).

On mRNA level, there was a large overlap between genes upregulated in the *wago-1(ok1074)* gravid adult and genes upregulated in *wago-4(tm1019)* L4 worms (**Figure 14B**). Apart from that, most of the overlap

between different mutants stemmed from the disproportionate number of genes regulated in the embryo in *wago-3(pk1673)* and *wago-4(tm1019)* (**Figure 14B**).

We can conclude that loss of WAGO-1, -3, or -4 have distinct effects on the RNA landscape.

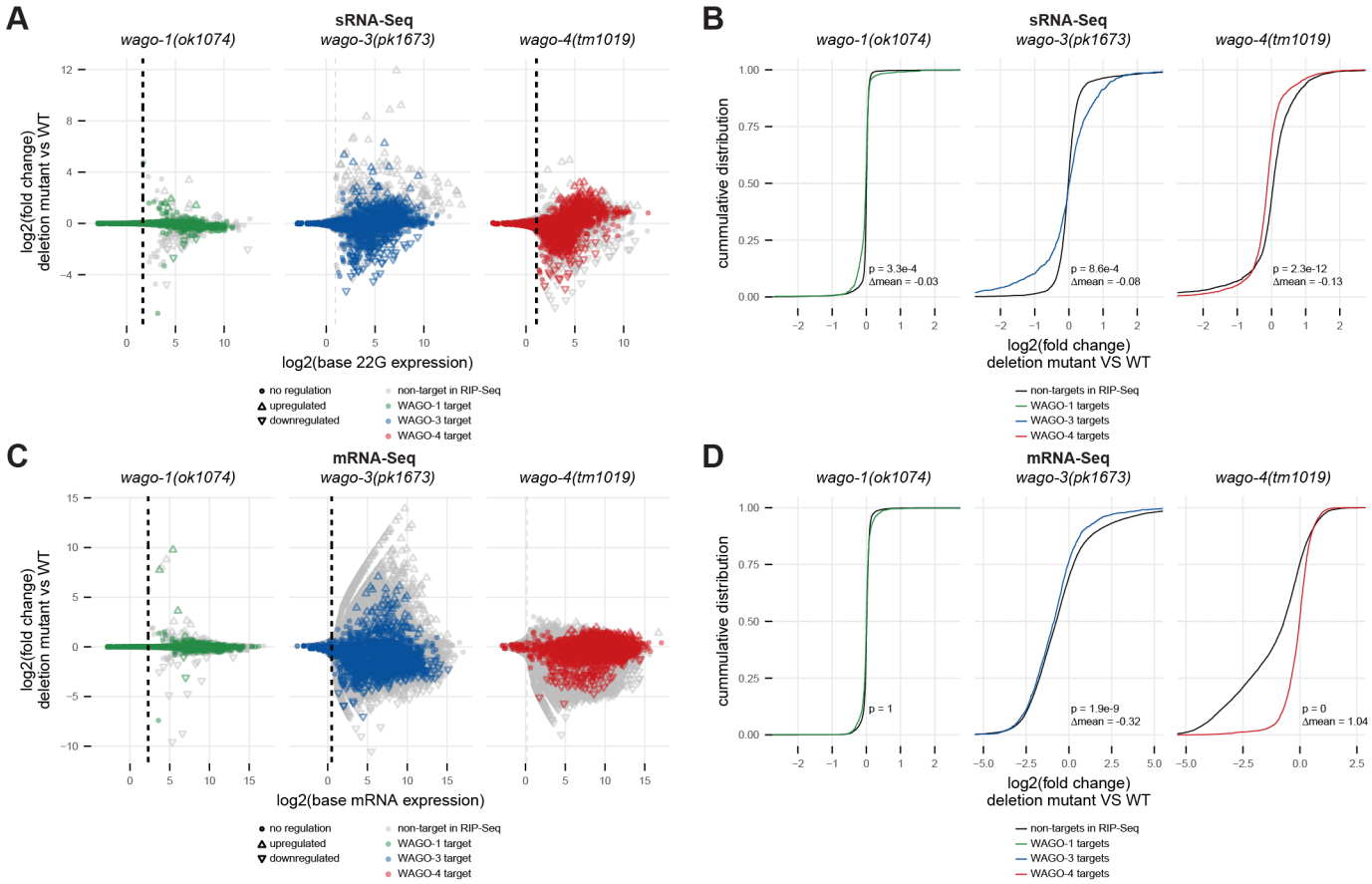
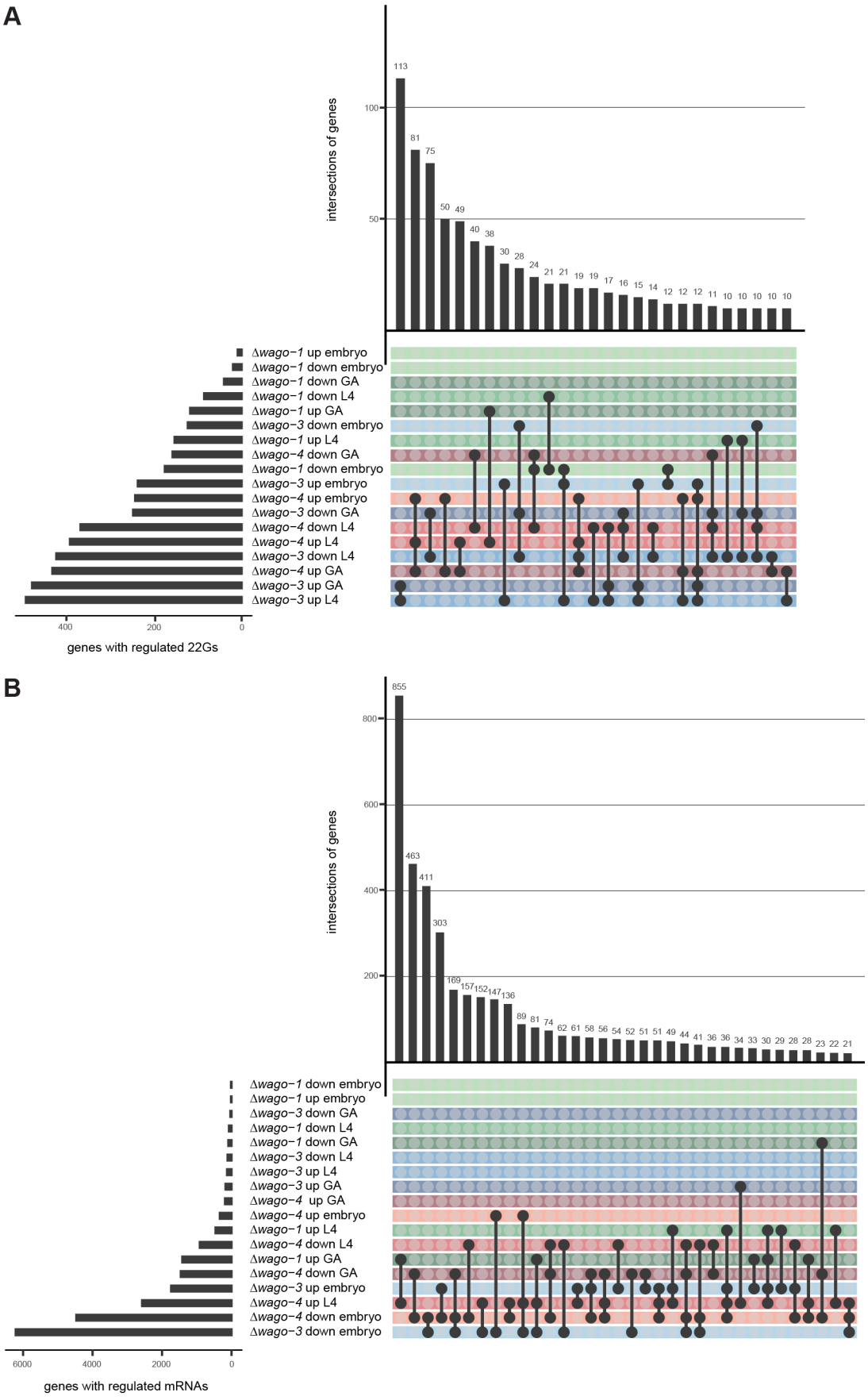


Figure 13: mRNA and 22G-RNA regulation in WAGO deletion mutant embryos compared to known targets. A *Scatterplots showing genes with regulated 22G-RNAs. Colours represent whether the gene has been found as a target in any hermaphrodite RIP-Seq experiment. Dotted line represents the cutoff used for quantification.* ***B*** *CDF plot quantifying panel A. X-axis has been shortened to show the relevant area and Bonferroni-adjusted p-values from one-way ANOVA indicated along with difference in mean, when* $p \leq 0.01$. ***C*** *Scatterplots showing genes with regulated mRNAs. For* wago-1(ok1074) *and* wago-4(tm1019)*. Colours represent whether the gene has been found as a target in any hermaphrodite RIP-Seq experiment. Dotted line represents the cutoff used for quantification.* ***D*** *CDF plot quantifying panel C. X-axis has been shortened to show the relevant area and Bonferroni-adjusted p-values from one-way ANOVA indicated along with difference in mean, when* $p \leq 0.01$.



Legend on next page

Figure 14: Overlap of genes regulated in different mutants and different lifestages. A *UpSet plot showing overlap of all up- or downregulated 22G-RNAs ("up"/"down") in* wago-1(ok1074) *("*Δwago-1*"),* wago-3(pk1673) *("*Δwago-3*"), or* wago-4(tm1019) *("*Δwago-4*") in embryo, L4 larvae, or gravid adults ("GA"). Unique values and overlaps below 10 genes have been omitted.* ***B*** *UpSet plot showing overlap of all up- or downregulated mRNAs ("up"/"down") in* wago-1(ok1074) *("*Δwago-1*"),* wago-3(pk1673) *("*Δwago-3*"), or* wago-4(tm1019) *("*Δwago-4*") in embryo, L4 larvae, or gravid adults ("GA"). Unique values and overlaps below 20 genes have been omitted.*

Interconnection between Argonaute pathways

The fact that neither WAGO deletion is lethal would suggest some redundancy between Argonaute pathways. Linkage between Argonaute pathways may be small or near complete, with all Argonautes affecting one another. If the latter is the case, our data is unlikely to provide any further insight into this. If only a few Argonautes act together, however, it would be possible that we could gain insights into this by looking at the targets of other Argonautes in out deletion mutants.

This section will go over the evidence we have for possible interaction partners of WAGO-1 , -3 and -4,

WAGO-1 AND WAGO-3 ARE SPATIALLY, BUT NOT FUNCTIONALLY, LINKED

Because WAGO-1 and WAGO-3 reside in the same granule (Phillips & Updike, 2022) and further co-immunoprecipitate with one another (**Figure 15A** and **B**), we reasoned that these two may show some redundancy. If WAGO-1 is indeed capable of binding WAGO-3 targets in the absence of WAGO-3, we would assume that loss of WAGO-3, which does not cause a loss of WAGO-3 target 22G-RNAs, would cause either a loss of WAGO-1 targets – since WAGO-1 would no longer be able to bind both its own targets *and* WAGO-3 targets – or an upregulation of WAGO-1 itself, or possibly both. The opposite would also be expected to be the case for WAGO-3 in a *wago-1* deletion mutant.

What we saw was that loss of WAGO-3 does not alter the expression level of WAGO-1, neither destabilizing the protein nor upregulating it to counteract the loss (**Figure 15C**). That the WAGO-1 protein itself is not upregulated, does not exclude the possiblity that WAGO-1 may partially counteract the loss of WAGO-3, so in order to see if the two have any redundancy, we investigated what happens with WAGO-3 target genes in *wago-1(ok1074)* and what happens with WAGO-1 target genes in *wago-3(pk1673)*.

In embryos and in gravid adult worms, WAGO-1 and WAGO-3 target genes are largely unaffected by either deletion. Most prominent in L4 larvae, both WAGO-3 target 22G-RNAs are upregulated in *wago-1(ok1074)* (**Figure 15D** and **E**) and WAGO-1 target 22G-RNAs are upregulated in *wago-3(pk1673)* (**Figure 15F** and **G**). However, if the two proteins were mutually redundant, we would expect to see a downregulation of the native target genes to accommodate the uptake of 22G-RNAs native to the other WAGO. With the added

observation that WAGO-1 target 22G-RNAs were also upregulated in L4 larvae in *wago-1(ok1074)*, we find it unlikely that WAGO-1 and WAGO-3 are mutually exclusive or that their targets are determined in such a way that one can accommodate loss of the other.

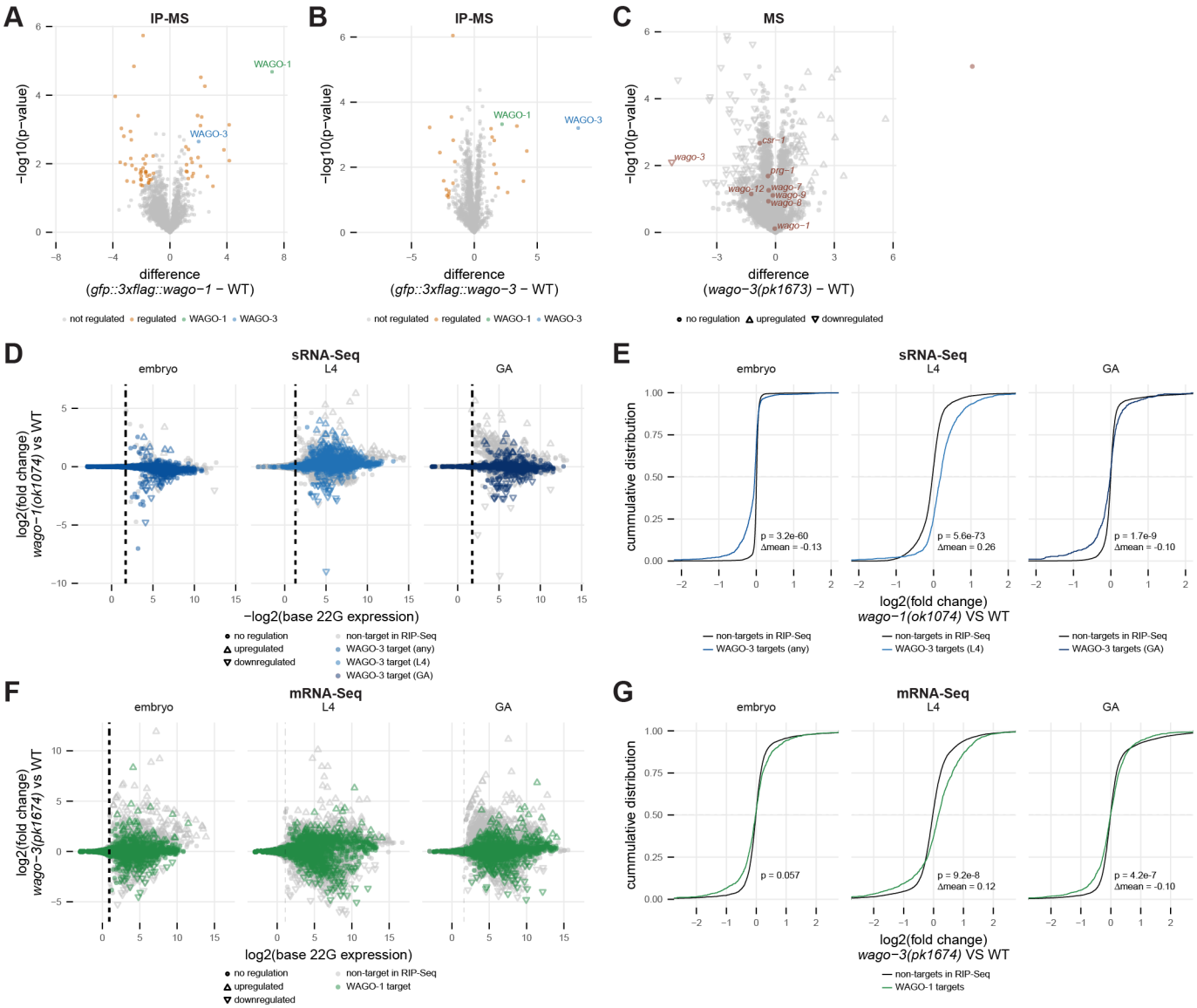


Figure 15: Correlations between WAGO-1 and WAGO-3. ***A*** *Scatterplot of mass spectrometry data from L4* gfp::3xflag::wago-1(jmc219). ***B*** *Scatterplot of mass spectrometry data from L4* gfp::3xflag::wago-3(xf119). ***C*** *Scatterplot of IP-MS data from L4* wago-3(pk1673). *All found Argonautes are marked.* ***D*** *Scatterplot showing 22G-RNAs regulated in* wago-1(ok1074), *marked for WAGO-3 RIP-Seq targets. Dotted line represents the cutoff used for quantification.* ***E*** *CDF plot quantifying panel D. X-axis has been shortened to show the relevant area and Bonferroni-adjusted p-values from one-way ANOVA indicated along with difference in mean, when p* ≤ 0.01. ***F*** *Scatterplot showing 22G-RNAs regulated in* wago-3(pk1673), *marked for WAGO-1 RIP-Seq targets. Dotted line represents the cutoff used for quantification.* ***G*** *CDF plot quantifying panel G. X-axis has been shortened to show the relevant area and Bonferroni-adjusted p-values from one-way ANOVA indicated along with difference in mean, when p* ≤ 0.01.

LOSS OF WAGO-3 MAY CAUSES UPREGULATION OF ERGO-1-RELATED SMALL RNA PATHWAYS

Each gene transported to a mutator focus may produce 22G-RNAs taken up by more than one Argonaute, and since the mechanisms by which 22G-RNAs are assigned to certain Argonautes is unknown, we also don't know the extent to which Argonautes may interact with one another or how they may be capable of counteracting the loss of one another. Seeing no evidence to support a direct link between WAGO-1 and WAGO-3, we went on to investigate possible other Argonaute pathways with which each WAGO may interact separately, starting with WAGO-3. First, because different Argonautes have different signatures of binding their targets, with 22G-RNAs of one Argonuate for instance being produced primarily at the transcription end site while 22G-RNAs of another Argonaute are produced along the entire gene body, we reasoned that we might be able to use the binding signatures to determine which Argonaute(s) take up WAGO-3 targets in the absence of WAGO-3 itself.

As shown in Chapter I (**Extended Figure 2**), WAGO-3 predominantly bind 22G-RNAs that map near the transcription end site of their respective genes, and this has been corroborated by others (Seroussi et al., 2023). If other Argonaute proteins with other signatures take up WAGO-3 targets in *wago-3(pk1673)*, we would be able to see this by looking at the gene coverage of 22G-RNAs in wildtype and *wago-3(pk1673)*. The metagene signature of 22G-RNAs mapping to WAGO-3 targets did not change in any lifestage, however (**Figure 16A**). Nonetheless, we did note that the signature we found for 22G-RNAs mapping to WAGO-3 target genes didn't match the signature of 22G-RNAs bound to WAGO-3 fully, with a less sharp peak at the transcription end site and more mapping to the whole gene body, suggesting that WAGO-3 may not be the only Argonaute binding WAGO-3 target 22G-RNAs in the native state.

We also investigated the signature of 22G-RNAs regulated by deletion of WAGO-3. Here, we found that downregulated 22G-RNAs *did* match the signature of WAGO-3-bound 22G-RNAs, especially in the post-embryonic lifestages, where a noticeable peak was found near the transcription end sites (**Figure 16B**). The 22G-RNAs that were upregulated upon WAGO-3 deletion were more scattered along the entire gene body of their corresponding genes (**Figure 16B**), which more resemble the signature of the 22G-RNAs mapping to WAGO-3 target genes in both wildtype and *wago-3(pk1673)*. This signature is akin to the one of WAGO-8 and WAGO-12 (Seroussi et al., 2023).

