Piwi and piRNAs act upstream of an endogenous siRNA pathway to suppress Tc3 transposon mobility in the *Caenorhabditis elegans* germline

Dissertation
zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultäten
der Georg-August-Universität zu Göttingen

vorgelegt von

Partha Pratim Das

aus Calcutta, India

Göttingen 2008

D7

Referent: Prof. Dr. Ernst Wimmer

Co-referent: Prof. Dr. Ralf Ficner

Tag der mündlichen Prüfung: 31.10.08

The experimental part was performed at the

Wellcome Trust/Cancer Research UK Gurdon Institute
The Henry Wellcome Building of Cancer and Developmental Biology, University of Cambridge, UK

Under supervision of Dr. Eric Miska

For my parents and pratibha……

Table of contents

Abbreviations

bp	base pair
BPB	bromophenol blue
C. elegans	*Caenorhabditis elegans*
cDNA	complementary DNA
dNTP	deoxynucleotide triphosphate
DNA	deoxyribonucleic acid
D. melanogaster	*Drosophila melanogaster*
D. rerio	*Danio rerio*
dsRNA	double-stranded RNA
E. coli	*Escherichia coli*
esiRNA	endogenous siRNAs
IAP	intracisternal A particle
LINE	long interspersed nucleotide elements
LNA	locked nucleic acid
LTR	long terminal repeats
mRNA	messenger RNA
miRNA	microRNA
miRNP	microRNA-containing RNP
M. musculus	*Mus musculus*
nt	nucleotide
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
piRNA	Piwi-interacting RNAs
pre-miRNA	precursor miRNA
pri-miRNA	primary precursor miRNA
PTGS	post transcriptional gene silencing
qRT-PCR	quantitative real time RT-PCR
PNK	polynucleotide kinase
rasiRNAs	repeat associated siRNA
RdRP	RNA dependent RNA polymerase
RISC	RNA induced silencing complex
RITS	RNA-induced transcriptional gene silencing
RNA	ribonucleic acid
RNAi	RNA interference
RNP	ribonucleoprotein
rpm	revolutions per minute
RT-PCR	Reverse-transcription PCR
SDS	sodium dodecylsulfate
SINE	small interspersed nucleotide elements
siRNA	small interfering RNA

S. pombe	*Schizosaccharomyces pombe*
S. cerevisiae	*Schizosaccharomyces cerevisiae*
ssRNA	single stranded RNA
TGS	transcriptional gene silencing
UTR	untranslated region

List of Figures

List of Tables

Abstract

Several hundred endogenous small RNAs, namely microRNAs (miRNAs), repeat-associated small interfering RNAs (rasiRNAs), small interfering RNAs (siRNAs) have been discovered in diverse organisms. These small RNAs are derived from different sources and they cause transcriptional gene silencing, translational repression and mRNA cleavage through effector complexes. Several functionally distinct RNA silencing effector complexes have already been isolated from different organisms. Effector complexes are composed of members of the Argonaute (AGO) protein family and single-stranded small RNAs, along with other associated proteins. Members of the Piwi subfamily of Argonaute proteins have conserved roles in germline development and stem cell maintenance. In *C. elegans*, PRG-1 and PRG-2 are the two members of Piwi sub-family of Argonaute proteins. In order to study the role of PRG-1 and PRG-2 in *C. elegans* germline development, we generated *prg-1* and *prg-2* single mutants as well as *prg-1; prg-2* double mutant worms, which show severe germline defects. I identify a class of 21 nucleotide RNAs, previously named 21U-RNAs, as the Piwi-interacting RNAs in *C. elegans*. Piwi and piRNA expression is restricted to the male and female germline and piRNA biogenesis is independent of many other proteins involved in small RNA pathways including DCR-1. I show that Piwi is specifically required for suppression of Tc3 transposons. The excision rate of Tc3 is 100-fold higher in *piwi* mutants compared to wild-type. There is no evidence for a 'ping-pong' mechanism in *C. elegans* for amplification of piRNAs. Finally, I demonstrate that a Piwi-piRNA complex acts upstream of the endogenous siRNA pathway for Tc3 silencing and that might suggest a link between the function of endogenous siRNAs and piRNAs.

1. Introduction

1.1 Introduction

The revolution of small RNA biology began after the ground-breaking discovery by Andrew Fire and Craig Mello injecting small double-stranded RNA (dsRNA) into *C. elegans* leads to specific silencing of a gene that highly homologous to the dsRNA sequence (Fire et al., 1998), this mechasims coined as RNA interference (RNAi). RNAi is performed through transcriptional gene silencing (TGS) and or post-transcriptional gene silencing (PTGS) by small RNAs (Bartel, 2004; Meister and Tuschl, 2004). Naturally occurring small RNAs are processed from long double stranded or single stranded RNA precursors, which are either generated from the genome or produced during viral replication. Several types of naturally occurring small RNAs have been discovered in recent years, namely, short interfering RNAs (siRNAs), repeat–associated small interfering RNAs (rasiRNAs), microRNAs (miRNAs), Piwi-interacting RNAs (piRNAs) and endogenous short interfering RNAs (esiRNAs) (Kim, 2006; Meister and Tuschl, 2004; Tam et al., 2008). siRNAs and rasiRNAs are generated from dsRNA precursors and either regulate transcriptional gene silencing or post-transcriptional gene silencing. Biogenesis of miRNAs has been studied quite extensively; most of the miRNA genes are transcribed by RNA Pol II as primary transcripts (Borchert et al., 2006; Lee et al., 2004) that are processed into stem-loop precursor miRNAs (pre-miRNAs) by the RNase III enzyme Drosha-containing Microprocessor complex (Denli et al., 2004; Gregory et al., 2004). Pre-miRNAs are transported from the nucleus to the cytoplasm via Exportin-5 (Bohnsack et al., 2004) and are recognized by another RNase III enzyme called Dicer that cleaves pre-miRNAs to 21-23nt double stranded mature miRNAs. Mature miRNAs are loaded into the Argonaute containing RNA-induced silencing complex (RISC) miRNA containing RISC complex either perform sequence specific cleavage (in plants) or translational repression (in animals) of the target mRNA (Peters and Meister, 2007) depending upon the degree of complementarity between mature miRNA and target mRNA (Bartel, 2004). Piwi-interacting RNAs (piRNAs) a

new class of small RNAs (26-31 nt) have been discovered recently by cloning of small RNAs associated with Piwi from mammalian testes (Kim et al., 2006). However, to date, the biogenesis of piRNAs is poorly understood. piRNA biogenesis is Dicer-independent (Houwing et al., 2007; Vagin et al., 2006). According to the piRNAs genomic distribution, it suggests that piRNAs might be generated from one long primary transcripts (Kim, 2006). Recently, esiRNAs have been discovered in *Drosophila* and mice (Czech et al., 2008; Ghildiyal et al., 2008; Kawamura et al., 2008; Okamura et al., 2008; Tam et al., 2008; Watanabe et al., 2008). esiRNAs are produced by Dicer from several precursors - (1) convergent transcripts and/or (2) dsRNA precursors formed by hybridization between spliced transcripts from protein coding genes and antisense transcripts from homologous pseudogenes and/or (3) stem-loop transcripts from transposon loci and regulate both protein coding genes and retrotransposons through unknown mechanisms.

Distinct class of small RNAs loaded into different ribonucleoprotein (RNP) complexes. These effector complexes consist of different combinations of RNAs, Argonautes and their associated proteins. The effector complex that mediates mRNA cleavage by siRNAs or miRNAs is called the RNA induced silencing complex (RISC), the effector complex that mediates translational repression through miRNA is called the miRNP complex, and the effector complex that mediates transcriptional silencing through chromatin modification is called the RNA-induced transcriptional gene silencing (RITS) complex (Meister and Tuschl, 2004).

1.2 The Argonaute Superfamily

1.2.1 The Argonaute protein family

The term Argonaute was originally used to describe a mutant of *Arabidopsis thaliana*, *AGO-1*, in which the morphology of the leaves resembled a small squid, the 'greater argonaut' or 'argonauta argo' (Bohmert et al., 1998). The Argonaute proteins are well

conserved and based on their sequence similarities; they are classified into three sub-families (Figure 1.1). The largest sub-family comprises the Argonautes (Ago), named after the founding Argonautes from *A. thaliana*. The second sub-family comprises the Piwis, named after Drosophila Piwi (P- element-induced wimpy testis). The third and last sub-family comprises nematode specifc Argonautes proteins (Yigit et al., 2006).

The number of Argonaute genes is highly variable between species (Table 1.1). For instance, there are 8 Argonaute genes in humans (4 Argonaute-like and 4 Piwi-like), 5 in *Drosophila* (2 Argonaute-like and 3 Piwi-like) and 27 Argonaute genes in *C. elegans* (5 Argonaute-like genes, 2 Piwi-like and rest belong to family 3). Argonaute sub-family of proteins are ubiquitously expressed whereas Piwi sub-family of Argonaute proteins are only expressed to germ cells (Peters and Meister, 2007). Most species have both Argonaute and Piwi proteins whereas some have either Argonaute or Piwi proteins, and others, like *Trypanosoma* and *Leishmania,* may have lost the PAZ domain and retain only a PIWI domain during evolution. In general, two types of silencing mechanisms are present in all organisms; Piwi-like proteins to regulate transcriptional gene silencing, such as transposon silencing, and Argonaute-like proteins to regulate translation by targeting mRNA. A high degree of gene duplication of Argonaute-like genes, especially in nematodes and plants, has allowed diversification in their function.
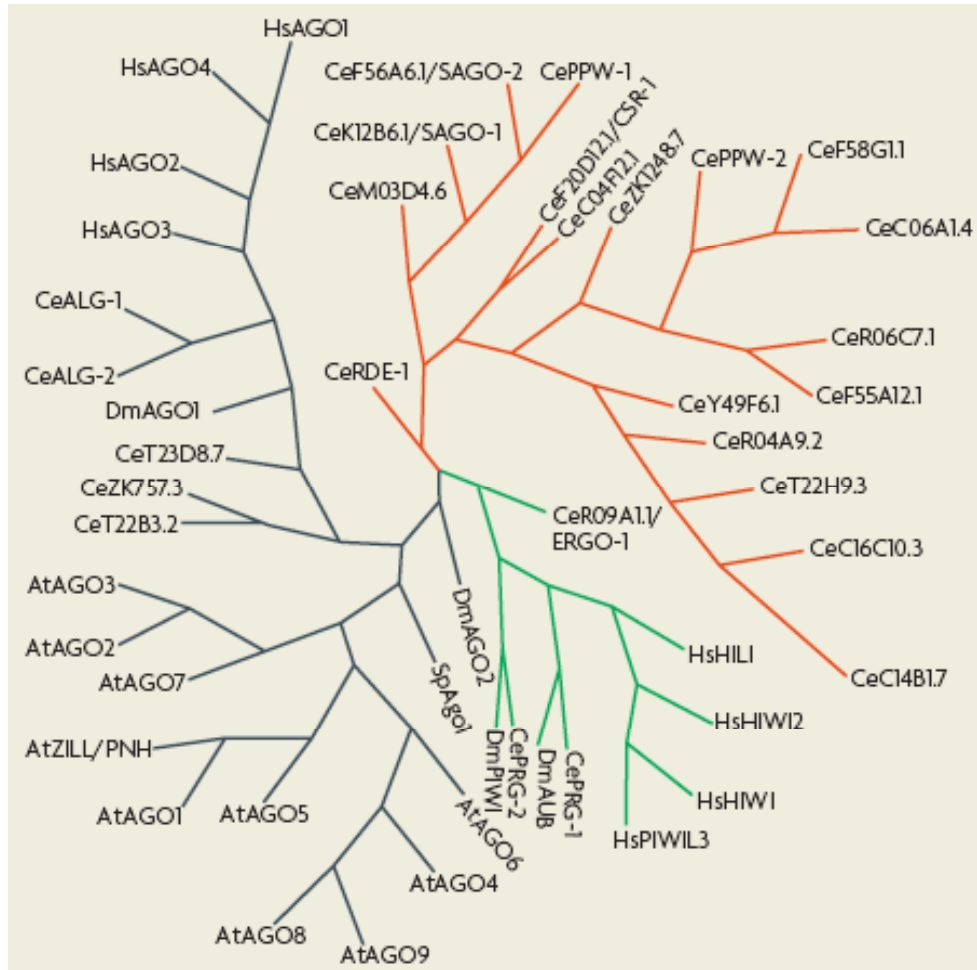
Figure 1.1 Phylogenetic tree of Argonaute proteins
Argonaute proteins from fungi, plants and animals may be divided into three families: (1) Argonaute sub-family (in dark blue), 2) Piwi family (in green) and 3) nematode-specific Argonaute proteins (in red). This figure was reproduced from Yigit et al., 2006.

Table 1.1 Argonaute proteins in different species and their small RNA partners

Argonaute proteins	Function	Small RNA class
Neurospora crass **(2)**		
QDE-2	Quelling	siRNA
Saccharomyces pombe **(1)**		
AGO-1	TGS, Heterochromatin silencing	siRNA
Tetrahymena thermophilia **(2)**		
TWI1	DNA elimination	siRNA
Arabidopsis thaliana **(10)**		
AGO1	Development, miRNA-mediated gene silencing	miRNA
AGO4	Heterochromatin silencing	rasiRNAs
AGO6	Heterochromatin silencing	rasiRNAs
AGO7	Leaf development	tasiRNA
Caenorhabditis elegans **(27)**		
RDE-1	Exogenous RNAi	siRNA
ALG-1	miRNA-mediated gene silencing	miRNA
ALG-2	miRNA-mediated gene silencing	miRNA
ERGO-1	endogenous RNAi	siRNA
CSR-1	Chromosome segregation, RNAi	siRNA
SAGO-1	Endogenous and exogenous RNAi	secondary siRNA
SAGO-2	Endogenous and exogenous RNAi	secondary siRNA
PPW-1	Endogenous and exogenous RNAi	secondary siRNA
PPW-2	Endogenous and exogenous RNAi	secondary siRNA
F58G1.1	Endogenous and exogenous RNAi	secondary siRNA
C16C10.3	Endogenous and exogenous RNAi	secondary siRNA
PRG-1	Germline maintenance	-
Drosophila melanogaster **(5)**		
AGO-1	miRNA-mediated gene silenciing	miRNA
AGO-2	RNAi	siRNA
AGO-3	Transposon silencing	piRNA
PIWI	Transposon silencing, germ stem cell maintenance	piRNA
Aubergine	Transposon silencing	piRNA
Danio rerio **(5)**		
ZIWI	Germline maintenance, Transposon silencing	piRNA
Mus musculus **(8)**		
AGO-2	Development, miRNA-mediated gene silencing	miRNA
MIWI	Transposon silencing, spermatogenesis	piRNA
MILI	Transposon silencing, spermatogenesis	piRNA
MIWI2	Transposon silencing	-
Rattus norvegicus **(7)**		
RIWI	-	piRNA
Homo sapiens **(8)**		
AGO-1	-	miRNA
AGO-2	-	miRNA, siRNA
AGO-3	-	miRNA
AGO-4	-	miRNA
HIWI	spermatogenesis?	-

The table contains Argonaute-like and Piwi-like proteins that have been associated with either different class of small RNAs or cellular functions. Numbers in parentheses indicate the total number of Argonautes found within the given species. *Arabidopsis* has 6 more and *C. elegans* has 15 more Argonaute proteins with unknown function/s. (-) indicate unknown function of Argonautes or unknown class of small RNAs. miRNA- microRNA, siRNA- small interfering RNA, piRNA- Piwi-interacting RNA, rasiRNA- repeat-associated siRNA, tasiRNA- trans-acting siRNAs, TGS- transcriptional gene silencing, RNAi- RNA interference.

1.2.2 Argonaute proteins and their functional domains

Argonaute proteins have a molecular weight of about 100KDa and consist of a PAZ (piwi-argonaute-zwille) domain, a PIWI domain and a MID domain (Figure 1.2A). The PAZ domain is found in both Dicer and Argonaute proteins, two protein families with a key role in RNAi. The PAZ domain consists of a nucleic-acid binding domain, termed the oligonucleotide/oligosaccharide – binding fold (OB fold). Nucleic acid binding studies have demonstrated that the PAZ domain has a high affinity for single-stranded 3′-ends and duplex siRNAs with dinucleotide 3′ overhangs, a weak affinity for mononucleotide 3′ overhangs and no affinity for blunt-ended duplexes (Lingel et al., 2003; Ma et al., 2004; Song et al., 2003; Yan et al., 2003). Mostly, miRNAs and siRNAs are loaded into the Argonaute containing effector complex, the 'passenger strand' or 'start strand' is cleaved by Argonaute and the 'guide strand' remains bound to the Argonaute and the effector complex become active (Matranga et al., 2005; Miyoshi et al., 2005).

The MID (Middle) domain structure is present between the PAZ and PIWI domains of Argonaute proteins. It is thought to be required to load small RNA into the effector complex presumably by receiving and binding the 5′ phosphate of the small RNA duplex (Nykanen et al., 2001). Structural studies of crystallized Argonaute proteins in the presence of either a single standed RNA or siRNA duplex show that the 5′ phosphate is anchored in between the MID and PIWI domains (called the MC domain) by divalent cation ions, such as magnesium (Ma et al., 2004; Ma et al., 2005; Parker et al., 2005; Yuan et al., 2005). The MC domain has striking sequence similarity to the cap-binding domain of eukaryotic initiation factor eIF4E. eIF4E binds to the mRNA cap by its two tryptophan (W56 and W102) residues. At the position equivalent to the tryptophan residues of eIF4E, AGO2 has phenylalanine (F470 and F505). Substitution of phenylalanine residues to valines abolishe both binding to m^{7}GTP-sepharose and silencing activity of AGO2 (Kiriakidou et al., 2007). In contrast, substitution of conserved phenylalanine residues to valines of AGO-1 in *Drosophila* abolishe silencing

effect of AGO-1 without affecting binding to the m^{7}GTP-sepharose (Eulalio et al., 2008).

The PIWI domain is the catalytic component of Argonaute proteins. The structure of full-length archaeal (*Pyrococcus furiosus*), eubacterial (*Aquifex aeolicus*) Argonautes and an Argonaute without its PAZ domain from *Archaeoglobus fulgidus* revealed that the PIWI domain has an RNase H fold (Ma et al., 2005; Parker et al., 2004; Song et al., 2004; Yuan et al., 2005) (Figure 1.2B). The RNase H domain is conserved in prokaryotes and eukaryotes and it cleaves target mRNAs determined by the guide strand in a sequence-specific manner (Figure 1.2C & D). The RNase H fold has a catalytic triad Asp-Asp-His/Asp (DDH) motif, which has a processing endonuclease activity with divalent cations (Tolia and Joshua-Tor, 2007). Mutation of the triad residues of human Ago2 abolishes slicer activity demonstrating the significance of the catalytic triad motif (DDH) of Argonaute proteins. *In vitro* assays also show that only hAgo2 has an endonuclease activity among the human Ago sub-family of proteins (Meister et al., 2004). Ago3 is catalytically inactive, even though it has DDH motif, indicating that other residues, post-translational modifications or associated proteins might also be important for catalytic activity. In addition, the Piwi subfamily Argonautes from *Drosophila* and rat have slicing activity *in vitro* (Lau et al., 2006; Saito et al., 2006). However, the catalytic potential of the rest of the Piwi subfamily of Argonautes proteins remains untested.

A

N PAZ Mid PIWI

PAZ domain- 2nt 3′ overhang
MID domain- 5′5′ phosphate binding
PIWI domain- RNase H catalytic activity

B

Slicer catalytic site

C

D
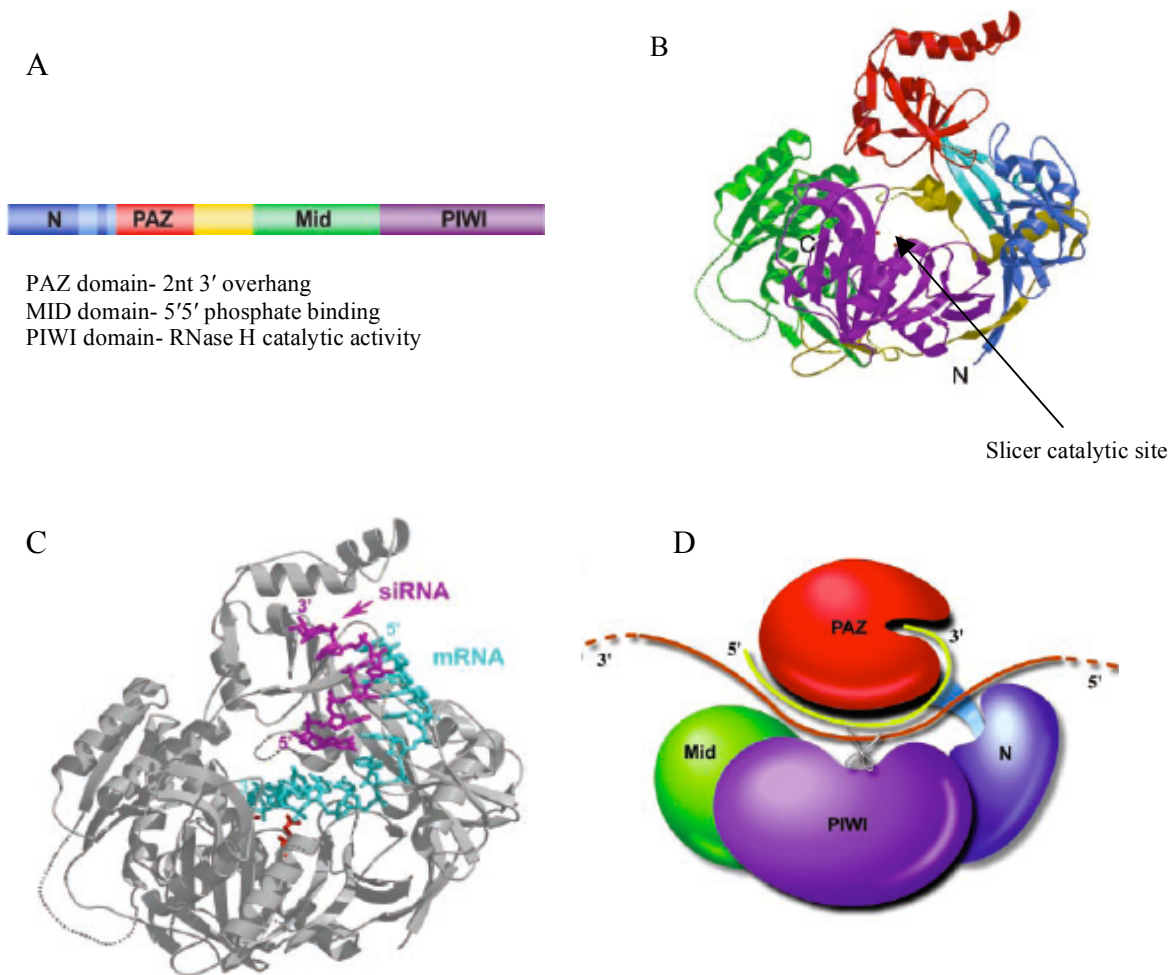
Figure 1.2 The Structure of Argonautes and a model for siRNA-guided mRNA cleavage by Argonautes. (A) A schematic diagram of the different domains of the Argonaute protein from *Pyrococcus furiosus* is shown and general functions of the different domains of the Argonaute protein are briefly described. (B) Crystal structure of the full-length Argonaute from *Pyrococcus furiosus* (pfAgo). The domains are coloured as follows: PAZ domain (red), MID domain (green) and PIWI domain (purple). The location of the Slicer catalytic site in the PIWI domain is shown by stick representation and by arrowhead projection. The overall fold looks like a clamp-like structure, with the PAZ domain situated over the PIWI domain, which contains the catalytic site (RNaseH fold with DDH motif) for Slicer activity. (C) The crystal structure of the Argonaute protein from *Pyrococcus furiosus*. The 3′ end of the siRNA (purple) is superimposed on the PAZ domain with part of the siRNA binding to the mRNA strand (light blue) and interacting with the active site of the PIWI domain (shown in red). (D) A model of siRNA-guided mRNA cleavage by Argonaute. The siRNA (yellow) binds with its 3′ end to the PAZ cleft and the 5′ end binds to the end of the MID domain. The mRNA (brown) enters between the N-terminal and PAZ domains and exits between the PAZ and middle (MID) domain. The active site in the PIWI domain (shown as scissors) cleaves the mRNA opposite the center nuceotide of the siRNA guide. These figures were adapted from Song et al., 2004.

1.2.3 Argonaute proteins and small RNA pathways

A siRNA-containing effector complex is commonly referred to as a RISC, whereas a miRNA-containing effector complex is referred to as a miRNP. Every RISC and miRNP complexes contain a member of the Argonaute protein (Hammond et al., 2000; Mourelatos et al., 2002).

It is sometime difficult to draw a line between several classes of small RNAs, because the nomencleature introduced early in the field not aware of the complexity of the small RNA pathways. The processing and loading of regulatory small RNAs onto distinct Argonaute proteins might differ among species (Table 1.1). For instance, in *A. thaliana*, different Argonaute proteins are associated with different classes of small RNAs, such as miRNAs, tasiRNAs and rasiRNAs (Dunoyer et al., 2007; Mi et al., 2008). In *C. elegans,* several Argonaute proteins are involved in different silencing pathways: ALG-1 and ALG-2 are essential for the miRNA pathway (Grishok et al., 2001); ERGO-1 and CSR-1 are essential for the endogenous RNAi pathway; RDE-1 is required for the exogenous RNAi pathway (Tabara et al., 1999; Yigit et al., 2006). Interestingly, *C. elegans* has numerous secondary Argonaute proteins, SAGO-1, SAGO-2, PPW-1, PPW-2, C16C10.3 and F58G1.1 (Yigit et al., 2006). Primary Argonautes (RDE-1, ERGO-1 and CSR-1) loaded with Dicer-derived primary siRNAs recognize target mRNAs and stimulate the production of secondary siRNAs through unprimed synthesis by RNA-dependent RNA polymerase (RdRP) (Pak and Fire, 2007; Sijen et al., 2007; Yigit et al., 2006). Secondary siRNAs have di- or tri-phosphates at their 5′ end compared to Dicer-derived primary siRNA, which carry 5′ monophosphate (Pak and Fire, 2007; Sijen et al., 2007), suggesting that secondary Argonautes may have a specificity for secondary siRNAs. Secondary Argonautes do not have a DDH catalytic triad in their PIWI domain compared to primary Argonautes. However, secondary Argonautes are important for both exogenous and endogenous RNAi pathways (Yigit et al., 2006). Recent studies in *Drosophila*, demonstrated that miRNAs are loaded into an AGO-1 complex and siRNAs loaded into the AGO-2 containing effector complex (Forstemann et al., 2007; Tomari et al., 2007). In human, miRNAs are incorporated into

AGO-1, 2, 3 and 4. However, AGO-2 is the only Argonaute that has endonuclease activity. Exogenously introduced siRNAs or endogenous miRNAs loaded into AGO-2 guide target cleavage by perfect complementarity (Meister et al., 2004). In mouse, AGO-2 also demonstrates catalytic activity, and *Ago-2 -/-* cells are unable to response to siRNAs to cleave a target (Liu et al., 2004).

In plants, miRNAs are thought to mostly mediate sequence specific mRNA target mediated by Argonautes. However, in animals miRNAs are partially complementary to their targets, and this is insufficient for sequence specific cleavage by Argonautes (Meister and Tuschl, 2004) . Extensive studies have been done to understand how miRNAs mediate translational regulation. Initial discovery in *C. elegans* showed that the miRNA *lin-4* binds to the 3′ untranslated region (UTR) of the *lin-14* mRNA and that this causes a dramatic reduction in LIN-14 protein level without affecting its mRNA level, indicating that miRNAs inhibit gene expression at a translational level (Lee et al., 1993; O'Donnell et al., 2005; Olsen and Ambros, 1999; Wightman et al., 1993). Additionally,, it has been shown that mammalian miRNAs binds to the 3′UTR of their targets and reduce the protein levels without changing the mRNA levels (O'Donnell et al., 2005). It is still unclear how miRNAs regulate gene expression.

It has been shown that miRNAs and Argonautes (miRNPs) co-sediment with polyribosomes, indicating translation is inhibited at the elongation step this might result in release off the nascent polypeptide chain from the polysome, followed by rapid degradation (Kim et al., 2004; Maroney et al., 2006; Nelson et al., 2004; Nottrott et al., 2006; Olsen and Ambros, 1999; Petersen et al., 2006). However, several studies suggest that Argonautes can inhibit translation at the initiation step. These studies demonstrate that for most of the miRNPs to mediate gene regulation, they requires a 5′ cap structure and a poly-A tail on the targeted mRNAs (Humphreys et al., 2005; Wang et al., 2006). Furthermore, they show that miRNPs do not affect cap-independent initiation using an IRES (internal ribosomal entry site) (Humphreys et al., 2005; Pillai et al., 2005). Recent studies strongly support miRNA-Argonaute complexes mediating translational inhibition at the initiation step in several ways. Firstly, it has been shown

in cell-free extract from flies that miRNAs inhibit the formation of the translationally active 80S ribosome by inhibiting the assembly of 43S initiation complex (Thermann and Hentze, 2007). Secondly, Argonaute (Ago2) binds to the m^7G cap through the MC domain and inhibits translation. Mutation of the MC domain of Ago2 results in a failure to bind to the m^7G cap and a loss of translational repression but without any affect on the catalytic activity of Ago2 and its assembly with the miRNA. Therefore, it suggests that Ago2 binding to the m^7G cap prevents binding of the eIF4E to the m^7G cap, disrupting assembly of the 43S initiation complex (Kiriakidou et al., 2007). Thirdly, eIF6, a ribosomal inhibitory protein inhibits assembly of 80S ribosome that leads to inhibition of translation (Chendrimada et al., 2007). Recently, work from the Izaurralde lab show that substitution of conserved binding site at MC domain of AGO-1 in *Droshophila* abolish AGO-1 silencing activity without affecting binding to the m7G cap and mutations of AGO1 prevents interacting with GW182, suggesting AGO1-GW182 interaction is essential for gene silencing by miRNAs (Eulalio et al., 2008). Aditionally, they show that eIF6 is not required for miRNA-mediated gene silencing in *Droshophila* (Eulalio et al., 2008).

Argonaute proteins that are complexed with miRNAs can also alter the stability of the targeted mRNA There is some evidence for RNA destabilization independent of translation inhibition, where a miRNA-targeted mRNA with a canonical cap and poly (A) tail is degraded by 5′→3′ decay after deadenylation (Giraldez et al., 2006; Wu et al., 2006). A study from *Drosophila* has shown that GW182 recruits the CCR4-NOT deadenylase and DCP1-DCP2 decapping complexes, which are responsible for the decay of miRNA-targeted transcripts (Behm-Ansmant et al., 2006). All decapping and deadenylating enzymes are localized at P-bodies, where mRNA turnover occurs (Behm-Ansmant et al., 2006). In yeast, it has been observed that mRNAs are moving between polyribosomes and P-bodies, and non-translating mRNA is localized to P-bodies for degradation (Brengues et al., 2005; Teixeira et al., 2005). The same scenario is observed in mammalian liver cells, where in normal conditions cationic amino acid transporter-1 (CAT-1) mRNA is repressed by the liver- specific miR-122 and is stored at P-bodies. CAT-1 mRNA is released from repression by HuR protein upon cellular

stress and localized to polyribosomes and is actively translated to produce CAT-1 protein (Bhattacharyya et al., 2006). An increasing amount of evidence suggests that miRNA-mediated translation repression and miRNA-targeted decay is uncoupled (Humphreys et al., 2005; Pillai et al., 2005; Wang et al., 2006). In flies, it has been shown that the level of RNA decay varies between individual miRNA-targeted transcripts, suggesting that other factors might regulate the stability of miRNP complexes and the turnover rate of the miRNA-targeted mRNA decay (Behm-Ansmant et al., 2006).

Number of computational and experimental approaches has been developed to find out miRNA targets (Bentwich, 2005; Easow et al., 2007; Hendrickson et al., 2008; Hofacker, 2007; Rajewsky, 2006; Sethupathy et al., 2006). 'Seed site' of the miRNAs, which are 2-7 nucleotides at the 5′ end of the miRNAs are base-paired with complementary sequences of the mRNA targets at their 3′UTR. The seed matches turn out important for miRNA-targeted mRNA regulation (Lewis et al., 2005). In order to understand the global endogeneous regulation of proteins by miRNAs, quantitative-mass-spectrometry-based approach was applied using SILAC (Stable isotope labelling with amino acids in cell culture) to investigate influcence of specific miRNAs (Baek et al., 2008; Selbach et al., 2008). Both studies show that a single miRNA downregulate production of several hundreds of proteins including miRNA predicted target proteins, suggesting a specific miRNAs can tune final scale of protein output.

1.3 The Piwi subfamily of Argonaute proteins

1.3.1 Expression of Piwi proteins and their roles in germ cells

The Piwi sub-family was named after the *Drosophila piwi* (P-element wimpy testis) gene, which has been implicated in germ cell development, stem cell self-renewal and maintenance of the division rate of germline stem cells (Cox et al., 2000). The

expression of the Piwi sub-family of Argonautes proteins is restricted to the germ cells in many organisms.

Tetrahymena, a ciliated protozoan, has two types of cell nuclei, a large somatic macronucleus and a small germline micronucleus. During development the macronucleus is formed from the micronucleus and specific regions of DNA are eliminated due to chromosomal rearrangement. TWI1, a *Tetrahymena* Piwi protein, is expressed prior to chromosomal rearrangement and has been implicated in the mechanism of DNA elimination (Mochizuki et al., 2002).

In the planarian *Schmidtea mediterranea*, two genes *smedwi-1* and *smedwi-2* both encode proteins that belong to the Piwi sub-family of Argonaute proteins. SMEDWI-1 and SMEDWI-2 are both expressed in germ cells and dividing adult stem cells (called the neoblast). RNAi of *smedwi-2* suggest their essential role in regeneration (Reddien et al., 2005).

In *Drosophila*, Piwis (Piwi, Aubergine (Aub) and Ago3) are expressed both in male and female germ cells. Piwi is also expressed in somatic cells surrounding the gonad (Cox et al., 1998; Cox et al., 2000). A mutation in *piwi* leads to severe defects in oogenesis, including loss of germline stem cells (Cox et al., 1998; Cox et al., 2000). Clonal studies indicate that Piwi is required for stem cell maintenance and division in the somatic cells. Loss of *piwi* in the gemline reduces stem cell division rates, but does not lead to a loss of stem cells or a block in oogenesis (Cox et al., 2000). Aubergine is required for the localization of posterior mRNA, such as oskar mRNA, but does not impinge on the localization of anterior mRNA, for instance bicoid mRNA. Therefore, it seems that Aubergine might play an important role in the axis specification development of *Drosophila* (Cook et al., 2004). A third Piwi gene in *Drosophila*, Ago3 has not been studied yet.

In zebrafish, Ziwi is expressed both in male and female gonads. Loss of *ziwi* results progressive loss of germ cells due to apoptosis, suggesting that Ziwi has a role in germ cell maintenance (Houwing et al., 2007).

Expression of three Piwi proteins in mice, MIWI (PIWIL1), MILI (PIWIL2) and MIWI2 (PIWIL4) is mainly restricted to the testes (Carmell et al., 2007; Deng and Lin, 2002; Kuramochi-Miyagawa et al., 2004; Kuramochi-Miyagawa et al., 2001). However, expression of Mili in prenatal ovaries has also been reported (Kuramochi-Miyagawa et al., 2001). Female *mili-/-* mice are fertile with no obvious abnormalities in the ovaries, whereas male *mili-/-* mice are sterile with smaller testes because of a spermatogenesis block at the first meiosis at the zygotene to early pachytene stage (Kuramochi-Miyagawa et al., 2004). Similarly, female *miwi-/-* mice are fertile, whereas males are sterile. Male *miwi-/-* null mice arrest spermatogenesis before the round spermatid stage (Deng and Lin, 2002). However, less is known about MIWI2. Expression of the rat Piwi protein, Riwi, is also restricted to the testes (Lau et al., 2006).

In Humans, Piwi protein is called HIWI, which is expressed in adult testes. The expression is restricted to spermatocytes and spermatids. Enhanced expression of Hiwi was found in seminomas, tumours derived from germ cells, whereas no expression of Hiwi was detectable in testicular tumours originating from somatic cells like sertoli cells and Leydig cells (Qiao et al., 2002).

In *C. elegans*, PRG-1 and PRG-2 are the two Piwi proteins. RNAi of *prg-1* leads to germ line defects and a shortening of the gonadal arms (Cox et al., 1998). However, extensive studies have not been done for these two Piwi proteins.

1.3.2 Piwi-interacting RNAs (piRNAs)

Recently, a novel class of small RNAs has been discovered that specifically interacts with the Piwi subfamily of Argonaute proteins. Therefore, these small RNAs are named

Piwi-interacting RNAs (piRNAs) (Aravin et al., 2006; Girard et al., 2006; Grivna et al., 2006a; Watanabe et al., 2006). piRNAs are 26-31 nucleotide long and are the most abundant small RNA species in testes. Miwi-associated piRNAs has been sequenced by high-throughput sequencing and around 26% of the sequences are found only once, indicating that population of piRNAs is very complex and diverse (Girard et al., 2006). 95% of the sequences start with a uridine (U) at their 5′ end position. In a parallel study, three small RNA libraries were prepared from MILI-interacting RNAs and small RNAs from mouse testes ranging between 18-26 nt and 24-33 nt were cloned and sequenced. Over 15000 sequences were identified and show that 26-28 nt RNAs are associated with MILI and additionally, 29-31 nt RNAs, which are not interacting with MILI are predominantly present in these libraries (Aravin et al., 2006). 85% of the MILI-interacting RNA contains a 5′ uridine residue. Few miRNAs, rasiRNAs and several piRNAs are also cloned from mouse oocytes (Watanabe et al., 2006). piRNA expression starts 14 days after birth (P14) during the pachytene stage of meiosis I of spermatogenesis, and these piRNAs are called pachytene piRNAs. Most recently, MILI-associated piRNAs have been discovered and these piRNAs start to be expressed before day 14 (P14), so they are called pre-pachytene piRNAs (Aravin et al., 2007b). Deep sequencing data from MILI–associated RNAs has shown that pre-pachytene piRNAs are less abundant compared to pachytene piRNAs and their genomic distribution is also different (Aravin et al., 2006; Aravin et al., 2007b). In rats, Piwi-containing complexes from testes were fractionated through chromatography, and these fractions contain mostly subpopulations of 29-30 nt RNAs, which are piRNAs and mostly begin with uridine (84%) (Lau et al., 2006). In *Drosophila*, Piwi-associated small RNAs are 25-29 nt in length. Most of the Piwi-associated small RNAs are derived from repetitive regions. Therefore, they have also been referred to as rasiRNAs (Saito et al., 2006). However, recently AGO2-associated endogenous siRNAs (~21 nt) have been discovered and a subset of endogenous siRNAs is also derived from transposons (Czech et al., 2008; Kawamura et al., 2008; Okamura et al., 2008), suggesting that piRNAs are a subset of rasiRNAs. Other Piwi proteins from *Drosophila*, Aub and AGO3 are also associated with small RNAs. Piwi and Aub bind to small RNAs that are predominantly antisense to transposons, whereas AGO3 binds

small RNAs that are sense to transposons. Piwi and Aub bound RNAs have a strong preference for 5′ uridine (83% and 72% respectively) but this trend is absent in AGO3 bound RNAs (37%) (Brennecke et al., 2007). In zebrafish, piRNA have been cloned from both testes and ovaries. piRNAs are ranging between 26-30 nt and they have a strong 'U' bias at their 5′ end (Houwing et al., 2007).

1.3.3 piRNA biogenesis

Mammalian piRNAs are present in the form of clusters and are enriched in intergenic regions. However, a small proportion of piRNAs map to repeat elements. The vast majority of piRNAs are clustered in genomic loci ranging between 1-100 kb and contain 10-4500 piRNAs (Aravin et al., 2006; Girard et al., 2006; Grivna et al., 2006a; Lau et al., 2006; Watanabe et al., 2006). However, pre-pachytene piRNAs are distributed over 900 smaller clusters, mostly mapping to retrotransposons (Aravin et al., 2006; Aravin et al., 2007b). The pachytene piRNA clusters and pre-pachytene clusters do not overlap with each other, suggesting that pachytene and pre-pachytene piRNAs are generated from distinct genomic loci (Aravin et al., 2007b). Mostly, mammalian piRNAs within a single cluster are derived from the same strand of DNA. This extreme strand bias indicates that piRNAs may be processed from a single long transcript. The regions covering piRNAs do not fold into stem-loop structures, suggesting they are distinct from miRNA pathways. In *Drosophila*, piRNAs are also clustered but they are derived mostly from repetitive regions, like retrotransposons and both pericentromeric and telomeric heterochromatic regions (Brennecke et al., 2007; Gunawardane et al., 2007; Saito et al., 2006). Piwi and Aub-interacting piRNAs have an antisense strand bias, whereas AGO3-interacting piRNAs that have a sense strand bias. AGO3-interacting piRNAs often have an adenine (A) residue at their 10th position from the 5′ end and this forms an A:U pair with the 5′ end uridine (U) residue at the 1st position of Piwi/Aub-interacting piRNAs. The first 10 nucleotides of Piwi/Aub-interacting piRNAs are complementary to the first 10 nucleotides of Piwi/Aub-interacting piRNAs. According to these observations, two groups independently proposed a speculative

"ping-pong" model by which piRNAs are amplified and maintain their piRNA pool (Brennecke et al., 2007; Gunawardane et al., 2007). However, this model does not provide any insight into how this amplification cycle starts, possibly by maternally contributed mature piRNAs. To date, there are no data available to support this hypothesis. Data from zebrafish and mice also suggest that a"ping-pong" mechanism might exist for a subset of piRNAs in these organisms (Aravin et al., 2007b; Houwing et al., 2007). In zebrafish, piRNAs are clustered to both repetitive and non-repetitive regions of the genome. Stretches of piRNA sequences map to one strand within a cluster, which may be up to several kilobases in size like mammalian piRNAs. Mostly piRNA clusters map to intergenic regions. However, piRNA clusters also map to repetitive regions of long-terminal repeats (LTRs) and DNA transposons (Houwing et al., 2007). 60% of the repeat-derived piRNAs map to the LTRs and they are antisense biased (in the case of a library prepared from ovaries compared to testes). Due to the strand bias of piRNAs in zebrafish, it is also thought that piRNAs are generated from a long primary transcript.

To date, studies from different organisms support that piRNAs are not generated from double-stranded precursors; rather that they are generated from primary transcripts and piRNA biogenesis is Dicer-independent (Houwing et al., 2007). Interestingly, individual piRNA sequences between mammals are not conserved, but the genomic regions from which they derive, particularly putative promoter regions, are highly conserved (Betel et al., 2007).

Studies from mammalian, zebrafish and *Drosophila* piRNAs demonstrate that the 3′ end of piRNAs are modified with a methyl group at the 2' or 3′ oxygen, most likely, 2' oxygen at the terminal ribose sugar (Horwich et al., 2007; Kirino and Mourelatos, 2007; Ohara et al., 2007; Saito et al., 2007). Because of this 2'-O-methyl modification at the 3′ end, piRNAs are insensitive to a β-elimination reaction (Vagin et al., 2006). This might serve to distinguish piRNAs from siRNAs and miRNAs. However, in plants siRNAs and miRNAs and in *Drosophila* endogenous siRNAs are modified at the 3′ end with a 2'-O-methyl group (Czech et al., 2008; Ghildiyal et al., 2008; Kawamura et al.,

2008; Li et al., 2005; Okamura et al., 2008). However, it is not reported whether mammalian endogenous siRNAs are modified. In plants, HEN-1, a RNA methyl transferase has been shown to be responsible for the 3′ end 2'-O-methylation (Li et al., 2005; Yu et al., 2005). A *Drosophila* HEN-1 homolog is DmHen-1/Pimet, which is a single-strand specific RNA methyltransferase that lacks the dsRNA-binding domain of plant HEN-1. Loss of *pimet* results loss of 2'-O-methylation at the 3′ end of the piRNAs. Recombinant Pimet is able to methylate specifically single stranded RNA *in vitro*. Additionally, recombinant Pimet binds to the piRNA-interacting Piwi, Aub and Ago3 complexes isolated from *pimet* mutant ovaries and methylates piRNAs within these complexes. However, Pimet does not methylate a Ago1-bound miRNA duplex or single stranded miRNAs (Horwich et al., 2007; Saito et al., 2007). These data suggest that 2'-O-methylation protects small RNAs from degradation and 3′ end modification is the final step of siRNAs or piRNAs in assembly into effector complexes. In contrast, the 5′ end of piRNAs is monophosphate as is the case for miRNAs and siRNAs (Kirino and Mourelatos, 2007; Ohara et al., 2007).

1.3.4 21U-RNAs

21U-RNAs are a novel class of small RNAs that were recently identified in *C. elegans* (Ruby et al., 2006). They are precisely 21 nt long and begin with a 5′ monophosphate uracil and are modified at their 3′ end. 21U-RNAs are clustered in the *C. elegans* genome. Uniquely, the vast majority of 21U RNAs map to two broad regions of chromosome IV, one region spanning between 4.5-7M and the other spanning between 13.5-17.2M. A third region between 9-9.7M contains few 21U-RNAs. Through high-throughput sequencing, almost 34,000 21U-RNA reads mapping to these three regions contained 5454 unique sequences. 21U-RNAs map to both strands of the DNA in a non-overlapping manner, suggesting they are not generated from a dsRNA precursor. Mostly 21U-RNAs map either between genes or within introns and with no strand bias. The density of annotated genes is lower in the 21U-RNA rich regions on chromosome IV compared to overall gene density. These 21U-RNAs are insensitive to a β-elimination reaction, suggesting their 3′ end terminal ribose sugar is modified at the 2'

or 3′ oxygen (Ruby et al., 2006). Other than uridine at the 5′ end, 21U-RNAs share no sequence similarity. Individual 21U-RNA sequences are not conserved between *C. elegans* and the related nematode *C. brigssae.* However most 21U-RNAs share an upstream sequence motif, which is conserved between *C. elegans* and *C. brigssae.* This motif is A-T rich and contains one large and one small motif. The large motif is 34 bp long and contains a central 8 nt core consensus sequence CTGTTTCA. The small motif has a YRNT core sequence, in which T corresponds to the 5′ U of the 21U-RNA. The large and small motifs are separated by 19-21 bp nucleotides. These upstream conserved motifs may serve as an individual promoter of a 21U-RNA or may act as a signal for precursor cleavage by an as yet undefined exo- or endonuclease.

