Characterization of RNAi Components Involved in Endogenous and Exogenous RNAi Pathways in *Paramecium tetraurelia*

Dissertation

Zur Erlangung des Doktorgrades der Naturwissenschaft (Dr. rer. nat.)

Der Fakultät für Mathematik und Naturwissenschaften der

Bergischen Universität Wuppertal

Angefertigt in der

Arbeitsgruppe für Molekulare Zellbiologie und Mikrobiologie

Vorgelegt von

Marcello Pirritano

Wuppertal, 2024

Dekan:in: Prof. Dr. Francesco Knechtli

Studiendekan:in: Prof. Dr. Fabian Mohr

Erste:r Gutachter:in: Prof. Dr. Martin Simon

Zweite:r Gutachter:in: Prof. Dr. Julia Bornhorst

Tag der mündlichen Prüfung: 06.12.2024

“You twine your life around people you love. And when they are gone, you grow around their absence instead. It is just another way they shape you.”

Jaheira, High Harper

Abstract

RNAi by Feeding in *Paramecium tetraurelia* has been used in molecular biology labs around the world as a tool to generate simple knockdowns of genes. But apart from the enzymes involved in this pathway, little is known about their actual functions. In this thesis, the various components of the RNAi machinery have been studied to shed light on their involvement during the different steps of feeding-associated siRNA biogenesis.

It was shown that the two RdRPs involved in the pathway, RdRP1 and RdRP2, interact with the exogenous dsRNA trigger and use it as a template to generate new RNA strands which are then further processed by Dicer1. This represents the first example of RdRPs within an exogenously triggered RNAi pathway being responsible for primary siRNA production and interacting with the exogenous trigger RNA directly.

Beside RdRPs, the function of Ptiwi proteins within the RNAi by Feeding pathway as well as the endogenous RNAi pathway was analyzed. While the three Ptiwis, Ptiwi12, Ptiwi13 and Ptiwi15, have been associated with the pathway in previous publications, this work showed that Ptiwi14 is also involved in this pathway, loading both primary siRNAs and secondary siRNAs, while the aforementioned three Ptiwis only associate with primary siRNAs. This increases the likelihood of the RNAi by Feeding pathway in *Paramecium tetraurelia* influencing the cellular chromatin organization, as already described in other species like *C. elegans*.

Additionally, this work analyzed the two *paramecium* specific proteins Pds1 and Pds2, so far only described as involved in the pathway, but not characterized at all. Pds1 localizes in the cytoplasm of the cell and remains elusive, while Pds2 was localized in cellular membranes, harboring potential to be involved in dsRNA uptake.

Last, untemplated nucleotides added to primary siRNAs produced in the RNAi by Feeding pathway were analyzed. This led to the hypothesis of poly-uridylation serving as an RNA degradation signal, supporting the degradation of the passenger strand from the primary siRNA duplex.

Contents

List of Figures

List of Supplementary Figures

List of Tables

List of Supplementary Tables

List of Abbreviations

approx. ..... approximately
APS ..... Ammonium Persulphate
BSA ..... Bovine serum albumin
C. elegans ..... Caenorhabditis elegans
DAPI ..... 4′,6-diamidino-2-phenylindole
DNA ..... Deoxy ribonucleic acid
dsDNA ..... Double stranded DNA
dsRNA ..... Double stranded RNA
E. coli ..... Escherichia coli
EDTA ..... Ethylenediaminetetraacetic Acid
EEJ ..... Exon-Exon-Junction, exon-exon junction
GFP ..... Green fluorescent protein
HD ..... Heteroduplex
IES ..... Internal eliminated sequences
IFA ..... Immunofluorescence assay
IPTG ..... Isopropyl-β-D-thiogalactopyranoside
KD ..... Knockdown
miRNA ..... microRNA
mRNA ..... messenger RNA
NEB ..... New England Biolabs
nt ..... Nucleotide
OD ..... Optical density
PAGE ..... Poly acrylamide gel electrophoresis
PCR ..... Polymerase chain reaction
piRNA ..... piwi-interacting RNA
pre-miRNA ..... precursor miRNA
PTGS ..... Post-transcriptional gene silencing
RdRP ..... RNA dependent RNA polymerase
RISC ..... RNA induced silencing complex
RITS ..... RNA induced transcriptional silencing
RNA ..... Ribonucleic acid
RNA IP ..... RNA immunoprecipitation
RNAi ..... RNA interference
RT ..... Room temperature
scnRNA ..... scan RNA
SDS ..... Sodium dodecyl sulphate
SHD ..... Switched Heteroduplex
siRNA ..... small interfering RNA
SRC ..... siRNA producing cluster
ssRNA ..... Single stranded RNA
TEMED ..... N,N,N′,N′-Tetramethyl ethylenediamine
vol ..... Volume
WGP ..... Wheat grass powder

1 Background

1.1 RNA Interference

RNA interference (or RNAi) is a broad term that is used to describe several different pathways that involve regulation of gene expression guided by a smallRNA (21-27 nucleotides in size). Usually, this regulation of gene expression results in a repression of the target gene, leading to a loss of the gene product, and therefore a loss of the gene's phenotype. This gene silencing is mediated by the smallRNA in a homology dependent manner, meaning that smallRNA and mRNA are reverse complementary to each other, which guides the RNAi machinery to its target.

The way this repression is achieved can vary between different RNAi pathways, which leads to a categorization of those pathways according to their target molecule and mode of operation.

The first RNAi pathway that was publicly recognized and later awarded with a Nobel prize in 2006 was studied in the nematode *Caenorhabditis elegans* and described by Andrew Fire and Craig Mello (Fire et al., 1998). They injected dsRNA sharing sequence homology with a target transcript into the animal, which led to a silencing of the targeted gene on a phenotypical level. Later, they found that the dsRNA required for this interference could also be applied via food bacteria, a mechanism that would become known as RNAi by Feeding (Timmons et al., 2001).

However, dsRNA induced RNAi in *C. elegans* was not the first RNAi mediated phenomenon ever described. In 1956, Alexander Brink described a phenomenon later called paramutation, where the inheritance of a specific allele responsible for anthocyanin synthesis in maize lead to an impairment of the synthesis (Brink, 1956). Inheritance of this allele didn't abide to Mendel's laws and it was later shown that it wasn't a genetic effect that was responsible for the impaired anthocyanin production, but a smallRNA mediated, inheritable epigenetic effect silencing both wildtype alleles present in the plant (Alleman et al., 2006).

In addition, two gene silencing phenomena were described in *Neurospora crassa* and Petunia, called quelling and co-suppression respectively. Both of these phenomena were induced by the introduction of transgene constructs, exogenous DNA that carries information of an open reading frame that is introduced into the nucleus of cells to induce the expression of the open reading frame. After further research, it was revealed that both of the described phenomena were mediated by smallRNA, too (Fagard et al., 2000; M. K. Montgomery & Fire, 1998; Napoli et al., 1990; Romano & Macino, 1992).

1.1.1 The Mode of Operation of RNAi

While all RNAi pathways are mediated by smallRNAs, their target for realizing phenotypical gene silencing can vary.

In general, the goal of RNA mediated silencing is to prevent the production of the gene product, usually protein, to realize the gene's phenotypical silencing. This can be achieved in two different ways, either preventing the formation of protein by degrading or blocking the mRNA in the cell, or by preventing the production of mRNA.

Inhibition of translation or degradation of mRNA is generally referred to as post-transcriptional gene silencing (PTGS), since it happens after mRNA is produced from the target. Here, the smallRNA responsible for target recognition is loaded by an argonaute protein, forming the so-called RNA induced silencing complex (RISC). This complex uses the loaded smallRNA as a guide and scours the mRNA population in the cell to find mRNA molecules sharing sequence homology with the smallRNA. If the RISC finds a mRNA matching this criterium, the mRNA gets either degraded or blocked, preventing translation and realizing the phenotypical silencing of the target gene (Hammond et al., 2001).

In contrast, RNAi can also prevent the production of mRNA, a mode of operation called co-transcriptional gene silencing (CTGS). This is realized by establishing a heterochromatic state at the genomic locus, lowering the accessibility of the DNA and therefore preventing the mRNA production by RNA Polymerase II. This process is best described in fission yeast, *Schizosaccharomyces pombe*, where smallRNAs are loaded into the argonaute protein, forming the RNA induced transcriptional silencing (RITS) complex. Here, the loaded smallRNA guides the RITS complex to nascent transcripts of the target's genomic locus in a homology dependent manner (Bühler et al., 2006). Binding of the RITS complex to the nascent transcript recruits several different chromatin remodeling enzymes, which establishes a heterochromatic state at the genomic locus and therefore, silences the gene (Martienssen & Moazed, 2015).

Interlude – Chromatin

DNA within the nucleus of a cell is not naked, but it is packaged by several proteins, forming the so-called nucleosome. These nucleosomes consist of two from each of the four histones, H2A, H2B, H3 and H4, forming a hetero-octamer that the DNA is wrapped around. DNA and nucleosomes together form the chromatin, a term coined to describe the packaging state of the DNA.

However, packaging of DNA is not the only function chromatin harbors. Instead, the chromatin state of the DNA is used to regulate and alter gene expression by making the DNA, and by proxy genes, more or less accessible to the transcriptional machinery. Chromatin states are not the same across the whole genome, making it possible to reduce accessibility of certain areas of the genome while at the same time enhancing accessibility at other loci. These different chromatin states are called heterochromatin, referring to the inaccessible state, and euchromatin, referring to the accessible state.

Establishment of a specific chromatin state as well as conversion of one state into the other is realized by modifications of the histone proteins. Many different modifications, especially placed at the N-terminal tail of histone H3 and H4 are known, however acetylation and methylation are the two modifications most commonly studied (Kouzarides, 2007). Here, acetylation is usually associated with accessible euchromatin, promoting gene expression, while methylation is usually associated with inaccessible heterochromatin, leading to transcriptional silencing of genes (Sterner & Berger, 2000; Z. Wang et al., 2008; Y. Zhang & Reinberg, 2001). However, this is not always a black and white situation, as these modifications can be placed at several different residues of the histone proteins and co-occur with each other. This leads to chromatin formation being orchestrated by a combination of multiple modifications at individual residues of the different histones, making understanding chromatin establishment very complicated (Liu et al., 2005; Pokholok et al., 2005). Despite that, certain modifications can be coupled to certain expression states, like acetylation of lysine 9 of histone 3 (denoted as H3K9ac) being linked to euchromatin and active gene expression or tri-methylation of histone 3 at lysine 27 (denoted as H3K27me3) being associated with heterochromatin (Agalioti et al., 2002; Boros et al., 2014; Z. Wang et al., 2008). These associations hint towards the existence of a so called “histone code”, linking specific histone modifications to specific chromatin states.

To convert chromatin from one state to the other and to propagate the information encoded in the histone modifications, specific reader and writer enzymes have to exist. Here, reader enzymes refer to enzymes capable of "reading" histone modifications and altering the transcriptional machinery according to the chromatin state, while writer enzymes refer to enzymes capable of placing histone modifications. Reader of histone methylation are proteins harboring chromo domains and proteins of the Tudor domain family, while histone acetylation is read by proteins with bromo domains (Bannister et al., 2001; M. H. Jones et al., 2000; Maurer-Stroh et al., 2003). Writer of histone acetylation are histone acetyltransferases, while methylation is placed by histone methyltransferases (Sterner & Berger, 2000; Y. Zhang & Reinberg, 2001). Both, placement and removal of histone modifications can be triggered by the cell in a variety of ways, one of which is the triggering of heterochromatin formation by RNAi, connecting the transcriptional control of gene expression with the pathway most famous for mRNA degradation (Volpe et al., 2002).

1.1.2 Different smallRNA Classes in RNAi and Their Biogenesis

As described previously, all RNAi pathways depend on a smallRNA guiding the effector complex to its target. However, not all smallRNAs are created equal, and different classes of smallRNAs can achieve different RNAi responses.

As a first class, small interfering RNAs (siRNAs) can be named. The precursor of siRNAs are long double stranded RNAs (dsRNAs). These dsRNAs can have various different origins, being either provided by the cell (due to e.g. bi-directional transcription of a genomic locus) or have exogenous origin (for example in the RNAi by Feeding pathway) (Czech et al., 2008). The dsRNA is cleaved by an enzyme called Dicer into 22-27nt long siRNA duplexes, carrying a 2nt 3' overhang (Czech et al., 2008). The specific size of this siRNA duplex varies between different organisms. One of the duplex strands is considered the guide strand, which is loaded into the argonaute protein, while the other, the passenger strand, gets degraded. siRNAs loaded into argonautes can mediate both, PTGS and CTGS, leading to the degradation of the target mRNA in the former, while changing the chromatin state of the target gene to heterochromatin in the latter (Czech et al., 2008; Kanno & Habu, 2011). In some organisms, PTGS mediated by siRNAs can result in the synthesis of a second batch of siRNAs. These siRNAs are referred to as secondary siRNAs, while the initial siRNAs produced in the pathway are called primary siRNAs. Secondary siRNAs are synthesized using the targeted mRNA as a template, meaning that secondary siRNAs of a gene can be detected outside of the sequence region of the initial

dsRNA that triggered primary siRNA production (Pak & Fire, 2007; Sanan-Mishra et al., 2021). Synthesis of secondary siRNAs shows dependency on RNA dependent RNA polymerases (RdRPs) with slight differences across organisms. For example, in *C. elegans*, RdRPs synthesize mature siRNAs, while in plants, RdRPs convert parts of the target mRNA into a double stranded RNA that can then be cleaved by a Dicer enzyme into secondary siRNAs (Pak & Fire, 2007; Sanan-Mishra et al., 2021).

A second class of smallRNAs represent microRNAs (miRNAs). These miRNAs are not cleaved out of ordinary dsRNAs, but have intricate hairpin RNAs as a precursor. Those hairpins are encoded in miRNA genes, which are transcribed by RNA polymerase II, producing a primary miRNA. The primary miRNA can fold back on itself, producing the hairpin structure needed for miRNA biogenesis (Kim, 2005). Still in the nucleus, the Drosha enzyme cuts at the base of the stem loop, converting the primary miRNA into a pre-miRNA (precursor miRNA) (Y. Lee et al., 2003). In animals, the pre-miRNA gets exported out of the nucleus via the Exportin-5 channel and cleaved by Dicer to produce the miRNA duplex, while this step still takes place in the nucleus in plants (Hutvágner et al., 2001; Yi et al., 2005). Similar to the siRNA biogenesis, the passenger strand of the miRNA duplex gets degraded and the guide strand gets loaded into the argonaute protein. miRNAs are usually only involved in PTGS, however, unlike siRNAs, they don't need full homology to their target mRNA, but only within the seed region of the miRNA (from nucleotide 2 to 7), with a specific number of mismatches outside of this region, making target prediction of miRNAs very difficult. Additionally, miRNAs usually don't degrade their target, but prevent translation of the mRNA by blocking it.

A third class of smallRNA in RNAi is represented by the piwi-interacting RNAs (piRNAs) These piRNAs represent a special class of smallRNAs, since they don't perform processes classically described as PTGS or CTGS, but are involved in silencing of transposable elements in the germline. Biogenesis of piRNAs is best described in the fly *Drosophila melanogaster*, so involved enzymes in piRNA synthesis will be called according to their given name in *drosophila*. Similar to miRNAs, the piRNA precursors are encoded in the genome, enriched in so called piRNA cluster. Here, the precursor piRNA gets transcribed and exported into the cytoplasm (Huang et al., 2017). The single stranded precursor piRNA gets cleaved by the endonuclease Zucchini into the mature piRNAs, which get loaded into the argonaute Aubergine and trimmed at the 3' end to their final form. Guided by the piRNA, Aubergine searches for transcripts of transposable elements in a homology dependent manner and cuts them by using its slicer activity. This cut represents the 5' end of a new piRNA, which gets matured by

Zucchini and loaded into the argonaute Piwi. With this new piRNA loaded, Piwi can attack transcripts of the piRNA precursor cluster, allowing Zucchini to mature the Piwi cleaved piRNA precursor into new piRNAs, which are in turn loaded into Aubergine (Huang et al., 2017). Here, this cycle repeats, leading to the synthesis of new piRNAs while simultaneously silencing transposable elements by cleaving their transcripts. Since piRNA biogenesis bounces between piRNA precursor transcripts and transcripts of transposable elements, this cycle was called the ping-pong cycle. Since piRNA cleavage is not dependent on Dicer enzymes (Vagin et al., 2006), but dependent on two argonaute proteins, two piRNAs that are synthesized by each other show an overlap of only 10nt and much larger 3' overhangs, compared to the 2nt overhangs that are produced by Dicer (Huang et al., 2017; Vagin et al., 2006).

1.1.3 Components of the RNAi Pathway – A Brief Overview

As displayed in the biogenesis of the different smallRNA classes, many enzymes are involved in RNAi pathways. The following chapter will describe the different enzyme types present and give a brief overview of their function.

Dicer

The Dicer enzyme is an endonuclease responsible for the maturation of siRNAs and miRNAs out of their precursor molecules. The main domains characteristic for Dicer enzymes are the PAZ domain, the tandem RNaseIII domain, and a dsRNA binding domain (MacRae et al., 2006; Vergani-Junior et al., 2021).

The PAZ domain mediates the directional cut of Dicer by enabling Dicer to recognize the 5' and 3' end of the RNA substrate and bind to it. Additionally, the architecture of the PAZ domain determines the length of smallRNAs generated by the Dicer enzyme, taking over the function of a molecular ruler to guide cleavage of the precursor molecule (MacRae et al., 2006, 2007; Park et al., 2011). The RNaseIII domain, capable of RNA cleavage as the name suggests, then cuts the precursor molecule, giving rise to the smallRNA.

Argonaute

Argonaute proteins are RNA-binding proteins that load smallRNAs and mediate their function by searching for the RNAi target in a homology dependent manner. They bind to the target and either block or degrade the target transcript, if operating in a PTGS mode, or they recruit effector proteins that establish heterochromatin formation, if operating in a CTGS mode (Bühler et al., 2006; Hammond et al., 2001; Martienssen & Moazed, 2015).

Argonautes can be divided in two phylogenetic subfamilies, Piwis and Agos (Carmell et al., 2002). Piwis are argonautes that display sequence similarity to the argonaute Piwi, an argonaute originally described in the germline of *Drosophila melanogaster*, which this subfamily is named after (Carmell et al., 2002; Saito et al., 2006). Here, the Piwi interacts with piRNAs and is responsible for the defense of the germline genome against transposable elements (Saito et al., 2006; Vagin et al., 2006), which is why argonautes of the Piwi subfamily are usually associated with transposon silencing in germline tissue (Juliano et al., 2011). Agos, on the other hand, are argonaute proteins sharing sequence similarity with the argonaute Ago1 described in *Arabidopsis thaliana* (Carmell et al., 2002). This argonaute subclass is involved in several different RNAi pathways and therefore associated with miRNAs and siRNAs.

Despite differences in their perceived function, the proteins of the Ago and Piwi subclasses share structural elements that are considered the hallmark domains of argonaute proteins. These include the PAZ domain, the MID domain, and the PIWI domain (Iwakawa & Tomari, 2022). Both, PAZ and MID domain are responsible for the loading of the single stranded smallRNA that guides the argonaute to its target. The PAZ domain forms a binding pocket that loads the 3’ end of the produced smallRNA (Lingel et al., 2004), while the MID domain represents a binding pocket of the 5’ end (Ma et al., 2005; Parker et al., 2004, 2005). On the other hand, the PIWI domain, forming an RNase H structure containing the catalytic tetrad of the amino acids DEDH (Nakanishi et al., 2013; J.-J. Song et al., 2004), facilitates the degradation of the target mRNA that is annealed to the guide smallRNA.

RNA dependent RNA Polymerase

RNA dependent RNA Polymerases (RdRPs) are, as the name suggests, RNA polymerases that use RNA as a template. While mostly studied in the context of retroviral replication, RdRPs are involved in several different RNAi pathways, usually responsible for dsRNA production as a substrate for the Dicer enzyme. This is the case, for example, in *Arabidopsis. thaliana*, where endogenous RNAs are converted into a dsRNA by the cellular RdRP, which is then cleaved into phased siRNAs (Feng et al., 2024). However, RdRPs are not only responsible for converting cellular RNAs into dsRNAs, but also for transitivity. As described earlier, transitivity is a process were mRNA attacked by primary siRNAs gets used as a template for the production of secondary siRNAs, that can extend outside of the region used for primary siRNA production, for example, in *C. elegans*. RdRPs are necessary for this phenomenon by converting the attacked mRNA into a dsRNA that can be used as Dicer substrate for secondary siRNA production, or synthesize secondary siRNAs directly from the dsRNA (Aoki et al., 2007; Moissiard et al., 2007).

smallRNA Modifying Enzymes

After synthesis of the smallRNA, RNAs are not only used as created, but are often modified to various degrees. These modifications include addition and removal of nucleotides and other chemical modifications placed by different enzymes.

The most studied chemical modification of smallRNAs is a 2'-O-methylation at the 3'-terminal nucleotide. This modification is often found in piRNAs but also in other smallRNA species, where it contributes to the stabilization of smallRNAs. This modification is placed by the Hen1 enzyme, an RNA methyltransferase. A lack of Hen1 thereby leads to a reduction of piRNAs, which in turn results in a decrease of genome stability and defects during development in many different eukaryotes, including *C. elegans* or zebra fish (Billi et al., 2012; Kamminga et al., 2010; T. A. Montgomery et al., 2012).

Besides its stabilizing function, 2'-O-methylation often competes with another modification of smallRNAs, the addition of nucleotides (Kamminga et al., 2010; Li et al., 2005). Here, one or multiple nucleotides are added to the 3' end of a smallRNA, extending its sequence. Since the added nucleotides are not present at the, for example, genomic locus the smallRNA originates from, they are referred to as untemplated nucleotides. The addition of untemplated nucleotides can be catalyzed by terminal nucleotidyltransferases, however different classes of terminal

nucleotidyltransferases have been described, like HESO1 or URT1 in *Arabidopsis. thaliana*, MUT68 in *Chlamydomonas reinhardtii*, or PUP-1 of the CID1-family in *C. elegans.* Most of the untemplated nucleotide modifications, usually poly-uridylation or poly-adenylation, have in common that their addition leads to the degradation of the smallRNA, making untemplation a destabilization factor (Ibrahim et al., 2010; Kelley et al., 2024; X. Wang et al., 2015; Zhao et al., 2012).

The last mayor modification of smallRNA molecules is the removal of nucleotides, called trimming. Trimming is caused by several different exonucleases, however in this case, the exonucleases in question do not digest the smallRNA in its entirety, but only remove a couple of bases from the 3' end of the RNA. Trimming of smallRNAs plays an important role in several different pathways, including the piRNA pathway as well as miRNA maturation. Different enzyme families are known to be involved in smallRNA trimming, including enzymes of the PARN family, well described in *C. elegans*, as well as the 3'-5' exonuclease Nibbler described in *Drosophila melanogaster* (B. W. Han et al., 2011; Pastore et al., 2021).

dsRNA Uptake Channels

For exogenous dsRNA to trigger RNAi, the dsRNA has to somehow enter the cytoplasm to be accessible to the RNAi components that are responsible for the conversion of dsRNA into a functional smallRNA. The two most popular channels that have been described to be involved in dsRNA transport are the two SID channels, SID1 and SID2.

SID1 was first described in *C. elegans* and has been recognized as the channel responsible for the transport of dsRNA between cells of the organism. This transport can be realized not only in the close vicinity of the cell that initially took up dsRNA, but across the entire organism, a process called systemic spreading of silencing, or systemic RNAi in short (Feinberg & Hunter, 2003; Shih et al., 2009; Shih & Hunter, 2011; Winston et al., 2002). This systemic spreading is thought to be a mechanism that allows multicellular organisms to respond to a RNAi trigger not only on a local level, but throughout the entire organism. This leads to an RNAi response in cells that did not encounter the dsRNA, providing them with an RNAi mediated immunity to the silencing trigger, for example a dsRNA virus. Systemic RNAi in plants can be realized by the plasmodesmata and the phloem transport, however systemic RNAi in animals requires the SID1 channel (Baulcombe, 2004; Hunter et al., 2006).

While SID1 only showed involvement in systemic RNAi, SID2 was localized in the intestine walls in *C. elegans*. Further studies of SID2 revealed that SID2 is the channel responsible for the uptake of dsRNA from the intestine lumen into surrounding cells, making dsRNA from the food available for the cellular RNAi machinery and enabling *C. elegans* to use exogenous dsRNA as an RNAi trigger (McEwan et al., 2012; Winston et al., 2007).

Both of the SID proteins are well characterized and widely conserved across different organisms, making them the canonical pathway of exogenous dsRNA uptake. For example, a homolog of SID1 can be found in the human genome, but no homolog of SID2 is present there, which is why humans are generally regarded as capable to transport smallRNAs between cells (systemic RNAi) but not capable of RNAi by Feeding or RNAi triggered by exogenous RNA. However, not all organisms capable of exogenous RNAi harbor a SID2 homolog. For example, the single cell organism *Paramecium tetraurelia* displays capability of RNAi by Feeding, but no SID2 homolog can be found in its genome (Galvani & Sperling, 2002). This implies that SID2-independent dsRNA uptake pathways exist. Indeed, for example a receptor-mediated endocytosis-based dsRNA uptake has been described in *Drosophila melanogaster* S2 cells (Saleh et al., 2006). Even though it was unclear for a long time, how the dsRNA escapes the endosomal vacuole, recent studies have shown that dsRNA induces a permeabilization of the endosomal membrane involving the Cl^-/H^+ antiporter Ostm1, which leads to an escape of the dsRNA into the cytoplasm, making it accessible to the RNAi machinery (Tanaka et al., 2024). This means that beside the canonical dsRNA uptake mechanism directed by the SID proteins, additional SID-independent dsRNA uptake mechanisms exist.

1.2 Ciliates as a Model Organism

Looking at the organism group labeled as “Ciliates”, one can see that this group is composed of approx. 8000 different species (Grattepanche et al., 2018). Even though unified by the ciliate label, those species are not closely related to each other, only sharing two distinct features regarded as the hallmark properties of ciliates. The first feature is the presence of cilia on the surface of the cellular cortex, which are used for movement and feeding. Second, they display a nuclear dimorphism, the presence of two different kinds of nuclei within the cell, one macronucleus, which is transcriptionally active and responsible for gene expression, and different amounts of micronuclei, transcriptionally silent during vegetative growth, but responsible for the genetic make-up of the next sexual generation (Grattepanche et al., 2018). This means that ciliates represent an interesting case, where a single-cell organism separates

germline and somatic nuclei from each other, similar to what can be observed in multicellular organisms. This leads to the fact that ciliates can usually undergo two different life cycle stages, the vegetative life cycle, and the sexual cycle, where the micronucleus undergoes meiosis and gives rise to the macronuclei of the next sexual generation.

Several different members of the ciliate group have been studied extensively and can even be regarded as well-established model organisms, including the two most famous members of the ciliate group, *Paramecium tetraurelia* and *Tetrahymena thermophila*.

Using ciliates as model organisms, many very important observations have been made in the past, two of which have even been credited with the noble prize. The first noble prize ciliate research contributed to was awarded in 1989 to Sidney Altman and Thomas Cech for the discovery of self-splicing introns and the concept of ribozymes, RNAs that can catalyze reactions just like enzymes, in *Escherichia coli* and *Tetrahymena thermophila*, respectively (Altman, 1990; Cech, 2004; Guerrier-Takada et al., 1983; Zaug & Cech, 1986). The second noble price awarded to ciliate research was received by Elizabeth Blackburn, Carol Greider and Jack Szostak, again for work performed with *Tetrahymena thermophila*, awarded for the discovery of the telomeres and the telomerase enzyme (Blackburn & Gall, 1978; Greider & Blackburn, 1985; Szostak & Blackburn, 1982).

Besides these groundbreaking contributions to our understanding of cellular processes, ciliates have a variety of different interesting features that are still extensively studied.

With the research field of epigenetic, the study of chromatin and information present in our genome beyond the DNA base composition, becoming a more important topic, a lot of research concerning chromatin was conducted fairly early in *Tetrahymena thermophila*, leading to the identification of the first histone acetyltransferase (Allis et al., 1985; Brownell et al., 1996; Vettese-Dadey et al., 1996).

Closely related to epigenetic is the topic of non-mendelian inheritance, a phenomenon that has been described in several different unicellular eukaryotes, including the ciliate *Paramecium tetraurelia*. Pioneer in this field of work was Tracy Sonneborn, who noticed that the inheritance of serotype expression, proteins present on the cell surface of *paramecium*, as well as the inheritance of the mating type of a cell, did not follow a mendelian distribution among offspring after conjugation, the sexual reproduction of two *paramecium* cells. Instead, inheritance of both traits followed the maternal cytoplasmatic line (Sonneborn, 1939; Sonneborn & Lesuer, 1948). He figured that some traits have to be inherited by a cytoplasmatic component present within

the cells, a concept he coined the word “plasmagene” for (Sonneborn, 1949). Little did he know that he described a phenomenon of epigenetic inheritance, that was later revealed to be directed by small regulatory RNAs (Nowacki & Landweber, 2009; Simon & Plattner, 2014). These descriptions of non-mendelian inheritance coupled with the nuclear dimorphism of *paramecium* and the ability to study informational flow between the soma and the germline made *paramecium* a valuable model organism for epigenetic studies.

1.2.1 RNAi in *Paramecium*

The more epigenetic mechanisms were uncovered and understood in *Paramecium tetraurelia*, the more pathways were revealed. Today, three different major RNAi pathways can be distinguished. (I) Endogenous RNAi is used by the cell to either regulate gene expression or maintain genome integrity. Different processes are in play, depending on the life cycle stage the cell is currently in, the vegetative cycle or the sexual development cycle. Besides endogenous RNAi, two additional pathways triggering RNAi can be separated. (II) The transgene induced RNAi is triggered by an aberrant transgene introduced into the cellular macronucleus while (III) the exogenous RNAi pathway can be triggered by a dsRNA present in the food bacteria.

These pathways seem to be separated from each other; however, they share key RNAi components, leading to a competition of the pathways if activated at the same time. *Paramecium tetraurelia* possesses 15 different Piwi-like argonaute proteins (Ptiwis), 8 different Dicer proteins, 3 of which are actual Dicer proteins and 5 are Dicer-like proteins, 22 different Cid terminal nucleotidyltransferases, and 4 different RNA dependent RNA polymerases (RdRPs) (Bouhouche et al., 2011; Hoehener et al., 2018; Marker et al., 2010, 2014; Sandoval et al., 2014). This high degree of different RNAi components is a result of three consecutive whole genome duplications that gave rise to the *Paramecium aurelia* complex, which is why many of the components are ohnologs, paralogs caused by whole genome duplications, of each other (Aury et al., 2006; Gout et al., 2023). Some of these ohnologs operate in the same pathway, like Ptiwi12 and Ptiwi15, while others are members of different pathways, like Ptiwi14 and Ptiwi08. On the other hand, Dicer1 does not have any ohnologs in *paramecium*, but is part of all three mentioned pathways.

1.2.1.1 RNAi During Development

RNAi processes during the sexual lifecycle of *paramecium* are one of the most extensively studied fields in *paramecium* research, even up to this date. Due to the mentioned nuclear dimorphism of ciliates, *paramecium* degrades its old macronucleus during sexual reproduction and the micronuclei go through meiosis, selecting two of the meiotic products to fuse and form the anlagen, macronuclei of the next generation. After amplification to a polyploid state, the new macronuclei of the next generation are formed and distributed to the new cells at the first cell division.

What sounds like a relatively ordinary sexual reproduction is in reality an intricate system due to the fact that many sequences present in the micronuclei, remnants of old transposons and internal eliminated sequences (IES), have to be removed upon formation of a new macronucleus. These IES are localized in coding and non-coding regions of the micronuclear genome, making their removal vital for formation of proper open reading frames. This removal is performed in two phases guided by RNAi mechanisms and using the old, parental macronucleus as a template.

During the first phase, the entire micronuclear genome is bi-directionally transcribed, leading to the formation of dsRNA (Gruchota et al., 2017). This dsRNA is cleaved by the two Dicer-like proteins Dcl2 and Dcl3, producing 25nt long scnRNAs (Hoehener et al., 2018; Sandoval et al., 2014). These scnRNAs are loaded by the Ptiwi proteins Ptiwi01 and Ptiwi09 and shuttle into the parental, degrading macronucleus (Bouhouche et al., 2011; Furrer et al., 2017). In the macronucleus, nascent transcripts of the entire genome are produced, which allow Ptiwi proteins to bind to these transcripts, if the scnRNAs loaded fit to sequences present in the parental MAC (Lepère et al., 2008, 2009; Mochizuki et al., 2002). Here, scnRNAs bound to the nascent transcripts degrade, while scnRNAs without a target belonging to sequences that derive from IES and that need to be removed from the macronucleus of the next generation, persist. These persisting scnRNAs loaded by the Ptiwi proteins shuttle into the developing anlagen of the next generation and identify IES that need to be removed (Lepère et al., 2008, 2009; Mochizuki et al., 2002). This identification is performed in a homology dependent manner; however, it is still unclear whether the scnRNAs interact with the DNA directly, or whether nascent transcripts are produced by the anlagen to facilitate the identification of IES. In any case, the IES gets marked by the RNAi machinery, involving the formation of heterochromatin at the IES locus (Miró-Pina et al., 2022; Singh et al., 2022). This leads to the precise excision of the IES by the introduction of double strand breaks by the PiggyMac protein, a domesticated

piggyBac transposase, and associated enzymes (Baudry et al., 2009; Bischerour et al., 2018). The *paramecium* DNA reparation machinery repairs the double strand break by non-homologous end joining, restoring genome integrity (Abello et al., 2020). The excised IES DNA does not get degraded immediately, but is used in phase two.

During phase two, the excised IES DNA fragments get concatenated and circulated, followed by transcription of the IES fragments (Allen et al., 2017). The resulting non-coding RNA gets cleaved by Dcl5, leading to the production of 25-30nt long iesRNAs (Hoehener et al., 2018; Sandoval et al., 2014). These iesRNAs are loaded by Ptiwi10 and Ptiwi11 and further promote excision of IES within the newly developing macronucleus, probably functioning as an amplification mechanism that is supposed to guarantee complete excision of all IES (Furrer et al., 2017).

1.2.1.2 RNAi During Vegetative Growth

In contrast to RNAi pathways during development, RNAi during the vegetative lifecycle is not as well described. Some aspects of the endogenous smallRNA pathway were known for some time. For instance, a particular endogenous cluster on scaffold 22, called Cluster22, that produces massive amounts of smallRNA but is placed outside of any known annotated gene, has been known and was used as a Northern Blot control for many years (Lepère et al., 2009; Marker et al., 2014). However, mechanistical insights into siRNAs produced during the vegetative lifecycle were only published in 2019 (Karunanithi, Oruganti, et al., 2019). Here, it was revealed that siRNAs are produced by many clusters across the genome, labeled siRNA producing clusters (SRCs). These SRCs often overlap with genes, but are not strictly antisense-orientated and show both, positive and negative correlation with expression level of the associated genes, suggesting different functions of the produced siRNAs (Karunanithi, Oruganti, et al., 2019). It has been shown that SRC produced siRNAs are dependent on the two RdRPs RdRP1 and RdRP2, as well as Dicer1 (Karunanithi, Oruganti, et al., 2019; Karunanithi et al., 2020). Additionally, some SRC produced siRNAs are loaded into the argonaute proteins Ptiwi13 and Ptiwi14, showing an involvement of both Ptiwi proteins in the vegetative RNAi pathway (Drews et al., 2021).

1.2.1.3 Transgene Induced RNAi

Also taking place during the vegetative lifecycle, but only triggered by transgenes that are truncated and introduced into the *paramecium* macronucleus at high copy number, the transgene induced RNAi pathway has been studied for some time. In this pathway, a transgene containing the sequence of an endogenous *paramecium* gene, but missing the 3'end of the gene, can trigger the phenotypical silencing of the target gene, suggesting the silencing of the endogenous versions of the gene *in trans* (Galvani & Sperling, 2001; Garnier et al., 2004).

Interestingly, the observed silencing seems to be dependent on multiple sets of RNAi components, including RdRP2 and RdRP3, Cid2, Dicer1, and Ptiwi13 and Ptiwi14 (Götz et al., 2016; Marker et al., 2010).

To date, it is presumed that the aberrant transcript of the truncated transgene activates an RNA surveillance mechanism that triggers antisense transcription of the truncated transgene and conversion of the single stranded RNA into a double strand, all dependent on RdRP2, RdRP3 and Cid2. This double strand is then cleaved by Dicer1, producing primary siRNAs, which are loaded into Ptiwi13 and Ptiwi14. These siRNAs can then attack mRNA transcripts, leading to a PTGS of the target gene and production of secondary siRNA across the entire gene body. Additionally, primary and secondary siRNAs establish a heterochromatic state at the endogenous gene loci, preventing the formation of new transcripts and establishing a CTGS (Drews et al., 2021; Götz et al., 2016).

1.2.1.4 The RNAi by Feeding Pathway: What We Know So Far

RNAi by Feeding is the only known RNAi pathway in *paramecium* that can be triggered by application of exogenous dsRNA. First described in 2002, this RNAi pathway leads to the phenotypical silencing of genes if a dsRNA that shares sequence homology with a target gene is applied to the cells, similar to the pathway described in *C. elegans* (Galvani & Sperling, 2002).

Tightly bound to this pathway is the food uptake of *paramecium*. *Paramecium* is a bacterivore organism, meaning that it uses bacteria as a food resource. It is a filter feeder, that uses its cilia to create vortexes in the culture media to transport particles and bacteria to the vestibulum, where the food particles get concentrated. Here, rows of specialized cilia move the food particle into the nascent digestive vacuole. After growing to a certain size, the digestive vacuole is released and starts to cycle through the cell on a determined circular path, the cyclosis. Right

after vacuole release, the digestive vacuole fuses with acidosomes to lower the vacuoles pH and initiate digestion. This acidification step is followed by the addition of lysosomes to the digestive vacuole, adding digestive enzymes as well as a rise in pH. Nutrients and other valuable resources get extracted from the vacuole, while indigestible material remains in the vacuole and gets excreted by the cell at the end of the cyclosis at the cytoproct (*Paramecium*, 2013). dsRNA present within the food vacuole during this digestive pathway leads to the activation of the RNAi by Feeding pathway. Usually, this is achieved by feeding *paramecium* with bacteria that are capable to produce dsRNA matching with a specific gene sequence, allowing the targeted silencing of genes (Galvani & Sperling, 2002).

Numerous different RNAi components of *paramecium* are thought to be involved in this pathway, including Ptiwi12, Ptiwi15 and Ptiwi13, RdRP1 and RdRP2, Dcr1, Cid1 and Cid2, as well as the two uncharacterized and *paramecium* specific proteins Pds1 and Pds2 (Bouhouche et al., 2011; Carradec, 2014; Carradec et al., 2015; Marker et al., 2010, 2014).

Harboring that many RNAi components, overlap with other already mentioned RNAi pathways cannot be omitted. Indeed, it has been shown that application of dsRNA, even dsRNA targeting genes not involved in smallRNA biogenesis or transcriptional regulation, leads to a genome wide decrease of vegetative SRC produced siRNAs and mRNAs (Karunanithi et al., 2020). This suggests a huge overlap of the RNAi by Feeding machinery with components of the vegetative RNAi pathway, showing a substantial competition of the two pathways over key RNAi enzymes.

Besides the large amount of RNAi components known to be associated with RNAi by Feeding, little is known about their actual function. So far, it is only known that RdRP1 and RdRP2, Cid1 and Cid2, as well as Dicer1 are necessary for the production of primary siRNAs from the trigger dsRNA molecules (Carradec et al., 2015). Here, it is speculated that RdRP1 and Cid1 as well as RdRP2 and Cid2 form two different RdRP complexes that are then used in the production of primary siRNAs. Additionally, it is speculated that the complex around RdRP2 is responsible for the production of secondary siRNAs, using the mRNA attacked by primary siRNAs as template (Carradec et al., 2015).

There is no evidence which siRNAs are loaded by which of the involved Ptiwi proteins. However, it is hypothesized that Ptiwi13 might load primary siRNAs and target mRNA transcripts, being mainly responsible for the PTGS phenotype, while Ptiwi12 and Ptiwi15, might load secondary siRNA, performing so far unknown functions (Figure 1). This speculation

is caused by the catalytic tetrad of the RNase H structure within the Ptiwi domains of these argonautes. The catalytic tetrad is present in Ptiwi13, but altered in Ptiwi12 and Ptiwi15, which is why it is speculated that Ptiwi12 and Ptiwi15 don't harbor slicer activity and are therefore not capable of cleaving mRNA targeted by the RNAi by Feeding machinery (Bouhouche et al., 2011; Carradec et al., 2015).

Apart from the mentioned information, nor further progress has been made in the mechanistical analysis of the RNAi by Feeding pathway in recent years, which is why knowledge about the specific steps that are carried out during primary siRNA production and PTGS mediated by the Ptiwi proteins is lacking.

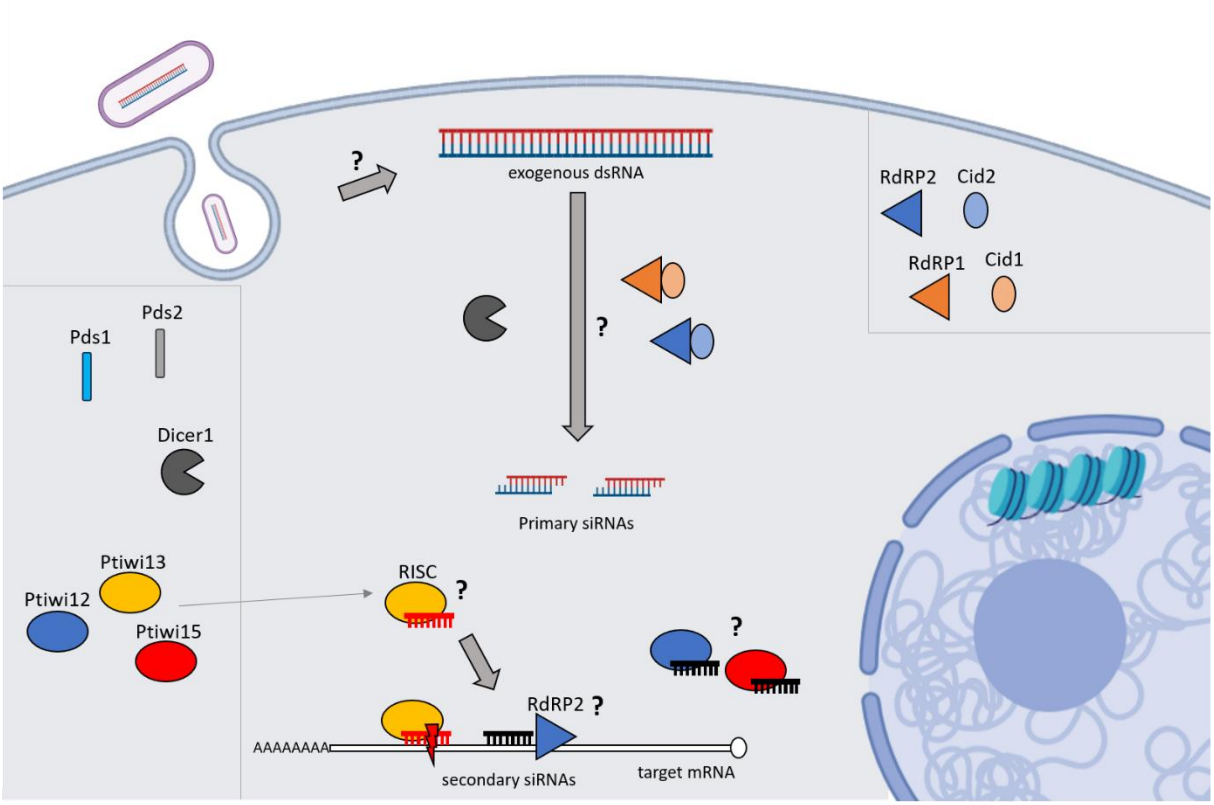

Figure 1 Schematic of our current knowledge of the RNAi by Feeding pathway in *Paramecium tetraurelia*. Displayed is the process of RNAi by Feeding according to current literature (Carradec et al., 2015; Marker et al., 2014). Steps that are not further characterized yet or are speculated, are labeled with a question mark. Enzymes involved in different steps are labeled accordingly and color coded to distinguish different enzymes of the same family. Omitted from the schematic are potential modifications of the smallRNAs, including nucleotide trimming and addition of untemplated nucleotides, to improve readability. A detailed description of the pathway can be taken from the text. Parts of the image were created using BioRender.

1.3 Scope of the Thesis

Aim of this thesis is the characterization of several different enzymes in the RNAi by Feeding pathway and their specific function at the individual steps of the siRNA biogenesis.

Pds1 and Pds2 are two proteins so far not characterized at all. By performing cellular localization and bioinformatical analysis of the amino acid sequence of both proteins, chapter 3 of this thesis will try to gain information about the two elusive proteins.

Previous work revealed that the two RdRPs, RdRP1 and RdRP2, are involved in the biogenesis of primary siRNAs, which raised the question of why two RdRPs are needed at this step of the RNAi pathway, if a trigger molecule already suitable for Dicer cleavage into siRNAs is supplied to the cell. By utilizing a molecular reporter in form of a specifically designed Heteroduplex dsRNA and applying this reporter to wildtype and RdRP-deficient cells, next generation sequencing data of produced siRNAs is analyzed to retrace the impact of the two RdRPs during primary siRNA production. Chapter 4 will discuss this RdRP impact on siRNA production and tries to formulate a hypothesis of the specific roles RdRPs are playing in this step.

In chapter 5, the role of the three Ptiwi proteins, Ptiwi12, Ptiwi13 and Ptiwi15 within the RNAi by Feeding pathway will be analyzed. Since the RNAi by Feeding mechanism and endogenous RNAi seem to display great overlap in terms of utilizing the same enzymes, Ptiwi14 will be included in this analysis as a Ptiwi so far not brought into contact with the feeding pathway, but with endogenous RNAi. Cellular localization of Ptiwi12 and Ptiwi15 as well as analysis of smallRNAs loaded by the four Ptiwis in wildtype situation and during application of dsRNA by RNA-immunoprecipitation will provide insight on whether the four Ptiwi proteins interact with specific subclasses of siRNAs, or whether Ptiwi proteins display a great overlap in their association with smallRNA species. Besides analyzing feeding-associated siRNAs, endogenous smallRNAs produced from SRCs will also be analyzed to shed light on the involvement of the four Ptiwi proteins in the endogenous RNAi processes.

Last but not least, chapter 6 will deal with siRNA modifications after their biogenesis by Dicer1. Here, feeding-associated siRNAs are analyzed in regards to untemplated nucleotides to answer the question of involvement of the Cid protein Cid1 in the placement of untemplated nucleotides and the extend of untemplation that can be found within smallRNA data. Additionally, by utilizing cells that are deficient of Cid1, the role of Cid1 in the biogenesis of primary siRNAs and their untemplation will be discussed.