Together, these analyses suggest that the distribution of 22G-RNAs over different Argonautes may shift upon loss of WAGO-3. The germline may lose a subset of 22G-RNAs that *only* bind WAGO-3 or Argonautes with similar signatures and may gain 22G-RNAs binding to WAGO-8 and/or WAGO-12, implicating a link between the pathways of WAGO-3, -8, and -12. We propose a shift in 22G-RNA distribution where WAGO-8 and WAGO-12 start binding more 22G-RNAs from annotated WAGO-3 target genes in the absence of WAGO-3.

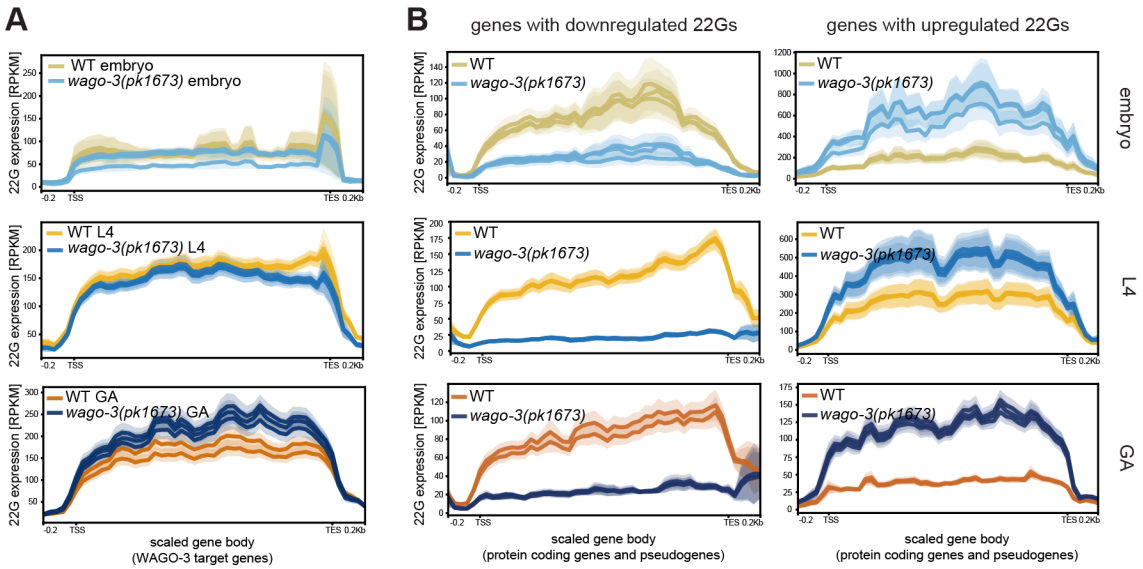
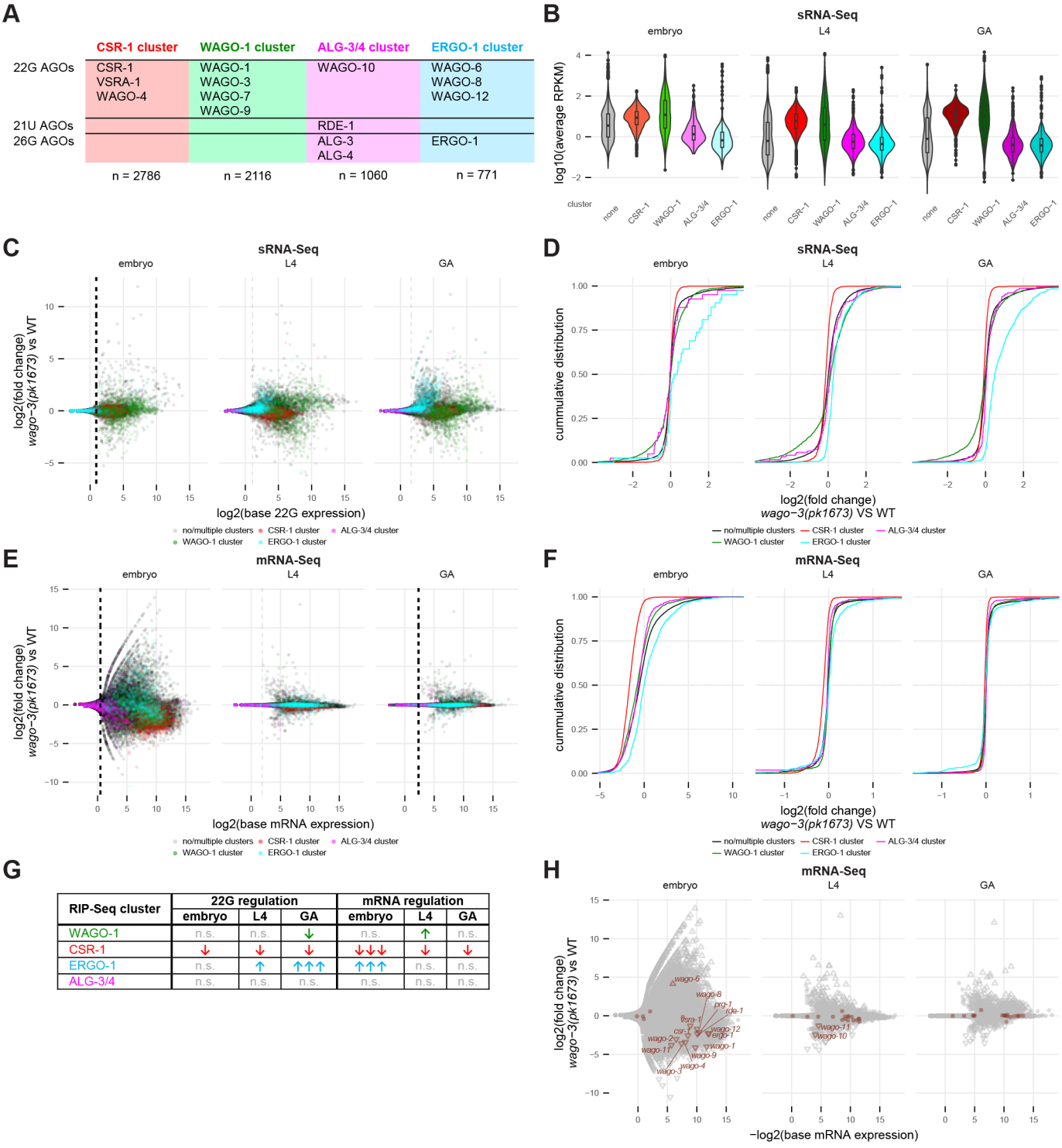


Figure 16: 22G-RNA mapping signatures in* wago-3(pk1673).** ***A *Average RPKMs of 22G-RNAs along the gene body of their corresponding genes (protein coding or pseudogenes) for genes found to be targets of WAGO-3 in L4 larvae (middle panel), gravid adult worms (bottom panel), or in either (top panel) in wildtype (WT) and* wago-3(pk1673) *embryo, L4 larvae, and gravid adult (GA) worms.* ***B*** *Average RPKMs of 22G-RNAs along the gene body of their corresponding genes (protein coding or pseudogenes) for genes up- or downregulated in* wago-3(pk1673) *embryo, L4 larvae, and gravid adult (GA) worms.*

To further answer the question of which Argonautes share redundancy with WAGO-3 and can work to counter its loss, we looked to the work of Seroussi et al. (Seroussi et al., 2023). They found that, based on RIP-Seq experiments, the Argonautes of *C. elegans* cluster into four subgroups, as summarised in **Figure 17A**. In the wildtype, 22G-RNAs from the WAGO-1 and CSR-1 clusters are more highly expressed than those from the ALG-3/4 and ERGO-1 clusters, with particularly low expression in the embryos (**Figure 17B**). Due to the low expression of ERGO-1 cluster 22G-RNAs, a large part of these fell below our shrinkage threshold (**Figure 17C**). Nonetheless, we did see a general upregulation of ERGO-1 cluster 22G-RNAs in *wago-3(pk1673)* (**Figure 17D** and **G**), which fits with our observation of upregulation of WAGO-8 and -12, as these both belong to this cluster. We also saw an increase of ERGO-1 cluster mRNAs in the embryo (**Figure 17E-G**), which could suggest that the upregulation of 22G-RNAs from this cluster may be a secondary effect caused by mRNA upregulation.

A

	CSR-1 cluster	WAGO-1 cluster	ALG-3/4 cluster	ERGO-1 cluster
22G AGOs	CSR-1 VSRA-1 WAGO-4	WAGO-1 WAGO-3 WAGO-7 WAGO-9	WAGO-10	WAGO-6 WAGO-8 WAGO-12
21U AGOs			RDE-1	
26G AGOs			ALG-3 ALG-4	ERGO-1
	n = 2786	n = 2116	n = 1060	n = 771



G

RIP-Seq cluster	22G regulation			mRNA regulation		
	embryo	L4	GA	embryo	L4	GA
WAGO-1	n.s.	n.s.	↓	n.s.	↑	n.s.
CSR-1	↓	↓	↓	↓↓↓	↓	↓
ERGO-1	n.s.	↑	↑↑↑	↑↑↑	n.s.	n.s.
ALG-3/4	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.

Figure 17: Correlations between* wago-3(pk1673) *and four predefined Argonaute clusters. ***A*** *Table showing cluster association of different Argonautes as defined by (Seroussi et al., 2023). Small RNA association of each Argonaute is indicated and numbers represent unique genes targeted in each cluster.* ***B*** *Violin plots showing expression of 22G-RNAs belonging to different clusters in wildtype embryo, L4, and gravid adult worms.* ***C*** *Scatterplot showing 22G-RNAs regulated in* wago-3(pk1673)*, marked for unique genes in each cluster, as indicated. Dotted line represents the cutoff used for quantification.* ***D*** *CDF plot quantifying panel C.*

Legend continues on next page

X-axis has been shortened to show the relevant area. ***E*** *Scatterplot showing mRNAs regulated in* wago-3(pk1673)*, marked for unique genes in each cluster, as indicated. Dotted line represents the cutoff used for quantification.* ***F*** *CDF plot quantifying panel E. X-axis has been shortened to show the relevant area.* ***G*** *Table summarizing cluster-wise regulation of 22G-RNAs and mRNAs in* wago-3(pk1673)*. Bonferroni-adjusted p-values from ANOVA test carried out for each group versus all others. ↕ = adjusted p-value ≤ 0.01, |Δ mean| < 0.5, ↕↕↕= adjusted p-value ≤ 0.01, |Δ mean| > 0.5,, n.s. = not significant (adjusted p-value > 0.01).* ***H*** *Scatterplot showing mRNAs regulated in* wago-3(pk1673)*, marked for Argonautes. All regulated Argonautes have been labelled.*

Upregulation of ERGO-1 cluster expression furthermore fit with our observation that we saw an increase in mRNA levels of WAGO-6, another participant in this cluster, in *wago-3(pk1673)* embryos (**Figure 17H**). We didn't see any upregulation of any Argonaute mRNAs in any other lifestage (**Figure 17H**) nor any upregulation of Argonaute peptides in L4 larvae (**Figure 15C**). We also noted a general downregulation of most Argonaute mRNAs in the embryo, including other participants of the ERGO-1 cluster (**Figure 17H**). Nonetheless, our data does suggest some sort of relationship between WAGO-3 and the ERGO-1 pathway.

Apart from ERGO-1 cluster genes, we saw a general downregulation of CSR-1 cluster genes, both on 22G-RNA level and on mRNA level, in *wago-3(pk1673)* (**Figure 17C-G**). This could also indicate a relationship between WAGO-3 and CSR-1 and/or WAGO-4. Although WAGO-3 is involved in paternal inheritance (Chapter I) as is ALG-3/4 (Conine et al., 2010), loss of WAGO-3 did not seem to affect ALG-3/4 cluster genes (**Figure 17C-G**), although there was a loss of WAGO-10 mRNAs at L4 (**Figure 17H**).

We conclude that there must be a link between WAGO-3 and the ERGO-1 cluster Argonautes WAGO-6, -8, and -12. Further experiments will be needed to elucidate the exact nature of this link.

Loss WAGO-1 or WAGO-4 causes lifestage-specific upregulation ALG-3/4 cluster Argonautes

We next turned our focus to the interaction partners of WAGO-1 and WAGO-4, starting with investigating the 22G-RNA binding signatures of their targets and of 22G-RNAs regulated by their loss. WAGO-1-bound 22G-RNAs have been shown to map predominantly near the transcription start site of their target genes (Gu et al., 2009; Seroussi et al., 2023) whilst WAGO-4-bound 22G-RNAs map predominantly near either end of their target genes (Seroussi et al., 2023).

As was the case for WAGO-3 target genes, 22G-RNAs mapping to WAGO-1 target genes in our samples did not match the general signature of WAGO-1-bound 22G-RNAs and the signature was unchanged upon deletion of *wago-1*. Instead, there was an increase of 22G-RNAs near the transcription end site, indication that both WAGO-3 and WAGO-1 target 22G-RNAs may be bound by other Argonautes in the native state (**Figure 18A**). For WAGO-4-bound 22G-RNAs, there was also no change in signature upon *wago-4* deletion, but the overall signature did match the one established for WAGO-4, with some preference for the

transcription end site (**Figure 18C**). This also matches the signature for CSR-1 (Seroussi et al., 2023), which is known to share targets with WAGO-4.

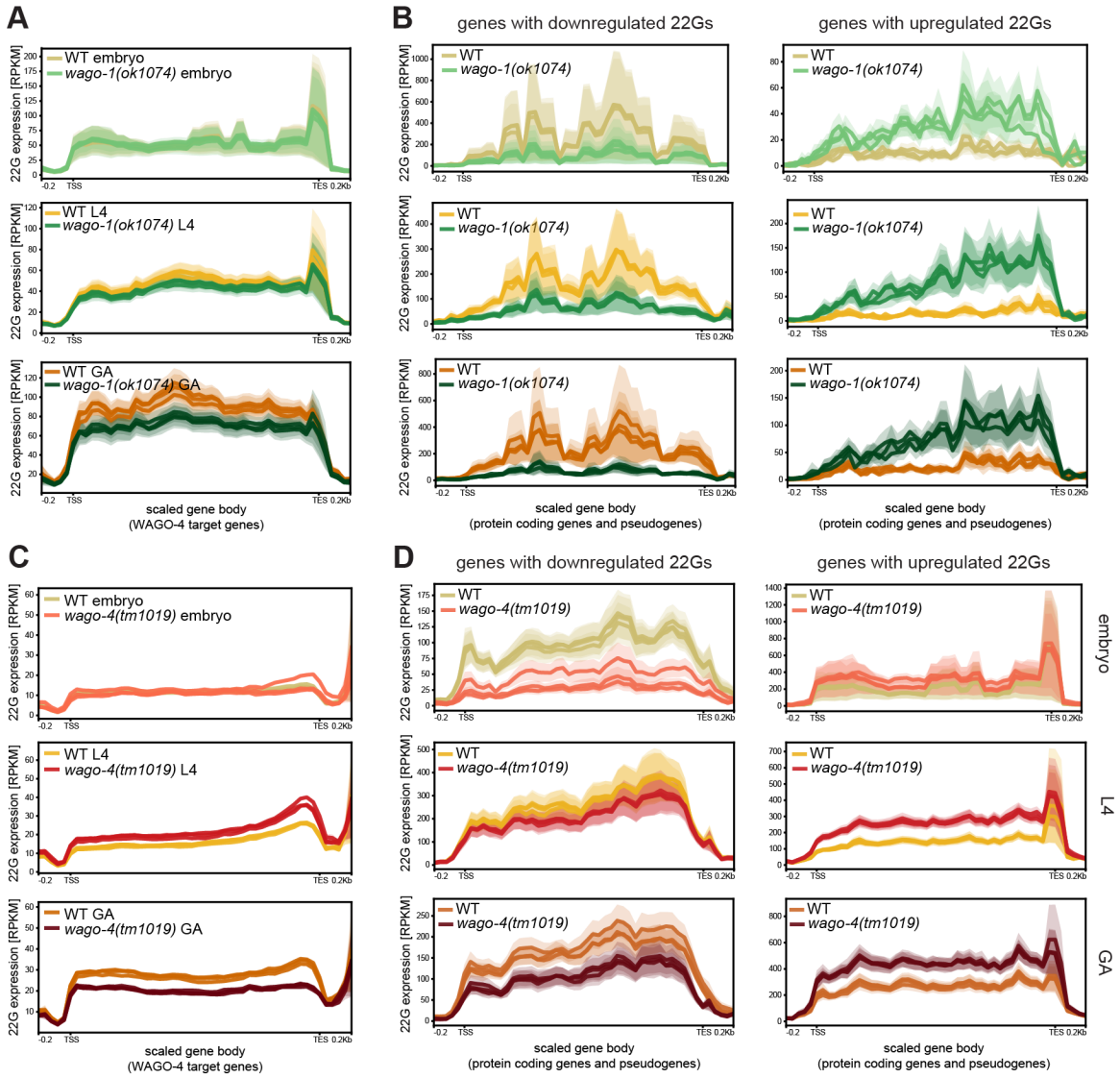


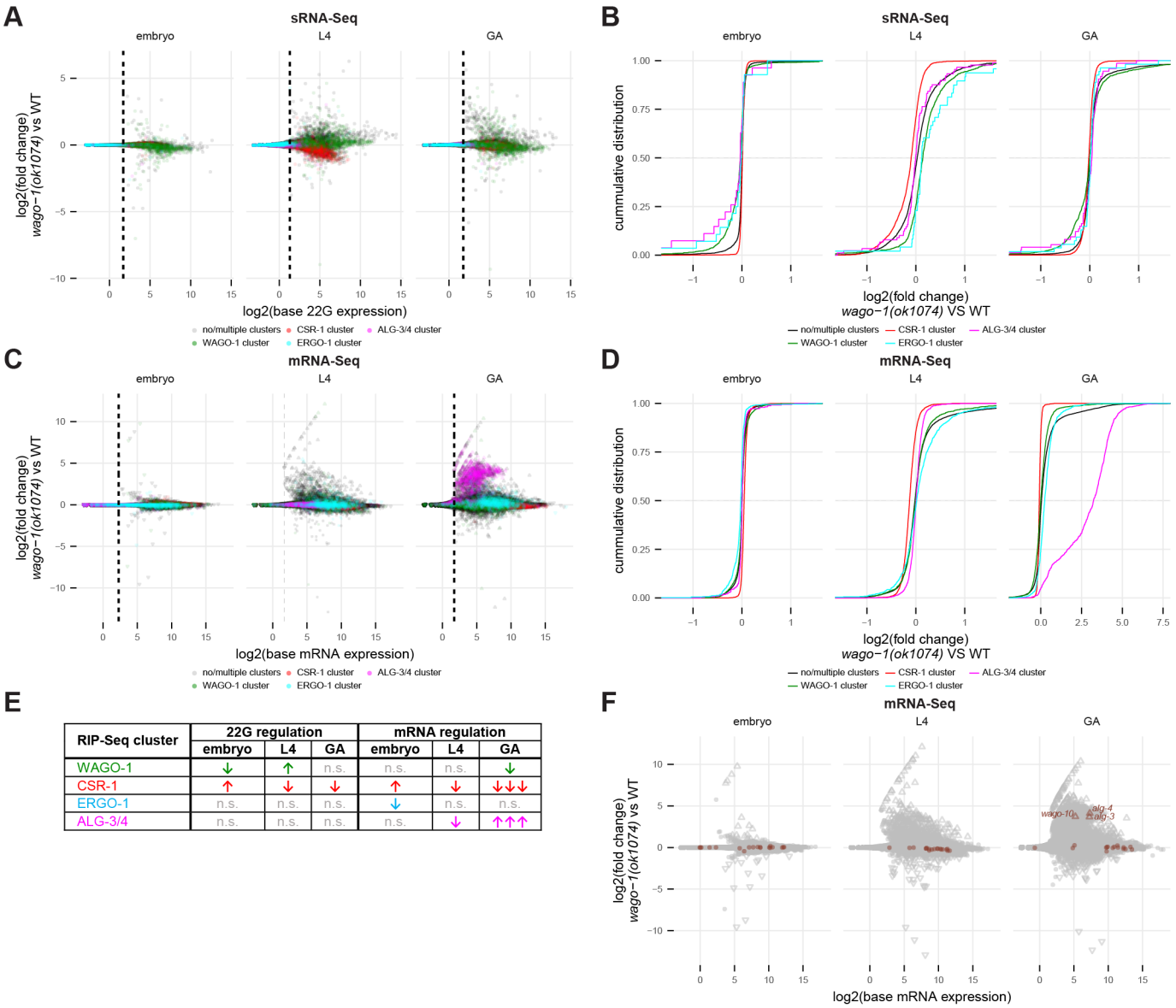
Figure 18: 22G-RNAs mapping signatures in* wago-1(ok1074) *and* wago-4(tm1019).** ***A *Average RPKMs of 22G-RNAs along the gene body of their corresponding genes (protein coding or pseudogenes) for genes found to be targets in WAGO-1 RIP-Seq in* wago-1(ok1074) *and wildtype.* ***B*** *Average RPKMs of 22G-RNAs along the gene body of their corresponding genes (protein coding or pseudogenes) for genes up- or downregulated in* wago-1(ok1074) *embryo, L4, and gravid adult (GA) worms.* ***C*** *Average RPKMs of 22G-RNAs along the gene body of their corresponding genes (protein coding or pseudogenes) for genes found to be targets in WAGO-4 RIP-Seq in* wago-4(tm1019) *and wildtype.* ***B*** *Average RPKMs of 22G-RNAs along the gene body of their corresponding genes (protein coding or pseudogenes) for genes up- or downregulated in* wago-4(tm1019) *embryo, L4, and gravid adult (GA) worms.*

Genes that lost 22G-RNAs in *wago-1(ok1074)* did so along the entire gene body. In contrast, genes that gained 22G-RNAs gained them with a bias at the transcription end site (**Figure 18B**). This is opposite to our observations in *wago-3(pk1673)*, suggesting that WAGO-1 and WAGO-3 could act in opposing manners. 22G-RNAs mapping to genes with regulated 22G-RNA levels in *wago-4(tm1019)* were predominantly present near the transcription end sites for both up- and downregulated 22G-RNAs (**Figure 18D**). This feature was less prevalent for 22G-RNAs downregulated in the embryo. This could implicate several Argonautes, but we cannot distinguish between different Argonautes interconnected with WAGO-4 based on this data.