1.4 Transposons

1.4.1 Different types of transposable elements

Transposons or transposable elements are nucleic acid sequences that are capable of both movements and propagation within host genomes (Kazazian, 2004). Their abundance and activity varies from species to species. For instance, in *S. cerevisiae*, 3% of the genome constitutes transposable elements, whereas the human and maize genomes comprise 50% and 80% transposons respectively (Flavell et al., 1974; Kim et al., 1998; Lander et al., 2001).

Transposable elements can be classified using several criteria; according to their strategy for movement, they are classified into class I or retrotransposons and class II or DNA transposons (Kazazian, 2004).

1) *Class I or retrotransposons*

Retrotransposons are transcribed into RNA, which is converted to DNA through a reverse transcriptase activity, followed by reintegration into the genome. Class I/ retrotransposons can be divided into two sub-types, on the basis of the presence or absence of long terminal repeats (LTRs) at the 5′ and 3′ ends of the transposable

elements, namely, long terminal repeats (LTR) and non-long terminal repeat (non-LTR) retrotransposons.

Long terminal repeats (LTR) retrotransposons has 'gag' and 'pol' genes, which are similar to to retroviruses. 'gag' encode for a virus particle coat and 'pol' encodes for a reverse transcriptase and a integrase for making cDNA from RNA and integrating into the genome. However, they do not code for the envelope protein that is required for packaging of virus particles and their exit from the cell (Figure 1.3A). LTR-retrotransposons comprise 8% of human DNA and most of all human LTR-retrotransposons are immobile, although a few HERV (human endogenous retrovirus) elements may remain active. In contrast, the mouse genome contains active LTR-retrotransposons, such as the intracisternal A particles (IAPs) (Medstrand and Mager, 1998).

Non-LTR retrotransposons are devoid of LTRs. Non-LTR retrotransposons are mostly LINE-1 (Long interspersed nucleotide elements-1) elements in mammals. LINE-1 constitutes 17% of the human genome (Lander et al., 2001). Full-length non-LTR retrotransposons are 4 to 6 kb in length and usually have two open reading frames (ORFs), one encoding an RNA binding protein and another encoding an endonuclease and a reverse transcriptase (RT) (Figure 1.3A). LINE-1 also has a 5′-untranslated region (5′UTR) containing an internal RNA polymerase II promoter, a 3′-untranslated region (3′UTR) and a poly-A tail. LINE or L1 usually flanked by 7-20 bp target site for duplication (TSD) (Deininger et al., 2003; Kazazian, 2004). Some of the non-LTR retrotransposons do not encode any proteins and are referred to as non-autonomous retrotransposons. Short interspersed nucleotide elements (SINE) represent non-autonomous retrotransposons in mammals (Slotkin and Martienssen, 2007). Alu elements are the major SINEs that constitute 11% of the human genome, whereas B1 and B2 elements are the major SINE families in the mouse genome. Both Alu and B1 elements are derived from the 7SL RNA, whereas B2 and most other SINEs are derived from tRNA genes. SINEs are generally very small (<300) bp and use their internal RNA polymerase III promoter for transcription (Figure 1.3A). These transcripts carry a

poly-A signal at the 3′ end and the sequence upstream of the poly-A signal is homologous to the sequence upstream of the poly-A signal in LINE-1 elements. Therefore, it is possible that the reverse transcriptase (RT) encoded by LINE-1 interacts with the shared upstream sequence of 3′ end and mobilizes SINEs in *trans*. Indeed, trans-mobilization of SINEs has been demonstrated by LINE-1 in cultured human cells (Kajikawa and Okada, 2002).

2) *Class II or DNA transposons*

DNA transposons do not require a reverse-transcription step to integrate into the genome. These elements are generally excised from one place in the genome and integrated into another site of the genome by a "copy-paste" mechanism. Active DNA transposons encode a transposase enzyme between inverted-termini (Figure 1.3B). The transposase enzyme recognizes the terminal inverted repeats (TIRs) of the DNA transposons at their 'old' donor sites and excises the transposons, then integrates them into their 'new' acceptor sites. The gap that is left at the donor site is repaired by DNA repair machinery (Kazazian, 2004). Some of the DNA transposons do not encode transposases and are called non-autonomous DNA transposons. Non-autonomous DNA transposons often consist of a pair of TIRs surrounding non-coding DNA transposon sequence. Additionally, in some of the non-autonomous DNA transposons, TIRs are present in a tail-tail orientation and they are called miniature inverted-repeat transposable elements (MITES) (Figure 1.3B) (Slotkin and Martienssen, 2007).

Helitrons are newly identified DNA transposons that replicate through a rolling-circle mechanism (Kapitonov and Jurka, 2007). Autonomous helitrons encode for a helicase as well as for a replicase (Figure 1.3B).

Transposons are also classified according to their self-sufficiency. Both retrotransposons and DNA transposons families have autonomous and non-autonomous elements (Figure 1.3A & B). Autonomous elements usually encode the protein required for their mobility whereas non-autonomous elements do not. Therefore, the mobility of

non-autonomous elements depends on other, autonomous elements (Slotkin and Martienssen, 2007).

The complexity of transposon poulations varies among organisms. For example, *S. cerevisiae* has only Ty LTR-retrotransposons. In mammals, DNA transposons are immobile, and non-LTR constitutes the most abundant class. *C. elegans* have both autonomous and non-autonomous DNA transposons but no retrotransposons. Plants possess both DNA and LTR retrotransposons. *Drosophila* and fish contain large numbers of DNA transposons and non-LTR retrotransposons.

A

Class I - Retrotransposons

B

Class II - DNA transposons
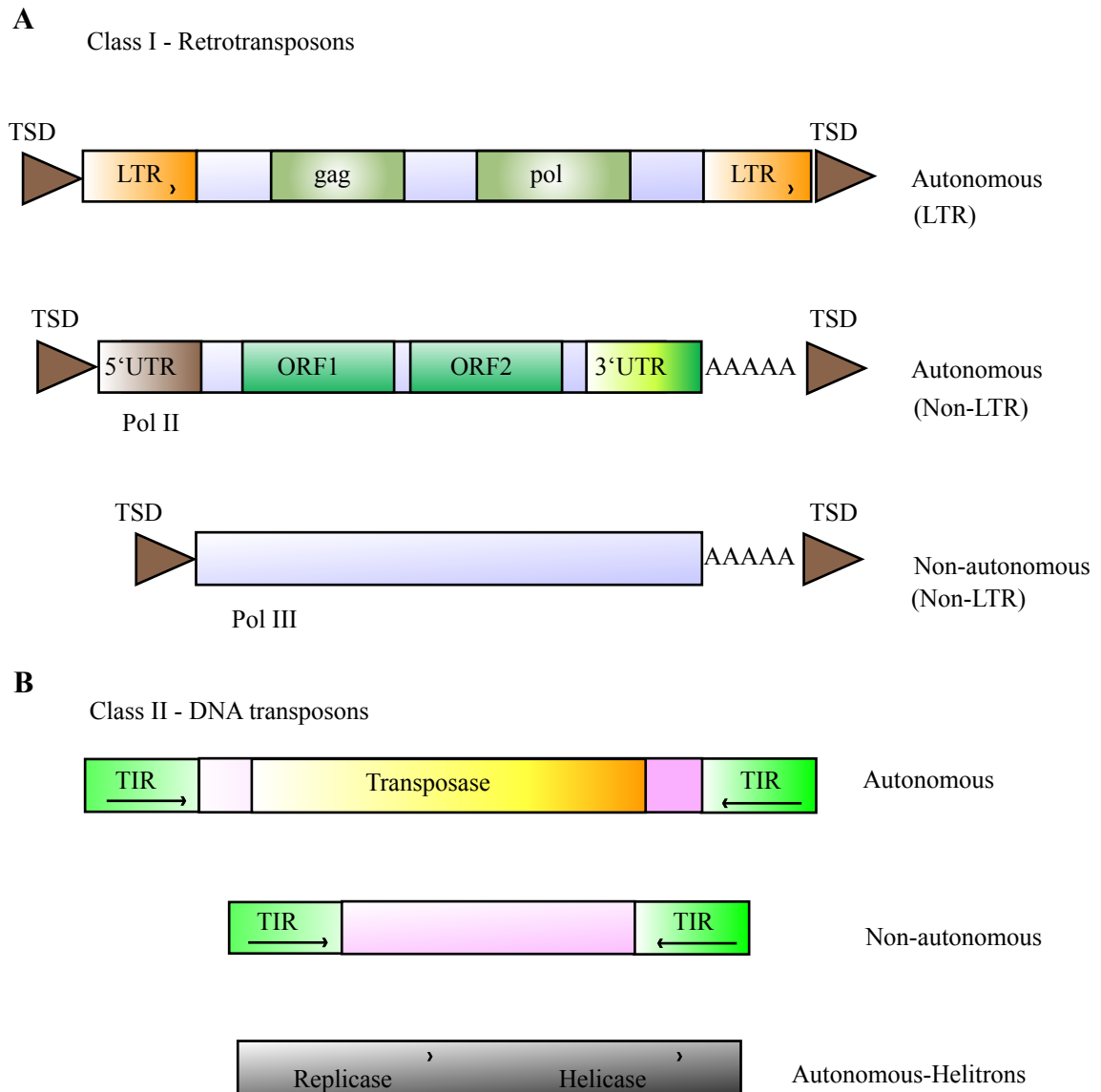
Figure 1.3 Different classes of transposons and their structures
(A) Class I or retrotransposons are divided into autonomous and non-autonomous retrotransposons according to the basis of their protein-coding capability and self-sufficiency. Autonomous retrotransposons are either LTR or non-LTR. LTR-retrotransposons encode gag and pol and they have LTR sequences at both ends of the transposable element. Non-LTR retrotransposons are devoid of LTRs at both ends. Usually they utilize internal Pol II promoters and produce a transcript, which encodes a RNA binding protein and a reverse transcriptase (RT). The transcript also carries a poly-A signal at the 3' end. Non-autonomous non-LTR transposons are usually very small (<300bp) and are transcribed using a RNA pol III promoter. The transcript does not encode a protein but carries a poly-A tail at the 3' end. TDS (terminal duplication sites). (B) Class II or DNA transposons either encode a transposase (autonomous) or do not encode for a transposase (non-autonomous), and are surrounded by terminal inverted repeats (TIRs). Autonomous Helitrons replicate through a rolling-circle mechanism and encode a replicase and a helicase protein.

1.4.2 Mechanisms of transposon silencing

Small RNAs are bound to the Argonaute proteins and Argonaute-small RNA complexes repress repeat sequences. Small RNA-directed cleavage of mRNA and transcriptional silencing through DNA and histone methylation are the mechanisms known to be involved in transposon silencing. Small RNAs can act *in trans* (for example, repression of *gypsy* by *flamenco*-derived piRNAs in *Drosophila*), called *trans*-silencing, or *in cis* (for example, centromeric silencing in *S. pombe*), called *cis*-silencing.

Tc1 Transposon regulation by siRNAs has been well characterized in *C. elegans*. Tc1 is one of the well-characterized DNA transposons in *C. elegan*s. 31 copies of Tc1 are present in Bristol N2 genome and they encode for a transposase for transposition. Tc1 transposition happens in somatic cells but in germ cells Tc1 transposition is completely suppressed (Sijen and Plasterk, 2003). Two flanking TIRs (54 bp) of a Tc1 DNA transposon fold back on each other and form a hairpin structure with a 54 nt double-stranded TIR, which is cleaved by Dicer (DCR-1) to produce 20-27 nt siRNAs. These siRNAs destroy the Tc1 mRNA by an RNAi mechanism and suppression of Tc1 is maintained in *C. elegans* germline. Mutants (for example *mut-7*) lacking Tc1 siRNAs allow transposition in germline (Sijen and Plasterk, 2003).

In *S. pombe*, transposon silencing is correlated with H3K9 methylation (Volpe et al., 2002). The RNA-induced transcriptional silencing (RITS) complex, contains Chp1, Ago1, Tas3, and centromeric siRNAs. A chromodomain protein-Chp1 binds to methylated lysine 9 of histone H3 (H3K9), and Argonaute binds to specific sequences of siRNA that base-pair with nascent centromeric mRNA. The RNA-directed RNA polymerase complex (RDRC), composed of Rdp1, the Hrr1 helicase, and the Cid12 Poly(A) polymerase family member, synthesizes double-stranded RNA and creates the substrate for Dicer to generate siRNAs. RDRC physically associates with RITS, and both complexes localize to noncoding centromeric RNAs and centromeric DNA repeats, suggesting that recognition of nascent RNA transcripts may be involved in

localization of these complexes to specific chromosome regions. The interaction between RITS-RDRC is Dicer and Clr4 methyltransferase dependent, suggesting siRNA and H3K9 methylation may be required for localization of RITS to the target chromosome regions (Motamedi et al., 2004). The C-terminal domain of Dicer physically interacts with RDRC. C-terminal truncations of Dcr1 that abolish its interaction with RDRC, but still generate siRNA *in vitro,* but abolish siRNA generation and heterochromatic gene silencing *in vivo*. Finally, reconstitution experiments show that the association of Dcr1 with RDRC strongly stimulates the dsRNA synthesis activity of RDRC, suggesting that heterochromatic dsRNA synthesis and siRNA generation are physically coupled processes and siRNA-mediated heterochromatin assembly is *cis*-restricted (Colmenares et al., 2007).

In *A. thaliana*, DNA methylation and siRNA-mediated chromatin modifications play important roles in transposon silencing. In plants, cytosines are methylated both symmetrically (CpG or CpNpG) and asymmetrically (CpNpNp), where N can be any nucleotide. These methylation marks are dependent on different methyltransferases. Met-1, a homolog of Dnmt-1 maintains CpG methylation, whereas plant-specific Cmt3 maintains CpNpG methylation. Dnmt3 homologs in Arabidopsis, Drm1 and Drm2, can methylate cytosines *de novo* in all sequence contexts. Met1 mutants gradually lose all CpG methylation, causing activation of transposons (Zhang et al., 2006). Transposon methylation is also strongly dependent upon chromatin modification and remodeling. SWI2/SNF2-like remodeling protein DDM-1 is required for transposon silencing. *ddm-1* mutants show a reduced level of DNA methylation on transposons (Jeddeloh et al., 1998; Lippman et al., 2004). Kryptonite (KYP) is a methyltransferase responsible for H3K9 methylation. *Kyp* mutants show a reduced level of non-CpG methylation, suggesting there is a connection between H3K9 modification and DNA methylation (Ebbs and Bender, 2006). However, the underlying mechanism is not fully understood. Heterochromatic silencing of transposons by siRNAs is achieved by plant-specific RNA polymerase IV. Loss of RNA PolIV can transcriptionally reactivate several silenced transposable elements. There are two RNA Pol IV polymerase, Pol IVa generates siRNAs and Pol IVb probably transfers the siRNAs to the Ago4 (Pontier et

al., 2005). It has been found that Pol IVb interacts with Ago4 in the nucleus presumably at DNA methylation target sites (Li et al., 2006), where the siRNA-loaded Ago4 complex promote silencing at different transposable elements (Qi et al., 2006).

In Drosophila, *piwi, aubergine* and *spindle-E* suppress position effect variegation (PEV) of a white reporter gene through H3K9 methylation. Mutations of *piwi, aubergine* and *spindle-E* genes lose H3K9 methylation and delocalize heterochromatin protein-1 (HP1) (Pal-Bhadra et al., 2004). HP1 is conserved in eukaryotes and it binds methylated H3K9, causing transcriptional repression (Bannister et al., 2001). HP1 also physically interacts with Piwi and the interaction is RNA dependent (Brower-Toland et al., 2007). However, the Lin group have been recently demonstrated that PIWI promotes euchromatic histone modifications and piRNA transcription at subtelomeric heterochromatin regions or telomere-associated sequences (TAS) at chromosome 3R (Yin and Lin, 2007). Piwi binds to 3R-TAS derived piRNAs, the expression of both 3R-TAS1 piRNAs and a white reporter gene in 3R-TAS becomes suppressed in *piwi* mutants. 3R-TAS loses euchromatic histone modifications and accumulates heterochromatic histone modifications and heterochromatin Protein 1a (HP1a) in *piwi* mutants. A P element inserted 128 base pairs downstream of the 3R-TAS1 piRNA coding sequence restores the euchromatic histone modifications of 3R-TAS and the expression of 3R-TAS1 piRNA in *piwi* mutants. These suggest that Piwi promotes the euchromatic character of 3R-TAS heterochromatin and its transcriptional activity, opposite to the known roles of Piwi and the RNA-mediated interference pathway in epigenetic silencing. In *Drosophila*, transposon regulation has also been shown by a master regulatory region containg piRNA clusters. *flamenco* was originally identified as a locus that controls the activity of the retroviral *gypsy* element (Pelisson et al., 1994). However, the molecular mechanism behind the *gypsy* regulation was a mystery. Surprisingly, piRNAs map to the *flamenco* region, which spans around 130 kb on chromosome X and 87% of the locus contains transposable elements, such as *gypsy*, idefix and ZAM. The *gypsy* RNA level is increased in *piwi* mutants compared to Aub and wild-type. piRNA expression is reduced specifically at the *flamenco* locus in

flamenco mutants compared to wild-type, suggesting piRNAs from the *flamenco* locus regulate *gypsy* elements (Brennecke et al., 2007).

Mammals also use DNA methylation and histone methylation to silence repetitive elements. In total 80% of the CpGs in the mouse genome are methylated and most of them are located close to repetitive regions (Gruenbaum et al., 1981). Unmethylated CpGs are grouped in clusters called 'CpG' islands that are present in the 5′ regulatory regions of many genes. Dnmt1 is proposed as a maintenance methyltransferase that is responsible for copying DNA methylation patterns to daughter strands during DNA replication. *Dnmt1*-null embryos reduce the methylation of IAP elements (Walsh et al., 1998). Dnmt3a and Dnmt3b are *de novo* methyltransferases, which create new methylation marks on DNA and are also responsible for methylation of IAP elements in male germ cells and embryos (Kato et al., 2007; Walsh et al., 1998). Dnmt3L is a DNA methyltransferase without catalytic activity that assists *de novo* methylation. Dnmt3L maintains methylation of LTR retrotransposons (IAP) and non-LTR transposons (LINE-1) in the male germline (Bourc'his and Bestor, 2004). To determine whether DNA methylation is linked to H3K9 modification to silence repetitive sequence, studies show that Suv39h HMTases are required for Dnmt3b-dependent DNA methylation at pericentric repeats but that DNA methylation at centromeric repeats occurs independently of Suv39h function, indicating additional complexity (Lehnertz et al., 2003). In addition, ES cells lacking the H3K9-specific Suv39h histone methyltransferase gene show upregulation of transposable elements (Martens et al., 2005). Similarly, like DNA methytransferases mutants, *mili* and *miwi2* knockout mice show a substantial reduction in DNA methylation of LINE-1 and IAP elements in testes, and activation of LINE-1 and IAP elements (Aravin et al., 2007b; Carmell et al., 2007). These observations suggest a role of mammalian Piwi proteins in TGS of transposable elements via an RNA-dependent DNA methylation mechanism.

1.5 *Caenorhabditis elegans*, its life cycle and germ line development

Caenorhabditis elegans is a free-living nematode (roundworm), about 1 mm in length and lives on soil. In the early 70s, Sydney Brenner started establishing *Caenorhabditis elegans* as a model organism. *C. elegans* is used as a model system in the laboratory for various reasons. In terms of genetic analysis, *C. elegans* is an attractive model for generating mutations in the genome. Several chemical mutagens are used for generating random mutaions in the genome. Therefore, multiple mutations such as complete deletion, partial deletion, insertion, point mutation can be generated for a gene, and that allows studying the complex gene functions. Overexpression study of a gene is also possible by making transgenic lines of *C. elegans*. In *C. elegans* it is relatively easy to disrupt the function of specific genes by RNA interference (RNAi) and that allow finding the function of genes. *C. elegans* have also been used successfully for biochemical analysis. *C. elegans* has the advantage of being a multicellular eukaryotic organism that is simple enough to be studied in great detail. It has around 1000 somatic cells in hermaphrodites and developmental fate of every single somatic cell has been mapped. Almost all individual *C. elegans* are hermaphrodite but very few are males. Therefore, major mode of reproduction is self-fertilization and easy to maintain a strain. However to generate a new strain, hermaphrodites can mate with males. Strains are cheap to maintain and can be frozen for long-term storage. Most importantly, *C. elegans* is the first multicellular organism, to have its genome completely sequenced. Complete genome sequence is approximately 100 Mb long and contains about 20,000 genes. Wormbase database attempts to collect all information about the *C. elegans* genome and published information on *C. elegans*. In addition, *Caenorahbditis* Genetics center is the great resources of all mutants and transgenic animals, which are available for experiments.

C. elegans are easy to grow under laboratory conditions and have a short life cycle. The life cycle is temperature-dependent. *C. elegans* goes through a reproductive life cycle in 5.5 days at 15°C, 3.5 days at 20°C, and 2.5 days at 25°C. *C. elegans* eggs are fertilized within the adult hermaphrodite (self-fertilization) and embryos layed within

few hours of fertilization. After hatching, development proceeds through 4 larval stages (L1-L2-L3-L4) to become young adults and then adults (Figure 1.4). Self-fertilizing adults produce about 300 progeny each. *C. elegans* can adopt an alternative life form, called the dauer larval stage. This stage is an adaptation to survive in extreme conditions such as starvation. Interestingly, in improved conditions, such as availability of food, daure worms resume normal development.
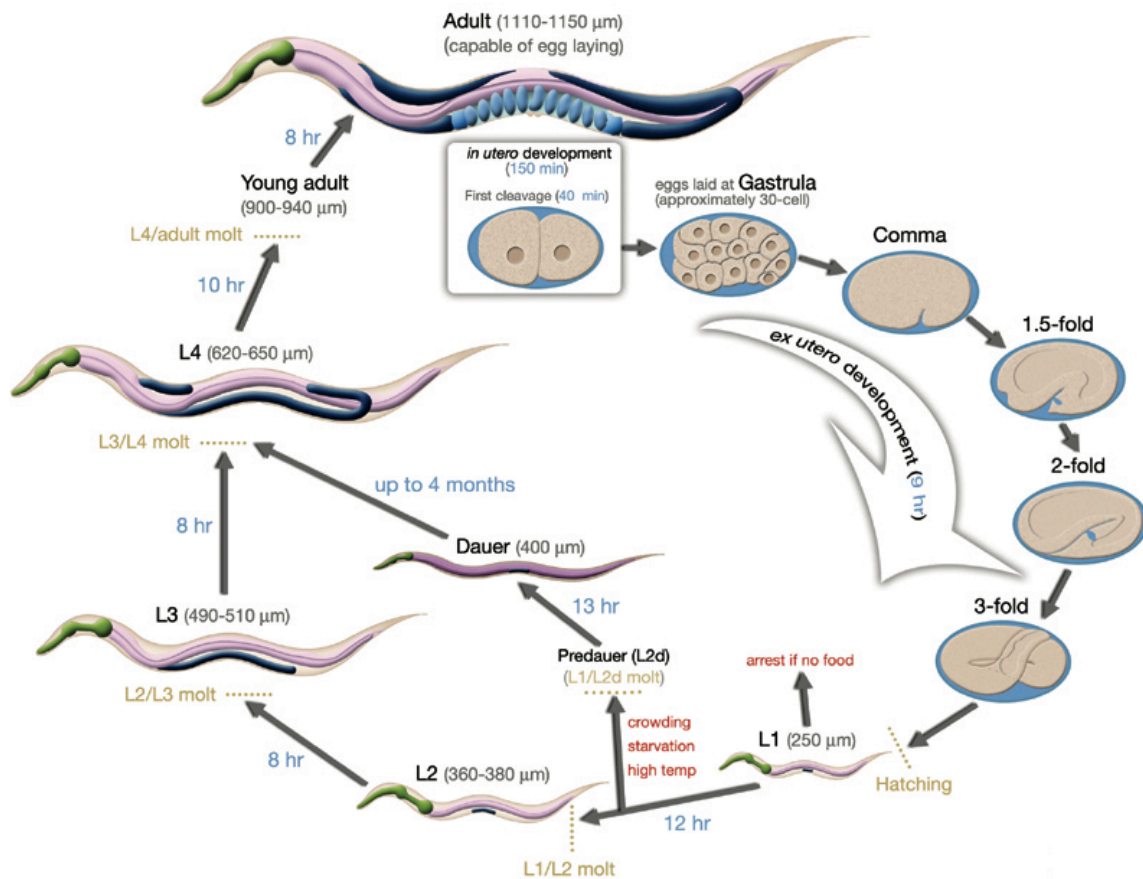
Figure 1.4 Life cycle of *C. elegans*
Life cycle of *C. elegans* at 20ºC. Numbers in blue along the arrows indicate length of the time animal spends at a certain stage. Length of the animals is indicated next to the stage name. Diferrent stages are indicated in bold letters namely, L1, L2, L3, L4, Young adult and Adult. This figure is adapted from Wormatlas.

C. elegans germline development can be divided into: specification, growth and maintenance. In early embryogenesis, germ cells are specified and become distinct from somatic cells. Specialized ribonucleoprotein particle, called P granules, which are maternally provided and segregated to the P blastomeres. Unequal divisions of germline blastomeres P_0, P_1, P_2, P_3 generate the AB, EMS, C and D somatic blastomeres and the first primordial germ cells P_4. P_4 divides equally into Z2 and Z3, which are not divide further until hatching (Sulston et al., 1983).The PIE-1 protein plays crucial role in specification of germline fate. PIE-1 is located in the nucleus each of the germline blastomeres, where it functions as a transcriptional repressor (Seydoux et al., 1996). MEP-1, a component of nucleosomal remodeling complex, is also required in somatic cells to repress germline fate (Unhavaithaya et al., 2002).

During larval stages, the germ line proliferates through mitotic divisions before entering into meiosis (Figure 1.5). At hatching, primordial germ cells Z2 and Z3 are flanked by Z1 and Z4 somatic gonad precursor cells. Z2 and Z3 cells begin dividing when nutrition is available. The number of germcells increases exponentially over the first two larval stages. During 3rd larval stage, most of the proximal germline nuclei enter into meiosis while distal nuclei are remain mitotic. Z1 and Z4 give rise to 12 cells; 2 distal tip cells and 10 proximal cells at the end of L1 stage. The 2 distal tip cells (DTC) are critical for mitotic proliferation of the germ line during development. Abalation of distal tip cells causes germ cells to leave mitotic cell cycle and enter into meiosis (Kimble and White, 1981), suggesting DTCs are necessary for germline proliferation during development and for maintenance of germline mitosis. Depletion of core components of the Notch signaling pathway, *lag-2* (ligand) and *glp-1* (receptor) causes germ cells to leave mitosis and enter into meiosis, suggesting role of Notch signaling pathway in germ cell proliferation (Austin and Kimble, 1987; Lambie and Kimble, 1991). 10 proximal cells form the hermaphrodite somatic gonad primordium in the late L2. The adult germ line possesses a 'mitotic region' at the distal end and a 'transition zone' at the proximal end (Figure 1.5). Germ cells in the mitotic region serve as stem cells, which has a self-renewal propety and producing differentiating gametes. There are few key regulators, GLD-1, 2, 3 and NOS-3 regulates the entry of germ cells

from mitosis into meiosis (Eckmann et al., 2004; Hansen et al., 2004; Kadyk and Kimble, 1998).

Gametogenesis starts at the late L4 and adult stage. Spermatogenesis began during the L4 stage and oogenesis during the adult stage. Spermatogenesis is similar in hermaphrodites and males but there are some significant differences, such as, male sperm are larger and they are stored in seminal vesicles whereas hermaphrodite sperm are stored in the spermatheca. *C. elegans* spermatozoa lack of flagella, they crawl during fertilization. In hermaphrodites, sperm are stored in the proximal gonadal arms, upon ovulation spermatids are pushed to the spermatheca, where spermatids become spermatozoa and fertilize the oocyte. Many spermatozoa are displaced to the uterus as the egg leaves the spermatheca and they crawl back into the spermatheca to fertilize the oocyte (Ward and Carrel, 1979). In males, sperm are stored in seminal vesicles and spermatids become spermatozoa when they are mixed with seminal fluid during ejaculation into the hermaphrodite uterus.
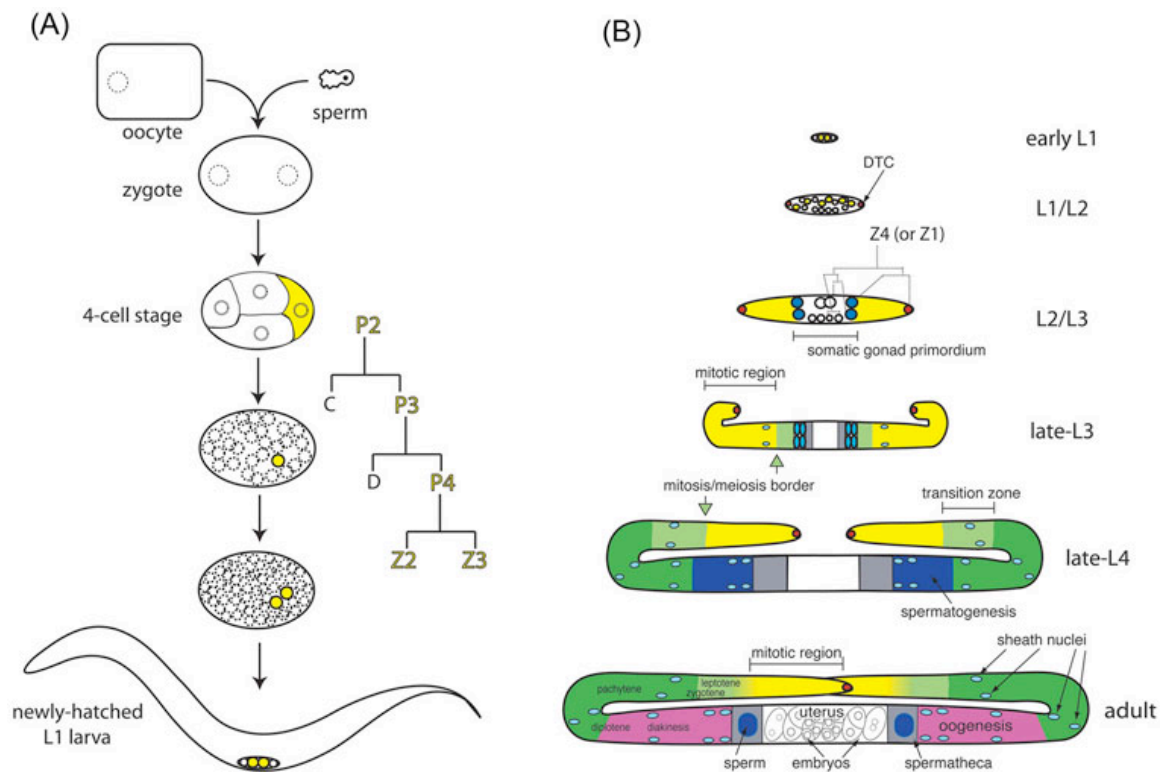
Figure 1.5 Schematic representation of development of germline
(A) Schematic representation of fertilization and the embryonic germ line. Oocyte and sperm meet and fertilization that initiates embryonic development. Germline lineages are depicted in yellow. (B) Schematic representation of post-embryonic hermaphrodite gonad development. Germlines are indicated in differen colours: yellow mitotic region, light green transition (early prophase of meiosis I), dark green pachytene, dark blue spermatogenesis, and pink oogenesis. The mitosis/meiosis border is indicated in the late-L3 and late-L4 by light green arrows. In the adult, the mitosis/meiosis border is not sharp (mitotic and meiotic nuclei are interspersed at the border) as indicated here by a yellow/green colour gradient. Somatic gonad color scheme: red DTC, blue sheath/spermatheca precursor cells, light blue sheath nuclei, grey spermatheca, and white uterus. This figue is adapted from Wormbook.

1.6 Aims of this thesis

So far I have discussed Arogonaute proteins, the Piwi sub-family of Argonaute proteins and their expression patterns and functions in development in different organisms. I have also discussed recently discovered Piwi-interacting RNAs (piRNAs) and their possible role in transposon silencing. However, in *C. elegans*, functions of Piwi proteins in development and their interacting RNAs are currently unknown. Therefore, I have been interested to investigate functions of Piwi proteins in *C. elegans*.

In this study I show that 1) *C. elegans* Piwi proteins have a conserved role in germline maintenance and development. 2) A class of 21 nucleotides RNAs, previously named as 21U-RNAs are piRNAs in *C. elegans* and they are associated with Piwi proteins. 3) piRNA biogenesis is independent of Dicer and genes from other small RNA pathways. 4) The expression of Piwi and piRNAs is restricted to both male and female germlines. 5) Piwi is specifically required for Tc3 transposon silencing in the germline. 6) There is no evidence of Ping-Pong model for piRNA amplification in C. elegans. 7) A Piwi-piRNA complex acts upstream of endogenous siRNAs in Tc3 silencing and there might be a link between piRNAs and endogenous siRNA function.

2. Materials and Methods

2.1 General materials

All chemicals were purchased from Sigma-Aldrich or Fisher Scientific. Restriction enzymes and other modifying enzymes were purchased from Roche and New England Biolabs (NEB). Radiochemicals were purchased from Amersham Biosciences or PerkinElmer. *C. elegans* strains used are shown in Table 2.2. The name and source of the antibodies used in this study are shown in Table 2.3. Oligonucleotide primers were synthesized by Sigma (Tables 2.4, 2.5, 2.6 and 2.7).

2.2 Bacteria related techniques

2.2.1 Bacterial media and growth

Autoclaved Luria-Bertani media (LB-media) (10 g/L tryptone, 5 g/L yeast extract, 5 g/L NaCl, pH 7.5) was used to grow liquid bacterial cultures by shaking at 225 revolutions per minute (rpm) at 37ºC for at least 12 hours. To select for bacteria transformed with a plasmid conferring resistance to ampicillin or kanamycin in liquid culture, these antibiotics were added to LB medium at room temperature prior to inoculation at final concentrations of 50 µg/ml or 100 µg/ml. To prepare plates containing ampicillin or kanamycin, these antibiotics were added to autoclaved LB-agar that had been cooled to less than 50ºC.

2.2.2 *E. coli* strains

Chemically competent XL-1 Blue (Stratagene), TOP10 (Invitrogen), DH5α (Invitrogen) and TOP10' (Invitrogen) cells were used in transformations for cloning

purposes. For expression of recombinant proteins, plasmid constructs were transformed with chemically competent BL21 (DE3)-CodonPlus cells (Stratagene).

2.2.3 Preparation of chemically competent bacteria

250 ml of SOB medium (0.5% yeast extract, 2% tryptone, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl2, 10 mM MgSO4 dissolved in water and autoclaved) was inoculated with a single colony of *E. coli* and the culture was grown at 19ºC with vigorous shaking until the OD_{600} reached 0.5 (normally it takes 20-24 hours). The culture was centrifuged at 4000 rpm for 10 min at 4ºC. The bacterial pellet was gently resuspended in 80 ml of ice-cold TB buffer (10 mM PIPES, 15 mM $CaCl_2$, 250 mM KCl were dissolved in nanopure water and the pH was adjusted to 6.7 and then 55 mM $MnCl_2$ was added. The final volume was adjusted with water, followed by filter sterilization with 0.45 µm filter and stored at 4ºC), incubated on ice for 10 min and pelleted again by centrifugation as before. The pellet was resuspended in 20 ml of ice-cold TB buffer and 1.4 ml of DMSO was added. Cells were aliquoted and stored at -80ºC.

2.2.4 Bacterial transformation

Cells were thawed on ice and the plasmid (usually 100 ng) was added to the 100 µl of competent cells in a tube, and incubated on ice for 30 min. The tube was then placed in a water bath at 42ºC for 45 sec and immediately chilled on ice minimum for 2 min. 200 µl of SOB medium was added and the bacterial culture incubated at 37ºC at 225 rpm for 1 hr before plating on LB-agar plates containing the appropriate antibiotics for selection.

2.2.5 Induction to bacterial cells for recombinant protein expression

A single colony of bacteria was used to inoculate a 20 ml LB culture with appropriate antibiotics and grown overnight at 37ºC. The following morning, 20 ml culture was

transferred to 500 ml - 2L LB medium containing suitable antibiotics and cells were grown at 37ºC until they reached OD_{600} of 0.4-0.6. Then IPTG (between 100 µM – 1mM, dependent upon expression of proteins) was added to the culture and it was grown at lower temperature (between 20-30ºC) for 4-6 hours. Cells were harvested by centrifugation.

2.3 *C. elegans* strains and culture

C. elegans was grown using standard conditions (W. B. Wood and the Community of *C. elegans* Researchers, 1988). Unless otherwise stated the food source used was *E. coli* strain HB101 (*Caenorhabditis* Genetics Center, University of Minnesota, Twin Cities, MN, USA). The wild-type strain was var. Bristol N2 (Brenner, 1974). Nematodes were grown at 20ºC, except where stated otherwise.

New deletion alleles for *prg-1* and *prg-2* were generated as described previously (Miska et al., 2007) and are listed in Table 3.1 and Figure 3.1. For a complete list of mutant strains used, see Table 3.2. All strains generated in this study have been submitted to the *Caenorhabditis* Genetics Center. For the developmental time course experiments animals were synchronized through bleaching followed by starvation-induced L1 arrest.

2.4 Molecular biology methods

2.4.1 Plasmid isolation (mini prep & midi prep)

3-5 ml (for mini prep) and 50-100 ml (for midi prep) of LB with appropriate antibiotics was inoculated with a single colony (after transformation) and grown at 37ºC overnight. Bacteria were pelleted by centrifugation at 4000 rpm for 10 min and the plasmid was isolated using QIAGEN plasmid mini kit (Qiagen) and QIAGEN plasmid midi kit (Qiagen) respectively according to the manufacture's instructions.

2.4.2 Genomic DNA isolation from *C. elegans*

Animals were harvested from plates by washing with M9 buffer (3g/L KH_2PO_4, 6g/L Na_2HPO_4, 5g/L NaCl, 1ml/L of 1MgSO4 dissolved in dH_20 and autoclaved) or from liquid cultures by floating on a sucrose cushion followed by washes with M9 buffer. Animals were pelleted by centrifugation at 2000 rpm for 2-3 min at RT (room temperature) and lysed in 5 volumes of lysis buffer (200 mM NaCl, 100 mM Tris-HCl pH 8.5, 50 mM EDTA and 0.5% SDS; aliquots were stored at -20ºC) with Proteinase K (final concentration was 0.1µg/µl) at 65ºC for 1-2 hrs and followed by incubation at 95ºC for 20-30 min. DNAse-free RNase (final concentration was 0.1µg/µl) was added to the lysate, incubated for 1hr at 37ºC. DNA was extracted with an equal volume of buffer saturated phenol, phenol:chloroform and phenol:cholorofom:isoamylalcohol. The aqueous phase was collected and DNA was precipitated with 0.1 volume of 3M sodium acetate and 2 volumes of absolute ethanol. DNA was pelleted by centrifugation at 14,000 rpm for 15 min at room temperature; the pellet was washed with 70% ethanol and dissolved in sterile water.

2.4.3 DNA amplification by polymerase chain reaction (PCR)

PCRs were performed using the either Taq DNA polymerase (NEB) or Phusion High-Fidelity DNA polymerase (NEB) according to the manufacturer's instructions. Phusion DNA polymerase was used for cloning. For Phusion High-Fidelity DNA polymerase, 50 µl of PCR reaction contained 1 µl of cDNA or DNA template (100 ng/ µl), 10 µl of 5X HF buffer, 1 µl of 10 mM dNTPs, 2.5 µl of 10 µM of forward primer and reverse primer, 0.5 µl (2 units/µl) of Phusion DNA polymerase and ddH_2O was added to a final volume 50 µl. Reactions were performed in an Eppendorf thermal cycler machine according to the following program: 98ºC – 30 sec (1 cycle); 98ºC – 10 sec, 55ºC – 30 sec, 72ºC – 1 min (10 cycles); 98ºC – 10 sec, 68ºC – 30 sec, 72ºC – 1 min (25 cycles); 72ºC – 7 min for final elongation.

For Taq DNA polymerase, 50 µl of PCR reaction contained 1 µl of DNA template (100 ng/ µl), 5 µl of 10X ThermoPol Reaction buffer, 1 µl of 10 mM dNTPs, 1 µl of 10 µM of forward primer and reverse primer, 0.5 µl (5 units/µl) of Taq DNA polymerase and ddH$_2$O was added to a final volume 50 µl.. Reactions were performed in a Eppendorf thermal cycler machine according to the following condition: 94ºC – 30 sec (1 cycle); 94ºC – 20 sec, 55ºC – 30 sec, 72ºC – 1 min (30 cycles); 72ºC – 7 min for final elongation. PCR products were visualised through agarose gel electrophoresis.

2.4.4 Agarose gel electrophoresis and purification of DNA

DNA was resolved using horizontal agarose gels. 0.8%-2.5% agarose gels were made in TAE buffer and ethidium bromide was added to this to a final concentration of 1 µg/ml. The gel was poured into a mould and allowed to set. The solidified gel was placed inside the electrophoresis tank and this was filled with 1X TAE buffer. Prior to electrophoresis, samples were diluted in 6X loading buffer (0.125% w/v bromophenol blue, 0.125% w/v xylene cyanol and 25% glycerol). DNA was visualised by UV illumination and DNA fragment size was estimated by comparing it to DNA markers (Bioline), which were resolved in parallel. For DNA purification, DNA was excised the gel with a clean razor blade. DNA was extracted from the agarose using QIAquick Gel Extraction kit (Qiagen) according to the manufacturer's instructions.

2.4.5 PCR product purification

PCR products were extracted from the PCR reactions using QIAquick PCR purification kit (Qiagen) according to the manufacturer's instructions.

2.4.6 Cloning

Genes were amplified by PCR from cDNA or genomic DNA by using specific primer pairs approximately 25 nucleotides long harbouring 5′ end restriction enzyme sites

appropriate for directional cloning into the target vector. The PCR product was purified with QIAquick PCR amplification kit (Qiagen) and then digested with suitable restriction enzymes (NEB / Roche) according to the manufacturer's instructions. The target vector was subsequently digested with the same restriction enzymes and dephosphorylated. The digested PCR product and target vector were loaded onto an agarose gel, resolved and purified via the QIAquick Gel Extraction Kit (Qiagen) according to the manufacturer's instructions. The purified PCR product and target vector were ligated using a Rapid DNA ligation Kit (Roche) according to the maufacturer's instructions. 7.5 µl of each ligation was transformed into XL-Blue or DH5α cells bacterial cells. Transformants were checked by either colony PCR or restriction digestion. Positive clones were confirmed by sequencing.

The ORF of *prg-1* was PCR amplified from cDNA (cDNA was made from total RNA using random primers (Invitrogen) using SuperScript II reverse transcriptase (Invitrogen)) and cloned into pDONR221 entry vector (Invitrogen) using Gateway Techology (Invitrogen) according to the manufacturer's instructions. Different fragments of the *prg-1* ORF were sub-cloned to pDEST-MBP using also Gateway technology.

2.4.7 DNA Sequencing

DNA sequencing reactions were performed by Cogenics (Essex, United Kingdom). Computational analysis was performed using the MacVector program.

2.4.8 RNA isolation

For total RNA isolation animals were harvested from plates by washing with M9 buffer or from liquid cultures by floating on a sucrose cushion followed by washes with M9 buffer. Animals were pelleted and frozen in liquid nitrogen and dissolved in ten pellet volumes of Trizol reagent (Invitrogen). Worms were broken up by three freeze-thaw cycles in liquid nitrogen and water followed by continuous vortexing for 30-45 min at

4ºC. 0.2 volumes of chloroform was added to the Trizol solution, mixed and incubated at room temperature for 3-5 min. The mixture was centrifuged at 14,000 rpm for 15 min at 4ºC and the aqueous phase was transferred to a new tube. Isopropanol was added to the aqueous solution to precipitae the RNA (50% of the volume of Trizol reagent used). Additionally, 1 µl of glycogen (20 µg/µl) was added as a carrier for RNA precipitation. Precipitation was carried out by centrifugation at 14,000 rpm for 30 min at 4ºC. The pellet was washed with 80% ethanol, dried and dissolved in RNase-free water.

The *mir*Vana miRNA isolation Kit (Ambion) was used to isolate small RNAs according to the manufacturer's instructions with some modifications. After worms were broken up, chloroform was added to the Trizol mixture, centrifuged and the aqueous phase was collected in a fresh tube. 1/3 volume of 100% ethanol was added to the aqueous phase and mixed thoroughly by vortexing. The sample was passed through the first filter cartridge (Ambion) by centrifugation at 10000 rpm for 15 sec at room temperature and the flow through (filtrate) was collected to a fresh tube. 2/3 volume of 100% ethanol was added to the filtrate, mixed thoroughly and passed through the second filter cartridge (Ambion). Flow-through was discarded and the filter was washed once with 700 µl of solution I (Ambion) followed by two washes with 500 µl of solution 2/3 (Ambion). RNA was eluted in pre-warmed (95ºC) RNase-free water from the filter. The concentration of RNAs was measured at OD_{260} by spectrophotometry using a NanoDrop ND-1000 spectrophotometer.