2 Material and Methods

2.1 Cultivation of Organisms

2.1.1 Cultivation and Handling of *Paramecium tetraurelia*

Paramecium tetraurelia cells were cultivated in wheat grass powder medium (WGP) at 16°C to 31°C depending on experimental needs. The day before usage, WGP was infused with bacteria (*Klebsiella pneumoniae*) and incubated at 31°C, enabling bacterial growth and allowing the bacteria to be used as a food source for *paramecium*. Depending on the specific culture conditions necessary for the experiment, β-sitosterol (1:5000) was added to promote cell division rates of *paramecium*.

Cells were grown in different glass ware, including depression slides and flasks of various sizes, to accommodate for culture volumes.

Klebsiella pneumoniae were grown on agar slants in 15ml tubes overnight at 31°C and stored at 4°C until usage.

The different cell lines used in this thesis are listed below (Table 1).

WGP Stock Buffer (20x)	
Tris	1.6g
Na_2HPO_4	15g
NaH_2PO_4	3.46g

Substances were dissolved in 1l water and pH was adjusted with HCl to pH 7. The solution was sterilized using an autoclave.

WGP Stock Solution (20x)	
Wheat Grass Powder (Pines International Co.)	66.6g

Powder was dissolved in 1l water and boiled for 20min, filtered through gaze and centrifuged with a cream separator (Wesfalia Separator AG). Particle-free solution was sterilized in an autoclave.

WGP Medium	
WGP Stock Buffer (20x)	50ml
WGP Stock Solution (20x)	50ml
Both components were mixed together and water was added to a final volume of 1l. The solution was sterilized using an autoclave.	

Agar Slants	
Nutrient Agar (Carl Roth)	3.7g
Agar was dissolved in 100ml water; aliquots were filled in 15ml falcon tubes and autoclaved. Tubes were cooled down lying on a table to create agar slants.	

β-sitosterol Solution (5000x)	
β-sitosterol	4mg
Dissolved in 1ml 100% ethanol.	

Table 1 Overview of used *Paramecium tetraurelia* strains.

Name	Mutated Gene	Gene Accession Number	Type of Mutation
51-7	none	none	none
Rdrp1 3.1	*rdrp1*	PTET.51.1.G0850056	R1361stop
Rdrp1 5.28	*rdrp1*	PTET.51.1.G0850056	del(374)/stop
Rdrp2 3.7	*rdrp2*	PTET.51.1.G1640013	E859K
Rdrp2 1.24	*rdrp2*	PTET.51.1.G1640013	1276(IES)/1330
Cid1 1.8	*cid1*	PTET.51.1.G910004 PTET.51.1.G910005	S174F

2.1.1.1 Daily Isolation Line and Aging of Cultures

To ensure working with *Paramecium tetraurelia* cells in their vegetative states, performing experiments with young cell cultures was crucial. Therefore, aging of cell cultures via daily isolation lines (Beisson et al., 2010 adapted from Sonneborn, 1938) was performed prior to experiments.

Under the binocular, a single *paramecium* cell was transferred to 250µl bacterized WGP in a depression slide and incubated at 31°C for 24h in a humid chamber to prevent evaporation. After 24h, a single cell was selected from the cells present in the depression slide and transferred to 250µl fresh WGP. This process was repeated 5 times to provide the cells with enough nutrition to maximize the cell division rate. After approx. 20 divisions, cells are capable to undergo autogamy, inducing their sexual reproduction cycle to produce a young generation of vegetative cells. Autogamy was induced by transferring 4-5 cells to 500µl WGP and incubating them at 31°C for 3 days. The increased cell density as well as the depletion of food bacteria over this time period triggers synchronous autogamy in all cells in a depression, leading to an arrest of the cells in the autogamous state.

Successfully induced autogamy was checked by nuclear staining, observing the fragmentation of the macronuclei as a sign for autogamy (2.1.1.2).

To start a young, vegetative *paramecium* culture, autogamous cells were transferred into fresh WGP and incubated overnight at 31°C. Replenishing the food source by adding fresh bacteria breaks autogamy and leads to the reformation of the macronucleus of a new vegetative generation of cells. After confirmation of new macronuclei formation by nuclei staining (2.1.1.2), cells were ready to be used for different experiments.

2.1.1.2 Staining of Cell Nuclei

Staining of the nuclei was achieved by incubating cells in a drop of 10µl medium on a slide mixed with 2µl EDTA (0.5M) and 1µl DAPI stock solution for 10min in the dark and by observing DAPI fluorescence under a fluorescence microscope using UV light. As DAPI binds to dsDNA in AT-rich regions, it can be used to make DNA visible and therefore allow for the examination of the macronucleus and micronuclei of *paramecium*.

DAPI Stock Solution	
DAPI	1µg

Dissolved in 1ml water.

2.1.1.3 Trichocyst Discharge Phenotype Check

Trichocysts are protein-filled granules which are located underneath the *paramecium* surface. *Paramecia* are able to discharge those trichocysts, thereby releasing the protein inside, which unrolls into a crystalline, needle-like structure. Those structures function as a defense mechanism against predators (Knoll et al., 1991) and can be triggered by a variety of different stimuli, including contact with other cells or shifts in pH.

To monitor the cells' capability to discharge trichocysts, 10µl of a cell culture were mixed with 10µl of a saturated picric acid solution (Morphisto) on a glass slide, which leads to trichocyst discharge and the death of cells. Using a light microscope, cells were monitored and the number of cells capable of trichocyst discharge was evaluated.

Cells were divided into three different phenotypical classes depending on their capability to discharge trichocysts: Trich+ for cells showing a wildtype-like discharge, Trich+/- for cells with a reduced number of triggered trichocysts and Trich- for no visible discharged trichocysts.

Trichocyst discharge is an easily observable phenotype which is used in RNAi by Feeding experiments to show successful knock-down of genes responsible for the secretion of trichocysts, naming *nd169* as an example of such genes.

2.1.1.4 Generation of Transgenic *Paramecium* Lines

Due to *paramecium*'s unique biological features, introducing foreign DNA to produce transgenic cell lines is fairly difficult. Since more common methods, like e.g. lipofection, fail to deliver DNA into the *paramecium* macronucleus, microinjection of prepared DNA is one of the only ways to reliably transform *paramecium* cells. Genetic manipulation like this is required to study proteins (which are usually fused to a specific amino acid tag sequence like FLAG or HA-tags or to other proteins like GFP) in their native environment and allows for further manipulation of the proteins themselves. Different plasmids used for microinjection are displayed in the table below (Table 2).

For DNA preparation prior to microinjection, 100µg plasmid were digested in a total volume of 400µl using the AdhI (NEB) restriction enzyme (2.2.3.1). The linearized plasmid was extracted with one volume basic phenol and centrifuged (13.000rpm, 5min) to separate the phases. The upper, aqueous phase was transferred into a fresh reaction tube and precipitated with 2.5 volume 100% ethanol and 1/10 volume sodium acetate (3M, pH 9) for 30min at room temperature. DNA pellets were washed with 80% ethanol and resuspended in 380µl water. 20µl sodium acetate (3M, pH 9) was added and the solution was filtered using an UltraFree MC filter (Millipore) to achieve a particle-free solution. 1ml of 100% ethanol was added, and DNA was precipitated (13.000rpm, 30min) and washed once with 80% ethanol. DNA pellets were air-dried and stored at -20°C. Prior to microinjection, DNA pellets were dissolved in 2-5µl water.

For microinjection, young *paramecium* cells were washed twice in Volvic water supplemented with BSA (0.2% w/v) and, in a volume of 1µl, transferred onto a coverslip which is mounted on a glass slide with paraffin oil. Paraffin oil was used to cover the droplets containing single cells, and water was removed with a glass capillary to immobilize the cells. Prepared DNA was injected into the macronucleus of the cell using a micromanipulator (Eppendorf) under a light microscope, and injected cells were recovered in Volvic-BSA (0.2% w/v). After being washed twice in Volvic-BSA (0.2% w/v), individual cells were transferred to depression slides with 100µl fresh WGP and incubated at RT. Surviving cells were screened for successful transformation by Western blotting (2.4.2.2) in case of FLAG-HA-tagged protein constructs, or by using a fluorescence microscope and screening for green fluorescence in case of GFP-tagged proteins.

Table 2 Overview of plasmids used for generation of transgenic *Paramecium tetraurelia* cell lines.

Name	Expressed Gene	Gene Accession Number	Tag
P12_N_FLAG	*ptiwi12*	PTET.51.1.G0040095	N-term FLAG
P12_N_GFP	*ptiwi12*	PTET.51.1.G0040095	N-term GFP
P13_N_FLAG	*ptiwi13*	PTET.51.1.G0480035	N-term FLAG
P13_N_GFP	*ptiwi13*	PTET.51.1.G0480035	N-term GFP
P14_N_FLAG	*ptiwi14*	PTET.51.1.G1630015	N-term FLAG
P15_N_FLAG	*ptiwi15*	PTET.51.1.G0120328	N-term FLAG
P15_N_GFP	*ptiwi15*	PTET.51.1.G0120328	N-term FLAG
Pds1_C_GFP	*pds1*	PTET.51.1.G0060311	C-term GFP
Pds2_C_GFP	*pds2*	Not annotated	C-term GFP
RdRP1_N_GFP	*rdrp1*	PTET.51.1.G0850056	N-term GFP
RdRP1_N_FLAG	*rdrp1*	PTET.51.1.G0850056	N-term FLAG
RdRP2_N_GFP	*rdrp2*	PTET.51.1.G1640013	N-term GFP
RdRP2_N_FLAG	*rdrp2*	PTET.51.1.G1640013	N-term FLAG

2.1.2 Cultivation of *Escherichia Coli*

Escherichia coli (*E. coli*) was incubated at 31°C in LB medium under constant shaking (200rpm) or on LB agar plates. To prevent growth of contaminants, a constant selection pressure was applied by adding tetracycline (1:1000) to the growth media. Depending on the experiment, ampicillin (1:1000) was added as well to ensure maintenance of transformed plasmids.

For storage, LB plates containing *E. coli* were sealed with parafilm and kept at 4°C.

In this work, two different kinds of *E. coli* strains were used. Top10 strain (Invitrogen/Takara) was used for cloning procedures, and HT115 (DE3) (provided by the Fire Lab) was used for RNAi by Feeding applications. Strains used in different experiments are indicated in the appropriate chapters.

LB Medium	
Peptone	10g
Yeast Extract	5g
NaCl	10g

Substances were dissolved in 1l water and solution was sterilized in an autoclave.

LB Agar plates	
Agar Powder	1.5g
LB Medium	100ml

Substances were dissolved using a microwave, supplemented with appropriate antibiotics and poured into petri dishes.

Ampicillin Stock Solution	
Ampicillin	100mg

Substances were dissolved in 1ml water and sterilized by filtration (0.22µm pore-size).

Tetracycline Stock Solution	
Tetracycline	12.5mg

Substances were dissolved in 1ml 100% ethanol.

2.2 Cloning

While some constructs necessary for this work were already available, others had to be cloned first.

2.2.1 DNA Amplification Using PCR

PCR is a standard molecular biology tool used to amplify DNA from various DNA templates. Small oligonucleotides, called primer, are designed to flank the targeted DNA region. Those primer bind to the edges of the desired PCR product to allow for the DNA polymerase to start DNA synthesis and elongate the primer, using the nucleotides provided in the PCR mix and the DNA as a template. After polymerization of the strands, the PCR reaction is heated to denature the PCR products, followed by a cooling step to facilitate primer binding, again. This cycle is repeated to exponentially amplify the desired region of the DNA until enough PCR product for further work is produced.

In this work, Q5 DNA polymerase (NEB) was used. PCR components as well as the PCR program for a default PCR using Q5 Polymerase were used according to the manufacturer's instructions and bought by the same supplier (NEB). Primer annealing temperatures were calculated using the NEB Tm web calculator (https://tmcalculator.neb.com/#!/main).

For downstream applications, PCR fragments were purified using the NucleoSpin Gel and PCR Clean-up XS Kit (Macherey-Nagel) according to the manufacturer's instructions and quantified using the Nanodrop One spectrophotometer (Thermo Fisher Scientific).

2.2.2 Agarose Gel Electrophoresis

By checking the size of DNA fragments produced by, for example, PCR or DNA restriction reactions, one can predict whether a reaction was successful, and whether the DNA fragments produced are the ones expected and needed for downstream applications.

DNA size determination was carried out using agarose gel electrophoresis. Agarose gels were prepared in different concentrations, ranging from 0.8-1.2% (w/v) agarose, by dissolving agarose in a running buffer. The resulting gel was submerged in running buffer in an electrophoresis chamber. DNA samples were mixed with a loading dye (6x loading dye purple, NEB; supplemented with 10xGelRed, Merck) and loaded into the pockets of the gel. By applying a voltage of 80-120V, DNA is forced through the polymer gel, separating different DNA fragments according to their size.

Visualization of DNA bands within the gel was facilitated by UV light on a gel documentation station by exciting the dye GelRed, which is bound to the migrated DNA.

DNA fragment sizes were estimated by comparing their migration length with a ladder, containing DNA fragments of known sizes. In this thesis, the Generuler 1kb ladder from Thermo Fisher Scientific was used.

For agarose gel electrophoresis of DNA fragments, 1xTAE buffer was used as a running buffer. Different buffers used for other substances beside DNA fragments are mentioned in the appropriate chapters

TAE Buffer Stock (50x)	
Tris	243g
Acetic Acid	57.1ml
EDTA	3.72g

Substances were dissolved in a total volume of 1l in water. For use, buffer was diluted 1:50 with water.

2.2.3 Ligation of DNA Fragments

2.2.3.1 Restriction Enzyme Digestion, DNA Purification and Blunt-End Ligation

Blunt-end ligation was used to join short PCR-products with pre-existing plasmids.

Pre-existing plasmids were linearized at the desired insertion region using a restriction enzyme digestion with the appropriate restriction enzyme (for ligation of RNAi by Feeding vectors, SmaI (NEB) was used. Other used enzymes are mentioned in the appropriate chapters). Digestion was achieved by assembling the default restriction enzyme reaction according to the manufacturer's instructions and by incubating the reaction at 37°C for 1 hour if not stated otherwise. Restriction enzymes were removed by phenol-chloroform extraction, and the linearized vector was precipitated by adding 1 vol of isopropanol and 1/10 vol of sodium acetate (3M, pH8) and incubation at -20°C for 1h. After centrifugation (13.000rpm, 30min at room temperature), the resulting pellet was washed twice in ethanol (80%) and resuspended in water. Linearized plasmid was quantified using the Nanodrop One (Thermo Fisher Scientific) instrument.

For ligation, T4 Ligase (NEB) was used, and the default ligation reaction was assembled according to the manufacturer's instructions. Blunt end ligation was carried out for 16h at 16°C. The ligation reaction was heat inactivated at 65°C for 10min and 1µl of the ligation reaction was used for transformation (2.2.4).

2.2.3.2 In-Fusion® Cloning

The In-Fusion© Snap Assembly Kit (TakaraBio) was used for cloning of more complex vectors, such as constructs for microinjection. Cloning was performed according to the manufacturer's instructions.

The plasmid backbone was linearized at the insertion region by PCR, using inverse-pointing primer (2.2.1), followed by digestion of the PCR reaction with DpnI (NEB) to digest the plasmid template (2.2.3.1).

DNA insert was prepared by PCR (2.2.1) using the genomic DNA from *paramecium* as a template. Primer for insert PCR were designed to ensure that PCR products carry 15nt overhangs that share sequence homology with the edges of the backbone PCR product.

After purification of the prepared PCR products, the In-Fusion cloning reaction was set up and incubated according to the manufacturer's instructions. 1µl of the reaction was then used for transformation (2.2.4).

2.2.4 Transformation of Competent Cells

In this thesis, both methods of transformation of competent *E.coli* cells, heat shock and electroporation, were used. Heat shock was used on cells of the Top10 strain whereas electroporation was used for HT115 (DE3) cells.

2.2.4.1 Preparation of Electrocompetent Cells

10ml of an overnight culture grown in LB medium supplemented with tetracycline were transferred in 1l LB medium and incubated (2.1.2) until an OD of 0.8 was reached. The culture was incubated on ice for 10min and cells were harvested in 50ml falcon tubes (4.300rpm, 10min at 4°C). Pellets were resuspended in glycerol (10%) and unified in four 50ml falcon tubes. Washing of the cells with 10% glycerol was repeated four times in total, unifying the cells in a single falcon tube during the process. After the last glycerol wash, the cell pellet was

resuspended in 3.5ml glycerol (10%) and cells were separated in 100µl aliquots. Aliquots were snap frozen in liquid nitrogen and stored at -80°C until needed for transformation.

2.2.4.2 Transformation by Electroporation

For electroporation, 100µl competent cells were thawed on ice and mixed with 1µl plasmid (10ng/µl). The mixture was transferred into a pre-cooled electroporation cuvette and a brief electric impulse of 1.8kV was applied using the *E. coli* Pulsor (BioRad). This impulse distorts the cell membrane and allows DNA to enter the bacteria.

500µl pre-warmed SOC medium was added to recover the cells and to transfer them to a 2ml reaction tube. The cells were incubated for 1h at 37°C under constant shaking (200rpm) to allow for expression of the plasmid-coded antibiotic resistance. Afterwards, 100µl of the cell suspension were plated onto LB plates supplemented with ampicillin and tetracycline (2.1.2)

SOC Medium	
Tryptone	20g
Yeast Extract	5g
NaCl	0.584g
KCl	0.18g
$MgCl_2$	0.95g
Glucose	3.6g

Substances (except $MgCl_2$ and glucose) were dissolved in a total volume of 1l in water and sterilized using an autoclave. $MgCl_2$ and glucose were added after autoclaving in a sterile manner.

2.2.4.3 Preparation of Heat shock Competent Cells

Heat shock competent cells were produced according to (C. T. Chung et al., 1989). 100ml LB medium was inoculated with 100µl overnight culture and incubated as described previously (2.1.2). After an OD of 0.4 - 0.7 was reached, cells were harvested (4.300rpm, 10min at 4°C) and resuspended in 1ml TSS buffer. Aliquots of 100µl volume were snap frozen and stored at -80°C until needed for transformation.

TSS Buffer	
PEG8000	1g
DMSO	0.5g
$MgCl_2$	4.76g

Substances were dissolved in a total volume of 100ml LB medium and sterilized using an autoclave.

2.2.4.4 Transformation by Heat Shock

For heat shock, 100µl of competent cells were thawed on ice and mixed with 1µl of plasmid DNA (2.2.3). Cells were incubated for 30min on ice and afterwards transferred to 42°C using a water bath to apply the heat shock, which triggers DNA uptake. After exactly 45s, cells were transferred to ice, chilled for 1min, and then 500µl pre-warmed SOC medium was applied for recovery. After SOC addition, cells were handled as described in section 2.2.4.2.

2.2.5 Plasmid Isolation

2.2.5.1 Isolation by Alkaline Lysis

To transfer plasmids from one *E. coli* strain to another or to check successful ligation of two DNA fragments, plasmid isolation from overnight cultures was performed.

First, overnight cultures of single *E. coli* colonies were inoculated by transferring a colony from a LB plate to 5ml of LB medium supplemented with ampicillin and tetracycline (2.1.2). Cultures were grown overnight at 37°C under constant shaking (200rpm).

2ml overnight culture were transferred to a 2ml reaction tube and harvested by centrifugation (13.000rpm, 5min). After discarding the supernatant, 340µl of resuspension buffer were added and the pellet was resuspended. 340µl of lysis buffer was added and carefully mixed by

inverting the tube. The presence of SDS and an increase of pH leads to the lysis of the *E. coli* cells, releasing the desired plasmid into solution. By adding 340µl of neutralization buffer, the pH quickly decreased, leading to the precipitation of unwanted molecules, like proteins and insoluble cell debris, together with larger DNA molecules, like genomic DNA, while keeping smaller molecules, like plasmids, in solution. After centrifugation (13.000rpm, 20min), 800µl of the supernatant was transferred to a fresh reaction tube and mixed with 1 volume of isopropanol for DNA precipitation (1h at -20°C). Precipitated plasmid DNA was centrifuged (13.000rpm, 30min), washed twice in 1ml of ethanol (80%) (13.000rpm, 5min) and resuspended in 100µl water after brief air-drying. Plasmid yield was determined with the Nanodrop One (Thermo Fisher Scientific).

After cloning, successful incorporation of the desired DNA fragments into the plasmid backbone was checked by agarose gel electrophoresis (2.2.2) and sanger sequencing, carried out by Macrogen, using appropriate sequencing primer depending on the used plasmid backbone.

Resuspension Buffer	
Tris	0.6g
EDTA	0.29g

Substances were dissolved in a total volume of 100ml in water.

Lysis Buffer	
NaOH	0.8g
SDS	1g

Substances were dissolved in a total volume of 100ml in water.

Neutralization Buffer	
Potassium Acetate	29.4g

Substances were dissolved in a total volume of 100ml in water and pH was adjusted to 5.5 with acetic acid.

2.2.5.2 Plasmid Isolation from Large Cultures

For some applications, large amounts of endotoxin free plasmids were required. In those cases, yield from plasmid isolation from overnight cultures was not enough.

To generate enough plasmid yield, the NucleoBond®xtra Midi EF Kit (Macherey-Nagel) was used according to the manufacturer's instructions.

2.3 RNAi by Feeding

RNAi by Feeding is a process, where an RNAi trigger (e.g. dsRNA) is introduced into an organism by adding the trigger to its food. This process was first described in 1998 by Timmons and Fire (Timmons & Fire, 1998), where the introduction of bacteria producing dsRNA in *C. elegans* targeting a specific reporter gene led to the silencing of this gene on a phenotypical level. Later, this protocol was adapted for *Paramecium tetraurelia* (Galvani & Sperling, 2002). In this thesis, RNAi by Feeding was triggered in two different ways, using *E. coli* producing dsRNA/ssRNA as well as using dsRNA packaged in nanoparticles, both of which will be explained in the following section.

The different plasmids used for RNAi induction are displayed in the table below (Table 3).

Table 3 Overview of plasmids used for RNAi induction.

Name	Target Gene	Gene Accession Number	Position within the Gene	RNAi Method
SHD_antisense	*nd169*	PTET.51.1.G0210080	1450-1860	Nanoparticles
SHD_sense	*nd169*	PTET.51.1.G0210080	1450-1860	Nanoparticles
HD	*nd169*	PTET.51.1.G0210080	1450-1860	Nanoparticles
L4440_P12	*ptiwi12*	PTET.51.1.G0040095	956-1404	*E. coli*
L4440_P13	*ptiwi13*	PTET.51.1.G0480035	105-787	*E. coli*
L4440_P14	*ptiwi14*	PTET.51.1.G1630015	285-1007	*E. coli*
L4440_P15	*ptiwi15*	PTET.51.1.G0120328	307-926	*E. coli*
L444T_HD_antisense	*nd169*	PTET.51.1.G0210080	1450-1860	*E. coli*
L444T_HD_sense	*nd169*	PTET.51.1.G0210080	1450-1860	*E. coli*
T444T_ND	*nd169*	PTET.51.1.G0210080	1486-1896	*E. coli*

2.3.1 RNAi by Feeding Mediated by RNA Producing *E. coli*

In this RNAi method, the RNAi trigger was applied by feeding of either dsRNA- or ssRNA-producing *E. coli* to the *Paramecium tetraurelia* cells. This is achieved by transforming (2.2.4) the strain HT115(DE3) with either the T444T or the L444T plasmid. The presence of the T7 promoter and T7 terminator sequences on those plasmids (Figure 2) grants the *E. coli* the ability to produce dsRNA or ssRNA, respectively.

First, overnight cultures of transformed *E. coli* cells were prepared (2.2.5.1) and used to inoculate the necessary volume of LB medium (containing ampicillin). Bacteria were cultivated as described (2.1.2) until an OD_{595} between 0.38 and 0.42 was reached. IPTG was added (1:400) to block the lac repressor and activate the lac operon in *E. coli*, which regulates the expression of the T7 polymerase in the HT115(DE3) strain. Inducing the culture like this allows the T7 polymerase to produce RNA from the provided plasmid by using the T7 promoter. Production of RNA was carried out for an additional 2.5h under cultivation conditions until induced bacteria were harvested by centrifugation (4.000rpm, 10min at 4°C). Bacteria pellets were

resuspended in ten times the harvested culture volume of WGP (supplemented with β-sitosterol, ampicillin and IPTG) and young *paramecia* were cultivated (2.1.1) in the feeding medium until the necessary cell number was reached, checking for autogamy during the cultivation period (2.1.1.2).

To ensure RNA production by *E. coli*, 2ml induced culture was harvested by centrifugation (8.000rpm, 5min) and the bacteria pellet was resuspended in 200µl 10mM Tris-HCl buffer (pH 8,4) and either stored at -20°C or processed immediately. Resuspended cells were incubated for 10min at 70°C in a water bath to ensure cell lysis and promptly mixed with 1 volume of phenol/chloroform/isoamylalcohol (Carl Roth). After vortexing and subsequent centrifugation (13.000rpm, 5min), the upper aqueous phase was transferred to a new reaction tube. 1 volume of isopropanol and 1/10 volume of sodium acetate (3M, pH8) was added and RNA was precipitated for a least 30min at -20°C. After additional centrifugation (13.000rpm, 30min), the RNA pellet was washed twice with 1ml ethanol (80%) and air dried for a few minutes. The RNA was resuspended in 50µl 10mM Tris-HCl (pH 8), quantified using the Nanodrop One (Thermo Fisher Scientific), and 1µg RNA was loaded onto a 1% TAE agarose gel (2.2.2) to verify the presence of dsRNA in the appropriate size range.

IPTG Stock Solution	
IPTG	50mg/ml
Dissolved in water.	

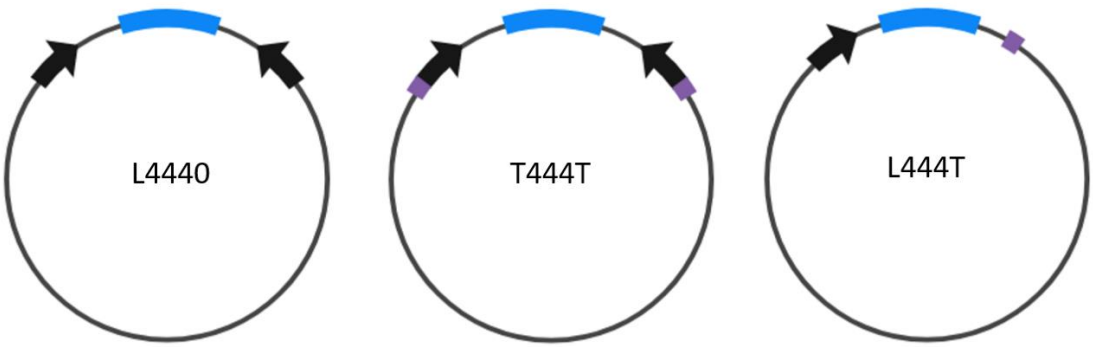

Figure 2 Schematic overview of different feeding plasmid backbones. Names of the different plasmid designs are displayed within each schematic. Black arrows represent T7 promoter sequences, purple squares represent T7 terminator sequences and blue squares represent the used feeding sequence. Figure created using BioRender.

2.3.2 RNAi by Feeding Induced by dsRNA Nanoparticles

For some of the molecular characterizations that were carried out in this thesis, the application of a specifically designed dsRNA, a so called Heteroduplex, was necessary. Since producing the Heteroduplex was not possible in the described *E. coli*-based RNA delivery system, a work-around using nanoparticles was found.

2.3.2.1 Production of dsRNA using *In Vitro* Transcription

For production of the Heteroduplex dsRNA, the HiScribe®T7 High Yield RNA Synthesis Kit (NEB) was used.

Templates for the *in vitro* transcription reaction for both strands of the Heteroduplex were produced by PCR (2.2.1) and mixed in a 1:1 molar ratio after purification. The *in vitro* reaction was assembled according to the manufacturer's instructions for standard RNA synthesis reactions including the recommended incubation times

After incubation, 70µl nuclease-free water was added and the reaction was heated at 85°C and slowly cooled down to 4°C (over a period of 45min, setting the cooling rate of the cycler to 0.1°C/s with 1.5min breaks after each 5°C step) to allow the two RNA strands to anneal and form a proper dsRNA. Following annealing, DNase I digestion and RNA purification using phenol/chloroform was performed as described by the manufacturer instructions apart from resuspending the produced RNA in water.

Yield of the synthesized Heteroduplex dsRNA was assessed by the NanoDrop One (Thermo Fisher Scientific) and presence of the dsRNA was verified by a 1% TAE agarose gel (2.2.2). The Heteroduplex dsRNA was stored at -20°C until production of the nanoparticles.

2.3.2.2 Production of Heteroduplex Nanoparticles

Production and processing of the nanoparticles was carried out by Johannes Büscher, Mark Sabura and Kristela Shehu in cooperation with the Schneider Lab working at the Saarland University department of biopharmaceutics and pharmaceutical technology.

dsRNA and DEAE-dextran were mixed in a mass ratio of 1:2 in aqueous solution to facilitate nanoparticle formation due to electrostatic interactions between negative charges of the dsRNA and positive charges of the DEAE-dextran polymer. This results in dsRNA-dextran-nanoparticles that carry a net positive charge. Size of particles and net charge were measured

using a Zetasizer Ultra (Malvern Panalytical) and particles were visualized by transmission electron microscopy using a JEOL JEM-2100 in bright-field mode with a slow-scan charge-coupled device camera at 200kV operating voltage. Pictures were taken from particle suspension drop casted onto carbon coated copper grids without any further staining after air drying.

To increase uptake rates in *Paramecium tetraurelia*, heteroduplex nanoparticles were labeled onto *E. coli* by harvesting an overnight culture of *E. coli* by centrifugation (4.000rpm, 10min), washing the bacteria twice in 1xPBS and resuspending the *E. coli* in 1xPBS in the same volume as initially harvested. Then, 1ml of the *E. coli* suspension were mixed with 50µl of the produced nanoparticle suspension and incubated for 30min at 37°C under constant shaking (180rpm).

Labeled bacteria suspended in 1xPBS were mixed with 20ml WGP medium and *paramecia* were grown using the heteroduplex nanoparticle medium for 24h at room temperature.

PBS Buffer(10x)	
Na_2HPO_4	1.78g
KH_2PO_4	0.24g
NaCl	8g
KCl	0.2g

Substances were dissolved in a total volume of 100ml in water. For use, buffer was diluted 1:10 in water.

2.4 Protein Associated Methods

For some of the methods described in the following section, different antibodies were needed. All antibodies used in this thesis are displayed in the table below (Table 4).

Table 4 Overview of different antibodies used in IFA and Western Blot.

Name	Host	Method	Target Peptide/Protein	Specification	Dilution	Origin
M2 FLAG	Mouse	WB	DYKDDDDK (FLAG-Tag)	Monoclonal	1:500	Sigma-Aldrich
Custom GFP Antibody	Rabbit	IFA/WB	GFP	Polyclonal	1:300/ 1:500	Helmut Plattner
Alpha-Tubulin Antibody	Mouse	WB	Alpha-Tubulin	Monoclonal	1:500	Sigma-Aldrich
Anti-Mouse IgG (H+L)	Goat	WB	Antibodies Produced in Mouse	Polyclonal, Peroxidase Conjugated	1:3000	Jackson Immuno Research
Anti-Rabbit IgG (H+L)	Goat	WB	Antibodies Produced in Rabbit	Polyclonal, Peroxidase Conjugated	1:3000	Jackson Immuno Research
Anti-Rabbit IgG (H+L)	Goat	IFA	Antibodies Produced in Rabbit	F(ab')2 Fragment, AlexaFluor594 Conjugated	1:3000	Thermo Fisher Scientific

2.4.1 Extraction of Bulk Protein from Vegetative *Paramecium* Cultures

Bulk protein samples from vegetative cultures were used to ensure presence of specifically tagged proteins from transgenic *Paramecium tetraurelia* cultures (e.g. FLAG-HA-tagged or GFP-tagged proteins of interest).

For bulk protein extraction, 10.000 cells were harvested by centrifugation in pear-shaped flasks using an oil test centrifuge (2.000rpm, 2min). The narrow part of the pear-shaped flask was sealed using a closed serologic pipette, and the supernatant was decanted. Volvic water was added to the cells. The cells were resuspended and starved for 30min to remove bacterial contaminations by allowing *paramecium* to digest food bacteria present in food vacuoles. After starvation, cells were harvested again, transferred to a 2ml reaction tube, and centrifuged to reduce the leftover volume to approx. 150µl. 50µl 4xLaemmli buffer (BioRad) supplemented with β-mercaptoethanol (1:10, BioRad) was added to the cells, mixed by vortexing, and boiled for 10min to ensure cell lysis. Laemmli-extracted proteins were either immediately used for SDS-PAGE and Western Blot (2.4.2) or stored at -20°C until further use.

2.4.2 SDS-PAGE and Western Blot

2.4.2.1 SDS-PAGE

Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) follows the principle of size separation by electrophoresis as described prior (2.2.2) but allows the separation of proteins instead of DNA/RNA. Here, the SDS denatures and binds to the protein, masking the proteins natural charge and providing a negative net charge. This allows for the separation of the proteins in an electrical field by size alone, making other differences between proteins neglectable. The matrix providing the “sieve-effect” used in a protein electrophoresis is a polyacrylamide gel matrix.

Discontinuous gels were prepared by mixing the separation gel (Table 5) and casting the gel between two glass plates in a casting station (Mini-PROTEAN Tetra cell casting Module, BioRad). The space between the glass plates was filled to about two-thirds and then topped with isopropanol. After polymerization of the separation gel, isopropanol was removed and the stacking gel was prepared and added on top of the separation gel, inserting a comb after completely filling up the room between the glass slides to create the pockets. After polymerization of the stacking gel, the casted gel was either used immediately or wrapped in wet paper towels and stored at 4°C until use.

Table 5. Recipe for the assembly of the stacking gel and the separation gel for SDS-PAGE

	Stacking Gel (5%)	Separation Gel (8%)
Acrylamide (40%, 29:1)	830µl	4.8ml
Stacking Gel Buffer	500µl	-
Separation Gel Buffer	-	3.5ml
Water	3.6ml	8.4ml
SDS (10% w/v)	50µl	167µl
APS(10% w/v)	10µl	100µl
TEMED	5µl	10µl

Stacking Gel Buffer	
Tris	6.05g

Substances were dissolved in a total volume of 100ml in water and pH was adjusted to 6.8.

Separation Gel Buffer	
Tris	18.17g

Substances were dissolved in a total volume of 100ml in water and pH was adjusted to 8.8.

Gels were placed in the electrophoresis chamber (Mini-PROTEAN Tetra vertical electrophoresis cell, BioRad), covered with SDS running buffer and equilibrated for 10min at 80V. Isolated bulk protein was centrifuged (13.000rpm, 10min) and 25µl supernatant containing soluble protein was loaded on the gel. 3µl of the Broad Range Color prestained Protein Standard ladder (NEB) was loaded as a size marker. Electrophoresis was performed at 80V until proteins entered the separation gel, after which electrophoresis continued at 120V until the blue colored dye front left the gel.

SDS Running Buffer (10x)	
Tris	3g
SDS	1g
Glycine	14.4g

Substances were dissolved in a total volume of 1l in water and pH was adjusted to 8.2. Before use, buffer was diluted 1:10 in water.

2.4.2.2 Western Blot

Western Blot is a method that allows for specific detection of proteins of interest within a sample by blotting the protein to a membrane (nitrocellulose) and detecting it with the use of specific antibodies.

Here, Western Blot is used to show that tagged versions of proteins are expressed by transgenic *paramecium* cell lines. This ensures that other downstream treatments of the samples/cells, like immunoprecipitations or immunofluorescence assays (2.4.3.1/ 2.4.4), can work.

For Western Blot, SDS-PAGE of samples was performed and SDS gels were transferred to 1x blotting buffer. Three sheets of Whatman filter paper were soaked in 1x blotting buffer and stacked onto a transfer cell. The nitrocellulose membrane was soaked in buffer as well and stacked onto the Whatman filter paper layer. The equilibrated SDS gel was placed in the center of the membrane and topped off with three additional layers of soaked Whatman filter paper. Air bubbles trapped in the different layers were removed, the transfer cell was wetted with 1xblotting buffer, excess buffer was removed, and the cell was sealed with the transfer cell lid. 150mA was applied to the transfer cell for 1.5-2h to facilitate the transfer of the proteins from the gel to the nitrocellulose membrane.

After blotting, the nitrocellulose membrane was transferred to a Ponceau S Staining solution, incubated for 10min, and the staining solution was removed from the membrane by rinsing the membrane twice in water. Air drying the membrane allows the staining solution to visualize blotted proteins in an unspecific manner to check for successful transfer of protein to the membrane.

For protein detection, the membrane was blocked thrice for 10min with 10ml blocking buffer under constant rotation. After removal of the blocking buffer, the primary antibody (diluted in

blocking buffer) was applied to the membrane and incubated overnight at 4°C under rotation. This antibody binds to the protein of interest, labeling it for subsequent detection.

After removal of the primary antibody, the membrane was washed thrice with 10ml TBST and then incubated with the secondary antibody (diluted in blocking buffer) for 1h, allowing it to bind to the primary antibody. To remove the unbound secondary antibody, the membrane was washed thrice with 10ml TBST and then developed by covering the membrane with Western-Bright Sirius substrate (Advansta) following the manufacturer's instructions. This substrate reacts with the horse radish peroxidase that is coupled with the secondary antibody and provides a chemiluminescent signal at the position of the blot, where the protein of interest is located. The chemiluminescent signal is detected using the CCD camera of the Amersham Imager600 (GE Healthcare) and documented.

Blotting Buffer(10x)	
Tris	30.3g
SDS	4.9g
Glycine	14.41g

Substances were dissolved in a total volume of 1l in water and pH was adjusted to 8.2. Before use, buffer was diluted 1:10 in water.

Ponceau S Staining Solution	
Ponceau S	10g
Acetic Acid	50ml

Substances were dissolved in a total volume of 1l in water.

TBS Buffer (20x)	
Tris	60.57g
NaCl	175.32g

Substances were dissolved in a total volume of 1l in water. Before use, buffer was diluted 1:20 in water.

TBST	
TBS (20x)	50ml
Tween-20	0.5ml

Substances were dissolved in a total volume of 1l in water.

Blocking Buffer	
TBST	100ml
Milk Powder	5g

Substances were mixed and filtered through a filter paper.

2.4.3 Localization of Tagged Proteins Within the Cell

2.4.3.1 Immunofluorescence Assay

Immunofluorescence assay (IFA) is an immunostaining technique in which antibodies are used to visualize structures within cells. In this case, antibodies against specifically tagged proteins were used to localize the tagged protein in *Paramecium tetraurelia* using a for *paramecium* adapted protocol (Frapporti et al., 2019).

In short, 10.000 transgenic *paramecium* cells were washed once in Volvic water and then starved for 30min to reduce background signals from bacteria inside food vacuoles. Cells were then centrifuged (2000rpm, 2min), collected in 500µl and then permeabilized by adding 500µl permeabilization buffer with subsequent incubation for 30min. Permeabilization as well as all subsequent steps were carried out under gentle rotation to avoid clumping of cells. Additional fixation was carried out by transferring the permeabilized and mildly fixated cells to 7ml of fixation buffer for 10min, followed by two wash steps using blocking buffer for 10min. Cells were then stored at 4°C in 1ml blocking buffer until further use.

For immunostaining, 50µl prepared cells were incubated in a final volume of 300µl blocking buffer containing the primary antibody directed against the protein of interest. After incubation overnight at 4°C to allow binding of the antibody to its epitope, cells were washed twice in 300µl blocking buffer and were then transferred to 300µl blocking buffer containing the secondary, fluorescently labeled antibody for 1h. After adding the secondary antibody, cells were kept in the dark and all subsequent steps were carried out in the dark to prevent bleaching

of the fluorophores. After an additional wash step in 300µl blocking buffer, cells were resuspended in a final volume of 10µl and transferred to a glass slide. After adding 1µl Vectashield (Vector Laboratories) and 1µl DAPI (0.2µg/ml), cells were sealed with a coverslip and nail polish.

Finished immunostainings were stored at 4°C and documented under a fluorescent microscope (Axio Observer, Zeiss). Exposure time as well as LED intensities and magnifications were kept the same for all related samples to ensure comparability.

PHEM Buffer (4x)	
EDTA	1.17g
HEPES	2.38g
PIPES	7.25g
$MgCl_2$	0.08g

Substances were dissolved in a total volume of 100ml in water, pH was adjusted to 6.9 and solution was sterilized by filtration.

Permeabilization Buffer	
PHEM buffer(4x)	2.5ml
Sucrose	0.4g
Triton-X100	250µl
Paraformaldehyde (4%)	5ml

Substances except paraformaldehyde were dissolved in a total volume of 5ml in water, then paraformaldehyde was added.

Fixation Buffer	
PHEM Buffer(4x)	7.5ml
Sucrose	1.2g
Triton-X100	240µl
Paraformaldehyde (4%)	21ml

Substances except paraformaldehyde were dissolved in a total volume of 9ml in water, then paraformaldehyde was added.

Blocking Buffer	
TBST	50ml
Bovine Serum Albumin	1g

Substances are dissolved and stored at 4°C until use.

2.4.3.2 Life Cell Imaging

While some tagged protein variants require immunostaining for detection, some GFP-tagged proteins were detectable by using the GFP fluorescence alone.

In this case, cells were washed once in Volvic water and starved for 10min in a depression slide. Single cells were transferred onto a coverslip and immobilized as previously described for microinjection (2.1.1.4). Immobilized cells were transferred to a fluorescent microscope (Axio Observer, Zeiss) and GFP fluorescent signals were documented.

Alternatively, cells were washed once in Volvic water and then transferred in 100µl Volvic water mixed with 1µl surface antigen serum and incubated for 10min. Antibodies within this serum bind to the surface antigen of the cells, clumping the cilia in the process and immobilizing the cells. Before being documented under the microscope, cells were transferred onto a coverslip, covered with oil and flattened as described previously (2.1.1.4) (Axio Observer, Zeiss).

While those methods reduce artificial fluorescent signals caused by unspecific binding of used antibodies, it is only suitable to document signals for cell lines showing strong GFP expression. Weaker cell lines might require immunostaining despite expressing GFP.

2.4.4 RNA Immunoprecipitation

RNA immunoprecipitation (RNA IP) is a method that allows analysis of RNA molecules that are bound to specific proteins of interest. It utilizes antibodies directed against the protein which are covalently bound to magnetic beads, allowing enrichment of the target protein from a cell lysate and subsequent extraction and analysis of associated RNA.

For RNA IP, 400.000 transgenic *paramecium* cells were washed in Volvic water (2.000rpm, 2min) and starved for 30min. After washing cells twice in 1x PBS, they were centrifuged (2.000rpm, 2min) and cell pellets were snap frozen in liquid nitrogen and stored at -80°C.

1ml of lysis buffer was added to the frozen cell pellet and the suspension was transferred into a glass douncer, homogenizing 30-40 times on ice to ensure proper lysis of cells and nuclei, which was checked under the microscope. The cell lysate was transferred to a fresh reaction tube and insoluble proteins and cell debris were removed by centrifugation (13.000rpm, 30min, 4°C). The cleared lysate was transferred to a fresh reaction tube, omitting the insoluble pellet. 50µl of the lysate were transferred to 375µl TRI Reagent® LS (Sigma-Aldrich) for total RNA extraction and the leftover volume was used for bead incubation.

Meanwhile, 50µl of antibody-beads (Anti-FLAG® M2 Magnetic Beads from Sigma Aldrich) per IP reaction were washed 3 times in IP buffer by pelleting the beads using a magnetic rack and resuspending the beads in 1ml IP buffer. After the final washing step, beads were resuspended in their original volume with IP buffer.

To the cleared cell lysate, 50µl of washed antibody-beads were added and samples were incubated overnight at 4°C under constant agitation to allow the protein of interest to bind to the antibody-beads. Samples were washed five times by pelleting the beads using a magnetic rack and resuspending them in 1ml IP buffer. Finally, beads were resuspended in 1ml IP buffer. From this IP buffer, the pellet of 100µl of bead suspension was transferred to 100µl 1xLaemmli buffer to perform Western blot (2.4.2), and the pellet of 900µl bead suspension was transferred to 750µl TRI Reagent® (Sigma-Aldrich) for RNA isolation (2.5.1).

Lysis Buffer	
Tris	0.6g
NaCl	0.87g
$MgCl_2$	0.047g
Dithiothreitol	0.015g
Sodium Deoxycholate	0.5g
Triton-X100	1ml
Vanadyl Ribonucleoside Complex (Sigma)	0.084g
Glycerol	10ml
Protease Inhibitor Complete (EDTA-free) (25x, Roche)	5ml

Substances were dissolved in a total volume of 100ml in water.

IP Buffer	
Tris	0.12g
NaCl	0.876g
NP40	10µl
$MgCl_2$	0.0095g
Glycerol	5ml

Substances were dissolved in a total volume of 100ml in water.

2.5 RNA Associated Methods

2.5.1 RNA Isolation from *Paramecium* Cells

50.000 cells were washed once in Volvic water and then starved for 30min if not stated otherwise. Cells were pelleted (2.000rpm, 2min) and lysed in 500µl TRI Reagent® (Sigma-Aldrich), preventing as much water as possible from being carried over. After vortexing, the samples were incubated for 5min at room temperature and then mixed with 100µl chloroform to trigger phase separation. After centrifugation (13.000rpm, 5min) the upper aqueous phase was transferred into a fresh, nuclease-free reaction tube. 1ml of cold isopropanol was added, mixed and RNA was precipitated for at least 30min at -20°C. RNA was pelleted (13.000rpm, 30min at 4°C), washed twice with 1ml 80% ethanol and resuspended in 50µl nuclease-free water after brief air drying.

RNA yield was quantified using the Nanodrop One (Thermo Fisher Scientific) and RNA was stored at -80°C until further use.

2.5.2 RNA Integrity Check

To prevent faulty results from downstream analysis caused by degraded RNA, RNA integrity was checked. For that, 1µg RNA in a total volume of 5µl was mixed with 10µl RNA loading dye and incubated for 5min at 65°C in a water bath. After heat denaturation, RNA was immediately transferred to ice to prevent re-annealing of RNA strands and loaded onto a denaturing agarose gel. After electrophoresis for at least 1.5h at 80V, RNA was visualized under UV light using a gel documentation station.

RNA Loading Dye	
Formamide	5ml
Formaldehyde (37%)	1.6ml
MOPS (10x)	1ml
Glycerol	1.75ml
Bromophenol Blue(0.1%)	0.5ml
GelRed(10000x)	10µl

Substances were dissolved in a total volume of 10ml in water.

MOPS Buffer (10x)	
MOPS	0.21g
NaCl	0.292g
EDTA	0.292g

Substances were dissolved in a total volume of 100ml in water and autoclaved. Before use, buffer was diluted 1:10 in water.