Seeing no indication of which other Argonautes WAGO-1 or WAGO-4 may affect or interact with based on 22G-RNA signatures, we next went on to look at regulation of Argonaute clusters, starting with the situation in *wago-1(ok1074)*. As shown in **Figure 7A** and **B**, loss of WAGO-1 causes a slight but significant upregulation of WAGO-1-bound 22G-RNAs. It also causes upregulation of WAGO-3-bound 22G-RNAs, as shown in **Figure D** and **E**, with WAGO-3 belonging to the same cluster as WAGO-1 (**Figure 17A**). It is therefore unsurprising, that WAGO-1 cluster 22G-RNAs were generally upregulated in *wago-1(ok1074)* (**Figure 19A-B** and **E**). In L4 larvae, CSR-1 cluster genes were downregulated both at 22G-RNA level and at mRNA level (**Figure 19A-E**). Interestingly, regulation happened in the same direction at mRNA and 22G-RNA level but in opposing directions in embryos or in L4/gravid adult worms (**Figure 19E**), emphasizing the importance of taking lifestage into account when defining Argonaute functions.

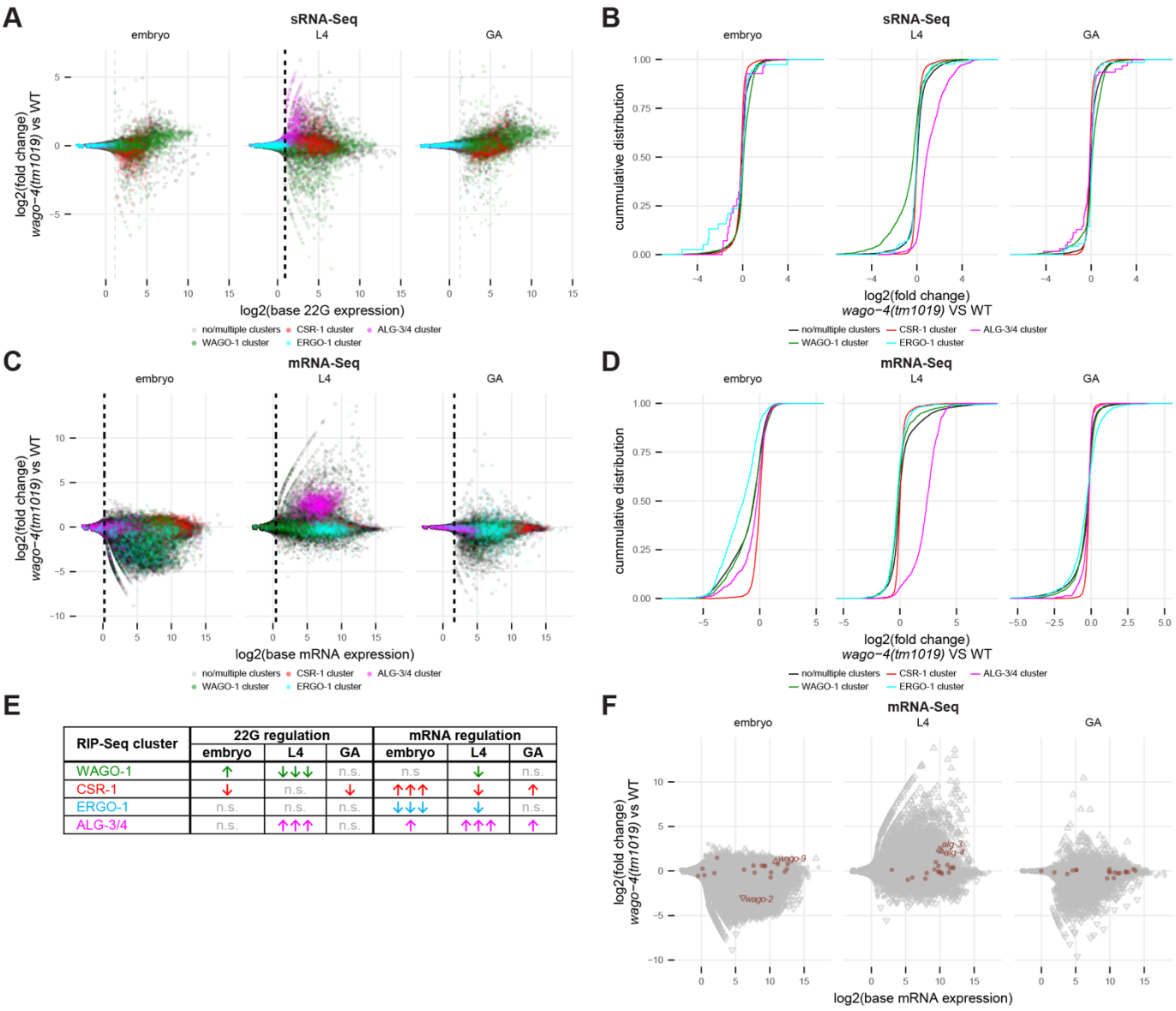
All of the differences mentioned above, however, were relatively small. Much more striking was the large upregulation of ALG-3/4 cluster mRNAs in gravid adult *wago-1(ok1074)* (**Figure 19C-E**), which was accompanied by an mRNA level upregulation of ALG-3 and -4 themselves along with WAGO-10, the only 22G-RNA-associating Argonaute in the ALG-3/4 cluster (**Figure 19F**). Equally as striking was the upregulation of the same cluster in *wago-4(tm1019)*, both at 22G-RNA level and at mRNA level, but at a different lifestage; namely in L4 larvae (**Figure 20A-E**). This was also accompanied by upregulation of ALG-3/4 mRNAs in the same lifestage (**Figure 20F**). This suggests a connection between WAGO-1 and WAGO-4, separated in time.

Also supporting a lifestage specific connection between WAGO-1 and WAGO-4 is the fact that *wago-4(tm1019)* causes regulation of the WAGO-1 cluster and the CSR-1 cluster in directions opposite of *wago-1(ok1074)* (**Figure 19E** and **Figure 20E**). In *wago-4(tm1019)*, regulation of WAGO-1 cluster 22G-RNAs also happened in different directions at different lifestages, as was the case in *wago-1(ok1074).* In the embryo, WAGO-1 cluster 22G-RNAs were upregulated, in L4 larvae they were downregulated. There was a further downregulation of ERGO-1 cluster mRNAs in the embryo (**Figure 20C-E**), which opposes *wago-3(pk1673)* regulation (**Figure 17E-G**). The connection between WAGO-1 and WAGO-4 at different lifestages is further supported by our previous observation of overlap between genes upregulated in *wago-1(ok1074)* gravid adult and genes upregulated in *wago-4(tm1019)* L4 larvae (**Figure 14B**).



RIP-Seq cluster	22G regulation			mRNA regulation		
	embryo	L4	GA	embryo	L4	GA
WAGO-1	↓	↑	n.s.	n.s.	n.s.	↓
CSR-1	↑	↓	↓	↑	↓	↓↓↓
ERGO-1	n.s.	n.s.	n.s.	↓	n.s.	n.s.
ALG-3/4	n.s.	n.s.	n.s.	n.s.	↓	↑↑↑

Figure 19: Argonaute cluster regulation in* wago-1(ok1074).** ***A *Scatterplot showing genes with regulated 22G-RNAs. Dotted line represents the cutoff used for quantification. Colours represent clusters as defined by (Seroussi et al., 2023) with genes attributed to more than one cluster grouped together with genes attributed to zero clusters.* ***B*** *CDF plot quantifying panel A. X-axis has been shortened to show the relevant area.* ***C*** *Scatterplot showing genes with regulated mRNAs. Dotted line represents the cutoff used for quantification. Colours represent clusters as defined by (Seroussi et al., 2023) with genes attributed to more than one cluster grouped together with genes attributed to zero clusters.* ***D*** *CDF plot quantifying panel C. X-axis has been shortened to show the relevant area.* ***E*** *Table summarizing cluster-wise regulation of 22G-RNAs and mRNAs in* wago-1(ok1074)*. Bonferroni-adjusted p-values from ANOVA test carried out for each group versus all others. ↕ = adjusted p-value ≤0.01, |Δ mean| < 0.5, ↕↕↕= adjusted p-value ≤0.01, |Δ mean| > 0.5,, n.s. = not significant (adjusted p-value > 0.01).* ***F*** *Scatterplot showing mRNAs regulated in* wago-1(ok1074)*, marked for Argonautes. All regulated Argonautes have been labelled.*



RIP-Seq cluster	22G regulation			mRNA regulation		
	embryo	L4	GA	embryo	L4	GA
WAGO-1	↑	↓↓↓	n.s.	n.s	↓	n.s.
CSR-1	↓	n.s.	↓	↑↑↑	↓	↑
ERGO-1	n.s.	n.s.	n.s.	↓↓↓	↓	n.s.
ALG-3/4	n.s.	↑↑↑	n.s.	↑	↑↑↑	↑

Figure 20: Argonaute cluster regulation in* wago-4(tm1019).** ***A *Scatterplot showing genes with regulated 22G-RNAs. Dotted line represents the cutoff used for quantification. Colours represent clusters as defined by (Seroussi et al., 2023) with genes attributed to more than one cluster grouped together with genes attributed to zero clusters.* ***B*** *CDF plot quantifying panel A. X-axis has been shortened to show the relevant area.* ***C*** *Scatterplot showing genes with regulated mRNAs. Dotted line represents the cutoff used for quantification. Colours represent clusters as defined by (Seroussi et al., 2023) with genes attributed to more than one cluster grouped together with genes attribu4ed to zero clusters.* ***D*** *CDF plot quantifying panel C. X-axis has been shortened to show the relevant area.* ***E*** *Table summarizing cluster-wise regulation of 22G-RNAs and mRNAs in* wago-4(tm1019)*. Bonferroni-adjusted p-values from ANOVA test carried out for each group versus all others. ↕ = adjusted p-value ≤0.01, |Δ mean| < 0.5, ↕↕↕= adjusted p-value ≤0.01, |Δ mean| > 0.5,, n.s. = not significant (adjusted p-value > 0.01).* ***F*** *Scatterplot showing mRNAs regulated in* wago-4(tm1019)*, marked for Argonaute proteins. All regulated Argonaute proteins have been labelled.*

We conclude there is a direct link between WAGO-1 and WAGO-4 and that this connects to ALG-3/4. The linkage between WAGO-1/4 and ALG-3/4 shifts somewhere between L4 larvae and gravid adult worms, with WAGO-4 having a function during spermatogenesis and WAGO-1 having a function later, either during oogenesis or embryogenesis. WAGO-10 may also be connected to this.

Can misregulation caused by loss of WAGO-4 be remembered when the protein is reintroduced?

mRNA (mis-)regulation caused by loss of WAGO-4 is remembered transgenerationally

Our final question arose from the observation that WAGO-4 is involved in epigenetic memory (Xu et al., 2018). We were interested to see whether the global changes to 22G-RNAs and to mRNA level could persist in a worm that had once lacked a functional copy of WAGO-4 but was currently wildtype. To answer that question, we included wildtype worms that had once been *wago-4(tm1019)* in our sequencing experiments (**Figure 4C** and **D** and **Figure 5**). This strain was generated by outcrossing *wago-4(tm1019)* to wildtype worms from both parental sides and selecting worms that were wildtype for the *wago-4* gene. Worms sequenced were the fifth generation homozygous, and in order to limit our analysis, we only sequenced gravid adult worms.

On 22G-RNA level, the recovered wildtype differed only slightly from the original wildtype (**Figure 21B**) and not nearly to the extent that the *wago-4(tm1019)* mutant differed from the original wildtype (**Figure 21A**), with 22G-RNAs against slightly more than 600 genes affected in the mutant compared to only roughly 80 genes in the recovered wildtype. Nonetheless, the overlap between the genes whose 22G-RNA levels were downregulated in either case was quite large, with 23 genes, corresponding to 59% of the genes with downregulated 22G-RNAs in the recovered wildtype, present in both cases (**Figure 21C**). Furthermore, all genes with downregulated 22G-RNAs in the recovered wildtype had fold changes close to or below zero in the comparison between *wago-4(tm1019)* and the original wildtype, even if the p-values were not significant (**Figure 21D**).Genes whose 22G-RNAs were upregulated in the recovered wildtype were less well conserved, with some even being downregulated in the mutant (**Figure 21C** and **D**), suggesting that there is some memory of 22G-RNA downregulation but 22G-RNA upregulation may be more sporadic.

Comparably more genes were regulated at mRNA level in both the *wago-4(tm1019)* mutant (**Figure 21E**) and in the recovered wildtype (**Figure 21F**), with about 1500 genes downregulated in *wago-4(tm1019)* and 850 genes downregulated in the recovered wildtype. The overlap of mRNAs either down- or upregulated in either case was large; 80% or above from the side of the recovered wildtype, and more strikingly, every gene

regulated in the recovered wildtype tended towards regulation in the same direction in *wago-4(tm1019)*, when not taking statistical significance into account (**Figure 21G**). This indicates that mRNA changes caused by loss of WAGO-4 are remembered at least for five generation after being reintroduced to WAGO-4.
Although we established that little-to-no correlation exists between genes with regulated mRNA levels and genes with regulated 22G-RNAs in any of our deletion mutants, we did check to see if there was any overlap between regulated 22G-RNAs and regulated mRNAs in our recovered wildtype, reasoning that changes that persisted after reintroduction of WAGO-4 might have higher correlation than changes which the worm could revert to wildtype within just five generations. This did not seem to be the case, however. Almost none of the genes that were regulated at mRNA level in the recovered wildtype also showed regulation at 22G-RNA level with only three exceptions. *asah-1*, an acylsphingosine amidohydrolase was downregulated at mRNA level but had upregulation of 22G-RNAs. This could be a direct effect of 22G-RNA silencing, although the role of the gene and its relation to WAGO-4 remains elusive. Two genes, *lea-1* and *skpo-2* were downregulated at both levels. These are both related to different types of stress; the former involved in response to hyperosmosis and heat and the latter related to oxidative stress. This could indicate a hyper susceptibility to stress sustained in the recovered wildtype, although it is unlikely to be a direct effect of silencing, wince the 22G-RNAs were also downregulated.

Finally, we were interested in seeing if there were any common features of the genes that were regulated in the recovered wildtype, so we used GO-term analysis to look for function of the affected genes. Due to the small sample size, we didn't find any significantly upregulated GO-terms among the genes whose 22G-RNA levels were affected in the recovered wildtype. On the other hand, the analysis of genes regulated at mRNA level revealed downregulation of genes related to neurogenesis and cell morphogenesis as well as transcription regulatory pathways (**Figure 21I**). Genes upregulated at mRNA level in the recovered wildtype were related to immune and defence responses and involved genes with different types of post-translational modification activities (**Figure 21J**). This suggests that differences in transcription can persist in the recovered wildtype due to changes both in the transcriptional and in the post-transcriptional proteome. It furthermore suggests a decreased capability of cells to differentiate into neurons along with an increased immune response, which could indicate a general change in the cellular landscape of worms having once lacked a functional WAGO-4.

In conclusion, mRNA misregulation caused by lack of WAGO-4 is remembered for at least five generation after reintroduction of the protein even if the 22G-RNA landscape is capable of reverting to a wildtype state earlier.

Figure 21: Features of gravid adult wildtype worms recovered from* wago-4(tm1019).** ***A *Scatterplot showing 22G-RNAs in* wago-4(tm1019) *versus a kept wildtype.* ***B*** *Scatterplot showing 22G-RNAs in a wildtype recovered from* wago-4(tm1019) *versus a kept wildtype.* ***C*** *Venn diagram showing all genes with 22G-RNAs regulated in either direction in panel A or B.* ***D*** *Scatterplot showing 22G-RNAs in* wago-4(tm1019) *versus a kept wildtype. Colours represent genes with regulated 22G-RNAs in the recovered versus kept wildtype.* ***E*** *Scatterplot showing mRNAs in* wago-4(tm1019) *versus a kept wildtype.* ***F*** *Scatterplot showing mRNAs in a wildtype recovered from* wago-4(tm1019) *versus a kept wildtype.* ***G*** *Venn diagram showing all genes with mRNAs regulated in either direction in panel E or F.* ***H*** *Scatterplot showing mRNAs in* wago-4(tm1019) *versus a kept wildtype. Colours represent genes with regulated mRNAs in the recovered versus kept wildtype.* ***I*** *Bar plot showing the ten most numerous 'molecular function' GO-terms and the ten most numerous 'biological*

Legend continues on next page

process' GO-terms all with adjusted p-values below 0.01 found in the genes with downregulated mRNA levels in the recovered versus kept wildtype. ***J*** *Bar plot showing the ten most numerous 'molecular function' GO-terms and the ten most numerous 'biological process' GO-terms all with adjusted p-values below 0.01 found in the genes with upregulated mRNA levels in the recovered versus kept wildtype.*

Discussion and conclusions

WAGO-3 responds to exo-siRNA, 21U-RNAs, and the maternal 26G-RNA pathway

WAGO-3 and WAGO-4 have been shown to be mutually necessary for exo-siRNAi (Gu et al., 2009), although WAGO-3 does not act in the first generation; rather, it is responsible for inheritance both via the sperm and the oocyte (Schreier et al., 2024). In **Chapter I**, we also showed that WAGO-3 is responsible for TEI of the 21U-RNA-mediated response (**Chapter I**, **Extended Figure 1**), and that this is paternally inherited (**Chapter I**, **Figure 2**). We did not find a maternal inheritance effect of WAGO-3, although our assay would not be capable of detecting such a phenomenon if other Argonautes acted in the same pathway (**Chapter I**, **Figure 2**). Rather, we showed that WAGO-3 is the only secondary Argonaute involved in paternal inheritance of 21U-RNA-mediated RNAe. WAGO-3 has also been demonstrated to be the only WAGO responsible for paternal inheritance of the exo-siRNA response (Schreier et al., 2024). Since WAGO-3 acts downstream of either 21U-RNA or exo-siRNA pathway, it would be unlikely for its inheritance to differ dependent on trigger; rather, any WAGO-3 response would be inherited both maternally and paternally, and the maternal inheritance acts in concert with other Argonaute proteins, including or limited to WAGO-4.