2.4.9 3′ end labelling of total RNA

100 ng of total RNA was 3′ end labeled with 5′[-α^{32}P]5′-3′ cytidine bis-phosphate (PerkinElmer). Each reaction (final volume 50 µl) was set up as follows: 1 µl of total RNA (100 ng/µl), 5 µl of 5′-32P 5′-3′ cytidine bis-phosphate (PerkinElmer), 5 µl of 10X RNA ligase buffer (MBI Fermentas), 15 µl of 100% DMSO, 2 µl of T4 RNA ligase (MBI Fermentas) and 24 µl of nuclease-free water was mixed and incubated overnight at 4°C. Unincorporated radioactivity was removed through G-25 microspin

columns (Amersham) and 2X loading buffer (8 M urea, 50 mM EDTA, bromophenol blue and xylene cyanol) was added. For size markers, Decade marker (Ambion) was 5′ end labeled according to the manufacture's instructions and 18/26 oligo mix was 5′ end labeled as follows: 2 µl of 18/26 oligo mix (10 pmole/µl), 5 µl of P^{32} γ-ATP (3000 Ci/mmole, 10 µCi/µl) (Amersham), 2 µl of 10X T4 PNK buffer (NEB), 2 µl of T4 PNK (NEB) and 9 µl of ddH$_2$O incubated at 37 ºC for 1 hr. The reaction was inactivated at 65ºC for 10 min and 2X loading buffer also added. 3′ end labeling reactions were resolved on a 18% denaturing polyacrylamide gel (Sequagel mix, National Diagnostics) along with 5′ end labeled Decade marker and the 18/26mer oligo mix. Gels were run at 20W for 4 hrs (initial voltage was 1000V and final voltage was 1498V), gels were wrapped in cling film and exposed to a PhosphoImager screen that was scanned through the Fujiflim FLA-5000 scanner (Fuji, Japan).

2.4.10 High-throughput sequencing

12-24-hour adult animals were harvested, total RNA was extracted using Trizol and small RNAs were isolated using the *mir*Vana miRNA isolation Kit (Ambion). For 5′-dependent libraries 10µg RNA of wild type and *piwi (n4357; n4358)*, *mut-7* and *dcr-1* mutants were size-selected to 18-30 nucleotides on a denaturing polyacrylamide gel and small RNA libraries were cloned through ligation of a 5′- and 3′-adapter according to the manufacturer's protocol (Illumina). After 5′ and 3′-adapter ligation, cDNA systhesis was performed using specific primer against 3-adapter through reverse transcription by superscript II reverse transcriptase (Invitrogen) according to the manufacturer's instructions. Then, cDNA was amplified through PCR using high fidelity Phusion DNA polymerase (NEB). For PCR primer pairs were designed against adapter sequences. PCR was performed according to the following condition: 98ºC – 30 sec (1 cycle); 98ºC – 10 sec, 58ºC – 30 sec, 72ºC – 20 sec (20 cycles); 72ºC – 5 min for final elongation. PCR product (~92 bp) was eluted from the gel, precipitated and dissolved in resuspension buffer (10mM Tris-Cl, pH 8.5). The concentration of the libraries was measured using NanoDrop at OD$_{260}$ and crosschecked by running 1 µl of the PCR

product on a gel. Libraries with 10nM final concentration were prepared for sequencing using an Illumina/Solexa 1G instrument.

Generation of these libraries as described above is dependent on the presence of a monophosphate at the 5′ end of the small RNAs. In *C. elegans* secondary endogenous siRNAs carry a triphosphate at their 5′ end. Therefore the method using to make the 5′-dependent library is not useful to clone these secondary siRNAs. To avoid this problem, a 5′-independent libraries of small RNAs were generated as described previously (Houwing et al., 2007) but with several modifications: RNA was size-selected to 19-29 nucleotides on a denaturing polyacrylamide gel. Small RNAs were treated with tobacco acid pyrophosphatase (Epicenter) to convert 5′-triphosphates to 5′-monophosphates, poly (A)-tailed using poly(A)-polymerase and ligated to a 5′-adapter. First-strand cDNA synthesis was performed using an oligo(dT) primer and M-MLV-RNase H⁻ reverse transcriptase (ABI). Incubation temperatures were 42°C for 15 min, ramp to 55°C followed by 55°C for 5 min. The resulting cDNAs were then PCR-amplified using a high fidelity DNA polymerase and 17 PCR cycles. All libraries were sequenced using an Illumina/Solexa 1G instrument.

2.4.11 Northern blotting of small RNAs

Small RNA northern blotting was performed using either DNA probes or LNA probes. The protocol used for Northern blotting using DNA probes is as follows: 15% sequaGel mix (15 ml of mix was prepared by adding 9 ml Concentrate, 4.5 ml of Diluent and 1.5 ml of Buffer (National Diagnostics) was prepared, poured the gel into criterion cassettes (Bio-Rad) and polymerized for 1-2 hrs. Gel was pre-run for 30 min-1 hr in 0.5X TBE buffer. 20-40 μg of total RNA or small RNA was mixed with 2X loading buffer (8 M urea, 50 mM EDTA, bromophenol blue and xylene cyanol). Samples were heated to 80ºC for 5-10 min, spun down and loaded to the wells after ensuring urea was washed off from the wells. The gel was run in 0.5X TBE buffer at 180-200 Volts until Bromophenol blue dye just runs off the gel and Xylene cyanol was still in the middle of the gel. After running the gel, it was stained with Ethidium bromide solution in 0.5X

TBE for 5 min and RNA was visualised by UV translluminator. Stains of tRNA (~78 nt) and 5S rRNA (~120 nt) bands were observed and used as a loading control. Gel was washed briefly with 0.5X TBE for 2-3 times. Nylon membrane (Genescreen Plus) and 3 mm Whatman chromatography filter paper (Fisher Scientific Ltd) was cut according to the size of the gel and soaked in 0.5X TBE buffer briefly just before the sandwich. Standard sandwich was prepared as following: 3 sheets of soaked Whatman filter papers were placed on the bottom platform of the semidry blotter (Hoefer system, Amersham Pharmacia) then gel (upside down), Nylon membrane and 3 sheets of Whatman filter papers. Any bubbles were removed rolling a plastic pipette over the sandwich. RNA was transferred to the Nylon membrane at constant current (3.3 mM/cm2) for 45 min to 1 hr. Membrane was marked to indicate orientation, cross-linked by UV radiation (1200 µJ, auto cross-linking setting) (UV Stratalinker, Stratagene) and baked for 2 hrs at 80ºC. Each membrane was pre-hybridized in 25 ml hybridization buffer (5X SSC, 20 mM Na_2HPO_4, pH 7.2, 7% SDS, 2X Denhardt's solution and 1 mg of denatured salmon sperm) for 1-2 hr at 50ºC. In the mean time, the probe was labelled by 5′ end labelling method as following in a 50 µl reaction: 2 µl of 10 µM oligo (DNA probe), 5 µl of 10X T4 Polynucleotide Kinase buffer (PNK buffer), 2 µl of T4 PNK (NEB), 5 µl of P^{32} γ-ATP (3000 Ci/mmole, 10 µCi/µl) (either GE healthcare or PerkinElmer) and 36 µl of water were mixed together and incubated for 1 hr at 37ºC. The reaction was heat inactivated at 65ºC for 10 min and purified through G-25 MicroSpin column (Amersham Pharmacia) according to the manufacturer's instructions. Half of the labelling reaction (25 µl) was added to the blot in hybridization buffer and hybridized for overnight at 50ºC. Each membrane was washed twice with 40 ml of 1st wash buffer (3X SSC, 25 mM Na2HPO4, 5% SDS and 10X Denhardt's solution) for 10 min each, followed by another two washes with same buffer for 30 min each. Finally, membrane was washed once with 80 ml of 2nd wash buffer (1X SSC and 1% SDS) for 10 min. All the washes were done at 50ºC. Membrane was dried, saran wrapped and exposed to the PhosphoImager screen and scanned through the Fujiflim FLA-5000 scanner. Signals were quantified through Mac-Bas program.

Expression level of 21U-RNAs is lower compared to other small RNA species. Therefore, conventional northern blotting method is less sensitive, time-consuming and it needs lot of initial stating material to detect 21U-RNAs. To overcome this problem, a newly developed northern method was used with some modifications to detect small RNAs (Pall and Hamilton, 2008). 15% of denaturing polyacrylamide gel [60 ml mix; 22.5 ml of 40% Acrylamide (19:1) - Accugel (National Diagnostics), 25.2 gm of urea (Fluka), 1.2 ml of 50X MOPS, pH 7 (1 M MOPS (Sigma) was dissolved in distilled water and pH was adjusted to 7 by NaOH, stored at 4ºC and water was added to make final volume 60 ml] was poured into the criterion cassettes (Bio-Rad), polymerized for 1-2hrs. Gel was pre-run for 3-4 hrs in 20 mM MOPS, pH 7) buffer. Sample was prepared in 2X urea loading buffer, heated at 80ºC for 5-10 min and loaded into the wells after removing urea from the wells. The gel was run at 180-200 Volt for 1-2 hrs or 60-90 Volt for overnight in 20 mM MOPS-NaOH (pH 7) buffer. The RNA was transferred to the neutral Nylon membrane, Hybond NX (Amersham/Pharmacia) in 20 mM MOPS-NaOH (pH 7) buffer as above. Membrane was briefly washed with water 2-3 times and placed RNA side up onto 3 Whatman filter papers saturated with freshly prepared EDC reagent [0.16 M EDC (Sigma) was prepared in 0.13 M 1-methylimidazole (Sigma) at pH 8. 245 µl of 12.5 M 1-methylimidazole was added to 9 ml DEPC-treated water, pH was adjusted to 8 with 1 M HCl. Immediately before use, 0.753 g EDC was added and made the volume up to 24 ml with DEPC treated water]. The membrane and Whatman filter papers were wrapped in Saran and incubated at 60ºC for 2 hrs. After cross-linking, the membrane was washed with RNAse-free water, dried. Pre-hybridization, hybridization and washing was performed as above.

Alternatively, LNA probe (Exiqon) was designed for 21UR-1 and due to their special modification in the nucleic acid; detection of the 21UR-1 through Northern blotting is sensitive. After RNA was transferred to the Nylon membrane (Genescreen Plus), cross-linked through UV cross-linking followed by baking at 80ºC for 2 hrs. Each membrane was pre-hybridized in 25 ml of hybridization buffer (0.36 M Na_2HPO_4, 0.14 M NaH_2PO_4, 1 mM EDTA, 7% SDS and 1mg of denatured salmon sperm DNA) at 60ºC for 1hr. 20 pmole of a P^{32} γ-ATP labelled LNA probe was added to the hybridization

buffer and hybridized at 60ºC for overnight. After hybridization, membrane was washed once with 2X SSC, 0.2% SDS at 50ºC for 15 min, once with 0.5X SSC, 0.1% SDS at 50ºC for 15 min and once with 0.1X SSC, 0.1% SDS at 65ºC for 5 min.

2.4.12 RT-PCR

For small RNA RT-PCR primer design was as described previously (Chen et al., 2005). 100ng of total RNA was used for each reverse transcription reaction with 50nM stem-loop RT primer, 1XRT buffer (Arrayscript, Ambion), 1 mM dNTPs, 3.33U/μl of MMCV reverse transcriptase (Arrayscript, Ambion) and 0.25U/μl RNase inhibitor (Ambion) in a final volume of 7.5μl. The reaction was incubated at 16ºC for 30 min, 42ºC for 30 min and 85ºC for 5 min. Real-time PCR was performed using Quantitect SYBR green PCR mix (Qiagen). For each reaction, 10μl of 2X Quantitect SYBR green PCR mix, 0.2μl of 100μM specific forward primer, 0.2μl of 100μM universal reverse primer, 8.27μl of RNase-free water and 1.33μl of RT product was incubated at 95ºC for 10 min, followed by 40 cycles of PCR (95ºC for 15s, 60ºC for 1 min) on a 7300 Real Time PCR System (Applied Biosystems, Foster City, CA, USA). All reactions were run either in duplicate or triplicate. Alternatively, a standard PCR was carried out and products were resolved on a 10% denaturing polyacrylamide gel and stained using SYBR green or PCR products were cloned and sequenced. For mRNA RT-PCR, total RNA was incubated with TURBO DNA-*free*, (Ambion) at 37ºC according to the manufacturer's instruction. 100ng of DNA-free RNA was used for reverse transcription, using random primers and Superscript RT II (Invitrogen) according to manufacturer's instructions and PCR was performed as described above. All primers used for RT-PCR are listed in Table 4.

2.4.13 RNase protection assay

Fragments of Tc1, Tc3 and U6 snRNA were either cloned into pGEMTeasy (Promega, Madison, WI, USA) or PCR product used as a template for *in vitro* transcription. In

case of plasmids, they were linearized, gel purified and phenol-chloroform extracted prior to *in vitro* transcription. PCR products were used directly after gel purification and phenol-choloform extraction. *In vitro* transcription was performed using MaxiScript III kit (Ambion). For each *In vitro* transcription reaction, 500 ng to 1 µg of DNA template, 2 µl of 10X transcription buffer (Ambion), 1 µl of 10 mM ATP, 1 µl of 10 mM CTP, 1 µl of 10 mM GTP, 5 µl of α-p32-UTP (3000Ci/mmole) (PerkinElmer), 2 µl of either T7 or SP6 promoter (20U/µl) (Ambion) and RNase-free water (up to final volume 20 µl) was incubated at 37ºC for 1 hr. After that, each reaction was incubated with 1 µl of TURBO-DNase to remove DNA templates, followed by addition of 1 µl of 0.5 M EDTA to stop the DNase activity. *In vitro* transcribed radiolabelled RNA probe was mixed with 2X loading buffer, resolved through denaturing polyacrylamide gel and exposed to the X-ray film. The desired band was excised from the gel and eluted in 350 µl of elution buffer (0.5 M NH_4 Ac, 0.2% SDS, 1 mM EDTA) at 37 ºC for overnight. The RNA was then ethanol precipitated with 3 volumes of absolute ethanol and 1 µl of glycogen at -80ºC for 30 min. The pellet was washed with 80% ethanol, dried and dissolved in 15 µl of RNase-free water.

Probes used were Tc3-N (bases 5-212 of transposase), Tc3-464 bp (bases 270-733 of transposase), Tc3-C (bases 698-913 of transposase), Tc1-370 bp (bases 523-893 transposase), U6 snRNA (bases 1-102). Each RNase protection assay (RPA) was performed as follows: 15-30μg of small RNAs (usually 4-5 µg/µl), 3 µl of probe, 1 µl of U6 probe (as an internal control) and hybridization buffer (53mM PIPES, pH6.4, 1.33mM EDTA, 0.53M NaCl, 67% formamide) (up to final volume 12 µl) was incubated at 46ºC for overnight. For each reaction, RNA was digested with 150 µl of RNase A/T1 (1:100 dilution in RNase digestion buffer (miRVana miRNA detection kit, Ambion)) for 30min at 37ºC. The reaction was stopped by adding 225 µl of RNase inactivation / PPT solution (miRVana miRNA detection kit, Ambion). RNA was precipitated, pellet was dissolved in urea loading buffer, resolved in 12% Sequagel (National Diagnostics) and exposed to X-ray film.

2.4.14 Microarray

Total RNA was prepared from synchronized young adults (12 hr adult) from wild-type and two individual *piwi* mutants using Trizol extraction method as described in section 2.4.8. Three biological samples were used. 5 μg of total RNA was amplified and labelled (Cogenics) according to their protocol, and hybridized to a GeneChip array (Affymetrix) contains probes for 22,500 transcripts from *Caenorhabditis elegans*. The expression of genes was normalized by expression of total number of genes and fold change of *piwi* mutants was calculated relative to wild-type.

2.5 Biochemical analysis

2.5.1 Purification of Recombinant proteins

PCR products corresponds to the N-terminal first 60 amino acids (1-60 aa), 1-241 aa, PAZ domain (218-354 aa), PIWI domain (499-810 aa) and C-terminal 738-824 aa of PRG-1 were sub-cloned from pDONR221-PRG-1-ORF plasmid to GST containing pGEX-5X-1 plasmid (Amersham). Simultaneously, PCR product corresponds to N-terminal first 60 amino acids (1-60 aa) and PAZ domain (218-354 aa)were also sub-cloned from pDONR221-PRG-1-ORF to pQE-30 and pMAL-c2X plasmid that carry N-terminal 6X-HIS tag and MBP tag sequences respectively. PCR products corresponds to 1-60 aa and PAZ domain of PRG-1 was cloned into different plasmids containing different tags because of antibody purification and to test the specificity of the antibodies against recombinant proteins. For *in vitro* interactions study of PRG-1 protein, various cDNA fragments of the PRG-1 protein (1-936 bp, 1-633 bp, 652-1062 bp, 1027-1575 bp, 1495-2430 bp, 1495-2007 bp, 1834-2475 bp and full-length 1-2475 bp) were sub-cloned by gateway cloning system from pDONR221-PRG-1-ORF to pDEST-MBP (pEM105) vector, which carries an N-terminal MBP tag sequence.

The constructs carrying pGEX-5X-1-(1-60aa) and pGEX-5X-1-PAZ were transformed into BL21 (DE3) CodonPlus (Statagene) *E. coli*. A single colony of bacteria was used to inoculate in a 20 ml LB culture with appropriate antibiotics and grown overnight at 37ºC as described in section 2.2.5. Bacterial cells were induced with 250 µM IPTG at 30ºC for 5 hrs and 100 µM IPTG at 25ºC for 24 hrs to express GST-(1-60aa) (pGEX-5X-1-(1-60aa)) and GST-PAZ (pGEX-5X-1-PAZ) recombinant proteins respectively. The bacterial cells were pelleted by centrifugation at 4500 rpm for 10 min at 4ºC and resuspended in 40ml of 1X PBS (150 mM Nacl, 2.5 mM KCl, 10 mM Na_2HPO_4, 2 mM K_2HPO_4) with EDTA-free protease inhibitor tablet (Roche). Lysozyme was added to final concentration of 0.25mg/ml to lyse the cells for 1 hr at cold room on rotation. The cell suspension was sonicated on ice with 30 sec pulse and 30 sec interval for 4 times using sonicator. The lysed cells were further incubated with Triton (final concentration 1%) for 30 min, centrifuged at 14000 rpm for 10-15 min at 4ºC. The supernantant was transferred to a 50 ml falcon tube containing 2ml of 50% slurry of glutathione-sepharose beads (GE Healthcare) and rotated for 1-2 hrs at 4ºC to bind the fusion protein to the beads. Beads were pelleted by centrifugation at 1500 rpm for 3 min at 4ºC, supernatant containing unbound fractions was removed and the beads were washed with chilled 1X PBS by centrifugation at 1500 rpm for 3 min. After that, beads were resuspended in 1X PBS gently and poured onto a protein purification column (Bio-Rad). Beads were washed with chilled 1X PBS (usually 10 times of the bed volume of the beads) and GST-fusion proteins were eluted with 50mM reduced Glutathione (Amersham Pharmacia) solution (Glutathione was dissolved in 1X PBS, pH was adjusted to 7.4) in 1ml fractions. Fusion proteins were aliquoted and frozen in liquid nitrogen, stored at -80ºC.

For MBP (Maltose binding protein) fusion proteins, 1L of LB-ampicillin culture containing bacterial cells were induced by 300 µM IPTG at 25ºC for 5-6 hrs. Cells were harvested and resuspended in 30 ml of column buffer (20 mM of Tris-HCl, pH 7.4; 200 mM of NaCl and 1mM EDTA) with 1 mM DTT and EDTA-free protease inhibitor tablet (1/50ml). The cells were lysed and supernantant was collected after centrifugation. The supernatant was incubated with 6 ml 50% slurry of Amylose resin

(NEB) to bind MBP-fusion protein to the beads and unbound fractions were removed by washing with chilled column buffer (10 times of the bed volume of the beads). The MBP fusion protein was eluted with 10mM maltose (Sigma) solution in 2ml fractions. Fusion proteins were frozen and stored as above.

Protein concentration was measured using Bio-Rad protein assay solution (2.5.2). Purity and quantity of fusion proteins were checked by SDS-PAGE, stained with SimplyBlue Safestain (Invitrogen).

2.5.2 Determination of protein concentration

Bio-Rad protein assay solution (5X) (Bio-Rad 500-0006) was used to determine protein concentrations. Bio-Rad protein assay solution was diluted to 1X with water. Highly concentrated proteins were diluted in a range of 1-3mg/ml with their respective buffers. 1, 5 and 10 µl of proteins were each added to 1ml of 1X Bio-Rad protein assay solution. For blank and control, water and sample buffer were each added to 1ml of 1X Bio-Rad protein assay solution. Reactions were incubated for 10 min at room temperature, transferred to plastic cuvettes and absorption was measured at 595 nm using a spectrophotometer (Pharmacia Ultraspec 2100). Alternatively, protein concentration was estimated by comparing band intensity with a dilution series of BSA (Bovine serum albumin) on a SDS-PAGE gel, stained with SimplyBlue Safestain (Invitrogen).

2.5.3 SDS-PAGE

Protein electrophoresis was performed using sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) gels and the Bio-Rad Mini-Protean II gel apparatus. Separating gels contained 330 mM Tris-HCl pH 8.8, 0.1% SDS, 6-15% polyacrylamide mixture (Protogel- 30%: 0.8% w/v acrylamide:bisacrylamide, from National Diagnostics). Gels were polymerized by addition of 0.1% ammonium persulphate

(APS) (Sigma) and 0.1% TEMED (Sigma). 70% ethanol or butanol was added top of the separating gel until gels were polymerized. After polymerization, ethanol was removed and gels were washed with water before pouring the stacking gels. Stacking gels layered above the separating gel contained 125 mM Tris-HCl pH 6.8, 0.1% SDS, 4% polyacrylamide mixture and water (upto final volume) polymerized with 0.1% APS and 0.1% TEMED. Bacterial extracts, recombinant proteins and worm extracts were mixed in 2x SDS sample buffer (100 mM Tris-HCl pH 6.8, 20% glycerol, 4% SDS, 20% β-Mercaptoethanol, 0.2% w/v bromophenol blue), denatured by incubation at 95ºC for 5 min immediately before loading. Electrophoresis was performed at 100-120 Volt in 1X running buffer (25mM Tris-base, 190 mM glycine, 0.1% SDS) until protein standard marker (prestained marker. MBI fermentas) migrated at the appropriate position on the separating gel.

2.5.4 Protein gel staining

Protein gels were stained with either commassie blue staining solution (0.25% w/v commassie brilliant blue R250, 10% methanol and 10% acetic acid) or using SimplyBlue Safestain (Invitrogen) for 1-2 hr at RT. The staining solution was decanted and replaced with either destaining solution (10% methanol, 10% acetic acid) (in case of commassie blue staining solution) or with water (in case of SimplyBlue Safestain).

2.5.5 Antibody generation

PRG-1 GST-(1-60aa) and GST-PAZ recombinant fusion proteins were dialysed in 1X PBS buffer to remove reduced glutathione, and injected individually into two rabbits to generate anti-PRG1 polyclonal antibodies (PickCell Laboratories, Amesterdam, The Netherlands).

Three different polyclonal anti-peptide antibodies were also generated by immunizing either rabbits or guineapigs. PRG-1 peptide sequences are shown in Table 2.3. First peptide antibody: Rabbit polyclonal PRG-1 antibody was generated injecting following

peptide YTRLDYSISPLSE (Abcam 15826). Second peptide antibody: QSKTGSSGQPQKC and RGRGSGSNNSGGKDQ peptides were co-injected into two rabbits for immunization (Eurogentec, Belgium) and the final serum was collected after 28 days. Third peptide antibody: SGRGRGRGSGSNNSGGKDQKYL & RQQGQSKTGSSGQPQKC peptides were co-injected into two rabbits and two guineapigs for immunization (Peptide Speciality Laboratories, Germany) and the final serum was collected after 28 days.

2.5.6 Western blot analysis

Proteins were resolved by SDS-PAGE and transferred onto nitrocellulose membrane (Hybond ECL membrane, Amersham) in transfer buffer (25 mM Tris-base, 190 mM glycine, 20% methanol) for 1 hr and 30 min at 300mA using Bio-Rad Mini-Trans-Blot cell apparatus at 4ºC. Protein transfer and loading were confirmed using Ponceau S solution (Sigma). Ponceau staining was removed by washing with TBS-T (150 mM NaCl, 20 mM Tris-HCl pH 7.6 with 0.5% Tween-20) buffer. Then the membrane was blocked in TBS-T buffer supplemented with 5% non-fat dried milk for 1 hr at RT or overnight at 4ºC with gentle shaking. The membrane was incubated for 2 hr at RT or overnight at 4ºC in primary antibody diluted in TBS-T supplemented with 5% non-fat dried milk with gentle shaking. Membrane was washed 3 times with TBS-T for 10 min each and then incubated with either anti-rabbit (Amersham) or anti-mouse (Amersham) horse radish peroxidase (HRP) conjugated secondary antibody diluted 1:10000 in TBS-T supplemented with 5% non-fat dried milk for 1 hr at RT. The membrane was washed 3 times with TBS-T for 10 min each. Protein bands were detected by using either Lumigen detection reagent (Amersham) or immobilon western kit (Millipore) according to the manufacturer's instructions.

Immunoblotted membranes were incubated in stripping buffer (2% SDS, 62.5 mM Tris-HCl pH 6.8, 100 mM β-Mercaptoethanol) for 30 min at 50 ºC. Stripped membranes were washed 4 times for 10 min at RT with TBS-T (gentle shaking) and then blocked in TBS-T supplemented with 5% milk as fresh membranes.

2.5.7 Preparation of *C. elegans* extract

Synchronized animals were grown in liquid culture at 15ºC and collected as young adults by floating on a sucrose gradient, followed by extensive washing with M9. Droplets of worm pellet were frozen in liquid nitrogen and then stored at -80ºC. Frozen droplets were crushed to powder using a chilled Biopulverizer with Spring-loaded hammer (capacity of 1-10 gm) (BioSpec Products Inc. Bartlesville, USA), followed by further grinding with a Cryo-cup grinder (BioSpec Products Inc. Bartlesville, USA) to make a smooth powder. The powder was transferred to a falcon tube and dissolved in lysis buffer (1X PBS with complete protease inhibitor tablets (Roche)), centrifuged at 14000 rpm for 15-20 min at 4ºC and the supernatant was collected.

Worm extract was also prepared by dissolving small balls in suitable lysis buffer and animals were disrupted using a One Shot cell disrupter (Constant Systems, Northants, UK) at 200 MPa pressure. The lysate was clarified by centrifugation at 14000 rpm for 20 min at 4ºC and supernatant was collected.

Small volume of worm pellet was directly dissolved in suitable lysis buffer, disrupterd using Bio rupture on ice for 10-15 min with 30 sec interval. Then, lysate was cleared by centrifugation at 14000 rpm for 20 min at 4ºC and supernatant was collected. In all cases clarified lysates were aliquoted, frozen in liquid nitrogen and subsequently stored at -80ºC.

2.5.8 Immunoprecipitation

Synchronized animals were grown in liquid culture at 15ºC and collected as young adults by floating on a sucrose gradient, followed by extensive washing with M9. Animals were resuspended in two volumes of lysis buffer (100 mM PIPES pH6, 100 mM NaCl, 3 mM MgCl$_2$, 1 mM EGTA, 1 mM DTT, 1 mM PMSF, 0.3M sucrose, 0.5% Triton X-100, complete protease inhibitor tablets (Roche), phosphatase inhibitor

cocktail 1 and 2 (Sigma)) and frozen in liquid nitrogen. Animals were disrupted using a One Shot cell disruptor at 200MPa (Constant Systems, Northants, UK). The lysate was clarified by centrifugation at 14,000rpm at 4°C for 20 min. For each immunoprecipitation 20 μl of protein A agarose beads (GE Healthcare) were coupled with 2 μl of serum (PPD2 or preimmune) by rotating for 2 hrs at 4°C in 500μl of 1XPBS, 0.5% Triton X-100. Beads were washed twice with 500 μl of IP buffer (30 mM HEPES, 1% NP40, 100 mM NaCl, 66 mM KCl, 1 mM MgCl$_2$, 1 mM DTT, EDTA free complete protease inhibitor tablet (Roche), 100U of RNaseOUT (Invitrogen)). 100μg of *C. elegans* lysate was diluted in IP buffer to 500μl and rotated with antibody-coupled beads for 2 hrs at 4°C. Beads were washed three times with 1ml of IP buffer, resuspended in 300μl of 1X proteinase K buffer (100mM Tris-HCl pH 7.4, 12mM EDTA, 150mM NaCl, 1% SDS) with 1 μg/μl of proteinase K, rotated for 30 min at room temperature and RNA was extracted using Trizol reagent. 10% of immunoprecipitated RNA was used for RT-PCR.

2.5.9 Purification of a PRG-1 complex through gel-filtration chromatography

A total worm extract was made from worm pellets using a Biopulverizer with Spring-loaded hammer (capacity of 1-10g) and Cryo-cup grinder as described above (Section 2.3.7). The cleared total worm extract was fractionated through SuperdexHR 200 (Amersham) and Superose 6 (Amersham) columns using an ÄKTA chromatography instrument (Pharmacia) according to the manufacturer's instructions. Fractions were analyzed by western blotting using an anti-PRG-1 antibody.

Table 2.1 Deletion alleles described in this study

allele	gene	mutangen	Lg[a]	cosmid	size deletion breakpoints
n4357	*prg-1*	EMS	I	D2030	GTTTTCTTTCCTTGGAGAGGT//GATGCTCATATTGTAATCT
n4503	*prg-1*	EMS	I	D2030	TTCAAGGCTTCTGAAAATTTAT//ACTTCTGCGACAGCAGC
n4358	*prg-2*	EMS	IV	C01G5	CGGTTCGTTTTCTTGAATCG//CCTTTAAGTTTTCATCTCAA
nDf57[b]	*prg-2*	EMS	IV	C01G5	ATCGGGATGAAGTTTGCAAA//AATCTAGAATACCGATTTCG

Strains with these deletion alleles have been submitted to the *Cenorhabditis* Genetics Center. [a] linkage group and [b] Deletion is upstream of the open reading frame.

Table 2.2 Strains used in this study

strain	genotype	comment
MT14034	*prg-1(n4357)*	4x outcrossed
MT14521	*prg-1(n4503)*	4x outcrossed
FX872	*prg-1(tm872)*	
MT14035	*prg-2(n4358)*	4x outcrossed
MT14531	*prg-2(nDf57)*	4x outcrossed
FX1094	*prg-2(tm1094)*	
MT14769	*prg-1(n4357); prg-2(n4358)*	*"piwi(n4357; n4358)"*[1]
SX9	*prg-1(n4503); prg-2(nDf57)*	*"piwi(n4503; nDf57)"*[1]
SX25	*avr-14(ad1302); dcr-1(ok247) / qC1*; avr-15(ad1051) glc-1(pk54)*	
WM27	*rde-1(ne219)*	
FX892	*csr-1(tm892) / unc-24(e138am)*	
WM158	*ergo-1(tm1860)*	
WM126	*sago-2(tm894) ppw-1(tm914); C06A1.4(tm887) F58G1.1(tm1019); M03D4.6(tm1144) sago-1(tm1195)*	
WM160	*sago-1(tm1195)*	
WM154	*sago-2(tm894)*	
NL2557	*ppw-1(pk1425)*	
NL5117	*ppw-2(pk1673)*	
NL2098	*rrf-1(pk1417)*	
EL476	*rrf-2(ok210)*	
MT12630	*rrf-3(pk1426)*	
SX496	*ego-1(om84) I / hT2 [bli-4(e937) let-?(q782) qIs48] (I;III)*	
TW332	*rde-3(r459)*	
GR1373	*eri-1(mg366)*	
NL917	*mut-7(pk204)*	
JK551	*unc-5(e53) fem-3(q22)*	ts allele[2]
BA17	*fem-1(hc17ts)*	ts allele[2]
CB4078	*glp-1(e2144)*	ts allele[2]
GC833	*glp-1(ar202)*	ts allele[2]
SS104	*glp-4(bn2ts)*	ts allele[2]
SX157	*prg-1(n4357); unc-22(st136)*	
SX166	*prg-1(n4357); prg-2(n4358); unc-22(st136)*	
SX171	*unc-32(e189); mut-7(pk204); unc-22(st136)*	
SX158	*prg-1(n4357); unc-22(r750)*	
SX165	*unc-32(e189); mut-7(pk204); unc-22(r750)*	
SX336	*prg-1(n4357); prg-2(n4358) unc-22(r750)*	
SX178	*prg-1(n4357); unc-22(r765)*	
SX179	*unc-32(e189); mut-7(pk204); unc-22(r765)*	
SX278	*prg-1(n4357); prg-2(n4358); unc-22(r765)*	
NL3643	*unc-22(st136)*	Tc1 insertion
TR1160	*unc-22(r750)*	Tc3 insertion
TR1175	*unc-22(r765)*	Tc4 insertion
SX339	*prg-1(n4357); prg-2(n4358) ; ccEx7271*	
SX340	*unc-32(e189) mut-7(pk204); ccEx7271*	
PD7271	*pha-1(e2123ts); ccEx7271*	

[1] Abbreviated genotypes.

[2] Temperature-sensitive allele.

Table 2.3 Antibodies used in this study

anti-PRG-1 antibodies (using peptides)

	Name	Species	Peptide synthesis & antibody generation
Peptides			
YTRLDYSISPLSE	Ab15826	Rabbit	Abcam (UK)
QSKTGSSGQPQKC	SPY877	Rabbit	Eurogentec (Belgium)
RGRGSGSNNSGGKDQ	SPY878	Rabbit	
SGRGRGRGSGSNNSGGKDQKYL	1932-1[x]	Guineapig	PSL (Germany)
RQQGQSKTGSSGQPQKC	1932-2[x]	Guineapig	PSL (Germany)
SGRGRGRGSGSNNSGGKDQKYL	1929-1[y]	Rabbit	PSL (Germany)
RQQGQSKTGSSGQPQKC	1929-2[y]	Rabbit	PSL (Germany)

anti-PRG-1 antibodies (using recombinant proteins)

	Name	Species	
Recombinant proteins			
GST-(1-60aa)	Rb5999	Rabbit	PickCell Laboratories (Netherlands)
	Rb6000	Rabbit	PickCell Laboratories (Netherlands)
GST-PAZ	Rb5802	Rabbit	PickCell Laboratories (Netherlands)
	Rb5803	Rabbit	PickCell Laboratories (Netherlands)
anti-tubulin antibody*	tubulin	Mouse	Dako (USA)

Six different antibodies were generated against PRG-1. PRG-1 specific peptides and PRG-1 recombinant proteins were injected to immunize rabbits or guineapigs. [x] antibodies highly specific to endogenous PRG-1. [Y] antibodies are specific to PRG-1 but still reognize some unspecific bands. * Anti-tubulin antibody was commercially available and purchased from Dako. Peptides and recombinant proteins were injected into two different rabbits or guineapigs for immunization apart from anti-PRG-1 antibody generated with Abcam.

Table 2.4 Oligonucleotides used for RT-PCR and northern blotting

Target	Sequence (5′->3′)	Type
21UR-1	CTCAACTGGTGTCGTGGAGTCGGCAATTCAGTTGAGGCACGGTT	RV
21UR-1	ACACTCCAGCTGGGTGGTACGTACGTTAAC	FW
21UR-1353	CTCAACTGGTGTCGTGGAGTCGGCAATTCAGTTGAGTACCATCA	RV
21UR-1353	ACACTCCAGCTGGGTCAGAACATCACATGA	FW
21UR-2171	CTCAACTGGTGTCGTGGAGTCGGCAATTCAGTTGAGGACTATCA	RV
21UR-2171	ACACTCCAGCTGGGTAGTCACGTAAATTGA	FW
21UR-3272	CTCAACTGGTGTCGTGGAGTCGGCAATTCAGTTGAGACTTGATG	RV
21UR-3272	ACACTCCAGCTGGGTGTTGAAGCTCGGCAT	FW
21UR-3442	CTCAACTGGTGTCGTGGAGTCGGCAATTCAGTTGAGCACAATCT	RV
21UR-3442	ACACTCCAGCTGGGTACTAGAGTGTTGAGA	FW
21UR-5045	CTCAACTGGTGTCGTGGAGTCGGCAATTCAGTTGAGGCCGTTAA	RV
21UR-5045	ACACTCCAGCTGGGTGCCAAACTCCATTTA	FW
21UR-5101	CTCAACTGGTGTCGTGGAGTCGGCAATTCAGTTGAGCCAAGTTC	RV
21UR-5101	ACACTCCAGCTGGGTAGGTACTACTTTGAA	FW
21UR-5166	CTCAACTGGTGTCGTGGAGTCGGCAATTCAGTTGAGGGCGAAAT	RV
21UR-5166	ACACTCCAGCTGGGTAACATAGCATTGATT	FW
miR-52	CTCAACTGGTGTCGTGGAGTCGGCAATTCAGTTGAGAGCACGGA	RV
miR-52	ACACTCCAGCTGGGCACCCGTACATATGTT	FW
(universal)	CTCAAGTGTCGTGGAGTCGGCAA	RV
Tc3	AATAGTCGCGGGTTGAGTTG	RV
Tc3	GAGCGTTCACGGAGAAGAAG	FW
Tc1	CACATGACGACGTTGAAACC	RV
Tc1	AACCGTTAAGCATGGAGGTG	FW
PRG-1	TGTTCCCACTCCTCCTTCAG	RV
PRG-1	ATGTTTCGCCAACTTCATCC	FW
PRG-2 (1)	CGTATCCGACGATCATTGTG	RV
PRG-2 (1)	CAATCAATGCGTGAATTTGC	FW
PRG-2 (2)	ACTCTGCTTGCTTTCCTTCG	RV
PRG-2 (2)	TCCACAGCAAAGACTTGACG	FW
ACTIN-1	GTACGTCCGGAAGCGTAGAG	RV
ACTIN-1	CACGGTATCGTCACCAACTG	FW
21UR-1	GCACGGTTAACGTACGTACCA	northern
siR-23-69	GACATCTCAGAAACATCGGCAAA	northern
siR-26-263	TAGCATATGCATGCACCATAAACAAC	northern
U6	ATCTTCTCTGTATTGTTCC	northern
let-7	AACTATACAACCTACTACCTCA	northern
miR-52	AGCACGGAAACATATGTACGGGTG	northern

Table 2.5 Primers used for RT-PCR of Class I and Class II genes

	Target	Sequence (5′->3′)	Type
Class I			
	Y105C5B.21	AACCCGAAGAACATCGTGAC	FW
		GCTGCCAGGTTTTGAATAGC	RV
	Y57G11C.24	AGTATCCGATGCAGGTGTCC	FW
		GCGGAGCAACTCTTCCATAG	RV
	H08M01.2	TGTCAGAGATGCTGCCAAAC	FW
		ATTGGGGAGCAATTGTTGAG	RV
	C27H2.2	AAGCTCGCTCGTCAAGTCTC	FW
		AGATTCGCCAATTCTGCATC	RV
	Y57G11C.32	ACCACTGGAATCCGATGAAG	FW
		ACGATTCGATTTGACCTTGC	RV
	F26D12.1	TCACAAGTGCTTCCAACGAG	FW
		TGTCTGTGGCTGACTTCTGG	RV
Class II	F22B3.4	TGGTCCGGAAATAGGAGTTG	FW
		TGCACGTCTCTTCATCTTCG	RV
	Y57G11C.16	CAAGGACGGAAAGACTGGAC	FW
		TCTTGGTGTGCTGTCCTCTG	RV
	Y51H4A.3	AGGATCAGTTCCCAGACGTG	FW
		GTGGACGCAGACGATCATAG	RV
	T06A10.4	ATGATCGAATGCGAGAATCC	FW
		CATTGGTGCAATACCAGTCG	RV
	ZK550.3	GCAAACTCGGTGGCTCTTAC	FW
		CAGCCAGAACTTTTGCTTCC	RV
	F52B11.2	ATCAATCCAGAAAGCCATCG	FW
		AGGCGACAAATTGATCATCC	RV

Table 2.6 Primers used for cloning of different fragments of PRG-1

Target	Sequence (5′->3′)	Type
PRG-1-(1-60 aa)	CCGGAATTCATGGCATCTGGAAGTGGTCG	FW
	CCGCTCGAGTTATTCAATCGGGATGAAGTTGG	RV
PRG-1-(1-241 aa)	CCGGAATTCATGGCATCTGGAAGTGGTCG	FW
	CCGCTCGAGTTATTTCTCTTGAACGCGTTGTGG	RV
PRG-1-PAZ (218-354 aa)	CGCGGATCCCCGAGTCGATGTACATTCTCTTTC	FW
	CCGCTCGAGTTACGGAGACATGCGAGTGTGC	RV
PRG-1-PIWI (499-810 aa)	CCGGAATTCATGCTCGTCGTAATGCTCG	FW
	CCGCTCGAGTTAGTGAAGAGATTGGGCTGTC	RV
PRG-1-(738-824 aa)	CCGGAATTCATGGACTTCTACTTGGTTCCA	FW
	CCGCTCGAGTTACAAGAAGAACAGCTTGTC	RV
PRG-1-(1-60 aa)	CGCGGATCCATGGCATCTGGAAGTGGTCG	FW
	CCCAAGCTTTTATTCAATCGGGATGAAGTTGG	RV
PRG-1-PAZ (218-354 aa)	CGCGGATCCATGTACATTCTCTTTCATAAGG	FW
	CCCAAGCTTTTACGGAGACATGCGAGTGTGC	RV
PRG-1-ORF	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGCATCTGGAAGTGGTC	FW
	GGGGACCACTTTGTACAAGAAAGCTGGGTCTTACAAGAAGAACAGCTTGTCACGAA	RV
PRG-1-(1-936 bp)	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGCATCTGGAAGTGGTCG	FW
	GGGGACCACTTTGTACAAGAAAGCTGGGTCTTATTGCTTTGGCTTTCCTTCAG	RV
PRG-1-(1-633 bp)	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGCATCTGGAAGTGGTCG	FW
	GGGGACCACTTTGTACAAGAAAGCTGGGTCTTAACGATTTTCAACGCAAAGC	RV
PRG-1-PAZ (652-1062 bp)	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGAGTCGATGTACATTCTCTTTC	FW
	GGGGACCACTTTGTACAAGAAAGCTGGGTCTTACGGAGACATGCGAGTGTGC	RV
PRG-1-(1027-1575 bp)	GGGGACAAGTTTGTACAAAAAAGCAGGCTTC ATGAAGGAGATCGCAAAGC	FW
	GGGGACCACTTTGTACAAGAAAGCTGGGTCTTAAATGGGGCACTCGACACATAGG	RV
PRG-1-PIWI (1495-2430 bp)	GGGGACAAGTTTGTACAAAAAAGCAGGCTTC ATGCTCGTCGTAATGCTCG	FW
	GGGGACCACTTTGTACAAGAAAGCTGGGTCTTA GTGAAGAGATTGGGCTGTC	RV
PRG-1-(1495-2007 bp)	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGCTCGTCGTAATGCTCG	FW
	GGGGACCACTTTGTACAAGAAAGCTGGGTCTTAGACTTCGGTGTTCTTGATATACG	RV
PRG-1-(1834-2475 bp)	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGCAAACAAGACCTCATGAGAATCC	FW
	GGGGACCACTTTGTACAAGAAAGCTGGGTCTTACAAGAAGAACAGCTTGTCACG	RV

Table 2.7 Primers used for cloning of the different fragments of Tc1, Tc3 and U6 for RNase Protection (RPA) assay

Probes name	Sequence (5′->3′)	Type	comment	RPA
Tc3-464	TGCTTCAAAAAGGACCATCC	FW	pGEMTeasy	yes
	AATAGTCGCGGGTTGAGTTG	RV		
Tc3-N ter	TAATACGACTCACTATAGGGAGA-CTCGAGGATCTGCCCTTTC	T7-FW	PCR product	yes
	ATTTAGGTGACACTATAGAAGGG-CGAATCACATTTCGTTCGTC	SP6-RV		
	CTCGAGGATCTGCCCTTTC	FW	pGEMTeasy	
	CGAATCACATTTCGTTCGTC	RV		
Tc3-C ter	TAATACGACTCACTATAGGGAGA-CGACAATCCATGTGAGCAAC	T7-FW	PCR product	yes
	ATTTAGGTGACACTATAGAAGGG-GCTGGTTGTCCGGAATAGAC	SP6-RV		
	CGACAATCCATGTGAGCAAC	FW	pGEMTeasy	
	GCTGGTTGTCCGGAATAGAC	RV		
Tc3-TIR	TAATACGACTCACTATAGGGAGA-CCCCCTAATTTGAAGGTTTG	T7-FW	PCR product	
	ATTTAGGTGACACTATAGAAGGG-GTTTTGGCATTACCGAGAGC	SP6-RV		
	CCCCCTAATTTGAAGGTTTG	FW	pGEMTeasy	yes
	GTTTTGGCATTACCGAGAGC	RV		
Tc1-370	TAATACGACTCACTATAGGGAGA-CCTGTTGGCTCTAGGTACTCTCC	T7-FW	PCR product	
	ATTTAGGTGACACTATAGAAGGG-GAAGCCCGAATACCTCCAAG	SP6-RV		
	CCTGTTGGCTCTAGGTACTCTCC	FW	pGEMTeasy	yes
	GAAGCCCGAATACCTCCAAG	RV		
Tc1-TIR	TAATACGACTCACTATAGGGAGA-GCTGGCCAAAAAGATATCCAC	T7-FW	PCR product	
	ATTTAGGTGACACTATAGAAGGG-CAAATTTGCGTTTCGTTTGC	SP6-RV		
	GCTGGCCAAAAAGATATCCAC	FW	pGEMTeasy	yes
	CAAATTTGCGTTTCGTTTGC	RV		
U6	TAATACGACTCACTATAGGGAGA-TCCGAGAACATATACTAAAATTGGAAC	T7-FW	PCR product	yes
	ATTTAGGTGACACTATAGAAGGG-AAAATTTGGAACGCTTCACG	SP6-RV		

Primers were used to clone different parts of the Tc1 and Tc3 either from genomic DNA or cDNA library (prepared from total RNA). T7-FW primers and SP6-RV primers are with T7 and SP6 promoter sequence respectively at their 5′ flanking end followed by a sequence that corresponding to their targets for amplification. PCR products carrying T7 and SP6 promoter sequences were used directly for *in vitro* transcription, or otherwise PCR product was cloned into pGEMTeasy plasmid (which contains T7 and SP6 promoter sequences flanking the MCS), linearized and then used for *in vitro* transcription. *In vitro* transcribed probes were used for RPA assay. 'yes' indicating probes, which were used for RPA assay.