Denaturing Agarose Gel	
Agarose	1.2g
Formaldehyde (37%)	7.5ml
MOPS Buffer (1x)	100ml

Agarose was dissolved in MOPS buffer and formaldehyde was added shortly before gel casting.

2.5.3 DNase Digestion for DNA Removal

During next generation sequencing experiments, DNA contaminations within RNA samples can lead to faulty results, so removal of leftover DNA is crucial for proper RNA sequencing and analysis.

DNA digestion was carried out using the RNA Clean&Concentrator-25 Kit (Zymo Research) following the manufacturer's instructions for total RNA clean-up, including the optional in-column DNase I treatment steps.

2.5.4 RNA Sequencing

In this thesis, sequencing of smallRNA and longRNA was performed. SmallRNA sequencing refers to the sequencing of small non-coding RNAs while longRNA sequencing can contain both, coding and non-coding longRNA depending on the longRNA sequencing method. RNA species sequenced for each method are described in more detail within the different sections.

2.5.4.1 smallRNA Enrichment

Before smallRNA libraries were produced, smallRNA molecules were enriched by urea-PAGE. This increased efficiency of the smallRNA library procedure compared to libraries where total RNA is used as input.

First, a urea-PAGE was prepared by assembling the gel mixture as described below and casting the gel between two glass plates in a casting station (Mini-PROTEAN Tetra cell casting Module, BioRad), adding a comb after filling the volume between the glass plates to create the pockets. After polymerization, the gel was placed in an electrophoresis chamber (Mini-PROTEAN Tetra vertical electrophoresis cell, BioRad), covered with 0.5x MOPS buffer and equilibrated for 30mins at 100V. Prior to sample loading, pockets were rinsed with 0.5x MOPS buffer.

Urea-PAGE	
Urea	4.2g
Acrylamide (40%, 19:1)	3.75ml
MOPS Buffer (10x)	0.5ml
Water	2.5ml
APS (10%w/v)	70µl
TEMED	3.5µl

Substances except APS and TEMED were dissolved in an ultrasonic bath, then the last two substances were added.

Samples were prepared by resuspending 10-20µg RNA in 4µl water and 13µl smallRNA loading buffer was added. Samples were incubated for 15min at 55°C in a water bath and immediately loaded onto the prepared urea-PAGE (adding 6µl of the NEB microRNA ladder as size reference).

50V was applied to the urea-PAGE until samples entered the gel matrix. Then voltage was increased to 150V until the loading dye reached the bottom of the gel. The gel was transferred to 50ml 0.5x MOPS buffer (supplemented with 5µl SYBR Gold dye, Life Technologies) and stained for 10min under constant shaking. After rinsing the gel twice with water, stained RNA lanes were visualized on a blue light table and the range corresponding to 18-30nt in sizes were excised from the gel with a scalpel.

Gel slices were transferred to a fresh 1.5ml reaction tube, 300µl NaCl (0.3M) was added, and slices were crushed using a pestle. Gel pieces were incubated overnight at 4°C while rotating to facilitate RNA extraction from the gel. After incubation, the suspension was transferred onto a Spin-X-column (Costar) and centrifuged (13.000rpm, 2min at 4°C). The gel pieces left in the filter module were rinsed once with the flowthrough and then, 1/10 volume 3M sodium acetate (pH 5.2), 1 volume isopropanol and 1µl GlycoBlue (Ambion) was added to the filtered smallRNA for precipitation.

Precipitation was carried out overnight at -20°C followed by subsequent centrifugation (13.000rpm, 30min at 4°C) and washing of the RNA pellet with 80% ethanol twice. Pellets were briefly air dried and then resuspended in the volume necessary for smallRNA library preparation (2.5.4.4).

2.5.4.2 Poly-A RNA Library Preparation

For transcriptomic analysis, poly-A RNA sequencing was performed. During the library preparation of this sequencing technique, RNAs with a poly-A tail are enriched and preferably sequenced, leading to sequencing of mainly mRNA molecules, thereby allowing for accurate gene expression level determination.

Transcriptome libraries were produced using the NEBNext® Ultra II Directional RNA Library Prep Kit for Illumina® Kit (NEB) combined with the NEBNext® Poly(A) mRNA Magnetic Isolation Module (NEB) according to the manufacturer's instructions, using 500ng DNase-treated total RNA as input (2.5.3). PCR amplification was performed with 14 PCR cycles and samples were indexed using the different index primer provided with the kit to allow

multiplexing. Library clean-up was performed by AMPure XP bead (Beckman coulter) purification according to the manufacturer's instructions.

2.5.4.3 RNA Ribodepletion Library Preparation

While poly-A-RNA sequencing enriches for RNAs carrying a poly-A-tail, ribodepletion sequencing uses specifically designed probes directed against the ribosomal RNA of the organism to degrade the ribosomal RNA, allowing for the sequencing of all other longRNA species.

Ribodepleted libraries were produced using the NEBNext® Ultra II Direction RNA library Prep Kit for Illumina® Kit (NEB) combined with the NEBNext® RNA Depletion Core Reagent Set (NEB) according to the manufacturer's instructions, using 500ng DNase-treated total RNA as input (2.5.3). Probes against the *Paramecium tetraurelia* ribosomal sequences were designed using the NEBNext® Custom RNA Depletion Design Tool and ordered at Microsynth AG. PCR amplification was performed with 14 PCR cycles and samples were indexed using the different index primer provided with the kit to allow multiplexing. Library clean-up was performed by AMPure XP bead (Beckman coulter) purification according to the manufacturer's instructions.

2.5.4.4 smallRNA Library Preparation

During smallRNA sequencing, only small non-coding RNA species like siRNAs are included into the library, allowing for specific sequencing of those regulatory smallRNAs.

Library preparation was carried out using the NEBNext® Small RNA Library Prep Set for Illumina® (NEB) with size-selected smallRNA as input (2.5.4.1). PCR amplification was performed with 14 PCR cycles and samples were indexed using the different index primer provided with the kit to allow multiplexing. Library clean-up was performed by acrylamide gel extraction according to the manufacturer's instructions.

2.5.4.5 Library Quality Control

Before libraries were subjected to next-generation sequencing, quality of the produced libraries was assessed to prevent sequencing and analysis of low-quality samples. This quality control procedure was independent from library type and was performed with every single library prior to sequencing.

After library preparation, yield of libraries was measured using the Qbit 4 Fluorometer in combination with the 1x HS dsDNA Kit (both from Invitrogen). Size distribution of the library molecules was estimated using the Bioanalyzer system in combination with the High Sensitivity DNA Kit (both from Agilent) or the QSep1 Bio-Fragment Analyzer in combination with the Standard Cartridge (S2) (both from Nippon Genetics).

Libraries showing both a suitable size distribution and enough yield were subjected to next-generation sequencing.

2.5.4.6 Next Generation Sequencing

Sequencing of libraries was carried out either at the department of epigenetics at Saarland university or at the Competence Centre for Genomic Analysis (CCGA) Kiel. Libraries were multiplexed at an equimolar ratio and sequencing pools were sent to the sequencing facilities for Illumina sequencing on an HiSeq2500, NextSeq500 or NovaSeq6000 platform, depending on library type. LongRNA libraries were usually sequenced in paired-end mode with 2x101bp read length while smallRNA libraries were sequenced in single-end mode with 1x50bp read length.

2.5.5 Bioinformatic Processing

2.5.5.1 Preprocessing of Reads

Prior to further downstream applications, reads derived from the sequencer had to be processed to allow for proper bioinformatic analysis.

Reads were adapter- and quality trimmed by using the Trim Galore tool which uses Cutadapt in adapter-autodetection mode and default quality score cut-offs (Krueger et al., 2023; Martin, 2011). LongRNA reads were discarded if their length after trimming fell below 70nt, while smallRNA reads were size-trimmed for keeping reads within a range of 18 to 30nt.

After trimming, read quality and quantity was assessed with the FastQC and MultiQC tools (Andrews, 2010; Ewels et al., 2016).

2.5.5.2 Mapping of Sequencing Reads

To sort reads to the genomic/plasmid sequence they originate from, different mapping algorithms were used. If not stated otherwise, mapping was performed using mapper plugins for Geneious Prime (https://www.geneious.com) with the most recent version at the time of data analysis.

For longRNA mapping, the Bowtie2 (Langmead & Salzberg, 2012) mapper in end-to-end mode, depending on the scientific question asked (specified at the appropriate results & discussion section), was used.

For smallRNA mapping, the Bowtie (Langmead et al., 2009) mapper was used, allowing zero mismatches within a seed region of 30nt.

Templates for mapping were chosen as necessary for different analysis. Specific templates and sequences used are specified in the appropriate sections.

If no deeper downstream analysis was performed, read counts for specific genes or templates were used to visualize differences between samples.

2.5.5.3 Sequence Logo Analysis

To get insight into different preferences of biogenesis pathways, sequence logo analysis was performed on mapped smallRNA reads to search for significantly over-represented bases at specific positions of smallRNA molecules. For sequence logo analysis, the weblogo tool (Crooks et al., 2004) was used. Sequence logos were normalized for the *Paramecium tetraurelia* genome base composition (using A:0.36; T/U:0.36; G:0.14; C:0.14 as nucleotide frequences as described in Lepère et al., 2009) and scaled to one or two bits.

2.5.5.4 smallRNA Overlap Analysis

Analyzing the length of an overlap between two smallRNA molecules derived from different strands of the same genomic locus might give further information on the biogenesis pathway of those smallRNAs and show differences between e.g. Dicer-processed or Ping-Pong-derived smallRNAs. SmallRNA overlap was calculated by using bam-files of mapped smallRNAs from different loci and analyzing them with the Small RNA Signatures tool by Christophe Antoniewski (Antoniewski, 2014) at the Mississippi 2 Galaxy webserver (https://mississippi.sorbonne-universite.fr).

2.5.5.5 Untemplated Nucleotides Within smallRNA Reads

Some smallRNA molecules carry modifications in the form of additional nucleotides, which are added after processing of the RNA. These nucleotides are not present in the templates the RNAs originate from and are therefore called untemplated nucleotides.

Presence of untemplated nucleotides in smallRNA reads were analyzed using a custom snakemake pipeline (https://www.github.com/greenjune-ship-it/untemplated-nucleotides-search). This pipeline allows for analysis of untemplated nucleotides at one locus for both sense and antisense directed reads separate from each other.

SmallRNA reads were mapped onto the template sequence without allowing any mismatches, using the Bowtie mapper (Langmead et al., 2009). Then, reads mapped to the sequences were extracted, sense and antisense directed reads were separated and sequence logos for each present read length were calculated with the weblogo tool (Crooks et al., 2004). Reads extracted this way were considered not modified with untemplated nucleotides.

In a second iteration, reads not mapped in the first step were trimmed by removing a single base from the 3' end of the smallRNA read using Cutadapt (Martin, 2011) and mapped to the template sequence again, in the same manner as in step one. Reads mapped in this iteration were extracted and processed the same way as reads mapped in step one and were considered to carry a single untemplated nucleotide, since this untemplated nucleotide stopped reads from being mapped in the first iteration prior to removal of the 3' base.

This process was repeated a total of four times, allowing for the analysis of up to three untemplated nucleotides.

2.5.5.6 Analysis of Transcriptomic Data

For transcriptomic analysis, longRNA reads were first mapped as previously described (2.5.5.2) and then, expression levels for each annotated gene were calculated by using the "Calculate expression level" function of Geneious Prime. Gene expression levels between treated and control samples were then compared using the DESeq2 (Love et al., 2014) plugin. Using this plugin, log2 ratios of gene expression level between control and treated (Ptiwi silenced) samples were calculated (log2 fold change) and used to compare the log2 fold change of long RNA corresponding to a specific gene against the log2 fold change of smallRNA levels of this gene in different Ptiwi silencing backgrounds.

2.5.5.7 Read Analysis Using the RAPID Pipeline

For some samples, read counts had to be calculated for many different annotated features, not limited to gene annotations. In those cases, reads were mapped and analyzed using both the RAPID stats and RAPID Vis module of the RAPID tool (Karunanithi, Simon, et al., 2019) with default setting

2.5.5.8 Protein Domain Prediction and Characterization

Protein domain prediction was carried out using the DeepTMHMM tool for transmembrane domain prediction (Hallgren et al., 2022) and the InterPro Scan tool for functional domain prediction (P. Jones et al., 2014) on the amino acid sequence of the protein of interest.

For additional information about a protein, amino acid sequences of proteins were blasted against the NCBI database, the Tetrahymena genome database (Stover et al., 2012) and the Paramecium Database (Arnaiz et al., 2020).

2.5.5.9 *De novo* Prediction of smallRNAs Loaded into Ptiwi Proteins

To analyses the function of different Ptiwi proteins, it was crucial to determine, with which smallRNAs these Ptiwi proteins associate. Therefore, smallRNAs produced from specific SRCs were considered loaded into a Ptiwi, if the SRCs could be predicted by using a cluster definition algorithm using smallRNA sequenced from Ptiwi-IP samples.

For smallRNA cluster definition, reads with a length of 23nt (or 22 and 23nt for Ptiwi12 and Ptiwi15) from Ptiwi-IP samples were extracted and down sampled to the lowest number of reads present between Ptiwi-IP samples to ensure comparability. After that, the ShortStack tool was used to perform cluster definition on the *paramecium* genome, using a minimal coverage of 50 for cluster definition (Axtell, 2013). Cluster predicted this way were overlapped with genomic coordinates of the previously described SRCs, to filter out the known SRCs associated with each Ptiwi-IP sample (Karunanithi, Oruganti, et al., 2019). SRCs predicted to associate with Ptiwis were merged between corresponding Ptiwi replicates and compared against the SRCs considered loaded in other Ptiwis. Visualization of the SRC overlap between different Ptiwi proteins were performed using a Venn diagram, utilizing the Venny webtool version 2.1.0 (Oliveros, 2007).

2.6 Separate Devices, Chemicals and Other Laboratory Equipment

This thesis won't list every single chemical and device used, since most of those chemicals and devices (like e.g. centrifuges, PCR cyclers, fridges, …) belong to the basic inventory of molecular biology laboratories. Specific equipment as well as kits used are mentioned in the appropriate sections of the methods chapter.

3 Characterization of Pds1 and Pds2

3.1 Background

As mentioned before, a forward genetic screen in *Paramecium tetraurelia* managed to reveal many of the core RNAi factors necessary for the induction of the RNAi by Feeding pathway (Marker et al., 2014). While most of the discovered RNAi factors were properly identifiable and therefore, information about their potential function is available through other species, this was not the case for Pds1 (*Paramecium* dsRNA-induced RNAi-specific protein 1). As its name suggests, this protein can only be found in organisms belonging to the *Paramecium* genus, including all species of the *aurelia* complex, *Paramecium multimicronucleatum*, *Paramecium caudatum* (Marker et al., 2014) and, as recently described, also in *Paramecium bursaria* (Jenkins et al., 2021) but not in other ciliates like *Tetrahymena thermophila*. Due to its lack of homologues in other organisms, assumptions about the function of Pds1 are very difficult. Next Generation Sequencing based screens of smallRNAs in mutants of the RNAi components revealed that, while some proteins involved in RNAi by Feeding have overlapping functions with endogenous smallRNA pathways, Pds1 seems to be specific for the exogenous RNAi pathway (Carradec et al., 2015).

Shortly after the publication of Pds1 by Marker et al., 2014, another *Paramecium* specific dsRNA-induced RNAi protein, Pds2, was discovered. However, it was never published publicly. The only information present for Pds2 is a protein domain prediction and a database screening showing that the *Pds2* gene seems to be present in the same organisms as *Pds1*, with the exception that *Pds2* has also been found in *Tetrahymena thermophila* and *Oxytricha trifallax* (Carradec, 2014).

Due to the lack of knowledge about the Pds1 and Pds2 proteins, this chapter aims to localize the two proteins within the cell and tries to use the localization information to further speculate about their potential function.

3.2 Results

3.2.1 Predicted Properties of Pds1 and Pds2

As a first step to gain information about the two proteins, the amino acid sequences of Pds1 and Pds2 were analyzed for potential conserved protein domains and clues for transmembrane domains or signal peptides (Figure 3).

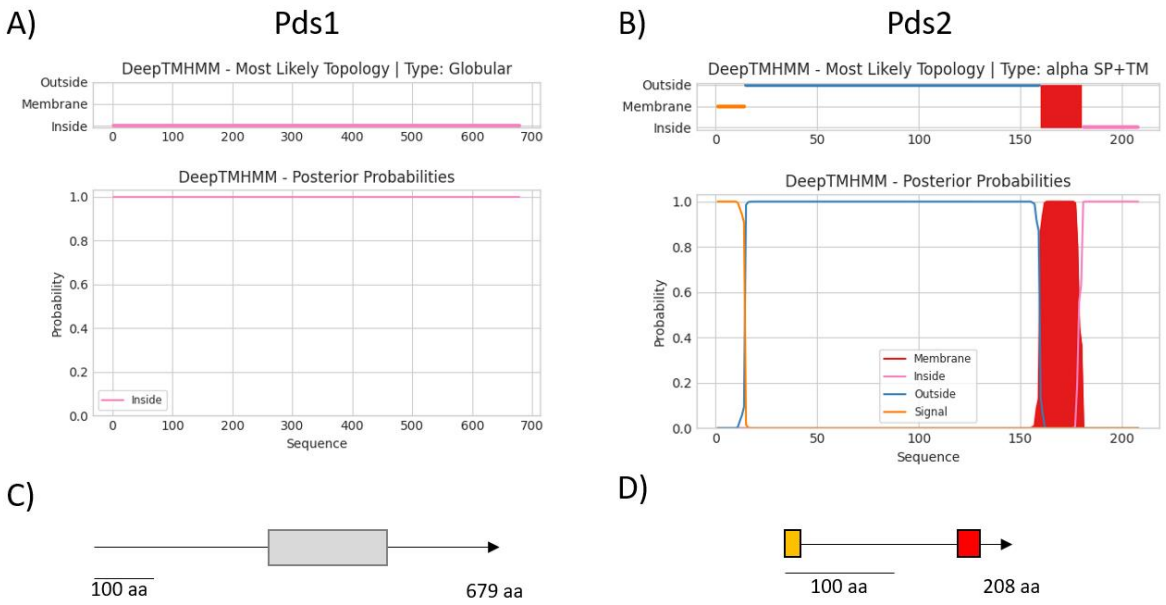

Figure 3 Predicted properties of Pds1 and Pds2.The tools DeepTMHMM and InterPro Scan were used to predict potential transmembrane domains and signal peptides (A,B) as well as other protein domains (C,D) for Pds1 (left) and Pds2 (right), respectively. The gray bar shows a predicted coiled-coil domain for Pds1, the orange bar indicates a signal peptide sequence and the red bar a transmembrane domain for Pds2 visualized as a stylized schematics of the two proteins (C,D).

While the prediction of different protein features for Pds1 did not result in any predicted domains with distinct functions, Pds2 shows a predicted signal peptide leader sequence and a transmembrane domain, with the N-terminus of the protein being assumed to be outside of the compartment while the C-terminal end of the protein probably remains inside (Figure 3). This makes Pds2 the first RNAi component in *Paramecium tetraurelia* that might show membrane localization. So far, an adequate transporter that enables uptake of RNA from the vacuole was not described, so the mode in which RNA enters the RNAi by Feeding pathway remained elusive. Therefore, this makes Pds2 a candidate for the potential RNA transporter that facilitates the entry of RNA from the food vacuole into the cell. However, no RNA interacting domains were predicted for Pds2.

Since domain prediction did not yield much result for both proteins, an additional uniprot blast search using the NCBI protein database was conducted to see whether up-to-date databases would provide similar proteins that could be used to gain insight on potential functions of Pds1 or Pds2. While no hit with a significant E-value (<0.01) and a protein description apart of "putative transmembrane protein" was found for Pds2 with either setting (data not shown), omitting "low complex sequence" filtering, and therefore relaxing search criteria a bit, led to some protein hits with significant E-value but lower alignment score and identity for Pds1. Filtering the hits for proteins associated with RNA led to three hits, one mRNA export factor, one RNA polymerase II associated protein and one ATP-dependent RNA helicase (Table S 1).

3.2.2 Localization of Pds1 and Pds2 Within the Cell

For uncharacterized proteins, knowledge about the cellular localization can help to interpret the protein of interest's function. Therefore, *paramecium* cell lines overexpressing Pds1 and Pds2 fused to a C-terminal GFP tag were generated and localization of the proteins were investigated using the natural GFP fluorescence (2.1.1.4/ 2.4.3.2).

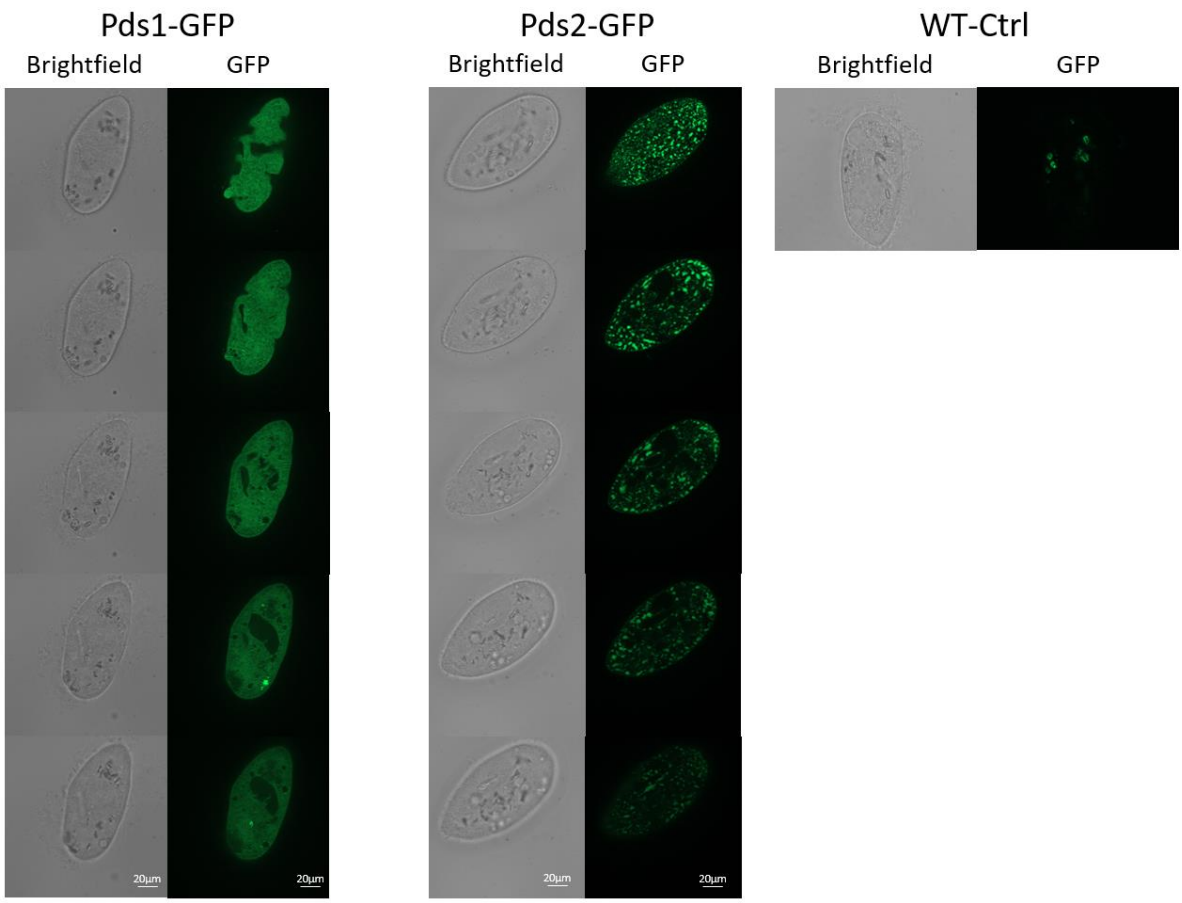

Figure 4 Localization of Pds1 and Pds2 using a GFP tagged fusion protein. Live cell images were taken using a fluorescent microscope (Axio Observer, Zeiss) in structured illumination mode (Apotome2, Zeiss) with an exposure time of 250ms for the GFP channel and 15ms for the bright light channel. Z-stack pictures were taken from transgene Pds1- and Pds2-GFP cells to properly display protein localization from near the surface of the cell (top) down to the middle of the cell (bottom). A wildtype (WT) cell line, not expressing any GFP protein, was used as control. Pictures were taken from cells fed with dsRNA producing *E.coli* to achieve native localization that might be triggered by dsRNA.

As shown (Figure 4), the localizations of the two Pds proteins widely differ from each other. The Pds1 signal shows a broad distribution within the cytosol of the cell, omitting the macronucleus, while the Pds2 signal forms specific loci, especially on the surface of the cell, but also within the cell cortex and cytoplasm. While a cytoplasmatic localization does not provide further insight into the specific function of Pds1, the localization of Pds2 within different membranes of the cell seems intriguing and confirms the protein feature prediction.

To verify protein localization, Western Blot analysis of total protein extracted from transgenic cells was performed to make sure that cells are expressing full length protein (Figure S 1). As shown, full length protein was detected for both Pds1 and Pds2 fusion proteins. However, Pds2-GFP showed an additional band at approx. 250kDa, which is significantly higher than the expected size of monomeric Pds2-GFP.

3.3 Discussion

Even though the presented data allowed for the collection of new information about the Pds1 and Pds2 proteins, for example the localization of Pds2, the function of these two proteins remains elusive. While it seems clear that Pds2 is a membrane protein and therefore prime candidate for the missing RNA transporter needed to import dsRNA into the cytoplasm, no domains have been predicted that suggest interaction with RNA. Additionally, harboring only one single transmembrane domain, Pds2 alone seems to be too small to be capable for membrane transport.

The most prominent RNA transporters in the light of RNAi by Feeding are Sid2 and Sid1. First described in *C. elegans,* Sid2 is necessary for the uptake of dsRNA from food, transporting the dsRNA from the intestine lumen into the cells (McEwan et al., 2012; Winston et al., 2007). While Sid2 is responsible for the initial uptake of dsRNA, Sid1 later on transports smallRNA produced from the dsRNA into other cells of the *C. elegans* body, spreading the silencing/trigger into other body cells, which is called systemic RNAi (Winston et al., 2002). Both transporters carry multiple transmembrane domains.

As a single-cell organism, systemic RNAi is not relevant for *paramecium*. However, no known homologue of Sid2 is present in the genome as well, so uptake of dsRNA must be realized by another protein or in another way.

One way for Pds2 to realize RNA transport despite its short size could be oligomerization, forming a functional channel by assembling with multiple copies of itself. This process of channel formation has been described for many proteins across different organisms, including humans, bacteria and plants (Bhatia et al., 2005; Ni & Hong, 2024; Xuan et al., 2013). Due to a single Pds2 protein only carrying one transmembrane domain, multiple monomers of the protein would be necessary to create a functional pore within the membrane.

One weak hint of oligomerization of Pds2 might be the Western Blot analysis (Figure S 1) of the fusion protein present in the transgenic cell lines. Here, a much larger band that has a similar intensity as the band created by the monomeric Pds2 molecule can be detected. Some protein-protein interactions can be resistant to the denaturation by SDS and the reducing properties of the β-mercaptoethanol present in the Laemmli loading dye, like sterically shielded disulfate-bridges or metal-ion mediated sulfate bonds (Stasser et al., 2005). However, these interactions are unlikely to play an important role for this protein due to the low number of cysteine and methionine present in the amino acid sequence (four and one, respectively).

Another possible protein-protein interaction responsible for oligomerization, dityrosine cross-linking, is also resistant to SDS and β-mercaptoethanol (Atwood et al., 2004) and more likely to be relevant for Pds2 due to large numbers of tyrosine (eighteen) present in in the protein. However, no proof of such interactions between multiple Pds2 molecules exist, so oligomerization of Pds2 remains speculation, for now.

Additionally, the possibility remains that the transporter responsible for RNA uptake escaped all genetic screenings so far and is not known yet. This might be the case if the transporter is essential for the cell or the transporter gene is located within regions of the genome that are not properly sequenced and annotated yet.

Speculation about the actual function of Pds1 remains even more difficult since the localization of Pds1 in the cytosol of the cell doesn't provide specific insights into its potential functions, except that it seems to be essential for dsRNA-induced siRNA production (Carradec et al., 2015). The blast hit reported for database comparison of the Pds1 amino acid sequence is the first clue that might point into a possible direction regarding Pds1 function. However, the low sequence similarity between the provided hit and Pds1, and the specific settings required for the hit being produced in the blast search don't provide hard evidence.

Interestingly enough, a recent publication about RNAi by Feeding in *Paramecium bursaria* collected the occurrence of the different described RNAi components of *Paramecium tetraurelia* within other related ciliate species and showed that Pds1 is present in all of the analyzed species, with the exception of *Tetrahymena thermophila* (Jenkins et al., 2021). With *Tetrahymena thermophila* being regarded as not capable of performing RNAi by Feeding by the scientific community and the emerging reports of the feeding pathway being successfully used in other *Paramecium* species, like *Paramecium bursaria* and *Paramecium caudatum* (Gao et al., 2024; Jenkins et al., 2021), Pds1 seems to be the linchpin component in ciliates that

provides a species with the capability to perform RNAi by Feeding. However, there is a single report claiming to use RNAi by Feeding successfully in *tetrahymena* (Najle et al., 2013), so the capability of *Tetrahymena thermophila* to perform RNAi by Feeding remains to be clarified by the scientific community for good.

4 The Role of Two RdRPs in The RNAi by Feeding Pathway

4.1 Background

RNA dependent RNA polymerases (RdRPs) are enzymes that are capable of synthesizing RNA by using another RNA strand as a template.

Their involvement in different RNA based pathways has been studied in various organisms, including *C. elegans*, *Arabidopsis thaliana* and the fungus *Neurospora crassa*.

Some of the pathways involving RdRPs are regulatory pathways that are triggered by endogenous mechanisms. Examples for such pathways are: (I) endogenous loci in *Arabidopsis thaliana* producing phased siRNAs capable of regulating gene expression in dependency of *Arabidopsis thaliana*'s RdRP RDR2 and RDR6 (Feng et al., 2024), (II) the phenomenon known as quelling in *Neurospora crassa*, that is dependent on the RdRP QDE-1, which leads to a paramutation-like silencing of affected genes (Chang et al., 2012; Cogoni & Macino, 1997, 1999), or (III) piRNA triggered RdRP dependent transitivity in *C. elegans* (Sapetschnig et al., 2015). In the last case, transitivity describes the production of smallRNAs, usually siRNAs, by RdRP activity, using the target of previously produced smallRNAs as a template. In other words, in this particular instance, piRNAs (which are referred to as primary RNAs) target a particular mRNA and RdRP activity synthesizes siRNAs, by using the targeted mRNAs as a template, which are then referred to as secondary siRNAs.

Apart from the endogenously triggered mechanism, RdRP activity was also described in pathways triggered by exogenous dsRNA, especially in *C. elegans*. Here, it was shown that initially applied dsRNA is cleaved by a Dicer enzyme into primary siRNAs, which then bind to mRNAs mediated by the argonaut RDE-1. However, primary siRNAs do not immediately degrade the target mRNA, but trigger RdRP-dependent secondary siRNA production, which is promoted by the RdRP RRF-1. Secondary siRNAs, usually loaded by the WAGO Argonaut, can either attack the target mRNA, leading to its degradation, or go into the nucleus of the cell, establishing chromatin modifications to silence genes on a CTGS level. Therefore, in *C. elegans*, RdRP activity is solely responsible for secondary siRNA production (Pak et al., 2012).

In *Paramecium tetraurelia,* the forward genetic screen revealed two RdRPs, RdRP1 and RdRP2, which are involved in the RNAi by Feeding pathway (Marker et al., 2014). Interestingly, later sequencing-based studies of smallRNAs produced in RdRP mutants revealed that unlike *C. elegans* or *Arabidopsis thaliana*, the lack of RdRPs in *paramecium* leads to a loss of primary siRNAs, showing that the processing of exogenous RNA requires RdRP activity,

whereas in other species, RdRP activity is only necessary for secondary siRNA production after transitivity on mRNA (Carradec et al., 2015).

Those findings led to the assumption that the exogenous dsRNA might be converted into different dsRNA molecules by RdRP activity. It also raises the questions of why RdRPs in *paramecium* are required for the production of primary siRNAs in general, and why two different RdRPs are needed for this step.

Two hypotheses originated from those questions. First, the two RdRPs might convert the two strands of the exogenous dsRNA into two separate dsRNA molecules, with one RdRP being responsible for one strand and the other RdRP for the other strand. Second, the two RdRPs might work together as a complex, requiring each other for full functionality and both being responsible for the conversion of one or both exogenous dsRNA strands into new dsRNA molecules.

4.2 Results

4.2.1 RdRP1 and RdRP2 Co-Localize in the Cytoplasm of Cells

To enhance the existing background information available for RdRP1 and RdRP2, localization of both proteins within *Paramecium tetraurelia* was performed to answer the question of whether RdRP1 and RdRP2 are localized in the same or in different cellular compartments within the *paramecium* cells.

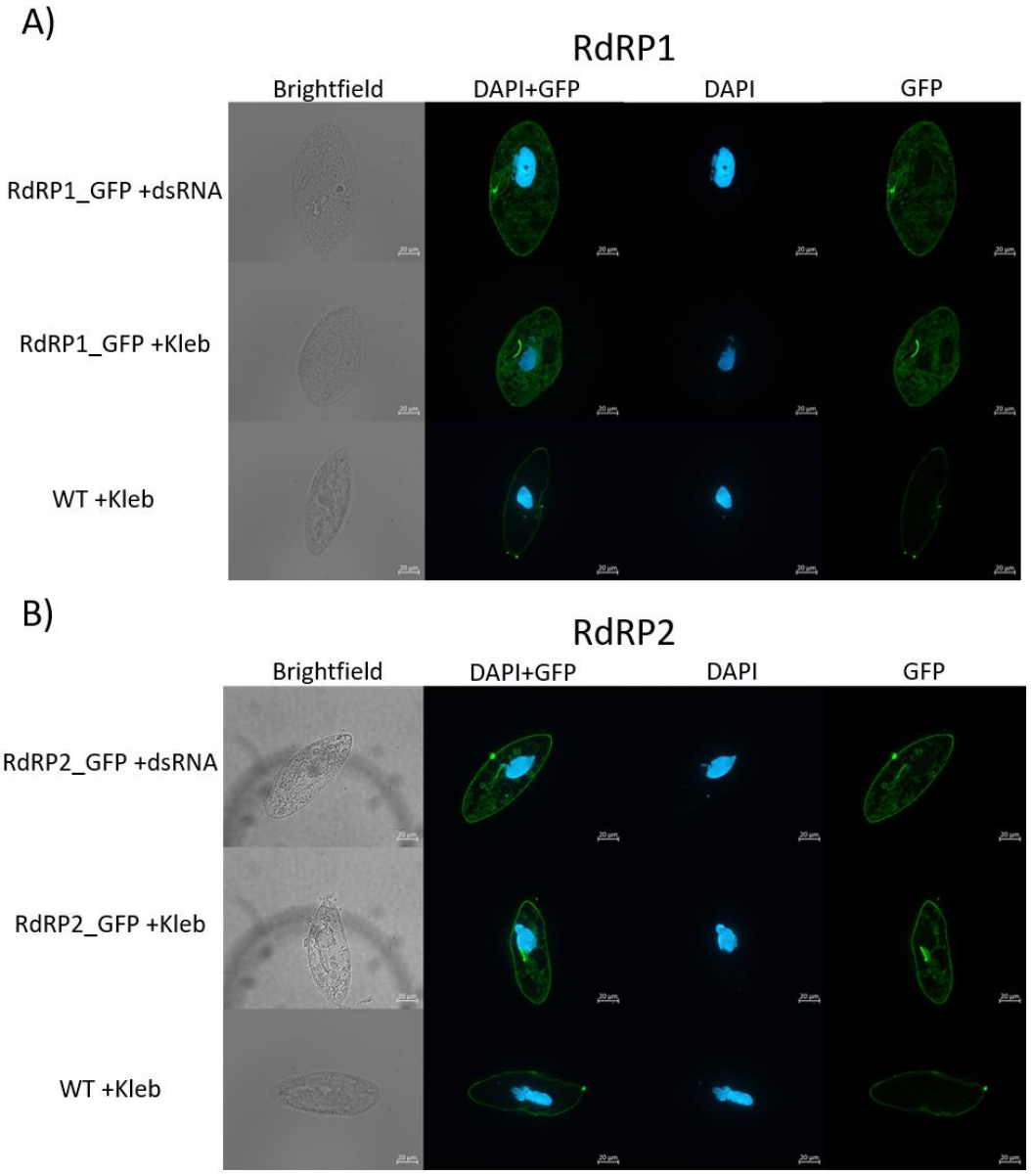

Figure 5 Localization of RdRP1 and RdRP2. Localization was carried out by IFA of cells expressing either RdRP1 (A) or RdRP2 (B) fused to a GFP tag and visualization of the localization using an Alexa594-coupled secondary antibody. Transgenic cells were fed with dsRNA producing *E.coli* (+dsRNA) or *Klebsiella pneumoniae* (+Kleb) to analyze RdRP localization in presence and absence of dsRNA. Wildtype cells (WT) not expressing any GFP were used as a control for unspecific signal of the αGFP-antibody. Exposure time for the different channels were kept the same for each cell to ensure comparability (GFP: 750ms; DAPI: 50ms; Brightfield: .45ms). Images were taken using a fluorescent microscope (Axio-Observer, Zeiss) and structured illumination (Apotome2; Zeiss).

As displayed (Figure 5), both RdRP1 and RdRP2, show a distribution within the cytosol of the cell, omitting the macronucleus. Applying dsRNA-containing and non-containing food to cells expressing the RdRP-GFP fusion proteins does not seem to change localization of the two proteins within the cells, suggesting that their localization is not depending on the presence of dsRNA.

4.2.2 Functionality of the Heteroduplex dsRNA

To answer the main question of whether the initial dsRNA trigger is the substrate of RdRP activity in the RNAi by Feeding pathway, feeding of a Heteroduplex dsRNA to *paramecium* cells was established.

In this study, a Heteroduplex dsRNA is an exogenous dsRNA where both strands of the double strands (Figure 6A, labeled in blue for sense and red for antisense) carry mismatches in relation to the target mRNA sequence. These mismatches can be used to distinguish the exogenous trigger, and RNAs derived from it, from endogenous RNAs of the same target sequence (Figure 6A, indicated by red bars). Additionally, one strand of the Heteroduplex RNA (here, the sense strand in blue) carries mismatches (indicated by blue bars) that are not present in the complementary strand, which distinguishes this strand, and all RNAs produced from it, from the other part of the Heteroduplex.

Both types of mismatches taken together allow for the identification of smallRNA species and their specific origin strand.

By design, the mismatches in the Heteroduplex are only present in certain possible combinations with strand orientations. For example, both types of mismatches together (indicated by blue and red bars) are only present on the sense orientated strand (labeled in blue) of the Heteroduplex while only one set of mismatches (labeled in red) are present on the antisense orientated strand (labeled in red). This means that application of the Heteroduplex and immediate processing of the Heteroduplex dsRNA by Dicer1 into smallRNAs can only yield two different kinds of smallRNA molecules; antisense directed molecules with one set of mismatches (red labeled smallRNAs with red indicated mismatches) and sense directed molecules with both sets of mismatches (blue labeled smallRNAs with blue and red indicated mismatches, Figure 6A lower path). Detection of only these two kinds of smallRNA molecules would therefore suggest immediate processing of the Heteroduplex dsRNA by Dicer.

However, RdRP activity on the initial Heteroduplex dsRNA would increase the possible combinations of mismatches present within a molecule and its orientation. For example, focusing on the sense strand of the Heteroduplex (labeled in blue), RdRP activity would use the sense strand as a template and create a complementary antisense strand (Figure 6A, upper path, labeled in green). This antisense RdRP product now carries both kinds of mismatches described earlier; a combination of mismatches and molecule orientation not possible without RdRP activity. The same is true for the antisense strand of the Heteroduplex (labeled in red) and the adequate RdRP product (labeled in orange). By detecting the Heteroduplex mismatches within smallRNA molecules produced from the Heteroduplex and analyzing the orientation of the smallRNAs carrying those mismatches, it can be concluded whether the Heteroduplex RNA was subjected to RdRP activity or not.

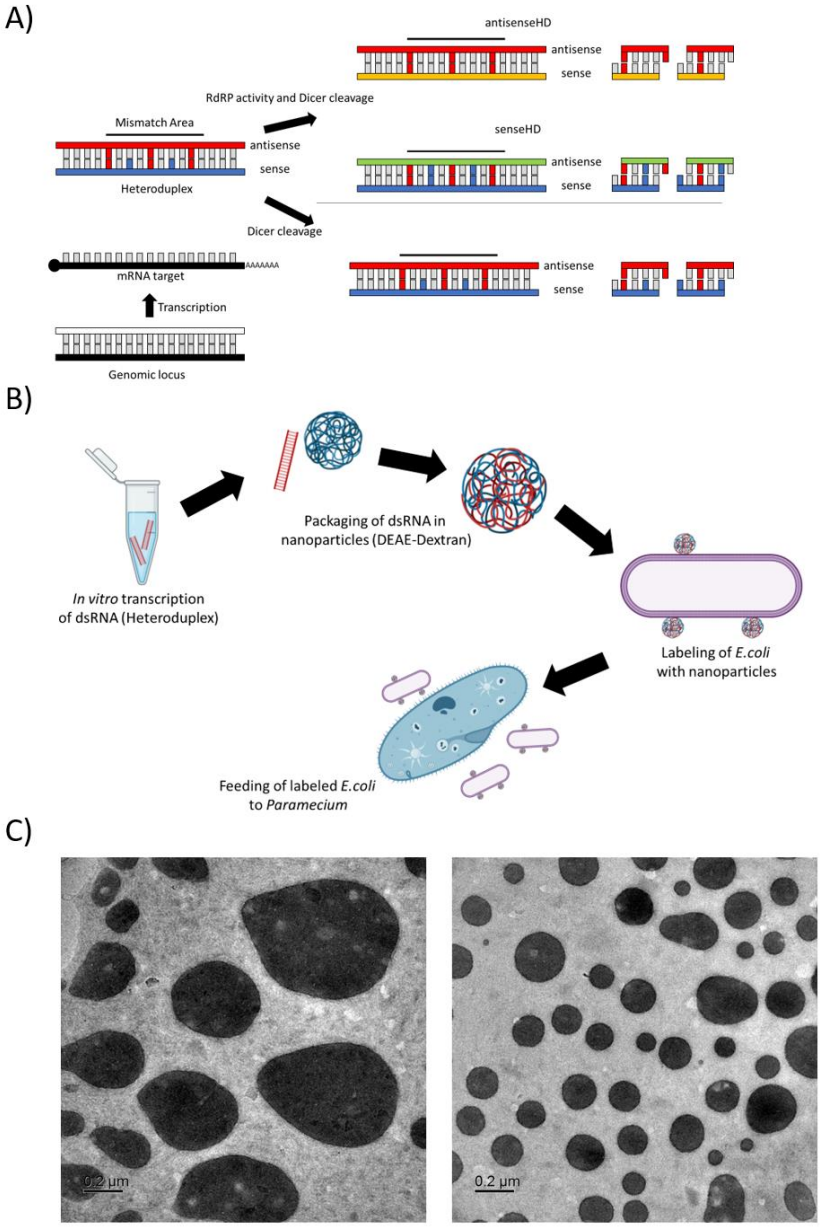

Figure 6 Concept of the Heteroduplex dsRNA delivery. A) Schematic of smallRNA molecules derived from the Heteroduplex trigger dsRNA depending on potential RdRP activity subjected to the Heteroduplex before Dicer cleavage. Mismatches distinguishing the different strands of the Heteroduplex from endogenous RNA and themselves are labeled in red and blue, respectively. B) Workflow of the Heteroduplex nanoparticle application, including Heteroduplex formation by *in vitro* transcription, packaging of Heteroduplex RNA into nanoparticles, labeling of *E. coli* with nanoparticles and feeding of labeled *E. coli* to *Paramecium tetraurelia*. C) Visualization of assembled Heteroduplex nanoparticles by electron microscopy. Parts of the image were created using the BioRender tool.

Application of the Heteroduplex RNA was carried out as described previously (2.3.2, visualized in Figure 6B). After Heteroduplex nanoparticle production, the mean size of the nanoparticles was estimated by electron microscopy, showing that most particles had a diameter of approx. 0.4µm, with some particles showing a larger diameter (Figure 6C).

4.2.3 Application of Heteroduplex Nanoparticles Reveal RdRP Activity on Trigger dsRNA

After establishing Heteroduplex production, packaging the Heteroduplex dsRNA into nanoparticles and application to the cells, smallRNA molecules from cells treated with Heteroduplex RNA were sequenced to analyze produced siRNAs.

As shown below (Figure S 2A), the overall abundance of siRNAs produced from the trigger Heteroduplex dsRNA delivered via nanoparticles (labeled as HD) is lower compared to ordinary dsRNA feeding using dsRNA-producing *E. coli* as a delivery vector (labeled as dsRNA_*Ecoli*). However, nanoparticle derived abundance approx. equals smallRNA abundance in samples were dsRNA-producing bacteria were diluted in a 1:10 ratio with non-dsRNA producing food bacteria (labeled as dsRNA_*Ecoli*_diluted).

Since both, bacteria-produced dsRNA and the nanoparticle-delivered Heteroduplex targets the *nd169* mRNA, a gene involved in the discharge of trichocysts, application of the dsRNA should lead to reduced trichocysts discharge if efficient gene silencing is triggered. For bacteria-produced dsRNA, a concentration-depended silencing effect in regards of trichocyst discharge can be observed, where feeding of higher concentrations of dsRNA results in the inability of cells to discharge their trichocysts (Figure S 2B left, labeled as Trich – in red) whereas dilution of the dsRNA trigger with non-dsRNA producing bacteria in different ratios leads to a steady decrease of the silencing phenotype and an increase of the wildtype phenotype (labeled as Trich + in blue). For nanoparticle-delivered Heteroduplex dsRNA, a slight increase in the number of cells displaying a silencing Trich – phenotype compared to the wildtype control can be observed (Figure S 2B right). However, the silencing effect of the Heteroduplex is not as pronounced as the silencing observed by the bacteria-delivered dsRNA. This fits to the observed abundance of produced smallRNA from the trigger molecules, as nanoparticle delivered Heteroduplex dsRNA seems to produce lower amounts of smallRNA molecules (Figure S 2A), which might lead to a decrease in phenotype.