As for a possible link between WAGO-3 and the 26G-RNA pathway, we did not find any suggestion of a paternal linkage, with no regulation of 22G-RNAs or mRNAs from the Argonaute cluster including the sperm and spermatogenesis expressed Argonautes ALG-3/4 (**Figure 17C-G**). No connection between ALG-3/4 and WAGO-3 has been reported, and our findings in **Chapter I** showed that only PRG-1 and the 21U-RNA response was involved in the paternal RNAe response mediated by WAGO-3, since loss of PRG-1 was sufficient to allow misloading of WAGO-3 in mutants that lacked paternally derived memory of WAGO-3-associated 22G-RNAs. Nonetheless, our data does suggest that WAGO-3 interacts with the maternal 26G-RNA pathway, even if 26G-RNAs do not elicit an inheritance response. For one, there is a sharp increase in ERGO-1 cluster mRNAs in the embryos of *wago-3* deletion mutants (**Figure 17E-G**). Even if we did not notice a decrease in ERGO-1 cluster 22G-RNAs, small effects falling below our detection might be enough to cause the effect, and an increase of mRNAs corresponding to ERGO-1 targets would coincide with a loss of the function of WAGO-3 to mediate the inheritance of ERGO-1 silencing. We also noticed that 22G-RNAs upregulated upon loss of WAGO-3 matched the binding signature of the two ERGO-1 cluster secondary Argonautes WAGO-8

and WAGO-12 as well as an increase in the mRNA levels of WAGO-6, which is a part of the same cluster, as defined by RIP-Seq experiments. We also note that nuclear Argonautes such as WAGO-12 might act downstream of other secondary Argonaute proteins (Ouyang et al., 2022), and our data supports a role of WAGO-3 acting in concert with either or all of the Argonautes WAGO-6, -8, and -12, to propagate the ERGO-1 response, and these being able to partially compensate loss of WAGO-3. While our data cannot explain the exact nature of the connection between these Argonautes, one possible explanation could be that WAGO-3, -6, -8, and -12 all are triggered, either directly or indirectly, by ERGO-1, and that loss of WAGO-3 causes upregulation of the other three Argonautes in an attempt to counter the loss.

WAGO-3 generally seems to be most important for mRNA regulation in the embryos, where several thousand genes were regulated in the deletion mutant (**Figure 12B**). Taken together with the observation that WAGO-3 is not necessary for exo-siRNA silencing in P0 (Schreier et al., 2024), this could support a model in which WAGO-3 acts primarily in the embryo, and otherwise only functions to mediate responses into sperm and oocytes in the L4 and adult worm.

In summary, we can conclude that while WAGO-3 is both paternally and maternally inherited, it associates only with the maternal, not the paternal, 26G-RNA pathway. It may do so either via direct interaction of the primary Argonaute, ERGO-1, or via interaction with one of the secondary Argonautes, WAGO-6, -8, or -12.

WAGO-1 may be responsible for paternal 26G-RNA inheritance

WAGO-1 has a negligible function in the embryo (**Figure 12**) and its main function is in the gravid adult, where its loss heavily upregulates over 1400 mRNAs (**Figure 9**), suggesting that it is responsible for silencing mRNAs in the gravid adult. This is consistent with the fact that loss of WAGO-1 has previously been shown to reduce germline 22G-RNA levels to near zero, with WAGO-3 and -4 rather being responsible for exo-siRNAs and having only minimal effects on the endogenous 22G-RNA pool (Gu et al., 2009). Our findings do somewhat contrast with this, however, as we did not see a general 22G-RNA depletion in our *wago-1* deletion mutants. While the effect is less prominent in L4 larvae, it appears that loss of WAGO-1 generally upregulates mRNA levels in worms but not in the embryo but that the regulation at 22G-RNA level is more modest (**Figure 6** and **Figure 12**). In L4 larvae, loss of WAGO-1 upregulates mRNAs and 22G-RNAs whose targets are expressed in L4 larvae but downregulate mRNA and 22G-RNAs whose targets are expressed later and in early embryos (**Figure 8E** and **H**), suggesting that WAGO-1 is responsible for controlling gene expression of expressed genes, possibly working as a limiting control. However, our data suggests that the mRNAs regulated may be somatic rather than germline expressed (**Figure 8M**), suggesting that these effects may be secondary and involve other pathways.

Loss of WAGO-1 causes a small increase in 22G-RNAs otherwise enriched in WAGO-1 RIP-Seq experiments (**Figure 7**, **Figure 10**, and **Figure 13**), although our own data has not been sufficient to explain where this increase comes from. One hypothesis would be that this comes from an increase in activity of the nuclear WAGO, WAGO-9, as 22G-RNAs upregulated in the *wago*-1 mutant tended toward mapping to the transcription end site of their corresponding genes (**Figure 18B**), which is the signature of WAGO-9 (Seroussi et al., 2023). WAGO-1 also shares 50% of its targets with WAGO-9, which makes it difficult to distinguish between the two Argonautes (Tyc et al., 2017). Furthermore, an intron-binding protein called EMB-4 has been shown to play a role in regulating the CSR-1 versus WAGO-9 response, although the study failed to distinguish between WAGO-9-bound 22G-RNAs and WAGO-1-bound 22G-RNAs due to their large overlap (Tyc et al., 2017). In our data, we saw that loss of WAGO-1 causes downregulation of CSR-1 target mRNAs (**Figure 19C-E**), which could be explained by either a direct or an indirect link to WAGO-9. While our data does not explore involvement of EMB-4, one hypothesis could be that loss of WAGO-1 causes increased activity of WAGO-9 to counter the loss, with WAGO-9 working downstream of WAGO-1 to silence the same targets, and that this is what disrupts the EMB-4-controlled balance between silencing WAGOs and the activating WAGO CSR-1.

Another line of evidence to support a possible direct link between WAGO-1 and WAGO-9, comes from the observation that WAGO-9 can function as a tertiary Argonaute, with its associated 22G-RNAs able to be generated from a primary response from other 22G-RNAs, otherwise only generated from siRNAs and 21U-RNAs (Ouyang et al., 2022). The study discovering this also showed that loss of WAGO-9 makes maternal transcripts subject to exo-RNAi accumulate at the pachytene region (Ouyang et al., 2022). Maternal transcripts usually start their expression in the pachytene region, and it was suggested that the effect of WAGO-9 lay in it working downstream of this with the transcripts accumulating at their starting point (Ouyang et al., 2022). Another recent study has shown that WAGO-1 is required neither for establishment nor for inheritance of the exo-siRNA response, but that loss of WAGO-1 did cause a small overactivity of exo-RNAi in the pachytene region of the germline; a phenotype which was maternally inherited (Schreier et al., 2024). This shared phenotype could suggest some functional linkage between WAGO-1 and WAGO-9. The same study also showed that WAGO-9 is needed for establishment and reestablishment of exo-siRNA mediated RNAi in the embryo (Schreier et al., 2024), where we have shown that WAGO-1 has less of an effect (**Figure 12**). All these observations together could all support a role of WAGO-9 working downstream of WAGO-1 to secure inheritance. Notably, however, the mRNA levels of WAGO-9 were not increased in any lifestage in the *wago-1* deletion mutant (**Figure 19F**), although activity levels could be increased regardless. More experiments will be needed in order to determine the nature of the relationship between WAGO-1 and

WAGO-9 and whether they do, indeed, act in the same pathway, or if they independently cause similar phenotypes.

While a previous study has suggested a link between WAGO-1 and inheritance of the ALG-3/4 response (Conine et al., 2010), the only basis for this hypothesis was the observation of a WAGO-1 in the sperm; a phenotype which we could not recapitulate in **Chapter I**, where we showed that at least N-terminally tagged GFP::3xFLAG::WAGO-1 fails to localise to mature sperm. Current studies have also failed to show any paternal RNAi inheritance phenomena in *wago-1* mutants (Schreier et al., 2024), further questioning a role of WAGO-1 in inheritance of the ALG-3/4 response. However, the finding that WAGO-1 does not contribute to inheritance from either maternal or paternal side relied on an exo-siRNA assay, and as it has already been established that WAGO-1 does not respond to exo-siRNA but rather contributes to endogenous silencing (Gu et al., 2009). Thus, it cannot be ruled out that WAGO-1 may have a role in endogenous inheritance of RNAi. Whether related to inheritance or not, our data does indeed support a role of WAGO-1 in interacting with the ALG-3/4 pathway. ALG-3/4 and WAGO-10, the only WAGO in the ALG-3/4 cluster as defined by RIP-Seq experiments, are all only expressed during spermatogenesis (Seroussi et al., 2023). Nonetheless, we found that loss of WAGO-1 causes a stark increase of ALG-3/4 cluster mRNAs in gravid adult worms, past the stage of spermatogenesis (**Figure 19C-E**). If WAGO-1 were to carry the ALG-3/4 signal into the sperm or into the adult worm, it would be expected that a loss of WAGO-1 would cause a loss of the ALG-3/4 cluster 22G-RNAs, which we did not see (**Figure 19A-B** and **E**). It is possible, however, that we did not see this because the changes were either too small for detection or because the expression of ALG-3/4 cluster 22G-RNAs is generally low (**Figure 17B**), and it possibly evaded detection for this reason. While ALG-3/4 cluster 22G-RNAs were not significantly decreased, ALG-3/4 cluster mRNA levels as well as mRNA levels of the *alg-3/4* genes themselves, along with *wago-10*, were increased in the gravid adult *wago-1* deletion mutants (**Figure 19F**). Assuming that these changes stem from minor, undetectable changes to 22G-RNA levels, this does not rule out a role of WAGO-1 in carrying the ALG-3/4 signal on to the next generation. ALG-3/4 have been shown to regulate their own expression (Almeida, de Jesus Domingues, et al., 2019), and their upregulation would be consistent with a malfunction of the ALG-3/4 pathway caused by failure of WAGO-1 to mediate the signal into the next generation.

Our data shows no indication that WAGO-1 may affect the ERGO-1 pathway or maternal inheritance of the 26G-RNA response. We did see a slight decrease of ERGO-1 cluster mRNAs in the embryos of *wago-1* mutant worms (**Figure 19C-E**), but it is unclear whether this has any significance. WAGO-1 has been shown to interact with the RNA-binding ATPase RDE-12, whose ATPase activity is necessary for the ERGO-1 pathway as well as for the RDE-1 (exo-siRNA) pathway (Shirayama et al., 2014). RDE-12 has been proposed to have a function in Argonaute loading (Shirayama et al., 2014), and with no function of WAGO-1 in the exo-siRNA

pathway, this could suggest that WAGO-1 is a player in the ERGO-1 pathway. ERGO-1, like WAGO-9, binds small RNAs accumulating at the transcription end site of their target genes (Seroussi et al., 2023), and it is not impossible that some of the increase in 22G-RNAs at transcription end sites we found in our *wago-1* deletion mutants stem from an upregulation of ERGO-1, although the exact nature of this upregulation is unclear. Since there is no regulation of ERGO-1 cluster 22G-RNAs (**Figure 19A-B** and **E**), or of ERGO-1 mRNA levels (**Figure 19F**), this might necessitate increased translation or activity of the ERGO-1 protein. Alternatively, since our data gives no reason to assume any connection between ERGO-1 and WAGO-1, it is more likely that the interaction of RDE-12 with ERGO-1 and the interaction of RDE-12 with WAGO-1 are not functionally related. Rather, RDE-12 may generally act in loading of Argonautes in different pathways.

In summary, our data shows that WAGO-1 is dispensable for endogenous RNAi in the embryo and that it has a function in paternal inheritance of the 26G-RNA response via the Argonaute proteins ALG-3/4. We fail to confirm a role in maternal inheritance via WAGO-1, although our data would not refute a role in inheritance, maternal or general, of WAGO-1 RNAi via the nuclear Argonaute protein WAGO-9.

WAGO-4 and WAGO-1 sequentially respond to ALG-3/4 26G-RNAs

Loss of WAGO-4 has strong effects on mRNA levels at all lifestages, but interestingly regulates mRNAs in different directions at different lifestages (**Figure 6B** and **Figure 12B** and **Figure 9B**), proving that the function of an Argonaute cannot be discussed without taking lifestage into account.

In L4 larvae, loss of WAGO-4 mimics loss of WAGO-1 in the gravid adult; mRNA levels of the *alg-3/4* genes as well as mRNA levels of ALG-3/4 cluster genes are both upregulated (**Figure 20**), which could suggest a functional link between WAGO-1 and WAGO-4 at different lifestages. While no link between these two Argonautes have previously been proven, current research does not refute such a claim. Here, we have shown that WAGO-1 and WAGO-4 both affect the ALG-3/4 pathway. The previous section detailed why our data suggests a role of WAGO-9 acting downstream of WAGO-1, and a possible link has been suggested between WAGO-4 and WAGO-9 in inheritance (Schreier et al., 2024) The proposed interaction between ALG-3/4 as a primary response, WAGO-1/4 as a secondary response, and WAGO-9 as a tertiary response is thus as follows:

ALG-3/4 is not expressed beyond the point of spermatogenesis (Conine et al., 2010), and must thusly trigger WAGO-1 and WAGO-4 both at the L4 lifestage. In L4 larvae, WAGO-1 does not elicit a response of this pathway, as evidenced by the unchanged levels of ALG-3/4 cluster 22G-RNAs and mRNAs upon loss of WAGO-1 (**Figure 9B**). Rather, WAGO-4 propagates the ALG-3/4 response in L4 larvae, possibly via interaction with WAGO-9. In adult worms, WAGO-4 no longer propagates this response, possibly because it is being triggered by other primary Argonautes taking up new 22G-RNAs. At this point, WAGO-1 will take over the function of silencing ALG-3/4 target mRNAs via interaction with WAGO-9.

In the absence of WAGO-4, we saw an increase both in 22G-RNAs and mRNAs of ALG-3/4 cluster genes (**Figure 20**). Since it has been shown that ALG-3/4 act both to silence and to activate gene expression, the latter via CSR-1 (Conine et al., 2013), and WAGO-4 share targets with CSR-1 (Xu et al., 2018), these increases could be explained by more WAGO-4/CSR-1 22G-RNAs being taken up by CSR-1 when WAGO-4 is lost. This is supported by two observations in our data. One, CSR-1/WAGO-4 22G-RNA and mRNA levels do not change drastically in *wago-4* mutant worms (**Figure 20**), suggesting that 22G-RNAs normally bound by WAGO-4 must be bound by another Argonaute. Two, we could not detect any discernible difference in 22G-RNA binding signature between the *wago-4* deletion mutant and the wildtype (**Figure 18**). With WAGO-4 and CSR-1 largely sharing binding signatures (Seroussi et al., 2023), CSR-1 taking up WAGO-4 targets in the absence of WAGO-4 could explain both these observations.

Overall, we conclude that WAGO-1 and WAGO-4 both interact with the ALG-3/4 pathway. We propose that they both may elicit their function via WAGO-9, and we suggest that propagation of the ALG-3/4 response is controlled on the one hand by silencing via WAGO-1 and WAGO-4 and other hand by activation via CSR-1. Loss of one of these Argonautes can change the balance between silencing and activating effects, thus having large effects on the mRNA levels of genes controlled by ALG-3/4.

WAGO-4 does not respond to ERGO-1

So far, we have demonstrated that WAGO-4 interacts with the ALG-3/4 pathway, either directly via WAGO-9, indirectly via overactivation of CSR-1, or through a combination of both of these mechanisms. We have not shown whether WAGO-4 interacts with the ERGO-1 pathway, but we did find that ERGO-1 cluster mRNAs were downregulated in the embryos of *wago-4* deletion mutant (**Figure 20C-E**). With CSR-1 cluster mRNAs simultaneously upregulated in the embryos of *wago-4* deletion mutants, it is not impossible that this could be caused by an effect of CSR-1 rather than directly by the loss of WAGO-4. Nonetheless, while CSR-1 has been shown to be directly involved in the ALG-3/4 response (Conine et al., 2013), no such link has been found between ERGO-1 and CSR-1. We offer a different explanation, prompted by the observation that *wago-9* mRNA is upregulated in *wago-4* deletion mutant embryos (**Figure 20F**).

As described in the previous section, there are several lines of evidence indicating a possible link between WAGO-4 and WAGO-9. Recently published data has shown that the inheritance of both WAGO-4 (via the oocyte) and WAGO-3 (via sperm and oocyte) necessitates reestablishment of the signal in the embryo via WAGO-9 (Schreier et al., 2024). My data implicates another nuclear Argonaute in the WAGO-3 pathway, namely WAGO-10, but does not stand in opposition to the findings by Schreier and coworkers. With WAGO-4 carrying ALG-3/4 supplied 22G-RNAs and WAGO-3 carrying ERGO-1 supplied 22G-RNAs, both Argonautes converging on the WAGO-9 pathway in the embryo, loss of WAGO-4 would cause WAGO-9 to predominantly

take up 22G-RNAs stemming from the maternal 26G-RNA pathway. This would both explain the downregulation of ERGO-1 cluster mRNAs in the embryos of *wago-4* deletion mutant and the slight upregulation of ALG-3/4 cluster mRNAs (**Figure 20C-E**). This lends further support to the notion that WAGO-4 would be directly involved in the ALG-3/4 pathway, and not only interfere with it indirectly via CSR-1.

While this hypothesis does not fully encapsulate the phenotypes of *wago-3* and *wago-4* deletions – with a remaining question of the possible involvement of CSR-1, whether direct or indirect – our data does cement that an interconnected pathway of several Argonautes, both primary and secondary/tertiary, exists in *C. elegans*. We conclude that WAGO-4 most likely interacts directly with ALG-3/4, taking up 22G-RNAs made from ALG-3/4 class 26G-RNAs, but does not interact with the ERGO-1 26G-RNA pathway.

Loss of WAGO-4 may cause early activation of an adult gene expression program

While loss of WAGO-1 or WAGO-3 affected certain lifestages more strongly than others, loss of WAGO-4 affected all lifestages. In L4 larvae, affected mRNAs were generally upregulated (**Figure 6B**), suggesting that WAGO-4 may act to suppress L4-type expression under normal conditions. Similarly, embryonic and adult mRNAs were downregulated in their corresponding lifestages (**Figure 9B** and **12B**). This could imply that WAGO-4 acts specifically to activate embryonic/adult expression and to supress L4 expression. However, this is contradicted by our observation that loss of WAGO-4 causes early activation of adult specific gene expression in L4 larvae (**Figure 8**). This suggests that WAGO-4 specifically inhibits adult-type expression in L4 larvae, possibly playing a direct role in when the switch from L4-type expression and spermatogenesis to adult-type expression and oogenesis occurs. As previously demonstrated, our data suggests that a shift may occur in the 22G-RNAs taken up by WAGO-4 between L4 and gravid adult lifestages, although RIP-Seq experiments in both lifestages would be needed to confirm this. The observation that mRNAs affected by loss of WAGO-4 are generally downregulated in adult worms may be indicative of such a switch. Alternatively, WAGO-4 may no longer be active in the adult worm, and the perceived activating effect may stem from a CSR-1 response.