3. Results

3.1 Introduction

Argonaute proteins associated with distinct classes of small RNAs form the core of the RNA induced silencing complex (RISC) (Carmell et al., 2002; Tolia and Joshua-Tor, 2007). The superfamily of Argonatue proteins can be divided into the Ago sub-family and the Piwi sub-family. The Ago sub-family is found in fission yeast, plants and animals, whereas the Piwi sub-family found in ciliates, slime moulds and animals.

Expression of the Piwi sub-family of Argonaute proteins is restricted mostly to germ cells in animals (Carmell et al., 2007; Deng and Lin, 2002; Houwing et al., 2007; Kuramochi-Miyagawa et al., 2004; Kuramochi-Miyagawa et al., 2001). In flatworms, Piwis are expressed in germ cells and somatic cells (Reddien et al., 2005). In *Drosophila*, Piwis (Piwi, Aubergine and AGO3) are expressed both in male and female germ lines. Additionally, Piwi is expressed in somatic cells (Cox et al., 1998; Cox et al., 2000; Lin and Spradling, 1997). *Drosophila* Piwi is required for germline stem cell maintenance. In *Tetrahymena*, a Piwi protein called Twi1 is required for DNA elimination in the macronucleus (Mochizuki et al., 2002). In Zebrafish, Ziwi is expressed both in male and female gonads and Ziwi is also required for germ cells maintenance (Houwing et al., 2007). In mice, expression of three Piwi proteins, Miwi (Piwil1), Mili (Piwil2) and Miwi2 (Piwil4) - is mainly restricted to testes (Carmell et al., 2007; Deng and Lin, 2002; Kuramochi-Miyagawa et al., 2004; Kuramochi-Miyagawa et al., 2001). However, expression of Mili in prenatal ovaries has also been reported (Kuramochi-Miyagawa et al., 2001). Loss of function alleles of the mouse Piwi genes, *miwi* and *mili*, affect meiotic progression of developing sperm but do not affect the maintenance of mitotic germ cells (Deng and Lin, 2002; Kuramochi-Miyagawa et al., 2004). Additionally, Piwi proteins are also required for retrotransposon silencing in *Drosophila* and mice (Aravin et al., 2001; Aravin et al., 2007b; Savitsky et al., 2006).

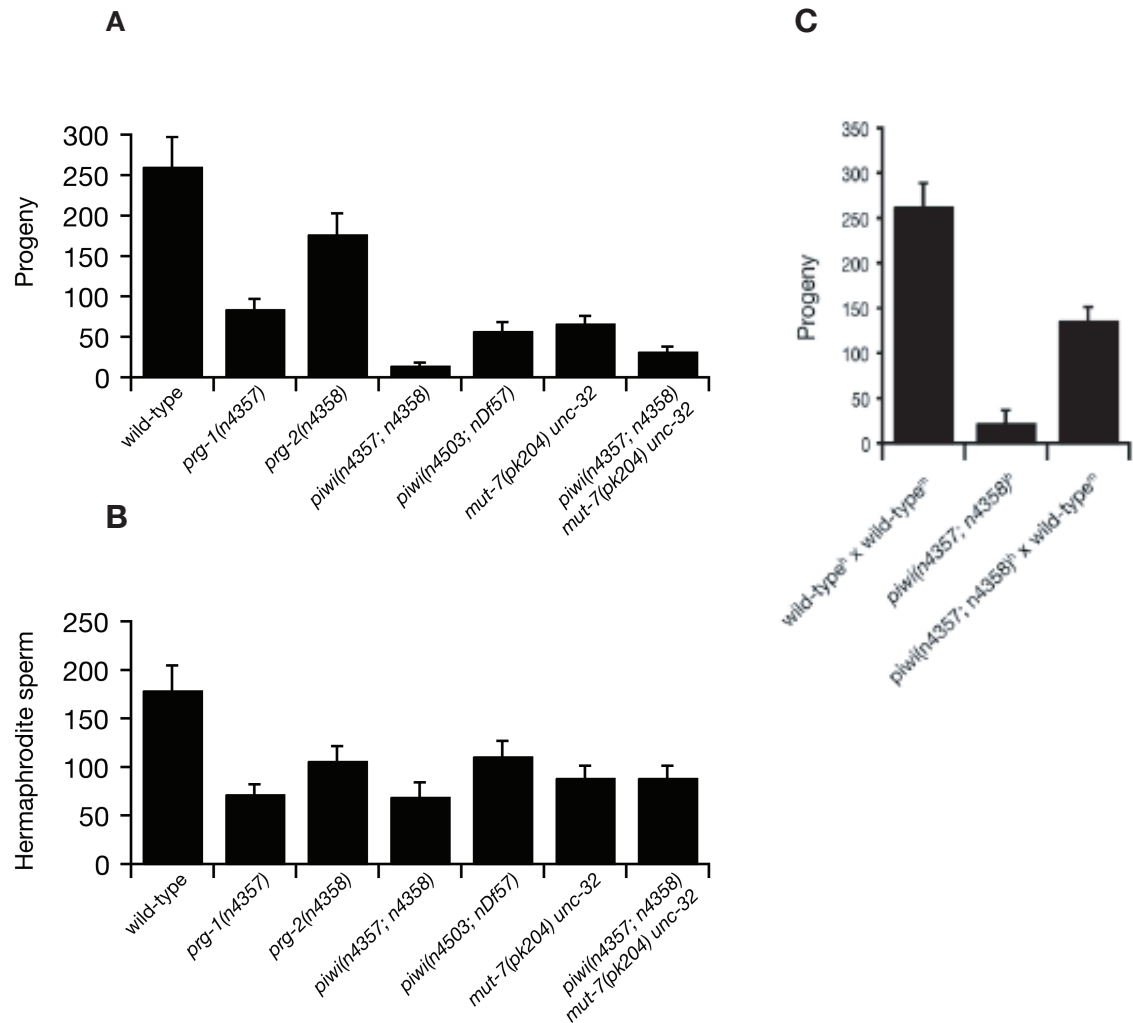## 3.2 Piwi is required for normal germline development

The Piwi proteins in *C. elegans*, PRG-1 and PRG-2 are 90% similar at their amino acid level. Due to their similarities at the amino acid level, it is thought that *prg-1* and *prg-2* might have been generated by a recent gene duplication and act redundantly. To investigate the roles of Piwi proteins in *C. elegans*, in collaboration with Eric Miska, I generated two deletion alleles of both *prg-1 (n4357, n4503)* and *prg-2 (n4358, nDf57)* (Figure 3.1A and 3.1B) by (ethane methyl sulfonate) EMS mutagenesis. This is a process by which EMS, a chemical mutagen, is used to introduce point mutations and/or insertions and/or deletions into the genome randomly. Then we mapped the deletions for *prg-1* and *prg-2,* followed by outcrossing to wild-type animals eight times to remove background mutations. I also used other alleles of *prg-1 (tm872)* and *prg-2 (tm1094)* (Figure 3.1C and 3.1D) obtained from the *Caenorhabditis* Genetics Center (CGC). In collaboration with Marloes Bagijn, I generated doubly mutant strains after outcrossing of the single mutants of *prg-1* and *prg-2*, which I will refer to as *prg-1; prg-2* double mutants or *piwi* mutants, *i.e. piwi (n4357; n4358)* and *piwi (n4503; nDf57)*. Single mutants of *prg-1 (n4357, n4503)* and *prg-2 (n4358, nDf57)* are viable. Progeny numbers were counted at 20ºC and the average progeny number was approximately 250 in the case of wild-type animals. However, progeny numbers were dramatically reduced in *prg-1* mutants to 20% - 25% of that of wild-type animals at the same temperature (Figure 3.2A). The fertility defects were enhanced in *piwi (n4357; n4358* and *n4503; nDf57)* mutants to 12% -16% of wild-type animals (Figure 3.2A). These observations agreed with previous observations where RNAi of *prg-1* showed a reduced number of progeny to 50% of wild-type animals (Cox et al., 1998). Previous studies of RNAi of *prg-1* also showed a dramatic shortening of the mitotic proliferative zone (MPZ) and a 50% reduction in the number of mitotic nuclei as compared to control animals (Cox et al., 1998). Therefore, hermaphrodite sperm was counted in wild-type and *prg-1* and *prg-2* mutant animals. Sperm counts were reduced to approximately 50% in both single and *piwi* mutants (Figure 3.2B). However, fertility of *piwi* mutants was not completely rescued by introducing wild-type sperm through mating (Figure 3.2C). This finding suggests that Piwi function is not restricted to

spermatogenesis but is also required during oogenesis and early germline development, including mitotic germ stem cell proliferation and maintenance. Altogether, these data confirm a conserved role of PRG-1 and PRG-2 in germline development in *C. elegans* like other Argonautes proteins in Piwi sub-family.
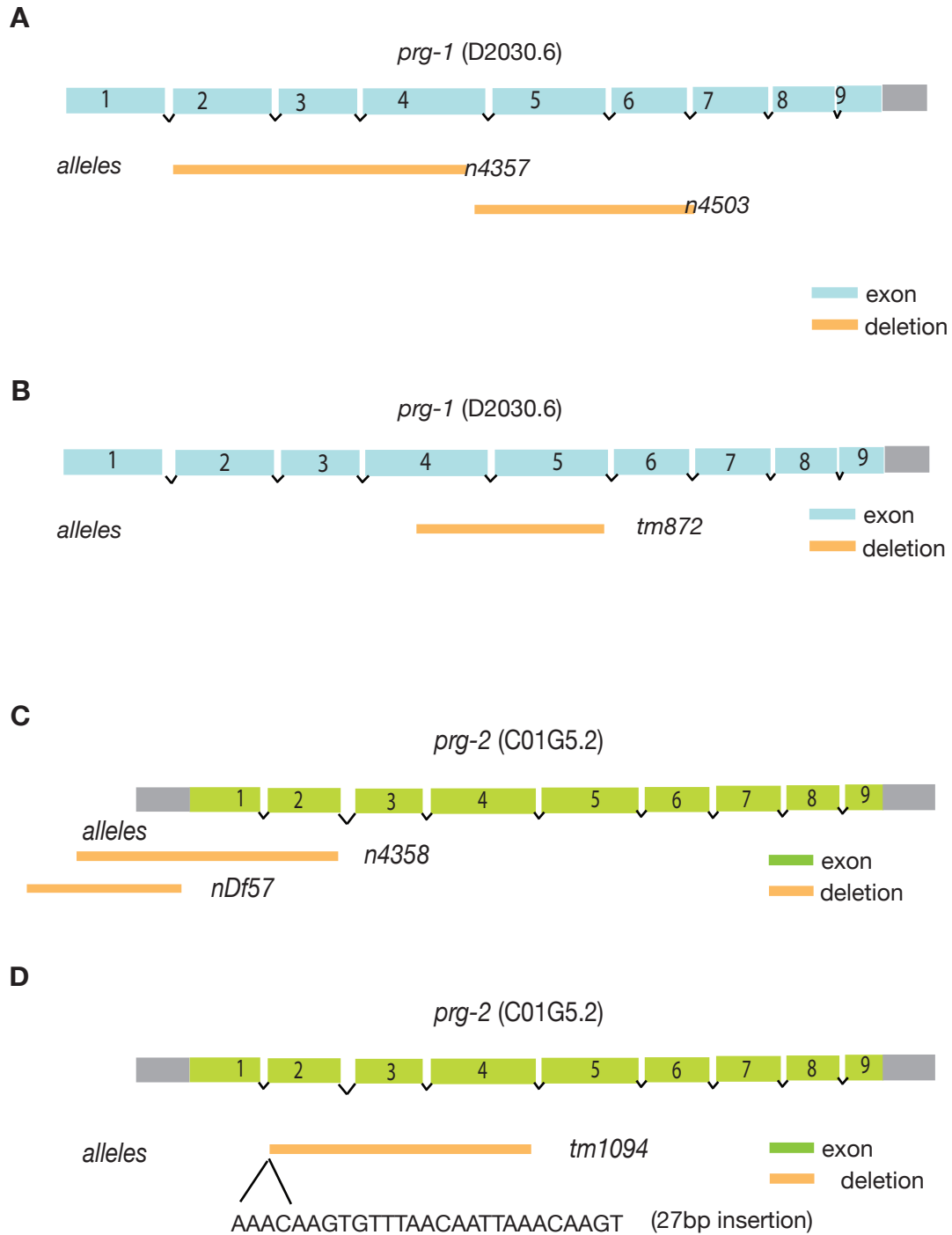
Figure 3.1 Schematic representation of *prg-1* and *prg-2* alleles

(A, B, C, D) Blue/Green boxes and black lines are representing exons and introns respectively. Orange boxes represent deletions, gray boxes represent untranslated regions. In D, 27bp sequence is an insertion between the deletion breakpoints.
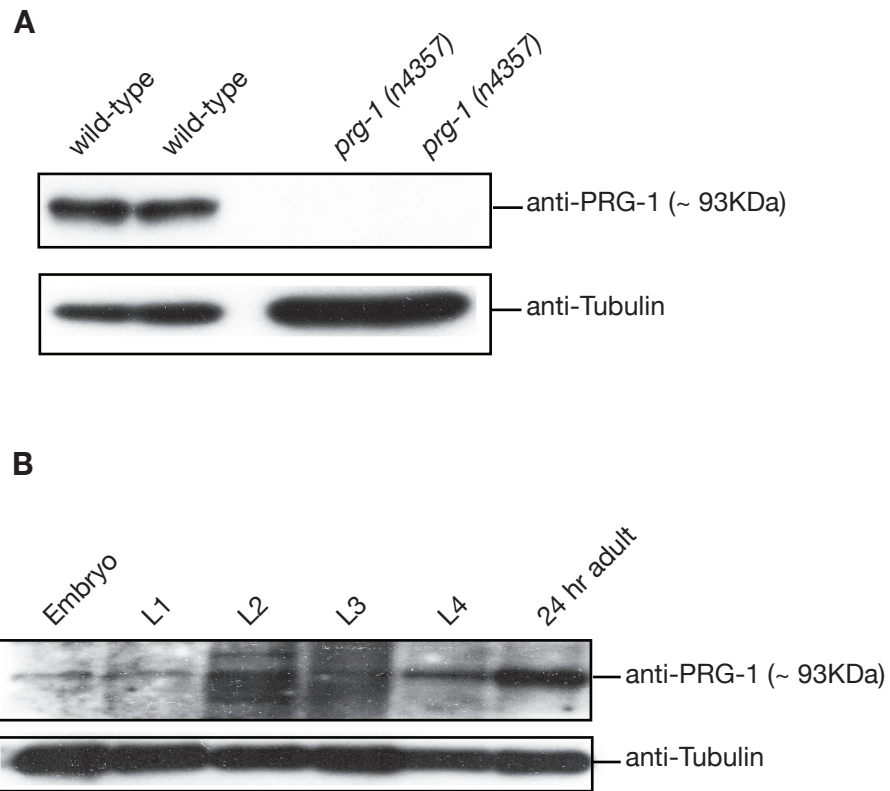
A

B

C

Figure 3.2 Piwi is required for normal germline development in *C. elegans*
(A) Total progeny were counted at 20°C by transferring adults to new plates every 24 hours. Progeny of at least 50 animals were counted for each strain. Error bars correspond to the standard error of the mean. (B) 12-hour adult animals were stained with DAPI and sperm were counted in both gonad arms. At least 40 animals were counted for each genotype shown. Error bars correspond to the standard error of the mean. (C) Wild-type and *piwi* mutants hermaphrodites were mated with sperms from wild-type males carrying a GFP transgene to distinguish cross- progeny. *piwi* mutants males very rarely mated successfully with *piwi* mutant hermaphrodites, therefore *piwi mutants* self-progeny are shown for comparison. Data shown are of atleast 20 different single-hermaphrodite matings. [h] hermaphrodites and [m] males. This figure was created in collaboration with Heeran Buhecha and Eric Miska.
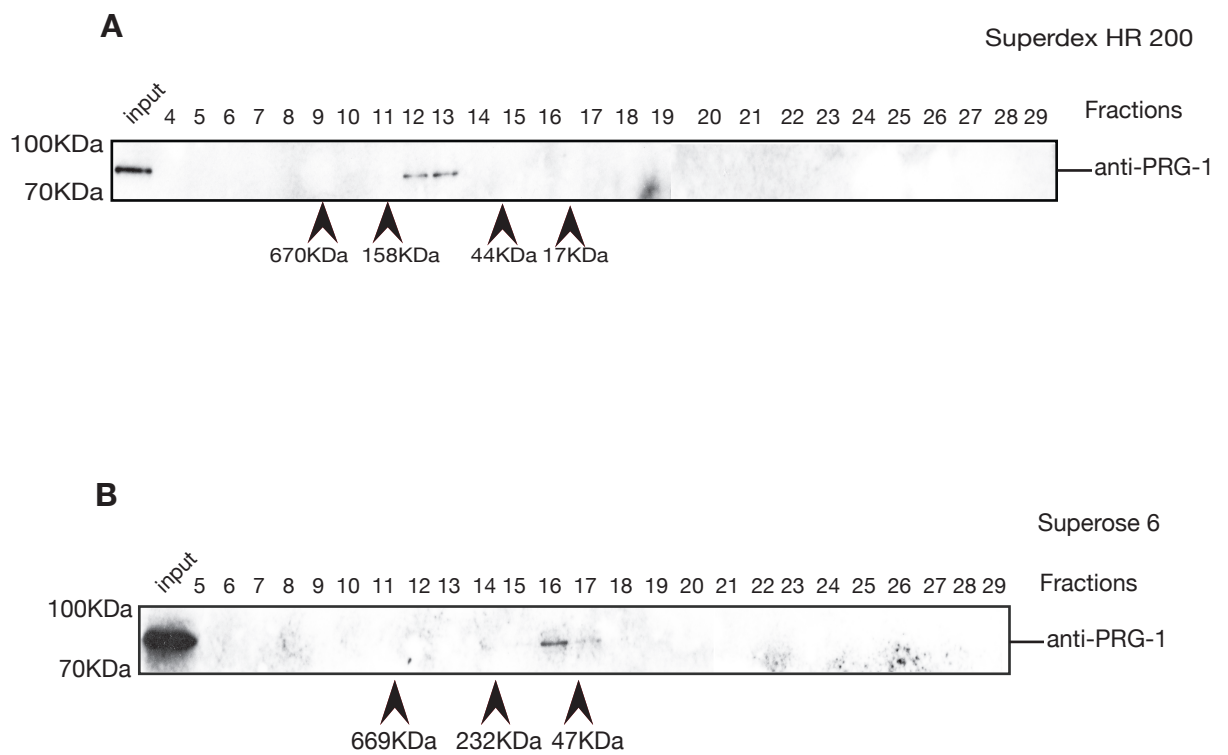
3.3 PRG-1 is not part of a stable protein complex

In order to invesitigate potential PRG-1 containing complexes, I generated a number of PRG-1 polyclonal antibodies in rabbits and guineapigs (Table 2.3). Specificity of α-PRG-1 antibodies was tested using wild-type extract and *prg-1 (n4357)* mutant extracts. I was able to detect endogenous PRG-1 protein at the expected size (~93KDa) in wild-type extract and not in *prg-1* mutant extract using anti-PRG-1 antibody (1932-1) (Table 2.3) (Figure 3.3A). Tubulin was used as an internal control and endogenous tubulin expression was detected by anti-tubulin antibody in both wild-type and *prg-1* mutant extract, suggesting that α-PRG-1 antibody is specific to endogenous PRG-1. To test the temporal expression of endogenous PRG-1 protein levels, I prepared total extract from different stages of wild-type animals and performed a western blot using α-PRG-1 antibody (1932-1). The result showed that PRG-1 expression is highest in young adult stages (24 hr) (Figure 3.3B). Therefore, for further experiments, I prepared extract from young adult animals (either 12 hr or 24-48 hr).

To investigate whether PRG-1 forms any endogenous protein complexes, wild-type total extract was passed through a Superdex 200HR gel-filtration column (optimal separation range 600kDa to 10kDa; exclusion limit 1.3MDa). Fractions were collected and a western blot was performed for PRG-1 (Figure 3.4A). PRG-1 eluted with an apparent molecular weight of 100kDa, close to its estimated molecular weight (93.8kDa), suggesting that PRG-1 might be present as a monomer. To verify if PRG-1 is contained in a higher molecular weight complex, I subjected wild-type total extract to a Superose 6 gel filtration column (optimal separation range 5000kDa to 5kDa; exclusion limit- 40MDa). Fractions were collected and a western blot of PRG-1 performed (Figure 3.4B), also indicating that PRG-1 eluted as a monomer. Taken together, these data suggest that PRG-1 might not be part of larger protein complex.

A

wild-type   wild-type   *prg-1 (n4357)*   *prg-1 (n4357)*

—anti-PRG-1 (~ 93KDa)

—anti-Tubulin

B

Embryo   L1   L2   L3   L4   24 hr adult

—anti-PRG-1 (~ 93KDa)

—anti-Tubulin

Figure 3.3 Specificity of anti-PRG-1 antibody (1932-1) and expression of endogenous PRG-1
(A) Total protein lysate was prepared from wild-type and *prg-1* mutants. Endogenous expression of PRG-1 was examined using anti-PRG-1 antibody (1932-1) by western blot analysis. Anti-Tubulin was used as loading control.
(B) Total protein lysate was prepared from different stages of wild-type animals. An equal amount of protein lyaste was used for western blot analysis. Anti-PRG-1 antibody was used to detect endogenous PRG-1 protein. Anti-Tubulin was used as loading control.

**A**   Superdex HR 200

input 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19   20 21 22 23 24 25 26 27 28 29 Fractions

100KDa

70KDa   anti-PRG-1

670KDa 158KDa   44KDa 17KDa

B   Superose 6

input 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 Fractions

100KDa

70KDa   anti-PRG-1

669KDa   232KDa 47KDa

Figure 3.4 Monomeric form of PRG-1

8mg of total extract from wild-type animals was prepared and subjected to Superdex 200 HR (A) and Superose 6 (B) gel filtration columns. Fractions were collected and endogenous PRG-1 was analyzed by western blot using an anti-PRG-1 antibody (1932-1). Fraction numbers are indicated on the top and size standards are on the bottom.

3.4 The 21U-RNAs are the piRNAs of *C. elegans*

In *Drosophila*, zebrafish, mice and rats, Piwi proteins are associated with 24-30 nt of non-coding small RNAs, namely piRNAs (Piwi-interacting RNAs) (Aravin et al., 2006; Brennecke et al., 2007; Girard et al., 2006; Gunawardane et al., 2007; Houwing et al., 2007; Lau et al., 2006; Saito et al., 2006; Vagin et al., 2006; Watanabe et al., 2006). To search for the piRNAs in *C. elegans* I prepared total RNA (40μg) from wild-type animals and *piwi* mutants, resolved the RNA on a denaturing gel and stained with SYBR green. I was not able to detect an abundant class of small RNA in the range of 26-31nt, the size range of piRNA (Figure 3.5A). The same amount of mouse testis total RNA was used as positive control, where I was able to detect an abundant 29-30nt small RNA population. Additionally, I performed 3′ end labeling of total RNA to increase the sensitivity of the assay. Again, I was not able to detect any small RNAs in the apropriate size range in RNA from *C. elegans* (Figure 3.5B). In contrast, I could detect populations of small RNAs ranging from 21-23 nt in *C. elegans*, which include miRNAs and endogenous siRNAs.

Failing to identify an abundant class of small RNAs at the correct size range I next searched for piRNAs among the small RNAs previously identified in *C. elegans*: miRNAs, endogenous siRNAs and 21U-RNAs (Ambros et al., 2003; Lee and Ambros, 2001; Ruby et al., 2006). As a first criterium for piRNAs I tested if any of the known classes of small RNAs was dependent on Piwi. Surprisingly, I found that the 21U-RNA 21UR-1 is absent in two different independent *piwi* mutants (*n4357; n4358* and *n4503; nDf57*) by northern blotting (Figure 3.6).
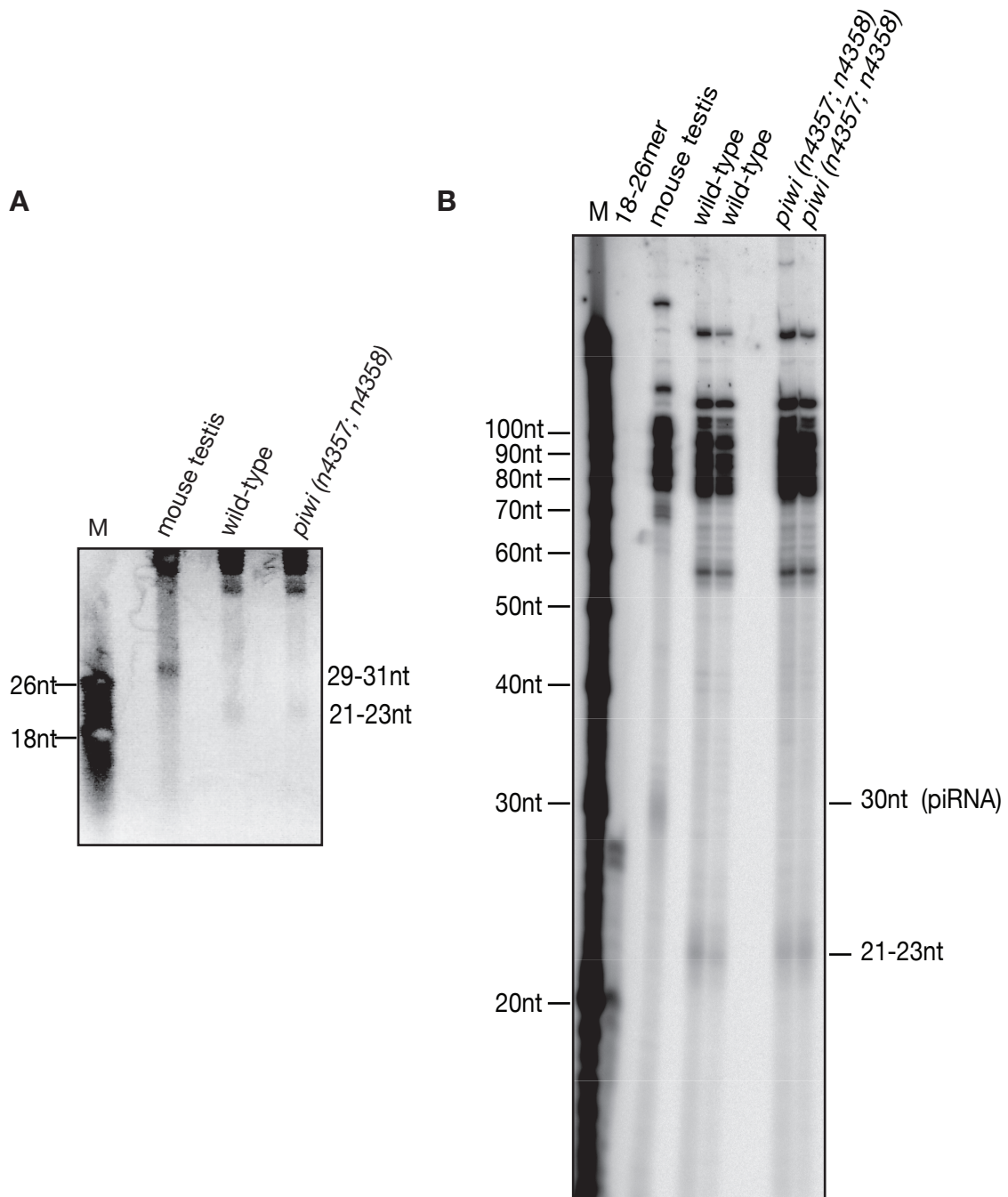
**Figure 3.5 There are no abundant 25-32 nt small RNAs in *C.elegans***
(A) 40 µg of total RNA from wild-type and *piwi (n4357; n4358)* mutant animals was resolved on 15% denaturing PAGE and stained with SYBR green. Mouse testes total RNA was used as positive control. (B) 100 ng of total RNA was labeled with pCp (5'-p32 5'-3' cytidine bis-phosphate) by 3' end labeling, purified through G-25 spin column and resolved through 18% denaturing PAGE and exposed to phophoimager screen.

The same blot was stripped and hybridized with a probe for the miRNA miR-52, a ubiquitously expressed miRNA in *C. elegans*. I found that the expression of miR-52 was unaffected (Figure 3.6). Expression levels of endogenous siRNAs were also checked by northern blotting. The expression levels of two additional endogenous siRNAs (siR23 and siR26) were unaffected in *piwi* mutants (Figure 3.7A & B). To check whether 21U-RNAs were generally absent in *piwi* mutants, I generated 5′ monophosphate small RNA libraries from wild-type animals and *piwi (n4357;n4358)* mutants. High-throughput sequencing of these libraries was performed using Solexa sequencing technology (Seo et al., 2004). The libraries were run three times individually. 2,719,949 reads (wild-type) and 764,960 reads (*piwi* mutants) aligned to the *C. elegans* genome (version WS170 - downloaded from UCSC genome browser) with at least one perfect match. Computaional analysis in collaboration with Leonard Goldstein, I found 1398 out of 5454 previously known 21U-RNAs (Ruby et al., 2006) and approximately 8000 novel 21U-RNAs. The global expression of 21U-RNAs was dramatically reduced or absent in *piwi* mutants compared to wild-type animals (Figure 3.8). Highest number of reads obtained was for 21UR-548 (3921 reads) in the wild-type library. Frequencies of the 400 most abundant 21U-RNAs in wild-type compared to *piwi* mutants library were plotted. The most abundant 21U RNA in a *piwi* mutant was 21UR-3224, which had 8 reads, whereas 21UR-3224 had 2127 reads in wild-type (Figure 3.8). I also used the Solexa sequencing data to check the expression levels of miRNAs, endogenous siRNAs including 26nt siRNAs (SiR26), in wild-type and *piwi* mutants and found no difference in their expression levels, as I had observed previously (Figure 3.6, 3.7A & B). Altogether these data suggested that 21U-RNAs might be the piRNAs of *C. elegans*.
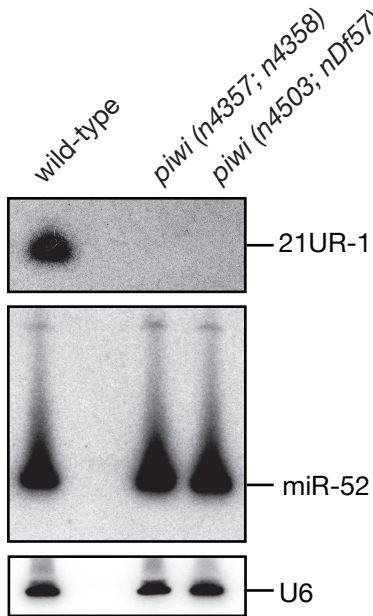
Figure 3.6 21U-RNA-1 is absent in *piwi* mutants
Northern blot showing that 21UR-1 is not detected in RNA (40 µg) isolated from two indepedent
piwi double-mutants: *piwi (n4357; n4358)* and *piwi (n4503; nDf57)*, whereas miR-52 is expressed
in both mutants (same blot re-probed). Antisense DNA probes were used for 21UR-1 and miR-52.
piwi(n4357; n4358) is an abbreviation for *prg-1(n4357); prg-2(n4358). piwi(n4503; nDf57)* is an
abbreviation for *prg-1(n4503); prg-2(nDf57)*.
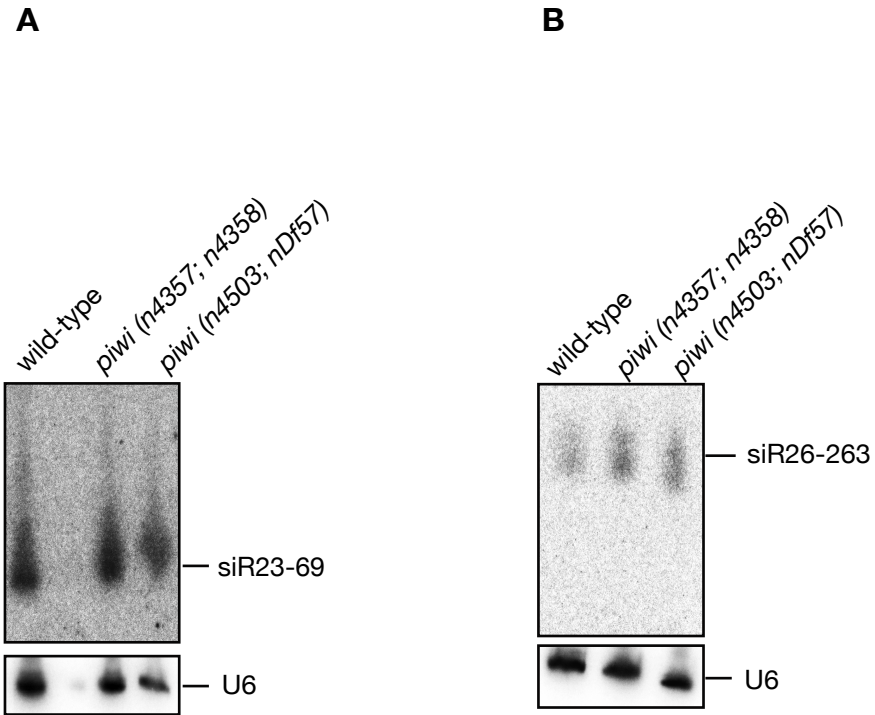
Figure 3.7 Expression of endogenous siRNAs were not affected in *piwi* mutants
(A & B) 40 µg of total RNA was resolved through 15% denaturing PAGE. Northern blot was performed by using specific probes to detect endogenous siRNAs (SiR23-69 and SiR26-263). Expression of a 23 nucleotide antisense RNA (siR23-69) and a 26 nucleotide antisense RNA (siR26-263) was not affected in *piwi* mutants. U6 was used as an internal control.
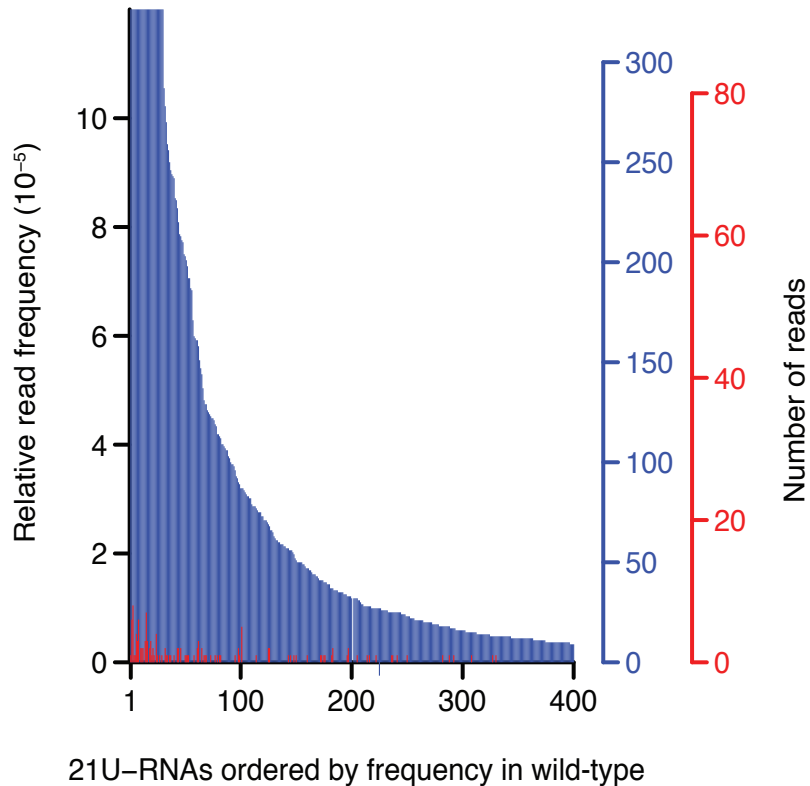
Figure 3.8 Expression of 21U RNAs was dramatically reduced in *piwi* mutants (high throughput sequencing)

High-throughput sequencing reveals that the expression of many 21U-RNAs is dramatically reduced in *piwi* mutants. 21U-RNAs detected by high-throughput sequencing of 5'-dependent wild-type and *piwi(n4357; n4358)* mutant libraries. Frequencies are shown for wild-type (blue) and *piwi* mutant (red) for the 400 most abundant 21U-RNAs in wild-type, plotted in the order of their wild-type frequency. Read frequencies were obtained by dividing the number of reads for a given 21U-RNA by the total number of reads from the same library (left-hand y-axis). The corresponding absolute number of reads are indicated in the right-hand y-axes. 21U-RNAs for which frequencies are shown include the 21U-RNA with most reads in the *piwi* mutant library (21UR-3224), which was sequenced 8 times in the *piwi* mutant and 2,127 times in wild-type. This figure was created in collaboration with Leonard Goldstein.

To further test this hypothesis, I used an anti-PRG-1 serum (Rb5999) (Table 2.3) to immunoprecipitate endogenous PRG-1 from wild-type total extract of young adult stage animals (12 hr adult), followed by RNA extraction and RT-PCR of specific 21U-RNAs, 21UR-5101 and 21UR-5045. I was able to detect specifically 21UR-5101 and 21UR-5045 in the immunoprecipitate from wild-type extract. However, expression of 21UR-5101 and 21UR-5045 was not detected from *piwi* mutant total extract or when using pre-immune serum (Figure 3.9).

High-throughput sequencing data suggested that levels of 21U-RNAs were reduced in *piwi* mutants, but not completely absent. Therefore, expression levels of seven individual 21U-RNAs were quantified by quantitative RT-PCR. The expression levels of most 21U-RNAs were undetectable in *piwi* mutants (Figure 3.10). However, expression of a few 21U-RNAs, such as 21UR-1, was still detectable at a very reduced level in *piwi* mutants (Figure 3.10). For 21UR-1, this was verified by cloning and sequencing of the RT-PCR product.

Studies from *Drosophila*, zebrafish and mice have shown that piRNAs are modified at the 3′ end with a 2-OMe group (Horwich et al., 2007; Houwing et al., 2007; Kirino and Mourelatos, 2007; Ohara et al., 2007; Saito et al., 2007; Vagin et al., 2006). To test whether 21U-RNAs are modified at the 3′ end, total RNA was treated with sodium periodate ($NaIO_4$) for β-elimination, followed by northern blotting with probes for miR-52 and 21UR-1. For miRNAs the 3′ end of the terminal ribose sugar carries a 2-hydroxyl group. Upon periodate treatment and β-elimination miRNAs (miR-52) lose the terminal ribose sugar and base from the 3′ end and therefore, miRNAs migrate faster in denaturing polyacrylamide gels than $NaIO_4$-untreated RNA (Figure 3.11). However, in the case of piRNAs the 2′-hydroxyl group is modified to a 2′-OMe group at the 3′ end. Therefore, piRNAs are insensitive to periodate treatment and the subsequent β-elimination reaction and migrate with the same speed as $NaIO_4$ untreated RNAs. A northern blot of 21UR-1 showed that migration of 21UR-1 was unaffected by $NaIO_4$ treatment and the β-elimination reaction (Figure 3.11), indicating 21U-RNAs are

modified at their 3′ terminal nucleotide. Taken together these data suggest that 21U-RNAs are the piRNAs of *C. elegans*. Henceforth, I will refer to them as piRNAs.
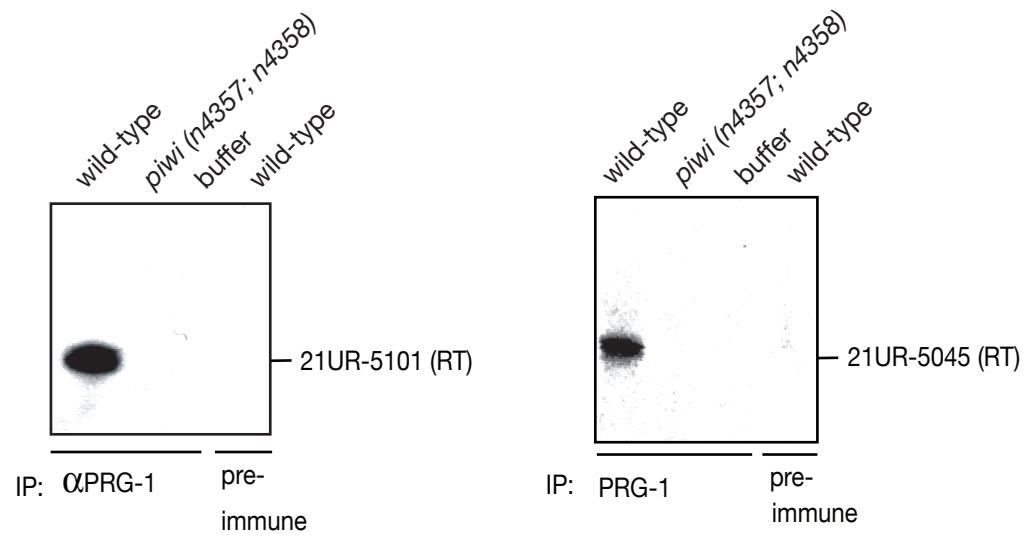
**Figure 3.9 21U-RNAs are associated with PRG-1**
Immunoprecipitation of endogenous PRG-1 using anti-PRG-1 serum (Rb5999) followed by RT-PCR of 21UR-5101 and 21UR-5045 reveals that 21U-RNAs are associated with PRG-1 in *C. elegans* extracts. Pre-immune serum and *piwi* mutant extracts were used as control.
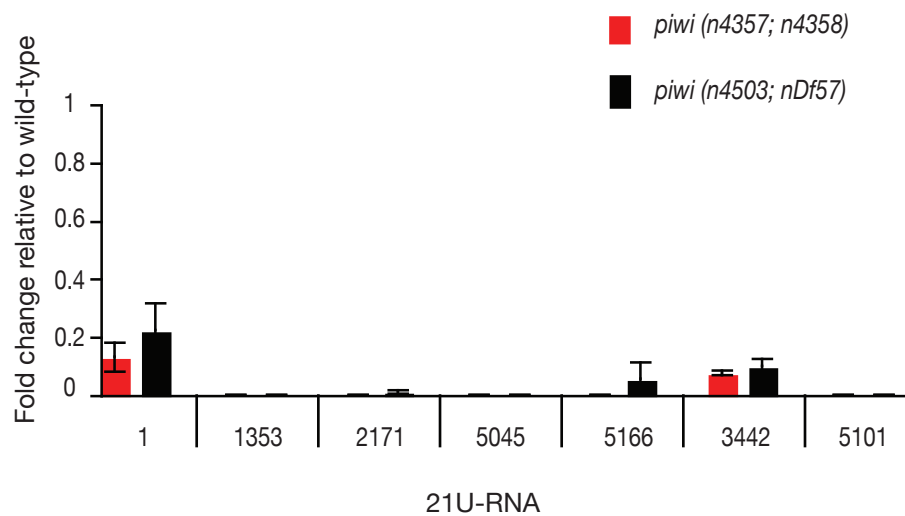
Figure 3.10 Piwi might not be essential for 21U-RNA biogenesis or stability
Quantitative RT-PCR of seven 21U-RNAs demonstrates that Piwi is not essential for 21U-RNA biogenesis or stability. Total RNA was extracted from 12 hour adult *C. elegans*. Expression levels shown are relative to levels in wild-type RNA. miR-52 expression was used as an internal control. Expression of 21U RNAs are normalized by miR-52 expression. After normalization, expression of 21U RNAs are set to 1 for wild-type and fold change is calculated in *piwi* mutants compared to wild-type. Data are from three independent biological replicates. Error bars represent standard error of the mean.
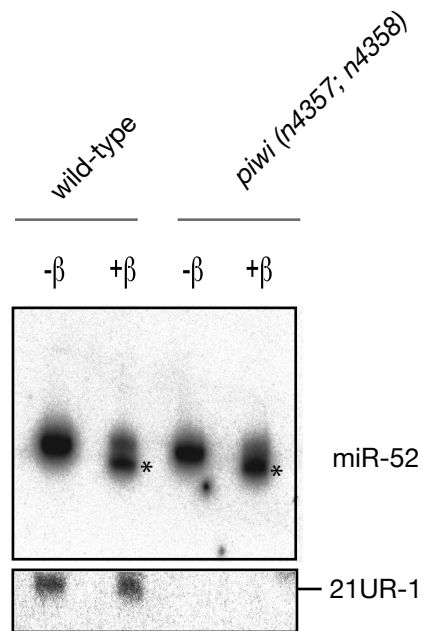
Figure 3.11 21U-RNAs are modified at their 3' end
20 µg of total RNA was treated with sodium periodate and subjected to beta-elimination reaction. RNA was precipitated and 10 µg of RNA was used for northern blotting. First, miR-52 was detected and then same blot was stripped and probed with anti-21UR-1. (*) indicates RNAs unmodified at their 3' terminal ribose sugar.