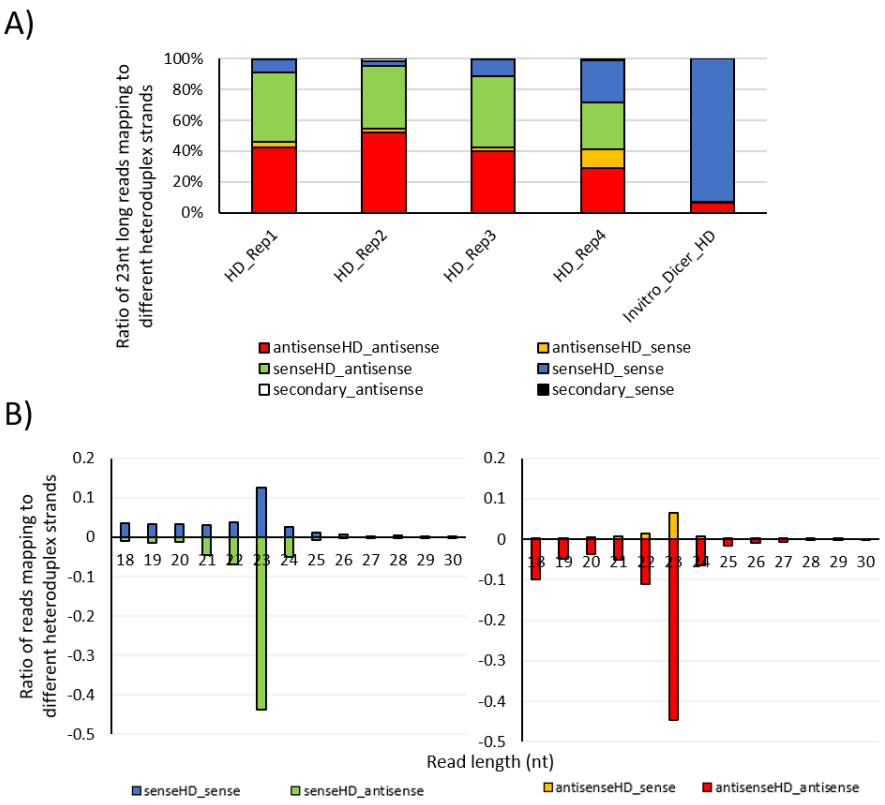

Figure 7 Heteroduplex derived smallRNAs in wildtype cells. A) Ratio of different Heteroduplex derived smallRNA species. Shown are smallRNAs mapping to the original Heteroduplex dsRNA strands (labeled in red and blue), to potential RdRP produced strands (labeled in green and orange), or to the corresponding endogenous sequence without mismatches, representing secondary siRNAs (labeled in black and white). *In vitro* diced Heteroduplex dsRNA served as a control. B) Mean read length distribution of smallRNAs derived from Heteroduplex dsRNA across all four replicates. smallRNAs corresponding to the different possible Heteroduplex strands are color coded as described above. Positive values correspond to sense-directed molecules whereas negative values correspond to antisense-directed molecules.

Analyzing the distribution of smallRNAs derived from the different strands of the Heteroduplex dsRNA, it can be observed that smallRNAs from both, the two original strands of the Heteroduplex (labeled in red and blue) and the two RdRP products (labeled in green and orange) are present within cells subjected to the Heteroduplex dsRNA particles (Figure 7A). Interestingly, the antisense-orientated smallRNA molecules (labeled in red and green) seem to be more abundant in the cells compared to the sense-orientated molecules (labeled in blue and orange), regardless of whether they originate from the original Heteroduplex strands or RdRP produced strands. As a control, Heteroduplex dsRNA was diced in an *in vitro* Dicer reaction to validate that putative RdRP products were a result of enzymatic activities within the cells after dsRNA uptake and not derived from faulty *in vitro* transcription. As suspected, *in vitro* dicing of produced Heteroduplex dsRNA only showed smallRNAs matching to the original Heteroduplex strands, showing that the putative RdRP produced strands and the antisense bias present in the cellular samples were indeed caused by cellular processes.

Finally, the correct processing of the Heteroduplex dsRNA to siRNAs was validated by analyzing the read length distribution of smallRNAs derived from the Heteroduplex. For all four different strands, the original Heteroduplex dsRNA strands and the RdRP products, the read length distribution of sequenced smallRNAs shows a peak at 23nt length (Figure 7B). With 23nt being the hallmark length of Dicer1-derived siRNAs in *Paramecium tetraurelia*, it can safely be assumed that the Heteroduplex dsRNA is accepted by the cellular RNAi by Feeding machinery and that it is properly processed into siRNAs.

To ensure that the reason for the different ratios of the four Heteroduplex strands is not caused by the asymmetrical distribution of the different mismatches across the two original strands, a second Heteroduplex, with the mismatch distribution between the two strands being switched (called SHD: Switched Heteroduplex), was applied to the cells (Figure S 2). As shown, the ratios between the smallRNAs derived from the four different Heteroduplex strands do not alter with relocation of the mismatches, showing that the strand preferences are not caused by stabilization of the strand with fewer numbers of mismatches, but that other cellular processes produce the observed imbalance in strand ratios.

RdRPs can operate in two different modes. They can either produce smallRNAs directly from the template RNA, or they can synthesize a full-length long RNA from the template, which can then form a long dsRNA with the complementary template and get cleaved by Dicer. Since smallRNA derived from the hypothetical RdRP product has been found, the question arises, whether the RdRP products can also be detected in longRNA sequencing data. Therefore, libraries from ribodepleted RNA samples using one of the four Heteroduplex replicates were prepared and sequenced (Figure 8).

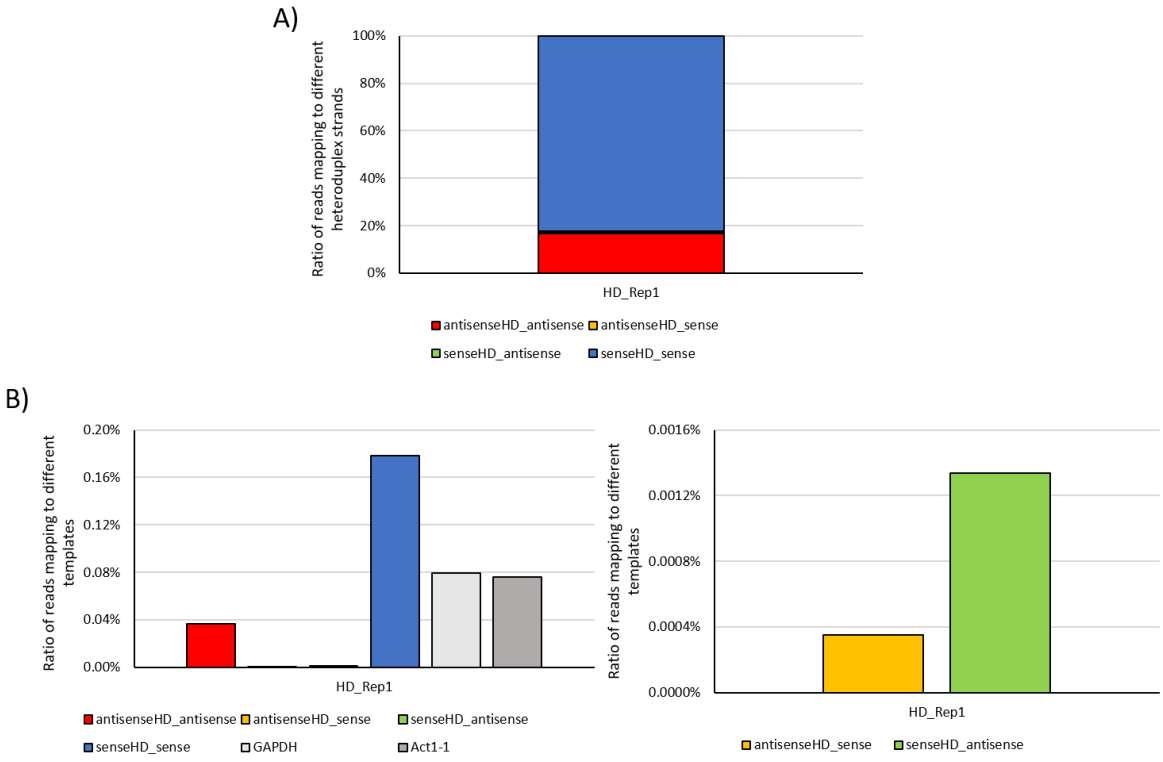

Figure 8 Heteroduplex derived longRNA in wildtype cells. A) Ratio of different Heteroduplex derived RNA species normalized to all Heteroduplex mapping reads. Shown are longRNAs mapping to the original Heteroduplex dsRNA strands (labeled in red and blue) or to RdRP produced strands (labeled in green and orange). B) Abundance of Heteroduplex derived reads mapping to the four different Heteroduplex derived strands (color coded as described before), as well as reads mapping to the two housekeeping genes glyceraldehyde-3-phosphate dehydrogenase (GAPDH, labeled in light grey) and actin1-1 (Act1-1, labeled in dark grey) (left), and the same graph showing only RdRP derived strands (right).

A first analysis of the Heteroduplex derived reads from the ribodepleted total RNA samples shows mainly reads that correspond to the two original Heteroduplex strands (Figure 8A). Quantifying the abundance of the four Heteroduplex derived strands in relation to the number of total reads sequenced shows that, while the two original strands show RNA levels within the cell that can even surpass levels of housekeeping genes (Figure 8B), RNA from the RdRP products can be detected, but is barely abundant, only representing approx. 0.0012% of all reads sequenced. Performing the same analysis with RNA from cells subjected to the switched Heteroduplex sequence revealed the same effect, showing again, that mismatch distribution does not influence the presence and absence of specific strands in the sequencing data (Figure S 3).

Being able to detect the two RdRP derived strands in longRNA sequencing data might suggest that the RdRPs synthesize longRNA using the original Heteroduplex strands as a template prior to Dicer1 cleaving the longRNA into siRNAs, but the low abundance of the RdRP products in the sequencing data makes quantification of the RdRP derived strands and drawing further

conclusions difficult. However, it might be the case that dsRNA, produced by the RdRP activity using one of the original Heteroduplex RNA strands as a template, is immediately cleaved by Dicer1 and converted into 23nt siRNAs. Since sequencing can only detect the steady-state level of RNA within a cell and does not provide information about catalytical dynamics, having the RdRP product being an intermediate product that is rapidly processed by the cellular machinery would explain the low abundance of those strands in the longRNA sequencing data, but their generally higher abundance in smallRNA data.

Summarizing those findings, since no other enzymes besides RdRPs are capable of producing the two strands not originally delivered by the Heteroduplex dsRNA, RdRP activity on the Heteroduplex trigger dsRNA has to occur. However, whether those RdRP products are only present in smallRNA data or whether the RdRP activity occurs before the Dicer1 cleavage on the longRNA level, is debatable.

4.2.4 Mutations in RdRP1 and RdRP2 Lead to Loss of Putative RdRP Products

After detection of the RdRP products in smallRNAs derived from Heteroduplex dsRNA in wildtype cells, the next step was to determine the required enzymes for their production. Since RdRP1 and RdRP2 have been reported to be involved in the RNAi by Feeding pathway, different mutant strains (Figure 9A) of both genes were subjected to Heteroduplex dsRNA feeding, and smallRNAs produced from the different strains were sequenced. For each RdRP, two mutant strains were examined. For RdRP1, one mutant carries a premature termination codon 374 amino acids deep into the protein upstream of the RdRP domain, resulting in a catalytical dead protein, while the other carries a premature termination codon 30 amino acids before the proteins native ending (Figure 9A). For RdRP2, one mutant carries an E to K amino acid substitution within the RdRP domain, which does not alter the catalytical motif, while the other displays a IES retention at codon 1276, leading to a frame shift and a resulting premature termination codon at codon 1330, 13 amino acids before the proteins native ending (Figure 9A).

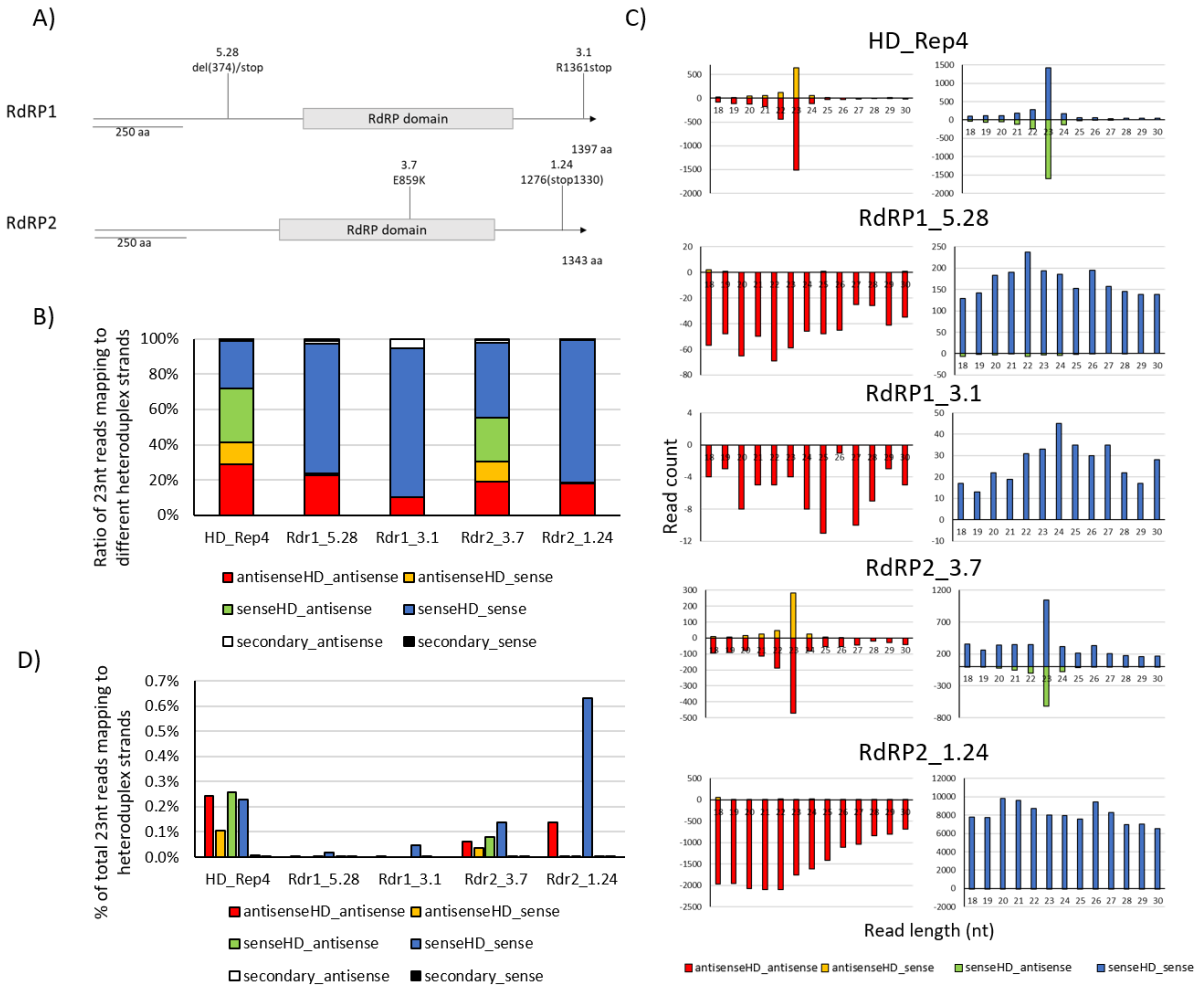

Figure 9 Heteroduplex derived smallRNAs in different RdRP mutants. A) Overview of different RdRP1 (top) and RdRP2 (bottom) mutant strains. Given is the number of the different strains, the position of the mutation within the RdRP protein and what kind of mutation is present. B) Ratio of different Heteroduplex derived smallRNA species. Shown are smallRNAs mapping to the original Heteroduplex dsRNA strands (labeled in red and blue), to potential RdRP produced strands (labeled in green and orange), or to the corresponding endogenous sequence without mismatches, representing secondary siRNAs (labeled in black and white). C) Read length distribution of smallRNAs derived from Heteroduplex dsRNA from different wildtype and mutant strains. smallRNAs corresponding to the different possible Heteroduplex strands are color coded as described above. Positive values correspond to sense-directed molecules whereas negative values correspond to antisense-directed molecules. D) Abundance of reads mapping to the different Heteroduplex strands or the corresponding endogenous sequence without mismatches, color coded as described above.

Analyzing the occurrence of smallRNAs derived from the different Heteroduplex strands in the mentioned mutant strains, it can be observed that only one of the analyzed mutants, mutant 3.7 of RdRP2, contains all four possible smallRNA species, similar to the wildtype. All other studied mutants, mutant 3.1 and 5.28 of RdRP1 and mutant 1.24 of RdRP2, show a loss of the RdRP produced strands while maintaining the presence of the two original Heteroduplex strands (Figure 9B). A similar observation is true for the read length distribution of sequenced smallRNAs from the different Heteroduplex strands. While the wildtype sample and mutant 3.7 of RdRP2 show the distinct 23nt siRNA peak for all four Heteroduplex smallRNA species, the

other three mutants show a loss of the siRNA peak (Figure 9C). Quantification of the strands present in the different samples normalized to the total read count reveal that all mutants show, to different extends, a decrease in Heteroduplex related smallRNA reads, even mutant 3.7 of RdRP2, which showed wildtype quality of sequenced smallRNAs, but a decrease in quantity.

Taking all those results into consideration, several different aspects about the involvement of RdRPs in the production of Heteroduplex derived smallRNAs can be observed.

First, the two RdRP produced strands derived from the Heteroduplex dsRNA show a dependency on both, RdRP1 and RdRP2. This means that both RdRPs operate in the production of RdRP products from both Heteroduplex strands, which refutes the hypothesis that one RdRP might be responsible for the production of the complementary strand from only one strand of the trigger dsRNA, while the other RdRP produces the complementary strand from the other.

Second, the presence of one of the RdRPs in its wildtype form is not sufficient to rescue the impaired function of the other, mutated RdRP, meaning that both RdRPs are not redundant in their function, but are probably either, performing individual functions that are both required for RdRP product formation, or are working together as a complex to perform the same function.

Third, the production of 23nt long siRNAs probably occurs after the synthesis of the RdRP derived strands, which means that the trigger dsRNA taken up by the cell from the food bacteria is not substrate for Dicer1. This can be concluded due to the loss of the 23nt siRNA peak in the read length distribution of all four Heteroduplex derived strands. If the initial trigger dsRNA would be the substrate for Dicer1 cleavage and siRNA production, loss of the ability to produce the RdRP derived strands should not impair Dicer1's ability to cleave the already existing dsRNA and should lead to maintenance of the 23nt peak in the original Heteroduplex strands (Figure 9C, labeled in blue and red) in RdRP mutants, which is not the case. This leads to the conclusion that Dicer1 needs the presence of the RdRP produced strand to facilitate dsRNA cleavage and siRNA formation, meaning that the two newly formed dsRNA molecules, consisting of one original Heteroduplex strand and one RdRP produced strand each, are probably the substrate for Dicer1 activity.

Fourth, the production of the RdRP produced strands is not solely dependent on the catalytic domain of the RdRPs, but is susceptible to structural anomalies of the two RdRPs. This can be concluded due to the fact that two of the studied mutants, namely mutant 3.1 of RdRP1 and mutant 1.24 of RdRP2, do not show alterations in the RdRP domain, but only an aberrant C-

terminus, due to a frame-shift creating mutation or the formation of a premature termination codon (Figure 9A).

Applying the newly gained insight, ribodepleted RNA from the different mutant strains was sequenced to see whether the different RdRP mutations have an impact on the composition and abundance of longRNA derived from the Heteroduplex RNA (Figure 10).

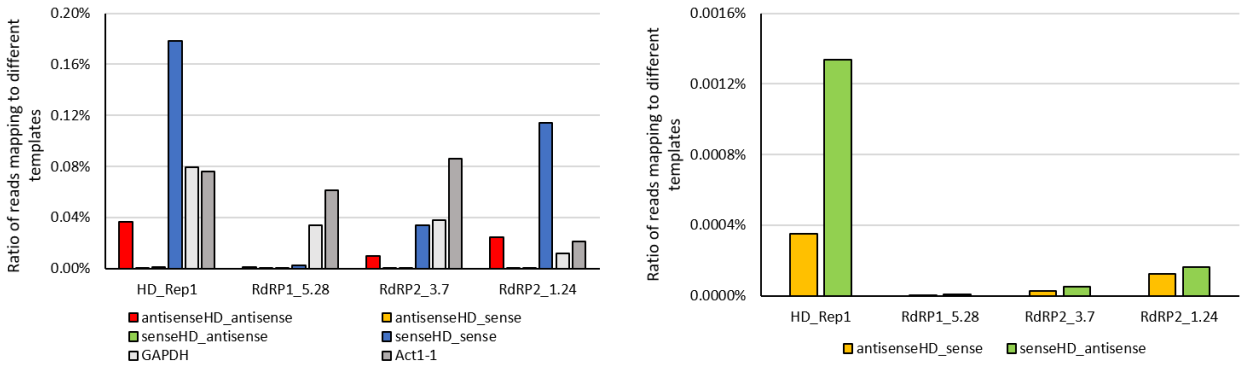

Figure 10 Heteroduplex derived longRNA in different RdRP mutants. Abundance of Heteroduplex derived reads mapping to the four different Heteroduplex derived strands (original Heteroduplex dsRNA strands labeled in red and blue, and RdRP produced strands labeled in green and orange), as well as reads mapping to the two housekeeping genes glyceraldehyde-3-phosphate dehydrogenase (GAPDH, labeled in light grey) and actin1-1 (Act1-1, labeled in dark grey) (left), and the same graph showing only RdRP derived strands (right).

Similar to the smallRNA analysis, sequencing of the longRNA present in the different mutants shows a reduction in abundance for the RdRP products in all mutants, with the mutant 1.24 of RdRP2 showing slightly higher levels of RdRP produced strands compared to the other mutants. Interestingly, levels of the original Heteroduplex strands are decreased in the mutants as well, when compared to the wildtype (Figure 10).

All in all, observations of the longRNA data coincide with the observations made for smallRNA data. All strains harboring mutations in the different RdRP genes contain lower amounts of RdRP produced strands in comparison to the wildtype sample in both smallRNA and longRNA sequencing data, which underlines the dependency of the production of those strands from RdRP1 and RdRP2.

Following this hypothesis of the role of RdRP function in RNAi by Feeding in *paramecium*, levels of the original Heteroduplex strands in longRNA sequencing data shouldn't be dependent on RdRP activity and therefore, shouldn't decrease if RdRP activity is lacking. However, this can be observed, especially in the RdRP1 mutant. It cannot be ruled out that after production of

the RdRP produced Heteroduplex strands, another round of RdRP amplification occurs, synthesizing the equivalent of the original strand from the RdRP produced strand as a template. The experimental setup doesn't allow for the differentiation between the original exogenous strand and potential copies of the same strand sequence produced by the cell, making it not possible to rule out this possibility. However, loss in abundance of the original Heteroduplex strands on longRNA level might also be caused by random degradation of the two strands, which might be promoted if the two strands are not "stabilized" by being subjected to RdRP activity.

4.2.5 RdRP Activity Allows Cells to Produce siRNA from Single Stranded RNA

Since previous data showed that RdRP activity is used to convert strands of the Heteroduplex dsRNA into their reverse complement counterparts, which are then subjected to Dicer cleavage, the question arises of whether single stranded RNA present in food can be subjected to the same mechanism.

Therefore, *E. coli* bacteria expressing part of the *paramecium nd169* gene sequence as a single stranded RNA in either antisense or sense orientation were fed to *paramecia*, and smallRNAs produced by those cells were sequenced (Figure 11).

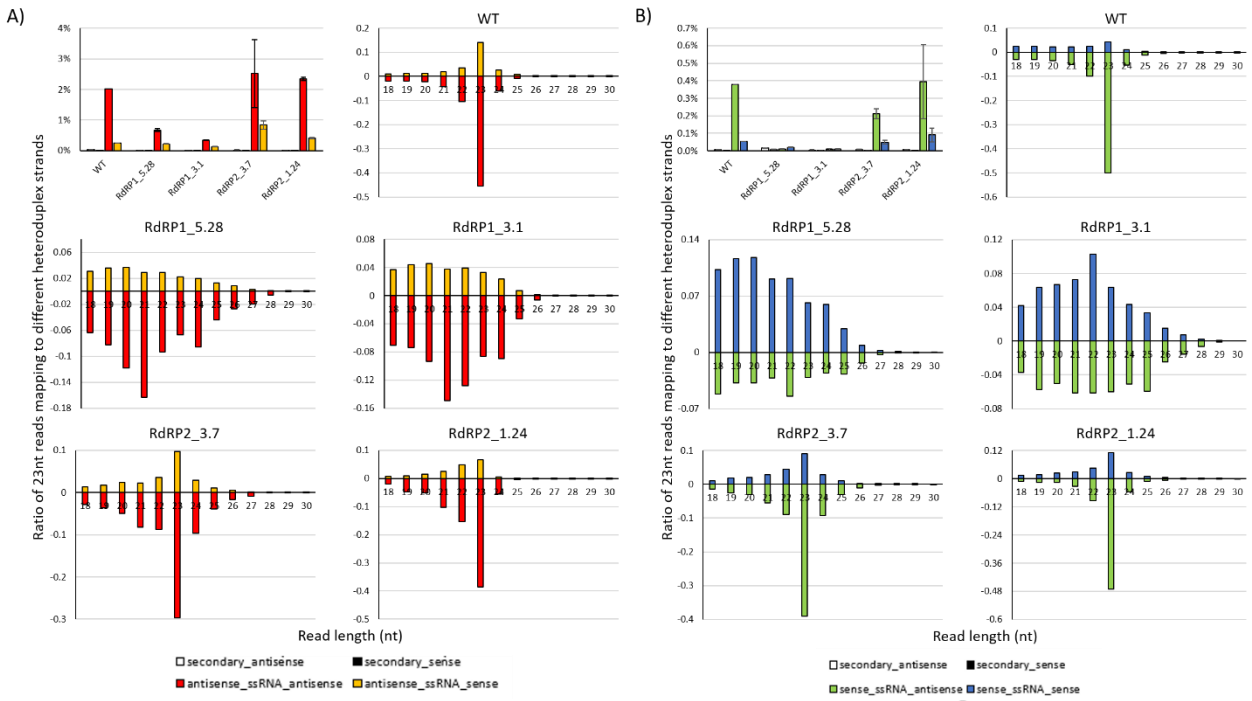

Figure 11 smallRNA species produced from single stranded RNA feeding. Shown are smallRNAs sequenced from paramecia cells fed with either antisense (A) or sense (B) orientated RNA sharing sequence homology with the nd169 gene. Displayed is abundance of smallRNAs derived from the single stranded RNA (upper left) as well as read length distributions of smallRNAs produced from wildtype cells or different RdRP mutants.

While smallRNA from the antisense-orientated ssRNA show generally higher abundance than smallRNAs from sense-orientated ssRNA, it can be observed that all smallRNA species, derived from the original ssRNA or the corresponding RdRP product, show lower abundance in the two RdRP1 mutants compared to the wildtype sample (Figure 11). However, both mutant strains of RdRP2 show smallRNA levels that equal or even slightly exceed wildtype levels, contrary to the observations made in the dsRNA Heteroduplex feeding. A similar observation is true regarding the smallRNA quality produced in the different strains, showing a loss of the 23nt siRNA peak in the two RdRP1 mutant strains, while maintaining the 23nt siRNA peak in both, wildtype and RdRP2 mutant samples. However, functionality of those siRNAs was not tested since no phenotypical analysis of the cells used in this experiment was performed, meaning that it is not clear whether smallRNAs produced by the ssRNAs are capable of establishing PTGS.

The results show that in general, *Paramecium tetraurelia* seems to be capable of producing smallRNA from single stranded RNA in food bacteria, something that was claimed to not work in past articles (Galvani & Sperling, 2002). However, past experiments only analyzed the phenotype caused by the produced siRNAs, not detecting the produced siRNAs themselves.

Interestingly, 23nt smallRNAs produced from the overexpressed ssRNA show dependency on RdRP1, but not on RdRP2, unlike smallRNAs that are produced from the double stranded Heteroduplex. Due to this difference, it is not sure, whether ssRNA is processed in a separate pathway, which would explain the dependency of RdRP1 but not RdRP2 compared to dsRNA, or whether ssRNA is processed in the same pathway described for dsRNA, but RdRP2's function is specific for dsRNA, making it not necessary for the processing of ssRNA.

Additionally, previous experiments in *Paramecium tetraurelia* have shown that 23nt antisense siRNAs are produced against the ribosomal RNA of food bacteria as well as against some bacterial transcripts, however, only in low levels for the latter (Carradec et al., 2015). In those experiments, ribosomal RNA derived siRNAs showed no dependencies on RNAi components involved in the dsRNA pathway, whereas a dependency on RdRP1 and RdRP2 for bacterial transcripts related smallRNA was shown, which implies that multiple pathways for the processing of food-related RNAs exist.

4.3 Discussion

4.3.1 RdRPs in *Paramecium tetraurelia* Act on Exogenous RNAs

As mentioned in the beginning of this chapter, RdRPs are involved in many different RNAi related pathways, triggered by either endogenous RNAs, like genome encoded aberrant transcripts or transgenes, or exogenous dsRNA (Chang et al., 2012; Feng et al., 2024; Pak et al., 2012; Sapetschnig et al., 2015).

For the endogenously triggered pathways, RdRPs usually convert a single stranded trigger RNA into a double strand, making it accessible for Dicer enzymes and therefore starting the production of siRNAs, which was described for many different organisms, including *Neurospora crassa, Arabidopsis thaliana,* and *Nicotiana benthamiana* (Chang et al., 2012; Dadami et al., 2013; Luo & Chen, 2007). Most of those trigger RNAs are produced by aberrant transcripts, either derived by a gene or introduced transgene, a phenomenon also described in *Paramecium tetraurelia*, where RdRP2 and RdRP3 are responsible for the production of a dsRNA molecule (Götz et al., 2016)

For the exogenously triggered pathways, a dsRNA is introduced into the organism, either by feeding, injection or other means, which triggers the production of primary siRNAs and can, depending on the organism, lead to the production of secondary siRNAs by transitivity, using the targeted mRNA as a template, as described, for example, in *C. elegans, Nicotiana*

benthamiana and the plant pathogen *Fusarium asiatiucm* (Pak et al., 2012; Petersen & Albrechtsen, 2005; Song et al., 2018).

Looking at all individual examples of induced RNAi, focusing on the specific roles RdRPs play in those pathways, it can be noticed that RdRPs usually act upon RNAs that have a cellular origin, independent of whether those RNAs represent the trigger RNA (like endogenously produced RNA from transgenes) or the template for secondary siRNA production (like endogenous mRNA transcripts). In none of these cases are exogenous RNAs used as RdRP templates.

Regarding this context of RdRP functions in other organisms, the presented data makes the involvement of RdRP1 and RdRP2 in the RNAi by Feeding pathway in *Paramecium tetraurelia* a novum, since it has been shown that the two RdRPs probably use the exogenous dsRNA trigger as a substrate before the first Dicer cleavage occurs.

The only other report of eukaryotic RdRP activity on exogenous single stranded RNA was conducted in wheat germs at the very beginning of RNAi research (Tang et al., 2003). However, this RdRP activity on exogenous RNAs was measured by incubating ssRNA in wheat germ extract, making those findings not comparable to natural processes in living cells. Apart from that, according to current knowledge, no other eukaryotic RdRP activity in living cells on exogenous RNA has been described, making the functions of RdRP1 and RdRP2 concluded from the Heteroduplex feeding in *paramecium* unique among eukaryotes.

4.3.2 Two RdRPs are Necessary for Primary siRNA Production

As suspected in previous studies (Carradec et al., 2015) and verified by the Heteroduplex feeding, both RdRPs, RdRP1 and RdRP2, are necessary for the production of primary siRNAs. Additionally, the Heteroduplex feeding data revealed that both strands of the trigger dsRNA are subjected to RdRP activity and both RdRPs are required for the processing of either original Heteroduplex strand. This refutes the hypothesis that each RdRP is responsible for processing only one trigger strand, while the other RdRP processes the other.

In reality, both RdRPs have to be present to convert each of the original Heteroduplex strands into their corresponding RdRP product, which can only then be processed by Dicer1 into primary siRNAs. This suggests that RdRP1 and RdRP2 might work together as a complex, or that they have functions that depend on each other.

RdRPs working together as a complex is not described in eukaryotes. However, there are some examples of viral RdRPs that have been reported to work together as a replication complex.

As one such example, the three non-structural proteins nsp7, nsp8 and nsp12 in SARS-CoV can be named. Here the protein nsp12 has been reported to be a functional RdRP (Ahn et al., 2012; te Velthuis et al., 2010) as it is the case for the complex formed by nsp7 and nsp8, which also displays RdRP activity (te Velthuis et al., 2012). In a later study, it has been shown that those two RdRPs form a complex together (Kirchdoerfer & Ward, 2019), showing the possibility of RdRPs working together.

Another possible reason for the requirement of two RdRPs might be specific properties these two RdRPs might have, that make two RdRP necessary for proper function. For example, two different modes of operation have been described for RdRPs, a primer-dependent mode of operation and a primer-independent one. As the name suggests, primer-dependent RdRPs need a short RNA primer bound to the template RNA to be able to start synthesis (Ferrer-Orta et al., 2004; te Velthuis et al., 2010), while primer-independent RdRPs can start their synthesis *de novo* (Govind & Savithri, 2010; Oh et al., 1999). It might be possible that one of the two RdRPs involved in the *Paramecium tetraurelia* pathway might be the one responsible for the synthesis of both RdRP produced strands. This RdRP might operate in a primer-dependent mode, with the needed primer being synthesized by the other RdRP in a primer-independent manner.

As both possibilities would explain the need for two distinct RdRPs involved in the same synthesis step, neither hypothesis can be verified using the present data. Further experiment in the form of interaction tests, co-immunoprecipitations or *in vitro* studies might shed light on whether one of the two hypotheses might be true.

4.3.3 The Importance of the RdRP C-Terminus

As observed for the different RdRP mutants, the two mutant strains resulting in shortened C-terminal amino acid sequences showed a severe impact on RdRP activity and loss of RdRP products even though the RdRP domain of the protein was unchanged and should therefore display catalytical activity.

For several viral RdRPs, studies suggest that the C-terminal amino acids, without being recognized as a conserved domain, have a profound impact on the activity of the RdRP. For example, it has been reported that RdRPs from the hepatitis c virus lacking parts of the C-terminus show an increase in activity due to C-terminal amino acids occupying the active site

cleft, which impairs RNA binding (Adachi et al., 2002; Lévêque et al., 2003). On the other hand, viral RdRPs from norwalk virus and mengovirus show a decrease in activity if the C-terminal amino acids are changed, speculating that the C-terminal amino acids reaching into the active site cleft might play a role in RNA synthesis initiation or interaction between two structural domains, the thumb and palm domains, of the polymerase (Dmitrieva et al., 2007; Ng et al., 2004).

Taking this into account, it is highly likely that *paramecium*'s two RdRPs might function in a similar way to the virus RdRPs, having the C-terminal amino acids interacting in some way with the catalytical active cleft, promoting their function. Structural analysis of the two proteins, either by structure prediction or by crystallization and solving the structure might give some further insight into the relevance of the C-terminal amino acids.

5 Description of Different Ptiwi Proteins and Their Function in Endogenous and Exogenous RNAi

5.1 Background

Argonaute proteins are the lynchpin of many different RNAi pathways, since their main role is the loading of small regulatory RNAs, the identification of the target molecule, and the initiation of the silencing mode, by either slicing of the target mRNA or the establishment of specific chromatin marks.

Argonautes can be separated in two different groups, Agos and Piwis, both of which are mainly distinguished by their mode of operation and their specific function in the cell. Here Piwis are the argonautes interacting with piRNA and silencing transposable elements in the germline, while Agos interact with siRNAs and miRNAs and regulate gene expression.

The organism *Paramecium tetraurelia* does not possesses any Ago proteins, but a total of 15 Piwi proteins can be found in its genome taking on the function of Agos. Multiple whole genome duplications during the evolvement of the *aurelia* clade are responsible for such a high number of different Piwis (Bouhouche et al., 2011), due to which many of the Piwis are ohnologs of each other.

Those 15 Piwis, called Ptiwis for *Paramecium tetraurelia* Piwi, can be separated in roughly two groups, Ptiwis expressed during the sexual reproduction of the cell, and Ptiwis expressed during the vegetative growth. Ptiwis involved in the complicated genome rearrangement processes during sexual reproduction and the smallRNAs associated to them have been extensively studied in the last couple of years (Bouhouche et al., 2011; Furrer et al., 2017; Singh et al., 2022; Solberg et al., 2023). However, Ptiwis involved in different vegetative RNAi pathways are not well understood.

In 2011, the most information available for vegetative Ptiwis were their association to two different RNAi pathways, Ptiwi12, Ptiwi13 and Ptiwi15 being associated with the RNAi by Feeding pathway, and Ptiwi13 and Ptiwi14 being associated with the transgene induced RNAi pathway (Bouhouche et al., 2011). After that, the function of Ptiwi13 and Ptiwi14 in the transgene induced RNAi pathway has been elucidated further and association to the endogenous gene regulation has been revealed (Drews et al., 2021; Götz et al., 2016), but Ptiwi12 and Ptiwi15 and their functions in the RNAi by Feeding mechanism have remained unstudied so far.

For all vegetative pathways, the question arises, why several Ptiwi proteins are needed, suggesting that the different Ptiwis might interact with subclasses of smallRNAs produced. While no information is present for the RNAi by Feeding pathway, it was revealed that Ptiwi13 and Ptiwi14 in the transgene induced RNAi pathways associate with similar classes of smallRNAs (Drews et al., 2021), even though it was speculated, that one Ptiwi might exclusively load primary siRNAs directly produced from the trigger RNA, while the other Ptiwi might exclusively interact with secondary siRNA (Drews et al., 2021). Apart from their interaction with smallRNAs, different functions are postulated for Ptiwi13 and Ptiwi14, due to their differences in cellular localization. While Ptiwi13 localizes in the cytoplasm of the cell, and is therefore thought to attack mRNA transcripts, Ptiwi14 mainly localizes to the macronucleus, providing it with the possibility to engage in chromatin remodeling (Drews et al., 2021). However, no interaction between Ptiwi14 and chromatin remodeling enzymes have been shown so far, remaining this hypothesis to be proven.

For Ptiwi12, Ptiwi13 and Ptiwi15 and their potential function in RNAi by Feeding, similar differences were proposed. It is thought, that some of the three Ptiwis might only interact with primary siRNAs produced from the dsRNA trigger, while other Ptiwis might specifically interact with secondary siRNAs. Interestingly, Ptiwi12 and Ptiwi15, being ohnologs of each other and therefore closely related, potentially lack slicer activity necessary for RNA cleavage due to an exchange of the catalytic triad from DDH to EDH in the Piwi domain. Due to this alteration, it is speculated that Ptiwi12 and Ptiwi15, interact with secondary siRNAs and mediate RNAi to the nucleus similar to pathways in *C. elegans*, while Ptiwi13 interacts with primary siRNAs and mediates target mRNA cleavage (Burton et al., 2011; Carradec et al., 2015)

Due to the lack of knowledge in the vegetative RNAi pathways, aim of the study was to shed light on the role of the three different Ptiwis, Ptiwi12, Ptiwi13 and Ptiwi15, in the RNAi by Feeding pathway, using the assumed non-involvement of Ptiwi14 in this pathway as a control. Additionally, smallRNAs associated to the four Ptiwi proteins during vegetative growth should be analyzed, to gather further insight into potential involvement of the four vegetative Ptiwis in the endogenous regulation of gene expression.

5.2 Results

5.2.1 Ptiwi12 and Ptiwi15 are Located in the Cytoplasm

Ptiwi12 and Ptiwi15 represent two of *paramecium*'s total of 15 Ptiwi proteins and have not been characterized yet. However, since they have been reported to be involved in in the RNAi by Feeding pathway (Bouhouche et al., 2011; Carradec et al., 2015; Marker et al., 2014), they seem to execute an important function for the cell, making a characterization of the two Ptiwis worthwhile. Due to their dissimilarity to other already characterized Ptiwis, like Ptiwi13, Ptiwi14 and Ptiwi08, highlighted by phylogenetic tree analysis (Figure 12A), *de novo* acquisition of data about Ptiwi12 and Ptiwi15 for their characterization was required.

Usually, argonaut proteins, of which the Ptiwi proteins in *Paramecium tetraurelia* are a part of, can be distinguished by their mode of operation, either working in PTGS or CTGS. Since the targets of their silencing effect is located at different parts of the cells, mRNA as a target of PTGS usually in the cytoplasm and DNA as target for CTGS triggered chromatin modifications in the macronucleus, solving the localization of the two so far uncharacterized Ptiwi proteins Ptiwi12 and Ptiwi15 can give a hint, whether they are involved in PTGS or CTGS.

For that, localization of both Ptiwis by overexpressing GFP-fused variants in different cell lines was performed (Figure 12C/D). Western Blot analysis of the transgenic cell lines used for Ptiwi-GFP overexpression validated the successful expression of the full-length fusion proteins prior to IFA (Figure 12B).

As displayed, both Ptiwi proteins show a broad distribution in the cytoplasm of the cell, clearly omitting the macronucleus, which is visualized by DAPI staining. Following this observation, the localization of the two Ptiwis might suggest, that their main function might be an PTGS resembling mode of action, using mRNA as target.

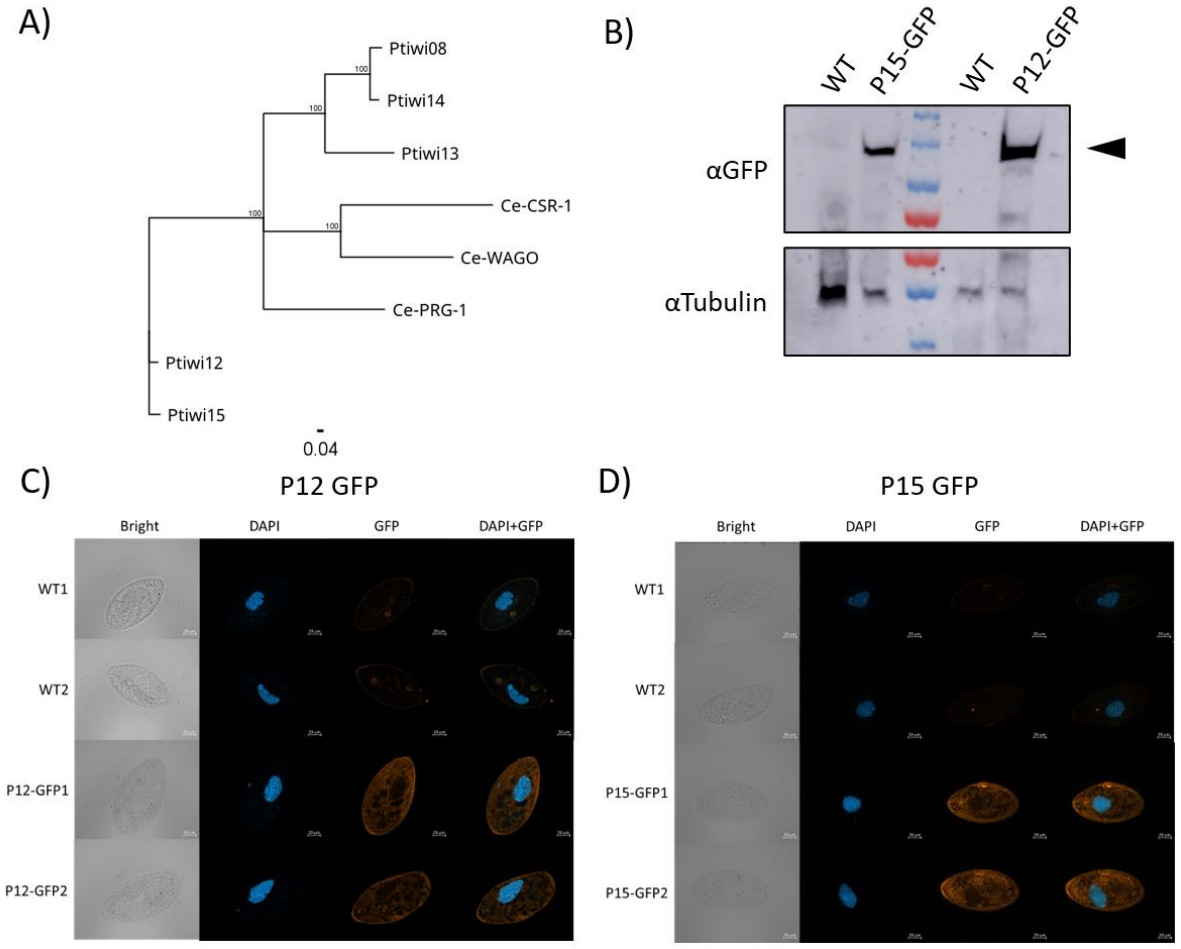

Figure 12 Localization of Ptiwi12 and Ptiwi15. Phylogenetic tree of Ptiwi12 and Ptiwi15 compared to *Paramecium tetraurelia*'s Ptiwi08/Ptiwi13 and Ptiwi14 as well as the argonauts CSR-1/PRG-1 and WAGO from *C. elegans*. Displayed is a neighbor joining consensus tree based on ClustalX aligned sequences. Bootstrap replicates are indicated at the different branches of the tree. B) Western Blot analysis of total protein from Ptiwi12 (P12) and Ptiwi15 (P15) -GFP fusion protein overexpressing cell lines compared to wildtype cells. Displayed are signals gathered by incubation with anti-GFP antibody (αGFP) as well as anti-alphaTubulin-antibody (αTubulin) as loading control. Expected signals from Ptiwi-GFP fusion proteins (approx. 115 kDa) are indicated by the black arrowhead. As marker, the color prestained protein standard, broad range (by NEB) was used. The red band corresponds to a size of 72 kDa. C/D) Localization of Ptiwi12 (C) and Ptiwi15 (D) fused to a GFP tag and visualization of the localization using an Alexa594-coupled secondary antibody. Wildtype cells served as a control for unspecific signals caused by the antibodies used. Macronuclei of cells were visualized by DAPI staining. Exposure time of all displayed channels were kept the same for all connected samples to ensure comparability (GFP: 950ms; DAPI: 50ms; Brightfield: .45ms). Images were taken using a fluorescent microscope (Axio-Observer, Zeiss) and structured illumination (Apotome2; Zeiss)

5.2.2 Different Ptiwi Proteins Load smallRNAs of Various Length from Different Endogenous Loci

To analyze the specific function of the four vegetative Ptiwis in *paramecium*, cell lines expressing the Ptiwi proteins Ptiwi12, Ptiwi13, Ptiwi14 and Ptiwi15 fused to a FLAG amino acid sequence were created and smallRNAs loaded by the different Ptiwi proteins were sequenced by utilizing the RNA-IP technique (Figure S 4). Before analyzing the association of the Ptiwi proteins to RNAi by Feeding derived smallRNAs, the first description of Ptiwi-bound smallRNAs was carried out with untreated cells, to decrease complexity of the initial dataset and focus on endogenously derived smallRNAs. To get a first overview of smallRNAs bound to the Ptiwi proteins, smallRNA sequencing data was mapped on different templates, representing various origins of smallRNAs that might be present within the cell, and visualized in dependency on the read length (Figure 13).