While we did also see some contradictory results with spermatogenetic mRNAs also being upregulated in a *wago*-4 deletion mutant, our data suggested that WAGO-4 acts to supress L4-type expression and activate adult-specific gene expression, with its loss causing early activation of adult-type expression in L4 larvae. It would be interesting to study spermatogenesis deficiencies in *wago-4* deletion worms, for instance investigating whether the L4 larval stage lasts shorter or, and possibly as a result thereof, less sperm resides in the adult spermatheca of *wago-4* deletion mutant worms. Such observations would prove a role of WAGO-

4 in suppressing early activation of adult type gene expression; more specifically, it would prove a role in suppressing a premature switch from spermatogenesis to oogenesis.

Loss of WAGO-1, -3, and -4 have different, lifestage dependent outcomes

Something that the data presented here undisputably showed, was that lifestage was related to the effects of losing either of the three Argonautes studied, with WAGO-1 and WAGO-4 even showing opposing effects at different lifestages (**Figure 19** and **Figure 20**). While WAGO-1 and WAGO-3 may physically interact or are at least closely linked in physical proximity, their functions do not seem to be related; rather, they appear to be competing for the same set of shared cofactors (**Figure 15**). Similarly, WAGO-1 and CSR-1 have been shown to share targets but enact different functions (Charlesworth et al., 2021; Claycomb et al., 2009; Xu et al., 2018), underlining how neither physical location nor targets as defined by RIP-Seq alone can predict the function or redundancy of any particular Argonaute. At the same time, we saw that WAGO-1 and WAGO-4, two Argonautes that have not previously been shown or proposed to interact, neither based on mass spectrometry of RIP-Seq data, are connected, even if it is via primary triggers or via downstream interactors. Importantly, we only managed to capture this connection because we analysed several lifestages simultaneously, once again highlighting the importance of taking this into account.

It was curious that our data failed to capture any correlation between changes to 22G-RNAs and mRNA levels. This could suggest that 22G-RNAs do not actually regulate their corresponding mRNAs, however, current knowledge of 22G-RNAs as well as the fact that germline 22G-RNAs are generated via RdRPs from transcripts already trapped in the germ granules would dispute this. Our data also showed that 22G-RNAs more abundant in L4 larvae correspond to mRNAs also with high relative abundance in L4 larvae, just as 22G-RNAs more abundant in gravid adult worms corresponded to mRNAs with high relative abundance in adult worms. This would suggest that 22G-RNAs are produced against genes that are expressed and possibly necessary in a certain lifestage and that they might engage in maintaining a steady state of expression rather than fully silencing their targets. For genes that are important, other mechanisms may also exist to influence the steady state. For instance, if 22G-RNAs against a gene is lost, this may cause an immediate upregulation of that gene. The cell may then react to the increased concentration of the protein by turning down its transcription, thus causing the mRNA levels to be unaffected in the end. Similarly, the cell might respond to low concentration of an important gene caused by overexpression of 22G-RNAs by turning up transcription. We may fail to find correlations between 22G-RNA levels and mRNA levels because the majority of 22G-RNAs affected by loss of WAGO-1, -3, or -4 are not solely reliant on RNAi. Perhaps some correlations do exist for certain small subgroups of genes yet to be ascertained or perhaps genes whose expression levels rely on 22G-RNA silencing are more likely to be bound by several Argonautes.

We also saw indications that a larger than expected part of the mRNAs affected by loss of WAGO-1, -3, or -4 may be somatic. Since neither of these WAGOs are somatic this would dictate that this is an indirect effect, and it could involve engagement of one or several nuclear Argonautes such as WAGO-9.

Altogether, the data presented here underlines the importance of taking lifestage into account when discussion RNAi and sheds light on the degree of interconnectivity that different *C. elegans* Argonautes may engage in.

WAGO-4 inheritance in recovered WT

The final thing we showed, was that a wildtype that was five generations homozygous for the *wago-4* gene after having had a *wago-4* deletion mutation, still partially resembled the *wago-4* deletion mutant when it came to mRNA expression (**Figure 21**). It would be interesting to investigate this on a larger scale – how many generations does this persist for? A detailed study including wildtypes at several different generations downstream of a deletion would be interesting to answer such a question. It would both answer the question of how many generations changes persist for, but also of whether the reversion to wildtype will ever be complete. It is also possible, that some changes would never revert to the original state.

While the changes to mRNA levels may not be a direct effect of WAGO-4-mediated RNAi, some sort of inheritance signal must be involved for it to persist after reintroduction of WAGO-4. Since the changes may not be directly related to the Argonaute itself, this also poses the question of how widespread this is as a phenomenon. Is it possible, that edits of genes not involved in RNAi affects downstream generations in similar fashions? This might be relevant for studies indicating small changes to *C. elegans* strains having been generated via crosses or even to strains having been generated via gene editing using a co-inserted marker gene. The full extent to which epigenetic memory affects offspring of worms with gene edits would need to be elucidated for confident interpretation of data generated from such strains.

Materials and methods

C. elegans strains and harvest

All strains used in this chapter were obtained from the *Caenorhabditis* Genetics Center (CGC), via personal correspondence with Prof. Dr. Julie M. Claycomb, or were generated in-house. All strains have been previously described and are summarized in **Table 1**. All strains were cultured on nematode growth medium (NGM) plates (90 mm diameter) seeded with *Escherichia coli* OP50 in temperature-controlled incubators at 20°C.

Before harvest, deletion mutants were outcrossed to a wildtype N2 Bristol strain from both parental sides with the males having been procured via heat shock at 30°C. The offspring was confirmed as homozygous for the deletion allele via polymerase chain reaction (PCR) using the primers specified in **Table 2**. Worms for any experiment were synchronized for lifestage via bleaching and grown on egg plates, the preparation of which is described below. Worms were bleached a second time before being transferred to standard NGM plates, in the case of harvesting of L4 larvae or adult worms, or to egg plates when harvesting embryos. Harvesting was always carried out using M9 buffer and aliquots were fast-frozen on dry ice before storage at -80°C

Deletion mutants
wago-1(ok1074)
wago-3(pk1673)
wago-4(tm1019)

Endogenously tagged mutants
gfp::3xflag::wago-1(jmc219)
gfp::3xflag::wago-3(xf119)

Table 1: Strains used in this study.

Egg plates were made by mixing egg yolk with 50 ml LB medium per egg. After incubation at 65 °C for 2–3 hours, the mixture was allowed to cool to room temperature before adding 10 ml OP50 culture per egg. About 10 ml was put on top of standard NGM plates (90 mm diameter) and incubated at room temperature. The next day, excess liquid was removed, and egg plates were incubated at room temperature for another 2 days.

genotype	forward primer	reverse primer	length [bp]
wago-1(ok1074)	CTTCGTCGTTCAGGATCGTC	GCAGTCGTCAGCAAGAACAG	816
wago-1 (wildtype)	CTTCGTCGTTCAGGATCGTC	CTTGTCCCGTGATGTCGAAC	688
wago-1 (exon 3)	AACGCCACCGAAGAAGGATT	CATCAGTTCAGCCGGGAAGT	495
wago-3(pk1673)	ACCTCGGAACACAAAACTGC	AATTCGCTGCTGATGGTGAC	603
wago-3 (wildtype)	ACCTCGGAACACAAAACTGC	ACGACTGGGTTCTTGGTAGG	656
wago-4(tm1019)	GTCATCAAAGGGCTCGACTG	CACTCAACGCCTTGGATAGC	385
wago-4 (wildtype)	GTCATCAAAGGGCTCGACTG	CACTCAACGCCTTGGATAGC	953

Table 2: Genotyping primers used in this study.

Immunoprecipitation

Previously collected aliquots of ~300 µL L4 larvae were thawed on ice and mixed with the same volume of 2× lysis buffer (50 mM Tris HCl pH 7.5, 300 mM NaCl, 3 mM MgCl2, 2 mM dithiothreitol (DTT), 0.2% Triton X-100, cOmplete Mini EDTA-free Protease Inhibitor Cocktail (11836170001, Roche)). For mass spectrometry analysis, samples were sonicated using a Bioruptor Plus device (B01020001, Diagenode) (4°C, 10 cycles, 30 seconds on and 30 seconds off). For sequencing, samples were dropped into liquid nitrogen in a mortar, forming little balls that were subsequently pulverised using a pestle. The powder from this was transferred to a douncer, where it was dounced 40 times before turning liquid.

For both types of experiments, samples were next centrifugated for 10 min at 4°C and 21,000*g* and supernatants were transferred into a fresh tube. Total protein concentrations of soluble worm extracts were determined using the Pierce BCA Protein-Assay (23225, Thermo Fisher Scientific) and an Infinite M200 Pro plate reader (Tecan), and extracts were diluted with 1× lysis buffer to reach similar concentrations and a volume of 500 µL in the case of mass spectrometry samples and 550 µL for the preparation of sequencing samples. In the latter case, 50 µL were removed to serve as input control.

For each experiment, 30 µl Novex DYNAL Dynabeads Protein G (10004D, Invitrogen) were washed three times with 500 µl 1× wash buffer (25 mM Tris HCl pH 7.5, 150 mM NaCl, 1.5 mM MgCl2, 1 mM DTT, cOmplete Mini EDTA-free Protease Inhibitor Cocktail), combined with the remaining 500 µl extract and incubated with rotation for 1 hour at 4 °C. In the meantime, 8 µg antibody (monoclonal anti-Flag M2, F3165, Sigma-Aldrich) was conjugated to another 30 µl Novex DYNAL Dynabeads Protein G according to the manufacturer's instructions. Extracts were separated from non-conjugated Dynabeads, combined with antibody-conjugated Dynabeads and incubated with rotation for 2 hours at 4 °C. Following three washes with 500 µl 1× wash buffer, antibody-conjugated Dynabeads were resuspended in 50 µl nuclease-free water. For mass

spectrometry analyses, extracts were instead resuspended in 25 µl 1.2× Novex NuPAGE LDS sample buffer (NP0007, Invitrogen) supplemented with 120 mM DTT and incubated at 70 °C for 10 minutes before fast-freezing in dry ice and storage at -80°C,

Mass spectrometry

Samples were prepared in quadruplicates and thawed on ice before being separated on a Novex NuPAGE 4–12% Bis-Tris Mini Protein Gel (NP0321, Invitrogen) in 1× Novex NuPAGE MOPS SDS Running Buffer (NP0001, Invitrogen) at 180 V for 10 minutes. After separation, the samples were processed by in-gel digest as described in (Shevchenko et al., 2006). After protein digest, the peptides were desalted using a C18 StageTip78. For measurement, the digested peptides were separated on a 25 cm reverse-phase capillary (inner diameter, 75 µm) packed with Reprosil C18 material (Dr. Maisch). Elution was carried out along a 2 hour gradient of 2–40% of a mixture of 80% acetonitrile/0.5% formic acid using the EASY-nLC 1000 system (LC120, Thermo Fisher Scientific). A Q Exactive Plus mass spectrometer (Thermo Fisher Scientific) operated with a Top10 data-dependent MS/MS acquisition method per full scan was used for measurement. Processing of the obtained results was performed with the MaxQuant software (v.1.5.2.8) against the Wormbase protein database (version WS270) for quantification.

RNA extraction

For samples not designated for immunoprecipitation experiments, harvested frozen worms were thawed on ice before being dropped into liquid nitrogen in a mortar, forming little balls that were subsequently pulverised using a pestle. The powder from this was transferred to a douncer, where it was dounced 40 times before turning liquid. This liquid was transferred to a new tube and was centrifugated for 10 min at 4°C and 21,000*g*. Supernatants were then transferred into a fresh tube.

RNA was extracted from this or from the extract from immunoprecipitation using TRIzol LS Reagenz (10296010, Invitrogen) according to the manufacturer's instructions, and resuspended in nuclease-free water. RNA quality and quantity was assessed initially using the Bioanalyzer RNA 6000 Nano Kit (5067-1511, Agilent Technologies) and subsequently using the Qubit RNA BR Assay Kit (Q10210, Invitrogen).

Small RNA library preparation and sequencing

RNA 5′ pyrophosphohydrolase (RppH) (M0356S, New England BioLabs) treatment was performed with a starting amount of 76-126 ng. NGS library prep was performed with NEXTflex Small RNA-Seq Kit V3 following Step A to Step G of Bioo Scientific`s standard protocol (V19.01) with a ligation of the 3' 4N Adenylated Adapter over night at 20°C. Libraries were prepared with a starting amount of 48 ng amplified

in 20 PCR cycles for immunoprecipitation experiments. For other small RNA sequencing experiments, libraries were prepared with a starting amount of 360 ng amplified in 15 PCR cycles

Amplified libraries were purified by running an 8% TBE gel and size-selected for 144 – 163 nt.

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 samples were pooled in equimolar ratio and sequenced on a NextSeq 500/550 Flowcell, SR for 1x 75 cycles plus 6-7 cycles for the index read.

mRNA library preparation and sequencing

NGS library prep was performed with Illumina's Stranded mRNA Prep Ligation Kit following Stranded mRNA Prep Ligation ReferenceGuide (June 2020) (Document # 1000000124518 v00). Libraries were prepared with a starting amount of 500 ng and amplified in 10 PCR cycles. For normalization, 1 µl of a 1:100 dilution of ERCC spike-ins (Ambion) was added. 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 or 4.0 Fluorometer (Life technologies). All samples were pooled in equimolar ratio and sequenced on a NextSeq500 Highoutput FC, SR for 1x 80 or 1x 79 cycles plus 10 cycles for the index read and 1 or 2 dark cycles upfront.

Read processing and mapping

Raw sequenced reads from high-quality libraries, as assessed by FastQC, were processed using Cutadapt (Martin, 2011) for adapter removal (-a TGGAATTCTCGGGTGCCAAGG -O 5 -m 26 -M 48) and low-quality reads were filtered out using the FASTX-Toolkit (fastq_quality_filter, -q 20 -p 100 -Q 33). Unique molecule identifiers were used to remove PCR duplicates using a custom script and were subsequently removed using seqtk (trimfq-l 4 – b 4). Finally, in the case of sRNA-Seq, reads shorter than 15 nucleotides were removed using seqtk (seq -L 15).

A custom GTF file was created by adding transposons retrieved from Wormbase (PRJNA13758.WS264) to the Ensembl reference WBcel235.84 and reads were aligned using bowtie (v.1.3.1) (Langmead et al., 2009) (--phred33-quals --tryhard --best --strata --chunkmbs 256 -v 2 -M 1) for sRNA-Seq or subread v.2.0.0 (Liao et al., 2013) featureCounts (--donotsort -t exon) for mRNA-Seq.

Coverage tracks

BigWig files were generated from mapped mRNA reads using deepTools v.3.5.5 (Ramírez et al., 2016) bamCoverage using default settings. These were visualised in the online tool Integrative Genomics Viewer

(IGV) (Robinson et al., 2011). Tracks were downloaded as single vector graphics (SVG) files and recoloured using Adobe Illustrator.

Small RNA classification

All mapped reads were divided into sense and antisense reads using BEDTools v.2.29.2 intersect (Quinlan & Hall, 2010), and reads of differing lengths and 5′ nucleotides were counted using a custom Python script which can be found at https://github.com/Tunphie/SequencingTools/blob/main/summarizeNucleotideByReadLength.py. Different classes of small RNAs were counted using a custom Python script available at https://github.com/Tunphie/SequencingTools/blob/main/smRNA_TypeCounter.py, which counts 22G-RNAs as any RNA of length 19-23 nt with a G at the 5'-position, 21U-RNAs as any RNA of length 19-22 with no requirement of 5'-nucleotde having not already been defined as a 22G-RNA, 26G-RNAs as any RNA of exact length 26 with a G at the 5'-position, and miRNAs as any RNA shorter than 27 nt which maps to a known miRNA gene. All reads longer than 27 nt were also counted.

22G-RNAs were extracted using a python script available at https://github.com/adomingues/filterReads and were defined as any read of length 20-23 nt with no bias at the 5'-position.

Differential analysis

Reads were counted using htseq-count (v.2.0.2) (Anders et al., 2015) (-s no -m intersection-nonempty). Reads per kb of transcript per million mapped reads (RPKM) values were calculated relative to all mapped reads.

For RIP-Seq experiments, targets were defined as genes that were positive in at least two out of three replicates, with positive meaning that the 22G-RNA RPKM was (1) above 4 in the IP; (2) at least twice as high in the IP relative to the input; and (3) non-zero in the input. For 22G-RNAs mapping to transposons, RPKM values were calculated relative to the predicted number and length of insertions in the custom annotation file and positives were defined using only requirements 2 and 3 above with no minimal RPKM requirement in the IP.

For all other experiments, targets were determined using DeSeq2 (Love et al., 2014) in R, with a p-value cutoff of 0.01 and a cutoff for log2(fold change) of 1. For mRNA-Seq, reads mapping to ERCCs were used for normalization by resetting the size factors using estimateSizeFactors() on the data with only ERCC reads. All data was shrunk using lfcShrink() with the generalized linear model from apeglm (Zhu et al., 2019).

Data comparisons and visualizations

Subsets of genes found in our analyses were compared to published data of RIP-Seq targets from (Gu et al., 2009; Schreier et al., 2022; Seroussi et al., 2023; Xu et al., 2018) and to a lists of genes expressed in the *C. elegans* germline of either *fem-3*- or *fog-2*-mutant animals (Ortiz et al., 2014). The data was also compared to RNA-Seq data from GExplore (Boeck et al., 2016; Hutter & Suh, 2016). Expression of each gene found in GEXplore was normalised to its own expression using R.

All comparions were carried out in R using the packages ggplot2, UpsetR (Conway et al., 2017), or VennDiagram (Chen & Boutros, 2011) for plotting. GO-term analyses were carried out using clusterProfiler (Yu et al., 2012). Where cutoffs were utilised during comparisons, these were determined as the lowest baseMean at which log2(fold change) reached a value of 1. Datapoint directly on the line were included.

Coverage of 22G-RNAs along targeted protein-coding genes was visualized by creating BigWig tracks using BEDTools (genomeCoverageBed -bg -scale -split) followed by bedGraphToBigWig from the UCSC Genome Browser's utilities. Coverage for each gene was determine with deepTools (computeMatrix scale-regions --metagene --missingDataAsZero -b 250 -a 250 --regionBodyLength 2000 --binSize 50 --averageTypeBins median) and plots were subsequently generated with plotProfile (--plotType se --averageType mean --perGroup).

General discussion

T. rugatulus does not fully turn off the piRNA pathway in inactive ovaries

Since RNAi has been found in all domains of life, our discovery of both piRNAs and miRNAs in *Temnothorax rugatulus* was unsurprising (**Chapter II**). Since piRNAs play a crucial role in the germ cells, we had also expected the relative increase in piRNAs over miRNAs in ovaries, with more relative expression of miRNAs in somatic tissues (**Chapter II**, **Figure 4a** and **Figure S1**). Our largest discovery came in the observation that rudimentary ovaries of workers without any capacity for fertility still produced piRNAs (**Chapter II**, **Figure 3**), suggesting that the importance of transposon silencing in the germline outweighs the energy disadvantages of keeping the mechanisms intact in organisms where it is no longer needed. Alternatively, it may also be that there has been no evolutionary drive to develop mechanisms for fully downregulating or turning off the piRNA pathway in ageing *T. rugatulus* workers.