3.5 piRNA biogenesis is independent of other small RNA pathways

To identify if additional genes are involved in the piRNA pathway in *C. elegans*, I tested several genes involved in oher small RNA pathways in *C. elegans* by using mutants or by RNAi. The expression of piRNAs was checked by northern blotting or quantitative RT-PCR (Table 3.1). First, I checked the expression of piRNAs in three independent mutants of *prg-1 (n4357, n4503* and *tm872)* and *prg-2 (n4358, nDf57* and *tm1094)*. I found that only PRG-1 is required for piRNA biogenesis or stability and not PRG-2 (Figure 3.12 and Table 3.1). Next, I checked whether other Argonaute proteins are involved in piRNA biogenesis or stability. Therefore, I tested RDE-1, which is required for exogenous RNAi (Tabara et al., 1999), ALG-1 and ALG-2, which are required for miRNA biogenesis (Grishok et al., 2001), ERGO-1 required for endogenous RNAi (Yigit et al., 2006), and CSR-1, an essential Argonaute required for chromosome segregation and germline RNAi (Yigit et al., 2006). None of the above Argonaute proteins were found to be involved in piRNA expression (Figure 3.12 and Table 3.1). Then, I tested secondary Argonaute proteins, including SAGO-1, SAGO-2, PPW-1, and PPW-2, which are involved in secondary siRNA production (Yigit et al., 2006). None of the secondary Argonautes was required for piRNA expression (Table 3.1). In addition, I tested all four RNA-dependent RNA polymerases (RdRPs) of *C. elegans*, RRF-1, RRF-2, RRF-3 and EGO-1, which are required for secondary siRNA generation (Sijen et al., 2001; Smardon et al., 2000). None of these was required for piRNA expression (Figure 3.12 and Table 3.1). I also tested several Dicer-related helicases DRH-1, DRH-2 and DRH-3. DRH-1 and DRH-2 are closely related helicases and act redundantly. DRH-1 physically interacts with RDE-4, which interacts with DCR-1 (Duchaine et al., 2006). None of these helicases was involved in piRNA expression (Table 3.1). Next, I checked several nucleases that give an enhanced RNAi (Eri) response when mutated, ERI-1, ERI-3, and ERI-5. It was shown previously that ERI-3 and ERI-5 promote an interaction between ERI-1 and DCR-1 (Duchaine et al., 2006). None of these nucleases were required for piRNA expression (Table 3.1). Drosha (DRSH-1) is an important component of the miRNA biogenesis pathway. Drosha is required for the processing of primary miRNA (pri-miRNA) transcripts into

precursor miRNA (pre-miRNA) hairpins (Denli et al., 2004). However, I found that DRSH-1 is also not involved in piRNA biogenesis or stability (Table 3.1). I checked whether the RNase III ribonuclease enzyme DCR-1 (Dicer) is required for piRNA expression. DCR-1 is essential for siRNA production and germline development (Grishok et al., 2001; Ketting et al., 2001). *dcr-1* mutants are sterile, therefore these experiments were performed using homozygous mutants animals from heterozygous mothers or using RNAi. I found that DCR-1 is not required for piRNA expression (Figure 3.13). Expression of the *let-7* miRNA, was checked in *dcr-1* mutants. An accumulation of precursor miRNA and a reduction of mature miRNA were observed in *dcr-1* mutants compared to N2 and *piwi* mutants as expected (Figure 3.13). However, mutation of *dcr-1* was not completely able to abolish the processing of pre-miRNA to mature miRNA due to maternal contribution of DCR-1. These data were also confirmed by high-throughput sequencing of *dcr-1* mutants (Figure 3.14).
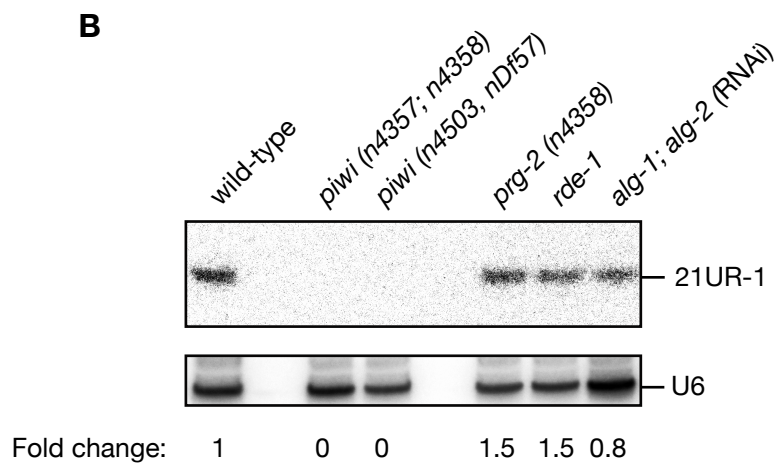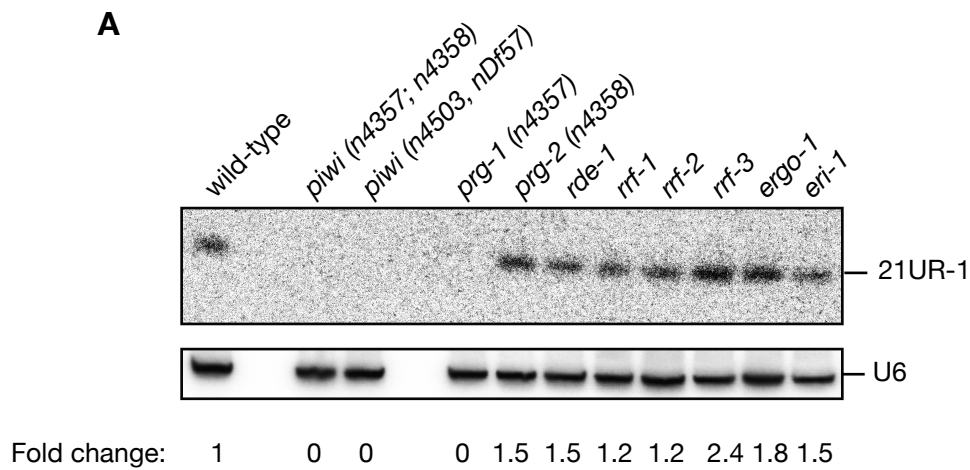
Figure 3.12 piRNA biogenesis is independent of many genes in other small RNA pathways
21UR-1 northern blotting of total RNA of wild-type and mutant young adult of *C. elegans*. For *alg-1; alg-2*(RNAi), *alg-1* mutant L1 larvae were transferred to *alg-2* RNAi feeding plates and young adult animals were harvested. A U6 northern blot is shown as loading control. Fold change is calculated after normalization with U6 expression. Expression of 21UR-1 after normalization is set to 1 for wild-type and relative fold change is calculated compared to wild-type. '0' represents no detectable expression of 21UR-1.

Table 3.1 Many proteins involved in small RNA pathways are not required for piRNA biogenesis or stability

| Genotype	Functional class	Fold change 21U-RNA	
		northern	qRT PCR
wild-type		1	1
prg-1(n4357)	Piwi	0	0
prg-1(n4503)	Piwi		0
prg-1(tm872)	Piwi		0
prg-1(RNAi)	Piwi	0.3	
prg-2(n4358)	Piwi	1.5	1.2
prg-2(nDf57)	Piwi		1.2
prg-2(tm1094)	Piwi		1.3
piwi(n4357; n4358)[a]	Piwi	0	0
piwi(n4503; nDf57)[a]	Piwi	0	0
dcr-1[b]	Type III RNase	0.9	0.8
drsh-1[b]	Type III RNase		0.8
drh-1(RNAi)	helicase	1.3	
drh-2(RNAi)	helicase	1.3	
drh-3(RNAi)	helicase	1.6	
rde-1	Argonaute	1.5	
rde-2	novel		0.9
rde-3(RNAi)	nucleotidyltransferase	2.6	
rde-4	dsRBP[c]		2.4
alg-1; alg-2(RNAi)	Argonaute	0.8	
csr-1	Argonaute		1
ergo-1	Argonaute	1.8	
sago-1	Argonaute	0.7	
sago-2	Argonaute	1	
mago	Argonaute		1.2
ppw-1	Argonaute		1.2
ppw-2	Argonaute		0.7
rrf-1	RdRP[c]	1.2	
rrf-2	RdRP[c]	1.2	
rrf-3	RdRP[c]	2.4	
ego-1[b]	RdRP[c]		1
eri-1	nuclease	1.5	
eri-3	nuclease		1
eri-5	nuclease		1.8
mut-2 (RNAi)	helicase	2.1	
mut-7	helicase	1.9	1

Relative 21U-RNA expression is shown as fold change with respect to wild type.
Empty field: not tested.
[a] Abbreviated genotypes.
[b] Homozygous animals derived from heterozygous mother.
[c] double-stranded RNA binding protein
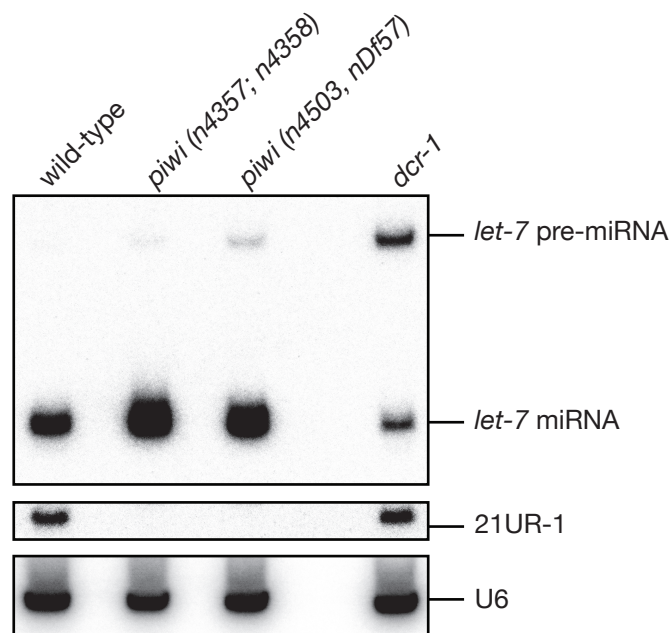
Figure 3.13 piRNA biogenes is Dicer independent
21UR-1 northern blotting of total RNA of wild-type and mutant young adult aqnimals of *C. elegans*. *dcr-1* mutant animals used were homozygous animals derived from heterozygous mothers. To test for loss of DCR-1 activity in *dcr-1* mutants, mature *let-7* miRNA and pre-miRNA is shown. A U6 northern blot is shown as loading control.

As most of the 21U-RNAs originated from chromosome IV (Ruby et al., 2006), the distribution of 21U-RNAs was mapped onto chromosome IV from the high-throughput sequencing analysis of 5′ dependent libraries of wild-type, *piwi* and *dcr-1* mutants (Figure 3.14). The result showed that dramatic reduction of 21U-RNAs in *piwi* mutants, as I mentioned previously (Figure 3.8). However, the frequency of 21U-RNAs was not affected in *dcr-1* mutant animals, supporting the conclusion that piRNA expression is independent of DCR-1 (Figure 3.14). Finally, I tested whether heterochomatin protein-1 (HP-1) was required for piRNA expression. In *Drosophila*, Piwi interacts directly with HP-1a, which is a key player of heterochomatin mediated gene silencing (Brower-Toland et al., 2007; Pal-Bhadra et al., 2004). In *C. elegans*, HPL-1 and HPL-2 encode two HP1-like proteins, HPL-1 and HPL-2. Northern blot analysis showed that there was a significant reduction of 21UR-1 expression in *hpl-1* mutants as compared to wild-type animals and *hpl-2* mutants (Figure 3.15). Therefore, these data suggest that piRNAs are independent of many other proteins involved in other small RNA pathways in *C. elegans,* including DCR-1. An initial result suggests that HPL-1 promotes piRNA biogenesis or stability.
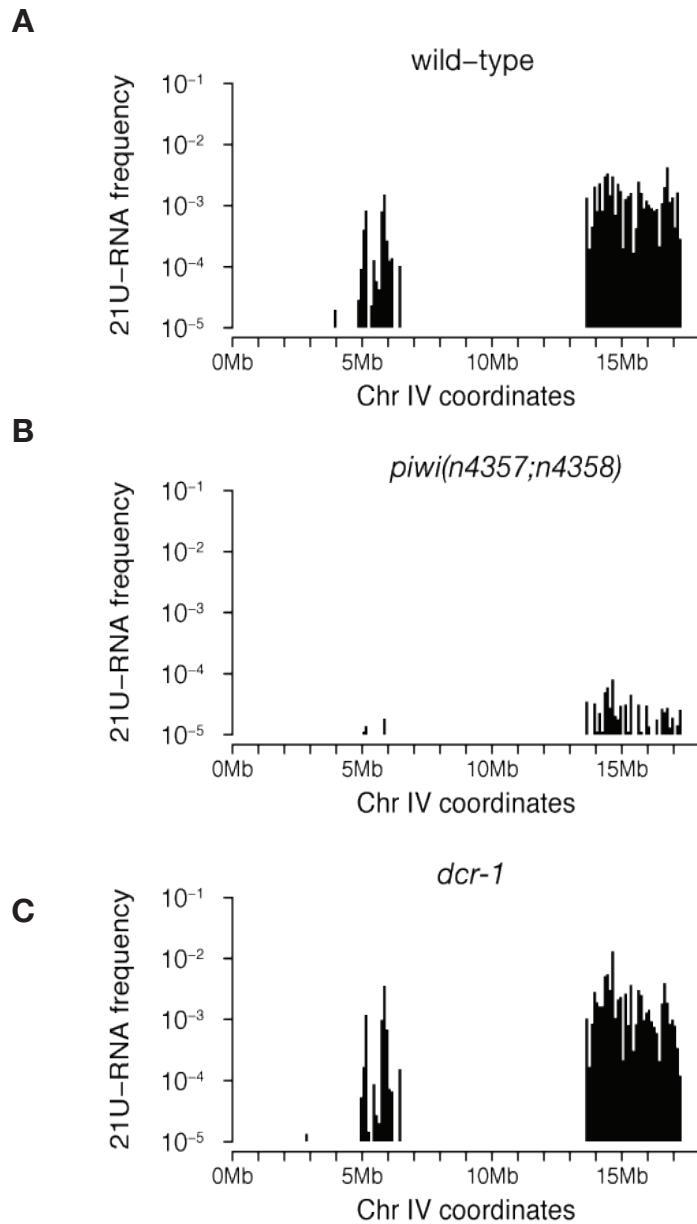
Figure 3.14 Expression of 21U-RNAs is Dicer independent
Distribution of 21U-RNAs on chromosome IV as detected by high-throughput sequencing of 5'
dependent wild-type, *piwi (n4357; n4358)* and *dcr-1* mutant libraries. The frequency for a given
21U-RNA and locus was obtained by correcting the number of reads for multiple alignments and
dividing by the total number of reads from the same library. Cumulative frequencies were plotted
for non-overlapping 100 kb window along chromosome IV. This figure was created in collabora-
tion with Leonard Goldstein.
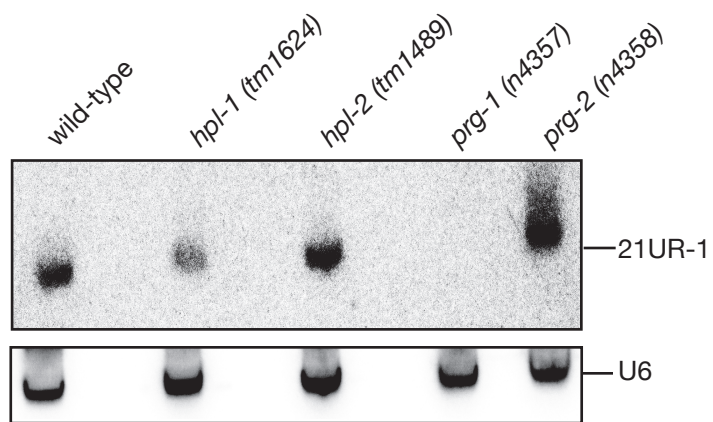
Figure 3.15 HPL-1 may be involved in the piRNA pathway
21UR-1 northern blotting of total RNA of wild-type, *hpl-1* and *hpl-2* mutant young adults (12 hr adult) of *C.elegans*. RNA from *prg-1* and *prg-2* mutants (12 hr adult) was also used. U6 is shown as loading control.

3.6 Piwi and piRNAs are expressed in the *C. elegans* male and female germline

The developmental dynamics of piRNAs was examined by northern blotting using probes specific to 21UR-1. The expression level was low in the L1 to L3 larval stages, higher in the L4 larval stage, and reached its highest expression level at the young adult and adult stages, the adult stages also containing embryos (Figure 3.16A). Quantitative RT-PCR (qRT-PCR) of another piRNA (21UR-5101) showed a similar pattern of expression (Figure 3.16B). Interestingly, qRT-PCR analysis of *prg-1* and *prg-2* mRNAs showed a similar expression pattern to piRNAs during development (Figure 3.17A & B). The temporal expression pattern of both piRNAs and *piwi* mRNAs correlated with the proliferation of the germline.

Previous studies showed that expression of piRNAs and Piwi are restricted to germ cells of most animals, apart from *Drosophila* (Carmell et al., 2007; Cox et al., 1998; Deng and Lin, 2002; Houwing et al., 2007; Kuramochi-Miyagawa et al., 2004; Kuramochi-Miyagawa et al., 2001; Lau et al., 2006). The above expression study of two individual piRNAs (21UR-1 and 21UR-5101) correlated with the proliferation of germ cells (Figure 3.16A & B). This observation encouraged me to investigate further whether piRNAs are expressed exclusively in the germline in *C. elegans*. Therefore, I tested piRNA expression in two temperature sensitive mutant animals, *glp-4(bn2ts)* and *glp-1(e2144lf),* both of which are almost completely devoid of germ cells at the restrictive temperature (25ºC) (Beanan and Strome, 1992). Expression of piRNA 21UR-1 was absent in these two mutants (Figure 3.18). This result suggests that piRNAs are expressed exclusively in the germ cells of *C. elegans*, as observed in other organisms.

To check whether piRNAs are expressed at the mitotic stage of early germline development, *glp-1 (ar202gf,ts)* temperature sensitive mutant animals were grown at the restrictive temperature (25ºC). At this temperature, mutants accumulate mitotic germ cells but these do not enter into meiosis (Pepper et al., 2003). Expression of

21UR-1 was unchanged in *glp-1 (ar202gf, ts)* mutants at the restrictive temperature compared to wild-type animals (Figure 3.18). Next, I tested whether the expression of piRNAs is male or female germline specific. Two temperature sensitive mutants, *fem-1 (hc17ts)* (Kimble et al., 1984) and *fem-3 (q22sd,ts)* (Barton et al., 1987), which are devoid of sperm and oocytes respectively at the restrictive temperature (25ºC), were grown at 25ºC. RNA was prepared and a northern blot was performed to detect 21UR-1. Expression of 21UR-1 was detected in both mutants (Figure 3.18).

Expression of *prg-1* mRNA was also examined in these germline mutants at the restrictive temperature and the same expression pattern was observed as for the 21UR-1 piRNA. Expression of *prg-1* mRNA was absent in *glp-4 (bn2ts)* mutants. However, expression of *prg-1* mRNA was detected in RNA from *fem-1 (hc17ts)*, *fem-3 (q22sd, ts)*, and *glp-1 (ar202gf, ts)* mutant animals at the restrictive temperature (Figure 3.19A). *prg-2* mRNA expression was also dramatically reduced in *glp-4 (bn2ts)* mutants (Figure 3.19B). Taken together, these data suggest that piRNAs and Piwi mRNAs are expressed specifically in the female and male germline.
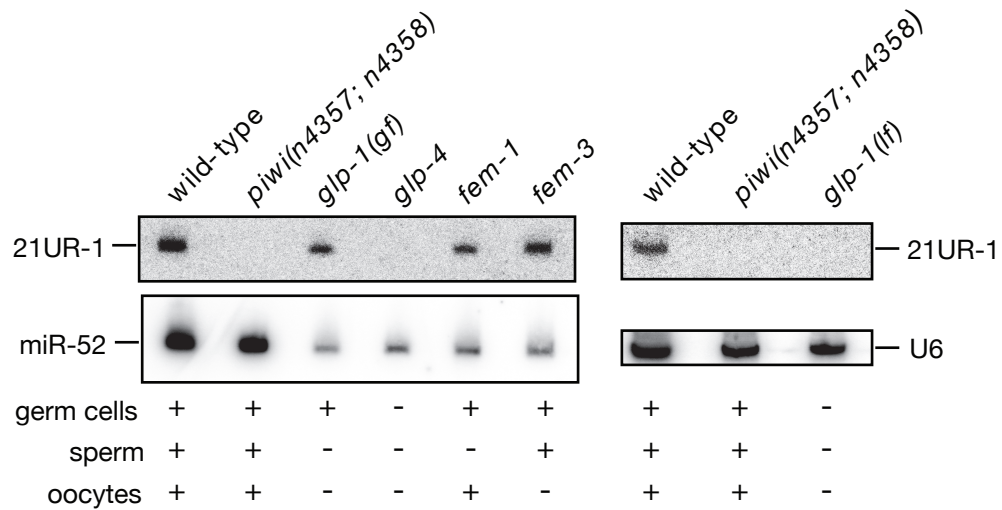
A

B
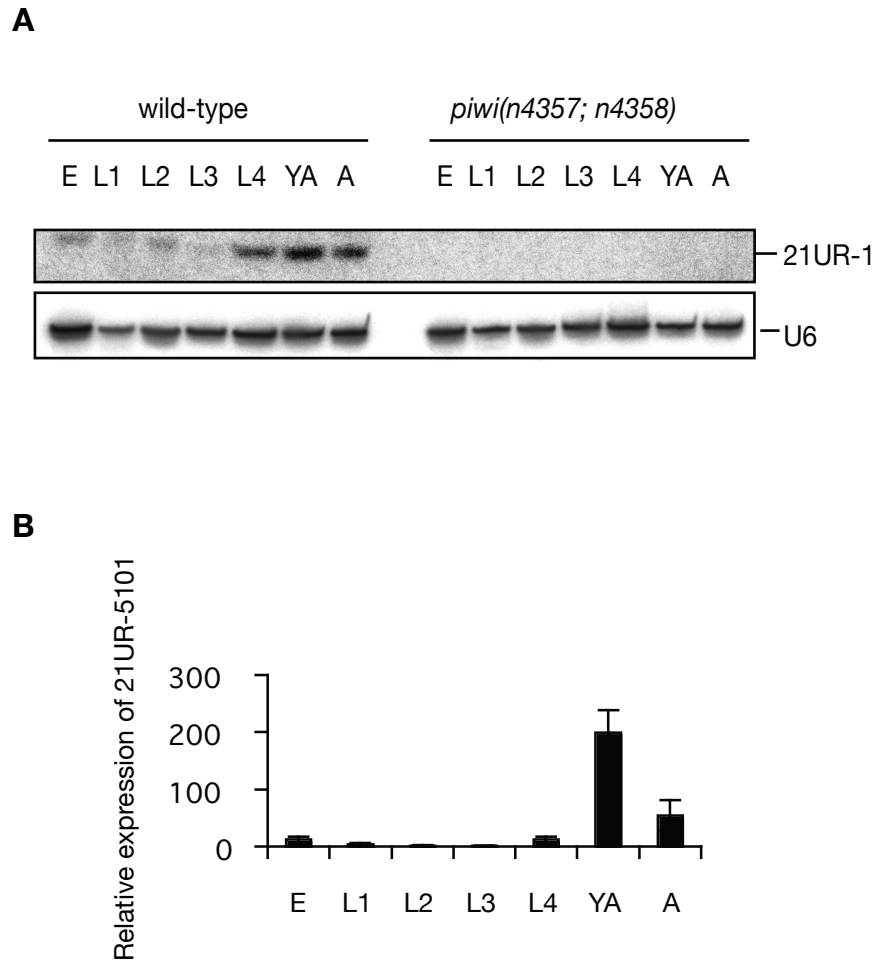
Figure 3.16 Profile of 21UR-1 expression during development
(A) Equal amount of total RNA was prepared from different developmental stages from wild-type
and *piwi* mutants. E, embryo. L1-L4, larval stages 1-4. YA, 12-hour adult. A, 24-48-hour adult.
Northern blot was performed and probed for 21UR-1. U6 used as loading control. (B) Quantitative
RT-PCR of 21UR-5101. miR-52 expression was used as internal control. Data are from three inde-
pendent biological replicates. Error bars represent standard error of the mean.
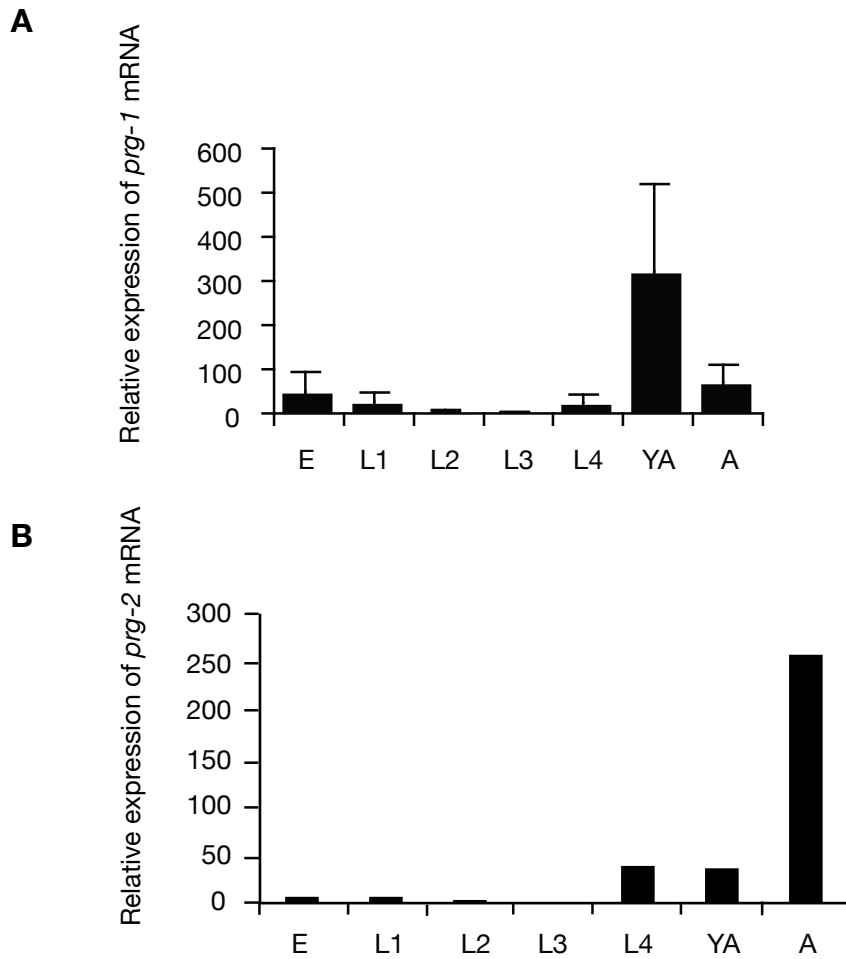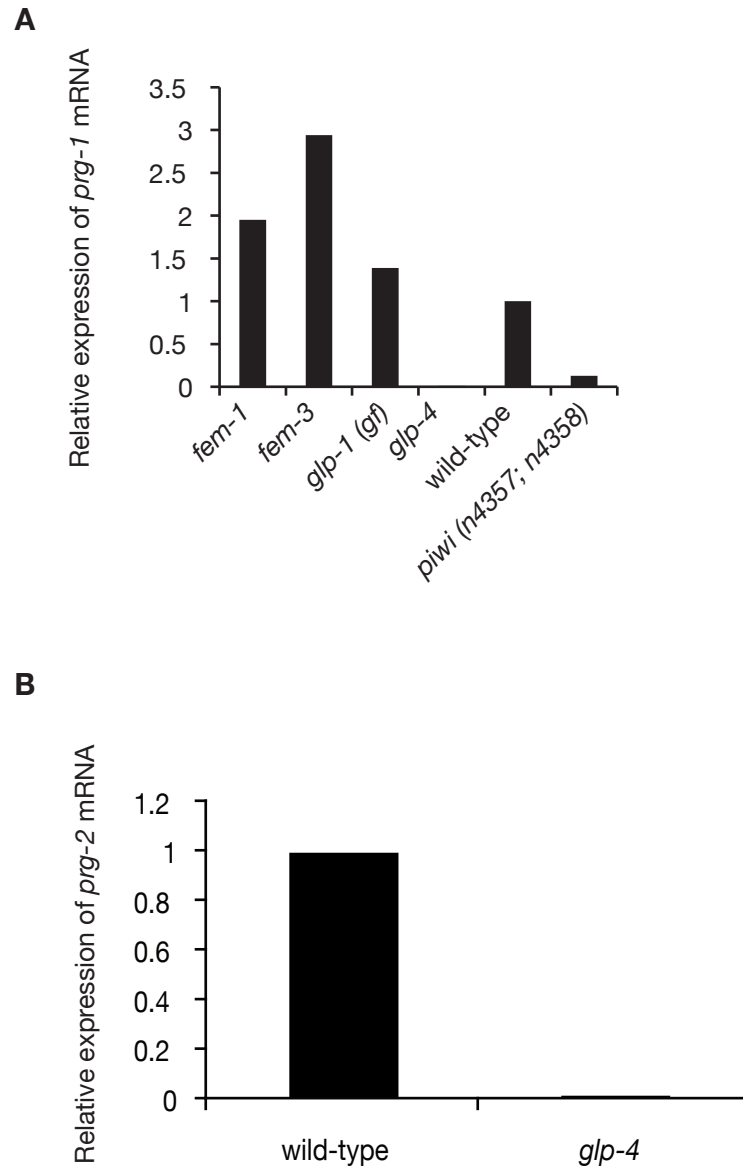
Figure 3.17 Temporal expression of PRG-1 and PRG-2 during development
Quantitative RT-PCR of *prg-1* mRNA (A) and *prg-2* mRNA (B). Actin mRNA was used as an internal control for normalization. Data are from three independent biological replicates. Error bars represent standard error of the mean. E, embryo. L1-L4, larval stages 1-4. YA, 12-hour adult. A, 24-48-hour adult.

	wild-type	piwi(n4357; n4358)	glp-1(gf)	glp-4	fem-1	fem-3	wild-type	piwi(n4357; n4358)	glp-1(lf)
germ cells	+	+	+	-	+	+	+	+	-
sperm	+	+	-	-	-	+	+	+	-
oocytes	+	+	-	-	+	-	+	+	-

Figure 3.18 piRNA expression is restricted to the male and female germline

Total RNA was prepared from young adult animals of wild-type and mutants. *glp-1(gf)*, *glp-1(lf)*, *glp-4*, *fem-1* and *fem-3* L1 larvae were grown to 12-hour adult stage at 25°C. 20 μg of total RNA was loaded in each lane, northern blot was performed using 21UR-1 probe. U6 and miR-52 northern blots are shown as loading controls.

A

B

**Figure 3.19 Expression of *prg-1* and *prg-2* is restricted to germline**
Quantitative RT-PCR of *prg-1* mRNA (A) and *prg-2* mRNA (B) was performed using total RNA isolated from 12 hour adult animals grown at 25°C. Actin mRNA was used as an internal control.

3.7 The role of Piwi and piRNAs on gene expression

I have demonstrated that 21U-RNAs or piRNAs are absent or dramatically reduced in *piwi* mutants compared to wild-type (Figure 3.8). On chromosome IV hundreds of protein-coding genes are present and 21U-RNAs are dispersed both intergenically and within introns of genes. The vast majority of 21U-RNAs were mapped to two broad regions of chromosome IV, 4.5-7Mbp and 13.5-17.2Mbp respectively (Ruby et al., 2006). Therefore, I decided to investigate the role of Piwi and piRNAs in the regulation of gene expression, either on chromosome IV or globally. To address this question, total RNA was prepared from synchronized young adult animals of wild-type and *piwi(n4357; n4358* and *n4503 nDf57)* mutant animals and hybridized to an Affymetrix gene expression array. This experiment was carried out using three independent biological samples. Surprisingly, I observed few changes in global gene expression between wild-type and *piwi* mutants (Figure 3.20A). Specifically, gene expression differences of biological replicates of the same genotype were of similar in magnitude to gene expression differences between samples of different gentotype. In parrticular, genes located within and around the 21U RNA loci on chromosome IV were not significantly altered their expression. However, major sperm protein (*msp*) genes were downregulated in *piwi* mutants compared to wild-type animals (Figure 3.20B). This observation may be explained by the reduced number of sperm in *piwi* mutant animals (Figure 3.2B). I was expecting to observe increased gene expression of some genes in the *piwi* mutants compared to wild-type animals; however, no such upregulation was observed. As mentioned above, the vast majority of 21U-RNAs were mapped to two clusters of chromosome IV (Ruby et al., 2006). To investigate this further, I tetsed expression of a panel of genes located within or around 21U RNA loci and classified them into 2 classes. Class I contained protein-coding genes present at one of the clusters of chromosome IV with a high number of 21U-RNAs within their genomic loci. Class II contained protein-coding genes present at one of the clusters of chromosome IV with no or few 21U- RNAs within their genomic loci. Expression of these genes was checked by qRT-PCR on isolated gonads of wild-type and *piwi* mutants. These experiments clould have revealed gene expression changes in the

gremline, which might have been masked by analyzing RNA from whole animals. No significant changes were observed among class II genes (Figure 3.21B). However, an increased level of gene expression was observed for two genes, *eps-8*, which encodes an epidermal growth receptor kinase substrate (Figure 3.21A), and the open reading frame H08M01.2, which is predicted to encode a RhoGAP protein. The significance to these results remains to be determined. However, there is no clear evidence suggesting that piRNAs regulate local gene expression in ther germline in *C. elegans*.
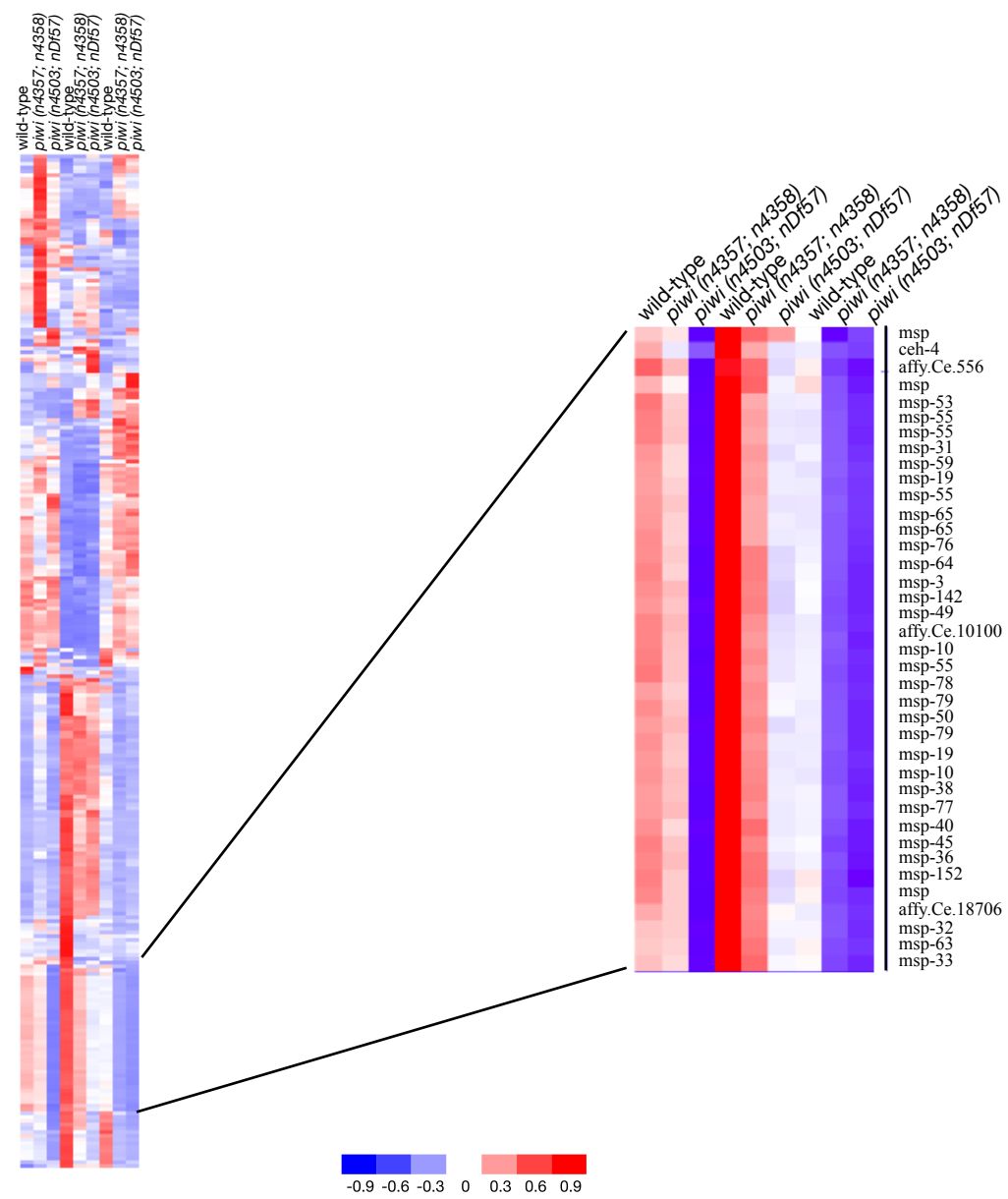
Figure 3.20 *C. elegans piwi* mutants show few differences in global gene expression
RNA was prepared from synchronized young adults (12 hr adult) from wild-type and *piwi* mutants, equal amount of total RNA used for Affymetrix analysis. Three different biological samples were used. Heat map is representing upregulated and downregulated genes. A significant part of the heat map is enlarged to show the unregulated genes and name of the genes are listed. Increased and decreased levels of gene expression are indicated in red and blue in colour. Scale bar represents log_2 intensity change. This figure is created in collaboration with Eric Miska.
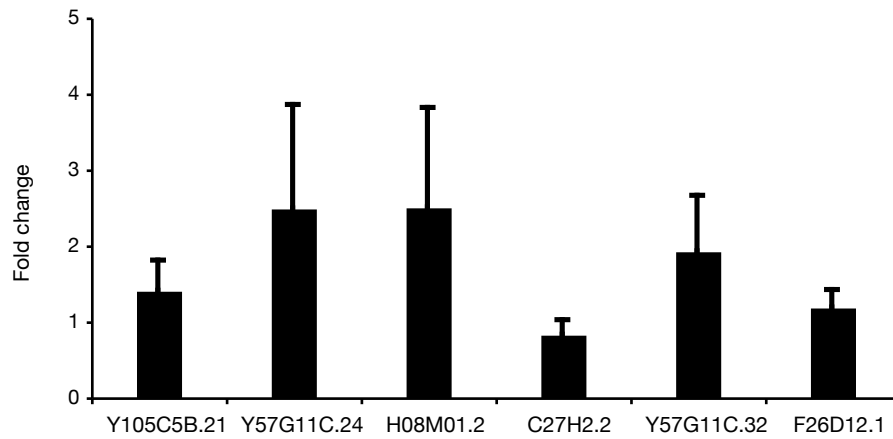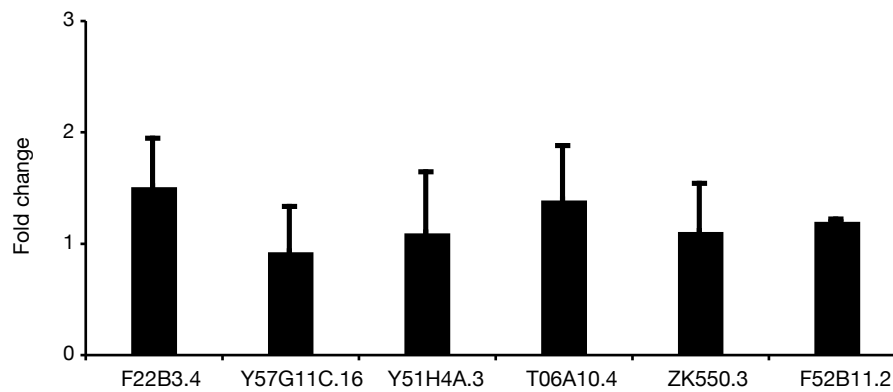
Figure 3.21 Gene expression within and around major clusters of 21U-RNA loci are unaffected by Piwi and piRNAs
lysate were prepared from gonads of wild-type and *piwi* mutants and used for quantitative RT-PCR analysis of class I and II genes. Expression of indiviual genes was normalized with Actin mRNA expression. Normalized values were set to one for wild-type and fold changes were calculated for *piwi* mutants relative to wild-type.

3.8 Piwi is specifically required to silence Tc3 DNA transposons

It has previously been shown that Piwi proteins are required for transposon silencing in *Drosophila* and mice (Aravin et al., 2001; Aravin et al., 2007b; Kuramochi-Miyagawa et al., 2008; Savitsky et al., 2006; Vagin et al., 2006). Therefore, I examined whether Piwi is required for transposon silencing in *C. elegans*.

Seven different families of DNA transposons exist in *C. elegans*, named Tc1 to Tc7. Tc1 and Tc3 are the most well-characterized and abundant transposons in *C. elegans* (Consortium, 1998). The Tc1 element is 1,610 bp long and contains two 54 bp terminal inverted repeats (TIRs) (Rosenzweig et al., 1983) (Figure 3.22A). The Tc3 element comprises 2,335 bp with 462 bp TIRs (Figure 3.22A). The genome of the Bristol N2 strain contains 31 and 22 copies of Tc1 and Tc3 respectively (Fischer et al., 2003) as shown in Figure 3.22B. Tc1 and Tc3 both encode transposases that bind to the TIRs of the respective transposon and catalyse excision and reinsertion of the transposon into target DNA (a TA dinucleotide). Transposition leaves behind double-strand breaks that are repaired by the DNA repair machinery (van Luenen et al., 1994; Vos et al., 1996; Vos and Plasterk, 1994; Vos et al., 1993).

I decided to focus on testing Tc1 and Tc3 DNA transposon regulation by Piwi proteins. First, expression of Tc1 transposase mRNA was checked by quantitative RT-PCR using a Tc1-specific pair of primers, which can recognize 15 copies of the 31 copies of the Tc1 transposase mRNA encoded in the genome. I did not observe any elevated expression of Tc1 transposase mRNA in three independent mutants of *prg-1* and *prg-2* nor in two independent mutants of *piwi* compared to wild-type (Figure 3.23A). However, expression of Tc1 transposase mRNA was elevated approximately 50-fold in *mut-7* mutants compared to wild-type. *mut-7* mutants have previously been shown to be defective in transposon silencing. (Ketting et al., 1999).

Next, I checked the expression of Tc3 transposase mRNA by qRT-PCR using a Tc3-specific primer pair, which can detect 20 copies out of 22 copies of the Tc3 transposase

mRNA encoded in the genome. The result showed that Tc3 transposase mRNA expression increased approximately 1.5-3 fold in two independent mutants of *prg-1 (n4357* and *n4503)* and 3-5 fold in two independent *piwi (n4357; n4358 and n4503; nDf57)* mutants when compared to wild-type (Figure 3.23B). However, the expression level of Tc3 transposase mRNA remained unchanged in three independent mutants of *prg-2*(*n4358, nDf57* and *tm1094*) compared to wild-type (Figure 3.23B). Additionally, I observed higher expression of Tc3 transposase mRNA in the *tm872* allele compared to the other two alleles of *prg-1*(*n4357* and *n4503)*, which might be an allele specific effect or perhaps due to background mutations. Tc3 transposase mRNA expression increased 5-fold in *mut-7* mutants compared to wild-type as expected (Ketting et al., 1999). Therefore, these data suggest that PRG-1 and Piwi are required for transposon silencing.

To strengthen the quantitative RT-PCR data, which suggest that PRG-1 and Piwi required for Tc3 transposon silencing, a genetic analysis of transposition was performed in collaboration with Marloes Bagijn. Animals carrying *unc-22::Tc1*, *unc-22::Tc3* or *unc-22::Tc4* alleles have a twitching phenotype due to the insertion of transposons into the *unc-22* muscle gene. This twitching phenotype can be easily scored. Therefore, *unc-22::Tc1, un-22c::Tc3 and unc-22::Tc4* transgenic animals were crossed to *prg-1 (n4357* and *n4503)* or *piwi (n4357; n4358* and *n4503; nDf57)* or *mut-7 (pk204)* mutants. Reversion of *unc-22* transposon insertion alleles by transposon excision from the *unc-22* gene caused restoration from a twitching phenotype to normal movement of animals, those moving animals were plated again to observe germline transmission of the transposition event. The reversion rate was calculated as the number of normal moving animals to the number of twitching animals on a plate. *unc-22* reversion rates of Tc1, Tc3 or Tc4 insertion alleles were less than 10^{-6} in a wild-type background (Table 3.2). In *mut-7* mutant animals, the reversion rate was increased 100 fold (10^{-4}) (Table 3.2), as reported previously (Ketting et al., 1999). Transposition rates of Tc1 and Tc4 were not affected in *prg-1 (n4357)* and *piwi (n4357;n4358)* mutants. However, the reversion rate of Tc3 was increased 100 fold in *prg-1 (n4357)* mutants and 1000 fold in two independent *piwi (n4357;n4358* and *n4503;nDf57)* mutants

(Table 3.2). The reversion rate of Tc3 in the *prg-1(tm872)* allele background was higher than for the *prg-1(n4357)* allele, again suggesting an allele-specific effect observed previously in qRT-PCR analysis.