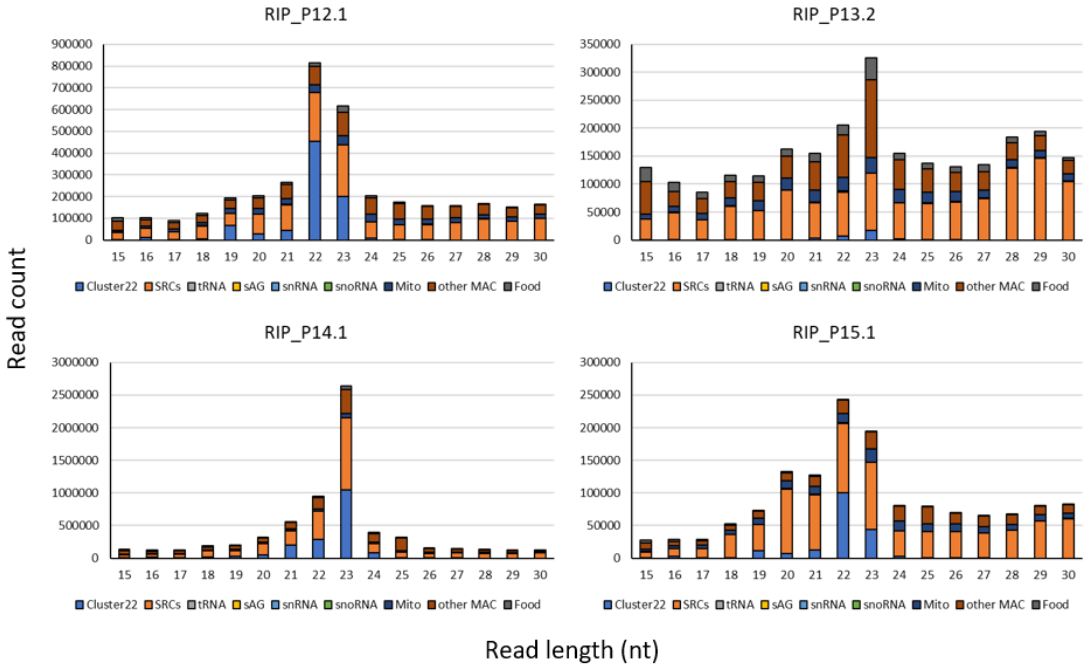

Figure 13 Read length of Ptiwi-bound smallRNA from different endogenous loci. Shown are the read counts of smallRNA sequencing data of each read length (in nucleotides) that is derived from different origins, as indicated by the color code displayed. Reads were generated from RNA immunoprecipitation data (RIP) from Ptiwi 12 (P12), Ptiwi13 (P13), Ptiwi14 (P14) and Ptiwi15 (P15). For Ptiwi12/13/15, data from only one available replicate is shown to enhance readability of the figure.

As observed in Figure 13, both, Ptiwi13 and Ptiwi14, show a read length distribution peak at 23nt size, a size canonically regarded as siRNAs in *paramecium*. However, Ptiwi12 and Ptiwi15 show an enrichment of 22nt smallRNA molecules in addition to the 23nt peak, suggesting that

Ptiwi12 and Ptiwi15 not only associate with canonical siRNAs, but also select for a 22nt smallRNA subgroup, which was so far not observed for *paramecium*.

Additionally, Ptiwi12, Ptiwi14 and Ptiwi15 associate with smallRNAs that are mainly derived from siRNA producing clusters (SRCs), that have been described previously for *paramecium* (Karunanithi, Oruganti, et al., 2019). These clusters include the cluster labeled as Cluster22 in the figure, which has been visualized separately due to its high abundance in sequencing data. However, even though SRC related reads are present in Ptiwi13, the majority of reads found in Ptiwi13 RIP data originate from other macronuclear associated sequences, suggesting that Ptiwi13 loads different smallRNAs compared to the other three Ptiwi proteins. Additionally. Ptiwi13 shows the largest ratio of reads mapping to the food bacteria genome.

5.2.3 Ptiwi Proteins Associate with Different SCRs

Due to the significance of 23nt reads in *paramecium* as well as the high abundance of 22nt reads in Ptiwi12 and Ptiwi15 data, all further analysis not displaying sequencing data of different length were carried out using only 23nt long reads for Ptiwi13 and Ptiwi14 as well as 22nt and 23nt long reads for Ptiwi12 and Ptiwi15.

To get further insight into the role of the different Ptiwis in the endogenously triggered RNAi pathways, the SRCs that are loaded into the different Ptiwi proteins were identified (Table S 2). SRCs were regarded as “loaded into a Ptiwi protein”, if *de novo* cluster prediction using RIP data from the Ptiwi of interest resulted in the prediction of a SRC that overlaps with one of the already described SCRs in *paramecium* (Karunanithi, Oruganti, et al., 2019; Solberg et al., 2023) (Figure 14).

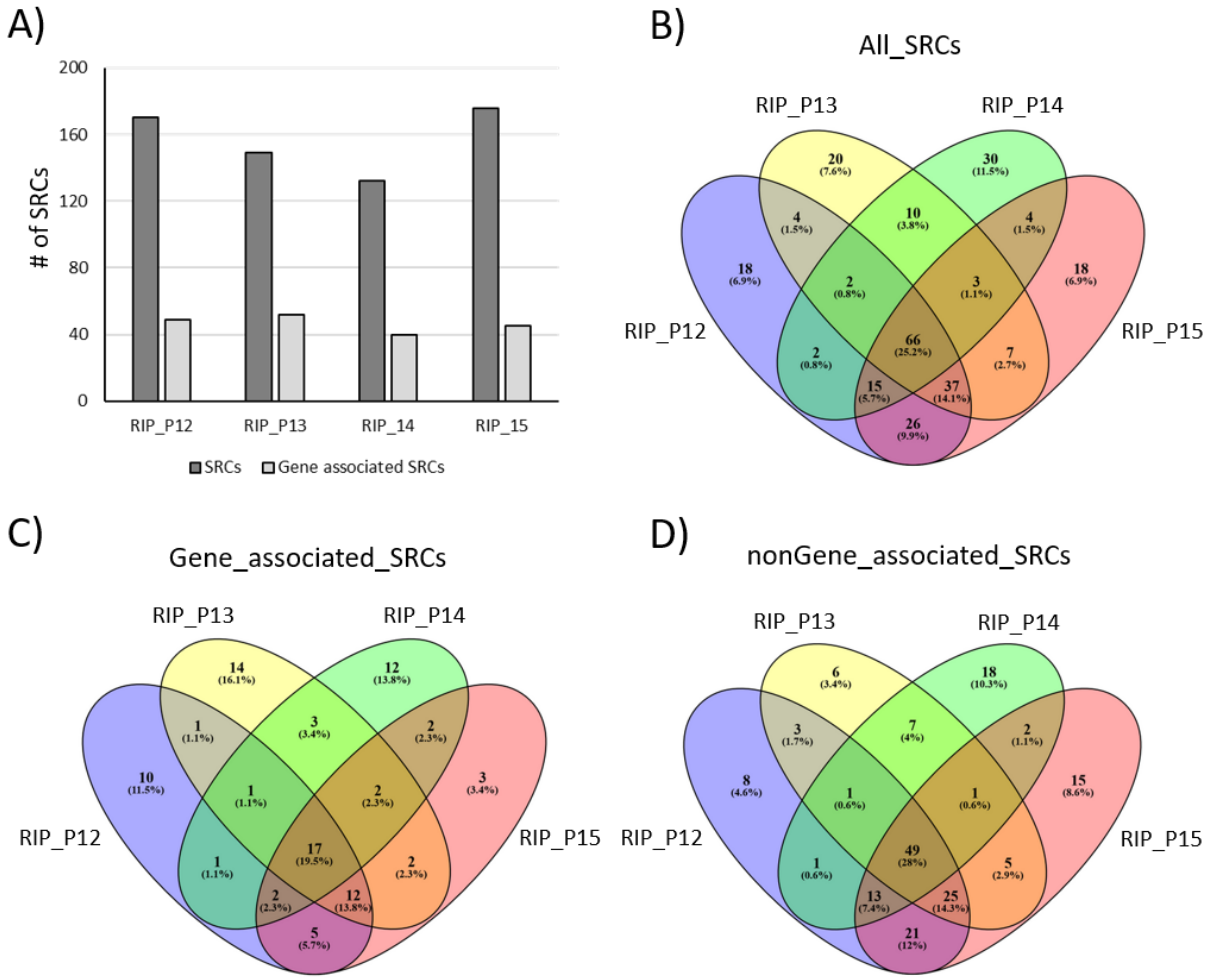

Figure 14 Overview of SRCs loaded into different Ptiwi proteins. A) Number of smallRNA producing clusters considered loaded into the corresponding Ptiwi that were predicted by using the pooled reads of two RIP replicates (with the exception of P14, which represents only one replicate) and overlapped with known SRCs (dark grey bars) or gene associated SRCs (light grey bars). B-D) Venn diagram of loaded SRCs between the different Ptiwi protein (Ptiwi12 in blue, Ptiwi13 in yellow, Ptiwi14 in green and Ptiwi15 in red) with focus on all SRCs (B), only gene associated SRCs (C), or non-gene associated SRCs (D).

For all four Ptiwi proteins, smallRNA from approx. 150 already described SRCs were found, actual numbers for individual Ptiwi proteins only deviating slightly from each other. Of those SRCs, approx. 40 overlapped with genes (Figure 14A).

To see whether all Ptiwis associate with the same SRC derived smallRNAs, Venn diagrams of the loaded SRCs from all Ptiwis were constructed (Figure 14B-D). As shown, even though Ptiwis had some overlap regarding loaded SRCs, there are SRCs which are uniquely loaded in individual Ptiwis.

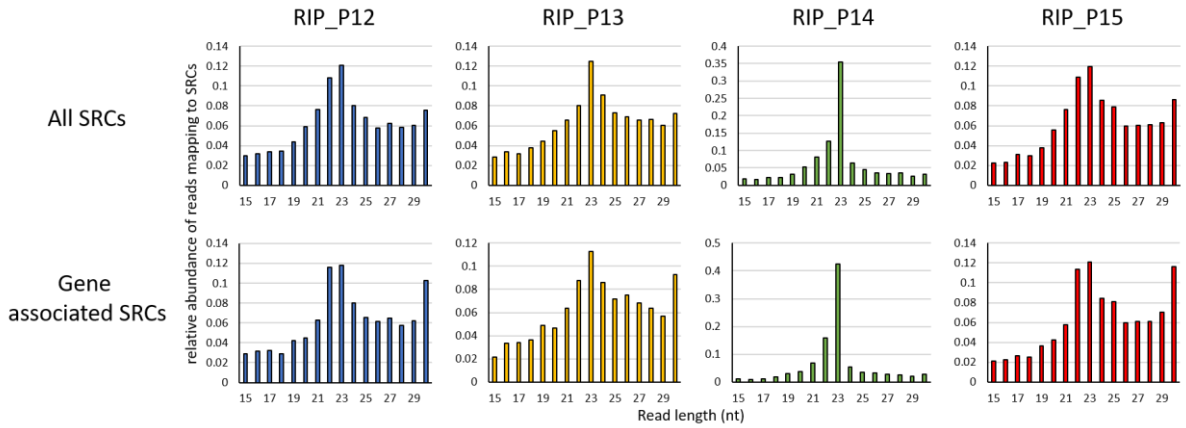

Figure 15 Mean read length distribution of smallRNAs derived from SRCs loaded into different Ptiwi proteins. Shown is the mean read length distribution of the dominant strand across two RIP replicates (except RIP_P14, which represents only one replicate) for either all SRCs (top) or only gene associated SRCs (bottom) that are considered loaded into the corresponding Ptiwi. Distributions were calculated for Ptiwi12 (P12), Ptiwi13 (P13), Ptiwi14 (P14) and Ptiwi15 (P15). Only SRCs displaying a clear strand bias of either >80% or <20% antisense bias were included in this analysis. Number of SRCs used for each replicate: RIP_P12_all: 153; RIP_P12_gene_associated: 41; RIP_P13_all: 125; RIP_P13_gene_associated: 44; RIP_P14_all: 98; RIP_P14_gene_associated: 31; RIP_P15_all: 150; RIP_P15_gene_associated: 33.

Analyzing the mean read length distribution of loaded SRCs in each of the RIP datasets, it can be observed that only Ptiwi14 displays a clear 23nt siRNA peak. For all other Ptiwi proteins, the read length distribution shows many smaller and longer fragments, making the 23nt siRNA peak not as prominent compared to Ptiwi14 data. Additionally, Ptiwi12 and Ptiwi15 display the enrichment of 22nt smallRNAs already described prior. This enrichment for 22nt smallRNAs gets more prominent if focused on the subgroup of gene associated SRCs, suggesting that this 22nt subgroup has an increased importance for SRCs that overlap with coding genes.

Again, read length distribution highlights the affinity of the different Ptiwi protein to smallRNAs with individual characteristics, which might suggest different functions in the grand scheme of endogenously trigger RNAi pathways.

5.2.4 Ptiwi14 Loads smallRNAs without Strand Specificity and Associate with Unspliced mRNA

To establish a gene silencing, argonaut proteins have to associate with smallRNAs in antisense orientation to the target gene to allow the smallRNA to bind to the sense orientated mRNA and mediate the silencing effect. To analyze, whether Ptiwi bound smallRNA can negatively regulate gene expression, the antisense ratio of smallRNAs from gene associated SRCs was calculated (Figure 16A)

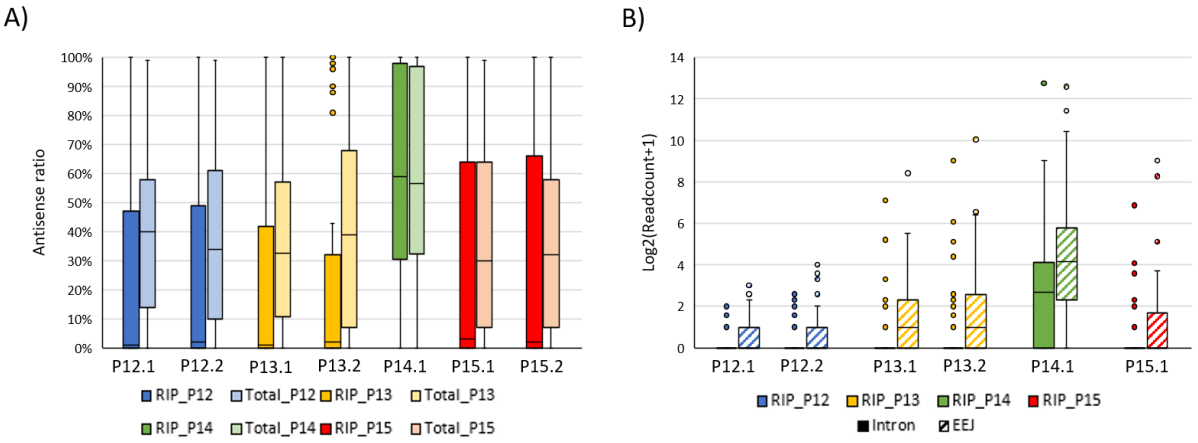

Figure 16 Properties of smallRNAs of gene associated SRCs. A) Antisense ratio across gene associated SRCs identified as loaded into the corresponding Ptiwi protein calculated from RIP data (RIP) and total RNA data (Total). Considered are antisense ratios for each replicate for Ptiwi12 (P12), Ptiwi13 (P13), Ptiwi14 (P14) and Ptiwi15 (P15). To ensure valid calculation of antisense ratio, SRCs with lower than 20 reads in either, the RIP or corresponding total RNA sample, were discarded from the analysis. Number of SRCs used for each replicate: P12.1: 47; P12.2: 47; P13.1: 50; P13.2: 51; P14.1: 40; P15.1: 43; P15.2: 43. B) Read count of reads mapping to either, Exon-Exon-Junctions (EEJ) (full bars) or the corresponding intron (striped bars) of gene associated SRCs, converted to the appropriate log2(Readcount+1) value taken from RIP data. Considered are SRCs that identified as loaded into the corresponding Ptiwi protein. Number of SRCs used for each replicate: P12.1: 40; P12.2: 40; P13.1: 44; P13.2: 44; P14.1: 39; P15.1: 43; P15.2: not enough mapping sequencing reads for this analysis.

Analyzing the antisense ratio of smallRNAs across different gene associated SRCs, it can be observed that Ptiwi12, Ptiwi13 and Ptiwi15 load predominantly sense oriented smallRNAs, with the ratio of antisense orientated smallRNAs being as low as approx. 5% in median across all considered SRCs. These antisense ratios differ largely from antisense ratios that can be overserved in total RNA sequencing from the corresponding Ptiwi-overexpressing cell lines, suggesting that the three Ptiwis don't select for antisense orientated smallRNAs, but for sense orientated smallRNAs.

Only Ptiwi14 shows an elevated antisense ratio compared to the other Ptiwis, loading approx. equal amounts of sense and antisense orientated siRNAs, with a median antisense ratio of 59%,

coinciding to the median antisense ratio of the corresponding total RNA control, which suggests a lack of strand preference or stabilization for Ptiwi14.

Taking prior knowledge about Ptiwi13 and Ptiwi14 into account, Ptiwi14 is the only Ptiwi protein analyzed in the present experiment, that doesn't show cytoplasmatic, but a nuclear localization (Drews et al., 2021). To account for the possibly different smallRNA pool that might be accessible to the different Ptiwis, gene associated SRCs were analyzed in regards of the presence of reads overlapping the exon-exon junction (EEJ) of spliced mRNA or mapping to the intron sequence of the annotated genes (Figure 16B). A higher number of siRNA reads associated to intron sequences would suggest an interaction of the Ptiwi proteins with unspliced mRNA, possibly in the macronucleus, while an enrichment of EEJ mapping reads would suggest association with spliced mRNA, probably in the cytoplasm.

For Ptiwi12, Ptiwi13 and Ptiwi15, no sample shows meaningful amounts of intron mapping reads, having the large majority of reads mapping to the EEJ. However, in comparison to the other Ptiwi RIP data, Ptiwi14 shows significantly higher amounts of intron mapping reads, underlining the differences between Ptiwi14 and the other three Ptiwis. Observing the higher amounts of intron mapping reads in the Ptiwi14 RIP sample, it can be speculated that Ptiwi14 either targets mRNA transcripts that are not completely spliced yet, and/or that Ptiwi14 loads siRNAs that are produced in the macronucleus from unspliced mRNA. Both possibilities fit to the observed localization of Ptiwi14 in the macronucleus.

Taking these observations together, siRNAs derived from gene associated SRCs and loaded into Ptiwi14 have different properties compared to siRNAs loaded by the other Ptiwi proteins. Additionally, the antisense ratio of Ptiwi12, Ptiwi13 and Ptiwi15 associated SRCs suggests, that mRNA targeting and silencing doesn't seem to be the main function of these three Ptiwi proteins, since sense orientated smallRNAs should not be able to interact with mRNA by base pair binding.

5.2.5 Level of smallRNAs and mRNA Do Not Correlate in Ptiwi Knock Down

Since analysis of the antisense ratio did not yield any hint of Ptiwi proteins being involved in RNA mediated silencing of genes, a knockdown (KD) of individual Ptiwi proteins was conducted to see, whether knockdown of the Ptiwi protein would result in changes regarding smallRNA and mRNA abundance. For Ptiwi12 and Ptiwi15, due to the high nucleotide sequence similarity between the two Ptiwi genes, a co-silencing was conducted.

For this, smallRNA and mRNA of wildtype and Ptiwi knockdown cells were sequenced, and the fold change of gene associated RNA levels between wildtype and knockdown cells were calculated. If smallRNAs influenced by one of the Ptiwi proteins would interfere with corresponding mRNA expression in any way, a loss of the Ptiwi protein should result in changed level of the smallRNA and therefore, changed level of the regulated corresponding mRNA as well, which should result in a correlation between smallRNA foldchange and mRNA foldchange (Figure 17).

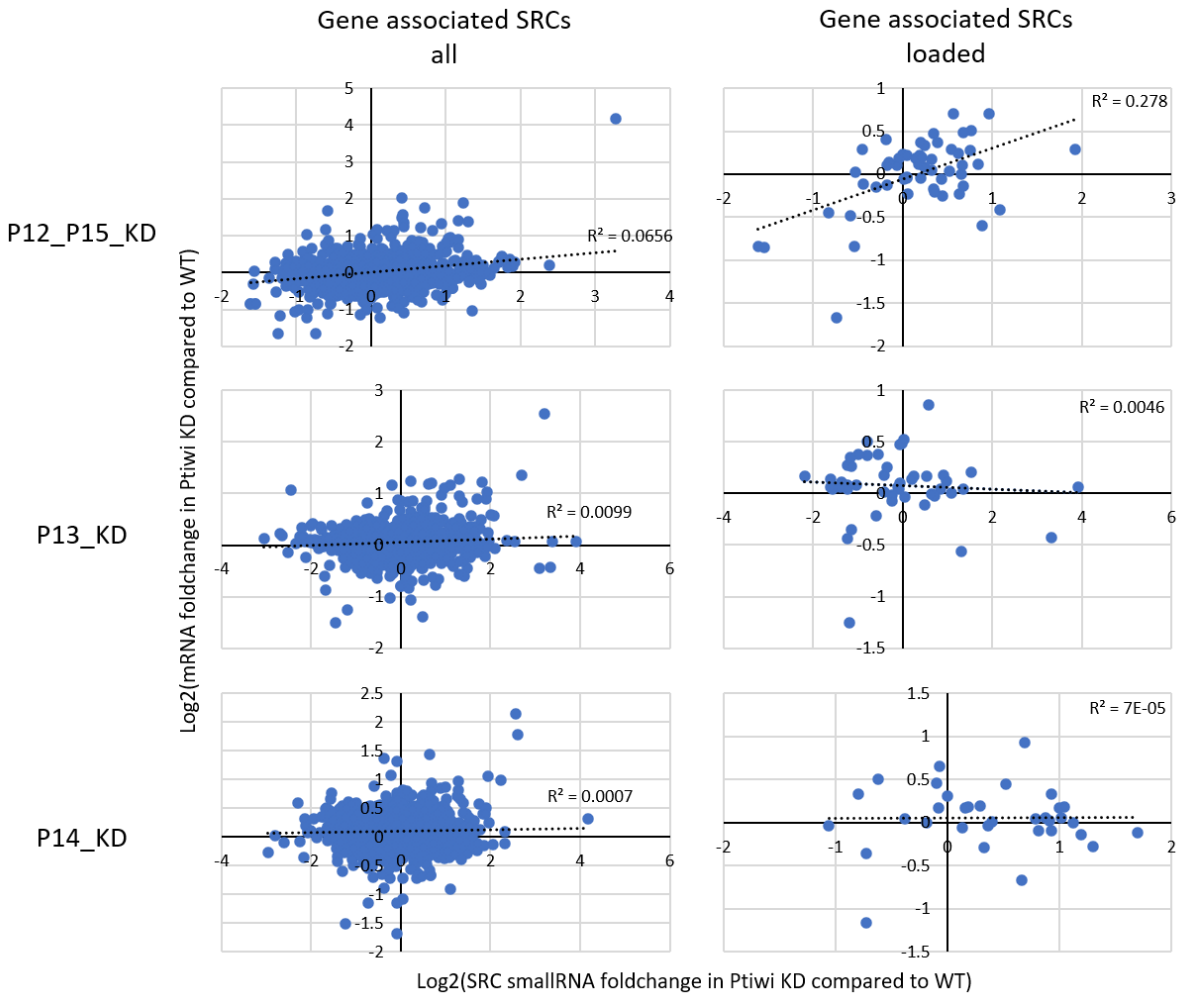

Figure 17 Correlation between smallRNA and mRNA foldchange in Ptiwi knockdown. Plotted are smallRNA foldchanges (log2, x-axis) calculated from smallRNA wildtype and Ptiwi KD data against mRNA foldchanges (log2, y-axis) calculated from smallRNA wildtype and Ptiwi KD data. For each plot, either Ptiwi13 (middle row), Ptiwi14 (bottom row) or Ptiwi12 and Ptiwi15 in combination (top row) were knocked down for the experiments. Plots were calculated using either all known gene associated SRCs (left) or a subset of gene associated SRCs that has been identified as loaded into the corresponding Ptiwis (right). SRCs were either, smallRNA foldchanges or mRNA foldchanges could not have been calculated were eliminated from the analysis. Number of SRCs included: P12_15KD_all: 1162; P12_15_KD_loaded: 54; P13_KD_all: 1166; P13_KD_loaded: 48; P14_KD_all: 1162; P13_KD_loaded: 34.

Plotting the calculated foldchanges for smallRNA and mRNA of all known gene associated SRCs against each other, no significant correlation can be observed between smallRNA and mRNA level changes upon silencing of either Ptiwi protein (Figure 17 left). This remains true, even if the analyzed SRCs are narrowed down to SRCs that are considered loaded by the corresponding knock downed Ptiwi, in an attempted to improve correlation by reducing the background noise of potential unregulated genes (Figure 17 right). This increases the R^2 score for SRCs loaded into Ptiwi12 and Ptiwi15 in the Ptiwi12-15 knockdown, however the correlation doesn't reach a level of statistically significance.

Taking this together, it seems like Ptiwi loaded smallRNAs do not regulate gene expression to an extend that can be observed by using sequencing data. However, it cannot be ruled out entirely, that the Ptiwi proteins are involved in gene regulation, since functions like fine-tuning of gene expression would result in very low alterations, that might not be detected by methods used in these experiments.

5.2.6 Primary and Secondary Feeding Associated siRNAs are Loaded into Different Ptiwi Proteins

Utilizing the same transgenic lines used for the analysis of endogenously triggered RNAi, Ptiwi-FLAG fusion protein expressing cells were subjected to dsRNA feeding against the *nd169* gene by applying dsRNA producing *E coli* to the cells. smallRNAs produced by the RNAi by Feeding pathway and loaded into different Ptiwi proteins were analyzed by RIP, to investigate the specific role of the four Ptiwi protein in this pathway (Figure 18).

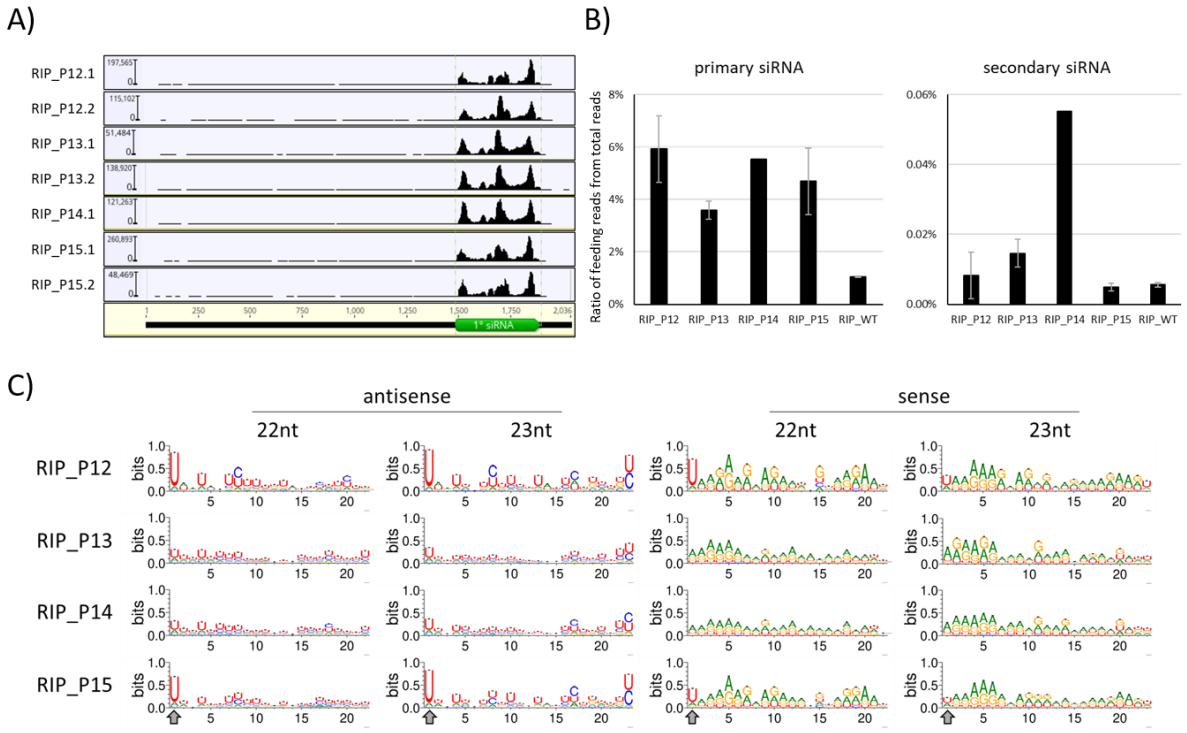

Figure 18 Overview of primary and secondary siRNAs loaded into Ptiwi proteins. A) Coverage plot of smallRNA sequencing reads on the *nd169* gene sequence. Shown are the coverages for replicates of RIP data from Ptiwi12 (P12), Pitwi13 (P13), Ptiwi14 (P14) and Ptiwi15 (P15). The *nd169* sequence corresponding to the dsRNA (and therefore primary siRNAs) is labeled in green. Regions outside the green labeled sequence correspond to secondary siRNAs. B) Mean abundance of primary (left) and secondary (right) siRNA in relation to the total number of reads sequenced for the different Ptiwi RIP datasets as well as a mock RIP performed with cells not expressing the FLAG-tagged version of Ptiwis. (RIP_WT). C) Sequence logos of 22nt and 23nt long primary smallRNA reads present in the RIP data of different Ptiwi IPs. Sequence logos were calculated using the weblogo tool for sense and antisense reads, normalizing the logo for the paramecium genome base composition (A/U: 0.36; G/C: 0.14 each, according to Lepère et al., 2009).

To gain a first impression of smallRNAs produced from the dsRNA trigger molecule, coverage tracks of smallRNAs mapping to the target gene were analyzed (Figure 18A). As displayed, feeding associated smallRNAs can be divided into two categories, primary smallRNAs (primary smallRNAs) which are produced from the dsRNA trigger and secondary smallRNAs (secondary smallRNAs) which are produced from the targeted mRNA. Both subclasses can be distinguished by mapping, primary smallRNAs being mapped to the sequence of the trigger dsRNA (labeled in green, Figure *18*A), while secondary smallRNAs map to regions outside of the primary smallRNA feeding fragment.

For all Ptiwi RIP data, it can be observed that primary smallRNAs represent the majority of the feeding associated RNAs, while secondary smallRNAs only represent a small fraction, barely being visible in the coverage track (Figure 18A).

Quantifying the two feeding-associated smallRNA subclasses in relation to the total reads sequenced, primary smallRNAs represent approx. 5% of all sequenced reads, showing a slightly lower abundance in the Ptiwi13 IP compared to the other analyzed Ptiwi proteins. However, all Ptiwis show a higher abundance of primary smallRNAs compared to the utilized mock RIP control (cells not expressing a tagged Ptiwi) showing that primary smallRNAs seem to be loaded in all four analyzed Ptiwis (Figure 18B).

Looking at the abundance of secondary smallRNAs, it can be noted that Ptiwi14 shows the highest level of secondary smallRNAs by far, with the other three Ptiwi proteins only showing secondary smallRNA level comparable to the wildtype mock RIP control, suggesting that secondary smallRNAs are preferentially and exclusively loaded into Ptiwi14 (Figure 18B).

To further characterize the loaded smallRNAs, a sequence logo of primary smallRNAs for 22nt and 23nt long molecules were calculated, to give further information about sequence preferences for siRNAs (23nt) and 22nt long, Ptiwi12 and Ptiwi15 preferred smallRNAs.

Searching for nucleotide preferences, it can be noted that both, Ptiwi12 and Ptiwi15, show a strong uracil preference for the very first base at the 5’ end of the smallRNA molecule (Figure 18C), which is not present in the sequence logo of the corresponding total RNA sequencing data (Figure S 5). This preference can be observed for antisense orientated smallRNAs of both sizes, 22nt and 23nt, but also sense orientated smallRNA show such preference. However, the preference for the sense orientated smallRNA strands is not as pronounced as for antisense orientated molecules, but due to the general preference of adenine and guanine in the sense strand, the slight uracil preference for the ultimate 5’ nucleotide is noticeable (Figure 18C).

None of the other two Ptiwi proteins, Ptiwi13 and Ptiwi14, show such preferences for either strand orientation or molecule size.

Preferences for the 5' nucleotides in argonaut IPs is usually a sign for specific selection and loading of smallRNAs by the argonaut protein, since many argonaut proteins show specific tendencies, usually for adenine or uracil, at the 5' end of the smallRNA strand that is selected for loading (Frank et al., 2010; Iwakawa & Tomari, 2022). However, having these preferences present in both, antisense and sense strand of the smallRNAs found in the Ptiwi IP, suggests that Ptiwi12 and Ptiwi15 actively load both, antisense and sense orientated smallRNAs derived from the trigger dsRNA.

5.2.7 Ptiwi Loaded Primary siRNAs Show Balanced Strand Bias

Since Ptiwi12 and Ptiwi15 showed 5' nucleotide preferences for sense and antisense orientated smallRNAs, read length distributions and antisense ratios were calculated for both, primary and secondary smallRNAs to see, whether mature 23nt siRNAs of both strands are loaded into the Ptiwi (Figure 19).

Looking at the read length distribution for both, primary and secondary smallRNAs, it can be observed that all Ptiwi proteins show a clear peak at 23nt size (Figure 19A-B). Additionally, the 22nt smallRNA fraction in Ptiwi12 and Ptiwi15 RIPs seem more abundant when focusing on the sense orientated reads compared to the antisense orientated reads, suggesting an enrichment of 22nt long sense orientated smallRNAs in these two Ptiwi proteins.

Calculating the antisense ratio of the feeding associated smallRNAs, it can be observed that all four Ptiwis show almost balanced antisense ratios of approx. 60% for the primary siRNAs. Comparing these antisense ratios to the antisense ratio of total RNAs sequenced from the same samples prior to IP shows that the antisense ratio in the RIP decreases from approx. 90% for all Ptiwis in total RNA, suggesting that all four Ptiwis load both strands, sense and antisense, in almost balanced amounts.

This effect seems to be specific for the primary smallRNAs, since RIP and total RNA sequencing show equal antisense ratios of approx. 95% for secondary smallRNAs.

Taking these results together, Ptiwi proteins seem to select primary and secondary smallRNAs in different manners, which results in distinct antisense ratios for both smallRNA subclasses. However, it can only be speculated, whether this difference is indeed due to selection criteria

that Ptiwi proteins apply to the two smallRNA subclasses, or whether this difference is due to reasons that are caused by the biogenesis of the two subclasses.

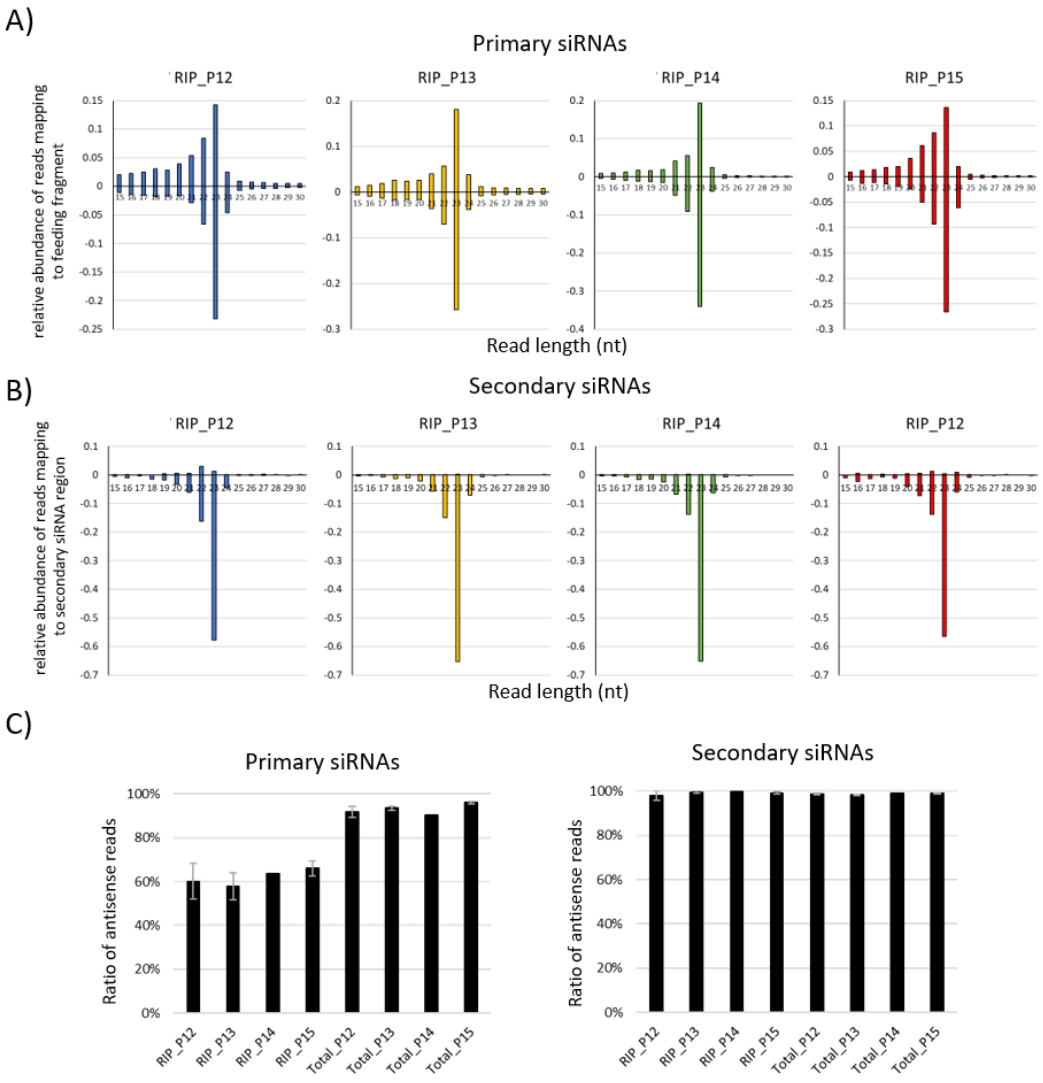

Figure 19 Read length distribution and strand bias of Ptiwi loaded smallRNAs. A/B) Mean read length distribution of primary (A) and secondary (B) smallRNAs mapped to the *nd169* gene sequence in RIP data from Ptiwi12 (P12), Ptiwi13 (P13), Ptiwi14 (P14) and Ptiwi15 (P15). For P12, P13 and P15, two replicates were considered while for P14, only one replicate was available. Positive values correspond to sense-directed molecules whereas negative values correspond to antisense-directed molecules. C) Antisense ratio of primary (left) and secondary (right) 23nt siRNA reads mapping to the *nd169* gene sequence for RIP data as well as total RNA sequencing data for the four mentioned Ptiwis.

5.2.8 Ptiwi12 and Ptiwi15 Associated with Dicer Cut siRNA Duplexes

Usually, argonaut proteins select one strand of the siRNA duplex that has been cut by Dicer, loading the selected strand and degrading the other, which leads to a strand bias of argonaut loaded siRNAs. However, Ptiwi proteins of *paramecium* seem to associate with both strands due to their balanced strand ratios. To see whether *paramecium* Ptiwis associate with both strands of the Dicer duplex at the same time, an overlap analysis of reads found in the RIP dataset was performed. This overlap analysis measures the distance between 5' ends of overlapping smallRNA molecules from opposite strands, reporting whether a particular distance is overrepresented in the data. The overlapping smallRNA molecules with opposite orientation thereby mimic the duplexes created from different smallRNA biogenesis pathways. The distance measured between the two 5' ends can be used as an indicator for biogenesis, since certain distances represent hallmarks for different biogenesis pathways. For instance, focusing on 23nt long siRNAs, an overlap of 21nt length indicates a Dicer dependent biogenesis, because Dicer creates siRNA duplexes with 2nt 3' overhangs, resulting in an overlap of 21nt. On the other hand, smallRNAs derived from a Ping-Pong-like biogenesis pathway show a 10nt overlap, making these two biogenesis pathways distinguishable (Antoniewski, 2014) (Figure 20).

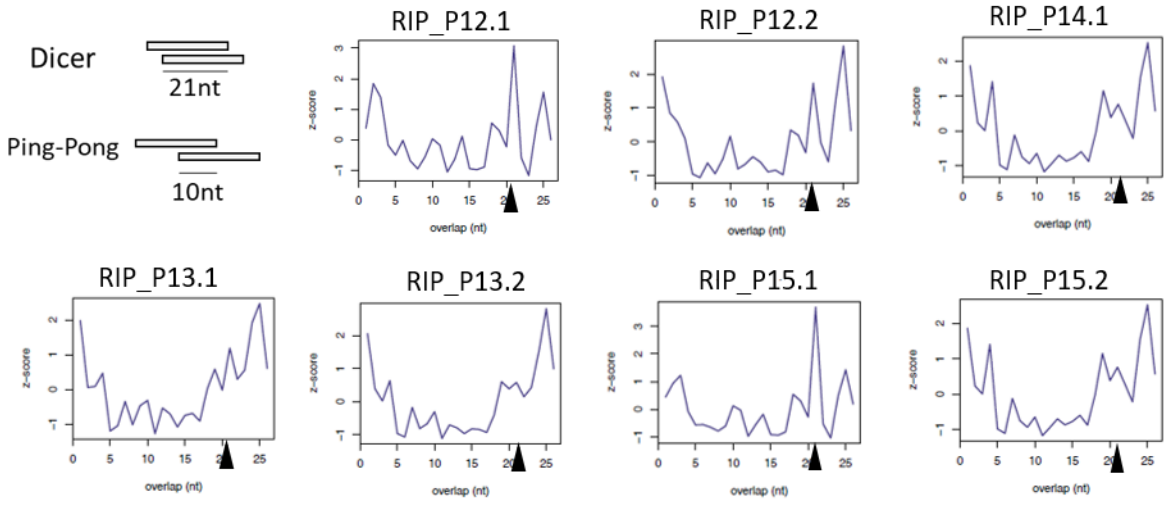

Figure 20 Overlap analysis of Ptiwi loaded 23nt siRNAs. Displayed are the measured distances between the 5' ends of smallRNAs on to opposite strands. An overlap distance of 21 nucleotides (nt) equals to a potential Dicer-dependent smallRNA biogenesis, while an overlap distance of 10nt equals to a Ping-Pong like biogenesis (upper left). SmallRNAs from several RNA IP replicates of the Ptiwis 12,13,14 and 15 were subjected to overlap analysis. The position of the potential Dicer-overlap peak is marked with a black arrowhead.

As shown, for two replicates, RIP_P12.1 and RIP_P15.1, a peak at an overlap distance of 21nt can be observed. This suggests not only a Dicer dependent biogenesis of the 23nt long primary siRNAs, but the fact that these overlaps can be observed in RIP data, suggests that the two Ptiwi proteins closely associate with the Dicer produced duplex, compared to the other two Ptiwi proteins, Ptiwi13 and Ptiwi14. However, the Dicer overlap peak can only be observed in one of the two replicates present for Ptiwi12 and Ptiwi15, suggesting either a rapid turn-over of the Dicer duplex or a low abundance just around the detection limit of the analysis.

5.3 Discussion

5.3.1 Ptiwi Proteins Load smallRNA From a Subset of Selected SRCs

In the described experiments, all of the analyzed Ptiwis loaded smallRNA from only approx. 150 of the described 2610 SRCs in *Paramecium tetraurelia* (Karunanithi, Oruganti, et al., 2019; Solberg et al., 2023). Earlier analysis of the two proteins Ptiwi13 and Ptiwi14 revealed an association of those Ptiwis with almost 1400 SRCs (Drews et al., 2021), highly exceeding the amount of SRCs identified in the current experiments. However, in the previous work, the method used to determine Ptiwi loaded SRCs differed from the method used in this work, which explains the differences in loading behavior. Here loaded SRCs were determined by *de novo* cluster prediction, while in the mentioned study, all SRCs for which enough reads were found to exceed a threshold of 1TPM in Ptiwi IP data were considered loaded by the Ptiwi protein. Experiments performed with Ptiwi08, an ohnolog of Ptiwi14, using the same loading definition as applied here, reported an association of Ptiwi08 with roughly 80 SRCs (Solberg et al., 2023), a number much closer to the results presented in this work.

Looking at the presented data, all four Ptiwis showed some similarities in smallRNA loading, but also differences regarding both, association with certain SRCs and loading of RNAs from different genomic or exogenous origins. Ptiwis associating with not only regulatory siRNAs, but also smallRNAs from other genomic loci have been described in previous studies for Ptiwi13 and Ptiwi14 (Drews et al., 2021) as well as Ptiwi08 (Solberg et al., 2023), but was also reported for different Twi proteins (Piwis in *Tetrahymena thermophila*), where some of the 12 existing Twis associate with smallRNAs, rDNA, mitochondria, pseudogenes, and even tRNA fragments (Couvillion et al., 2009, 2012). Interestingly, in previous studies Ptiwi13 has been associated with smallRNA derived from food bacteria (Drews et al., 2021), which is fitting because Ptiwi13 is involved in the RNAi by Feeding mechanism (Bouhouche et al., 2011),

associating Ptiwi13 with exogenous RNAs in general, which can be confirmed by the results presented in this work.

5.3.2 Ptiwi Associated Endogenous smallRNAs Show Sense Orientation

Analyzing the antisense ratio of Ptiwi bound smallRNAs mapping to SRCs, it has been revealed that most Ptiwis enrich for sense orientated smallRNAs. Sense bias for SRC associated smallRNAs has been described in the initial SRC definition (Karunanithi, Oruganti, et al., 2019) and has been shown for some *Paramecium tetraurelia* genes, most prominent of which is the surface antigen gene family, which are included in the SRCs. However, this high abundance of sense orientated smallRNAs in argonaut bound RNA is unusually, because most pathways associate argonauts with antisense orientated smallRNA, for example in *C. elegans* or *Tetrahymena thermophila* (Gu et al., 2009; T. Han et al., 2009; S. R. Lee & Collins, 2006; Ruby et al., 2006).

This indicates, that the endogenous smallRNA pathways in *paramecium* seem to be very complex and different from other RNAi pathways. Apart from SRCs, one additional endogenous pathway capable of gene regulation has been described in *paramecium*, which is transgene induced silencing. Here, a truncated transgene introduced in the *paramecium* macronucleus produces an aberrant mRNA transcript which triggers silencing of the endogenous gene copy *in trans*. Even though those two pathways share key components, Ptiwi13 and Ptiwi14 (Drews et al., 2021), transgene induced smallRNAs show a strict antisense bias compared to the sense bias displayed by the SRC associated smallRNAs, indicating that different criteria are applied to strand selection, depending on the pathway that is funneling smallRNAs into the Ptiwi proteins.