In *D. melanogaster*, loss of PIWI causes the ovaries to become rudimentary (Jin et al., 2013; Lin & Spradling, 1997), indicating a link between PIWI and ovary physiology. piRNAs have also been proposed to play a role in fish, whose ovarian activation and general fertility is impacted by them being kept in captivity, although miRNAs are the more studied of the small RNAs in this role (Papadaki et al., 2020; Papadaki et al., 2024). Nonetheless, PIWI Argonautes and piRNAs affecting ovary maturation in other organisms is not necessarily indicative of ovary maturation affecting PIWI Argonautes and piRNAs. Since social insects are unique in their need for inactivating ovaries, it is not possible to say whether the piRNAs in inactive ovaries of *T. rugatulus* is specific for this ant, all ants, all social insects or if piRNA activity would remain in the ovaries of even more organisms, if they, too, developed mechanisms for ovary inactivation matching that of *T. rugatulus*. While it would be hard to address the cost-benefit of keeping an active piRNA pathway in rudimentary ovaries of animals that do not rely on ovary inactivation as a part of their social organisation, it would be interesting to investigate the undeveloped ovaries of workers of other social insects, to assess whether the incomplete loss of piRNAs that we see in *T. rugatulus* is common to all social insects.

Caste and behaviour are partially determined by miRNAs

Also interesting was our finding that several miRNAs were differentially expressed between female ants of different castes and that some of the specific miRNAs found in our analyses coincided with miRNAs found to differ between castes of other species (**Chapter II**). While the miRNA miR-133 showed the same tendency between worker and queen castes in our analyses of *T. rugatulus* and in a previous study on the termite *Reticulitermes speratus* (Matsunami et al., 2019), miR-12 showed opposing expression in workers versus queens in *T. rugatulus* and *Apis mellifera* (Macedo et al., 2016). In an evolutionary context, it is possible that

some miRNAs have preserved both their function in regulating caste and the gene regulatory pathways in which they partake, whereas others may have preserved a function in regulating caste but have adapted to target new genes or impart different functions. Notably, we have not scoured the genome for possible targets of any of the miRNAs found in our study, nor have we investigated the Argonaute protein – or, indeed, Argonaute proteins – with which miRNAs may associate in *T. rugatulus*. Many differentiation steps can have occurred over the time of evolution to both miRNA sequences, miRNA target sequences, and functions of any protein in the RISC or otherwise involved in RNAi. However, although we cannot say for certain what the function of miRNAs is in caste-determination, it is clear that such a function must exist. This function is also not limited to *T. rugatulus*; rather caste-specific miRNA expression has been found in a variety of other social insects including two other species of ants (Roberto Bonasio et al., 2010), in two species of bumblebees (Collins et al., 2017; Liu et al., 2019), in the honeybee *A. mellifera* (Ashby et al., 2016; Chen et al., 2010), and in the termite *R. speratus* (Matsunami et al., 2019). This all indicates that miRNAs play a crucial role in the organization of social insects.

In *A. mellifera*, miRNAs have further been shown to correlate to caste- and sex-specific differentiation in several tissues including brain (Vieira et al., 2021) and wing discs (Soares et al., 2021). As touched upon in the introduction, small RNAs have been found to play a role in sex determination in both *A. mellifera* and the jewel wasp *Nasonia vitripennis* (Verhulst et al., 2010; Watson et al., 2022). Males are often overlooked in the studies of ant or other social insects due to their short lifespans and low contribution to the colony, but an interesting aspect that our study failed to shed light on is the presence of small RNAs in male ants and in *T. rguatulus* sperm. It would be interesting to know whether the male miRNA population differs vastly from the female miRNA population or whether male miRNAs might more closely resemble either of the female castes. The degree to which piRNAs are active in the male reproductive organs might also be of interest, although male ants do not have caste systems and it would therefore not be possible to investigate ants with rudimentary sex organs, like we did for the females in **Chapter II**.

A further area of interest might be the difference between colonies from different areas. A previous study has shown that miRNAs in the larval food of developing *A. mellifera* larvae was related to caste differentiation (Guo et al., 2013), and it might be curious to test whether *T. rugatulus* ants having lived in different environments or feeding on different food supplies show different characteristics regarding caste differentiation or if they even start expressing different subsets of small RNAs themselves.

In general, epigenetics including DNA methylations and histone modifications but also small RNAs has been shown to control insect behaviour, development, adaptability, and stress response (Gupta & Nair, 2025). Female ants of different castes have distinct phenotypical traits, but they share the same DNA, so it follows that epigenetics must be the cause of the phenotypical differences, but it is possible that phenotype

and behaviour is linked via small RNAs and/or epigenetics as a whole. Even in humans, who are evolutionarily very distant from social insects, epigenetics have been shown to play a role in behaviour and personality (Yan, 2025), with miRNAs for instance playing a role in attention deficit hyperactivity disorder (Martinez & Peplow, 2024) and depression (Zheng & Jin, 2024). The microbiome has also been shown to play a role in both human behaviour (Yan, 2025) and caste differences of ants (Favreau et al., 2018). Perhaps, the mechanisms controlling the behaviours of complex animals and of simpler species such as social insects are far more conserved than the level of behavioural complexity might suggest.

PEI-granules infer paternal inheritance of RNAi in *C. elegans*

In **Chapter I**, we showed that the secondary *C. elegans* Argonaute WAGO-3 is required for paternal inheritance of the PRG-1 (21U-RNA) pathway of RNAi, and that WAGO-3 segregates properly into sperm via the PEI-granule, containing the proteins PEI-1 and PEI-2.

We showed that the PEI-granules segregate to sperm via anchoring to the fibrous body-membranous organelle (FB-MO), which is a sperm-specific organelle (**Chapter I**, **Figure 6**). All germ granules exist perinuclearly past the 4-cell stage, and they form there via their association with the NPCs, which are transmembranal (Phillips & Updike, 2022). NPCs are not present in other membranes than the nucleus, however, which would necessitate either inclusion of another membranal protein in the PEI-granule or PTMs in the form of fatty acid anchors to one of the proteins in the granule. The addition of the C16:0 fatty acid palmitate to cysteine is one of the most common PTMs involved in protein trafficking (Tabaczar et al., 2017), and we did indeed show that *S*-palmitoylation via the palmitoyltransferase SPE-10, most likely of PEI-2, is involved in PEI-granule association with FB-MOs (**Chapter I**, **Figure 8A-B**).

Disruption of the PEI-granules interferes with the paternal inheritance of RNAi, because WAGO-3 fails to segregate into the budding spermatids (**Chapter I**, **Figure 1**). Generally, the assembly and structure of nuage such as PEI-granules, P-granules, Z-granules, and other germ granules are necessary for the function and inheritance of RNAi (Phillips & Updike, 2022). Dispersion or enlargement of P-granules disrupt RNAi (Phillips & Updike, 2022), but it was also shown that the RNA-binding protein MEG-3 recruit mRNAs to P-granules not for silencing, but for its own stability and propensity for stabilizing the P-granules (Scholl et al., 2024). While intrinsically disordered protein regions are often thought to be the main drivers of phase separation in nuage, factors such as RNA concentration also play a role in nuage formation and stability. PGL-1 and PGL-3 are required for P-granule assembly but not for the stability thereof (Phillips & Updike, 2022), so it is not their disordered nature alone that cause phase separation. Our results indicate that the structured regions, the BTB and BACK domains of PEI-1 and PEI-2, are the drivers of phase separation in the PEI-

granules (**Chapter I**, **Figure 3**). Our data also suggests that PEI-granules are more static and less liquid-like than P granules (**Chapter I**, **Figure 6**).

What drives WAGO-3 into P-granules and what drives WAGO-3 into PEI-granules? It could be direct interactions with proteins in either granule and it most likely depends on a tight regulation of other pathways. Loss of the intrinsically disordered region of PEI-1 disrupted the recruitment of WAGO-3 to PEI-granules (**Chapter I**, **Figure 4**), which could suggest a direct interaction between the two proteins via this region of PEI-1. Alternatively, WAGO-3 could fail to locate to the PEI-granules lacking the intrinsically disordered region of PEI-1 due to a change in fluidity of the granule. This further begs the question of how WAGO-3 manages to locate to two different granules with different physiochemical properties.

In general, the information of how different proteins are directed into different nuage is lacking, as is the information of how the tightly associated germ granules manage to move proteins and RNA between them without mixing. Much more research is needed in order to fully understand the assembly and recruitment of phase-separated structures, although it is undisputed that they play an important role in RNAi and inheritance hereof.

WAGO-3 and WAGO-4 are responsible for inheritance via nuclear WAGOs

Something that our data in **Chapter I** and **Chapter III**, along with published data, more specifically the research by Schreier et al. (2024), has shown, is that WAGO-3 and WAGO-4 are both responsible for inheritance and that this requires interaction of at least one nuclear WAGO, WAGO-9. In **Chapter III**, we also demonstrated how WAGO-3 and WAGO-4 are most likely triggered by two different 26G-RNA pathways. We offered evidence to demonstrate how WAGO-3 (inherited both via sperm and oocyte and triggered by ERGO-1) and WAGO-4 (inherited only via the oocyte and triggered by ALG-3/4) both converge on the same pathway, utilising WAGO-9 to establish inheritance in the embryo. Interestingly, we could also reveal that WAGO-1 interacts with both the ALG-3/4 pathway and WAGO-9, but no inheritance phenotype has yet been described in *wago-1* mutant animals. It is also interesting, that deletion of *wago-4* only appears to affect the ALG-3/4 pathway in L4 larvae, with WAGO-1 taking over the function in gravid adult worms. WAGO-1 and WAGO-4 are not mutually exclusive, as the distinction of lifestage-specificity seems to play a role, and yet they seem to interact. More research would need to be conducted in order to determine the exact nature of this interaction. In **Chapter III**, we proposed that WAGO-4 may take up new targets past L4 lifestage, and that this is why we no longer see an effect on the ALG-3/4 pathway in adult worms, but we also cannot rule out that WAGO-1 and WAGO-4 interact directly, with the latter triggering production of 22G-RNAs in the former. Double mutants of both WAGOs, RIP-Seq of one WAGO in a deletion mutant of the other WAGO, and RIP-Seq of both WAGOs at different lifestages may help to better understand interaction.

WAGO-4 is not paternally inherited (Schreier et al., 2024), and yet it interacts with the ALG-3/4 pathway, which is active during spermatogenesis and can thus be characterised as the paternal 26G-RNA pathway. Simultaneously, the only WAGO confirmed to act in paternal inheritance, WAGO-3 (Schreier et al., 2022; Schreier et al., 2024), interacts with the maternal 26G-RNA pathway (governed by ERGO-1 which is present during oogenesis and in oocytes). So why does a maternally inherited WAGO interact with a paternal primary Argonaute, and a paternally inherited WAGO interact with a maternal primary Argonaute? It is possible, that interaction between maternal and paternal responses is needed in order to ensure proper inheritance of both 26G-RNA responses from both sides. Furthermore, we have demonstrated that WAGO-3 does not only interact with WAGO-9, but its loss generally influences WAGOs which share targets with ERGO-1 (**Chapter III**). It would be interesting to further study the interaction pathway of WAGO-3 in order to determine whether it interacts directly with multiple secondary/tertiary Argonautes, or whether our observations are secondary effects, stemming from disruptions to the ERGO-1 pathway.

Withstanding questions

What are the precise functions of the somatic germline-expressed WAGOs? How is their uptake of 22G-RNAs determined and what do their silencing pathways look like? While the data we have provided in **Chapter I** and **Chapter III** certainly help shed some light on this, leading to the conclusions expressed above and in the '**Discussion and conclusions'** in **Chapter III**, current research still fails to fully answer these questions.

Many of the conclusions we have drawn here are based on observations in mutants lacking a single *wago* gene. While this can explain general patterns and have enabled us to reveal interaction of various WAGO pathways, the interconnectivity of these pathways also makes it impossible to draw conclusions for the exact modes of action. When loss of one WAGO causes disturbances to 22G-RNA and/or mRNA levels of another WAGO pathway, this can be due to direct interaction of these WAGOs, it can be caused by a general interaction of their pathways, or it can even be caused by whole other WAGOs changing which 22G-RNAs they associate with. In **Chapter III**, we have used the characteristics of the 22G-RNA binding signatures to infer which Argonautes compensate loss of another, but with many Argonautes largely sharing signatures (Seroussi et al., 2023) and the signatures of subgroups of 22G-RNAs from sRNA-Seq not following these signatures perfectly (**Chapter III**, **Figure 16** and **Figure 18**), we still cannot fully prove the compensatory effects of certain Argonautes. In order to determine whether compensation for loss of one Argonaute is carried out by one specific Argonaute, RIP-Seq experiments in the deletion mutants would be necessary. With several Argonautes presumed to act in the same pathway(s), mutants with multiple *wago* deletion might also warrant such studies. However, if one wishes to examine the effects of all Argonautes, this does pose a logistic issue. It would not currently be feasible to perform RIP-Seq experiments on all Argonautes in all single

mutants, and much less so if all combinations of double or even triple mutants were to be included as well. Even with only using a select few double or triple mutants or none at all, this would be a cumbersome experiment, that might still prove futile as we do not know how the Argonaute pathways interact. Further complicating large-scale experiments is our observation that lifestage undisputably affects function (**Chapter III**), and performing large-scale experiments of all Argonautes simultaneously and at all lifestages may not be the best approach. Rather, with the current knowledge of *C. elegans* Argonaute pathways, focusing on a single WAGO might prove more helpful in determining overall pathways than trying to tackle all at once. Perhaps looking further upstream and determining how a WAGO is loaded would also be more beneficial than investigating the downstream effects. Knowing how certain 22G-RNAs get loaded into a certain WAGO and which cofactors are necessary for this may provide better understanding of how an Argonaute can be misloaded and how it can interact with some primary or tertiary Argonautes but not with others.

Also yet unknown is the role of CSR-1. While the research presented in this thesis has not focused on this rather controversial Argonaute, changes to CSR-1 cluster 22G-RNA and/or mRNA levels could be observed in all three deletion mutants studied in **Chapter III** (**Figure 17**, **Figure 19**, and **Figure 20**). Since CSR-1 is peculiar in having both silencing and activating activities, along with other proposed functions, it is difficult to fully ascertain how much of the mis-regulation caused by *wago* deletions has to do with CSR-1 and the CSR-1 versus WAGO homeostasis.

RIP-Seq experiments may pose another yet undiscussed problem. Current research has shown that findings in RIP-Seq experiments are reproducible, which we demonstrate on small scale in **Chapter III** (**Figure 2F** and **Figure 3B** and **C**). Nonetheless, we saw some differences in 22G-RNAs taken up by WAGO-3 in L4 larvae and gravid adult worms (**Chapter III**, **Figure 2**), but as we have not carried the RIP-Seq experiments out in tandem, we can only hint at the extent of the lifestage-specific uptake of 22G-RNAs by WAGO-3. More pressing as a concern, however, is the fact that we do not currently have antibodies against the endogenous WAGOs and that RIP-Seq experiments are generally carried out using a tagged version of the Argonaute of interest. While the worms with endogenous tags show no phenotype, these tags are always placed at the N-terminus; a region which is known to be processed by DPF-3 in WAGO-1 and WAGO-3 (Gudipati et al., 2021). In order to confirm several of our observations in **Chapter III**, it would be necessary to confirm whether WAGO-1 and WAGO-3 change the 22G-RNAs which they bind upon modification of their N-termini.

The deletion mutants used in **Chapter III** were the fifth generation homozygous. This generation was chosen to accommodate for genotyping and growth while doing it early enough to minimize secondary effects. As demonstrated, however, small RNA misregulation caused by changes to the genetic landscape can persist after five generations (**Chapter III**, **Figure 21**). With changes to mRNA levels and changes to 22G-

RNA levels not coinciding, it is clear that secondary effects must also play a role in the deletion mutants. Rather than doing large-scale transcriptomics, it might be more interesting to perform single-worm sequencing of an earlier generation. It might even be interesting to see if any detectable changes exists between newly outcrossed descendants of a worm having priorly been kept homozygous for a *wago* deletion and newly outcrossed descendant of a worm recently having obtained the same deletion.

Due to interconnectivity and possible partial redundancy of Argonaute pathways, studying deletion mutants without accompanying RIP-Seq experiments can hint at functions, but do not provide the full image. A possible other way to investigate Argonaute specificity would be to alter Argonautes rather than to remove them from the system altogether. As briefly touched upon in the introduction, an Argonaute can be forced to change which 22G-RNAs it associates with. Gudipati and coworkers (2021) found that loss of DPF-3 causes WAGO-1 to take up WAGO-4 targets. They also found that 22G-RNAs normally associated with WAGO-1 were generally lost in these mutants, suggesting that no other Argonaute compensated the loss of binding of WAGO-1. This is in opposition to our findings in **Chapter III**, where there was no general loss of WAGO-1-associated 22G-RNAs in *wago-1* deletion mutants. Nonetheless, Gudipati et al. did not further analyse changes to 22G-RNA landscape, and their results were also affected by changes to WAGO-3. It would be interesting to see if chemical modifications of WAGO-1, already shown by Gudipati et al. to be capable of recapitulating the *dpf-3* deletion phenotype when WAGO-1 can no longer works as a substrate for the protease, would affect the WAGO-9 and ALG-3/4 pathways in similar fashion to the *wago-1* deletion mutant. Similarly, would it be possible to change the 22G-RNAs taken up by WAGO-4? WAGO-4 is not a substrate of DPF-3 (Gudipati et al., 2021), so a method for forcing such a change is currently not evident. Nonetheless, it would be interesting to study different mutations of WAGO-1 and WAGO-4 to find out why they can act similarly but at different lifestages. If their physical location is related to this difference, would it be possible to force WAGO-1 into the Z-granule and WAGO-4 into the P-granule, for instance via switching their intrinsically disordered regions? It is still unclear exactly what makes a protein favour one granule over another, but more study into this may provide insights that would allow generation of new tools to force granular re-localisation, which might help shed more light on the Argonaute pathways.

Much is still unknown about the specifics of the worm-specific *C. elegans* RNAi pathway(s), but the data presented in **Chapter I** and **Chapter III** here has provided several hints at the functions of WAGO-1, -3, and -4 as well as their interconnectivity with other Argonaute pathways. These findings can aid in designing new experiments, as we now better understand the right questions to ask.