Taken together, quantitative RT-PCR and genetic analysis data demomstrate that PRG-1 and Piwi are specific and powerful suppressors of Tc3 transposition in the germline of *C. elegans*.
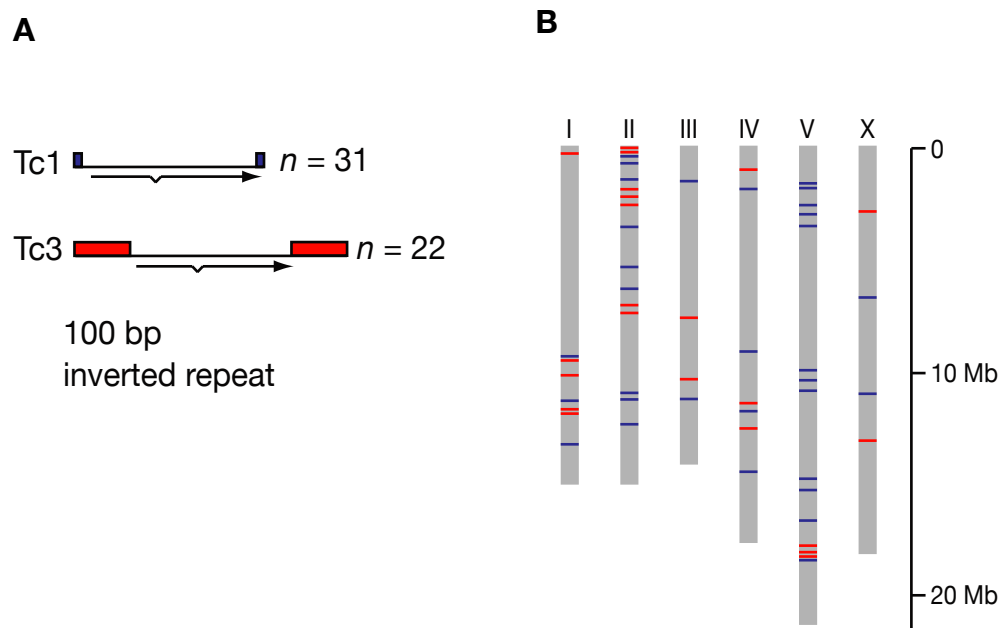
Figure 3.22 Schematic diagram of the genomic structure of Tc1 and Tc3
(A) Tc1 and Tc3 are encode a single spliced transcript for transposase, and flanked by terminal inverted repeats (TIRs) (solid coloured boxes). bp, base pairs. n, number of copies in the wild-type strain N2. (B) Distribution of Tc1 and Tc3 transposons in the *C.elegans* genome. Red and blue lines are indicating genomic positions of Tc3 and Tc1 on the chromosomes of Bristol N2 strain.
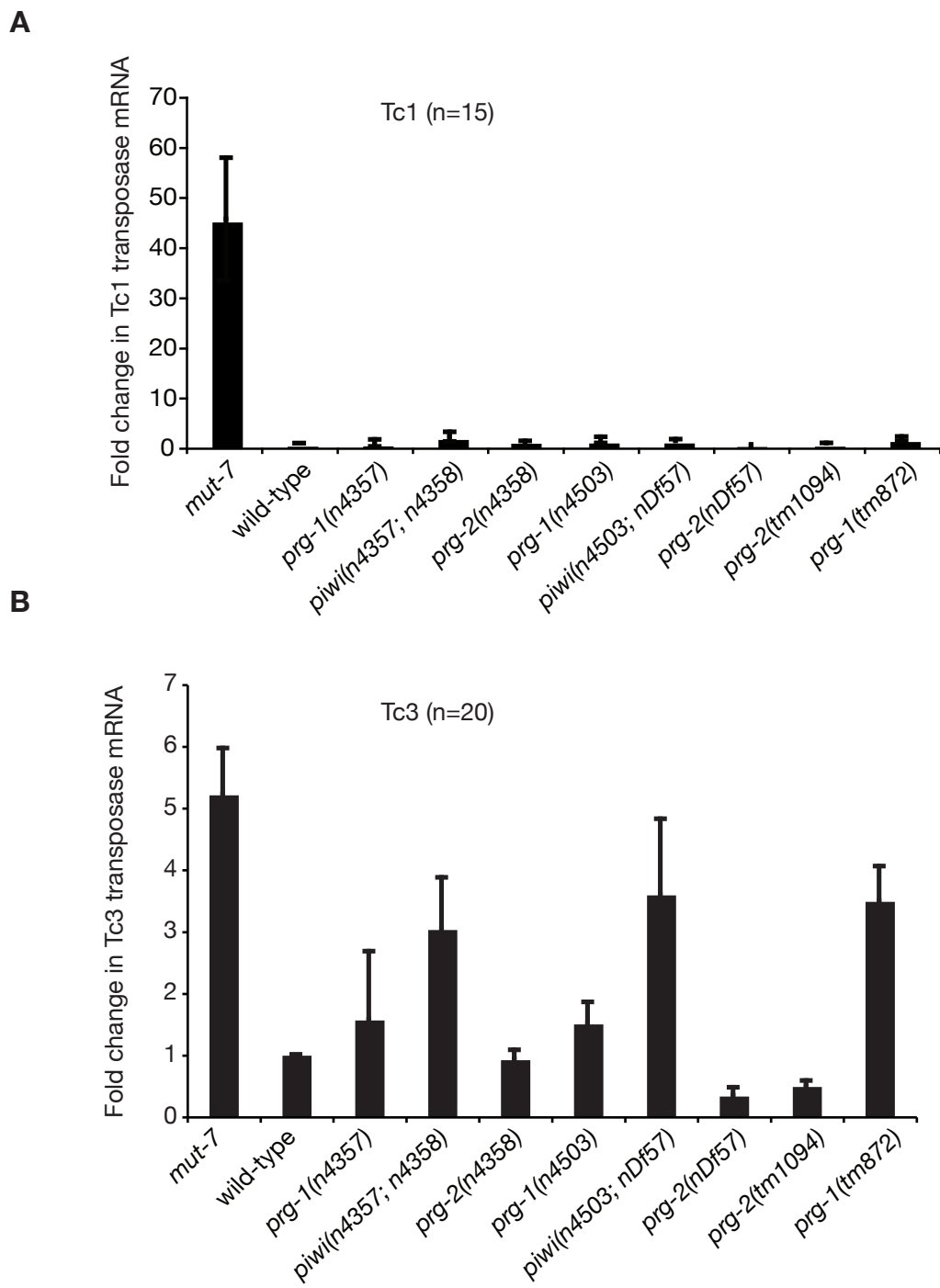
Figure 3.23 (A & B) Quantitative RT-PCR of Tc1 and Tc3 transposase mRNA
Quantitative RT-PCR of Tc1 or Tc3 transposase mRNA. As the genomic copies of Tc1 and Tc3 have minor sequence variations, the number of transposon loci amplified by each qRT-PCR primer pair are shown (*n*). Actin mRNA was used as internal control. Expression levels shown are relative to levels from wild-type RNA. Data are from three independent biological replicates. Error bars represent standard error of the mean.

**Table 3.2 PIWI is required to inhibit Tc3 transposition in the germline**

	unc-22::Tc1	unc-22::Tc3	unc-22::Tc4
wild-type	$<10^{-6\,a}$	$<10^{-6\,a}$	$<10^{-6\,a}$
prg-1(n4357)	10^{-6}	10^{-4}	10^{-6}
prg-1(tm872)	n.d.	10^{-3}	n.d.
piwi(n4357; n4358)	10^{-6}	10^{-3}	10^{-6}
piwi(n4503; nDf57)	n.d.	10^{-3}	n.d.
mut-7	10^{-4}	10^{-4}	10^{-4}

Transposition rates were estimated by scoring *unc-22* reversion rates. Germline excision events were verified by scoring the progeny of revertants. Animals were grown at 20°C. [a] No *unc-22* reversions were observed. n.d.: not determined.

## 3.9 Piwi acts upstream of an endogenous siRNA pathway

Previous work demonstrated a role of MUT-7 (Ketting et al., 1999) and endogenous siRNAs in germline transposon silencing (Sijen and Plasterk, 2003). Our findings suggested that Piwi proteins and piRNAs are also required for germline transposon silencing. To determine how these two pathways might be connected, I first tested whether MUT-7 and endogenous siRNAs act upstream of piRNAs. Northern blotting and high-throughput Solexa sequencing data showed that piRNA expression was not affected in *mut-7* mutants (Figure 3.24A & data not shown). Next, I examined whether Piwi proteins and piRNAs might act upstream of MUT-7 and endogenous siRNAs. For that, I mapped piRNAs to transposons in the genome and none of the piRNAs mapped to the Tc1 and Tc4 DNA transposons. However, a single piRNA (21UR-139) mapped to the sense strand of the Tc3 DNA transposon at 20 of the 22 copies of the insertion in the wild-type genome.

Two distinct populations of endogenous siRNAs are involved in RNAi in *C. elegans*: primary siRNAs and secondary siRNAs. Secondary siRNAs constitute the vast majority of endogenous siRNAs (Pak and Fire, 2007; Sijen et al., 2007). Secondary siRNAs carry 5′ di- or tri-phosphates and are thought to be generated by RdRPs through unprimed RNA synthesis (Pak and Fire, 2007; Sijen et al., 2007). To determine how siRNAs are distributed throughout the genome, 5′ independent small RNA libraries of wild-type and *piwi* (*n4357*; *n4358*) mutants were prepared, which capture also secondary siRNAs. After high-throughput sequencing, analysis was done in collaboration with Leonard Goldstein. A large number of endogenous siRNAs were found to map to Tc1 and Tc3 DNA transposons. The distribution of endogenous siRNAs to the Tc1 transposase and TIRs was not affected in *piwi (n4357; n4358)* mutants when compared to wild-type (Figure 3.25B). However, endogenous siRNAs mapping to the Tc3 transposase and TIRs was dramatically reduced in *piwi (n4357; n4358)* mutants compared to wild-type (Figure 3.25A).
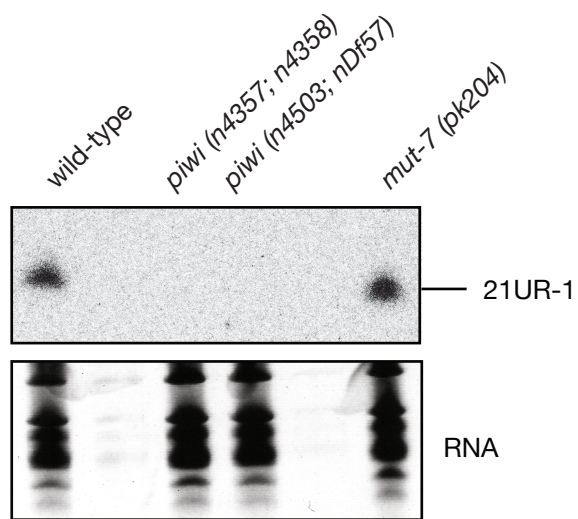
Figure 3.24 piRNA expression is independent of MUT-7
21UR-1 northern blotting of total RNA (40 μg) of wild-type, two indepedent *piwi* double mutants: *piwi (n4357; n4358)* and *piwi (n4503; nDf57)*, *mut-7* mutants (12 hr adult) of *C. elegans.* Total RNA is shown as loading control (GelStar).
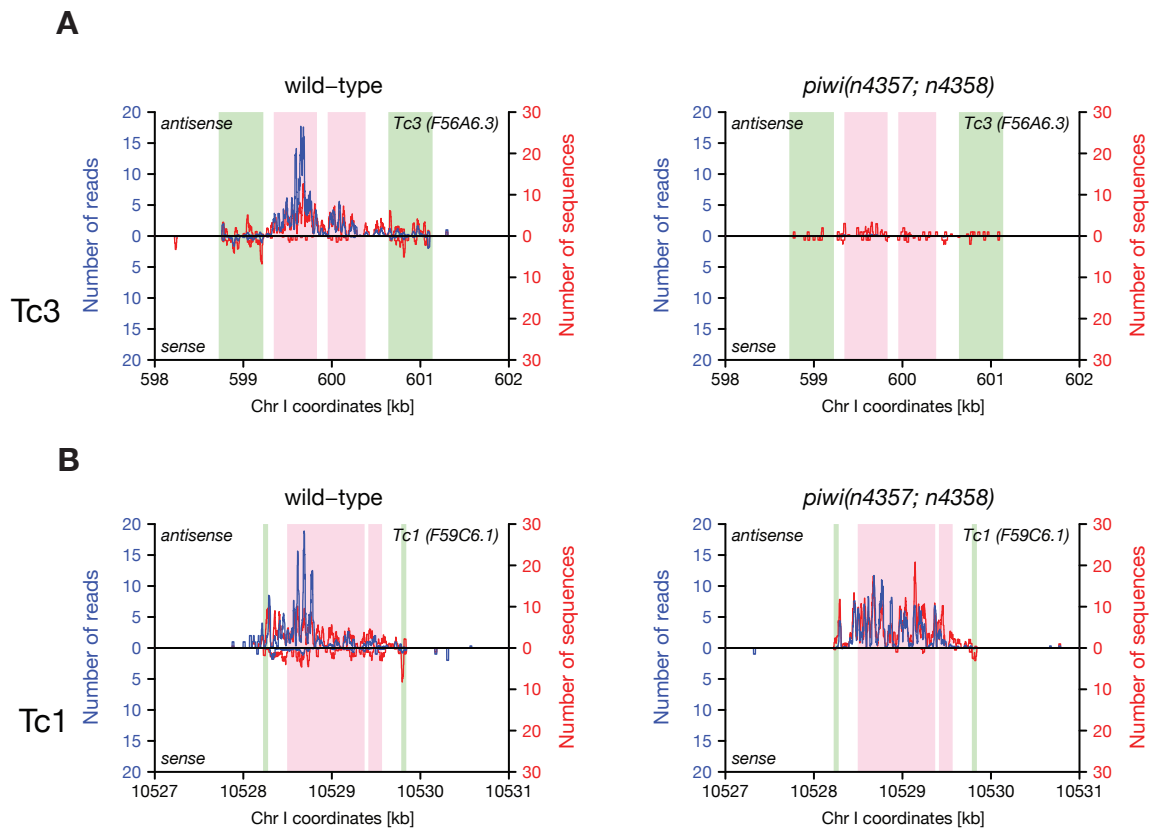
Figure 3.25 Tc3-associated endogenous siRNAs are almost absent in *piwi* mutants
Endogenous siRNAs mapping to the loci of Tc3 (A) and Tc1 (B) on chromosome I as identified by high-throughput sequencing of 5′ independent wild-type (left) and *piwi* mutant (right) libraries. Inverted repeat and exon sequences are indicated in green and pink respectively. The number of aligned sequence reads (blue) and number of aligned unique sequences (red) were plotted for each base pair position, with the top and bottom graph in each panel corresponding to the antisense and sense strand relative to the transposase transcript. Read and sequence counts were corrected for multiple alignments to the genome. The total number of reads from wild-type and *piwi* mutant librar-ies were comparable (2,963,895 and 3,017,027 of reads with perfect matches to the reference genome respectively).This figure was created in collaboration with Leonard Goldstein.

To confirm these data, I performed RNase protection assays (RPA) using sense and anti-sense probes for both the Tc1 and Tc3 transposons. Both sense and antisense probes were generated through *in vitro* transcription of cloned DNA templates of Tc1 and Tc3 using T7 or SP6 RNA polymerase (Figure 3.26A). Probes were gel purified and hybridized against small RNAs from wild-type animals. Anti-sense endogenous siRNAs were detected using sense probes for both Tc1 (370) and Tc3 (464) transposases. However, endogenous sense siRNAs were not detected using antisense probes of Tc1 and Tc3 (Figure 3.27). These results demonstrated that most of the endogenous siRNAs mapping to Tc1 and Tc3 transposases are antisense, which is in agreement with my high-thorughput sequencing data (Figure 3.25A & B). To control small RNA loading for each RPA assay, a U6 probe was used as an internal control (Figure 3.26C). RPA conditions were optimized in the absence or presence of the U6 probe with two independent Tc3 probes (Figure 3.26C & 3.28). Next, I performed a similar RPA experiment using small RNAs either from wild type, *piwi (n4357; n4358)* mutants or *mut-7* mutants using both the Tc1 or Tc3 sense probes. I found that both Tc1 and Tc3 endogenous siRNAs are almost absent in *mut-7* mutants (Figure 3.29), as shown previously for Tc1 (Sijen and Plasterk, 2003). Tc1 endogenous siRNAs were present in both wild-type and *piwi (n4357; n4358)* mutants. However, Tc3 endogenous siRNAs were absent in *piwi (n4357; n4358)* mutants compared to wild-type (Figure 3.29). Unfortunately, I was not able to detect endogenous siRNAs from Tc3-TIR using the same assay (data not shown), which might indicate a lower number of endogenous siRNAs generated from the TIR region of Tc3. Indeed, high-throughput sequencing identified a lower number of siRNAs mapping to the TIRs as compared to the transposases (Figure 3.25). Therefore, it might be technically difficult to detect TIR endogenous siRNAs by RPA.

Taken altogether, both high-throughput sequencing data and biochemical data suggest that Tc3 siRNAs are Piwi dependent. I therefore speculate that Piwi and piRNAs act upstream of a MUT-7 dependent endogenous siRNA pathway for Tc3 silencing.
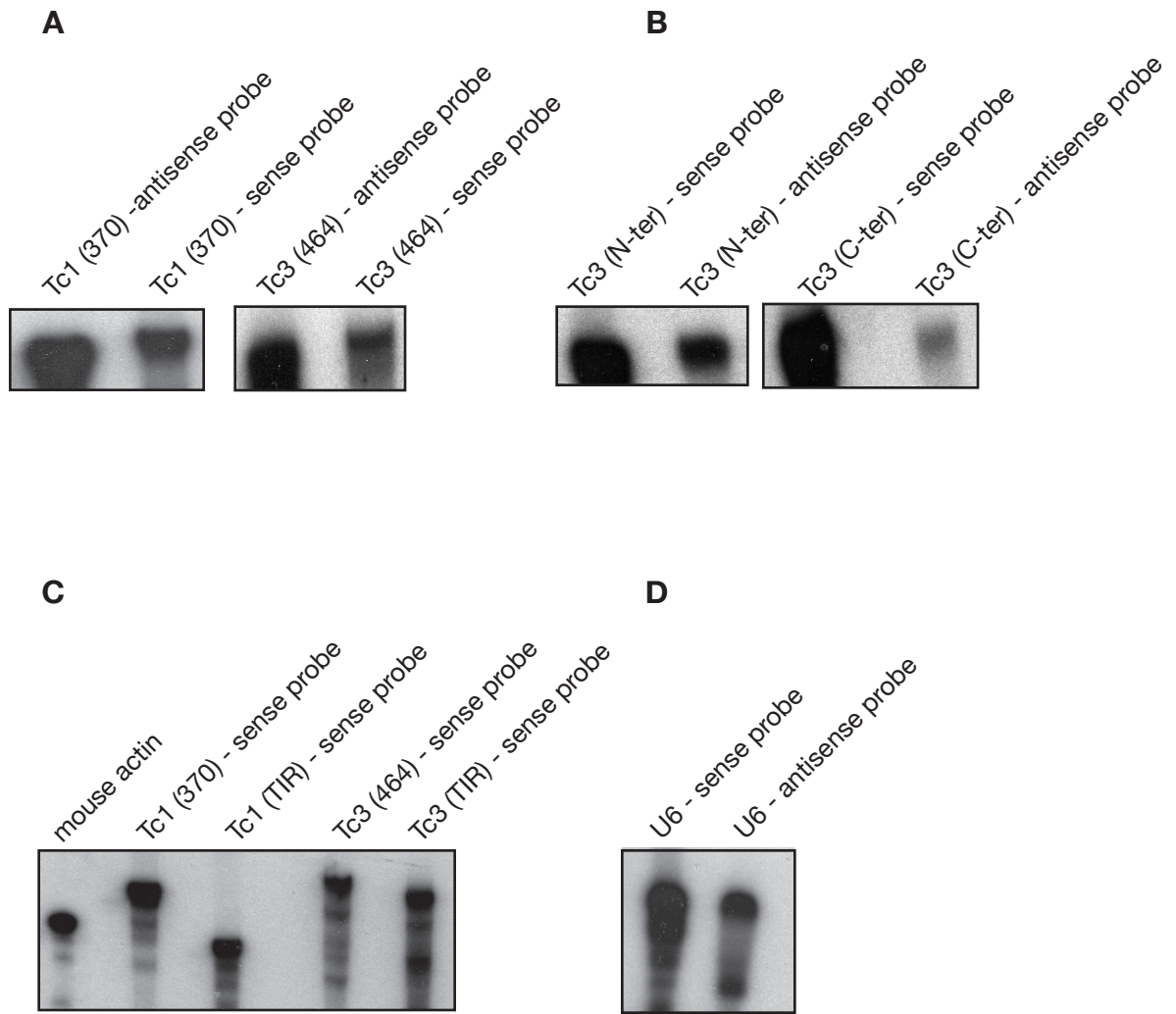
Figure 3.26 Tc1, Tc3 and U6 probes used for RNase Protection A assay (RPA)
(A, B, C and D) alpha-p32 UTP labeled sense and antisense probes were prepared by *in vitro* transcription (IVT). Probes were resolved through denaturing PAGE and exposed to X-ray film. All probes were gel purified prior to hybridization to RNA. Tc1 (370, TIR) represents 370bp of the transposase and TIR of the Tc1 transposon and Tc3 (464, N-ter, C-ter and TIR) represents 464 bp, N-terminal, C- terminal of Tc3 transposase and TIR of the Tc3 transposon, used as a DNA template for IVT.
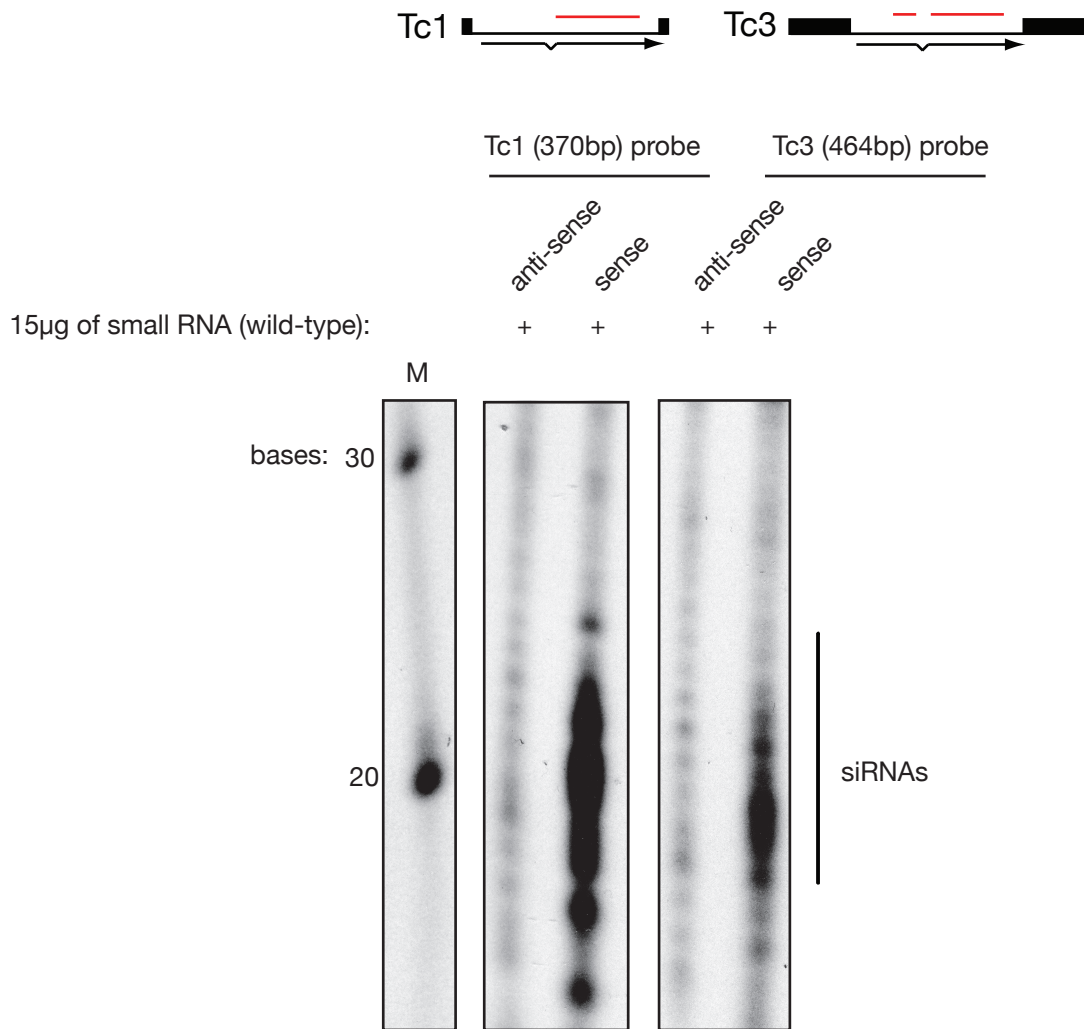
Figure 3.27 Tc1 and Tc3 antisense siRNAs were detected by sense probes of Tc1 and Tc3
RNase protection assay was performed using wild-type RNA and sense and antisense probes for Tc1 and Tc3 as indicated (red). Endogenous sense siRNAs were not detected above background levels in this assay.
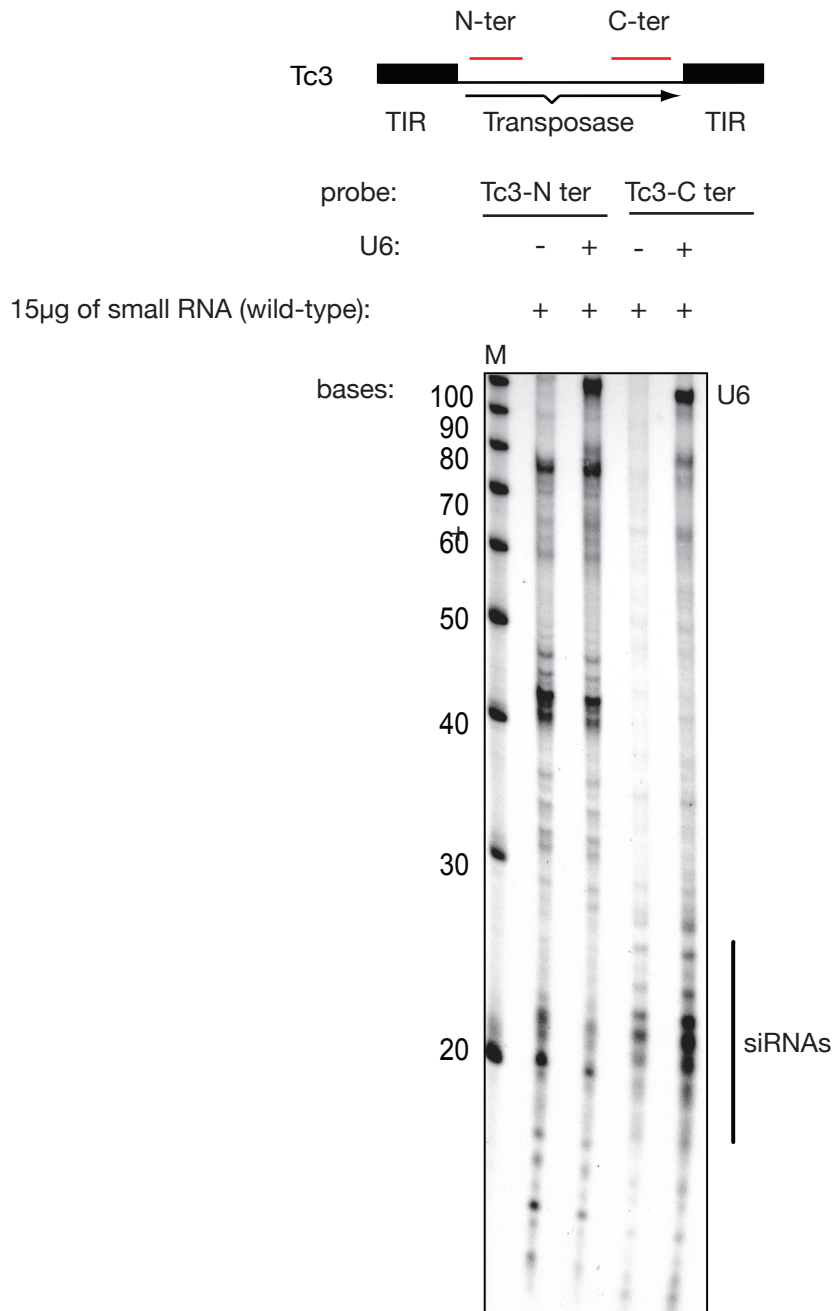
Figure 3.28 Optimization of RNase protection assays using U6 as an internal control
RNase protection assays were performed using wild-type RNA and two different sense probes for the Tc3 transposase (Tc3-N ter and Tc3-C ter), shown in red. U6 used as an internal control. Tc3 probes hybridized in presence or absence of U6 probe against small RNAs in solution to check the specificity of the probes.
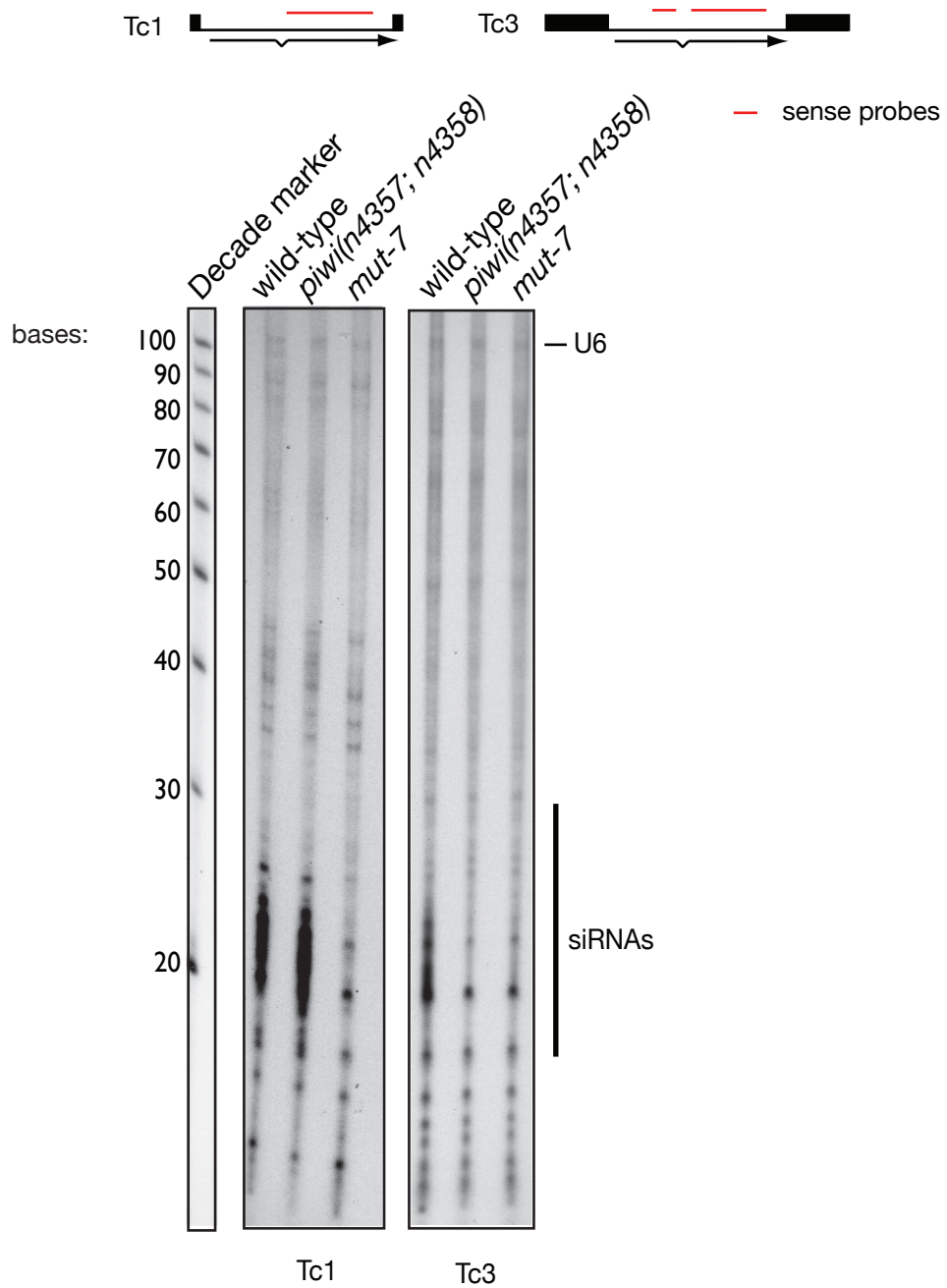
Figure 3.29 Endogenous Tc3 antisense siRNAs are dramatically reduced in *piwi* mutants
Tc3 transposase antisense siRNAs are dramatically reduced in *piwi* mutants. RNase protection assay using sense fragments of Tc1 and Tc3 transposase. Sense siRNAs were not detected above background levels using this assay. U6 was used as an internal control.

3.10 Regulation of endogenous siRNAs by Piwi-piRNA complex

These findings raised the possibility that piRNAs might be generally linked to endogenous siRNAs. To test this hypothesis, in collaboration with Leonard Goldstein, all cloned endogenous siRNAs were mapped relative to all piRNA loci (6021 loci for 5454 21U-RNAs) (Figure 3.30A). From the deep sequencing data (5′ independent libraries), small RNAs were categorized according to their length, and most of them have G residues at their 5′ end, indicating they are secondary endogenous siRNAs (Figure 3.31). Each row corresponds to each 21U RNA locus. Different colours represent different chromosomes. Each dot corresponds to the relative position of the nearby 5′ end of a cloned anti-sense endogenous siRNAs (Figure 3.30A) or sense endogenous RNAs (Figure 30B) to the 5′ end of the 21U RNA/piRNA (sense). The graph above the map represents the frequency of distances between the 5′ end of 21U RNA/ piRNA (sense) and 5′ end of anti-sense or sense endogenous siRNAs (Figure 3.30A & B). We observed anti-sense endogenous siRNAs mostly positioned 25-100bp downstream of the sense piRNA loci in wild-type and the number of anti-sense endogenous siRNAs was significantly reduced in *piwi* mutants (Figure 3.30A). However, the distributions of sense endogenous siRNAs either centered or equally positioned both up and downstream of piRNA loci in wild-type, and the number of sense endogenous siRNAs were also reduced in *piwi* mutants (Figure 3.30B). Therefore, it indicates that the endogenous siRNA pathway might be Piwi and piRNA dependent.
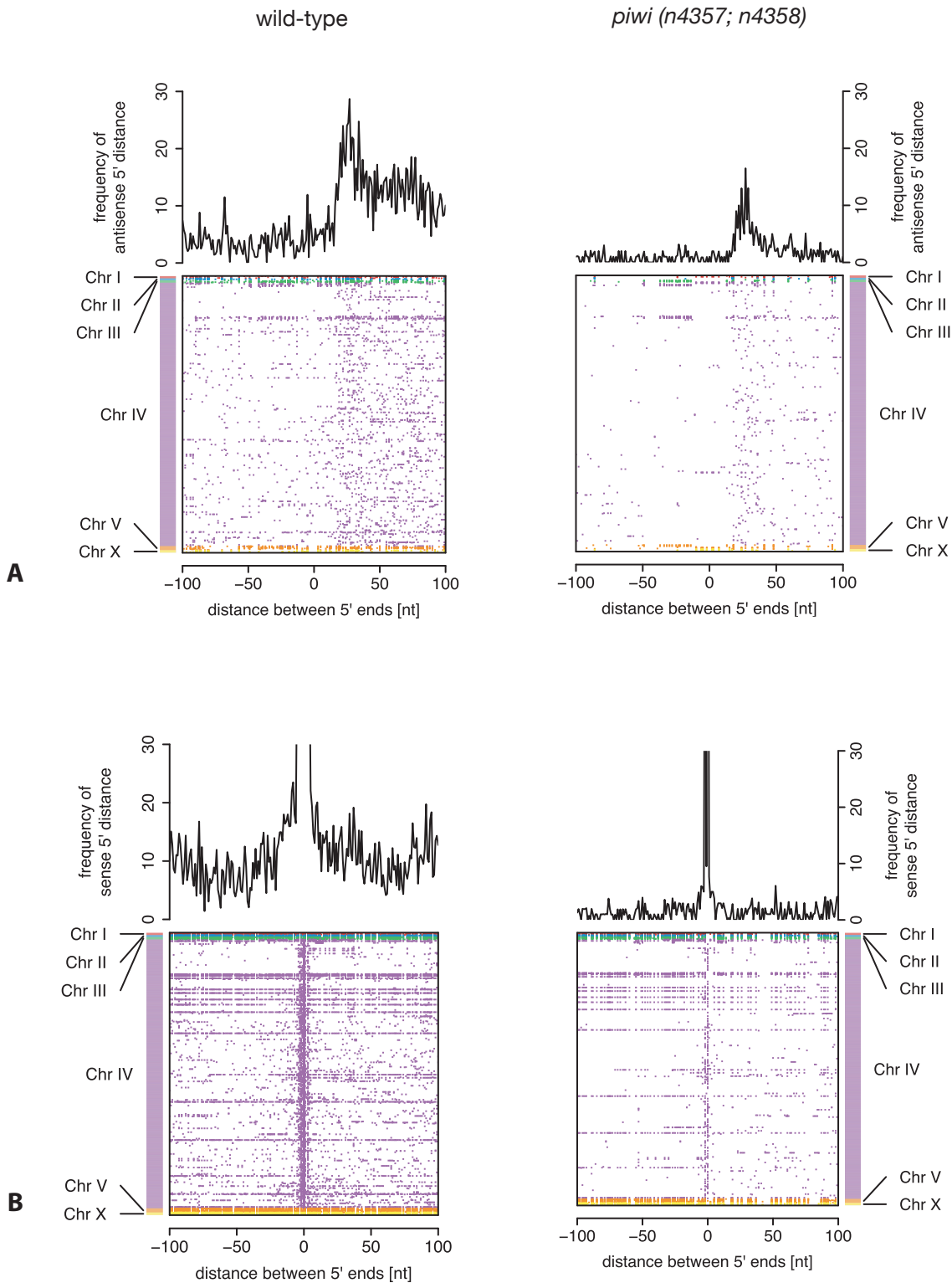
Figure 3.30 Mapping of antisense (A) and sense (B) endogenous siRNAs relative to piRNA loci
Rows corresponds to each 6021 piRNA loci, ordered by genomic position with colours represents different chromosomes. For a given row (21U RNA locus), dots correspond to the relative position of nearby antisense (A) and sense (B) small RNAs as defined by the distance of the 5' end of the cloned endogenous siRNAs relative to 21U RNA 5' end. This figure was created in collaboration with Leonard Goldstein.

113
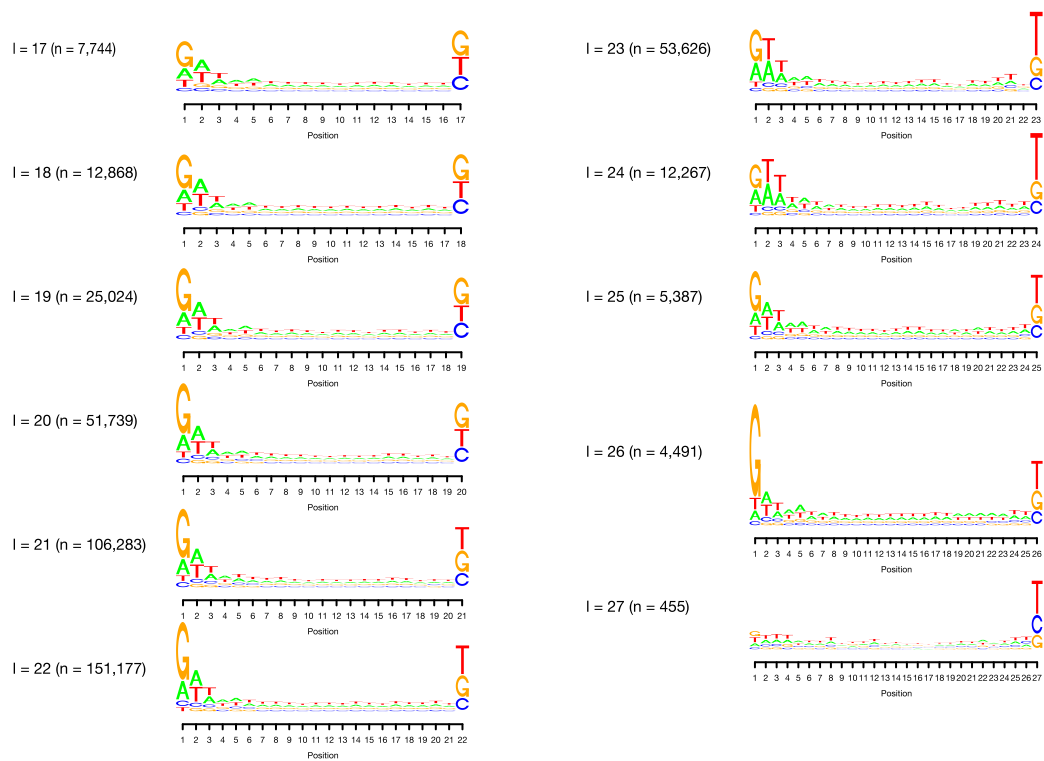
Figure 3.31 Nucleotide bias at 5' independent small RNA sequences
Small RNAs were identified by 5' independent sequencing from wild-type libraries show no A bias at 10 th position. Sequence Logos were generated separately for sets of unique sequences with identical length (l). The number of unique sequences (n) is given in brackets. 28 nt sequence reads were not shown.

4. Discussion

4.1 Introduction

To date, studies of the Piwi sub-family of Argonaute proteins in several animals have shown that their expression is restricted to germ cells. Piwi proteins are associated with non-coding small RNAs (26-31nt), called piRNAs (Piwi-interacting RNAs). Piwi and piRNAs have conserved roles in germ cell maintenance and genome stability through transposon silencing (Aravin et al., 2007a; O'Donnell and Boeke, 2007). Plants and yeast, which lack Piwi proteins, employ siRNA pathways to silence repetitve elements (Buhler and Moazed, 2007; Slotkin and Martienssen, 2007; Zhang et al., 2008). In *C. elegans*, PRG-1 and PRG-2 are the only members of the Piwi sub-family of Argonaute proteins. This Piwi sub-family of Argonautes and their associated small RNAs have not been explored. In this study I have shown that PRG-1, a Piwi sub-family Argonaute interactes with 21U-RNAs, which are 21 nt long, begin with a uridine monophosphate and are modified at their 3′ terminal ribose, and these 21U-RNAs are the piRNAs in *C. elegans*. Piwi proteins and piRNAs are both expressed in the germline and play important roles in germline development. In addition, I have shown that Piwi and piRNAs are involved in silencing Tc3 transposons in the *C. elegans* germline. The molecular mechanism of Tc3 transposon silencing is not well understood. To explore how Tc3 transposon might be regulated by piRNAs, all piRNAs were mapped to the *C. elegans* genome, but only one piRNA mapped to Tc3 transposons. Interestingly, the sequence of each piRNA is unique; there is no sequence similarity between individual piRNAs and no sequence-specific targets for piRNAs have been found in the *C. elegans* genome. Therefore, it might be possible that piRNAs regulate Tc3 transposon silencing via endogenous siRNAs. I have shown by both high-throughput sequencing and RNase protection experiments that endogenous siRNAs mapping to the Tc3 transposase and terminal inverted repeats are present in wild-type animals but are absent in *piwi* mutants, suggesting a link between a piRNA and endogenous siRNAs. Additionally, I have shown that piRNAs can act upstream of an

endogenous siRNA pathway. Remarkably, many of these endogenous siRNAs mapped antisense and 25-100nt downsteam of piRNAs and these endogenous siRNAs were found to be dramatically reduced in *piwi* mutants. These data indicated that the link between piRNAs and siRNAs might not be restricted to Tc3 transposon silencing but might also play a potential role in global gene regulation.

4.2 21U-RNAs are piRNAs with surprising features

I have identified 21U-RNAs as the piRNAs of *C. elegans* (Figure 3.5-3.11). Immunoprecipitation of anti-PRG-1 was able to precipitate endogenous 21U-RNAs (Figure 3.9). A parallel study in the Mello lab also found that PRG-1 immunoprecipitates are 100-fold enriched in 21U-RNAs, strongly supporting our findings. *C. elegans* 21U-RNAs/piRNAs share several features with piRNAs from other species. Firstly, piRNAs from *C. elegans* carry a 5′ end uridine monophosphate, as do piRNAs from *Drosophila*, zebrafish and mice (Farazi et al., 2008; Klattenhoff and Theurkauf, 2008). Secondly, *C. elegans* piRNAs are modified with 2′ or 3′-O-methyl groups at their 3′ terminal ribose (Figure 3.9) (Ruby et al., 2006). Again, this feature is shared by piRNAs from *Drosophila*, zebrafish and mice (Horwich et al., 2007; Houwing et al., 2007; Kirino and Mourelatos, 2007; Ohara et al., 2007; Saito et al., 2007). However, 3′ end modification is not an unique feature of piRNAs as plant miRNAs, siRNAs and *Drosophila* siRNAs also carry this modification (Ghildiyal et al., 2008; Kawamura et al., 2008; Li et al., 2005; Okamura et al., 2008).

Otherwise piRNAs from *C. elegans* are quite different from piRNAs of other species. piRNAs from *C. elegans* are smaller (21nt) as compared to piRNAs from *Drosophila* (23-28nt), zebrafish (~29nt) , mice (26-31nt) and rat (29-30nt). Most of the piRNAs from mice and rat are clustered in intergenic regions. The distribution of mammalian piRNAs is very different from piRNAs in *C. elegans*. Uniquely, *C. elegans* piRNAs map to two broad but distinct regions of chromosome IV, one spanning between 4.5-7M and other spanning between 13.5-17.2M (Ruby et al., 2006). The role of these clusters is still unclear and also not conserved in related nematode species (data not

shown). Finally, unlike in other species, *C. elegans* piRNAs are not generally derived from transposons. Only a single piRNA (21UR-139) maps to the sense strand of the Tc3 transposon.

4.3 Biogenesis of piRNAs

It is still unclear how piRNAs are generated and what their precursors are. To date there is no evidence that piRNAs are generated from a dsRNA precursor. The strong strand bias of clustered piRNAs from mice suggests that they might be generated from a single-stranded precursor. Supporting this notion, strand specific RT-PCR experiments detected putative sense primary transcripts but failed to detect any antisense transcripts (Watanabe et al., 2006). Unlike miRNA precursors, the regions covering piRNAs do not fold into stem-loop structures. Therefore, the biogenesis pathway of piRNAs appears to be distinct from that of miRNAs. The striking difference in the genomic organization of *C. elegans* piRNA clusters compared to vertebrates and *Drosophila* might suggest a divergent mechanism of biogenesis. Northern blotting and high-throughput sequencing data suggest that *C. elegans* piRNAs are Dicer independent, as is the case for piRNAs in mice, zebrafish and *Drosophila* (Grivna et al., 2006b; Houwing et al., 2007; Vagin et al., 2006) (Figure 3.13 & 3.14). The distribution of piRNAs in the genome and the lack of a local strand bias in *C. elegans* suggests that it is unlikely that piRNAs are generated from double stranded precursors.

In *C. elegans*, RNA-dependent RNA polymerase (RdRP) is required for the unprimed synthesis of secondary siRNAs (Pak and Fire, 2007; Sijen et al., 2001; Sijen et al., 2007) and could be involved in piRNA biogenesis. However, two observations do not support this hypothesis. Firstly, northern blotting and quantitative RT-PCR analysis suggests that piRNA expression is independent of RdRPs (Table 3.1 and Figure 3.12). Secondly, RdRPs generate secondary siRNAs with a triphosphate at their 5′ end and this does not fit with the characteristics of piRNAs, which carry a 5′ monophosphate.