5.3.3 Ptiwi14 Might Associate with Nascent Transcripts and Might be Involved in CTGS

Ptiwi14 is the only vegetative Ptiwi analyzed so far, that shows localization in the macronucleus, enabling Ptiwi14 to be potentially associated with chromatin modifications and CTGS. Indeed, it has been shown for transgene induced silencing, that siRNAs produced by the truncated transgene can trigger chromatin alterations at the target gene locus, including a higher nucleosome occupancy, increased levels of the repressive mark H3K27me3, and a decrease of H3K9ac and H3K4me3 (Götz et al., 2016). In this context, it was speculated, that un-spliced nascent transcripts are used as a template for siRNA production, which can then be loaded into Ptiwi14, implying that Ptiwi14 might be responsible for the establishment of a heterochromatic state by attacking nascent transcripts at the target locus (Götz et al., 2016).

In this work, it was shown that Ptiwi14 associates with endogenous smallRNAs that map to intron sequences, allowing Ptiwi14 to attack nascent transcripts of genes, similar to its speculated function in the transgene induced silencing pathway. The association of Ptiwi14 with intron mapping molecules might provide further hints, that a CTGS-like mechanism is used in *paramecium* to establish chromatin modification mediated by Ptiwi14 for endogenous gene regulation.

Besides its potential function in endogenous and transgene-induced RNAi, results presented in this work showed that Ptiwi14 also associates with primary siRNAs and, as the only Ptiwi, with secondary siRNAs triggered by RNAi by Feeding. So far, Ptiwi14 was not considered to be involved in the RNAi by Feeding pathway (Bouhouche et al., 2011). However, the previous analysis were limited on a phenotypical screening, meaning that Ptiwi14 can still be involved in the pathway, but is apparently not the primary driving force for the silencing phenotype. The true function of Ptiwi14 in the RNAi by Feeding pathway remains elusive, since potential chromatin modification caused by RNAi by Feeding in *paramecium* has not been studied yet.

However, it is noteworthy that earlier work studying the RNAi by Feeding pathway mentioned, that dsRNA feeding during the cell's sexual development, can cause deletions of the targeted gene in the next sexual generation (Garnier et al., 2004). These deletions are not limited to the dsRNA target sequence, but expand over the entire gene body, suggesting that the deletions are not caused by primary siRNAs, but more likely secondary siRNAs. Ptiwi14 being associated with secondary siRNAs, located in the macronucleus and the only vegetative Ptiwi that not only shows expression during vegetative growth, but also during autogamy, might suggest that Ptiwi14 could be responsible for the aforementioned deletions.

RNAi by Feeding triggered siRNAs making modifications in the cell nucleus is not a new concept. In *C. elegans*, for example, several different RNAi pathways, including RNAi by Feeding, cumulate in the production of secondary siRNAs, which are loaded into the argonaute WAGO and establish RNA polymerase II stalling and a heterochromatic state, which can even be passed on to new generations (Burton et al., 2011; Frolows & Ashe, 2021). Deletion of the target gene in *paramecium* might therefore be the next step to not transiently and reversibly silence a gene, but permanently delete it, only storing a copy of the target gene in the micronuclear genome.

5.3.4 22nt smallRNAs are Loaded in Ptiwi12 and Ptiwi15

So far, three different sizes of smallRNAs were described in *paramecium*, siRNAs of predominantly 23nt length involved in different RNAi pathways during vegetative growth, like transgene induced RNAi, feeding induced RNAi or endogenous RNAi (Carradec et al., 2015; Götz et al., 2016; Karunanithi, Oruganti, et al., 2019) and 25nt long scnRNAs/ 27nt long iesRNAs, involved in DNA elimination during sexual reproduction (Lepère et al., 2008; Sandoval et al., 2014). For all those pathways, no 22nt long sense siRNAs were described, even though presented data in this work showed an association of smallRNAs of this length with Ptiwi12 and Ptiwi15 in both, endogenous and exogenous RNAi.

Analysis of the Dicer and Dicer-like enzymes in *paramecium* revealed, that most Dicer-like proteins produce either smallRNA of varying length or 25nt, leaving Dicer1, the only Dicer with an annotated RNAse III domain, responsible for the production of 23nt siRNAs (Hoehener et al., 2018; Sandoval et al., 2014). This leaves the synthesis of 22nt siRNAs elusive, since no biogenesis pathways and enzymatic requirements for their production are known so far.

If 22nt siRNAs are not produced by any Dicer enzyme, then biogenesis of the 22nt siRNAs has to be realized in another manner. Several enzymes have been described to be capable to trim smallRNAs after biogenesis to alter their length. Examples for those enzymes are the 3'to 5' exonuclease Eri1, involved in smallRNA biogenesis of several different RNAi pathways in different organisms (Duchaine et al., 2006; R. C. Lee et al., 2006; Thomas et al., 2014), Nibbler, a 3' to 5' exonuclease involved in miRNA trimming (Han et al., 2011; Xie et al., 2020), or PARN proteins, Poly-A specific ribonucleases that were also reported to trim miRNAs at the 3' end (Lee et al., 2019). Even though processes like that have not been described in *paramecium* yet, at least two different Eri proteins and three different PARN proteins can be found in the

paramecium genome, which could be potential candidates for smallRNA trimming, opening the possibility that smallRNA trimming processes might be of importance in *paramecium*.

5.3.5 Ptiwi12 and Ptiwi15 Select for Primary smallRNAs and Might Associate with Dicer Duplexes

Sequence logo analysis of primary siRNAs revealed a strong 5' uracil bias of antisense directed smallRNAs bound to Ptiwi12 and Ptiwi15 not present in the corresponding total RNA. 5' nucleotide preferences are something often observed in argonaute associated smallRNAs. These preferences are usually caused by interaction with the MID domain of the argonaut, selecting for 5' U smallRNAs specifically (Brennecke et al., 2007; Stein et al., 2019). Presence of such 5'U preference for antisense smallRNAs in Ptiwi12 and Ptiwi15 might indicate an active selection of those smallRNAs for loading. However, absence of this preference doesn't mean, that the primary siRNAs are not loaded into the other Ptiwis. Interestingly, a slight 5'U preference can be observed in Ptiwi12 and PTiwi15 IPs for sense RNAs as well, suggesting that Ptiwi12 and 15 might not online actively load antisense RNAs, but also sense RNA.

Loading of both, sense and antisense orientated smallRNAs was something observed for all four analyzed Ptiwi proteins. This balanced strand ratio is something unique for the RNAi by Feeding induced primary smallRNAs, since secondary smallRNA induced by the RNAi by Feeding pathway showed a strict antisense strand bias. Additionally, other analyzed pathways for those Ptiwis, endogenous RNAi and transgene induced RNAi (Drews et al., 2021) also showed strand biases, showing the exclusivity of the strand balance for the primary smallRNAs of the exogenous RNAi pathway. Interestingly, Ptiwi12's and Ptiwi15's lack of strand selection resulted, in some replicates, in detection of siRNAs showing a Dicer cleavage overlap signature. This suggests the association of Dicer cut duplexes with the two Ptiwis, which underlines the Dicer-dependent biogenesis of the primary siRNAs, but highlights the lack of strand selection in this particular pathway.

Usually, selection of one strand is caused by, beside 5' nucleotide preference, thermodynamic properties that are either sensed by the argonaute protein directly, or determined by the Dicer complex, assisting the loading of the argonaute protein with the guide strand (Boland et al., 2010; Iwakawa & Tomari, 2022; Suzuki et al., 2015; Tomari et al., 2004). This means that guide strand selection is a process influenced by different factors, including the physical properties of the specific duplex, chemical modifications carried by the RNA, the Dicer complex involved

in biogenesis, and individual properties of the argonaute itself (Iwakawa & Tomari, 2022; Medley et al., 2020; Steiner et al., 2009; Tomari et al., 2004; Varley et al., 2020). Differences in either factor specific for the RNAi by Feeding pathway for primary siRNAs might influence the strand selection process and therefore explain the balanced strand bias.

6 Untemplated Nucleotide Modifications of smallRNAs

6.1 Background

Apart from the mentioned core RNAi components, the forward genetic screen uncovering several enzymes involved in the RNAi by Feeding pathway in *paramecium* showed the involvement of two nucleotidyltransferases, Cid1 and Cid2, in the pathway (Marker et al., 2014). While no further studies were conducted uncovering the function of Cid2, due to the lack of mutant cell lines, it has been shown that mutations in the *cid1* gene leads to a complete loss of primary siRNAs, suggesting that Cid1 plays a role in the biogenesis of those smallRNAs (Carradec et al., 2015).

Nucleotidyltransferases are enzymes that are capable of adding untemplated nucleotides to the 3'end of RNAs. They are involved in several different smallRNA pathways, including miRNA and siRNA pathways across different organisms (Burroughs et al., 2010; C. Z. Chung et al., 2016; X. Wang et al., 2015; Wyman et al., 2011). Here, the addition of untemplated nucleotides to the 3' end of a smallRNA usually competes with 2'-O-methylation of the terminal 3' nucleotide, the latter known to be a stabilization factor for piRNAs (Billi et al., 2012), while the former can lead to degradation of smallRNAs (Ibrahim et al., 2010; X. Wang et al., 2015; Zhao et al., 2012).

However, addition of untemplated nucleotides can not only lead to smallRNA degradation, but can fulfil other functions as well. As it was reported in *Tetrahymena thermophila*, nucleotidyltransferase activity is required for the initiation of the RdRP mediated synthesis of RNA, with the RdRP using the untemplated nucleotides provided by the nucleotidyltransferases as a primer (Talsky & Collins, 2010), giving an example of untemplated nucleotides not causing RNA degradation.

While in theory, nucleotidyltransferases can add any nucleotide to the 3' end of smallRNAs, most transferases display a preference towards a specific nucleotide. The most common preferences observed for transferases are the addition of uracil or adenine containing nucleotides (Burroughs et al., 2010; C. Z. Chung et al., 2016; Ibrahim et al., 2010; X. Wang et al., 2015), with cytosine or guanine representing only niche cases (Hyde et al., 2010; Malik et al., 2020).

For *paramecium*, no further studies have been conducted to characterize Cid1 or Cid2, so nucleotide preferences as well as other characteristics are still unknown. However, untemplated nucleotides analyzed in earlier studies revealed the addition of mainly uracil containing

nucleotides to siRNAs derived from the RNAi by Feeding pathway, suggesting that one or both of the two nucleotidyltransferases display a preference for uracil (Carradec et al., 2015).

In the following chapter, the question of whether untemplation can be found on smallRNAs generated by the RNAi by Feeding pathway will be answered by using a custom pipeline. Apart from general presence of the untemplation, the analysis will focus on whether untemplation is more abundant on specific strands of the feeding smallRNAs, or whether untemplation can be found across all strands at similar levels. Last, the question of whether smallRNAs carrying untemplated nucleotides are associated to Ptiwi proteins will be addressed. For that, data that has been generated and analyzed in previous chapters, namely the Heteroduplex dataset (see Chapter 4) and the Ptiwi IP dataset (see Chapter 5), will be subjected to the untemplated nucleotide analysis.

6.2 Results

6.2.1 smallRNAs Produced from Exogenous RNA Carry Untemplated Nucleotides

Previous work on smallRNAs derived from exogenous dsRNA revealed the presence of untemplated nucleotides added to smallRNA, presumably after Dicer cut. Those untemplated nucleotides were reported to be primarily added to sense smallRNA and consist of uracils added in various numbers (Carradec et al., 2015). Using the new established technique of Heteroduplex feeding (see Chapter 4), untemplated nucleotides were studied in wildtype cells to see whether the ability to distinguish original strands and RdRP produced strands might deepen the knowledge about untemplation (Figure 21).

As displayed and in concordance to the reports in Carradec et al., 2015, untemplated nucleotides represent a minority compared to the number of reads carrying no modifications (Figure 21A). Looking at the overall distribution of modified reads for each read length, it can be noted that more untemplated reads are present in molecules with longer read lengths compared to shorter reads, suggesting that the presence of untemplated nucleotides increases with length (Figure 21B).

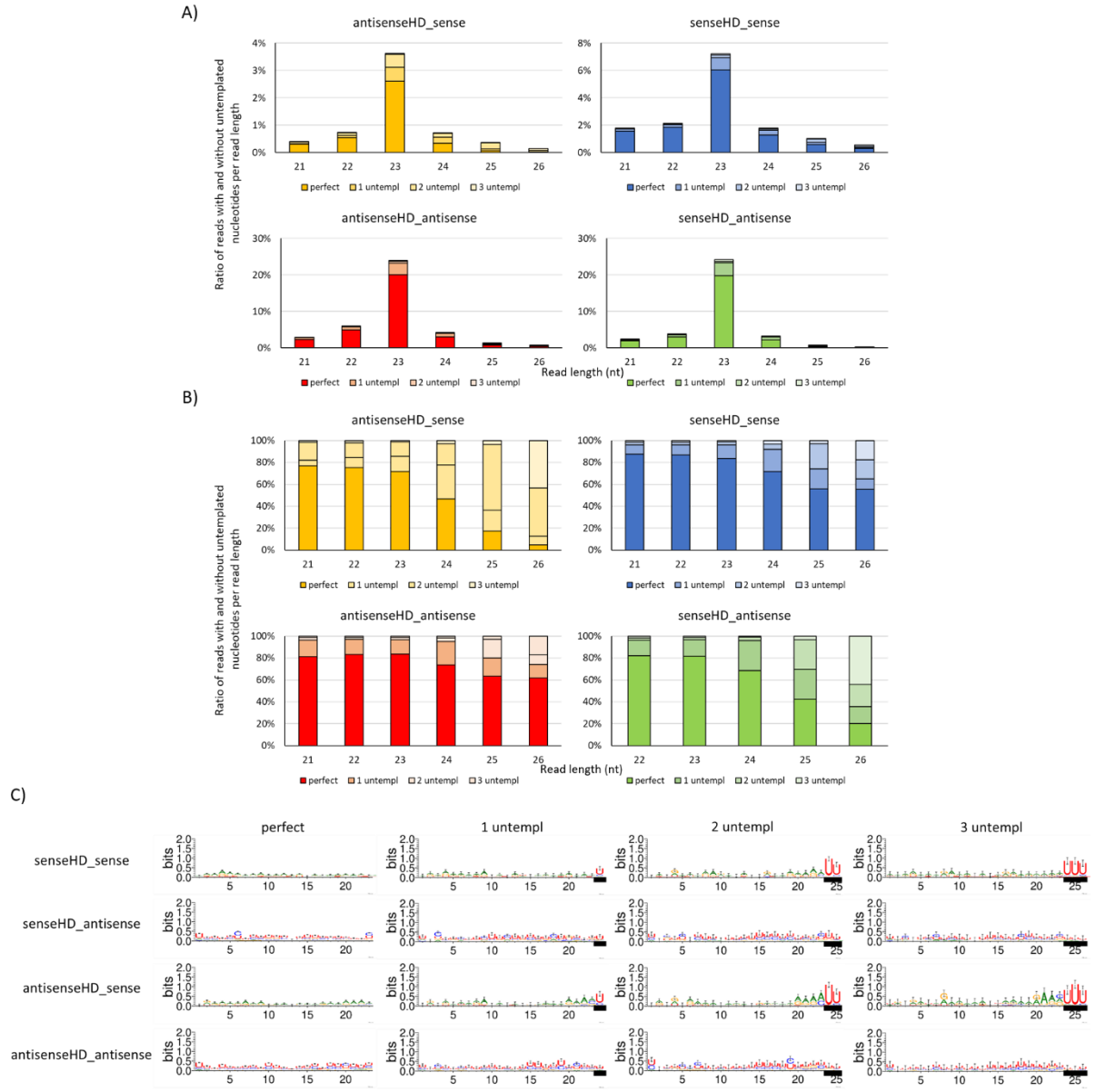

Figure 21 Untemplated nucleotides added to Heteroduplex derived smallRNAs in wildtype cells. A) Read length distribution of smallRNAs produced from the four Heteroduplex strands with and without a maximum number of three untemplated nucleotides. Strands displayed are labeled using the following color code: original antisense strand in red, original sense strand in blue, RdRP derived antisense strand in green, RdRP derived sense strand in orange. Ratio of untemplated nucleotides is labeled by lighter shades of the corresponding color. B) Display of the ratio of untemplated nucleotides present for each individual read length from smallRNAs derived from the four Heteroduplex strands. C) Sequence logo analysis of 23nt+x reads, carrying zero to three untemplated nucleotides derived from the four Heteroduplex strands. Untemplated nucleotides displayed are underlined with a black bar. Sequence logos were calculated using the weblogo tool, normalizing the logo for the paramecium genome base composition (A/U: 0.36; G/C: 0.14 each), according to (Lepère et al., 2009).

Interestingly, untemplation seems to be more frequent on strands produced from the RdRP proteins, senseHD_antisense and antisenseHD_sense, compared to the original strands, suggesting that the smallRNAs derived from RdRP products carry more untemplated nucleotides than smallRNAs derived from the original strand.

Additionally, the number of untemplated nucleotides seem to follow a pattern of 23nt+x, meaning that one untemplated nucleotide is the majority of untemplation observed for 24nt long molecules, while the number of reads with two untemplated nucleotides increases in the group of 25nt long molecules and so on, specifically observable for the RdRP products (Figure 21B).

Performing a sequence logo analysis of reads following this 23nt+x patterns shows, that untemplation on the sense strand is mainly consisting of uracils, being either single uracil additions or addition of a poly-uracil tail, matching to the previous observation of poly-uridylation (Carradec et al., 2015), while the antisense strands don't show any nucleotide preference, indicating addition of random nucleotides (Figure 21C). Poly-uridylation does not follow the observation of untemplation frequency being higher at RdRP produced strands, but seems to be more dependent on strand orientation, since it is found on both, the original sense strand and the sense directed RdRP product. This poly-uridylation is not found on reads with a total length of 23nt (Figure S 6), indicating that the poly-uridylation is probably added to siRNAs with a length of 23nt.

Taking these results together, it can be noted that, while the smallRNAs derived from RdRP produced strands show higher degrees of untemplation, the sense orientated smallRNA independent of their origin show specific modification consisting of poly-uridylation.

6.2.2 Poly-Uridylation of Exogenous siRNAs is Present in Ptiwi IPs

With evidence of poly-uridylation of exogenously triggered smallRNA being present in bulk data, the question of whether the modified smallRNAs can also be found bound to the corresponding Ptiwi proteins arises. Data from dsRNA fed transgene cell lines (see Chapter 5) were therefore analyzed for the presence of untemplated nucleotides in primary feeding associated smallRNAs (Figure 22).

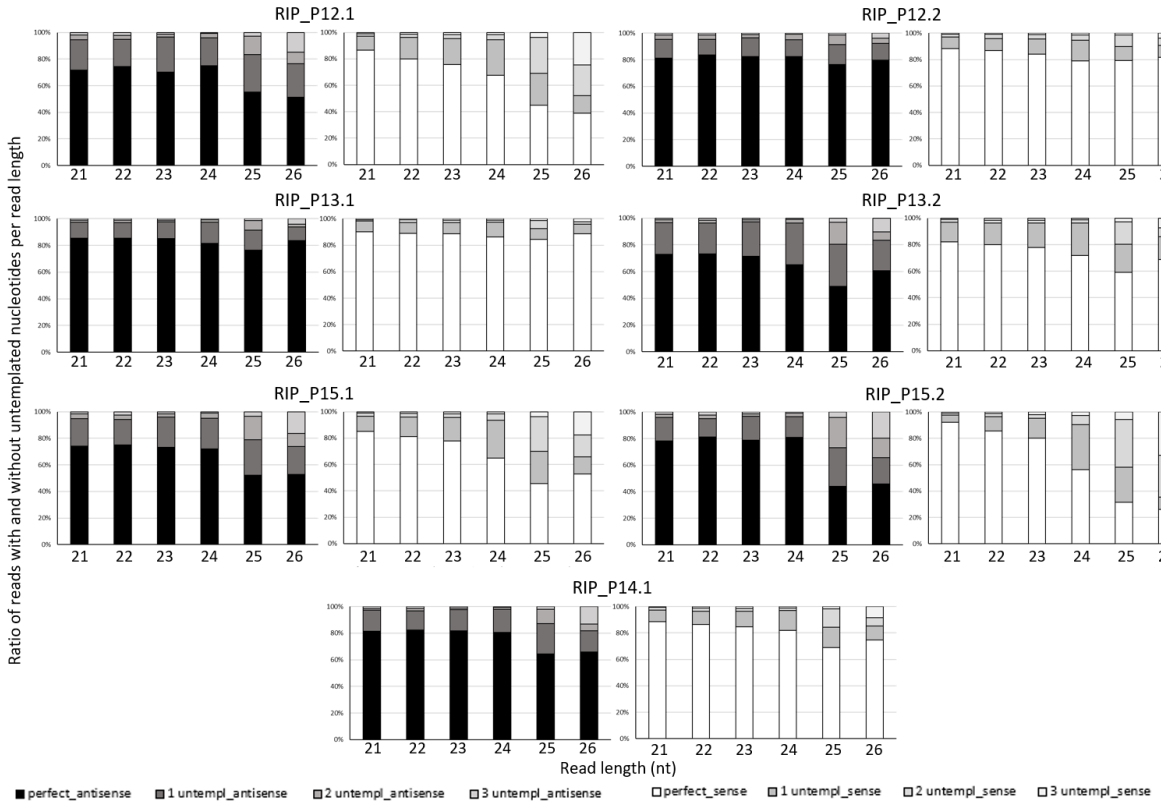

Figure 22 Untemplated nucleotides of *nd169* feeding associated primary smallRNAs present in Ptiwi RIP data. Displayed is the ratio of untemplated nucleotides present for each individual read length from feeding associated smallRNAs of sense or antisense orientation. Shown are ratios of molecules carrying from zero up to three untemplated nucleotides.

Analyzing untemplation in RNA IP data for different Ptiwis, it can be noted that the abundance of untemplation shows large variations, not only between different Ptiwis, but also between replicates of individual Ptiwis (Figure 22). The lowest amount of untemplation can be found in both Ptiwi13 replicates, with Ptiwi12 and Ptiwi15 showing increased, but varying amounts of untemplation between replicates. For most replicates, sense and antisense orientated smallRNAs show similar degrees of untemplation. However, the experimental setup of the Ptiwi RIPs not using a Heteroduplex dsRNA does not allow for separation of original and RdRP

produced strands, masking the asymmetrical distribution of general untemplation between those two products, which was observed prior (Figure 21).

Analyzing the corresponding sequence logos of sense orientated 23nt+x reads, it can be observed that almost all Ptiwi replicates carry the poly-uridylation signature of 23nt siRNAs being elongated using multiple uracils (Figure S 7), while similar patterns cannot be observed for antisense orientated molecules or molecules of 23nt length (data not shown). This suggests that either, Ptiwi proteins load poly-uridylated smallRNAs, or that smallRNAs get uridylated after Ptiwi loading.

6.2.3 Cid1 Mutant Data Suggest Deficiency in RdRP Product Formation

Dealing with untemplated nucleotides, nucleotidyltransferases usually are the class of enzymes responsible for the addition of untemplated nucleotides to smallRNAs. In *paramecium*, two nucleotidyltransferases, Cid1 and Cid2, were reported to be involved in the RNAi by Feeding pathway (Marker et al., 2014). Since only strains carrying a mutation in the *cid1* gene were available, smallRNA biogenesis using the Heteroduplex dsRNA with cells lacking functional Cid1 were analyzed (Figure 23A).

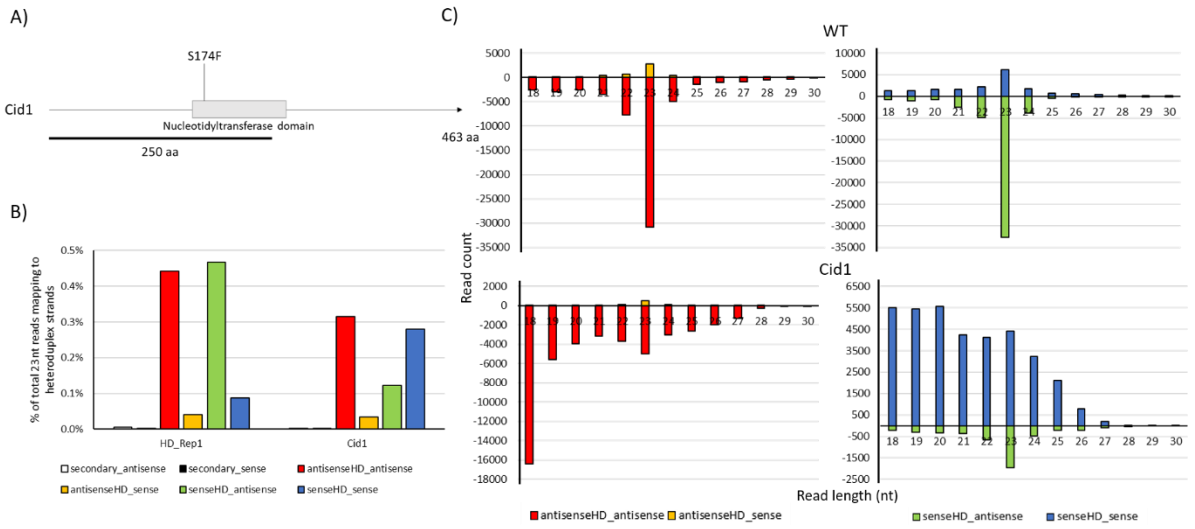

Figure 23 Heteroduplex derived smallRNA in Cid1 mutant. A) Visualization of the mutant Cid1 strain used in the Cid1 Heteroduplex analysis. B) Abundance of reads mapping to the different Heteroduplex strands (original antisense strand in red, original sense strand in blue, RdRP derived antisense strand in green, RdRP derived sense strand in orange) or the corresponding endogenous sequence without mismatches, representing secondary siRNAs (antisense orientated in white, sense orientated in black). Abundances are displayed for a wildtype cell line (HD_Repl1) and a cell line carrying a Cid1 mutation (Cid1). B) Read length distribution of smallRNAs mapping to the different Heteroduplex strands, color coded as described prior, derived from wildtype cells (top) or cells carrying a cid1 mutation (bottom).

Looking at the smallRNAs produced in Cid1 mutants, it can be noted that mutants produce lower levels of smallRNAs than wildtype cells. This is especially true for RdRP produced strands, showing a large decrease in level most noticeable for the usually highly abundant senseHD_antisense RdRP product (Figure 23B).

Looking at the smallRNA read length, it can be noted that smallRNAs derived from the Cid1 mutant follow an interesting distribution (Figure 23C). In comparison to the wildtype read length distribution displayed, Cid1 mutants show a loss of the 23nt siRNA peak for both original strands (red and blue), but a clear 23nt siRNA peak for both of the RdRP produced strands (green and orange), even though their levels are heavily decreased, as already noted prior. This suggests that loss of Cid1 leads to the inability of efficiently dicing the original strands and an impaired production of the RdRP derived strand. However, small amounts of RdRP derived strands, probably produced by background activity, can still be diced properly.

Due to the low levels of smallRNAs produced in the Cid1 mutant, not enough reads were available to perform the untemplated nucleotide analysis with reads from the Cid1 mutants. However, since behavior in poly-uridylation for wildtype heteroduplex derived smallRNAs so far matched to the already described behavior in literature, it can be speculated that poly-uridylation level in the Cid1 mutant would decrease to a similar extend as primary siRNA production decreased, implying no direct effect of Cid1 on the poly-uridylation level and therefore not providing any evidence of Cid1 being responsible for poly-uridylation of smallRNAs (Carradec et al., 2015).

6.3 Discussion

6.3.1 Poly-Uridylation of Sense Directed siRNA Might Lead to Degradation

As observed in the data, several different cases of untemplation can be noted for exogenously triggered smallRNA. However, in most cases, especially when looking at 23nt siRNAs or antisense orientated smallRNAs following the 23nt+x pattern, untemplation can be detected, but no nucleotide preference. The same is true for antisense orientated RdRP derived smallRNAs, making speculation about the origin of the untemplation and its function difficult.

Only the sense orientated 23nt+x molecules displayed a heavy uracil bias in the sequence logo, providing a specific characteristic and making speculation possible. So far, two prime candidates for untemplation in the RNAi by Feeding mechanism were provided, Cid1 and Cid2 (Marker et al., 2014). Since data for both Cid enzymes were either missing, or not enough

sequencing reads were generated, no new insight on the involvement of either Cid enzyme in the poly-uridylation of smallRNAs can be provided. However, it was already shown in previous experiments that Cid1 is probably not involved in the poly-uridylation of the smallRNAs (Carradec et al., 2015). Additionally, it has been shown that lack of Cid1 and Cid2 results in the inability of the cell to produce primary siRNAs, making it unlikely that Cid1 or Cid2 are responsible for their poly-uridylation after biogenesis as well, suggesting that one of the other Cid proteins encoded in the *paramecium* genome might be responsible for this poly-uridylation (Carradec et al., 2015; Marker et al., 2014).

With the poly-uridylation being present in the RNA IP data, and also following the pattern of 23nt+x, meaning that 23nt siRNAs are elongated with the addition of uridines, it is highly likely that this poly-uridylation is placed after the biogenesis of the siRNA duplex by Dicer and loading of the duplex by the Ptiwi proteins. Since poly-uridylation is only present on the sense strand, the strand that should be regarded as the passenger strand, since PTGS requires an antisense orientated siRNA to bind to the mRNA and thereby attack the target transcript, it can be speculated that the poly-uridylation is a mark to facilitate degradation of the passenger strand. RNA uridylation as a signal for RNA degradation is a phenomenon that has been described many times in different organisms, including various different plant species, *Schizosaccharomyces pombe*, and humans (de Almeida et al., 2018; Katoh et al., 2009; Scott & Norbury, 2013; Zhao et al., 2012). Since usage of the Heteroduplex dsRNA didn't show any correlation of the 23nt+x poly-uridylation of sense directed smallRNAs with their origin, it is likely that uridylation does not play any particular role in the biogenesis of the smallRNAs and might therefore be involved in strand selection after biogenesis.

6.3.2 Cid1 Has an Impact on RdRP Derived Strand Production

Since it is known that both, Cid1 and Cid2, are involved in the RNAi by Feeding pathways, and that lack of both leads to a missing silencing phenotype and to reduced/no production of primary siRNAs, it is very likely that both of the Cid proteins are involved in the biogenesis of the primary siRNAs (Carradec et al., 2015; Marker et al., 2014).

So far, Cid1 was not connected to untemplated nucleotides observed in smallRNA sequencing data, implying that Cid1 might not work on smallRNA, but longRNA. Looking at the biogenesis of Heteroduplex derived smallRNAs, it was shown that mutants of Cid1 showed a loss of 23nt siRNA peaks for both original strands, but not for RdRP produced strands, even though the

level of RdRP produced strands decreased in the mutant, bringing Cid1 into connection with RdRP1 and RdRP2.

In *Tetrahymena thermophila*, it has been shown that the single RdRP present associates not only with Dicer, but also with different nucleotidyltransferases (S. R. Lee et al., 2009; S. R. Lee & Collins, 2007), strengthening the potential connection between Cid1 and RdRP1 and/or RdRP2. Additionally, it has been shown in *Tetrahymena thermophila*, that RdRP activity on single stranded RNA required the activity of a nucleotidyltransferase, that catalyzes the addition of multiple uridines to the end of the single stranded RNA, leading to looping of the poly-U tail and priming of the RdRP activity on the single strand (Talsky & Collins, 2010). In a scenario like that, lack of the nucleotidyltransferase would lead to a loss of the RdRP product, which is what can be observed in the Cid1 mutant. Residual Cid activity, either of the mutated Cid1 or of Cid2, might still lead to production of low levels of RdRP produced strands, which might explain the presence of clear 23nt siRNA peaks for the RdRP produced strands in the mutant data, being the result of low background levels of diceable duplexes. Low level of 23nt siRNAs in the original strands might get overshadowed by random degradation products.

7 General Discussion and Perspectives

In the following chapter, the results of the individual chapters are compared with each other to give an overview of implications resulting from the individual findings and their influence on other protein groups not discussed in prior chapters.

7.1 RdRP Complexes Might Be Responsible for Exogenous dsRNA Processing

In this work, it has been shown that the two RdRPs, RdRP1 and RdRP2, are involved in the production of primary siRNAs by transcribing the two strands of exogenous dsRNA into new complementary strands. Most likely, the resulting new dsRNA, presumably consisting of one exogenous and one synthesized strand, is then processed further into primary siRNAs.

It has been shown that loss of either RdRP1 or RdRP2 results in the loss of both RdRP produced strands (see Chapter 4.2.4), indicating that both enzymes are necessary for their production, refuting prior speculation that both RdRPs work independently from each other at this step of the pathway (Carradec et al., 2015). Additionally, it has been shown that loss of the nucleotidyltransferase Cid1 leads to reduced amounts of RdRP products (see Chapter 6.2.3), suggesting that Cid1, might also be involved in this step of the RNAi by Feeding pathway. Interestingly, the low levels of RdRP produced strands present in the Cid1 mutant show a proper Dicer cleavage pattern, displaying a 23nt peak in their read length distribution (see Chapter 6.2.3), suggesting that these low levels of RdRP-produced strands, probably binding to the complementary original strand, are Dicer substrates. At the same time, the loss of the 23nt peak for the original strands suggests that even if the RdRP products cannot be synthesized, the exogenous double strand is never substrate for Dicer. This implies that either, the exogenous strands are strictly handled by the cell, preventing their contact with the Dicer enzyme, or the exogenous strands are kept single-stranded upon import into the cell from the food vacuole, to prevent an imported dsRNA to be subjected to Dicer cleavage.

Looking at these results, it can be speculated that RdRP1, RdRP2 and at least Cid1 might work as a complex, with its main function being the transcription of the exogenous RNA into a new dsRNA, consisting of partially host derived RNA.

RdRPs in general are mainly studied in viruses, since they represent the main component of retroviral replication machineries. Indeed, many viral RdRPs seem to form complexes with several different co-factors to facilitate retroviral genome replication, some of those co-factors

being nucleotidyltransferases or the RdRP itself displaying nucleotidyltransferase activity (Pitsillou et al., 2021; Z. Wang et al., 2013; W.-F. Zhang et al., 2020), highlighting the potential importance of such activity in a replication-like process.

Additionally, co-operation of nucleotidyltransferases and RdRPs has not only been demonstrated in viruses, but also in the ciliate *Tetrahymena thermophila*, where terminal nucleotidyltransferase activity is connected to RdRP activity for several different RNAi pathways, and in one case even required for the synthesis initiation of the RdRP (S. R. Lee et al., 2009; Talsky & Collins, 2010).

To proof interaction between RdRPs and Cid in *paramecium*, further characterization of Cid1 as well as RdRP1 and RdRP2 and their cellular functions would be necessary. Performing interaction studies of the RdRPs with Cid1, either *in vitro* or *in vivo*, in form of co-immunoprecipitations or yeast two hybrid assays, as an example, could shed some light on the interacting of the three proteins with each other. This complex might even include Cid2, since it was shown once that Cid2 is also involved in the RNAi by Feeding pathway. However, since then, no further data concerning Cid2 was generated (Carradec et al., 2015; Marker et al., 2014).

Considering the potential interaction between RdRPs and Cids, it is an interesting strategy of *paramecium* to convert exogenous RNA into cellular synthetized RNA before introducing it into a cellular pathway, giving *paramecium* additional agency and control over the foreign RNA. The concept of RNAi as a virus defense mechanism is not new, but has been discussed for many different organisms, including plants in general, but also *C. elegans* and *Drosophila melanogaster* (Pumplin & Voinnet, 2013; Wilkins et al., 2005; Zambon et al., 2006). Not involving the exogenous, potentially virus-derived RNA in a pathway, but gaining control over it by converting it into a dsRNA using cellular components might increase the effectiveness of this pathway, which might be the reason why so far, no viruses have been described that are capable to infect *Paramecium tetraurelia*, suggesting an extraordinary defense mechanism. In that case, RdRP activity would serve as a mechanism to distinguish self from non-self RNA, and converting non-self RNA into a less dangerous intermediate product before guiding it into cellular pathways.

7.2 Uridylation of smallRNA Might Help Ptiwis with Strand Selection

While smallRNAs associated to SRCs displayed a clear sense bias, Ptiwi loaded primary siRNAs from the RNAi by Feeding pathway displayed no strand bias (see Chapter 5.2.7). Usually, argonaute protein select for one strand of the smallRNA duplex produced by Dicer, in most cases the antisense strand, to mediate their function. The antisense strand is thereby preferred, because it can bind to the sense directed mRNA, providing the ability to attack specific mRNA targets and cleave them, conveying the PTGS.

Beside this balanced loading preference, untemplation was observed for primary siRNAs in both total RNA data and Piwi selected RIP data. This untemplation showed a tendency of 23nt siRNAs being modified with poly-uracil tail of various length (see Chapter 6.2.1 and 6.2.2).

So far, the balanced strand bias in Ptiwi data is probably not a specific characteristic of the Ptiwi proteins, since strand selection worked for endogenous smallRNAs as shown prior (see Chapter 5.2.4), and for transgene induced smallRNA as described in literature (Drews et al., 2021). Therefore, the lack of obvious strand selection must be a property of the RNAi by Feeding pathway. Generally, smallRNAs have to be reverse complement to the target RNAs to ensure base pairing with the target, which promotes degradation. Considering the source of the primary siRNAs and the goal of RNAi to destroy RNA sharing the same sequence as the trigger, loading both strands does make sense in the context of RNAi by Feeding, since RNA molecules of both strand orientations are present in the food and serve as potential targets. However, only antisense orientated smallRNAs are probably capable of attacking endogenous mRNA transcripts, connecting only antisense orientated smallRNAs to the silencing phenotype. Sense orientated smallRNAs not regularly finding a target might be the reason why predominantly sense orientated siRNAs get modified with untemplated uridines, a modification generally considered to be an RNA degradation trigger (de Almeida et al., 2018; Katoh et al., 2009; Scott & Norbury, 2013; Zhao et al., 2012), which would lead to the elimination of sense loaded RISCs, while antisense loaded RISCs persist. This theory presumes that the balanced strand bias present in Ptiwi IP data is caused by Ptiwi proteins randomly selecting one strand of the generated siRNAs duplex, ending up with half of the Ptiwi proteins being loaded with sense orientated siRNAs, while the other half of the Ptiwi proteins load antisense orientated siRNAs.

On the other hand, the balanced strand bias might be achieved by the Ptiwi proteins being associated with the Dicer provided siRNAs duplex, thus not being loaded with a single stranded RNA but a double stranded one, which would make the RISC not functional. In that case, the uridylation of the sense siRNA might be a signal for degradation to promote strand selection,

leading to the degradation of the sense passenger strand and the persistence of the antisense guide strand. A hint for this hypothesis might be the detected Dicer overlap length present in some of the Ptiwi12 and Ptiwi15 replicates (see Chapter 5.3.5), suggesting the association of both Ptiwis with the entire Dicer cleaved siRNA duplex.

However, both hypotheses are not mutually exclusive and to the current point in time, neither of them can be regarded as proven.

In any way, it is unlikely that the prominent uridylation of 23nt siRNAs in the RNAi by Feeding pathway serves any purposes essential for the pathway, since the poly-uridylation seem to be a cell cycle dependent property of primary siRNAs. Even though poly-uridylation was observed in this and in previous works (Carradec et al., 2015), a very early study investigating primary siRNAs of the RNAi by Feeding pathway conducted during the sexual reproduction cycle of *paramecium* showed that during this time period, primary siRNAs are not poly-uridylated at the sense strand, but adenylated at the antisense strand, suggesting that the untemplated nucleotide preference changes with the life cycle stage of the cell for the same pathway (Lepère et al., 2009).

However, the genetic requirements and the proteins involved in the degradation pathway that might be triggered by the addition of untemplated nucleotides to are not know so far, requiring further work into this pathway to proof this theory.

7.3 RNAi by Feeding Induced smallRNAs Might Alter Chromatin Modifications

Prior to this work, it was speculated that Ptiwi13 loads primary siRNAs and mediates PTGS by attacking mRNA transcripts while Ptiwi12 and Ptiwi15 load secondary siRNAs and might wander into the MAC and establish chromatin changes (Bouhouche et al., 2011; Carradec et al., 2015). The reason for this hypothesis was the altered catalytic Slicer triad of Ptiwi12 and Ptiwi15 that might result in the inability of Ptiwi12 and Ptiwi15 to cleave target mRNA.

Sequencing of the smallRNAs associated to the different vegetative Ptiwis in *Paramecium tetraurelia* showed that all three Ptiwis involved in the RNAi by Feeding pathway (Bouhouche et al., 2011), Ptiwi12, Ptiwi13 and Ptiwi15, associate with primary siRNAs, while Ptiwi14, so far not connected with the Feeding pathway, associates with both, primary and secondary siRNAs (see Chapter 5.2.6). Ptiwi14 being localized in the macronucleus of *paramecium* and already playing a role in a pathway that involves changes to the chromatin landscape, the

transgene induced RNAi pathway (Drews et al., 2021; Götz et al., 2016), suggests that the RNAi by Feeding pathway might also influence the chromatin. So far, the chromatin dynamics during dsRNA Feeding in *paramecium* have not been studied, suggesting chromatin IP experiments with cells undergoing dsRNA feeding as a logical next step to shed further light onto Ptiwi14's involvement in the pathway.

Indeed, feeding induced secondary siRNAs altering chromatin modifications is something already described and extensively studied in *C elegans*. Here, chromatin states established by secondary siRNAs are even observed to be inherited to next generations (Burton et al., 2011; Frolows & Ashe, 2021; Spracklin et al., 2017). Similarly, RNAi by Feeding in *paramecium* can lead to deletions of the target gene in subsequent sexual generations (Garnier et al., 2004), mirroring the phenomena observed in *C. elegans* to higher extremes.

Besides Ptiwi14, differences in function for Ptiwi13 and Ptiwi12/15 could not be observed so far, which leaves the question of why three different Ptiwi proteins associate exclusively with primary siRNAs in the RNAi by Feeding pathways. Studying protein interactions between the four different Ptiwis and other cellular proteins might reveal more about differences between the Ptiwis involved. For example, association of Ptiwi12/15 and Ptiwi13 with different co-factors might reveal potential functional differences, while association of Ptiwi14 with chromatin remodeling components might strengthen its role as a chromatin influencing Ptiwi.

7.4 The RNAi by Feeding Pathway Might Involve More Components Not Known So Far

Besides Ptiwis and RdRPs, several additional enzymes can still be involved in this pathway, whose role is so far not understood at all.

Two proteins known to be involved (Carradec, 2014; Marker et al., 2014) but still elusive, are Pds1 and Pds2. For both proteins, it was possible to perform localization within the cell, revealing a cytoplasmatic localization for Pds1 and a membrane-association for Pds2 (see Chapter 3.2.2). Apart from that, no additional information regarding protein domain structures or similar predictions have been revealed, leaving the potential function of those two proteins unknown. It can only be speculated that Pds2 might serve as the so far unknown importer that is responsible for RNA uptake. However, it would have to oligomerize to achieve a channel size large enough for RNA import.

Pds1 remains even more elusive, since its cytosolic localization does not provide any further hints towards its function. The BLAST search conducted with relaxed criteria revealed different possible hits of proteins bearing functions connected to RNA pathways (see Chapter 3.2.1). From those proteins, the ATP dependent RNA helicase hit might be the most interesting one. Argonaute proteins in general can be divided into two classes, Ago and Piwi proteins, depending on the RNAs they interact with and the mode of function they operate under. *Paramecium*'s Ptiwi are proteins that resemble Piwis, but display Ago-like functions (Drews et al., 2021), making the Ptiwis bridge the gap between these two argonaute categories. Especially for Piwi proteins and the piRNA pathway, it has been reported for several different organisms, including *Drosophila melanogaster*, mouse and human, that RNA helicases interact with these Piwis and help in the silencing of RNA transposons or the biogenesis of piRNAs (Goodier et al., 2012; Guo & Wu, 2013; Tomari et al., 2004; Wenda et al., 2017; Zheng et al., 2010). Since *paramecium*'s Ptiwis show properties of the argonaute Piwi sub group, their association with putative RNA helicases might be possible. However, whether Pds1 fulfills such an RNA helicase role remains unclear until further characteristics of Pds1 are revealed experimentally.

Apart from these two proteins, it can further be speculated about the involvement of proteins in the RNAi by Feeding pathway that have not been considered yet. The preference of Ptiwi12 and Ptiwi15 to associate with 22nt long smallRNAs in both, the RNAi by Feeding and the endogenous RNAi pathway (see Chapter 5.2.2 and 5.2.7), suggests either, the production of so far not characterized 22nt smallRNAs by a Dicer enzyme, or a trimming of Dicer1 produced 23nt siRNAs to a length of 22nt. For the latter, smallRNA trimming enzymes would be required, of which members of the PARN and the Eri family can be found in the *paramecium* genome. Them not showing up in the forward genetic screen that revealed several of the other known RNAi by Feeding components (Marker et al., 2014) might suggest that those enzymes represent essential proteins, which is why mutations of those genes would not result in phenotypes detectable by the screening, but dead cells. This is probably the reason why none of the Ptiwi proteins were detected during the screening, even though their involvement in the RNAi by Feeding pathway have been confirmed prior (Bouhouche et al., 2011).

That being said, further studies of the smallRNA biogenesis pathways, including complexes that are formed with Dicer1 and the Ptiwi proteins, might give further insight into potential proteins modifying smallRNAs during and after their biogenesis.

7.5 The RNAi by Feeding Pathway: What We Know Now

Taking all the presented data together, several changes can be made to the scheme of the RNAi by Feeding pathway in *Paramecium tetraurelia* (Figure 24).