References

Almeida, M. V., Andrade-Navarro, M. A., & Ketting, R. F. (2019). Function and Evolution of Nematode RNAi Pathways. *Noncoding RNA*, *5*(1). 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 Genet*, *15*(2), e1007784. https://doi.org/10.1371/journal.pgen.1007784

Ambros, V., & Ruvkun, G. (2018). Recent Molecular Genetic Explorations of Caenorhabditis elegans MicroRNAs. *Genetics*, *209*(3), 651-673. https://doi.org/10.1534/genetics.118.300291

Anders, S., Pyl, P. T., & Huber, W. (2015). HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics*, *31*(2), 166-169. https://doi.org/10.1093/bioinformatics/btu638

Aoki, S. T., Lynch, T. R., Crittenden, S. L., Bingman, C. A., Wickens, M., & Kimble, J. (2021). C. elegans germ granules require both assembly and localized regulators for mRNA repression. *Nat Commun*, *12*(1), 996. https://doi.org/10.1038/s41467-021-21278-1

Ashby, R., Forêt, S., Searle, I., & Maleszka, R. (2016). MicroRNAs in Honey Bee Caste Determination. *Scientific Reports*, *6*(1), 18794. https://doi.org/10.1038/srep18794

Bartel, D. P. (2009). MicroRNAs: target recognition and regulatory functions. *Cell*, *136*(2), 215-233. https://doi.org/10.1016/j.cell.2009.01.002

Bessereau, J. L. (2006). Transposons in C. elegans. *WormBook*, 1-13. https://doi.org/10.1895/wormbook.1.70.1

Blumenthal, T. (2012). Trans-splicing and operons in C. elegans. *WormBook*, 1-11. https://doi.org/10.1895/wormbook.1.5.2

Boeck, M. E., Huynh, C., Gevirtzman, L., Thompson, O. A., Wang, G., Kasper, D. M., Reinke, V., Hillier, L. W., & Waterston, R. H. (2016). The time-resolved transcriptome of C. elegans. *Genome Res*, *26*(10), 1441-1450. https://doi.org/10.1101/gr.202663.115

Bonasio, R., Zhang, G., Ye, C., Mutti, N. S., Fang, X., Qin, N., Donahue, G., Yang, P., Li, Q., Li, C., Zhang, P., Huang, Z., Berger, S. L., Reinberg, D., Wang, J., & Liebig, J. (2010). Genomic comparison of the ants Camponotus floridanus and Harpegnathos saltator. *Science*, *329*(5995), 1068-1071. https://doi.org/10.1126/science.1192428

Bonasio, R., Zhang, G., Ye, C., Mutti, N. S., Fang, X., Qin, N., Donahue, G., Yang, P., Li, Q., Li, C., Zhang, P., Huang, Z., Berger, S. L., Reinberg, D., Wang, J., & Liebig, J. (2010). Genomic Comparison of the Ants Camponotus floridanus and Harpegnathos saltator. *Science*, *329*(5995), 1068-1071. https://doi.org/10.1126/science.1192428

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

Bressendorff, S., Kausika, S., Sjøgaard, I. M. Z., Oksbjerg, E. D., Michels, A., Poulsen, C., & Brodersen, P. (2023). The N-coil and the globular N-terminal domain of plant ARGONAUTE1 are interaction hubs for regulatory factors. *Biochem J*, *480*(13), 957-974. https://doi.org/10.1042/bcj20230025

Brutscher, L. M., & Flenniken, M. L. (2015). RNAi and Antiviral Defense in the Honey Bee. *J Immunol Res*, *2015*, 941897. https://doi.org/10.1155/2015/941897

Cecchetelli, A. D., & Cram, E. J. (2017). Regulating distal tip cell migration in space and time. *Mech Dev*, *148*, 11-17. https://doi.org/10.1016/j.mod.2017.04.003

Charbonneau, D., Sasaki, T., & Dornhaus, A. (2017). Who needs 'lazy' workers? Inactive workers act as a 'reserve' labor force replacing active workers, but inactive workers are not replaced when they are removed. *PLoS One*, *12*(9), e0184074. https://doi.org/10.1371/journal.pone.0184074

Charlesworth, A. G., Seroussi, U., Lehrbach, N. J., Renaud, M. S., Sundby, A. E., Molnar, R. I., Lao, R. X., Willis, A. R., Woock, J. R., Aber, M. J., Diao, A. J., Reinke, A. W., Ruvkun, G., & Claycomb, J. M. (2021). Two isoforms of the essential C. elegans Argonaute CSR-1 differentially regulate sperm and oocyte fertility. *Nucleic Acids Res*, *49*(15), 8836-8865. https://doi.org/10.1093/nar/gkab619

Chen, H., & Boutros, P. C. (2011). VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. *BMC Bioinformatics*, *12*(1), 35. https://doi.org/10.1186/1471-2105-12-35

Chen, X., Ma, C., Chen, C., Lu, Q., Shi, W., Liu, Z., Wang, H., & Guo, H. (2017). Integration of lncRNA-miRNA-mRNA reveals novel insights into oviposition regulation in honey bees. *PeerJ*, *5*, e3881. https://doi.org/10.7717/peerj.3881

Chen, X., Yu, X., Cai, Y., Zheng, H., Yu, D., Liu, G., Zhou, Q., Hu, S., & Hu, F. (2010). Next-generation small RNA sequencing for microRNAs profiling in the honey bee Apis mellifera. *Insect Mol Biol*, *19*(6), 799-805. https://doi.org/10.1111/j.1365-2583.2010.01039.x

Choppin, M., Feldmeyer, B., & Foitzik, S. (2021). Histone acetylation regulates the expression of genes involved in worker reproduction in the ant Temnothorax rugatulus. *BMC Genomics*, *22*(1), 871. https://doi.org/10.1186/s12864-021-08196-8

Choy, J. S., Wei, S., Lee, J. Y., Tan, S., Chu, S., & Lee, T. H. (2010). DNA methylation increases nucleosome compaction and rigidity. *J Am Chem Soc*, *132*(6), 1782-1783. https://doi.org/10.1021/ja910264z

Claycomb, J. M., Batista, P. J., Pang, K. M., Gu, W., Vasale, J. J., van Wolfswinkel, J. C., Chaves, D. A., Shirayama, M., Mitani, S., Ketting, R. F., Conte, D., Jr., & Mello, C. C. (2009). The Argonaute CSR-1 and its 22G-RNA cofactors are required for holocentric chromosome segregation. *Cell*, *139*(1), 123-134. https://doi.org/10.1016/j.cell.2009.09.014

Collins, D. H., Mohorianu, I., Beckers, M., Moulton, V., Dalmay, T., & Bourke, A. F. (2017). MicroRNAs Associated with Caste Determination and Differentiation in a Primitively Eusocial Insect. *Sci Rep*, *7*, 45674. https://doi.org/10.1038/srep45674

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. *Proc Natl Acad Sci U S A*, *107*(8), 3588-3593. https://doi.org/10.1073/pnas.0911685107

Conine, C. C., Moresco, J. J., Gu, W., Shirayama, M., Conte, D., Jr., Yates, J. R., 3rd, & Mello, C. C. (2013). Argonautes promote male fertility and provide a paternal memory of germline gene expression in C. elegans. *Cell*, *155*(7), 1532-1544. https://doi.org/10.1016/j.cell.2013.11.032

Conway, J. R., Lex, A., & Gehlenborg, N. (2017). UpSetR: an R package for the visualization of intersecting sets and their properties. *Bioinformatics*, *33*(18), 2938-2940. https://doi.org/10.1093/bioinformatics/btx364

Cordeiro Rodrigues, R. J., de Jesus Domingues, A. M., Hellmann, S., Dietz, S., de Albuquerque, B. F. M., Renz, C., Ulrich, H. D., Sarkies, P., Butter, F., & Ketting, R. F. (2019). PETISCO is a novel protein complex required for 21U RNA biogenesis and embryonic viability. *Genes Dev*, *33*(13-14), 857-870. https://doi.org/10.1101/gad.322446.118

Corrêa, R. L., Steiner, F. A., Berezikov, E., & Ketting, R. F. (2010). MicroRNA-directed siRNA biogenesis in Caenorhabditis elegans. *PLoS Genet*, *6*(4), e1000903. https://doi.org/10.1371/journal.pgen.1000903

Corsi, A. K., Wightman, B., & Chalfie, M. (2015). A Transparent window into biology: A primer on Caenorhabditis elegans. *WormBook*, 1-31. https://doi.org/10.1895/wormbook.1.177.1

Czech, B., & Hannon, G. J. (2016). One Loop to Rule Them All: The Ping-Pong Cycle and piRNA-Guided Silencing. *Trends in Biochemical Sciences*, *41*(4), 324-337. https://doi.org/https://doi.org/10.1016/j.tibs.2015.12.008

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. *Annu Rev Genet*, *52*, 131-157. https://doi.org/10.1146/annurev-genet-120417-031441

DiRienzo, N., & Dornhaus, A. (2017). Temnothorax rugatulus ant colonies consistently vary in nest structure across time and context. *PLoS One*, *12*(6), e0177598. https://doi.org/10.1371/journal.pone.0177598

Elsner, D., Meusemann, K., & Korb, J. (2018). Longevity and transposon defense, the case of termite reproductives. *Proc Natl Acad Sci U S A*, *115*(21), 5504-5509. https://doi.org/10.1073/pnas.1804046115

Fabian, M. R., Sonenberg, N., & Filipowicz, W. (2010). Regulation of mRNA translation and stability by microRNAs. *Annu Rev Biochem*, *79*, 351-379. https://doi.org/10.1146/annurev-biochem-060308-103103

Fang, W., & Bartel, D. P. (2020). MicroRNA Clustering Assists Processing of Suboptimal MicroRNA Hairpins through the Action of the ERH Protein. *Mol Cell*, *78*(2), 289-302.e286. https://doi.org/10.1016/j.molcel.2020.01.026

Favreau, E., Martínez-Ruiz, C., Rodrigues Santiago, L., Hammond, R. L., & Wurm, Y. (2018). Genes and genomic processes underpinning the social lives of ants. *Curr Opin Insect Sci*, *25*, 83-90. https://doi.org/10.1016/j.cois.2017.12.001

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

Fitz-James, M. H., & Cavalli, G. (2022). Molecular mechanisms of transgenerational epigenetic inheritance. *Nature Reviews Genetics*, *23*(6), 325-341. https://doi.org/10.1038/s41576-021-00438-5

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

Gu, W., Shirayama, M., Conte, D., Jr., Vasale, J., Batista, P. J., Claycomb, J. M., Moresco, J. J., Youngman, E. M., Keys, J., Stoltz, M. J., Chen, C. C., Chaves, D. A., Duan, S., Kasschau, K. D., Fahlgren, N., Yates, J. R., 3rd, Mitani, S., Carrington, J. C., & Mello, C. C. (2009). Distinct argonaute-mediated 22G-RNA pathways direct genome surveillance in the C. elegans germline. *Mol Cell*, *36*(2), 231-244. https://doi.org/10.1016/j.molcel.2009.09.020

Gudipati, R. K., Braun, K., Gypas, F., Hess, D., Schreier, J., Carl, S. H., Ketting, R. F., & Großhans, H. (2021). Protease-mediated processing of Argonaute proteins controls small RNA association. *Mol Cell*, *81*(11), 2388-2402.e2388. https://doi.org/10.1016/j.molcel.2021.03.029

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

Guo, X., Su, S., Skogerboe, G., Dai, S., Li, W., Li, Z., Liu, F., Ni, R., Guo, Y., Chen, S., Zhang, S., & Chen, R. (2013). Recipe for a busy bee: microRNAs in Honey Bee caste determination. *PLoS One*, *8*(12), e81661. https://doi.org/10.1371/journal.pone.0081661

Gupta, A., & Nair, S. (2025). Epigenetic processes in insect adaptation to environmental stress. *Curr Opin Insect Sci*, *67*, 101294. https://doi.org/10.1016/j.cois.2024.101294

Ha, M., & Kim, V. N. (2014). Regulation of microRNA biogenesis. *Nat Rev Mol Cell Biol*, *15*(8), 509-524. https://doi.org/10.1038/nrm3838

Hargitai, B., Kutnyánszky, V., Blauwkamp, T. A., Steták, A., Csankovszki, G., Takács-Vellai, K., & Vellai, T. (2009). xol-1, the master sex-switch gene in C. elegans, is a transcriptional target of the terminal sex-determining factor TRA-1. *Development*, *136*(23), 3881-3887. https://doi.org/10.1242/dev.034637

Heimpel, G. E., & de Boer, J. G. (2008). Sex Determination in the Hymenoptera. *Annual Review of Entomology*, *53*(Volume 53, 2008), 209-230. https://doi.org/https://doi.org/10.1146/annurev.ento.53.103106.093441

Hoogstrate, S. W., Volkers, R. J. M., Sterken, M. G., Kammenga, J. E., & Snoek, L. B. (2014). Nematode endogenous small RNA pathways. *Worm*, *3*(1), e28234. https://doi.org/10.4161/worm.28234

Hutter, H., & Suh, J. (2016). GExplore 1.4: An expanded web interface for queries on Caenorhabditis elegans protein and gene function. *Worm*, *5*(4), e1234659. https://doi.org/10.1080/21624054.2016.1234659

Jin, Z., Flynt, A. S., & Lai, E. C. (2013). Drosophila piwi mutants exhibit germline stem cell tumors that are sustained by elevated Dpp signaling. *Curr Biol*, *23*(15), 1442-1448. https://doi.org/10.1016/j.cub.2013.06.021

Kaufman, E. J., & Miska, E. A. (2010). The microRNAs of Caenorhabditis elegans. *Semin Cell Dev Biol*, *21*(7), 728-737. https://doi.org/10.1016/j.semcdb.2010.07.001

Ketting, R. F. (2011). The many faces of RNAi. *Dev Cell*, *20*(2), 148-161. https://doi.org/10.1016/j.devcel.2011.01.012

Kimble, J., & Hirsh, D. (1979). The postembryonic cell lineages of the hermaphrodite and male gonads in Caenorhabditis elegans. *Developmental Biology*, *70*(2), 396-417. https://doi.org/https://doi.org/10.1016/0012-1606(79)90035-6

Knittel, T. L., Montgomery, B. E., Tate, A. J., Deihl, E. W., Nawrocki, A. S., Hoerndli, F. J., & Montgomery, T. A. (2024). A low-abundance class of Dicer-dependent siRNAs produced from a variety of features in C. elegans. *Genome Res*. https://doi.org/10.1101/gr.279083.124

Kohlmeier, P., Feldmeyer, B., & Foitzik, S. (2023). Histone acetyltransferases and external demands influence task switching in Temnothorax ants. *Biol Lett*, *19*(7), 20230176. https://doi.org/10.1098/rsbl.2023.0176

Kumar, A. (2020). Jump around: transposons in and out of the laboratory. *F1000Res*, *9*. https://doi.org/10.12688/f1000research.21018.1

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

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., & Higgins, D. G. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics*, *23*(21), 2947-2948. https://doi.org/10.1093/bioinformatics/btm404

Leighton, G. M., Charbonneau, D., & Dornhaus, A. (2017). Task switching is associated with temporal delays in Temnothorax rugatulus ants. *Behav Ecol*, *28*(1), 319-327. https://doi.org/10.1093/beheco/arw162

Letunic, I., & Bork, P. (2024). Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Research*, *52*(W1), W78-W82. https://doi.org/10.1093/nar/gkae268

Liao, Y., Smyth, G. K., & Shi, W. (2013). The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Res*, *41*(10), e108. https://doi.org/10.1093/nar/gkt214

Lin, H., & Spradling, A. C. (1997). A novel group of pumilio mutations affects the asymmetric division of germline stem cells in the Drosophila ovary. *Development*, *124*(12), 2463-2476. https://doi.org/10.1242/dev.124.12.2463

Liu, L., Wang, X., Zhao, W., Li, Q., Li, J., Chen, H., & Shan, G. (2023). Systematic characterization of small RNAs associated with C. elegans Argonautes. *Sci China Life Sci*, *66*(6), 1303-1322. https://doi.org/10.1007/s11427-022-2304-8

Liu, M., Huang, J., Zhang, G., Liu, X., & An, J. (2019). Analysis of miRNAs in the Heads of Different Castes of the Bumblebee Bombus lantschouensis (Hymenoptera: Apidae). *Insects*, *10*(10). https://doi.org/10.3390/insects10100349

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

Macedo, L. M., Nunes, F. M., Freitas, F. C., Pires, C. V., Tanaka, E. D., Martins, J. R., Piulachs, M. D., Cristino, A. S., Pinheiro, D. G., & Simões, Z. L. (2016). MicroRNA signatures characterizing caste-independent ovarian activity in queen and worker honeybees (Apis mellifera L.). *Insect Mol Biol*, *25*(3), 216-226. https://doi.org/10.1111/imb.12214

Madeira, F., Madhusoodanan, N., Lee, J., Eusebi, A., Niewielska, A., Tivey, A. R. N., Lopez, R., & Butcher, S. (2024). The EMBL-EBI Job Dispatcher sequence analysis tools framework in 2024. *Nucleic Acids Research*, *52*(W1), W521-W525. https://doi.org/10.1093/nar/gkae241

Makeyeva, Y. V., Shirayama, M., & Mello, C. C. (2021). Cues from mRNA splicing prevent default Argonaute silencing in C. elegans. *Dev Cell*, *56*(18), 2636-2648.e2634. https://doi.org/10.1016/j.devcel.2021.08.022

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal; Vol 17, No 1: Next Generation Sequencing Data Analysis*. https://doi.org/10.14806/ej.17.1.200

Martinez, B., & Peplow, P. V. (2024). MicroRNAs as potential biomarkers for diagnosis of attention deficit hyperactivity disorder. *Neural Regen Res*, *19*(3), 557-562. https://doi.org/10.4103/1673-5374.380880

Matsunami, M., Nozawa, M., Suzuki, R., Toga, K., Masuoka, Y., Yamaguchi, K., Maekawa, K., Shigenobu, S., & Miura, T. (2019). Caste-specific microRNA expression in termites: insights into soldier differentiation. *Insect Mol Biol*, *28*(1), 86-98. https://doi.org/10.1111/imb.12530

Meyer, B. J. (2022). The X chromosome in C. elegans sex determination and dosage compensation. *Curr Opin Genet Dev*, *74*, 101912. https://doi.org/10.1016/j.gde.2022.101912

Nan, X., Ng, H. H., Johnson, C. A., Laherty, C. D., Turner, B. M., Eisenman, R. N., & Bird, A. (1998). Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature*, *393*(6683), 386-389. https://doi.org/10.1038/30764

Negroni, M. A., Feldmeyer, B., & Foitzik, S. (2021). Experimental increase in fecundity causes upregulation of fecundity and body maintenance genes in the fat body of ant queens. *Biol Lett*, *17*(2), 20200909. https://doi.org/10.1098/rsbl.2020.0909

Negroni, M. A., Foitzik, S., & Feldmeyer, B. (2019). Long-lived Temnothorax ant queens switch from investment in immunity to antioxidant production with age. *Sci Rep*, *9*(1), 7270. https://doi.org/10.1038/s41598-019-43796-1

Negroni, M. A., Segers, F., Vogelweith, F., & Foitzik, S. (2020). Immune challenge reduces gut microbial diversity and triggers fertility-dependent gene expression changes in a social insect. *BMC Genomics*, *21*(1), 816. https://doi.org/10.1186/s12864-020-07191-9

Ortiz, M. A., Noble, D., Sorokin, E. P., & Kimble, J. (2014). A new dataset of spermatogenic vs. oogenic transcriptomes in the nematode Caenorhabditis elegans. *G3 (Bethesda)*, *4*(9), 1765-1772. https://doi.org/10.1534/g3.114.012351

Ouyang, J. P. T., Zhang, W. L., & Seydoux, G. (2022). The conserved helicase ZNFX-1 memorializes silenced RNAs in perinuclear condensates. *Nat Cell Biol*, *24*(7), 1129-1140. https://doi.org/10.1038/s41556-022-00940-w

Palli, S. R. (2023). RNAi turns 25:contributions and challenges in insect science. *Front Insect Sci*, *3*, 1209478. https://doi.org/10.3389/finsc.2023.1209478

Papadaki, M., Kaitetzidou, E., Mylonas, C. C., & Sarropoulou, E. (2020). Non-coding RNA Expression Patterns of Two Different Teleost Gonad Maturation Stages. *Mar Biotechnol (NY)*, *22*(5), 683-695. https://doi.org/10.1007/s10126-020-09991-2

Papadaki, M., Mylonas, C. C., & Sarropoulou, E. (2024). MicroRNAs are involved in ovarian physiology of greater amberjack (Seriola dumerili) under captivity. *Gen Comp Endocrinol*, *357*, 114581. https://doi.org/10.1016/j.ygcen.2024.114581