In *C. elegans*, 21U-RNAs/piRNAs map to both sense and antisense strands but rarely overlap with each other, suggesting that piRNAs are not generated from a dsRNA precursor (Ruby et al., 2006). Instead they may be generated from a single stranded primary transcript. Most of the 21U-RNA genomic loci share two upstream sequence motifs. The large motif is A-T rich, 34 bp long and contains an 8 nt core consensus sequence CTGTTTCA. The small motif has a YRNT core sequence, in which T corresponds to the 5′ U of the 21U RNA. The large and small motifs are separated by approximately 20 bp (Ruby et al., 2006). These upstream motifs are conserved between *C. elegans* and *C. briggsae*, but the 21U-RNA sequences themselves are not conserved (Ruby et al., 2006). The conserved upstream motifs may together serve as individual piRNA promoters to drive expression of each 21U-RNA or may act as a signal for targeted cleavage of a precursor RNA by an as yet unidentified exo- or endonuclease. Interestingly, piRNA loci in mice are also conserved among mammals, but individual piRNA sequences are not (Aravin et al., 2006; Girard et al., 2006). Having linked a conserved sequence motif to piRNAs in *C. elegans* it will be important to search for such motifs in other organisms and this might help our understanding of piRNA biogenesis. If 21U-RNAs act as individual transcription units, searching for transcription factors binding to this upstream motif would be of great interest.

It is equally mysterious how the 3′ end of *C. elegans* piRNAs is generated. Mature piRNAs may be generated by two cleavage events and then loaded into the Piwi complex or could be generated by initial 5′ end formation, loading into Piwi, followed by resection at their 3′ end by a 3′ exonuclease.

I have shown using northern blotting, quantitative RT-PCR and high-throughput sequencing that the expression of piRNAs is dramatically reduced in *piwi* mutants (Figure 3.6, 8 and Table 3.1). However, piRNAs are not completely absent in *piwi* mutants (Figure 3.8 and Figure 3.10), suggesting Piwi is not absolutely essential for piRNA biogenesis. This observation is similar to *alg-1; alg-2* mutants, in which miRNA expression is reduced but not completely absent (Grishok et al., 2001). These

data suggest that Piwi might not be an essential component of the piRNA biogenesis pathway, but that it is required for piRNA stability.

To search for more genes, which might be involved in piRNA biogenesis, I tested genes from small RNA pathways in *C. elegans* (Figure 3.12 & Table 3.1). It appears that piRNA biogenesis is independent of many other genes that are involved in small RNA pathways including *dcr-1*. Therefore, it will be interesting to look for new genes that are required for piRNA biogenesis. Genome-wide RNAi screening in *C. elegans* might identify these genes.

4.4 No evidence for ping-pong in *C. elegans*

Two groups have proposed a 'ping-pong' model for piRNA production in *Drosophila* (Brennecke et al., 2007; Gunawardane et al., 2007). In this model Piwi and Aub bind to antisense strand piRNAs, while Ago3 binds to sense strand piRNAs. Piwi- and Aub-interacting piRNAs have a 5′ uridine (U), while Ago3 interacting piRNAs often have adenine (A) at their 10th nucleotide. Some Ago3-interacting piRNAs show complementary to the first 10 nt of the Piwi/Aub-interacting piRNAs. Based on these observations, the 'ping-pong' model proposes that Ago3-interacting sense piRNAs mediate antisense strand cleavage at an A:U base pair that generates the 5′ end of the antisense piRNAs, which are associated with Piwi/Aub followed by 3′ end processing to generate mature antisense piRNAs. The mature antisense piRNAs together with Piwi/Aub then generate sense strand piRNAs. However, this model does not explain how the 3′ end of mature piRNAs is generated. Mutants for two *Drosophila* nucleases, *zucchini* and *squash*, showed a loss of piRNA production (Pane et al., 2007) indicating zucchini and squash may process the 3′ end of the precursor of mature piRNAs. Data from mice and zebrafish suggest a similar 'ping-pong' mechanism might exist in these organisms for a subset of piRNAs (Aravin et al., 2007b; Gunawardane et al., 2007).

To investigate whether a 'ping-pong' mechanism exists in *C. elegans*, high-throughput sequencing was performed of 5′ independent small RNA libraries prepared from wild-

type animals. Sequence analysis showed no 'A' bias at the tenth position of the endogenous small RNAs (Figure 3.31) and there was only rare overlap of piRNAs (data not shown). Therefore, I conclude that there is no exsistance of 'ping-pong' mechanism in *C. elegans* for the ampification of piRNAs. Mammals, fish and flies are lacking RdRPs, which mediate secondary siRNA production in fission yeast, plants and *C. elegans*. Therefore, mammals, fish and flies might have developed a 'ping-pong' mechanism for the amplification of piRNAs as an alternative to a secondary siRNA amplification pathway.

4.5 Piwi and piRNAs acts upstream of secondary siRNA pathways

I have demonstrated that Tc3 transposon silencing is associated with endogenous siRNAs that depend on Piwi and piRNAs. Mapping of piRNAs to transposons showed that only one piRNA (21UR-139) mapped to the Tc3 transposase and may act in *cis*. A parallel study also found another piRNA, 21UR-15703 that also maps to Tc3 transposons, but to the 3′ UTR of the transposase mRNA, just upstream of the TIR (terminal inverted repeats) (Batista et al., 2008). Both piRNAs map to the sense strand of the Tc3 transposon. Endogenous siRNAs map to the transposase and TIRs of the Tc3 transposon in wild-type, but are dramatically reduced in *piwi* mutants (Figure 3.25 & 3.29). A parallel study found a loss of endogenous siRNAs exclusively from the TIR but not from the transposase of Tc3 in *prg-1* mutants compared to wild-type (Batista et al., 2008). These results strongly suggest that PRG-1 and PRG-2 have different functions in Tc3 transposon silencing. To explain these data, a speculative model was proposed to explain how sense piRNAs might be required for antisense siRNA production at the Tc3 locus. Sense piRNAs loaded onto Piwi might stimulate RdRP activity using Tc3 antisense transcripts as a template, in a manner analogous to siRNAs in RNAi amplification. The sense transcripts would be targeted by antisense siRNAs, and might act as a template for a second round of RdRP activity (Figure 4.1). Such a loop would function to maintain high levels of Tc3 siRNAs. TIR-associated siRNAs might be functionally distinct from the siRNAs derived from the Tc3 transposase and they might act through two independent pathways. In the case of the Tc3 transposase,

siRNAs might regulate Tc3 transposons through PTGS. In addition, TIR siRNAs might regulate Tc3 through chromatin-mediated TGS. Both quantitative RT-PCR and transposition analysis show synergistic effects of PRG-1 and PRG-2 in the regulation of Tc3, which need to be investigated further. The amplification loop described here is independent of the sequence of the initiating 21U-RNAs. Therefore, the link between piRNAs and siRNAs might not be restricted to Tc3. Indeed, a positional bias was observed for antisense siRNAs downstream of piRNA loci (Figure 3.30A).
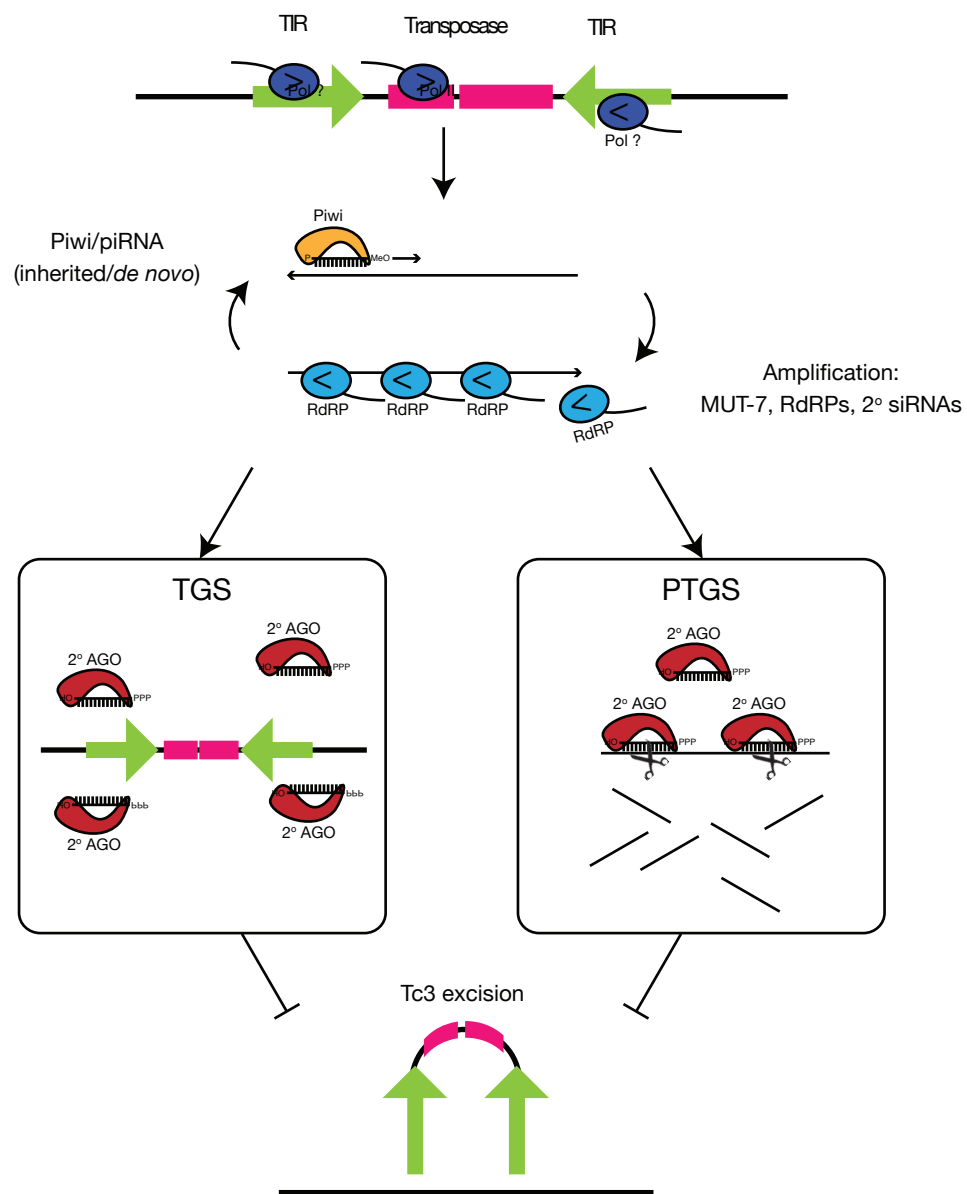
Figure 4.1 A speculative model of the role of Piwi in Tc3 silencing.
TGS, transcriptional gene silencing. PTGS, post-transcriptional gene silencing. 2⁰ AGO, secondary Argonautes. TIR, terminal inverted repeats. Model is described in discussion (section 4.5).

4.6 Potential roles of piRNAs *in cis* and *in trans*

In *C. elegans*, uniquely, 21U-RNAs/piRNAs are clustered to two broad regions of chromosome IV and few map elsewhere in the genome. These piRNAs are either present in intergenic regions or within introns. Therefore, piRNAs might regulate gene expression locally (*in cis*), or globally (*in trans*), or both. To test this hypothesis, first global gene expression profiling was performed comparing wild-type and *piwi (prg-1; prg-2)* mutants. No striking changes in gene expression were observed (Figure 3.20A & B). However, parellel studies in the Mello lab showed that expression of number of genes might be altered in *prg-1* mutants (Batista et al., 2008). This discrepancy may reflect either a genuine difference between *prg-1; prg-2* double mutants and *prg-1* single mutants or possibly due to differences in the gene expression profiling technologies used.

Next, I tested gene expression of few class I (genes with high number of piRNAs in their introns) and class II genes (genes with no or few piRNAs in their introns) by qRT-PCR from isolated gonads. No significant changes in these genes expression was found between wild-type and *piwi* mutants (Figure 3.21A & B). A parallel study from the Reinke lab, using dissected gonads from wild-type and *prg-1* mutant males show that some spermatogenesis-enriched genes on chromosome IV are downregulated, and many of the spermatogenesis-enriched genes on chromosome IV and genes located other than chromosome IV are unaffected in *prg-1* mutant males, including *msp* genes (Wang and Reinke, 2008). However, my studies demonstrated downregulation of *msp* genes in *piwi* mutant hermaphrodites (Figure 3.20B). This might be due to the low sperm counts (Figure 3.2), or analyzing gene expression from RNA prepared from whole animals instead of isolated gonads.

It might be possible that piRNAs regulate gene expression both in *cis* and *trans*. To explain this paradox, I speculate that PRG-1/piRNA complex may base pair imperfectly with its targets, like animal miRNAs on their target mRNAs. Deep sequencing studies from this work and the Mello lab proposed almost 15000 different piRNAs encoded in

C. elegans. If all of the piRNAs are involved in piRNA-mediated regulation through partial matches, then the entire transcriptome is probably under piRNA-directed regulation. Perhaps piRNAs collectively regulate gene expression globally through partial matches. Therefore, it remains to be answered how piRNAs regulate gene expression, whether in *cis,* or in *trans,* and what are the rules governing piRNAs and their target interactions.

4.7 A potential role for *C. elegans* piRNAs in transcriptional gene silencing?

To address the possible mechanism of Piwi function, their localization has been studied in different species. In *Drosophila*, Piwi is localized in the nucleus of germ cells and somatic cells, whereas Ago3 and Aubergine are mostly detected in the cytoplasm and in perinuclear germanium granules called the nuage (Brennecke et al., 2007; Cox et al., 2000; Gunawardane et al., 2007). A zebrafish Piwi protein, Ziwi, is also localized to nuage (Houwing et al., 2007). The mouse Piwi proteins Miwi, Mili and Miwi2 are also localized to the cytoplasm (Deng and Lin, 2002). It has also been shown that Miwi is a component of the chromatoid body (Kotaja et al., 2006). In rat, Riwi has been found associated with a RecQ helicase and piRNAs, suggesting that they may function at the DNA level (Lau et al., 2006). In *C. elegans*, both endogenous PRG-1 and PRG-1::GFP localized to the germline nuage equivalent, the P granules (Batista et al., 2008; Wang and Reinke, 2008) and this expression pattern is consistent with the localization of the Piwi family of proteins to germ granules in *Drosophila*, zebrafish and mammals. Similar to other organisms, piRNA expression also tightly correlated with the expression of the Piwi subfamily of Argonaute proteins, including PRG-1 (Figure 16 and 17).

Both transposition assays and quantitative RT-PCR analysis show that Piwi protein in *C. elegans* specifically suppresses Tc3 transposons (Figure 3.23 and Table 3.2). Endogenous siRNAs are dramatically reduced from the Tc3 transposase and TIR

regions in *piwi* mutants (Figure 3.25 & 3.29), whereas endogenous siRNAs are absent exclusively from the TIRs of the Tc3 transposon but not from the transposase in *prg-1* mutants (Batista et al., 2008). These data suggest that endogenous siRNAs might regulate Tc3 transposons through TGS or PTGS in a Piwi/piRNA dependent manner. In additon, I have shown that piRNAs expression was reduced dramatically in *hpl-1* mutants (Figure 3.15), suggesting there may be a connection between Piwi and HPL-1 (HP1) in TGS of Tc3 transposons. Interestingly, the reduced number of TIR siRNAs in *prg-1* mutants suggests that PRG-1 might regulate Tc3 at TIRs through chromatin-mediated TGS. Perhaps PRG-1 suppresses Tc3 mobility at TIRs along with TIR siRNAs and HP1, maintaining a heterochromatic state. Loss of PRG-1 or HPL-1 might lead to alteration of chromatin structure at TIRs and activation of Tc3. In *Drosophila* it has been shown that Piwi physically interacts with HP1 in an RNA-dependent manner and it has been suggested that this complex might regulate heterochromatin formation (Brower-Toland., 2007; Pal-Bhadra., 2004). In contrats, the Lin group has demonstrated that Piwi promotes the euchromatic character of 3R-TAS, a telomeric repeat (Yin and Lin, 2007). Ago proteins in yeast and plants are nuclear and cytoplasmic and play crucial roles in TGS through heterochromatin formation and/or DNA methylation at repetitive regions of the genome (Buhler and Moazed, 2007) (Lippman and Martienssen, 2004; Slotkin and Martienssen, 2007). In mice, *Mili* and *Miwi2* null mutants show derepression of LINE-1 and IAP retrotransposons in male germ cells (Kuramochi-Miyagawa et al., 2008). In contrast, immunohistochemisry and biochemical data show that Miwi, Mili and Miwi2 are localized to the cytoplasm. Therefore, it is possible that these proteins might shuttle between nucleus and cytoplasm, and are involved in both TGS and PTGS.

In *C. elegans*, PRG-1 localizes to the perinuclear region of germ cells. In order to understand the potential role of PRG-1 in TGS, further studies are needed. For example, it would be interesting to test whether HPL-1 and PRG-1 co-localize in the nucleus. If so, how does PRG-1 shuttle from the nucleus to the cytoplasm? Perhaps post-translational modification of PRG-1 regulates shuttling. In addition, it will be interesting to further investigate the mechanism of Tc3 transoposon silencing by Piwi

proteins in *C. elegans*. Finally, it would be important to determine the function of Piwi and piRNAs in the control of germline development.

5. References

Ambros, V., Lee, R.C., Lavanway, A., Williams, P.T., and Jewell, D. (2003). MicroRNAs and other tiny endogenous RNAs in C. elegans. Curr Biol *13*, 807-818.

Aravin, A., Gaidatzis, D., Pfeffer, S., Lagos-Quintana, M., Landgraf, P., Iovino, N., Morris, P., Brownstein, M.J., Kuramochi-Miyagawa, S., Nakano, T.*, et al.* (2006). A novel class of small RNAs bind to MILI protein in mouse testes. Nature *442*, 203-207.

Aravin, A.A., Hannon, G.J., and Brennecke, J. (2007a). The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race. Science *318*, 761-764.

Aravin, A.A., Naumova, N.M., Tulin, A.V., Vagin, V.V., Rozovsky, Y.M., and Gvozdev, V.A. (2001). Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the D. melanogaster germline. Curr Biol *11*, 1017-1027.

Aravin, A.A., Sachidanandam, R., Girard, A., Fejes-Toth, K., and Hannon, G.J. (2007b). Developmentally regulated piRNA clusters implicate MILI in transposon control. Science *316*, 744-747.

Austin, J., and Kimble, J. (1987). glp-1 is required in the germ line for regulation of the decision between mitosis and meiosis in C. elegans. Cell *51*, 589-599.

Baek, D., Villen, J., Shin, C., Camargo, F.D., Gygi, S.P., and Bartel, D.P. (2008). The impact of microRNAs on protein output. Nature.

Bannister, A.J., Zegerman, P., Partridge, J.F., Miska, E.A., Thomas, J.O., Allshire, R.C., and Kouzarides, T. (2001). Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. Nature *410*, 120-124.

Bartel, D.P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. Cell *116*, 281-297.

Barton, M.K., Schedl, T.B., and Kimble, J. (1987). Gain-of-function mutations of fem-3, a sex-determination gene in Caenorhabditis elegans. Genetics *115*, 107-119.

Batista, P.J., Ruby, J.G., Claycomb, J.M., Chiang, R., Fahlgren, N., Kasschau, K.D., Chaves, D.A., Gu, W., Vasale, J.J., Duan, S.*, et al.* (2008). PRG-1 and 21U-RNAs interact to form the piRNA complex required for fertility in C. elegans. Mol Cell *31*, 67-78.

Beanan, M.J., and Strome, S. (1992). Characterization of a germ-line proliferation mutation in C. elegans. Development *116*, 755-766.

Behm-Ansmant, I., Rehwinkel, J., Doerks, T., Stark, A., Bork, P., and Izaurralde, E. (2006). mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes. Genes Dev *20*, 1885-1898.

Bentwich, I. (2005). Prediction and validation of microRNAs and their targets. FEBS Lett *579*, 5904-5910.

Betel, D., Sheridan, R., Marks, D.S., and Sander, C. (2007). Computational analysis of mouse piRNA sequence and biogenesis. PLoS Comput Biol *3*, e222.

Bhattacharyya, S.N., Habermacher, R., Martine, U., Closs, E.I., and Filipowicz, W. (2006). Stress-induced reversal of microRNA repression and mRNA P-body localization in human cells. Cold Spring Harb Symp Quant Biol *71*, 513-521.

Bohmert, K., Camus, I., Bellini, C., Bouchez, D., Caboche, M., and Benning, C. (1998). AGO1 defines a novel locus of Arabidopsis controlling leaf development. Embo J *17*, 170-180.

Bohnsack, M.T., Czaplinski, K., and Gorlich, D. (2004). Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. Rna *10*, 185-191.

Borchert, G.M., Lanier, W., and Davidson, B.L. (2006). RNA polymerase III transcribes human microRNAs. Nat Struct Mol Biol *13*, 1097-1101.

Bourc'his, D., and Bestor, T.H. (2004). Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. Nature *431*, 96-99.

Brengues, M., Teixeira, D., and Parker, R. (2005). Movement of eukaryotic mRNAs between polysomes and cytoplasmic processing bodies. Science *310*, 486-489.

Brennecke, J., Aravin, A.A., Stark, A., Dus, M., Kellis, M., Sachidanandam, R., and Hannon, G.J. (2007). Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. Cell *128*, 1089-1103.

Brenner, S. (1974). The genetics of Caenorhabditis elegans. Genetics *77*, 71-94.

Brower-Toland, B., Findley, S.D., Jiang, L., Liu, L., Yin, H., Dus, M., Zhou, P., Elgin, S.C., and Lin, H. (2007). Drosophila PIWI associates with chromatin and interacts directly with HP1a. Genes Dev *21*, 2300-2311.

Buhler, M., and Moazed, D. (2007). Transcription and RNAi in heterochromatic gene silencing. Nat Struct Mol Biol *14*, 1041-1048.

Carmell, M.A., Girard, A., van de Kant, H.J., Bourc'his, D., Bestor, T.H., de Rooij, D.G., and Hannon, G.J. (2007). MIWI2 is essential for spermatogenesis and repression of transposons in the mouse male germline. Dev Cell *12*, 503-514.

Carmell, M.A., Xuan, Z., Zhang, M.Q., and Hannon, G.J. (2002). The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis. Genes Dev *16*, 2733-2742.

Chen, P.Y., Manninga, H., Slanchev, K., Chien, M., Russo, J.J., Ju, J., Sheridan, R., John, B., Marks, D.S., Gaidatzis, D*., et al.* (2005). The developmental miRNA profiles of zebrafish as determined by small RNA cloning. Genes Dev *19*, 1288-1293.

Chendrimada, T.P., Finn, K.J., Ji, X., Baillat, D., Gregory, R.I., Liebhaber, S.A., Pasquinelli, A.E., and Shiekhattar, R. (2007). MicroRNA silencing through RISC recruitment of eIF6. Nature *447*, 823-828.

Colmenares, S.U., Buker, S.M., Buhler, M., Dlakic, M., and Moazed, D. (2007). Coupling of double-stranded RNA synthesis and siRNA generation in fission yeast RNAi. Mol Cell *27*, 449-461.

Cook, H.A., Koppetsch, B.S., Wu, J., and Theurkauf, W.E. (2004). The Drosophila SDE3 homolog armitage is required for oskar mRNA silencing and embryonic axis specification. Cell *116*, 817-829.

Cox, D.N., Chao, A., Baker, J., Chang, L., Qiao, D., and Lin, H. (1998). A novel class of evolutionarily conserved genes defined by piwi are essential for stem cell self-renewal. Genes Dev *12*, 3715-3727.

Cox, D.N., Chao, A., and Lin, H. (2000). piwi encodes a nucleoplasmic factor whose activity modulates the number and division rate of germline stem cells. Development *127*, 503-514.

Czech, B., Malone, C.D., Zhou, R., Stark, A., Schlingeheyde, C., Dus, M., Perrimon, N., Kellis, M., Wohlschlegel, J.A., Sachidanandam, R*., et al.* (2008). An endogenous small interfering RNA pathway in Drosophila. Nature *453*, 798-802.

Deininger, P.L., Moran, J.V., Batzer, M.A., and Kazazian, H.H., Jr. (2003). Mobile elements and mammalian genome evolution. Curr Opin Genet Dev *13*, 651-658.

Deng, W., and Lin, H. (2002). miwi, a murine homolog of piwi, encodes a cytoplasmic protein essential for spermatogenesis. Dev Cell *2*, 819-830.

Denli, A.M., Tops, B.B., Plasterk, R.H., Ketting, R.F., and Hannon, G.J. (2004). Processing of primary microRNAs by the Microprocessor complex. Nature *432*, 231-235.

Duchaine, T.F., Wohlschlegel, J.A., Kennedy, S., Bei, Y., Conte, D., Jr., Pang, K., Brownell, D.R., Harding, S., Mitani, S., Ruvkun, G*., et al.* (2006). Functional proteomics reveals the biochemical niche of C. elegans DCR-1 in multiple small-RNA-mediated pathways. Cell *124*, 343-354.

Dunoyer, P., Himber, C., Ruiz-Ferrer, V., Alioua, A., and Voinnet, O. (2007). Intra- and intercellular RNA interference in Arabidopsis thaliana requires components of the microRNA and heterochromatic silencing pathways. Nat Genet *39*, 848-856.

Easow, G., Teleman, A.A., and Cohen, S.M. (2007). Isolation of microRNA targets by miRNP immunopurification. Rna *13*, 1198-1204.

Ebbs, M.L., and Bender, J. (2006). Locus-specific control of DNA methylation by the Arabidopsis SUVH5 histone methyltransferase. Plant Cell *18*, 1166-1176.

Eckmann, C.R., Crittenden, S.L., Suh, N., and Kimble, J. (2004). GLD-3 and control of the mitosis/meiosis decision in the germline of Caenorhabditis elegans. Genetics *168*, 147-160.

Eulalio, A., Huntzinger, E., and Izaurralde, E. (2008). GW182 interaction with Argonaute is essential for miRNA-mediated translational repression and mRNA decay. Nat Struct Mol Biol *15*, 346-353.

Farazi, T.A., Juranek, S.A., and Tuschl, T. (2008). The growing catalog of small RNAs and their association with distinct Argonaute/Piwi family members. Development *135*, 1201-1214.

Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E., and Mello, C.C. (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature *391*, 806-811.

Fischer, S.E., Wienholds, E., and Plasterk, R.H. (2003). Continuous exchange of sequence information between dispersed Tc1 transposons in the Caenorhabditis elegans genome. Genetics *164*, 127-134.

Flavell, R.B., Bennett, M.D., Smith, J.B., and Smith, D.B. (1974). Genome size and the proportion of repeated nucleotide sequence DNA in plants. Biochem Genet *12*, 257-269.

Forstemann, K., Horwich, M.D., Wee, L., Tomari, Y., and Zamore, P.D. (2007). Drosophila microRNAs are sorted into functionally distinct argonaute complexes after production by dicer-1. Cell *130*, 287-297.

Ghildiyal, M., Seitz, H., Horwich, M.D., Li, C., Du, T., Lee, S., Xu, J., Kittler, E.L., Zapp, M.L., Weng, Z., and Zamore, P.D. (2008). Endogenous siRNAs derived from transposons and mRNAs in Drosophila somatic cells. Science *320*, 1077-1081.

Giraldez, A.J., Mishima, Y., Rihel, J., Grocock, R.J., Van Dongen, S., Inoue, K., Enright, A.J., and Schier, A.F. (2006). Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs. Science *312*, 75-79.

Girard, A., Sachidanandam, R., Hannon, G.J., and Carmell, M.A. (2006). A germline-specific class of small RNAs binds mammalian Piwi proteins. Nature *442*, 199-202.

Gregory, R.I., Yan, K.P., Amuthan, G., Chendrimada, T., Doratotaj, B., Cooch, N., and Shiekhattar, R. (2004). The Microprocessor complex mediates the genesis of microRNAs. Nature *432*, 235-240.

Grishok, A., Pasquinelli, A.E., Conte, D., Li, N., Parrish, S., Ha, I., Baillie, D.L., Fire, A., Ruvkun, G., and Mello, C.C. (2001). Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control C. elegans developmental timing. Cell *106*, 23-34.

Grivna, S.T., Beyret, E., Wang, Z., and Lin, H. (2006a). A novel class of small RNAs in mouse spermatogenic cells. Genes Dev *20*, 1709-1714.

Grivna, S.T., Pyhtila, B., and Lin, H. (2006b). MIWI associates with translational machinery and PIWI-interacting RNAs (piRNAs) in regulating spermatogenesis. Proc Natl Acad Sci U S A *103*, 13415-13420.

Gruenbaum, Y., Stein, R., Cedar, H., and Razin, A. (1981). Methylation of CpG sequences in eukaryotic DNA. FEBS Lett *124*, 67-71.

Gunawardane, L.S., Saito, K., Nishida, K.M., Miyoshi, K., Kawamura, Y., Nagami, T., Siomi, H., and Siomi, M.C. (2007). A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in Drosophila. Science *315*, 1587-1590.

Hansen, D., Wilson-Berry, L., Dang, T., and Schedl, T. (2004). Control of the proliferation versus meiotic development decision in the C. elegans germline through regulation of GLD-1 protein accumulation. Development *131*, 93-104.

Hendrickson, D.G., Hogan, D.J., Herschlag, D., Ferrell, J.E., and Brown, P.O. (2008). Systematic identification of mRNAs recruited to argonaute 2 by specific microRNAs and corresponding changes in transcript abundance. PLoS ONE *3*, e2126.

Hofacker, I.L. (2007). How microRNAs choose their targets. Nat Genet *39*, 1191-1192.

Horwich, M.D., Li, C., Matranga, C., Vagin, V., Farley, G., Wang, P., and Zamore, P.D. (2007). The Drosophila RNA methyltransferase, DmHen1, modifies germline piRNAs and single-stranded siRNAs in RISC. Curr Biol *17*, 1265-1272.

Houwing, S., Kamminga, L.M., Berezikov, E., Cronembold, D., Girard, A., van den Elst, H., Filippov, D.V., Blaser, H., Raz, E., Moens, C.B.*, et al.* (2007). A role for Piwi and piRNAs in germ cell maintenance and transposon silencing in Zebrafish. Cell *129*, 69-82.

Humphreys, D.T., Westman, B.J., Martin, D.I., and Preiss, T. (2005). MicroRNAs control translation initiation by inhibiting eukaryotic initiation factor 4E/cap and poly(A) tail function. Proc Natl Acad Sci U S A *102*, 16961-16966.

Jeddeloh, J.A., Bender, J., and Richards, E.J. (1998). The DNA methylation locus DDM1 is required for maintenance of gene silencing in Arabidopsis. Genes Dev *12*, 1714-1725.

Kadyk, L.C., and Kimble, J. (1998). Genetic regulation of entry into meiosis in Caenorhabditis elegans. Development *125*, 1803-1813.

Kajikawa, M., and Okada, N. (2002). LINEs mobilize SINEs in the eel through a shared 3' sequence. Cell *111*, 433-444.

Kapitonov, V.V., and Jurka, J. (2007). Helitrons on a roll: eukaryotic rolling-circle transposons. Trends Genet *23*, 521-529.

Kato, Y., Kaneda, M., Hata, K., Kumaki, K., Hisano, M., Kohara, Y., Okano, M., Li, E., Nozaki, M., and Sasaki, H. (2007). Role of the Dnmt3 family in de novo methylation of imprinted and repetitive sequences during male germ cell development in the mouse. Hum Mol Genet *16*, 2272-2280.

Kawamura, Y., Saito, K., Kin, T., Ono, Y., Asai, K., Sunohara, T., Okada, T.N., Siomi, M.C., and Siomi, H. (2008). Drosophila endogenous small RNAs bind to Argonaute 2 in somatic cells. Nature *453*, 793-797.

Kazazian, H.H., Jr. (2004). Mobile elements: drivers of genome evolution. Science *303*, 1626-1632.

Ketting, R.F., Fischer, S.E., Bernstein, E., Sijen, T., Hannon, G.J., and Plasterk, R.H. (2001). Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans. Genes Dev *15*, 2654-2659.

Ketting, R.F., Haverkamp, T.H., van Luenen, H.G., and Plasterk, R.H. (1999). Mut-7 of C. elegans, required for transposon silencing and RNA interference, is a homolog of Werner syndrome helicase and RNaseD. Cell *99*, 133-141.

Kim, J., Krichevsky, A., Grad, Y., Hayes, G.D., Kosik, K.S., Church, G.M., and Ruvkun, G. (2004). Identification of many microRNAs that copurify with polyribosomes in mammalian neurons. Proc Natl Acad Sci U S A *101*, 360-365.

Kim, J.M., Vanguri, S., Boeke, J.D., Gabriel, A., and Voytas, D.F. (1998). Transposable elements and genome organization: a comprehensive survey of retrotransposons revealed by the complete Saccharomyces cerevisiae genome sequence. Genome Res *8*, 464-478.

Kim, V.N. (2006). Small RNAs just got bigger: Piwi-interacting RNAs (piRNAs) in mammalian testes. Genes Dev *20*, 1993-1997.

Kimble, J., Edgar, L., and Hirsh, D. (1984). Specification of male development in Caenorhabditis elegans: the fem genes. Dev Biol *105*, 234-239.

Kimble, J.E., and White, J.G. (1981). On the control of germ cell development in Caenorhabditis elegans. Dev Biol *81*, 208-219.

Kiriakidou, M., Tan, G.S., Lamprinaki, S., De Planell-Saguer, M., Nelson, P.T., and Mourelatos, Z. (2007). An mRNA m7G cap binding-like motif within human Ago2 represses translation. Cell *129*, 1141-1151.

Kirino, Y., and Mourelatos, Z. (2007). Mouse Piwi-interacting RNAs are 2'-O-methylated at their 3' termini. Nat Struct Mol Biol *14*, 347-348.

Klattenhoff, C., and Theurkauf, W. (2008). Biogenesis and germline functions of piRNAs. Development *135*, 3-9.

Kotaja, N., Lin, H., Parvinen, M., and Sassone-Corsi, P. (2006). Interplay of PIWI/Argonaute protein MIWI and kinesin KIF17b in chromatoid bodies of male germ cells. J Cell Sci *119*, 2819-2825.

Kuramochi-Miyagawa, S., Kimura, T., Ijiri, T.W., Isobe, T., Asada, N., Fujita, Y., Ikawa, M., Iwai, N., Okabe, M., Deng, W.*, et al.* (2004). Mili, a mammalian member of piwi family gene, is essential for spermatogenesis. Development *131*, 839-849.

Kuramochi-Miyagawa, S., Kimura, T., Yomogida, K., Kuroiwa, A., Tadokoro, Y., Fujita, Y., Sato, M., Matsuda, Y., and Nakano, T. (2001). Two mouse piwi-related genes: miwi and mili. Mech Dev *108*, 121-133.

Kuramochi-Miyagawa, S., Watanabe, T., Gotoh, K., Totoki, Y., Toyoda, A., Ikawa, M., Asada, N., Kojima, K., Yamaguchi, Y., Ijiri, T.W.*, et al.* (2008). DNA methylation of retrotransposon genes is regulated by Piwi family members MILI and MIWI2 in murine fetal testes. Genes Dev *22*, 908-917.

Lambie, E.J., and Kimble, J. (1991). Two homologous regulatory genes, lin-12 and glp-1, have overlapping functions. Development *112*, 231-240.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W.*, et al.* (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860-921.

Lau, N.C., Seto, A.G., Kim, J., Kuramochi-Miyagawa, S., Nakano, T., Bartel, D.P., and Kingston, R.E. (2006). Characterization of the piRNA complex from rat testes. Science *313*, 363-367.

Lee, R.C., and Ambros, V. (2001). An extensive class of small RNAs in Caenorhabditis elegans. Science *294*, 862-864.

Lee, R.C., Feinbaum, R.L., and Ambros, V. (1993). The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell *75*, 843-854.

Lee, Y., Kim, M., Han, J., Yeom, K.H., Lee, S., Baek, S.H., and Kim, V.N. (2004). MicroRNA genes are transcribed by RNA polymerase II. Embo J *23*, 4051-4060.

Lehnertz, B., Ueda, Y., Derijck, A.A., Braunschweig, U., Perez-Burgos, L., Kubicek, S., Chen, T., Li, E., Jenuwein, T., and Peters, A.H. (2003). Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. Curr Biol *13*, 1192-1200.

Lewis, B.P., Burge, C.B., and Bartel, D.P. (2005). Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell *120*, 15-20.

Li, C.F., Pontes, O., El-Shami, M., Henderson, I.R., Bernatavichute, Y.V., Chan, S.W., Lagrange, T., Pikaard, C.S., and Jacobsen, S.E. (2006). An ARGONAUTE4-containing nuclear processing center colocalized with Cajal bodies in Arabidopsis thaliana. Cell *126*, 93-106.

Li, J., Yang, Z., Yu, B., Liu, J., and Chen, X. (2005). Methylation protects miRNAs and siRNAs from a 3'-end uridylation activity in Arabidopsis. Curr Biol *15*, 1501-1507.

Lin, H., and Spradling, A.C. (1997). A novel group of pumilio mutations affects the asymmetric division of germline stem cells in the Drosophila ovary. Development *124*, 2463-2476.

Lingel, A., Simon, B., Izaurralde, E., and Sattler, M. (2003). Structure and nucleic-acid binding of the Drosophila Argonaute 2 PAZ domain. Nature *426*, 465-469.

Lippman, Z., Gendrel, A.V., Black, M., Vaughn, M.W., Dedhia, N., McCombie, W.R., Lavine, K., Mittal, V., May, B., Kasschau, K.D.*, et al.* (2004). Role of transposable elements in heterochromatin and epigenetic control. Nature *430*, 471-476.

Lippman, Z., and Martienssen, R. (2004). The role of RNA interference in heterochromatic silencing. Nature *431*, 364-370.

Liu, J., Carmell, M.A., Rivas, F.V., Marsden, C.G., Thomson, J.M., Song, J.J., Hammond, S.M., Joshua-Tor, L., and Hannon, G.J. (2004). Argonaute2 is the catalytic engine of mammalian RNAi. Science *305*, 1437-1441.

Ma, J.B., Ye, K., and Patel, D.J. (2004). Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain. Nature *429*, 318-322.

Ma, J.B., Yuan, Y.R., Meister, G., Pei, Y., Tuschl, T., and Patel, D.J. (2005). Structural basis for 5'-end-specific recognition of guide RNA by the A. fulgidus Piwi protein. Nature *434*, 666-670.

Maroney, P.A., Yu, Y., Fisher, J., and Nilsen, T.W. (2006). Evidence that microRNAs are associated with translating messenger RNAs in human cells. Nat Struct Mol Biol *13*, 1102-1107.

Martens, J.H., O'Sullivan, R.J., Braunschweig, U., Opravil, S., Radolf, M., Steinlein, P., and Jenuwein, T. (2005). The profile of repeat-associated histone lysine methylation states in the mouse epigenome. Embo J *24*, 800-812.

Matranga, C., Tomari, Y., Shin, C., Bartel, D.P., and Zamore, P.D. (2005). Passenger-strand cleavage facilitates assembly of siRNA into Ago2-containing RNAi enzyme complexes. Cell *123*, 607-620.

Medstrand, P., and Mager, D.L. (1998). Human-specific integrations of the HERV-K endogenous retrovirus family. J Virol *72*, 9782-9787.

Meister, G., Landthaler, M., Patkaniowska, A., Dorsett, Y., Teng, G., and Tuschl, T. (2004). Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. Mol Cell *15*, 185-197.

Meister, G., and Tuschl, T. (2004). Mechanisms of gene silencing by double-stranded RNA. Nature *431*, 343-349.

Mi, S., Cai, T., Hu, Y., Chen, Y., Hodges, E., Ni, F., Wu, L., Li, S., Zhou, H., Long, C.*, et al.* (2008). Sorting of small RNAs into Arabidopsis argonaute complexes is directed by the 5' terminal nucleotide. Cell *133*, 116-127.

Miska, E.A., Alvarez-Saavedra, E., Abbott, A.L., Lau, N.C., Hellman, A.B., McGonagle, S.M., Bartel, D.P., Ambros, V.R., and Horvitz, H.R. (2007). Most Caenorhabditis elegans microRNAs are individually not essential for development or viability. PLoS Genet *3*, e215.

Miyoshi, K., Tsukumo, H., Nagami, T., Siomi, H., and Siomi, M.C. (2005). Slicer function of Drosophila Argonautes and its involvement in RISC formation. Genes Dev *19*, 2837-2848.

Mochizuki, K., Fine, N.A., Fujisawa, T., and Gorovsky, M.A. (2002). Analysis of a piwi-related gene implicates small RNAs in genome rearrangement in tetrahymena. Cell *110*, 689-699.

Motamedi, M.R., Verdel, A., Colmenares, S.U., Gerber, S.A., Gygi, S.P., and Moazed, D. (2004). Two RNAi complexes, RITS and RDRC, physically interact and localize to noncoding centromeric RNAs. Cell *119*, 789-802.

Nelson, P.T., Hatzigeorgiou, A.G., and Mourelatos, Z. (2004). miRNP:mRNA association in polyribosomes in a human neuronal cell line. Rna *10*, 387-394.

Nottrott, S., Simard, M.J., and Richter, J.D. (2006). Human let-7a miRNA blocks protein production on actively translating polyribosomes. Nat Struct Mol Biol *13*, 1108-1114.

Nykanen, A., Haley, B., and Zamore, P.D. (2001). ATP requirements and small interfering RNA structure in the RNA interference pathway. Cell *107*, 309-321.

O'Donnell, K.A., and Boeke, J.D. (2007). Mighty Piwis defend the germline against genome intruders. Cell *129*, 37-44.

O'Donnell, K.A., Wentzel, E.A., Zeller, K.I., Dang, C.V., and Mendell, J.T. (2005). c-Myc-regulated microRNAs modulate E2F1 expression. Nature *435*, 839-843.

Ohara, T., Sakaguchi, Y., Suzuki, T., Ueda, H., Miyauchi, K., and Suzuki, T. (2007). The 3' termini of mouse Piwi-interacting RNAs are 2'-O-methylated. Nat Struct Mol Biol *14*, 349-350.

Okamura, K., Chung, W.J., Ruby, J.G., Guo, H., Bartel, D.P., and Lai, E.C. (2008). The Drosophila hairpin RNA pathway generates endogenous short interfering RNAs. Nature *453*, 803-806.

Olsen, P.H., and Ambros, V. (1999). The lin-4 regulatory RNA controls developmental timing in Caenorhabditis elegans by blocking LIN-14 protein synthesis after the initiation of translation. Dev Biol *216*, 671-680.

Pak, J., and Fire, A. (2007). Distinct populations of primary and secondary effectors during RNAi in C. elegans. Science *315*, 241-244.

Pal-Bhadra, M., Leibovitch, B.A., Gandhi, S.G., Rao, M., Bhadra, U., Birchler, J.A., and Elgin, S.C. (2004). Heterochromatic silencing and HP1 localization in Drosophila are dependent on the RNAi machinery. Science *303*, 669-672.

Pall, G.S., and Hamilton, A.J. (2008). Improved northern blot method for enhanced detection of small RNA. Nat Protoc *3*, 1077-1084.

Pane, A., Wehr, K., and Schupbach, T. (2007). zucchini and squash encode two putative nucleases required for rasiRNA production in the Drosophila germline. Dev Cell *12*, 851-862.

Parker, J.S., Roe, S.M., and Barford, D. (2004). Crystal structure of a PIWI protein suggests mechanisms for siRNA recognition and slicer activity. Embo J *23*, 4727-4737.

Parker, J.S., Roe, S.M., and Barford, D. (2005). Structural insights into mRNA recognition from a PIWI domain-siRNA guide complex. Nature *434*, 663-666.

Pelisson, A., Song, S.U., Prud'homme, N., Smith, P.A., Bucheton, A., and Corces, V.G. (1994). Gypsy transposition correlates with the production of a retroviral envelope-like protein under the tissue-specific control of the Drosophila flamenco gene. Embo J *13*, 4401-4411.

Pepper, A.S., Killian, D.J., and Hubbard, E.J. (2003). Genetic analysis of Caenorhabditis elegans glp-1 mutants suggests receptor interaction or competition. Genetics *163*, 115-132.

Peters, L., and Meister, G. (2007). Argonaute proteins: mediators of RNA silencing. Mol Cell *26*, 611-623.

Petersen, C.P., Bordeleau, M.E., Pelletier, J., and Sharp, P.A. (2006). Short RNAs repress translation after initiation in mammalian cells. Mol Cell *21*, 533-542.

Pillai, R.S., Bhattacharyya, S.N., Artus, C.G., Zoller, T., Cougot, N., Basyuk, E., Bertrand, E., and Filipowicz, W. (2005). Inhibition of translational initiation by Let-7 MicroRNA in human cells. Science *309*, 1573-1576.

Pontier, D., Yahubyan, G., Vega, D., Bulski, A., Saez-Vasquez, J., Hakimi, M.A., Lerbs-Mache, S., Colot, V., and Lagrange, T. (2005). Reinforcement of silencing at transposons and highly repeated sequences requires the concerted action of two distinct RNA polymerases IV in Arabidopsis. Genes Dev *19*, 2030-2040.

Qi, Y., He, X., Wang, X.J., Kohany, O., Jurka, J., and Hannon, G.J. (2006). Distinct catalytic and non-catalytic roles of ARGONAUTE4 in RNA-directed DNA methylation. Nature *443*, 1008-1012.

Qiao, D., Zeeman, A.M., Deng, W., Looijenga, L.H., and Lin, H. (2002). Molecular characterization of hiwi, a human member of the piwi gene family whose overexpression is correlated to seminomas. Oncogene *21*, 3988-3999.

Rajewsky, N. (2006). microRNA target predictions in animals. Nat Genet *38 Suppl*, S8-13.

Reddien, P.W., Oviedo, N.J., Jennings, J.R., Jenkin, J.C., and Sanchez Alvarado, A. (2005). SMEDWI-2 is a PIWI-like protein that regulates planarian stem cells. Science *310*, 1327-1330.