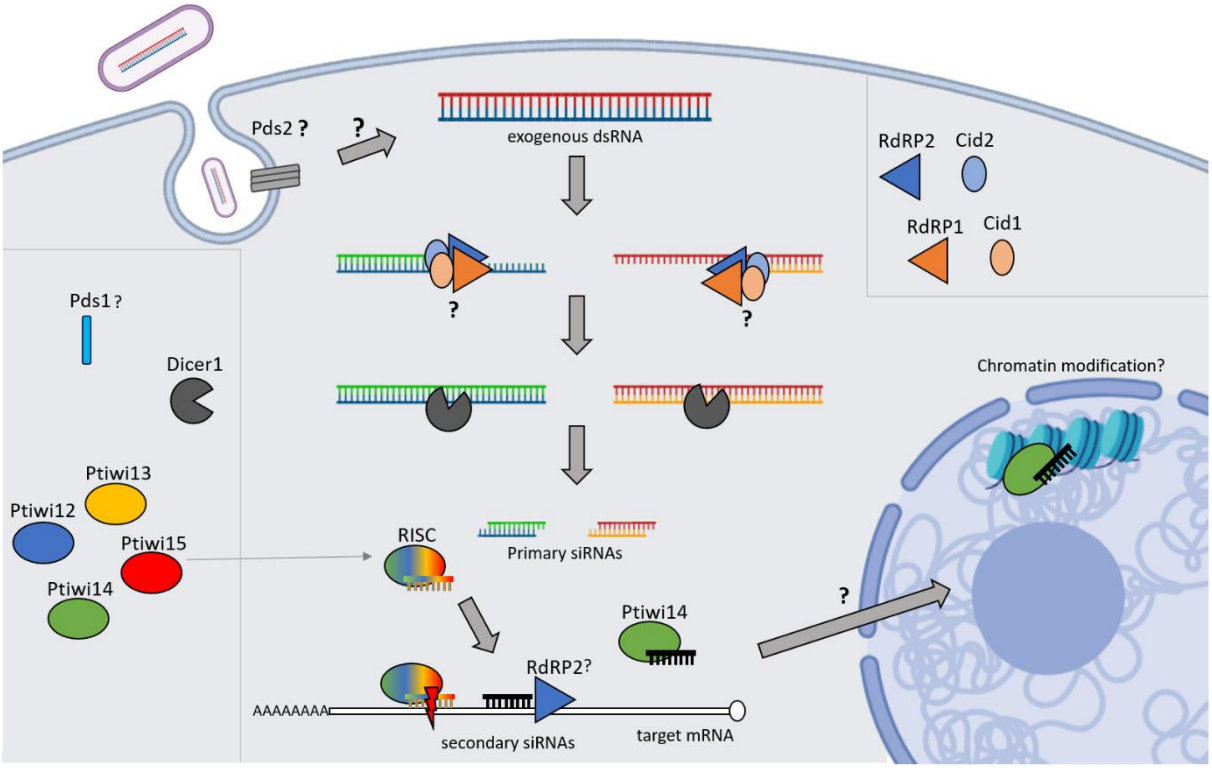

Figure 24 Updated schematic of the RNAi by Feeding pathway in *Paramecium tetraurelia*. Displayed is the process of RNAi by Feeding according to the presented data. Steps that are not further characterized yet or are speculated, are labeled with a question mark. Enzymes involved in different steps are labeled accordingly and color coded to distinguish different enzymes of the same family. Omitted from the schematic are potential modifications of the smallRNAs, including nucleotide trimming and addition of untemplated nucleotides, to improve readability. A detailed description of the pathway can be taken from the text. Parts of the image were created using BioRender.

As the first step, the dsRNA trigger has to be imported from the bacteria within the food vacuole. So far, it is unknown how the dsRNA escapes the food vacuole, but Pds2 might be involved in this step, due to its association with the cellular membranes. Additionally, it was not shown yet whether dsRNA or ssRNA is imported into the cell and subjected to the next steps of the RNAi by Feeding pathway.

After RNA import, the individual strands of the exogenous trigger RNA are transcribed by RdRP1 and RdRP2, potentially forming a complex together with Cid1 and Cid2. Here, Cid1 and Cid2 might add untemplated nucleotides to the exogenous RNA strand, which can then

loop around and prime the RdRP activity, as discussed prior. After RdRP activity, the two exogenous RNA strands are converted into new dsRNA, forming hybrids between exogenous and cellular produced RNA strands. Whether additional rounds of replication are performed at this step is not known yet.

The newly synthesized dsRNA is handed over to Dicer1, which cleaves the dsRNA into primary 23nt long siRNA duplexes. These duplexes are loaded into the involved Ptiwi proteins, Ptiwi12, Ptiwi15, Ptiwi13 and Ptiwi14. Here, strand selection of the guiding strand might be performed by the Ptiwi protein itself, or degradation of the (sense directed) passenger strand might be realized by poly-uridylation mediated RNA degradation. Alternatively, sense directed siRNAs loaded by Ptiwis might be subjected to poly-uridylation mediated RNA degradation due to a lack of complementary target mRNA.

Why these primary siRNAs are loaded into four different Ptiwi proteins and whether these proteins perform different task within in the cell is not understood yet. However, it is most likely that the primary siRNAs loaded into Ptiwis mediate PTGS, attacking mRNA transcripts of the target gene, leading to the RNAi silencing phenotype.

Using the attacked mRNA as a template, secondary siRNAs are synthesized, probably by RdRP2 (according to Carradec et al., 2015). These secondary siRNAs are loaded into Ptiwi14. So far, the specific function of the secondary siRNAs is unknown. However, it can be speculated that Ptiwi14 imports the secondary siRNAs into the MAC of the cells, where it might establish chromatin changes or deletions of the genetic locus, if feeding is performed during autogamy.

At some point during this process, it is possible that other enzymes, like PARN or Eri, trim the 3’ ends of siRNAs, leading to the production of 22nt long smallRNAs. This trimming might take place after loading of the siRNAs by the Ptiwis 12 and 15, since they associate the most with this elusive smallRNAs. Whether these 22nt smallRNAs display different functions than the canonical 23nt siRNAs is not understood.

Bibliography

Abello, A., Régnier, V., Arnaiz, O., Bars, R. L., Bétermier, M., & Bischerour, J. (2020). Functional diversification of Paramecium Ku80 paralogs safeguards genome integrity during precise programmed DNA elimination. *PLOS Genetics*, *16*(4), e1008723. https://doi.org/10.1371/journal.pgen.1008723

Adachi, T., Ago, H., Habuka, N., Okuda, K., Komatsu, M., Ikeda, S., & Yatsunami, K. (2002). The essential role of C-terminal residues in regulating the activity of hepatitis C virus RNA-dependent RNA polymerase. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, *1601*(1), 38–48. https://doi.org/10.1016/S1570-9639(02)00433-8

Agalioti, T., Chen, G., & Thanos, D. (2002). Deciphering the Transcriptional Histone Acetylation Code for a Human Gene. *Cell*, *111*(3), 381–392. https://doi.org/10.1016/S0092-8674(02)01077-2

Ahn, D.-G., Choi, J.-K., Taylor, D. R., & Oh, J.-W. (2012). Biochemical characterization of a recombinant SARS coronavirus nsp12 RNA-dependent RNA polymerase capable of copying viral RNA templates. *Archives of Virology*, *157*(11), 2095–2104. https://doi.org/10.1007/s00705-012-1404-x

Alleman, M., Sidorenko, L., McGinnis, K., Seshadri, V., Dorweiler, J. E., White, J., Sikkink, K., & Chandler, V. L. (2006). An RNA-dependent RNA polymerase is required for paramutation in maize. *Nature*, *442*(7100), 295–298. https://doi.org/10.1038/nature04884

Allen, S. E., Hug, I., Pabian, S., Rzeszutek, I., Hoehener, C., & Nowacki, M. (2017). Circular Concatemers of Ultra-Short DNA Segments Produce Regulatory RNAs. *Cell*, *168*(6), 990-999.e7. https://doi.org/10.1016/j.cell.2017.02.020

Allis, C. D., Chicoine, L. G., Richman, R., & Schulman, I. G. (1985). Deposition-related histone acetylation in micronuclei of conjugating Tetrahymena. *Proceedings of the National Academy of Sciences*, *82*(23), 8048–8052. https://doi.org/10.1073/pnas.82.23.8048

Altman, S. (1990). Enzymatic Cleavage of RNA by RNA (Nobel Lecture). *Angewandte Chemie International Edition in English*, *29*(7), 749–758. https://doi.org/10.1002/anie.199007491

Andrews, S. (2010). *A Quality Control Tool for High Throughput Sequence Data [Online]. Available online at: Http://www.bioinformatics.babraham.ac.uk/projects/fastqc/*. https://qubeshub.org/resources/fastqc

Antoniewski, C. (2014). Computing siRNA and piRNA Overlap Signatures. In A. Werner (Ed.), *Animal Endo-SiRNAs: Methods and Protocols* (pp. 135–146). Springer. https://doi.org/10.1007/978-1-4939-0931-5_12

Aoki, K., Moriguchi, H., Yoshioka, T., Okawa, K., & Tabara, H. (2007). In vitro analyses of the production and activity of secondary small interfering RNAs in C. elegans. *The EMBO Journal*, *26*(24), 5007–5019. https://doi.org/10.1038/sj.emboj.7601910

Arnaiz, O., Meyer, E., & Sperling, L. (2020). ParameciumDB 2019: Integrating genomic data across the genus for functional and evolutionary biology. *Nucleic Acids Research*, *48*(D1), D599–D605. https://doi.org/10.1093/nar/gkz948

Atwood, C. S., Perry, G., Zeng, H., Kato, Y., Jones, W. D., Ling, K.-Q., Huang, X., Moir, R. D., Wang, D., Sayre, L. M., Smith, M. A., Chen, S. G., & Bush, A. I. (2004). Copper mediates dityrosine cross-linking of Alzheimer's amyloid-beta. *Biochemistry*, *43*(2), 560–568. https://doi.org/10.1021/bi0358824

Aury, J.-M., Jaillon, O., Duret, L., Noel, B., Jubin, C., Porcel, B. M., Ségurens, B., Daubin, V., Anthouard, V., Aiach, N., Arnaiz, O., Billaut, A., Beisson, J., Blanc, I., Bouhouche, K., Câmara, F., Duharcourt, S., Guigo, R., Gogendeau, D., … Wincker, P. (2006). Global trends of whole-genome duplications revealed by the ciliate Paramecium tetraurelia. *Nature*, *444*(7116), 171–178. https://doi.org/10.1038/nature05230

Axtell, M. J. (2013). ShortStack: Comprehensive annotation and quantification of small RNA genes. *RNA*, *19*(6), 740–751. https://doi.org/10.1261/rna.035279.112

Bannister, A. J., Zegerman, P., Partridge, J. F., Miska, E. A., Thomas, J. O., Allshire, R. C., & Kouzarides, T. (2001). Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature*, *410*(6824), 120–124. https://doi.org/10.1038/35065138

Baudry, C., Malinsky, S., Restituito, M., Kapusta, A., Rosa, S., Meyer, E., & Bétermier, M. (2009). PiggyMac, a domesticated piggyBac transposase involved in programmed genome rearrangements in the ciliate Paramecium tetraurelia. *Genes & Development*, *23*(21), 2478–2483. https://doi.org/10.1101/gad.547309

Baulcombe, D. (2004). RNA silencing in plants. *Nature*, *431*(7006), 356–363. https://doi.org/10.1038/nature02874

Beisson, J., Bétermier, M., Bré, M.-H., Cohen, J., Duharcourt, S., Duret, L., Kung, C., Malinsky, S., Meyer, E., Preer, J. R., & Sperling, L. (2010). Maintaining Clonal Paramecium tetraurelia

Cell Lines of Controlled Age through Daily Reisolation. *Cold Spring Harbor Protocols*, *2010*(1), pdb.prot5361. https://doi.org/10.1101/pdb.prot5361

Bhatia, A., Schäfer, H.-J., & Hrycyna, C. A. (2005). Oligomerization of the Human ABC Transporter ABCG2: Evaluation of the Native Protein and Chimeric Dimers. *Biochemistry*, *44*(32), 10893–10904. https://doi.org/10.1021/bi0503807

Billi, A. C., Alessi, A. F., Khivansara, V., Han, T., Freeberg, M., Mitani, S., & Kim, J. K. (2012). The Caenorhabditis elegans HEN1 Ortholog, HENN-1, Methylates and Stabilizes Select Subclasses of Germline Small RNAs | PLOS Genetics. *PLOS Genetics*. https://doi.org/10.1371/journal.pgen.1002617

Bischerour, J., Bhullar, S., Denby Wilkes, C., Régnier, V., Mathy, N., Dubois, E., Singh, A., Swart, E., Arnaiz, O., Sperling, L., Nowacki, M., & Bétermier, M. (2018). Six domesticated PiggyBac transposases together carry out programmed DNA elimination in Paramecium. *eLife*, *7*, e37927. https://doi.org/10.7554/eLife.37927

Blackburn, E. H., & Gall, J. G. (1978). A tandemly repeated sequence at the termini of the extrachromosomal ribosomal RNA genes in *Tetrahymena*. *Journal of Molecular Biology*, *120*(1), 33–53. https://doi.org/10.1016/0022-2836(78)90294-2

Boland, A., Tritschler, F., Heimstädt, S., Izaurralde, E., & Weichenrieder, O. (2010). Crystal structure and ligand binding of the MID domain of a eukaryotic Argonaute protein. *EMBO Reports*, *11*(7), 522–527. https://doi.org/10.1038/embor.2010.81

Boros, J., Arnoult, N., Stroobant, V., Collet, J.-F., & Decottignies, A. (2014). Polycomb Repressive Complex 2 and H3K27me3 Cooperate with H3K9 Methylation To Maintain Heterochromatin Protein 1α at Chromatin. *Molecular and Cellular Biology*, *34*(19), 3662–3674. https://doi.org/10.1128/MCB.00205-14

Bouhouche, K., Gout, J.-F., Kapusta, A., Bétermier, M., & Meyer, E. (2011). Functional specialization of Piwi proteins in Paramecium tetraurelia from post-transcriptional gene silencing to genome remodelling. *Nucleic Acids Research*, *39*(10), 4249–4264. https://doi.org/10.1093/nar/gkq1283

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

Brink, R. A. (1956). A Genetic Change Associated with the R Locus in Maize Which Is Directed and Potentially Reversible. *Genetics*, *41*(6), 872–889.

Brownell, J. E., Zhou, J., Ranalli, T., Kobayashi, R., Edmondson, D. G., Roth, S. Y., & Allis, C. D. (1996). Tetrahymena Histone Acetyltransferase A: A Homolog to Yeast Gcn5p Linking Histone Acetylation to Gene Activation. *Cell*, *84*(6), 843–851. https://doi.org/10.1016/S0092-8674(00)81063-6

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

Burroughs, A. M., Ando, Y., Hoon, M. J. L. de, Tomaru, Y., Nishibu, T., Ukekawa, R., Funakoshi, T., Kurokawa, T., Suzuki, H., Hayashizaki, Y., & Daub, C. O. (2010). A comprehensive survey of 3′ animal miRNA modification events and a possible role for 3′ adenylation in modulating miRNA targeting effectiveness. *Genome Research*, *20*(10), 1398–1410. https://doi.org/10.1101/gr.106054.110

Burton, N. O., Burkhart, K. B., & Kennedy, S. (2011). Nuclear RNAi maintains heritable gene silencing in Caenorhabditis elegans. *Proceedings of the National Academy of Sciences*, *Vol. 108*(No. 49). https://doi.org/10.1073/pnas.1113310108

Carmell, M. A., Xuan, Z., Zhang, M. Q., & Hannon, G. J. (2002). The Argonaute family: Tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis. *Genes & Development*, *16*(21), 2733–2742. https://doi.org/10.1101/gad.1026102

Carradec, Q. (2014). *Mécanismes et fonctions de la voie d'ARN interférence induite par ARN double brin chez Paramecium tetraurelia* [Phdthesis, Université Pierre et Marie Curie - Paris VI]. https://theses.hal.science/tel-01077957

Carradec, Q., Götz, U., Arnaiz, O., Pouch, J., Simon, M., Meyer, E., & Marker, S. (2015). Primary and secondary siRNA synthesis triggered by RNAs from food bacteria in the ciliate Paramecium tetraurelia. *Nucleic Acids Research*, *43*(3), 1818–1833. https://doi.org/10.1093/nar/gku1331

Cech, T. R. (2004). Self-Splicing and Enzymatic Activity of an Intervening Sequence RNA from Tetrahymena. *Bioscience Reports*, *24*(4), 362–385. https://doi.org/10.1007/s10540-005-2738-3

Chang, S.-S., Zhang, Z., & Liu, Y. (2012). RNA Interference Pathways in Fungi: Mechanisms and Functions. *Annual Review of Microbiology*, *66*(Volume 66, 2012), 305–323. https://doi.org/10.1146/annurev-micro-092611-150138

Chung, C. T., Niemela, S. L., & Miller, R. H. (1989). One-step preparation of competent Escherichia coli: Transformation and storage of bacterial cells in the same solution. *Proceedings of the National Academy of Sciences*, *86*(7), 2172–2175. https://doi.org/10.1073/pnas.86.7.2172

Chung, C. Z., Jo, D. H. S., & Heinemann, I. U. (2016). Nucleotide specificity of the human terminal nucleotidyltransferase Gld2 (TUT2). *RNA*, *22*(8), 1239–1249. https://doi.org/10.1261/rna.056077.116

Cogoni, C., & Macino, G. (1997). Isolation of quelling-defective (qde) mutants impaired in posttranscriptional transgene-induced gene silencing in Neurospora crassa. *Proceedings of the National Academy of Sciences*, *94*(19), 10233–10238. https://doi.org/10.1073/pnas.94.19.10233

Cogoni, C., & Macino, G. (1999). Gene silencing in Neurospora crassa requires a protein homologous to RNA-dependent RNA polymerase. *Nature*, *399*(6732), 166–169. https://doi.org/10.1038/20215

Couvillion, M. T., Bounova, G., Purdom, E., Speed, T. P., & Collins, K. (2012). A Tetrahymena Piwi Bound to Mature tRNA 3′ Fragments Activates the Exonuclease Xrn2 for RNA Processing in the Nucleus. *Molecular Cell*, *48*(4), 509–520. https://doi.org/10.1016/j.molcel.2012.09.010

Couvillion, M. T., Lee, S. R., Hogstad, B., Malone, C. D., Tonkin, L. A., Sachidanandam, R., Hannon, G. J., & Collins, K. (2009). Sequence, biogenesis, and function of diverse small RNA classes bound to the Piwi family proteins of Tetrahymena thermophila. *Genes & Development*, *23*(17), 2016–2032. https://doi.org/10.1101/gad.1821209

Crooks, G. E., Hon, G., Chandonia, J.-M., & Brenner, S. E. (2004). WebLogo: A sequence logo generator. *Genome Research*, *14*(6), 1188–1190. https://doi.org/10.1101/gr.849004

Czech, B., Malone, C. D., Zhou, R., Stark, A., Schlingeheyde, C., Dus, M., Perrimon, N., Kellis, M., Wohlschlegel, J. A., Sachidanandam, R., Hannon, G. J., & Brennecke, J. (2008). An endogenous small interfering RNA pathway in Drosophila. *Nature*, *453*(7196), 798–802. https://doi.org/10.1038/nature07007

Dadami, E., Moser, M., Zwiebel, M., Krczal, G., Wassenegger, M., & Dalakouras, A. (2013). An endogene-resembling transgene delays the onset of silencing and limits siRNA accumulation. *FEBS Letters*, *587*(6), 706–710. https://doi.org/10.1016/j.febslet.2013.01.045

de Almeida, C., Scheer, H., Gobert, A., Fileccia, V., Martinelli, F., Zuber, H., & Gagliardi, D. (2018). RNA uridylation and decay in plants. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1762), 20180163. https://doi.org/10.1098/rstb.2018.0163

Dmitrieva, T. M., Alexeevski, A. V., Shatskaya, G. S., Tolskaya, E. A., Gmyl, A. P., Khitrina, E. V., & Agol, V. I. (2007). Significance of the C-terminal amino acid residue in mengovirus RNA-dependent RNA polymerase. *Virology*, *365*(1), 79–91. https://doi.org/10.1016/j.virol.2007.02.038

Drews, F., Karunanithi, S., Götz, U., Marker, S., deWijn, R., Pirritano, M., Rodrigues-Viana, A. M., Jung, M., Gasparoni, G., Schulz, M. H., & Simon, M. (2021). Two Piwis with Ago-like functions silence somatic genes at the chromatin level. *RNA Biology*, *18*(sup2), 757–769. https://doi.org/10.1080/15476286.2021.1991114

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

Ewels, P., Magnusson, M., Lundin, S., & Käller, M. (2016). MultiQC: Summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*, *32*(19), 3047–3048. https://doi.org/10.1093/bioinformatics/btw354

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

Feinberg, E. H., & Hunter, C. P. (2003). Transport of dsRNA into Cells by the Transmembrane Protein SID-1. *Science*, *301*(5639), 1545–1547. https://doi.org/10.1126/science.1087117

Feng, Z., Ma, X., Wu, X., Wu, W., Shen, B., Li, S., Tang, Y., Wang, J., Shao, C., & Meng, Y. (2024). Genome-wide identification of phasiRNAs in Arabidopsis thaliana, and insights into biogenesis, temperature sensitivity, and organ specificity. *Plant, Cell & Environment*, *n/a*(n/a). https://doi.org/10.1111/pce.14974

Ferrer-Orta, C., Arias, A., Perez-Luque, R., Escarmís, C., Domingo, E., & Verdaguer, N. (2004). Structure of Foot-and-Mouth Disease Virus RNA-dependent RNA Polymerase and Its Complex with a Template-Primer RNA *. *Journal of Biological Chemistry*, *279*(45), 47212–47221. https://doi.org/10.1074/jbc.M405465200

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

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

Frapporti, A., Miró Pina, C., Arnaiz, O., Holoch, D., Kawaguchi, T., Humbert, A., Eleftheriou, E., Lombard, B., Loew, D., Sperling, L., Guitot, K., Margueron, R., & Duharcourt, S. (2019). The Polycomb protein Ezl1 mediates H3K9 and H3K27 methylation to repress transposable elements in Paramecium. *Nature Communications*, *10*(1), 2710. https://doi.org/10.1038/s41467-019-10648-5

Frolows, N., & Ashe, A. (2021). Small RNAs and chromatin in the multigenerational epigenetic landscape of Caenorhabditis elegans | Philosophical Transactions of the Royal Society B: Biological Sciences. *Philosophical Transactions of The Royal Society B*, *376*(1826). https://royalsocietypublishing.org/doi/full/10.1098/rstb.2020.0112

Furrer, D. I., Swart, E. C., Kraft, M. F., Sandoval, P. Y., & Nowacki, M. (2017). Two Sets of Piwi Proteins Are Involved in Distinct sRNA Pathways Leading to Elimination of Germline-Specific DNA. *Cell Reports*, *20*(2), 505–520. https://doi.org/10.1016/j.celrep.2017.06.050

Galvani, A., & Sperling, L. (2001). Transgene-mediated post-transcriptional gene silencing is inhibited by 3′ non-coding sequences in Paramecium. *Nucleic Acids Research*, *29*(21), 4387–4394. https://doi.org/10.1093/nar/29.21.4387

Galvani, A., & Sperling, L. (2002). RNA interference by feeding in Paramecium. *Trends in Genetics*, *18*(1), 11–12. https://doi.org/10.1016/S0168-9525(01)02548-3

Gao, Y., Solberg, T., Wang, R., Yu, Y., Al-Rasheid, K. A. S., & Gao, F. (2024). Application of RNA interference and protein localization to investigate housekeeping and developmentally regulated genes in the emerging model protozoan Paramecium caudatum. *Communications Biology*, *7*(1), 1–11. https://doi.org/10.1038/s42003-024-05906-2

Garnier, O., Serrano, V., Duharcourt, S., & Meyer, E. (2004). RNA-Mediated Programming of Developmental Genome Rearrangements in Paramecium tetraurelia. *Molecular and Cellular Biology*, *24*(17), 7370–7379. https://doi.org/10.1128/MCB.24.17.7370-7379.2004

Goodier, J. L., Cheung, L. E., & Jr, H. H. K. (2012). MOV10 RNA Helicase Is a Potent Inhibitor of Retrotransposition in Cells. *PLOS Genetics*, *8*(10), e1002941. https://doi.org/10.1371/journal.pgen.1002941

Götz, U., Marker, S., Cheaib, M., Andresen, K., Shrestha, S., Durai, D. A., Nordström, K. J., Schulz, M. H., & Simon, M. (2016). Two sets of RNAi components are required for heterochromatin formation in trans triggered by truncated transgenes. *Nucleic Acids Research*, *44*(12), 5908–5923. https://doi.org/10.1093/nar/gkw267

Gout, J.-F., Hao, Y., Johri, P., Arnaiz, O., Doak, T. G., Bhullar, S., Couloux, A., Guérin, F., Malinsky, S., Potekhin, A., Sawka, N., Sperling, L., Labadie, K., Meyer, E., Duharcourt, S., & Lynch, M. (2023). Dynamics of Gene Loss following Ancient Whole-Genome Duplication in the Cryptic Paramecium Complex. *Molecular Biology and Evolution*, *40*(5), msad107. https://doi.org/10.1093/molbev/msad107

Govind, K., & Savithri, H. S. (2010). Primer-independent initiation of RNA synthesis by *SeMV* recombinant RNA-dependent RNA polymerase. *Virology*, *401*(2), 280–292. https://doi.org/10.1016/j.virol.2010.02.025

Grattepanche, J.-D., Walker, L. M., Ott, B. M., Paim Pinto, D. L., Delwiche, C. F., Lane, C. E., & Katz, L. A. (2018). Microbial Diversity in the Eukaryotic SAR Clade: Illuminating the Darkness Between Morphology and Molecular Data. *BioEssays*, *40*(4), 1700198. https://doi.org/10.1002/bies.201700198

Greider, C. W., & Blackburn, E. H. (1985). Identification of a specific telomere terminal transferase activity in tetrahymena extracts. *Cell*, *43*(2, Part 1), 405–413. https://doi.org/10.1016/0092-8674(85)90170-9

Gruchota, J., Denby Wilkes, C., Arnaiz, O., Sperling, L., & Nowak, J. K. (2017). A meiosis-specific Spt5 homolog involved in non-coding transcription. *Nucleic Acids Research*, *45*(8), 4722–4732. https://doi.org/10.1093/nar/gkw1318

Gu, W., Shirayama, M., Conte, D., Vasale, J., Batista, P. J., Claycomb, J. M., Moresco, J. J., Youngman, E. M., Keys, J., Stoltz, M. J., Chen, C.-C. G., Chaves, D. A., Duan, S., Kasschau, K. D., Fahlgren, N., Yates, J. R., Mitani, S., Carrington, J. C., & Mello, C. C. (2009). Distinct

Argonaute-Mediated 22G-RNA Pathways Direct Genome Surveillance in the C. elegans Germline. *Molecular Cell*, *36*(2), 231–244. https://doi.org/10.1016/j.molcel.2009.09.020

Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N., & Altman, S. (1983). The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell*, *35*(3), 849–857. https://doi.org/10.1016/0092-8674(83)90117-4

Guo, M., & Wu, Y. (2013). Fighting an old war with a new weapon—Silencing transposons by Piwi-interacting RNA. *IUBMB Life*, *65*(9), 739–747. https://doi.org/10.1002/iub.1192

Hallgren, J., Tsirigos, K. D., Pedersen, M. D., Armenteros, J. J. A., Marcatili, P., Nielsen, H., Krogh, A., & Winther, O. (2022). *DeepTMHMM predicts alpha and beta transmembrane proteins using deep neural networks* (p. 2022.04.08.487609). bioRxiv. https://doi.org/10.1101/2022.04.08.487609

Hammond, S. M., Caudy, A. A., & Hannon, G. J. (2001). Post-transcriptional gene silencing by double-stranded RNA. *Nature Reviews Genetics*, *2*(2), 110–119. https://doi.org/10.1038/35052556

Han, B. W., Hung, J.-H., Weng, Z., Zamore, P. D., & Ameres, S. L. (2011). The 3′-to-5′ Exoribonuclease Nibbler Shapes the 3′ Ends of MicroRNAs Bound to Drosophila Argonaute1. *Current Biology*, *21*(22), 1878–1887. https://doi.org/10.1016/j.cub.2011.09.034

Han, T., Manoharan, A. P., Harkins, T. T., Bouffard, P., Fitzpatrick, C., Chu, D. S., Thierry-Mieg, D., Thierry-Mieg, J., & Kim, J. K. (2009). 26G endo-siRNAs regulate spermatogenic and zygotic gene expression in Caenorhabditis elegans. *Proceedings of the National Academy of Sciences*, *Vol. 106*(No. 44). https://doi.org/10.1073/pnas.0906378106

Hoehener, C., Hug, I., & Nowacki, M. (2018). Dicer-like Enzymes with Sequence Cleavage Preferences. *Cell*, *173*(1), 234-247.e7. https://doi.org/10.1016/j.cell.2018.02.029

Huang, X., Tóth, K. F., & Aravin, A. A. (2017). piRNA Biogenesis in Drosophila melanogaster. *Trends in Genetics*, *33*(11), 882–894. https://doi.org/10.1016/j.tig.2017.09.002

Hunter, C. P., Winston, W. M., Molodowitch, C., Feinberg, E. H., Shih, J., Sutherlin, M., Wright, A. J., & Fitzgerald, M. C. (2006). Systemic RNAi in Caenorhabditis elegans. *Cold Spring Harbor Symposia on Quantitative Biology*, *71*, 95–100. https://doi.org/10.1101/sqb.2006.71.060

Hutvágner, G., McLachlan, J., Pasquinelli, A. E., Bálint, É., Tuschl, T., & Zamore, P. D. (2001). A Cellular Function for the RNA-Interference Enzyme Dicer in the Maturation of the let-7 Small Temporal RNA. *Science*, *293*(5531), 834–838. https://doi.org/10.1126/science.1062961

Hyde, S. J., Eckenroth, B. E., Smith, B. A., Eberley, W. A., Heintz, N. H., Jackman, J. E., & Doublié, S. (2010). tRNAHis guanylyltransferase (THG1), a unique 3′-5′ nucleotidyl transferase, shares unexpected structural homology with canonical 5′-3′ DNA polymerases. *Proceedings of the National Academy of Sciences*, *107*(47), 20305–20310. https://doi.org/10.1073/pnas.1010436107

Ibrahim, F., Rymarquis, L. A., Kim, E.-J., Becker, J., Balassa, E., Green, P. J., & Cerutti, H. (2010). Uridylation of mature miRNAs and siRNAs by the MUT68 nucleotidyltransferase promotes their degradation in Chlamydomonas. *Proceedings of the National Academy of Sciences*, *107*(8), 3906–3911. https://doi.org/10.1073/pnas.0912632107

Iwakawa, H., & Tomari, Y. (2022). Life of RISC: Formation, action, and degradation of RNA-induced silencing complex. *Molecular Cell*, *82*(1), 30–43. https://doi.org/10.1016/j.molcel.2021.11.026

Jenkins, B. H., Maguire, F., Leonard, G., Eaton, J. D., West, S., Housden, B. E., Milner, D. S., & Richards, T. A. (2021). Characterization of the RNA-interference pathway as a tool for reverse genetic analysis in the nascent phototrophic endosymbiosis, Paramecium bursaria. *Royal Society Open Science*, *8*(4), 210140. https://doi.org/10.1098/rsos.210140

Jones, M. H., Hamana, N., Nezu, J., & Shimane, M. (2000). A Novel Family of Bromodomain Genes. *Genomics*, *63*(1), 40–45. https://doi.org/10.1006/geno.1999.6071

Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., Pesseat, S., Quinn, A. F., Sangrador-Vegas, A., Scheremetjew, M., Yong, S.-Y., Lopez, R., & Hunter, S. (2014). InterProScan 5: Genome-scale protein function classification. *Bioinformatics*, *30*(9), 1236–1240. https://doi.org/10.1093/bioinformatics/btu031

Juliano, C., Wang, J., & Lin, H. (2011). Uniting Germline and Stem Cells: The Function of Piwi Proteins and the piRNA Pathway in Diverse Organisms. *Annual Review of Genetics*, *45*(Volume 45, 2011), 447–469. https://doi.org/10.1146/annurev-genet-110410-132541

Kamminga, L. M., Luteijn, M. J., den Broeder, M. J., Redl, S., Kaaij, L. J. T., Roovers, E. F., Ladurner, P., Berezikov, E., & Ketting, R. F. (2010). Hen1 is required for oocyte development

and piRNA stability in zebrafish. *The EMBO Journal*, *29*(21), 3688–3700. https://doi.org/10.1038/emboj.2010.233

Kanno, T., & Habu, Y. (2011). siRNA-mediated chromatin maintenance and its function in *Arabidopsis thaliana*. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, *1809*(8), 444–451. https://doi.org/10.1016/j.bbagrm.2011.05.002

Karunanithi, S., Oruganti, V., de Wijn, R., Drews, F., Cheaib, M., Nordström, K., Simon, M., & Schulz, M. H. (2020). Feeding exogenous dsRNA interferes with endogenous sRNA accumulation in Paramecium. *DNA Research*, *27*(1), dsaa005. https://doi.org/10.1093/dnares/dsaa005

Karunanithi, S., Oruganti, V., Marker, S., Rodriguez-Viana, A. M., Drews, F., Pirritano, M., Nordström, K., Simon, M., & Schulz, M. H. (2019). Exogenous RNAi mechanisms contribute to transcriptome adaptation by phased siRNA clusters in Paramecium. *Nucleic Acids Research*, *47*(15), 8036–8049. https://doi.org/10.1093/nar/gkz553

Karunanithi, S., Simon, M., & Schulz, M. H. (2019). Automated analysis of small RNA datasets with RAPID. *PeerJ*, *7*, e6710. https://doi.org/10.7717/peerj.6710

Katoh, T., Sakaguchi, Y., Miyauchi, K., Suzuki, T., Kashiwabara, S., Baba, T., & Suzuki, T. (2009). Selective stabilization of mammalian microRNAs by 3′ adenylation mediated by the cytoplasmic poly(A) polymerase GLD-2. *Genes & Development*, *23*(4), 433–438. https://doi.org/10.1101/gad.1761509

Kelley, L. H., Caldas, I. V., Sullenberger, M. T., Yongblah, K. E., Niazi, A. M., Iyer, A., Li, Y., Tran, P. M., Valen, E., Ahmed-Braimah, Y. H., & Maine, E. M. (2024). Poly(U) polymerase activity in Caenorhabditis elegans regulates abundance and tailing of sRNA and mRNA. *Genetics*, iyae120. https://doi.org/10.1093/genetics/iyae120

Kim, V. N. (2005). MicroRNA biogenesis: Coordinated cropping and dicing. *Nature Reviews Molecular Cell Biology*, *6*(5), 376–385. https://doi.org/10.1038/nrm1644

Kirchdoerfer, R. N., & Ward, A. B. (2019). Structure of the SARS-CoV nsp12 polymerase bound to nsp7 and nsp8 co-factors. *Nature Communications*, *10*(1), 2342. https://doi.org/10.1038/s41467-019-10280-3

Knoll, G., Haacke-Bell, B., & Plattner, H. (1991). Local trichocyst exocytosis provides an efficient escape mechanism for Paramecium cells. *European Journal of Protistology*, *27*(4), 381–385. https://doi.org/10.1016/S0932-4739(11)80256-7

Kouzarides, T. (2007). Chromatin Modifications and Their Function. *Cell*, *128*(4), 693–705. https://doi.org/10.1016/j.cell.2007.02.005

Krueger, F., James, F., Ewels, P., Afyounian, E., Weinstein, M., Schuster-Boeckler, B., Hulselmans, G., & sclamons. (2023). *FelixKrueger/TrimGalore: V0.6.10 - https://github.com/FelixKrueger/TrimGalore* (Version 0.6.10) [Computer software]. Zenodo. https://doi.org/10.5281/zenodo.7598955

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*(4), 357–359. https://doi.org/10.1038/nmeth.1923

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

Lee, D., Park, D., Park, J. H., Kim, J. H., & Shin, C. (2019). Poly(A)-specific ribonuclease sculpts the 3′ ends of microRNAs. *RNA*, *25*(3), 388–405. https://doi.org/10.1261/rna.069633.118

Lee, R. C., Hammell, C. M., & Ambros, V. (2006). Interacting endogenous and exogenous RNAi pathways in Caenorhabditis elegans. *RNA*, *12*(4), 589–597. https://doi.org/10.1261/rna.2231506

Lee, S. R., & Collins, K. (2006). Two classes of endogenous small RNAs in Tetrahymena thermophila. *Genes & Development*, *20*(1), 28–33. https://doi.org/10.1101/gad.1377006

Lee, S. R., & Collins, K. (2007). Physical and functional coupling of RNA-dependent RNA polymerase and Dicer in the biogenesis of endogenous siRNAs. *Nature Structural & Molecular Biology*, *14*(7), 604–610. https://doi.org/10.1038/nsmb1262

Lee, S. R., Talsky, K. B., & Collins, K. (2009). A single RNA-dependent RNA polymerase assembles with mutually exclusive nucleotidyl transferase subunits to direct different pathways of small RNA biogenesis. *RNA*, *15*(7), 1363–1374. https://doi.org/10.1261/rna.1630309

Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Rådmark, O., Kim, S., & Kim, V. N. (2003). The nuclear RNase III Drosha initiates microRNA processing. *Nature*, *425*(6956), 415–419. https://doi.org/10.1038/nature01957

Lepère, G., Bétermier, M., Meyer, E., & Duharcourt, S. (2008). Maternal noncoding transcripts antagonize the targeting of DNA elimination by scanRNAs in Paramecium tetraurelia. *Genes & Development*, *22*(11), 1501–1512. https://doi.org/10.1101/gad.473008

Lepère, G., Nowacki, M., Serrano, V., Gout, J.-F., Guglielmi, G., Duharcourt, S., & Meyer, E. (2009). Silencing-associated and meiosis-specific small RNA pathways in Paramecium tetraurelia. *Nucleic Acids Research*, *37*(3), 903–915. https://doi.org/10.1093/nar/gkn1018

Lévêque, V. J.-P., Johnson, R. B., Parsons, S., Ren, J., Xie, C., Zhang, F., & Wang, Q. M. (2003). Identification of a C-Terminal Regulatory Motif in Hepatitis C Virus RNA-Dependent RNA Polymerase: Structural and Biochemical Analysis. *Journal of Virology*, *77*(16), 9020–9028. https://doi.org/10.1128/jvi.77.16.9020-9028.2003

Li, J., Yang, Z., Yu, B., Liu, J., & Chen, X. (2005). Methylation Protects miRNAs and siRNAs from a 3′-End Uridylation Activity in Arabidopsis. *Current Biology*, *15*(16), 1501–1507. https://doi.org/10.1016/j.cub.2005.07.029

Lingel, A., Simon, B., Izaurralde, E., & Sattler, M. (2004). Nucleic acid 3′-end recognition by the Argonaute2 PAZ domain. *Nature Structural & Molecular Biology*, *11*(6), 576–577. https://doi.org/10.1038/nsmb777

Liu, C. L., Kaplan, T., Kim, M., Buratowski, S., Schreiber, S. L., Friedman, N., & Rando, O. J. (2005). Single-Nucleosome Mapping of Histone Modifications in S. cerevisiae. *PLOS Biology*, *3*(10), e328. https://doi.org/10.1371/journal.pbio.0030328

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

Luo, Z., & Chen, Z. (2007). Improperly Terminated, Unpolyadenylated mRNA of Sense Transgenes Is Targeted by RDR6-Mediated RNA Silencing in Arabidopsis. *The Plant Cell*, *19*(3), 943–958. https://doi.org/10.1105/tpc.106.045724

Ma, J.-B., Yuan, Y.-R., Meister, G., Pei, Y., Tuschl, T., & Patel, D. J. (2005). Structural basis for 5′-end-specific recognition of guide RNA by the A. fulgidus Piwi protein. *Nature*, *434*(7033), 666–670. https://doi.org/10.1038/nature03514

MacRae, I. J., Zhou, K., & Doudna, J. A. (2007). Structural determinants of RNA recognition and cleavage by Dicer. *Nature Structural & Molecular Biology*, *14*(10), 934–940. https://doi.org/10.1038/nsmb1293

MacRae, I. J., Zhou, K., Li, F., Repic, A., Brooks, A. N., Cande, W. Z., Adams, P. D., & Doudna, J. A. (2006). Structural Basis for Double-Stranded RNA Processing by Dicer. *Science*, *311*(5758), 195–198. https://doi.org/10.1126/science.1121638

Malik, D., Kobyłecki, K., Krawczyk, P., Poznański, J., Jakielaszek, A., Napiórkowska, A., Dziembowski, A., Tomecki, R., & Nowotny, M. (2020). Structure and mechanism of CutA, RNA nucleotidyl transferase with an unusual preference for cytosine. *Nucleic Acids Research*, *48*(16), 9387–9405. https://doi.org/10.1093/nar/gkaa647

Marker, S., Carradec, Q., Tanty, V., Arnaiz, O., & Meyer, E. (2014). A forward genetic screen reveals essential and non-essential RNAi factors in Paramecium tetraurelia. *Nucleic Acids Research*, *42*(11), 7268–7280. https://doi.org/10.1093/nar/gku223

Marker, S., Le Mouël, A., Meyer, E., & Simon, M. (2010). Distinct RNA-dependent RNA polymerases are required for RNAi triggered by double-stranded RNA versus truncated transgenes in Paramecium tetraurelia. *Nucleic Acids Research*, *38*(12), 4092–4107. https://doi.org/10.1093/nar/gkq131

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

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

Maurer-Stroh, S., Dickens, N. J., Hughes-Davies, L., Kouzarides, T., Eisenhaber, F., & Ponting, C. P. (2003). The Tudor domain 'Royal Family': Tudor, plant Agenet, Chromo, PWWP and MBT domains. *Trends in Biochemical Sciences*, *28*(2), 69–74. https://doi.org/10.1016/S0968-0004(03)00004-5

McEwan, D. L., Weisman, A. S., & Hunter, C. P. (2012). Uptake of Extracellular Double-Stranded RNA by SID-2. *Molecular Cell*, *47*(5), 746–754. https://doi.org/10.1016/j.molcel.2012.07.014

Medley, J. C., Panzade, G., & Zinovyeva, A. Y. (2020). microRNA strand selection: Unwinding the rules. *WIREs RNA*, *12*(3). https://doi.org/10.1002/wrna.1627

Miró-Pina, C., Charmant, O., Kawaguchi, T., Holoch, D., Michaud, A., Cohen, I., Humbert, A., Jaszczyszyn, Y., Chevreux, G., Maestro, L. D., Ait-Si-Ali, S., Arnaiz, O., Margueron, R., & Duharcourt, S. (2022). Paramecium Polycomb repressive complex 2 physically interacts with

the small RNA-binding PIWI protein to repress transposable elements. *Developmental Cell*, *57*(8), 1037-1052.e8. https://doi.org/10.1016/j.devcel.2022.03.014

Mochizuki, K., Fine, N. A., Fujisawa, T., & Gorovsky, M. A. (2002). Analysis of a piwi-Related Gene Implicates Small RNAs in Genome Rearrangement in Tetrahymena. *Cell*, *110*(6), 689–699. https://doi.org/10.1016/S0092-8674(02)00909-1

Moissiard, G., Parizotto, E. A., Himber, C., & Voinnet, O. (2007). Transitivity in Arabidopsis can be primed, requires the redundant action of the antiviral Dicer-like 4 and Dicer-like 2, and is compromised by viral-encoded suppressor proteins. *RNA*, *13*(8), 1268–1278. https://doi.org/10.1261/rna.541307

Montgomery, M. K., & Fire, A. (1998). Double-stranded RNA as a mediator in sequence-specific genetic silencing and co-suppression: Trends in Genetics. *Trends in Genetics*, *14*(7), 255–258. https://doi.org/10.1016/S0168-9525(98)01510-8

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

Najle, S. R., Nusblat, A. D., Nudel, C. B., & Uttaro, A. D. (2013). The Sterol-C7 Desaturase from the Ciliate Tetrahymena thermophila Is a Rieske Oxygenase, Which Is Highly Conserved in Animals. *Molecular Biology and Evolution*, *30*(7), 1630–1643. https://doi.org/10.1093/molbev/mst076

Nakanishi, K., Ascano, M., Gogakos, T., Ishibe-Murakami, S., Serganov, A. A., Briskin, D., Morozov, P., Tuschl, T., & Patel, D. J. (2013). Eukaryote-Specific Insertion Elements Control Human ARGONAUTE Slicer Activity. *Cell Reports*, *3*(6), 1893–1900. https://doi.org/10.1016/j.celrep.2013.06.010

Napoli, C., Lemieux, C., & Jorgensen, R. (1990). Introduction of a Chimeric Chalcone Synthase Gene into Petunia Results in Reversible Co-Suppression of Homologous Genes in trans. *The Plant Cell*, *2*(4), 279–289. https://doi.org/10.1105/tpc.2.4.279

Ng, K. K.-S., Pendás-Franco, N., Rojo, J., Boga, J. A., Machín, À., Alonso, J. M. M., & Parra, F. (2004). Crystal Structure of Norwalk Virus Polymerase Reveals the Carboxyl Terminus in the Active Site Cleft *. *Journal of Biological Chemistry*, *279*(16), 16638–16645. https://doi.org/10.1074/jbc.M400584200

Ni, C., & Hong, M. (2024). Oligomerization of drug transporters: Forms, functions, and mechanisms. *Acta Pharmaceutica Sinica B*, *14*(5), 1924–1938. https://doi.org/10.1016/j.apsb.2024.01.007

Nowacki, M., & Landweber, L. F. (2009). Epigenetic inheritance in ciliates. *Current Opinion in Microbiology*, *12*(6), 638–643. https://doi.org/10.1016/j.mib.2009.09.012

Oh, J.-W., Ito, T., & Lai, M. M. C. (1999). A Recombinant Hepatitis C Virus RNA-Dependent RNA Polymerase Capable of Copying the Full-Length Viral RNA. *Journal of Virology*, *73*(9), 7694–7702. https://doi.org/10.1128/jvi.73.9.7694-7702.1999

Oliveros, J. C. (2007). *Venny. An Interactive Tool for Comparing Lists with Venn's Diagrams.* [Computer software]. https://bioinfogp.cnb.csic.es/tools/venny/index.html

Pak, J., & Fire, A. (2007). Distinct Populations of Primary and Secondary Effectors During RNAi in C. elegans. *Science*, *315*(5809), 241–244. https://doi.org/10.1126/science.1132839

Pak, J., Maniar, J. M., Mello, C. C., & Fire, A. (2012). Protection from Feed-Forward Amplification in an Amplified RNAi Mechanism. *Cell*, *151*(4), 885–899. https://doi.org/10.1016/j.cell.2012.10.022

Paramecium (with Görtz, H.-D., Adoutte, A., & Preer, J. R. J.). (2013). Springer Berlin.

Park, J.-E., Heo, I., Tian, Y., Simanshu, D. K., Chang, H., Jee, D., Patel, D. J., & Kim, V. N. (2011). Dicer recognizes the 5′ end of RNA for efficient and accurate processing. *Nature*, *475*(7355), 201–205. https://doi.org/10.1038/nature10198

Parker, J. S., Roe, S. M., & Barford, D. (2004). Crystal structure of a PIWI protein suggests mechanisms for siRNA recognition and slicer activity. *The EMBO Journal*, *23*(24), 4727–4737. https://doi.org/10.1038/sj.emboj.7600488

Parker, J. S., Roe, S. M., & Barford, D. (2005). Structural insights into mRNA recognition from a PIWI domain–siRNA guide complex. *Nature*, *434*(7033), 663–666. https://doi.org/10.1038/nature03462

Pastore, B., Hertz, H. L., Price, I. F., & Tang, W. (2021). Pre-piRNA trimming and 2′-O-methylation protect piRNAs from 3′ tailing and degradation in C. elegans. *Cell Reports*, *36*(9). https://doi.org/10.1016/j.celrep.2021.109640

Petersen, B. O., & Albrechtsen, M. (2005). Evidence implying only unprimed RdRP activity during transitive gene silencing in plants. *Plant Molecular Biology*, *58*(4), 575–583. https://doi.org/10.1007/s11103-005-7307-4

Pitsillou, E., Liang, J., Yu Meng Huang, H., Hung, A., & Karagiannis, T. C. (2021). *In silico* investigation to identify potential small molecule inhibitors of the RNA-dependent RNA polymerase (RdRp) nidovirus RdRp-associated nucleotidyltransferase domain. *Chemical Physics Letters*, *779*, 138889. https://doi.org/10.1016/j.cplett.2021.138889

Pokholok, D. K., Harbison, C. T., Levine, S., Cole, M., Hannett, N. M., Lee, T. I., Bell, G. W., Walker, K., Rolfe, P. A., Herbolsheimer, E., Zeitlinger, J., Lewitter, F., Gifford, D. K., & Young, R. A. (2005). Genome-wide Map of Nucleosome Acetylation and Methylation in Yeast. *Cell*, *122*(4), 517–527. https://doi.org/10.1016/j.cell.2005.06.026

Pumplin, N., & Voinnet, O. (2013). RNA silencing suppression by plant pathogens: Defence, counter-defence and counter-counter-defence. *Nature Reviews Microbiology*, *11*(11), 745–760. https://doi.org/10.1038/nrmicro3120

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

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

Saito, K., Nishida, K. M., Mori, T., Kawamura, Y., Miyoshi, K., Nagami, T., Siomi, H., & Siomi, M. C. (2006). Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the Drosophila genome. *Genes & Development*, *20*, 2214–2222.