Peixoto, P., Cartron, P. F., Serandour, A. A., & Hervouet, E. (2020). From 1957 to Nowadays: A Brief History of Epigenetics. *Int J Mol Sci*, *21*(20). https://doi.org/10.3390/ijms21207571

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 Dev*, *26*(13), 1433-1444. https://doi.org/10.1101/gad.193904.112

Phillips, C. M., & Updike, D. L. (2022). Germ granules and gene regulation in the Caenorhabditis elegans germline. *Genetics*, *220*(3). https://doi.org/10.1093/genetics/iyab195

Podvalnaya, N., Bronkhorst, A. W., Lichtenberger, R., Hellmann, S., Nischwitz, E., Falk, T., Karaulanov, E., Butter, F., Falk, S., & Ketting, R. F. (2023). piRNA processing by a trimeric Schlafen-domain nuclease. *Nature*, *622*(7982), 402-409. https://doi.org/10.1038/s41586-023-06588-2

Portman, D. S. (2007). Genetic control of sex differences in C. elegans neurobiology and behavior. *Adv Genet*, *59*, 1-37. https://doi.org/10.1016/s0065-2660(07)59001-2

Post, F., Bornberg-Bauer, E., Vasseur-Cognet, M., & Harrison, M. C. (2023). More effective transposon regulation in fertile, long-lived termite queens than in sterile workers. *Mol Ecol*, *32*(2), 369-380. https://doi.org/10.1111/mec.16753

Quarato, P., Singh, M., Cornes, E., Li, B., Bourdon, L., Mueller, F., Didier, C., & Cecere, G. (2021). Germline inherited small RNAs facilitate the clearance of untranslated maternal mRNAs in C. elegans embryos. *Nat Commun*, *12*(1), 1441. https://doi.org/10.1038/s41467-021-21691-6

Quinlan, A. R., & Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, *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., Heyne, S., Dündar, F., & 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

Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., & Mesirov, J. P. (2011). Integrative genomics viewer. *Nature Biotechnology*, *29*(1), 24-26. https://doi.org/10.1038/nbt.1754

Rüppell, O., Heinze, J., & Hölldobler, B. (2001). Complex determination of queen body size in the queen size dimorphic ant Leptothorax rugatulus (Formicidae: Hymenoptera). *Heredity (Edinb)*, *87*(Pt 1), 33-40. https://doi.org/10.1046/j.1365-2540.2001.00904.x

Rüppell, O., Strätz, M., Baier, B., & Heinze, J. (2003). Mitochondrial markers in the ant Leptothorax rugatulus reveal the population genetic consequences of female philopatry at different hierarchical levels. *Mol Ecol*, *12*(3), 795-801. https://doi.org/10.1046/j.1365-294x.2003.01769.x

Scholl, A., Liu, Y., & Seydoux, G. (2024). Caenorhabditis elegans germ granules accumulate hundreds of low translation mRNAs with no systematic preference for germ cell fate regulators. *Development*, *151*(13). https://doi.org/10.1242/dev.202575

Schreier, J., Dietz, S., Boermel, M., Oorschot, V., Seistrup, A. S., de Jesus Domingues, A. M., Bronkhorst, A. W., Nguyen, D. A. H., Phillis, S., Gleason, E. J., L'Hernault, S. W., Phillips, C. M., Butter, F., & Ketting, R. F. (2022). Membrane-associated cytoplasmic granules carrying the Argonaute protein WAGO-3 enable paternal epigenetic inheritance in Caenorhabditis elegans. *Nat Cell Biol*, *24*(2), 217-229. https://doi.org/10.1038/s41556-021-00827-2

Schreier, J., Kielisch, F., & Ketting, R. F. (2024). A genetic framework for RNAi inheritance in <em>Caenorhabditis elegans</em&gt. *bioRxiv*, 2024.2010.2002.616260. https://doi.org/10.1101/2024.10.02.616260

Sendoel, A., Subasic, D., Ducoli, L., Keller, M., Michel, E., Kohler, I., Singh, K. D., Zheng, X., Brümmer, A., Imig, J., Kishore, S., Wu, Y., Kanitz, A., Kaech, A., Mittal, N., Matia-González, A. M., Gerber, A. P., Zavolan, M., Aebersold, R.,...Hengartner, M. O. (2019). MINA-1 and WAGO-4 are part of regulatory network coordinating germ cell death and RNAi in C. elegans. *Cell Death Differ*, *26*(10), 2157-2178. https://doi.org/10.1038/s41418-019-0291-z

Seroussi, U., Lugowski, A., Wadi, L., Lao, R. X., Willis, A. R., Zhao, W., Sundby, A. E., Charlesworth, A. G., Reinke, A. W., & Claycomb, J. M. (2023). A comprehensive survey of C. elegans argonaute proteins reveals organism-wide gene regulatory networks and functions. *Elife*, *12*. https://doi.org/10.7554/eLife.83853

Seth, M., Shirayama, M., Gu, W., Ishidate, T., Conte, D., Jr., & Mello, C. C. (2013). The C. elegans CSR-1 argonaute pathway counteracts epigenetic silencing to promote germline gene expression. *Dev Cell*, *27*(6), 656-663. https://doi.org/10.1016/j.devcel.2013.11.014

Shevchenko, A., Tomas, H., Havlis, J., Olsen, J. V., & Mann, M. (2006). In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat Protoc*, *1*(6), 2856-2860. https://doi.org/10.1038/nprot.2006.468

Shi, C., & Murphy, C. T. (2021). Sex and death. *Curr Top Dev Biol*, *144*, 353-375. https://doi.org/10.1016/bs.ctdb.2020.08.004

Shirayama, M., Stanney, W., 3rd, Gu, W., Seth, M., & Mello, C. C. (2014). The Vasa Homolog RDE-12 engages target mRNA and multiple argonaute proteins to promote RNAi in C. elegans. *Curr Biol*, *24*(8), 845-851. https://doi.org/10.1016/j.cub.2014.03.008

Shukla, A., Yan, J., Pagano, D. J., Dodson, A. E., Fei, Y., Gorham, J., Seidman, J. G., Wickens, M., & Kennedy, S. (2020). poly(UG)-tailed RNAs in genome protection and epigenetic inheritance. *Nature*, *582*(7811), 283-288. https://doi.org/10.1038/s41586-020-2323-8

Soares, M. P. M., Pinheiro, D. G., de Paula Freitas, F. C., Simões, Z. L. P., & Bitondi, M. M. G. (2021). Transcriptome dynamics during metamorphosis of imaginal discs into wings and thoracic dorsum in Apis mellifera castes. *BMC Genomics*, *22*(1), 756. https://doi.org/10.1186/s12864-021-08040-z

Suter, B. (2018). RNA localization and transport. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, *1861*(10), 938-951. https://doi.org/https://doi.org/10.1016/j.bbagrm.2018.08.004

Tabaczar, S., Czogalla, A., Podkalicka, J., Biernatowska, A., & Sikorski, A. F. (2017). Protein palmitoylation: Palmitoyltransferases and their specificity. *Exp Biol Med (Maywood)*, *242*(11), 1150-1157. https://doi.org/10.1177/1535370217707732

Tabara, H., Mitani, S., Mochizuki, M., Kohara, Y., & Nagata, K. (2023). A small RNA system ensures accurate homologous pairing and unpaired silencing of meiotic chromosomes. *Embo j*, *42*(11), e105002. https://doi.org/10.15252/embj.2020105002

Tate, P. H., & Bird, A. P. (1993). Effects of DNA methylation on DNA-binding proteins and gene expression. *Curr Opin Genet Dev*, *3*(2), 226-231. https://doi.org/10.1016/0959-437x(93)90027-m

Trifinopoulos, J., Nguyen, L.-T., von Haeseler, A., & Minh, B. Q. (2016). W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Research*, *44*(W1), W232-W235. https://doi.org/10.1093/nar/gkw256

Tsai, H. Y., Chen, C. C., Conte, D., Jr., Moresco, J. J., Chaves, D. A., Mitani, S., Yates, J. R., 3rd, Tsai, M. D., & Mello, C. C. (2015). A ribonuclease coordinates siRNA amplification and mRNA cleavage during RNAi. *Cell*, *160*(3), 407-419. https://doi.org/10.1016/j.cell.2015.01.010

Tsuboyama, K., Tadakuma, H., & Tomari, Y. (2018). Conformational Activation of Argonaute by Distinct yet Coordinated Actions of the Hsp70 and Hsp90 Chaperone Systems. *Mol Cell*, *70*(4), 722-729.e724. https://doi.org/10.1016/j.molcel.2018.04.010

Tyc, K. M., Nabih, A., Wu, M. Z., Wedeles, C. J., Sobotka, J. A., & Claycomb, J. M. (2017). The Conserved Intron Binding Protein EMB-4 Plays Differential Roles in Germline Small RNA Pathways of C. elegans. *Dev Cell*, *42*(3), 256-270.e256. https://doi.org/10.1016/j.devcel.2017.07.003

Valentini, G., Masuda, N., Shaffer, Z., Hanson, J. R., Sasaki, T., Walker, S. I., Pavlic, T. P., & Pratt, S. C. (2020). Division of labour promotes the spread of information in colony emigrations by the ant Temnothorax rugatulus. *Proc Biol Sci*, *287*(1924), 20192950. https://doi.org/10.1098/rspb.2019.2950

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., Duchaine, T. F., Mello, C. C., & Conte, D., Jr. (2010). Sequential rounds of RNA-dependent RNA transcription drive endogenous small-RNA biogenesis in the ERGO-1/Argonaute pathway. *Proc Natl Acad Sci U S A*, *107*(8), 3582-3587. https://doi.org/10.1073/pnas.0911908107

Vastenhouw, N. L., Fischer, S. E., Robert, V. J., Thijssen, K. L., Fraser, A. G., Kamath, R. S., Ahringer, J., & Plasterk, R. H. (2003). A genome-wide screen identifies 27 genes involved in transposon silencing in C. elegans. *Curr Biol*, *13*(15), 1311-1316. https://doi.org/10.1016/s0960-9822(03)00539-6

Verhulst, E. C., Beukeboom, L. W., & van de Zande, L. (2010). Maternal control of haplodiploid sex determination in the wasp Nasonia. *Science*, *328*(5978), 620-623. https://doi.org/10.1126/science.1185805

Vieira, J., de Paula Freitas, F. C., Santos Cristino, A., Guariz Pinheiro, D., Aguiar, L. R., Framartino Bezerra Laure, M. A., Rosatto Moda, L. M., Paulino Simões, Z. L., & Barchuk, A. R. (2021). Molecular underpinnings of the early brain developmental response to differential feeding in the honey bee Apis mellifera. *Biochim Biophys Acta Gene Regul Mech*, *1864*(9), 194732. https://doi.org/10.1016/j.bbagrm.2021.194732

Wan, G., Fields, B. D., Spracklin, G., Shukla, A., Phillips, C. M., & Kennedy, S. (2018). Spatiotemporal regulation of liquid-like condensates in epigenetic inheritance. *Nature*, *557*(7707), 679-683. https://doi.org/10.1038/s41586-018-0132-0

Watson, O. T., Buchmann, G., Young, P., Lo, K., Remnant, E. J., Yagound, B., Shambrook, M., Hill, A. F., Oldroyd, B. P., & Ashe, A. (2022). Abundant small RNAs in the reproductive tissues and eggs of the honey bee, Apis mellifera. *BMC Genomics*, *23*(1), 257. https://doi.org/10.1186/s12864-022-08478-9

Wedeles, C. J., Wu, M. Z., & Claycomb, J. M. (2013). Protection of germline gene expression by the C. elegans Argonaute CSR-1. *Dev Cell*, *27*(6), 664-671. https://doi.org/10.1016/j.devcel.2013.11.016

Wolffe, A. P., & Guschin, D. (2000). Review: chromatin structural features and targets that regulate transcription. *J Struct Biol*, *129*(2-3), 102-122. https://doi.org/10.1006/jsbi.2000.4217

Wu, J., Yang, J., Cho, W. C., & Zheng, Y. (2020). Argonaute proteins: Structural features, functions and emerging roles. *J Adv Res*, *24*, 317-324. https://doi.org/10.1016/j.jare.2020.04.017

Xu, F., Feng, X., Chen, X., Weng, C., Yan, Q., Xu, T., Hong, M., & Guang, S. (2018). A Cytoplasmic Argonaute Protein Promotes the Inheritance of RNAi. *Cell Rep*, *23*(8), 2482-2494. https://doi.org/10.1016/j.celrep.2018.04.072

Yan, Q. (2025). Personality and Psychoneuroimmunology: A Systems Biology Perspective. *Methods Mol Biol*, *2868*, 15-36. https://doi.org/10.1007/978-1-0716-4200-9_2

Yu, G., Wang, L.-G., Han, Y., & He, Q.-Y. (2012). clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *OMICS: A Journal of Integrative Biology*, *16*(5), 284-287. https://doi.org/10.1089/omi.2011.0118

Zheng, Y. B., & Jin, X. (2024). Evidence for the Contribution of the miR-206/BDNF Pathway in the Pathophysiology of Depression. *Int J Neuropsychopharmacol*, *27*(10). https://doi.org/10.1093/ijnp/pyae039

Zhu, A., Ibrahim, J. G., & Love, M. I. (2019). Heavy-tailed prior distributions for sequence count data: removing the noise and preserving large differences. *Bioinformatics*, *35*(12), 2084-2092. https://doi.org/10.1093/bioinformatics/bty895

Zhu, K. Y., & Palli, S. R. (2020). Mechanisms, Applications, and Challenges of Insect RNA Interference. *Annu Rev Entomol*, *65*, 293-311. https://doi.org/10.1146/annurev-ento-011019-025224

Acknowledgements

I would like to start by thanking René for hiring me, since it's very difficult to finish a Ph.D. thesis without a Ph.D. position to begin with. When I came to Mainz for my interview back in 2019, I did not think that I would actually accept a position should it be offered to me, and I had never even heard the name René Ketting. Still, as faith would have it, we did find each other, and I will always be grateful that he decided to take a chance with me. René has always been an excellent supervisor, an even greater scientist, and a genuinely nice person and boss. The atmosphere at work has always been pleasant, for which I owe my thanks both to René and to all of the members of the Ketting lab along with IMB and the IPP as a whole.

I hope that the members of the Ketting lab not mentioned by name here will forgive my omission, and know that their presence, inputs, and support were valued no less from it.

The first person from our lab that I do wish to mention by name is [redacted]. Although he left the lab before my joining – a betrayal for which I can hardly judge him, as one should always put obligations to family and personal happiness above all else – I owe a lot to the joined efforts of him and [redacted] along with [redacted]'s entire team at the bioinformatics core facilities at IMB. In general, I would like to thank all the core facilities – bioinformatics, flow cytometry, proteomics, and, in particular, genomics, for their support. Although not all of their work has made it to publications, their efforts have been greatly appreciated. Had I been a more religious person, I'd might have had to seek repentance from the Lord for all of the projects I have scrapped over time. Instead, I hope to obtain the forgiveness of [redacted] for not making better use of all their hard work.

In addition, I owe my thanks to my TAC members, [redacted], for always providing me with valuable insights.

I should also want to extend my gratitude towards [redacted], who has been both mentor and sparring partner on my journey. Although I would like to believe that I have taught him a few things along the way, there is no mistaken that I owe him much more than he does me. A huge thanks to [redacted] both, for allowing me to send way too much work their ways.

Thanks to [redacted] for helping me with my mass spectrometry and for always being very patient with me.

Thanks to [redacted] for all of her work in the lab, for showing me the ropes, and for all of our adventures in Gloomhaven.

On the more social side of things, I have to thank [redacted] for being my friend from day one. His silly antics and pleasant person made coming to work every day a joy. I am sorry that I am not sorrier for so often preferring the comforts of working from home.

To all of the people that started their PhD journeys alongside me – [redacted] – I also owe thanks. Even if they all managed to conclude their journeys ahead of me, they have been my base, providing me with better support and friendship than one could hope to ask for.

Speaking of friendships, I also have to extend my gratitude to my friends and family who stuck with me back in [redacted]. Even in the age of internet and videocalls, maintaining a relationship over long distances and across country borders can be difficult and I know that I have not made it easy. A special thanks goes to [redacted] for somehow believing me better than I am and for forgiving my many character flaws.

[redacted]

With all of my official thanks concluded, I would like to end by thanking myself for sticking with it. It is a surprise to no one that a lot of work goes into concluding a Ph.D. thesis, but even so, it can be hard to grasp the underlying hardships when only presented with the finished product. If a doctoral researcher who has never thought about quitting exist, I have yet to meet them, with many kooky ideas having been presented as valid alternatives. I doubt that many will have made it this far through these acknowledgements, but if you have, and if you are currently in the midst of things: *You can do it!* Try to measure your time in successes rather than failures and remember that you are worth so much more than the sum of all your shortcomings.

Even if you, after nearly six years, still cannot confidently genotype a single strain.

– CURRICULUM VITAE –

Contact Information

Name: Ann-Sophie Seistrup

[redacted]

Work Experience

2019 – Now **Institute of Molecular Biology (IMB) Mainz**
Doctoral Researcher

2014 – 2018 **University of Southern Denmark**
Teaching Assistant

Education

2019 – Now **Ph.D., Biology**
Johannes Gutenberg University, Mainz &
International PhD Programme on Gene Regulation, Epigenetics & Genome Stability, Mainz, Germany
Ph.D. thesis: "Regulatory small RNA pathways in the nematode *C. elegans* and the ant *T. rugatulus*"
Supervisor: Prof. Dr. René F. Ketting

2016 – 2018 **M.Sc., Computational Biomedicine**
University of Southern Denmark
Master's thesis: "A Statistical Approach to Pseudoexon Investigation"
Supervisor: [redacted]

2013 – 2016 **B.Sc., Biochemistry and Molecular Biology**
University of Southern Denmark
Bachelor's thesis: "Homologues of the Methyltransferase RsmF in Thermophilic Bacteria"
Supervisor: [redacted]

Publications

2023 **Age- and caste-independent piRNAs in the germline and miRNA profiles linked to caste and fecundity in the ant *Temnothorax rugatulus***

Ann-Sophie Seistrup, Marina Choppin, Shamitha Govind, Barbara Feldmeyer, Marion Kever, Emil Karaulanov, Alice Séguret, Sivarajan Karunanithi, Miguel V. Almeida, René F. Ketting, Susanne Foitzik

Mol Ecol. 2023 Nov;32(22):6027-6043.
doi: 10.1111/mec.17162. Epub 2023 Oct 13. PMID: 37830492.

2022 **Membrane-associated cytoplasmic granules carrying the Argonaute protein WAGO-3 enable paternal epigenetic inheritance in Caenorhabditis elegans**

Jan Schreier, Sabrina Dietz, Mandy Boermel, Viola Oorschot, **Ann-Sophie Seistrup**, Antonio M. de Jesus Domingues, Alfred W. Bronkhorst, Dieu An H. Nguyen, Stephanie Phillis, Elizabeth J. Gleason, Steven W. L'Hernault, Carolyn M. Phillips, Falk Butter, René F. Ketting

Nat Cell Biol. 2022 Feb;24(2):217-229.
doi: 10.1038/s41556-021-00827-2. Epub 2022 Feb 7. PMID: 35132225

Volunteer Work

2022 – 2024 **Pint of Science Germany e.V.**
Event coordinator and website manager using the platform Silverstripe (Mainz).

2019 **International Medical Cooperation Committee (IMCC)**
Local board member and financial controller (Esbjerg).

2015 – 2020 **Sexekspressen**
Sex educator of Danish school grades 7 through 10 as well as at-risk adults (Odense/Esbjerg).
Administrative internal position; responsible for booking and planning (Esbjerg).