Saleh, M.-C., van Rij, R. P., Hekele, A., Gillis, A., Foley, E., O'Farrell, P. H., & Andino, R. (2006). The endocytic pathway mediates cell entry of dsRNA to induce RNAi silencing. *Nature Cell Biology*, *8*(8), 793–802. https://doi.org/10.1038/ncb1439

Sanan-Mishra, N., Abdul Kader Jailani, A., Mandal, B., & Mukherjee, S. K. (2021). Secondary siRNAs in Plants: Biosynthesis, Various Functions, and Applications in Virology. *Frontiers in Plant Science*, *12*. https://doi.org/10.3389/fpls.2021.610283

Sandoval, P. Y., Swart, E. C., Arambasic, M., & Nowacki, M. (2014). Functional Diversification of Dicer-like Proteins and Small RNAs Required for Genome Sculpting. *Developmental Cell*, *28*(2), 174–188. https://doi.org/10.1016/j.devcel.2013.12.010

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

Scott, D. D., & Norbury, C. J. (2013). RNA decay *via* 3′ uridylation. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, *1829*(6), 654–665. https://doi.org/10.1016/j.bbagrm.2013.01.009

Shih, J. D., Fitzgerald, M. C., Sutherlin, M., & Hunter, C. P. (2009). The SID-1 double-stranded RNA transporter is not selective for dsRNA length. *RNA*, *15*(3), 384–390. https://doi.org/10.1261/rna.1286409

Shih, J. D., & Hunter, C. P. (2011). SID-1 is a dsRNA-selective dsRNA-gated channel. *RNA*, *17*(6), 1057–1065. https://doi.org/10.1261/rna.2596511

Simon, M., & Plattner, H. (2014). Chapter Three - Unicellular Eukaryotes as Models in Cell and Molecular Biology: Critical Appraisal of Their Past and Future Value. In K. W. Jeon (Ed.), *International Review of Cell and Molecular Biology* (Vol. 309, pp. 141–198). Academic Press. https://doi.org/10.1016/B978-0-12-800255-1.00003-X

Singh, A., Maurer-Alcalá, X. X., Solberg, T., Häußermann, L., Gisler, S., Ignarski, M., Swart, E. C., & Nowacki, M. (2022). Chromatin remodeling is required for sRNA-guided DNA elimination in Paramecium. *The EMBO Journal*, *41*(22), e111839. https://doi.org/10.15252/embj.2022111839

Solberg, T., Mason, V., Wang, C., & Nowacki, M. (2023). Developmental mRNA clearance by PIWI-bound endo-siRNAs in Paramecium. *Cell Reports*, *42*(3). https://doi.org/10.1016/j.celrep.2023.112213

Song, J.-J., Smith, S. K., Hannon, G. J., & Joshua-Tor, L. (2004). Crystal Structure of Argonaute and Its Implications for RISC Slicer Activity. *Science*, *305*(5689), 1434–1437. https://doi.org/10.1126/science.1102514

Song, X.-S., Gu, K.-X., Duan, X.-X., Xiao, X.-M., Hou, Y.-P., Duan, Y.-B., Wang, J.-X., & Zhou, M.-G. (2018). A *myosin5* dsRNA that reduces the fungicide resistance and pathogenicity

of *Fusarium asiaticum*. *Pesticide Biochemistry and Physiology*, *150*, 1–9. https://doi.org/10.1016/j.pestbp.2018.07.004

Sonneborn, T. M. (1938). The delayed occurrence and total omission of endomixis in selected lines of paramecium aurelia. *The Biological Bulletin*, *74*(1), 76–82. https://doi.org/10.2307/1537886

Sonneborn, T. M. (1939). Paramecium aurelia: Mating Types and Groups; Lethal Interactions; Determination and Inheritance. *The American Naturalist*, *73*(748), 390–413. https://doi.org/10.1086/280850

Sonneborn, T. M. (1949). Beyond the gene. *American Scientist*, *37*(1), 33–59.

Sonneborn, T. M., & Lesuer, A. (1948). Antigenic Characters in Paramecium aurelia (Variety 4): Determination, Inheritance and Induced Mutation. *The American Naturalist*, *82*(802), 69–78. https://doi.org/10.1086/281566

Spracklin, G., Fields, B., Wan, G., Becker, D., Wallig, A., Shukla, A., & Kennedy, S. (2017). The RNAi Inheritance Machinery of Caenorhabditis elegans. *Genetics*, *206*(3), 1403–1416. https://doi.org/10.1534/genetics.116.198812

Stasser, J. P., Eisses, J. F., Barry, A. N., Kaplan, J. H., & Blackburn, N. J. (2005). Cysteine-to-Serine Mutants of the Human Copper Chaperone for Superoxide Dismutase Reveal a Copper Cluster at a Domain III Dimer Interface. *Biochemistry*, *44*(9), 3143–3152. https://doi.org/10.1021/bi0478392

Stein, C. B., Genzor, P., Mitra, S., Elchert, A. R., Ipsaro, J. J., Benner, L., Sobti, S., Su, Y., Hammell, M., Joshua-Tor, L., & Haase, A. D. (2019). Decoding the 5′ nucleotide bias of PIWI-interacting RNAs. *Nature Communications*, *10*(1), 828. https://doi.org/10.1038/s41467-019-08803-z

Steiner, F. A., Okihara, K. L., Hoogstrate, S. W., Sijen, T., & Ketting, R. F. (2009). RDE-1 slicer activity is required only for passenger-strand cleavage during RNAi in Caenorhabditis elegans. *Nature Structural & Molecular Biology*, *16*(2), 207–211. https://doi.org/10.1038/nsmb.1541

Sterner, D. E., & Berger, S. L. (2000). Acetylation of Histones and Transcription-Related Factors. *Microbiology and Molecular Biology Reviews*, *64*(2), 435–459. https://doi.org/10.1128/mmbr.64.2.435-459.2000

Stover, N. A., Punia, R. S., Bowen, M. S., Dolins, S. B., & Clark, T. G. (2012). Tetrahymena genome database Wiki: A community-maintained model organism database. *Database*, *2012*, bas007. https://doi.org/10.1093/database/bas007

Suzuki, H. I., Katsura, A., Yasuda, T., Ueno, T., Mano, H., Sugimoto, K., & Miyazono, K. (2015). Small-RNA asymmetry is directly driven by mammalian Argonautes. *Nature Structural & Molecular Biology*, *22*(7), 512–521. https://doi.org/10.1038/nsmb.3050

Szostak, J. W., & Blackburn, E. H. (1982). Cloning yeast telomeres on linear plasmid vectors. *Cell*, *29*(1), 245–255. https://doi.org/10.1016/0092-8674(82)90109-X

Talsky, K. B., & Collins, K. (2010). Initiation by a Eukaryotic RNA-dependent RNA Polymerase Requires Looping of the Template End and Is Influenced by the Template-tailing Activity of an Associated Uridyltransferase *. *Journal of Biological Chemistry*, *285*(36), 27614–27623. https://doi.org/10.1074/jbc.M110.142273

Tanaka, T., Yano, T., Usuki, S., Seo, Y., Mizuta, K., Okaguchi, M., Yamaguchi, M., Hanyu-Nakamura, K., Toyama-Sorimachi, N., Brückner, K., & Nakamura, A. (2024). Endocytosed dsRNAs induce lysosomal membrane permeabilization that allows cytosolic dsRNA translocation for Drosophila RNAi responses. *Nature Communications*, *15*(1), 6993. https://doi.org/10.1038/s41467-024-51343-4

Tang, G., Reinhart, B. J., Bartel, D. P., & Zamore, P. D. (2003). A biochemical framework for RNA silencing in plants. *Genes & Development*, *17*(1), 49–63. https://doi.org/10.1101/gad.1048103

te Velthuis, A. J. W., Arnold, J. J., Cameron, C. E., van den Worm, S. H. E., & Snijder, E. J. (2010). The RNA polymerase activity of SARS-coronavirus nsp12 is primer dependent. *Nucleic Acids Research*, *38*(1), 203–214. https://doi.org/10.1093/nar/gkp904

te Velthuis, A. J. W., van den Worm, S. H. E., & Snijder, E. J. (2012). The SARS-coronavirus nsp7+nsp8 complex is a unique multimeric RNA polymerase capable of both de novo initiation and primer extension. *Nucleic Acids Research*, *40*(4), 1737–1747. https://doi.org/10.1093/nar/gkr893

Thomas, M. F., L'Etoile, N. D., & Ansel, K. M. (2014). Eri1: A conserved enzyme at the crossroads of multiple RNA-processing pathways. *Trends in Genetics*, *30*(7), 298–307. https://doi.org/10.1016/j.tig.2014.05.003

Timmons, L., Court, D. L., & Fire, A. (2001). Ingestion of bacterially expressed dsRNAs can produce specific and potent genetic interference in *Caenorhabditis elegans*. *Gene*, *263*(1), 103–112. https://doi.org/10.1016/S0378-1119(00)00579-5

Timmons, L., & Fire, A. (1998). Specific interference by ingested dsRNA. *Nature*, *395*(6705), 854–854. https://doi.org/10.1038/27579

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

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

Varley, A. J., Hammill, M. L., Salim, L., & Desaulniers, J.-P. (2020). Effects of Chemical Modifications on siRNA Strand Selection in Mammalian Cells. *Nucleic Acid Therapeutics*, *30*(4), 229–236. https://doi.org/10.1089/nat.2020.0848

Vergani-Junior, C. A., Tonon-da-Silva, G., Inan, M. D., & Mori, M. A. (2021). DICER: Structure, function, and regulation. *Biophysical Reviews*, *13*(6), 1081–1090. https://doi.org/10.1007/s12551-021-00902-w

Vettese-Dadey, M., Grant, P. A., Hebbes, T. R., Crane- Robinson, C., Allis, C. D., & Workman, J. L. (1996). Acetylation of histone H4 plays a primary role in enhancing transcription factor binding to nucleosomal DNA in vitro. *The EMBO Journal*, *15*(10), 2508–2518. https://doi.org/10.1002/j.1460-2075.1996.tb00608.x

Volpe, T. A., Kidner, C., Hall, I. M., Teng, G., Grewal, S. I. S., & Martienssen, R. A. (2002). Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science (New York, N.Y.)*, *297*(5588), 1833–1837. https://doi.org/10.1126/science.1074973

Wang, X., Zhang, S., Dou, Y., Zhang, C., Chen, X., Yu, B., & Ren, G. (2015). Synergistic and Independent Actions of Multiple Terminal Nucleotidyl Transferases in the 3' Tailing of Small RNAs in Arabidopsis. *PLOS Genetics*, *11*(4), e1005091. https://doi.org/10.1371/journal.pgen.1005091

Wang, Z., Qiu, Y., Liu, Y., Qi, N., Si, J., Xia, X., Wu, D., Hu, Y., & Zhou, X. (2013). Characterization of a Nodavirus Replicase Revealed a de Novo Initiation Mechanism of RNA

Synthesis and Terminal Nucleotidyltransferase Activity *. *Journal of Biological Chemistry*, *288*(43), 30785–30801. https://doi.org/10.1074/jbc.M113.492728

Wang, Z., Zang, C., Rosenfeld, J. A., Schones, D. E., Barski, A., Cuddapah, S., Cui, K., Roh, T.-Y., Peng, W., Zhang, M. Q., & Zhao, K. (2008). Combinatorial patterns of histone acetylations and methylations in the human genome. *Nature Genetics*, *40*(7), 897–903. https://doi.org/10.1038/ng.154

Wenda, J. M., Homolka, D., Yang, Z., Spinelli, P., Sachidanandam, R., Pandey, R. R., & Pillai, R. S. (2017). Distinct Roles of RNA Helicases MVH and TDRD9 in PIWI Slicing-Triggered Mammalian piRNA Biogenesis and Function. *Developmental Cell*, *41*(6), 623-637.e9. https://doi.org/10.1016/j.devcel.2017.05.021

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

Winston, W. M., Molodowitch, C., & Hunter, C. P. (2002). Systemic RNAi in C. elegans Requires the Putative Transmembrane Protein SID-1. *Science*, *295*(5564), 2456–2459. https://doi.org/10.1126/science.1068836

Winston, W. M., Sutherlin, M., Wright, A. J., Feinberg, E. H., & Hunter, C. P. (2007). Caenorhabditis elegans SID-2 is required for environmental RNA interference. *Proceedings of the National Academy of Sciences*, *104*(25), 10565–10570. https://doi.org/10.1073/pnas.0611282104

Wyman, S. K., Knouf, E. C., Parkin, R. K., Fritz, B. R., Lin, D. W., Dennis, L. M., Krouse, M. A., Webster, P. J., & Tewari, M. (2011). Post-transcriptional generation of miRNA variants by multiple nucleotidyl transferases contributes to miRNA transcriptome complexity. *Genome Research*, *21*(9), 1450–1461. https://doi.org/10.1101/gr.118059.110

Xie, W., Sowemimo, I., Hayashi, R., Wang, J., Burkard, T. R., Brennecke, J., Ameres, S. L., & Patel, D. J. (2020). Structure-function analysis of microRNA 3′-end trimming by Nibbler. *Proceedings of the National Academy of Sciences*, *117*(48), 30370–30379. https://doi.org/10.1073/pnas.2018156117

Xuan, Y. H., Hu, Y. B., Chen, L.-Q., Sosso, D., Ducat, D. C., Hou, B.-H., & Frommer, W. B. (2013). Functional role of oligomerization for bacterial and plant SWEET sugar transporter

family. *Proceedings of the National Academy of Sciences*, *110*(39), E3685–E3694. https://doi.org/10.1073/pnas.1311244110

Yi, R., Doehle, B. P., Qin, Y., Macara, I. G., & Cullen, B. R. (2005). Overexpression of Exportin 5 enhances RNA interference mediated by short hairpin RNAs and microRNAs. *RNA*, *11*(2), 220–226. https://doi.org/10.1261/rna.7233305

Zambon, R. A., Vakharia, V. N., & Wu, L. P. (2006). RNAi is an antiviral immune response against a dsRNA virus in Drosophila melanogaster. *Cellular Microbiology*, *8*(5), 880–889. https://doi.org/10.1111/j.1462-5822.2006.00688.x

Zaug, A. J., & Cech, T. R. (1986). The Intervening Sequence RNA of Tetrahymena Is an Enzyme. *Science*, *231*(4737), 470–475. https://doi.org/10.1126/science.3941911

Zhang, W.-F., Stephen, P., Thériault, J.-F., Wang, R., & Lin, S.-X. (2020). Novel Coronavirus Polymerase and Nucleotidyl-Transferase Structures: Potential to Target New Outbreaks. *The Journal of Physical Chemistry Letters*, *11*(11), 4430–4435. https://doi.org/10.1021/acs.jpclett.0c00571

Zhang, Y., & Reinberg, D. (2001). Transcription regulation by histone methylation: Interplay between different covalent modifications of the core histone tails. *Genes & Development*, *15*(18), 2343–2360. https://doi.org/10.1101/gad.927301

Zhao, Y., Yu, Y., Zhai, J., Ramachandran, V., Dinh, T. T., Meyers, B. C., Mo, B., & Chen, X. (2012). The Arabidopsis Nucleotidyl Transferase HESO1 Uridylates Unmethylated Small RNAs to Trigger Their Degradation. *Current Biology*, *22*(8), 689–694. https://doi.org/10.1016/j.cub.2012.02.051

Zheng, K., Xiol, J., Reuter, M., Eckardt, S., Leu, N. A., McLaughlin, K. J., Stark, A., Sachidanandam, R., Pillai, R. S., & Wang, P. J. (2010). Mouse MOV10L1 associates with Piwi proteins and is an essential component of the Piwi-interacting RNA (piRNA) pathway. *Proceedings of the National Academy of Sciences*, *107*(26), 11841–11846. https://doi.org/10.1073/pnas.1003953107

Supplementary Material

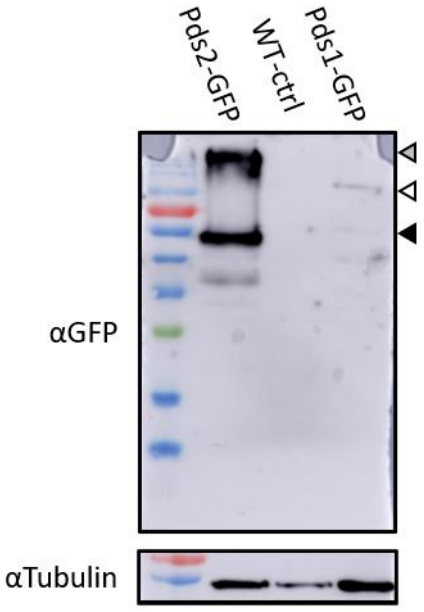

Figure S 1 Western Blot of Pds1 and Pds2 fusion protein. Total protein of transgenic Pds1-GFP, Pds2-GFP or wildtype (WT) culture was blotted and decorated with anti-GFP antibody (αGFP) or anti-alpha-Tubulin antibody (αTubulin) as loading control. Estimated sizes for fusion proteins are: Pds1-GFP: 107kDa (white arrowhead); Pds2-GFP: 51 kDa (black arrowhead) and alpha-Tubulin: 55kDa. The red band of the prestained protein ladder (NEB) corresponds to 72kDa.

Table S 1 NCBI Blast hits for Pds1 without "low complex sequence" filtering.

Hit #	Accession	Description	Organism	Length	Score (Bits)	Identities (%)	Positives (%)	E-value
118	A0A9Q0LUK8	mRNA export factor GLE1	Anaeramoeba ignava	1055	160	24.3	43.9	1.2E-7
221	A0A9Q0LX29	ATP-dependent RNA helicase	Anaeramoeba ignava	1207	152	26.2	42.4	1.1E-6
237	A0A9Q0LVE3	RNA polymerase ii-associated protein	Anaeramoeba ignava	2449	151	20.8	40.3	1.7E-66

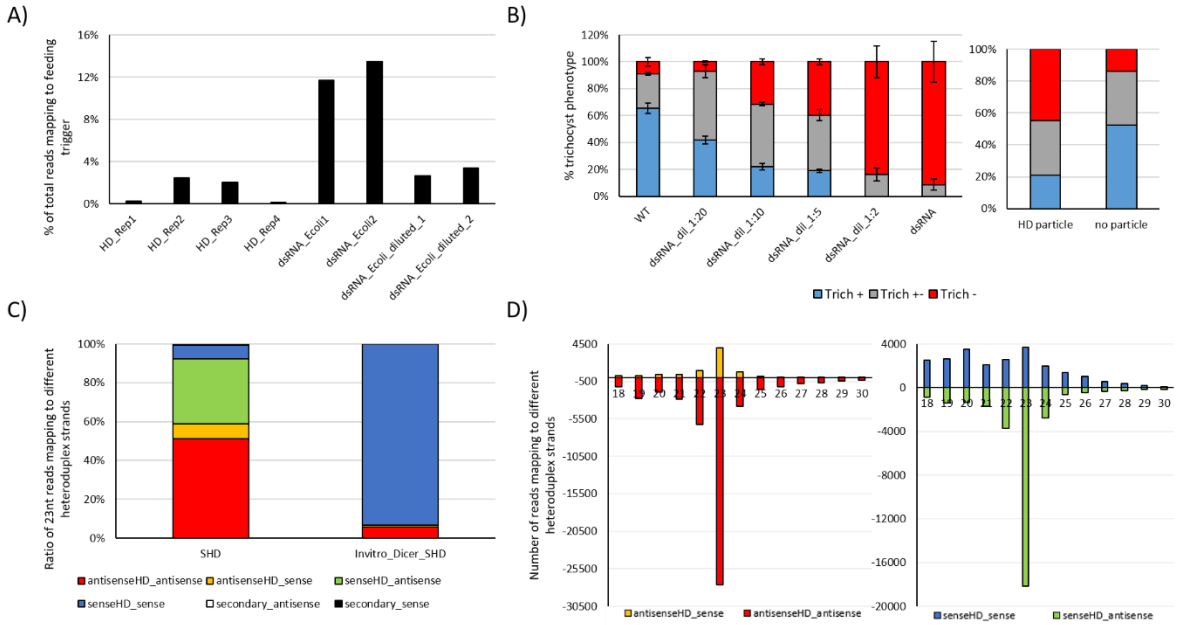

Figure S 2 Phenotype of different dsRNA feeding methods and switched HD smallRNA data. A) Abundance of smallRNAs derived from dsRNA trigger in different experiments normalized to total reads sequenced. Shown are experiments were Heteroduplex dsRNA bound to nanoparticles (HD), dsRNA produced in *E. coli* (dsRNA_*Ecoli*) and dsRNA produced in *E. coli* diluted 1:10 with other food bacteria (dsRNA_*Ecoli*_diluted) were applied to cells. B) Trichocyst phenotype of wildtype cells and cells fed with dsRNA targeting the nd169 gene. dsRNA was delivered either by dsRNA producing *E. coli* diluted in different ratios with other food bacteria (left) or via Heteroduplex nanoparticles (right). C) Ratio of different Heteroduplex derived smallRNA species. Shown are smallRNAs mapping to the original Heteroduplex dsRNA strands (labeled in red and blue), to potential RdRP produced strands (labeled in green and orange), or to the corresponding endogenous sequence without mismatches, representing secondary siRNAs (labeled in black and white). In vitro diced Heteroduplex dsRNA served as a control. D) Mean read length distribution of smallRNAs derived from Heteroduplex dsRNA across all four replicates. smallRNAs corresponding to the different possible Heteroduplex strands are color coded as described above. Positive values correspond to sense-directed molecules whereas negative values correspond to antisense-directed molecules.

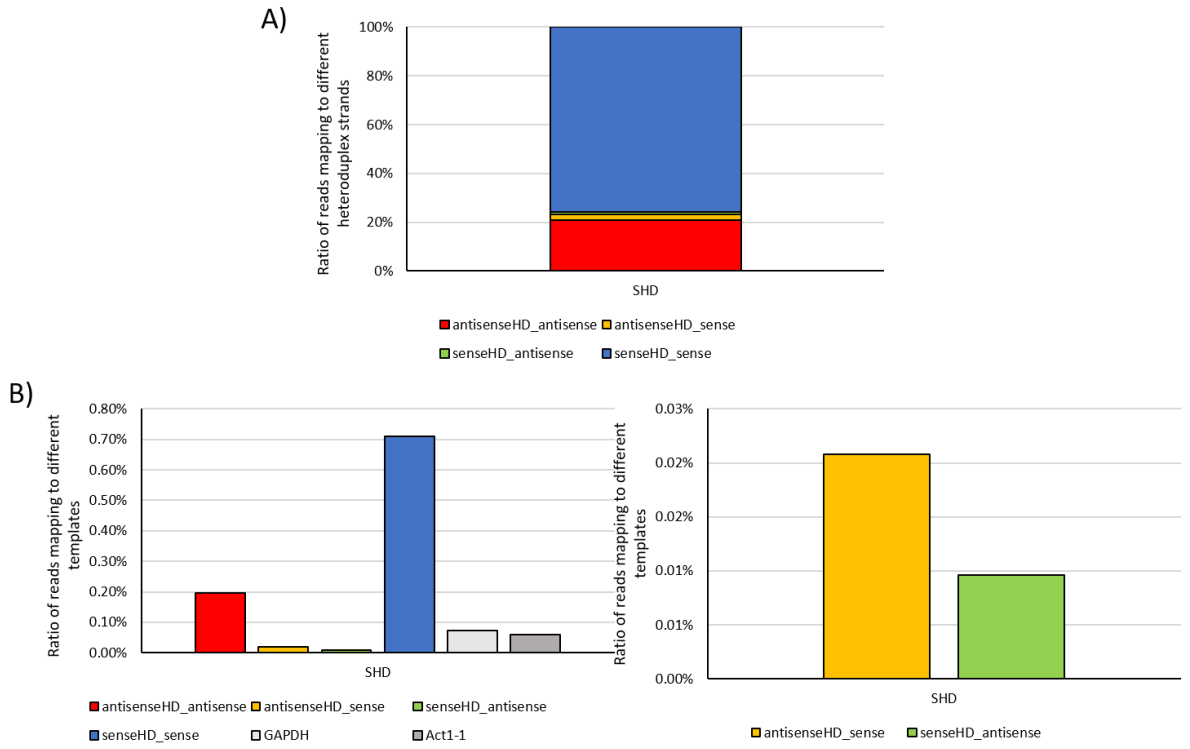

Figure S 3 Switched Heteroduplex (SHD) longRNA data. A) Ratio of different SHD derived RNA species normalized to all Heteroduplex mapping reads. Shown are longRNAs mapping to the original SHD dsRNA strands (labeled in red and blue) or to RdRP produced strands (labeled in green and orange). B) Abundance of Heteroduplex derived reads mapping to the four different Heteroduplex derived strand (color coded as described prior), as well as reads mapping to the two housekeeping genes glyceraldehyde-3-phosphate dehydrogenase (GAPDH, labeled in light grey) and actin1-1 (Act1-1, labeled in dark grey) (left), and the same graph showing only RdRP derived strands (right).

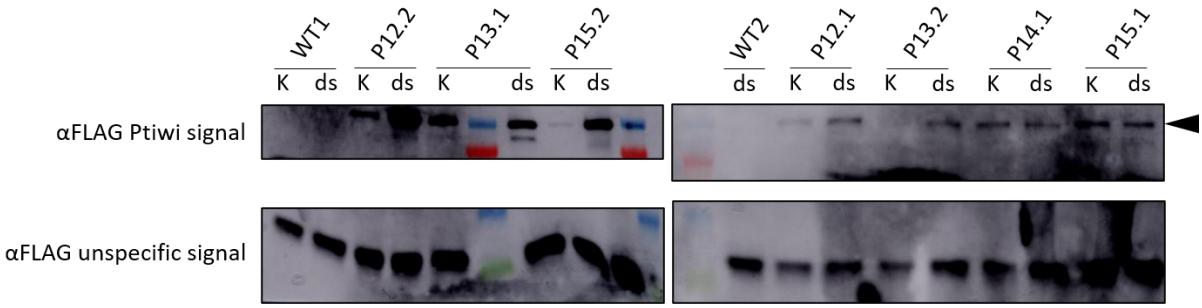

Figure S 4 Western Blot analysis of RNA IP samples. Shown are Western Blot signals derived from a FLAG-antibody (αFLAG) directed against the FLAG-tag different Ptiwi fusion proteins expressed in transgenic cell lines. Cells were either fed with dsRNA producing *E.coli* (dsRNA) or non-dsRNA containing *Klebsiella pneumoniae* bacteria (K), depending on whether the IP sample was used for endogenous SRC analysis or feeding smallRNA analysis. Lysate of cells were used for RNA IP and presence of the Ptiwi fusion protein was detected at the expected size (approx. 90kDa, indicated by the black arrowhead). RNA IP samples from wildtype cells not expressing the fusion protein were used as a negative control. Unspecific signals of the FLAG antibody were used as a loading control. The color prestained protein standard (broad range, NEB) was used, the red band corresponds to 72kDa while the green band corresponds to 23kDa.

Table S 2 List of SRCs that are considered loaded into different Ptiwis. Loaded SRCs were determined by *de novo* cluster prediction using reads from Ptiwi RIP sequencing data and overlapping the predicted SRCs with the existing SRCs described previously (Karunanithi, Oruganti, et al., 2019). Gene associated SRCs are displayed in green, while non gene associated SRCs are displayed in black.

RIP_P12	RIP_P13	RIP_P14	RIP_P15
C1020 C1086 C1096	C1020 C1114 C1126	C1086 C1096 C1114	C1020 C1096 C1114
C1114 C1126 C1149	C1148 C1149 C1191	C1118 C1149 C1191	C1126 C1148 C1149
C1172 C1191 C1205	C1205 C125 C1311	C1205 C125 C1377	C1191 C1205 C1311
C1311 C1350 C136	C136 C143 C1478	C1447 C156 C158	C136 C1447 C1478
C1478 C1691 C1712	C156 C158 C1712	C1639 C1899 C1976	C156 C158 C1691
C1756 C1839 C1899	C1805 C1838 C1839	C2038 C2080 C2103	C1805 C1839 C1899
C1908 C1909 C1913	C1899 C1905 C1913	C2104 C2172 C2225	C1913 C192 C2104
C192 C2104 C2135	C192 C1976 C2098	C232 C2322 C2334	C2135 C232 C2322
C2225 C2304 C232	C2169 C2170 C2225	C2338 C2405 C241	C2328 C2338 C241
C2322 C2328 C2338	C2257 C232 C2322	C2518 C2586 C263	C2473 C2486 C2518
C241 C2473 C2486	C2328 C2334 C2338	C325 C344 C477	C2546 C2586 C263
C2518 C2546 C263	C241 C2452 C2486	C495 C577 C657	C277 C340 C344
C340 C344 C418	C2518 C263 C299	C673 C694 C745	C418 C495 C594
C495 C657 C673	C32 C340 C344	C893 C1092 C1095	C657 C673 C745
C745 C781 C822	C418 C43 C495	C1107 C1163 C1166	C875 C893 C901
C875 C880 C893	C556 C657 C673	C1203 C1212 C1233	C1001 C1057 C1091
C966 C1057 C1091	C745 C893 C948	C1269 C1273 C1278	C1107 C1163 C1166
C1107 C1127 C1163	C990 C1057 C1107	C1319 C1348 C1380	C1203 C1212 C1254
C1166 C1203 C1212	C1127 C1163 C1166	C1389 C141 C1426	C1273 C1278 C1375
C1233 C1240 C1254	C1203 C1212 C1254	C1428 C1430 C1431	C1380 C141 C1426
C1278 C1283 C1375	C1278 C1283 C1389	C150 C1501 C153	C1427 C1428 C1429
C1380 C141 C1425	C141 C1426 C1428	C157 C1573 C1588	C1430 C1431 C1434
C1426 C1428 C1430	C1431 C150 C1511	C1621 C1638 C1811	C1454 C150 C1516
C1431 C1454 C150	C1516 C157 C1588	C1822 C1911 C1914	C153 C157 C1621
C1505 C1516 C153	C1621 C1638 C1703	C1922 C193 C1934	C1638 C1650 C1686
C1621 C1638 C1650	C1777 C1784 C1818	C1947 C1965 C1970	C1689 C1703 C1710
C1686 C1703 C1710	C1822 C1842 C1863	C1975 C1989 C199	C1711 C1784 C1818
C1743 C1818 C1822	C1880 C193 C1947	C1991 C2024 C2035	C1822 C1842 C1868
C1842 C1872 C1880	C1965 C1970 C1975	C204 C2052 C2056	C1872 C1880 C190
C190 C1907 C1922	C1989 C199 C1991	C2061 C2062 C2078	C1914 C1922 C193

C193 C1934 C1947 C1965 C1969 C1970 C1975 C1988 C1989 C199 C1991 C2010 C2035 C204 C2061 C2070 C2077 C2078 C210 C211 C2204 C2226 C2232 C2242 C2271 C2288 C2307 C2336 C2344 C236 C2375 C2398 C24 C2403 C2430 C2459 C2505 C2555 C2556 C2560 C2572 C2580 C265 C28 C29 C31 C323 C338 C345 C361 C384 C433 C448 C473 C492 C504 C57 C571 C59 C599 C643 C652 C7 C720 C740 C754 C762 C783 C788 C789 C806 C814 C824 C826 C861 C871 C909 C940 C955 C957	C2035 C204 C2052 C2056 C2061 C210 C211 C2204 C2226 C2232 C224 C2246 C225 C2271 C2299 C2307 C2336 C236 C2375 C2398 C24 C2403 C2430 C2552 C2556 C2560 C2572 C2580 C265 C28 C29 C31 C338 C341 C345 C384 C42 C433 C448 C473 C504 C596 C599 C652 C7 C720 C721 C753 C754 C762 C783 C789 C806 C824 C861 C871 C909 C955 C957	C2181 C2201 C2226 C2232 C2271 C2284 C2307 C236 C2379 C2398 C2430 C2459 C2529 C2552 C2555 C2556 C2560 C2572 C265 C28 C338 C345 C384 C417 C448 C473 C492 C504 C571 C596 C599 C643 C652 C753 C754 C806 C824 C826 871 C88 C909 C940	C1934 C1947 C1965 C1969 C1970 C1975 C1988 C1989 C199 C1991 C2010 C2035 C204 C2061 C2070 C2077 C2078 C210 C211 C2129 C2224 C2226 C2232 C224 C2242 C2246 C225 C2271 C2288 C2307 C2336 C2344 C2346 C236 C2375 C2398 C24 C2403 C2430 C2457 C2459 C2555 C2556 C2560 C2572 C2580 C265 C28 C29 C31 C338 C345 C361 C384 C42 C433 C473 C492 C504 C513 C571 C59 C599 C643 C652 C7 C720 C740 C754 C762 C783 C788 C789 C80 C806 C814 C824 C826 C838 C861 C871 C909 C940 C955 C957 C959

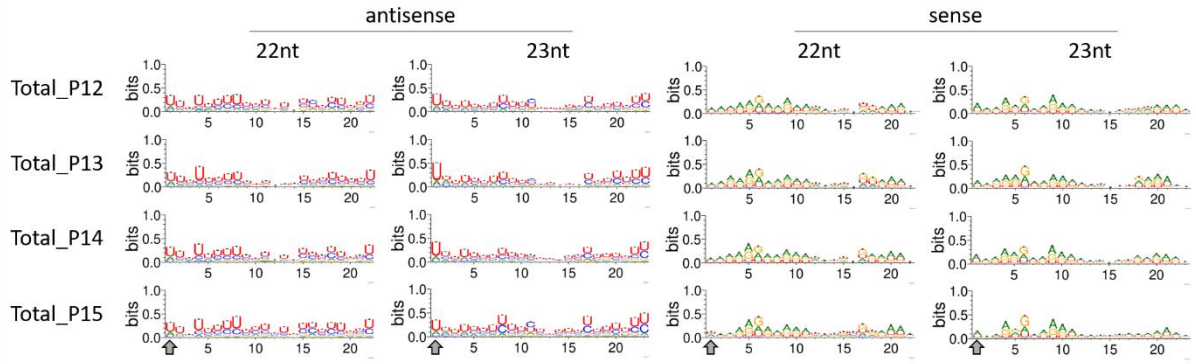

Figure S 5 Sequence logo analysis of smallRNA reads derived from total RNA. Sequence logos of 22nt and 23nt long primary smallRNA reads present in the total RNA data of different Ptiwi-fusion protein expressing cell lines. Sequence logos were calculated using the weblogo tool for sense and antisense reads, normalizing the logo for the paramecium genome base composition (A/U: 0.36; G/C: 0.14 each, according to (Lepère et al., 2009)).

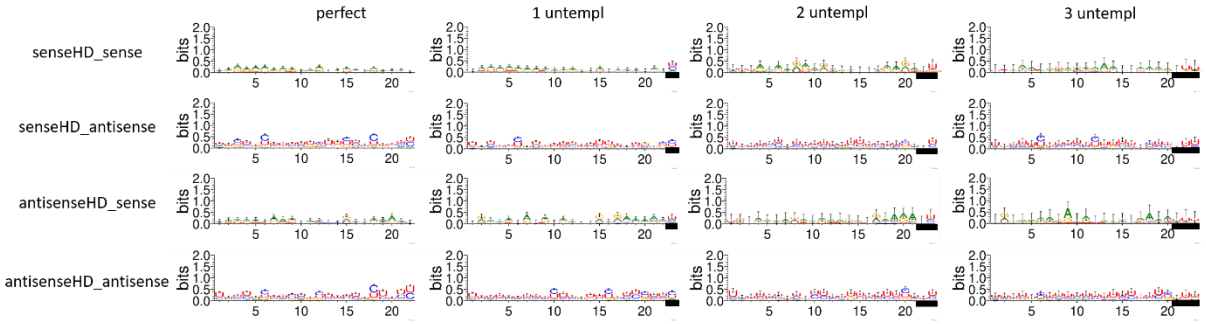

Figure S 6 Sequence logo analysis of untemplated nucleotides in 23nt siRNAs derived from Heteroduplex strands. Displayed are sequence logos from 23nt long siRNAs carrying varying amounts of untemplated nucleotides. Untemplated nucleotides displayed in the sequence logos are underlined with a black bar. Sequence logos were calculated using the weblogo tool, normalizing the logo for the paramecium genome base composition (A/U: 0.36; G/C: 0.14 each, according to (Lepère et al., 2009)

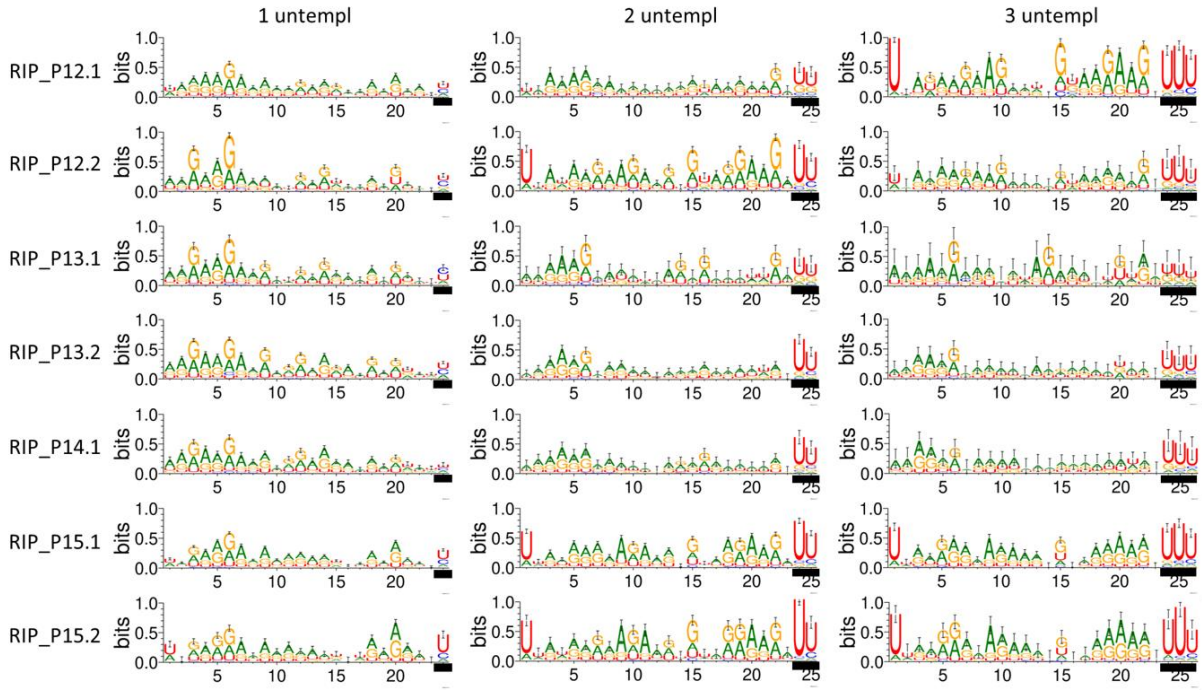

Figure S 7 Sequence logo analysis of untemplated 23nt+x smallRNA reads in Ptiwi RIP data. Displayed are sequence logos from 23nt+x long siRNAs carrying varying amounts of untemplated nucleotides. Untemplated nucleotides displayed in the sequence logos are underlined with a black bar. Sequence logos were calculated using the weblogo tool, normalizing the logo for the paramecium genome base composition (A/U: 0.36; G/C: 0.14 each, according to (Lepère et al., 2009)

Curriculum vitae

Aus Datenschutzgründen ist der abgedruckte Lebenslauf nicht in der elektronischen Version dieser Dissertation enthalten.

Aus Datenschutzgründen ist der abgedruckte Lebenslauf nicht in der elektronischen Version dieser Dissertation enthalten.

Aus Datenschutzgründen ist der abgedruckte Lebenslauf nicht in der elektronischen Version dieser Dissertation enthalten.

Aus Datenschutzgründen ist der abgedruckte Lebenslauf nicht in der elektronischen Version dieser Dissertation enthalten.

Aus Datenschutzgründen ist der abgedruckte Lebenslauf nicht in der elektronischen Version dieser Dissertation enthalten.

Acknowledgments

Zunächst einmal möchte ich mich bei **Martin Simon** für die Möglichkeit bedanken, nicht nur meine Dissertation, sondern auch alle meine anderen Abschlussarbeiten in seiner Arbeitsgruppe zu schreiben. Martin, wir arbeiten jetzt seit 10 Jahren miteinander und haben uns auf einigen Etappen unseres jeweiligen Werdegangs begleitet. Ich danke dir für alle deine Hilfe, deinem Beitrag zu meiner Entwicklung als Mensch und Wissenschaftler, sowie für all deine Bemühen und die Gedanken, die du dir machst, damit es nicht nur mir, sondern auch all den Leuten, die für und mit dir arbeiten, gut geht.

Außerdem möchte ich mich bei **Julia Bornhorst** für die Übernahme des Zweitgutachtens dieser Arbeit bedanken. Danke, dass du und deine AG uns nach unserem Umzug so offen und freundlich an der Uni Wuppertal empfangen habt!

Ich möchte mich bei **Gilles Gasparoni** von der Universität des Saarlandes sowie bei **Sören Franzenburg** vom IKMB für die Unterstützung bei der Sequenzierung all meiner verschiedenen Proben bedanken.

Bei **Marc Schneider** und seiner Arbeitsgruppe an der Universität des Saarlandes, allen voran **Johannes Büscher** sowie **Kristela Shelu** und **Mark Sabura**, möchte ich mich für die Hilfe bei der Etablierung der Heteroduplex-Nanopartikel bedanken. Es hat viel Zeit und Arbeit gekostet, den Heteroduplex ans Laufen zu bekommen und ohne eure Ideen und dem Knowhow, was ihr mitgebracht habt, wären diese Experimente so nicht möglich gewesen.

Bei **Eric Meyer**, **Alexey Potekhin** and **Sandra Duharcourt** möchte ich mich für das Teilen der verschiedenen RNAi Mutanten bedanken, ohne die viele Teile dieser Arbeit deutlich schwerer gewesen wären.

Niemals genug kann ich mich bei den Leuten meiner Arbeitsgruppe in Wuppertal bedanken. Danke an **Melanie** und **Natalie** für alle die Unterstützung, die ihr im Laboralltag leistet und das offene Ohr, was ihr beide für die Probleme anderer Leute habt. Kommt man zu euch ins Büro, kann man immer ein nettes Gespräch, Hilfsbereitschaft und allem voran einen gut gedeckten Tisch mit Nervennahrung erwarten. Vor allem Letzterer hat schon so manchen Tag gerettet. Danke auch an **Katinka**, für alle die Hilfe beim Navigieren des universitären Bürokratie-Dschungels sowie das Herumschlagen mit Moodle und Studilöwe, ohne dich wüsste man oft nicht, wo einem da gerade der Kopf steht.

Auch Teil der Arbeitsgruppe, aber nochmal extra hervorheben, möchte ich **Franzi**. Wir teilen uns jetzt seit meiner Masterarbeit, ca. 7 Jahre also, ein Büro (naja, mittlerweile Büro, früher nur ne lange Bank im Labor) und ich glaube, man kann sich nur schwer eine bessere Kollegin wünschen. Nur wenige Leute hatten einen so großen Einfluss auf mich als Mensch, und ich bin froh, dass du nicht nur eine grandiose Kollegin, sondern auch eine meiner besten Freunde geworden bist! Danke dir für alle die Unterstützung, die du mir entgegenbringst, dass du meine neuen Drink-Ideen probierst und all die schönen Stunden, die wir über diese verdammten Rätsel-Spiele verzweifelt sind! Bleib wie du bist!

Ich möchte mich bei **Steffi** und **Andrea** für die wunderschöne Zeit in Saarbrücken und all die langen Samstag-Abende in der Nautilus Bar, sowie bei **Conny**, **Angi** und **Tobi** für die vielen Abende im Discord bei verschiedensten Spielen und Pen&Paper-Runden bedanken!

Ein weiterer besonderer Dank gilt **Tamara**. Zunächst einmal danke ich dir, dass du dich zum Korrekturlesen der Arbeit bereiterklärt hast, auch wenn du die meisten Inhalte vermutlich bereits ziemlich gut kanntest, denn in den letzten zwei Jahren musstest du dir oft anhören, was alles nicht so funktioniert, wie es soll. Danke, dass du dabei nie die Nerven verloren, geduldig zugehört, mit Rat zur Seite gestanden und mir immer gut zugeredet hast. Du warst mir im Bezug auf die Arbeit, aber auch im Privatleben eine riesengroße Hilfe und ohne deinen Support hätte ich einige Dinge nicht hinbekommen! Ich bin unglaublich froh, dass wir uns gefunden haben und möchte mich für unsere bisherige gemeinsame Zeit, sowie all die Zeit die noch kommen wird, bedanken.

Zu guter Letzt richtet sich mein Dank an **meine Familie**, allen voran an **meine Eltern** sowie **meine Schwester**. Mein Leben lang habt ihr mir bei meinem Weg geholfen, mich bei meinen Entscheidungen unterstützt und mit eurem Engagement dafür gesorgt, dass viele Sachen in meinem Leben einfacher waren, als sie hätten sein müssen. Dafür möchte ich mich bei euch bedanken. Ich weiß, dass ich mich nur halb so oft melde wie ich sollte, und nur ein Bruchteil so oft wie ihr es verdienen würdet, aber ich danke euch, dass ihr mich so akzeptiert wie ich bin. Ich bin stolz auf alles, was ihr drei leistet und bisher erreicht habt!

Declaration of Authorship

Hiermit versichere ich an Eides statt, dass ich diese Arbeit selbständig angefertigt und keine anderen als die angegebenen Hilfsmittel und Quellen verwendet habe.

Die vorliegende Arbeit wurde bisher weder im In- noch Ausland in gleicher oder ähnlicher Form bei einem Verfahren zur Erlangung eines akademischen Grades vorgelegt.

Ich erkläre darüber hinaus, dass alle Methoden, Daten und Arbeitsabläufe wahrheitsgetreu dokumentiert und keinerlei Daten manipuliert wurden.

Unterschrift: _____

Datum: _____