Rosenzweig, B., Liao, L.W., and Hirsh, D. (1983). Sequence of the C. elegans transposable element Tc1. Nucleic Acids Res *11*, 4201-4209.

Ruby, J.G., Jan, C., Player, C., Axtell, M.J., Lee, W., Nusbaum, C., Ge, H., and Bartel, D.P. (2006). Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in C. elegans. Cell *127*, 1193-1207.

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

Saito, K., Sakaguchi, Y., Suzuki, T., Suzuki, T., Siomi, H., and Siomi, M.C. (2007). Pimet, the Drosophila homolog of HEN1, mediates 2'-O-methylation of Piwi-interacting RNAs at their 3' ends. Genes Dev *21*, 1603-1608.

Savitsky, M., Kwon, D., Georgiev, P., Kalmykova, A., and Gvozdev, V. (2006). Telomere elongation is under the control of the RNAi-based mechanism in the Drosophila germline. Genes Dev *20*, 345-354.

Selbach, M., Schwanhausser, B., Thierfelder, N., Fang, Z., Khanin, R., and Rajewsky, N. (2008). Widespread changes in protein synthesis induced by microRNAs. Nature.

Seo, T.S., Bai, X., Ruparel, H., Li, Z., Turro, N.J., and Ju, J. (2004). Photocleavable fluorescent nucleotides for DNA sequencing on a chip constructed by site-specific coupling chemistry. Proc Natl Acad Sci U S A *101*, 5488-5493.

Sethupathy, P., Megraw, M., and Hatzigeorgiou, A.G. (2006). A guide through present computational approaches for the identification of mammalian microRNA targets. Nat Methods *3*, 881-886.

Seydoux, G., Mello, C.C., Pettitt, J., Wood, W.B., Priess, J.R., and Fire, A. (1996). Repression of gene expression in the embryonic germ lineage of C. elegans. Nature *382*, 713-716.

Sijen, T., Fleenor, J., Simmer, F., Thijssen, K.L., Parrish, S., Timmons, L., Plasterk, R.H., and Fire, A. (2001). On the role of RNA amplification in dsRNA-triggered gene silencing. Cell *107*, 465-476.

Sijen, T., and Plasterk, R.H. (2003). Transposon silencing in the Caenorhabditis elegans germ line by natural RNAi. Nature *426*, 310-314.

Sijen, T., Steiner, F.A., Thijssen, K.L., and Plasterk, R.H. (2007). Secondary siRNAs result from unprimed RNA synthesis and form a distinct class. Science *315*, 244-247.

Slotkin, R.K., and Martienssen, R. (2007). Transposable elements and the epigenetic regulation of the genome. Nat Rev Genet *8*, 272-285.

Smardon, A., Spoerke, J.M., Stacey, S.C., Klein, M.E., Mackin, N., and Maine, E.M. (2000). EGO-1 is related to RNA-directed RNA polymerase and functions in germ-line development and RNA interference in C. elegans. Curr Biol *10*, 169-178.

Song, J.J., Liu, J., Tolia, N.H., Schneiderman, J., Smith, S.K., Martienssen, R.A., Hannon, G.J., and Joshua-Tor, L. (2003). The crystal structure of the Argonaute2 PAZ domain reveals an RNA binding motif in RNAi effector complexes. Nat Struct Biol *10*, 1026-1032.

Song, J.J., Smith, S.K., Hannon, G.J., and Joshua-Tor, L. (2004). Crystal structure of Argonaute and its implications for RISC slicer activity. Science *305*, 1434-1437.

Sulston, J.E., Schierenberg, E., White, J.G., and Thomson, J.N. (1983). The embryonic cell lineage of the nematode Caenorhabditis elegans. Dev Biol *100*, 64-119.

Tabara, H., Sarkissian, M., Kelly, W.G., Fleenor, J., Grishok, A., Timmons, L., Fire, A., and Mello, C.C. (1999). The rde-1 gene, RNA interference, and transposon silencing in C. elegans. Cell *99*, 123-132.

Tam, O.H., Aravin, A.A., Stein, P., Girard, A., Murchison, E.P., Cheloufi, S., Hodges, E., Anger, M., Sachidanandam, R., Schultz, R.M., and Hannon, G.J. (2008). Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. Nature *453*, 534-538.

Teixeira, D., Sheth, U., Valencia-Sanchez, M.A., Brengues, M., and Parker, R. (2005). Processing bodies require RNA for assembly and contain nontranslating mRNAs. Rna *11*, 371-382.

Thermann, R., and Hentze, M.W. (2007). Drosophila miR2 induces pseudo-polysomes and inhibits translation initiation. Nature *447*, 875-878.

Tolia, N.H., and Joshua-Tor, L. (2007). Slicer and the argonautes. Nat Chem Biol *3*, 36-43.

Tomari, Y., Du, T., and Zamore, P.D. (2007). Sorting of Drosophila small silencing RNAs. Cell *130*, 299-308.

Unhavaithaya, Y., Shin, T.H., Miliaras, N., Lee, J., Oyama, T., and Mello, C.C. (2002). MEP-1 and a homolog of the NURD complex component Mi-2 act together to maintain germline-soma distinctions in C. elegans. Cell *111*, 991-1002.

Vagin, V.V., Sigova, A., Li, C., Seitz, H., Gvozdev, V., and Zamore, P.D. (2006). A distinct small RNA pathway silences selfish genetic elements in the germline. Science *313*, 320-324.

van Luenen, H.G., Colloms, S.D., and Plasterk, R.H. (1994). The mechanism of transposition of Tc3 in C. elegans. Cell *79*, 293-301.

Volpe, T.A., Kidner, C., Hall, I.M., Teng, G., Grewal, S.I., and Martienssen, R.A. (2002). Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. Science *297*, 1833-1837.

Vos, J.C., De Baere, I., and Plasterk, R.H. (1996). Transposase is the only nematode protein required for in vitro transposition of Tc1. Genes Dev *10*, 755-761.

Vos, J.C., and Plasterk, R.H. (1994). Tc1 transposase of Caenorhabditis elegans is an endonuclease with a bipartite DNA binding domain. Embo J *13*, 6125-6132.

Vos, J.C., van Luenen, H.G., and Plasterk, R.H. (1993). Characterization of the Caenorhabditis elegans Tc1 transposase in vivo and in vitro. Genes Dev *7*, 1244-1253.

W. B. Wood and the Community of *C. elegans* Researchers (1988). The Nematode *Caenorhabditis elegans* (Cold Spring Harbor, New York: Cold Spring Harbor Press).

Walsh, C.P., Chaillet, J.R., and Bestor, T.H. (1998). Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. Nat Genet *20*, 116-117.

Wang, B., Love, T.M., Call, M.E., Doench, J.G., and Novina, C.D. (2006). Recapitulation of short RNA-directed translational gene silencing in vitro. Mol Cell *22*, 553-560.

Wang, G., and Reinke, V. (2008). A C. elegans Piwi, PRG-1, regulates 21U-RNAs during spermatogenesis. Curr Biol *18*, 861-867.

Ward, S., and Carrel, J.S. (1979). Fertilization and sperm competition in the nematode Caenorhabditis elegans. Dev Biol *73*, 304-321.

Watanabe, T., Takeda, A., Tsukiyama, T., Mise, K., Okuno, T., Sasaki, H., Minami, N., and Imai, H. (2006). Identification and characterization of two novel classes of small RNAs in the mouse germline: retrotransposon-derived siRNAs in oocytes and germline small RNAs in testes. Genes Dev *20*, 1732-1743.

Watanabe, T., Totoki, Y., Toyoda, A., Kaneda, M., Kuramochi-Miyagawa, S., Obata, Y., Chiba, H., Kohara, Y., Kono, T., Nakano, T*., et al.* (2008). Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. Nature *453*, 539-543.

Wightman, B., Ha, I., and Ruvkun, G. (1993). Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. Cell *75*, 855-862.

Wu, L., Fan, J., and Belasco, J.G. (2006). MicroRNAs direct rapid deadenylation of mRNA. Proc Natl Acad Sci U S A *103*, 4034-4039.

Yan, K.S., Yan, S., Farooq, A., Han, A., Zeng, L., and Zhou, M.M. (2003). Structure and conserved RNA binding of the PAZ domain. Nature *426*, 468-474.

Yigit, E., Batista, P.J., Bei, Y., Pang, K.M., Chen, C.C., Tolia, N.H., Joshua-Tor, L., Mitani, S., Simard, M.J., and Mello, C.C. (2006). Analysis of the C. elegans Argonaute family reveals that distinct Argonautes act sequentially during RNAi. Cell *127*, 747-757.

Yin, H., and Lin, H. (2007). An epigenetic activation role of Piwi and a Piwi-associated piRNA in Drosophila melanogaster. Nature *450*, 304-308.

Yu, B., Yang, Z., Li, J., Minakhina, S., Yang, M., Padgett, R.W., Steward, R., and Chen, X. (2005). Methylation as a crucial step in plant microRNA biogenesis. Science *307*, 932-935.

Yuan, Y.R., Pei, Y., Ma, J.B., Kuryavyi, V., Zhadina, M., Meister, G., Chen, H.Y., Dauter, Z., Tuschl, T., and Patel, D.J. (2005). Crystal structure of A. aeolicus argonaute, a site-specific DNA-guided endoribonuclease, provides insights into RISC-mediated mRNA cleavage. Mol Cell *19*, 405-419.

Zhang, K., Mosch, K., Fischle, W., and Grewal, S.I. (2008). Roles of the Clr4 methyltransferase complex in nucleation, spreading and maintenance of heterochromatin. Nat Struct Mol Biol *15*, 381-388.

Zhang, X., Yazaki, J., Sundaresan, A., Cokus, S., Chan, S.W., Chen, H., Henderson, I.R., Shinn, P., Pellegrini, M., Jacobsen, S.E., and Ecker, J.R. (2006). Genome-wide high-resolution mapping and functional analysis of DNA methylation in arabidopsis. Cell *126*, 1189-1201.

6. Appendix   (I)

6.1 Acknowledgements

I would like to thank Dr. Eric Miska for his help, support and guidance throughout my PhD period. I am grateful for his supervision, mentorship and foremost for providing me with the opportunity to work as a part of his research group.

I am thankful to Prof. Dr. Ernst Wimmer and Prof. Dr. Ralf Ficner for their support, hospitality and valuable time.

I would like to thank Rob Shaw for his great job in strain management. I thank Erica Havecker, Prof. David Baulcombe and James Hadfield for their help with solexa sequencing. I thank Leonard Goldstein for his excellent job in analyzing the high throughput sequencing data. I thank Fuchou Tang for his help with qRT-PCR assays for small RNAs. I would like to thank Max and Tim for their help in gel filtration chromatography. I thank to *Caenorhabditis* Genetics Center, University of Minnesota and Dr. Shohei Mitani for providing *C. elegans* strains.

I would like to thank all members of the Miska lab for many stimulating scientific discussions, especially Marloes, Rob, Alex and Nic. I am thankful to Eric, Rob and Nic for correcting this manuscript. I would like to thank my good friend and fellow bench member Marloes for her help in this project. Without her help, this work would not have been possible to finish. I would also like to thank all the lab members for making this lab a very comfortable place to work for my entire PhD. I want to thank my 'brother' Javi for cheering me up all the time. I am thankful to Javi, Ale and Funda, who gave me shelter when I had no place to stay in Cambridge!

I would like to thank Alex and Andrea for their help during thesis submission.

6.2 Declaration

The work described in this dissertation "Piwi and piRNAa act upstream of an endogenous siRNA pathway to suppress Tc3 transposon mobility in the *Caenorhabditis elegans* germline" was carried out in the Wellcome Trust/Cancer Research UK Gurdon Institute, The Henry Wellcome Building of Cancer and Developmental Biology, University of Cambridge between November, 2005 to September, 2008.

This dissertation is a result of my own work, except where specific reference is made to the work by others. Any work done as a part of collaboration is clearly stated in the text. No part of this dissertation has already been, or is presently being, submitted for any other degree or qualification.

Parts of this thesis were published in:

Das PP, Bagijn MP, Goldstein LD, Woolford JR, Lehrbach NJ, Sapetschnig A, Buhecha HR, Gilchrist MJ, Howe KL, Stark R, Matthews N, Berezikov E, Ketting RF, Tavaré S, Miska EA. Piwi and piRNAs act upstream of an endogenous siRNA pathway to suppress Tc3 transposon mobility in the *Caenorhabditis elegans* germline. **Mol Cell (2008 Jul 11; 31(1): 79-90)**

Partha Pratim Das
Göttingen, 15.09.2008

(III)

6.3 Publications

Lykke-Andersen K, Gilchrist MJ, Grabarek JB, **Das PP**, Miska EA, Zernicka-Goetz M. Maternal Argonaute 2 is essential for early mouse development at the maternal-zygotic transition. **MBC (2008 Aug 13)**

Das PP, Bagijn MP, Goldstein LD, Woolford JR, Lehrbach NJ, Sapetschnig A, Buhecha HR, Gilchrist MJ, Howe KL, Stark R, Matthews N, Berezikov E, Ketting RF, Tavaré S, Miska EA. Piwi and piRNAs act upstream of an endogenous siRNA pathway to suppress Tc3 transposon mobility in the *Caenorhabditis elegans* germline. **Mol Cell (2008 Jul 11; 31(1): 79-90)**

Hayashi K, Chuva de Sousa Lopes SM, Kaneda M, Tang F, Hajkova P, Lao K, O'Carroll D, **Das PP**, Tarakhovsky A, Miska EA, Surani MA. MicroRNA biogenesis is required for mouse primordial germ cell development and spermatogenesis. **PLoS ONE (2008 Mar 5; 3(3): e1738)**

Vigorito E, Perks KL, Abreu-Goodger C, Bunting S, Xiang Z, Kohlhaas S, **Das PP**, Miska EA, Rodriguez A, Bradley A, Smith KG, Rada C, Enright AJ, Toellner KM, Maclennan IC, Turner M. microRNA-155 regulates the generation of immunoglobulin class-switched plasma cells. **Immunity (2007 Dec; 27(6): 847-59)**

O'carroll D, Mecklenbrauker I, **Das PP**, Santana A, Koenig U, Enright AJ, Miska EA, Tarakhovsky.A Slicer-independent role for Argonaute 2 in hematopoiesis and the microRNA pathway. **Genes Dev (2007 Aug 15; 21(16): 1999-2004)**

Rodriguez A, Vigorito E, Clare S, Warren MV, Couttet P, Soond DR, van Dongen S, Grocock RJ, **Das PP**, Miska EA, Vetrie D, Okkenhaug K, Enright AJ, Dougan G, Turner M, Bradley. A Requirement of bic/microRNA-155 for normal immune function. **Science (2007 Apr 27; 316(5824): 608-11)**

Lujambio A, Ropero S, Ballestar E, Fraga MF, Cerrato C, Setien F, Casado S, Suarez-Gauthier A, Sanchez-Cespedes M, Gitt A, Spiteri I, **Das PP**, Caldas C, Miska EA, Esteller M. Genetic unmasking of an epigenetically silenced microRNA in human cancer cells. **Cancer Res. (2007 Feb 15; 67(4): 1424-9)**

6.4 CURRICULUM VITAE

NAME: PARTHA PRATIM DAS

NATIONALITY: Indian.

DATE OF BIRTH: 30/11/1977

SEX: Male.

ADDRESS FOR CORRESPONDENCE: Wellcome Trust / Cancer Research UK Gurdon Institute, The Henry Wellcome Building of Cancer and Developmental Biology, University of Cambridge, Tennis court Road, Cambridge CB2 1QN, United Kingdom

HOME ADDRESS: 35/1C Gopal Misra Road, Newpark, Bahela, Calcutta-700034, West Bengal, India.

Phone no: +44-7706922686 (M)
        +44-1223767221 (O)

E-MAIL:
p.das@gurdon.cam.ac.uk

EDUCATIONAL PROFILE

PhD student, Wellcome Trust/ Cancer Research UK Gurdon Institute, University of Cambridge, Cambridge, UK and University of Gottingen, Germany (2005 April – Present)

Research student, Max Planck Institute of Biochemistry, Martinsried near Munich, Germany (2004 April – 2005 Mar)

MSc in Molecular Biology, International Max Planck Research School For Molecular Biology, Gottingen, Germany (2002 Sep – 2004 Mar)

MSc in Zoology and Molecular Biology, Department of Zoology, University of Pune, Pune, India (2000 June – 2002 July)

BSc in Zoology, botany and Chemistry, University of Calcutta, Calcutta, India (1996 Aug – 1999 Dec)

RESEARCH EXPERIENCE

1. PhD project: Role of short RNAs in Caenorhabditis elegans germline development (Eric Miska's Lab, Wellcome Trust/Cancer Research UK Gurdon Institute, University of Cambridge, Cambridge, UK)

Piwi and piRNAs Act Upstream of an Endogenous siRNA Pathway to Suppress Tc3 Transposon Mobility in the *Caenorhabditis elegans* Germline. *Mol cell* (2008 Jul 11; 31(1): 79-90)

The Piwi proteins of the Argonaute superfamily are required for normal germline development in *Drosophila*, zebrafish, and mice and associate with 24–30 nucleotide RNAs termed piRNAs. We identify a class of 21 nucleotide RNAs, previously named 21U-RNAs, as the piRNAs of *C. elegans*. Piwi and piRNA expression is restricted to the male and female germline and independent of many proteins in other small-RNA pathways, including DCR-1. We show that Piwi is specifically required to silence Tc3, but not other Tc/mariner DNA transposons. Tc3 excision rates in the germline are increased at least 100-fold in *piwi* mutants as compared to wild-type. We find no evidence for a Ping-Pong model for piRNA amplification in *C. elegans*. Instead, we demonstrate that Piwi acts upstream of an endogenous siRNA pathway in Tc3 silencing. These data might suggest a link between piRNA and siRNA function.

2. MSc project: Translational regulation of p27 by RPL18a, a component of the 60S ribosome (Ludger Hengst's Lab, Department of Regulation of Cell Proliferation, Max Planck Institute For Biochemistry, Martinsried, Germany)

p27 (Kip1) regulates cell proliferation by binding to and modulating the activity of cyclin-dependent kinases. In normal cells, p27 protein increases in growth arrest but also oscillates during cell cycle progression. Expression of p27 is regulated through mechanisms including transcription, translation and ubiquitin-mediated degradation. Each of these pathways may contribute to deregulation of p27 in hyperproliferative diseases. Protein levels of p27 are oscillating periodically in proliferating cells but mRNA level remain constant throughout the cell cycle. The protein level of p27 reaches maximum in G_0 phase and declines as cells enter into S phase. To investigate the mechanisms of p27 translational regulation, proteins has been isolated that bind to the 5´UTR of p27mRNA. In a first analysis, we investigated their ability to regulate p27 expression by using a reporter translation system under control of the p27 5´UTR. Our group already identified some of the specific p27 5'UTR binding proteins and which regulates p27 translation (For example HuR (Kullmann et al., 2002), which inhibits p27 translation). In my master thesis, I found out a protein (RPL18a), which is the component of 60S ribosome also negatively regulates p27 translation.

3. *Title: Functional identity of promoter motifs of 122 amino acids ORF from Drosophila melanogaster (MSc thesis project)*

(Department of Zoology, University of Pune, India.)

SCHOLARSHIPS AND ACADEMIC ACHIEVEMENTS:

- Selection for MSc in Department of Zoology, University of Pune, India (through All India based entrance examination)

- Selection for MSc-PhD Program in International Max Planck Research School For Molecular Biology, Goettingen, Germany (Global selection through entrance test and interview; total seat-20)

- Max Planck Research fellowships (Sep 2002- Aug 2003) during the study of MSc

- Max Planck Research fellowships (2003 Sep -2005 Mar)

- Cancer Research UK (2005- present)

PUBLICATIONS

Lykke-Andersen K, Gilchrist MJ, Grabarek JB, **Das PP**, Miska EA, Zernicka-Goetz M. Maternal Argonaute 2 is essential for early mouse development at the maternal-zygotic transition. **MBC (2008 Aug 13)**

Das PP, Bagijn MP, Goldstein LD, Woolford JR, Lehrbach NJ, Sapetschnig A, Buhecha HR, Gilchrist MJ, Howe KL, Stark R, Matthews N, Berezikov E, Ketting RF, Tavaré S, Miska EA. Piwi and piRNAs act upstream of an endogenous siRNA pathway to suppress Tc3 transposon mobility in the *Caenorhabditis elegans* germline. **Mol Cell (2008 Jul 11; 31(1): 79-90)**

Hayashi K, Chuva de Sousa Lopes SM, Kaneda M, Tang F, Hajkova P, Lao K, O'Carroll D, **Das PP**, Tarakhovsky A, Miska EA, Surani MA. MicroRNA biogenesis is required for mouse primordial germ cell development and spermatogenesis. **PLoS ONE (2008 Mar 5; 3(3): e1738)**

Vigorito E, Perks KL, Abreu-Goodger C, Bunting S, Xiang Z, Kohlhaas S, **Das PP**, Miska EA, Rodriguez A, Bradley A, Smith KG, Rada C, Enright AJ, Toellner KM, Maclennan IC, Turner M. microRNA-155 regulates the generation of immunoglobulin class-switched plasma cells. **Immunity (2007 Dec; 27(6): 847-59)**

O'carroll D, Mecklenbrauker I, **Das PP**, Santana A, Koenig U, Enright AJ, Miska EA, Tarakhovsky.A Slicer-independent role for Argonaute 2 in hematopoiesis and the microRNA pathway. **Genes Dev (2007 Aug 15; 21(16): 1999-2004)**

Rodriguez A, Vigorito E, Clare S, Warren MV, Couttet P, Soond DR, van Dongen S, Grocock RJ, **Das PP**, Miska EA, Vetrie D, Okkenhaug K, Enright AJ, Dougan G, Turner M, Bradley. A Requirement of bic/microRNA-155 for normal immune function. **Science (2007 Apr 27; 316(5824): 608-11)**

Lujambio A, Ropero S, Ballestar E, Fraga MF, Cerrato C, Setien F, Casado S, Suarez-Gauthier A, Sanchez-Cespedes M, Gitt A, Spiteri I, **Das PP**, Caldas C, Miska EA, Esteller M. Genetic unmasking of an epigenetically silenced microRNA in human cancer cells. **Cancer Res. (2007 Feb 15; 67(4): 1424-9)**

Article

Piwi and piRNAs Act Upstream of an Endogenous siRNA Pathway to Suppress Tc3 Transposon Mobility in the *Caenorhabditis elegans* Germline

Partha P. Das,[1,2,3] Marloes P. Bagijn,[1,4] Leonard D. Goldstein,[4,5,7] Julie R. Woolford,[1,2] Nicolas J. Lehrbach,[1,2] Alexandra Sapetschnig,[1,2] Heeran R. Buhecha,[1,2] Michael J. Gilchrist,[1] Kevin L. Howe,[5] Rory Stark,[5] Nik Matthews,[5] Eugene Berezikov,[6] René F. Ketting,[6] Simon Tavaré,[4,5,7] and Eric A. Miska[1,2,*]

[1]Wellcome Trust Cancer Research UK Gurdon Institute, University of Cambridge, The Henry Wellcome Building of Cancer and Developmental Biology, Tennis Court Road, Cambridge CB2 1QN, UK
[2]Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge CB2 1GA, UK
[3]PhD Program, Faculty of Biology, GZMB, Georg-August-Universität Göttingen, Justus-von-Liebig-Weg 11, 37077 Göttingen, Germany
[4]Department of Oncology, University of Cambridge, Hills Road, Cambridge CB2 2XZ, UK
[5]Cancer Research UK, Cambridge Research Institute, Li Ka-Shing Centre, Robinson Way, Cambridge CB2 0RE, UK
[6]Hubrecht Institute, Uppsalalaan 8, 3584 CT, Utrecht, The Netherlands
[7]Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Centre for Mathematical Sciences, Wilberforce Road, Cambridge CB3 0WA, UK
*Correspondence: eam29@cam.ac.uk
DOI 10.1016/j.molcel.2008.06.003

SUMMARY

The Piwi proteins of the Argonaute superfamily are required for normal germline development in *Drosophila*, zebrafish, and mice and associate with 24–30 nucleotide RNAs termed piRNAs. We identify a class of 21 nucleotide RNAs, previously named 21U-RNAs, as the piRNAs of *C. elegans*. Piwi and piRNA expression is restricted to the male and female germline and independent of many proteins in other small-RNA pathways, including DCR-1. We show that Piwi is specifically required to silence Tc3, but not other Tc/mariner DNA transposons. Tc3 excision rates in the germline are increased at least 100-fold in *piwi* mutants as compared to wild-type. We find no evidence for a Ping-Pong model for piRNA amplification in *C. elegans*. Instead, we demonstrate that Piwi acts upstream of an endogenous siRNA pathway in Tc3 silencing. These data might suggest a link between piRNA and siRNA function.

INTRODUCTION

Piwi proteins are part of a superfamily of Argonaute proteins that is defined by the presence of PAZ and Piwi domains (Cerutti et al., 2000). The PAZ domain has been shown to bind to 3′ ends of short RNAs, and the Piwi domain is similar to the catalytic domain of RNase H. The Argonaute superfamily can easily be divided into two clades according to sequence similarity (Carmell et al., 2002). The AGO clade is found in fission yeast, plants, and animals, whereas the Piwi clade is found in ciliates, slime molds, and animals. Members of the AGO clade directly bind to siRNAs and miRNAs, are involved in many aspects of transcriptional (TGS) and posttranscriptional gene silencing (PTGS) in many

species, and are generally ubiquitously expressed (Joshua-Tor, 2006).

The expression of proteins of the Piwi clade might be restricted to germ cells in vertebrates (Carmell et al., 2007; Deng and Lin, 2002; Houwing et al., 2007; Kuramochi-Miyagawa et al., 2001, 2004), but not in *Drosophila*, where Piwis are also found in somatic cells associated with the germline (Brennecke et al., 2007; Cox et al., 1998; Gunawardane et al., 2007; Saito et al., 2006). In flatworms, Piwis are expressed in germ cells and somatic stem cells (Reddien et al., 2005). A number of functions have been assigned to Piwi proteins: *Drosophila* Piwi is required for oogenesis and stem cell maintenance (Cox et al., 1998, 2000; Lin and Spradling, 1997); a *Tetrahymena* Piwi protein, Tiwi, is required for DNA elimination in the macronucleus (Mochizuki et al., 2002); in planaria the Piwi protein SMEDWI-1 is required for stem cell regulation (Reddien et al., 2005); and, in vertebrates, Piwi proteins are required for germline development in the zebrafish (Houwing et al., 2007) and male germline development in the mouse (Aravin et al., 2006; Deng and Lin, 2002; Girard et al., 2006; Grivna et al., 2006; Kuramochi-Miyagawa et al., 2001). Piwi proteins are also required for transposon silencing in the *Drosophila* germline (Aravin et al., 2001; Savitsky et al., 2006; Vagin et al., 2006), and several studies suggest that many of the phenotypes observed in *piwi* mutants might be due to loss of germline integrity (Chen et al., 2007; Klattenhoff et al., 2007; Pane et al., 2007).

Recently, a class of 24–30 nucleotide RNAs has been found to interact with Piwi proteins in *Drosophila*, zebrafish, mice, and rats and has been named *Pi*wi-interacting *RNAs* or piRNAs (Klattenhoff and Theurkauf, 2008). piRNA populations are complex; there are hundreds of thousands of unique piRNAs in mammals. The identification of piRNAs has improved our understanding of Piwi function, in particular with regard to transposon silencing: for example, in *Drosophila* the piRNAs of the *flamenco* locus control the *gypsy* retrotransposon (Brennecke et al., 2007; Desset et al., 2003; Prud'homme et al., 1995).

The best-studied transposable elements of *C. elegans* are the DNA transposons of the Tc/*mariner* superfamily (Moerman and Waterston, 1984), in particular Tc1 and Tc3 (van Luenen et al., 1994; Vos et al., 1996). Tc1 and Tc3 are also the most abundant transposons of the Tc family in the N2 Bristol strain of *C. elegans* with 31 and 22 copies, respectively (*C. elegans* Sequencing Consortium, 1998). Tc1 and Tc3 are autonomous elements encoding a transposase specific to each element (van Luenen et al., 1993; Vos et al., 1993).

Here we demonstrate a role for *C. elegans* Piwi in germline development and germline transposon silencing. We identify the recently named 21U-RNAs as the piRNAs of *C. elegans*. We demonstrate that in *C. elegans* piRNAs act upstream of an endogenous siRNA pathway for Tc3 silencing. These data shed light on piRNA function.

RESULTS

Piwi Is Required for Normal Germline Development

To investigate the roles of Piwi proteins in *C. elegans* we generated mutants lacking Piwi function. The *C. elegans* genome encodes two *Piwi*-related genes, *prg-1* and *prg-2*. These genes are likely the result of a recent gene duplication, as the genomes of the related nematode species *C. briggsae* and *C. remanei* each contain a single *prg* gene. PRG-1 and PRG-2 are 91% identical at the amino acid level, which suggests that they might act redundantly. We generated two deletion alleles each for *prg-1* (*n4357*, *n4503*) and *prg-2* (*n4358*, *nDf57*) and generated double mutant strains after outcrossing of the single mutants (see the Supplemental Data available online). For brevity, we will refer to PRG-1 and PRG-2 as Piwi, and we will refer to *prg-1; prg-2* double mutants as *piwi* mutants, i.e., *piwi(n4357; n4358)* instead of *prg-1(n4357); prg-2(n4358)*. *prg-1* and *prg-2* single mutants and *piwi* mutants were homozygous viable and showed neither a defect in exogenous RNAi in either the soma or germline nor defects in miRNA biogenesis or function (data not shown). However, all mutant strains showed reduced fertility (Figure S1A). Of the single mutants, *prg-1* mutant animals showed the most pronounced effect, with fertility reduced to 25% of that of wild-type animals (Figure S1A). The fertility defect is enhanced in *piwi* mutants (Figure S1A). These observations agreed with previous studies of *prg-1* using RNAi and an independent *prg-1* allele (Cox et al., 1998; Yigit et al., 2006). As *prg-1* RNAi had suggested a role for PRG-1 in spermatogenesis (Cox et al., 1998), we counted hermaphrodite sperm in wild-type and *prg-1* and *prg-2* mutant animals. Sperm counts were reduced to approximately 50% in both single and double mutants (Figure S1B). However, fertility of *piwi* mutants was not restored to wild-type levels by introducing wild-type sperm through mating, suggesting that Piwi function is not restricted to spermatogenesis (Figure S1C). Finally, *piwi* mutant germlines showed abnormal mitotic to meiotic transitions (data not shown). These data confirm a conserved role for PRG-1 and PRG-2 in *C. elegans* germline development (Cox et al., 1998).

The piRNAs of *C. elegans* Are 21 Nucleotide RNAs

In *Drosophila*, zebrafish, mice, and rats, Piwi proteins are associated with 24–30 nucleotide piRNAs (Aravin et al., 2006; Bren-

necke et al., 2007; Girard et al., 2006; Grivna et al., 2006; Gunawardane et al., 2007; Houwing et al., 2007; Lau et al., 2006; Saito et al., 2006; Vagin et al., 2006; Watanabe et al., 2006). Searching for piRNAs in *C. elegans*, we were unable to identify an abundant class of RNAs in this size range (data not shown). We therefore searched for piRNAs among the small RNAs previously identified in *C. elegans*: miRNAs, tncRNAs, endogenous siRNAs, and 21U-RNAs (Ambros et al., 2003; Lau et al., 2001; Lee and Ambros, 2001; Lim et al., 2003; Ruby et al., 2006). To identify candidate piRNAs, we tested if any of these short RNAs were dependent on Piwi. Surprisingly, we found that a 21U-RNA, 21UR-1, which was detected in RNA from wild-type animals, was absent in RNA from two independent *piwi* mutants by northern blotting (Figure 1A). In contrast, expression of a ubiquitous miRNA, miR-52, was unaffected. To test if 21U-RNAs were generally absent in *piwi* mutants we generated libraries of 5′ monophosphate small RNAs from wild-type and *piwi(n4357; n4358)* mutants. High-throughput sequencing identified 1398 out of 5454 previously known 21U-RNAs (Ruby et al., 2006). We also identified a large number of candidate 21U-RNAs (data not shown). 21U-RNAs were either absent or dramatically underrepresented in the *piwi* mutant library as compared to the wild-type library (Figure 1B). The most abundant 21U-RNA in the *piwi* sample had eight reads as compared to 2127 reads in wild-type. We also assessed the expression levels of other small RNAs in wild-type versus *piwi* mutants and found no differences in miRNA expression, tncRNA expression, or a number of siRNA species including a 26 nucleotide siRNA (Figure 1C and data not shown). These data suggest that 21U-RNAs might be the piRNAs of *C. elegans*. To test this hypothesis directly, we generated a rabbit polyclonal antibody against PRG-1. After immunoprecipitation of PRG-1 using αPRG-1 serum from wild-type *C. elegans* adult whole-cell extracts, we detected 21U-RNAs by RT-PCR but were not able to detect them from *piwi* mutant extracts or when using preimmune serum (Figure 1D). Overall, 21U-RNAs are 100-fold enriched in PRG-1 immunoprecipitates (Batista et al., 2008 [this issue of *Molecular Cell*]; see Discussion). As the high-throughput sequencing data suggested that 21U-RNAs were dramatically reduced in *piwi* mutants, we independently quantified the expression of seven 21U-RNAs by quantitative RT-PCR. As shown in Figure 1E, while the expression of a number of 21U-RNAs is dramatically reduced, some 21U-RNAs, including 21UR-1, were still detected in *piwi* mutants. For 21UR-1 we verified that the signal was specific by cloning and sequencing of RT-PCR products (data not shown). Taken together, these data suggest that the 21U-RNAs are the piRNAs of *C. elegans*, and we will refer to them as piRNAs below.

piRNA Biogenesis Is Independent of Many Genes in Other Small-RNA Pathways

To identify additional genes involved in piRNA pathways in *C. elegans*, we tested a panel of genes using mutants and RNAi for their effect on piRNA expression by northern blotting or quantitative RT-PCR (Figure 2 and Table S1). First, we checked piRNA expression in *prg-1* and *prg-2* single mutants. Using three independent alleles and RNAi experiments, we found that only PRG-1 but not PRG-2 is required for piRNA
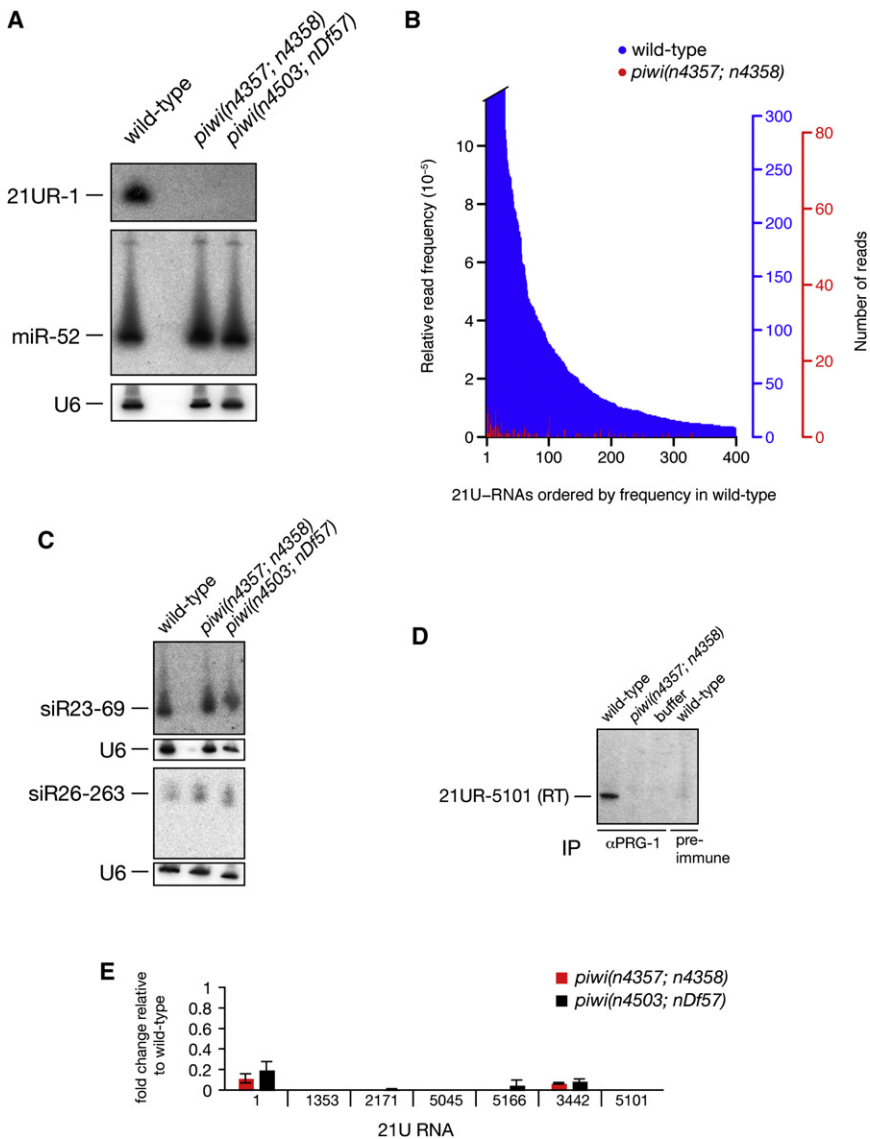
Figure 1. The 21U-RNAs Are *C. elegans* piRNAs

(A) Northern blot showing that 21UR-1 is not detected in RNA (40 μg) isolated from two independent *piwi* double mutants: *piwi(n4357; n4358)* and *piwi(n4503; nDf57)*, whereas miR-52 is expressed in both mutants (same blot reprobed). Antisense DNA probes were used for 21UR-1 and miR-52. A U6 northern blot is shown as loading control. *piwi(n4357; n4358)* is an abbreviation for *prg-1(n4357); prg-2(n4358)*. *piwi(n4503; nDf57)* is an abbreviation for *prg-1(n4503); prg-2(nDf57)*.

(B) High-throughput sequencing reveals that the expression of many 21U-RNAs is dramatically reduced in *piwi* mutants. 21U-RNAs cloned from 5′ dependent wild-type and *piwi(n4357; n4358)* mutant libraries. Frequencies are shown for wild-type (blue) and *piwi* mutant (red) for the 400 most abundant 21U-RNAs in wild-type, plotted in the order of their wild-type frequency. Read frequencies were obtained by dividing the number of reads for a given 21U-RNA by the total number of reads from the same library (left-hand y axis). The corresponding absolute number of reads are indicated in the right-hand y axes. 21U-RNAs for which frequencies are shown include the 21U-RNA with most reads in the *piwi* mutant library (21UR-3224), which was sequenced eight times in the *piwi* mutant and 2127 times in wild-type.

(C) Expression of a 23 nucleotide antisense RNA (siR23-69) and a 26 nucleotide antisense RNA (siR26-263) is not affected in *piwi* mutants (northern blotting, 40 μg total RNA, antisense DNA probes).

(D) Immunoprecipitation followed by RT-PCR for 21UR-5101 reveals that 21U-RNAs are associated with PRG-1 in *C. elegans* extracts.

(E) Quantitative RT-PCR of seven 21U-RNAs demonstrates that Piwi is not essential for 21U-RNA biogenesis. Total RNA was extracted from 12 hr adult *C. elegans*. Expression levels shown are relative to levels in wild-type RNA. miR-52 expression was used as an internal control. Data are from three independent biological replicates. Error bars represent standard error of the mean.

expression (Figure 2 and Table S1). Next, we tested other Argonaute proteins; these included RDE-1, which is required for exogenous RNAi (Tabara et al., 1999), ALG-1 and ALG-2, which are redundantly required for miRNA function (Grishok et al., 2001), ERGO-1, SAGO-1, SAGO-2, and a group of five "MAGO" Argonaute proteins associated with endogenous siRNA pathways (Yigit et al., 2006). None of these Argonaute proteins were required for piRNA expression, emphasizing the specificity of the requirement for PRG-1. In addition, we tested a number of other proteins involved in small-RNA biology in *C. elegans*, including all four known RNA-dependent RNA polymerases (RdRPs), RRF-1, RRF-2, RRF-3, and EGO-1 (Simmer et al., 2002; Smardon et al., 2000), none of which were required for piRNA expression. We also tested if the RNase III enzyme DCR-1, which is essential for the generation of siRNAs and germline development (Grishok et al., 2001; Ketting et al., 2001; Knight and Bass, 2001), is required for piRNA expression.

As *dcr-1* mutants are sterile, these experiments were carried out using homozygous mutant animals derived from heterozygous mothers or using RNAi. We found that DCR-1 is not required for piRNA expression (Figure 2B; Table S1). These data were also confirmed by high-throughput sequencing in *dcr-1* mutants (Figure 2C). These data suggest that piRNAs are independent of many proteins involved in other small-RNA pathways, including DCR-1.

Piwi and piRNAs Are Restricted to the Male and Female Germline

We found that 21UR-1 expression is developmentally regulated with highest expression detected in young adults and adults by northern blot (Figure 3A). Additional piRNAs showed a similar expression pattern, as determined by quantitative RT-PCR (e.g., 21UR-5101; Figure 3B). Interestingly, *prg-1* and *prg-2* mRNAs show a similar pattern (Figure 3C; Figure S2). The observed
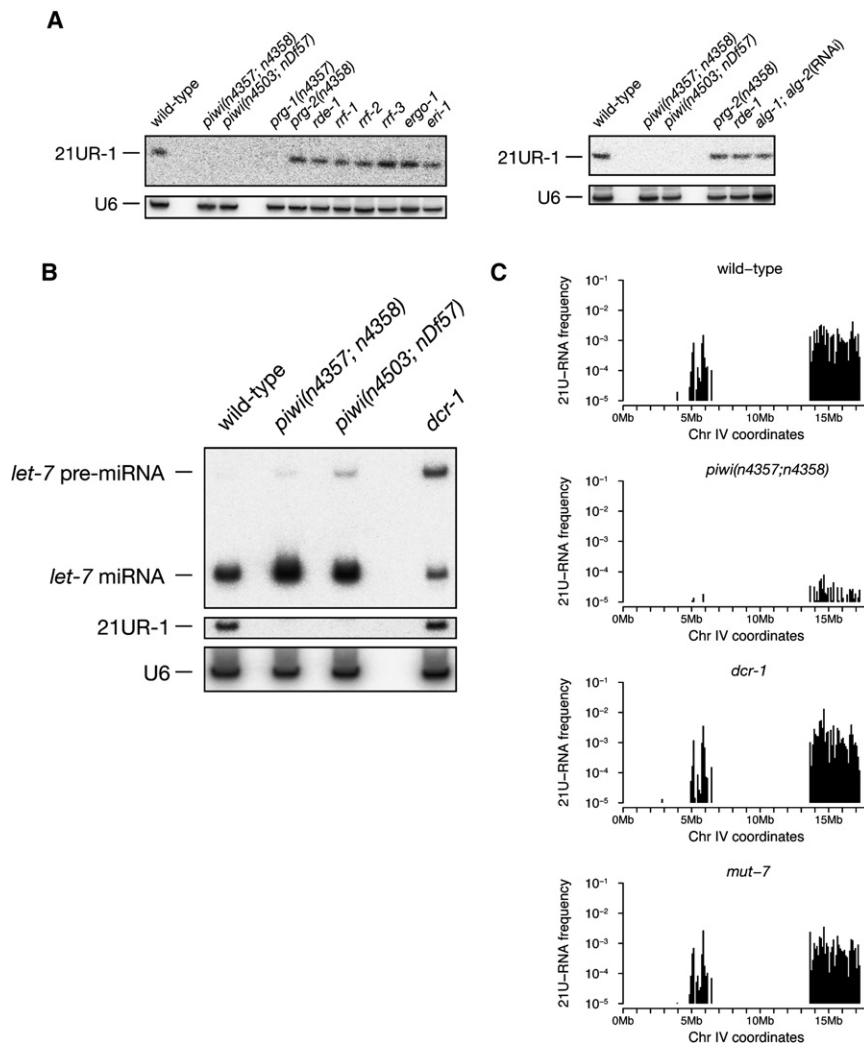
Figure 2. piRNA Biogenesis Does Not Require Many Known Small-RNA Pathway Proteins

(A) 21UR-1 northern blotting of total RNA of wild-type and mutant young adult *C. elegans*. In the case of *alg-1; alg-2*(RNAi), *alg-1* mutant L1 larvae were transferred to *alg-2* RNAi feeding plates, and young adult animals were harvested. A U6 northern blot is shown as loading control. See Table S1 in the Supplemental Data for quantification of these results.

(B) 21UR-1 northern blotting of total RNA of wild-type and mutant young adult *C. elegans*. *dcr-1* mutant animals used were homozygous animals derived from heterozygous mothers. To test for loss of DCR-1 activity in *dcr-1* mutants, *let-7* miRNA and pre-miRNA is shown. A U6 northern blot is shown as loading control.

(C) Distribution of 21U-RNAs on chromosome IV as detected by high-throughput sequencing of 5′ dependent wild-type, *piwi(n4357; n4358)*, *dcr-1*, and *mut-7* mutant libraries. Frequencies for a given 21U-RNA and locus were obtained by correcting the number of reads for multiple alignments and dividing by the total number of reads from the same library. Cumulative frequencies were plotted for nonoverlapping 100 kb windows along chromosome IV.

Piwi Is Specifically Required for Silencing of Tc3 DNA Transposons

As Piwi proteins are required for transposon silencing in *Drosophila* (Aravin et al., 2001; Savitsky et al., 2006; Vagin et al., 2006), we examined transposon silencing in *piwi* mutants in *C. elegans*. We focused on the two most abundant DNA transposons in *C. elegans*, Tc1 and Tc3 (Figures 4A and 4B) (*C. elegans* Sequencing Consortium, 1998). First, we examined Tc1 transposase expression by quantitative RT-PCR using primers specific for 15 Tc1 loci (Figure 4C). We observed an approximately 50-fold increase in Tc1 transposase mRNA levels in *mut-7* mutants, which show an elevated rate of transposition of Tc/mariner elements in the germline (Ketting et al., 1999). We did not, however, observe an increase in Tc1 transposase mRNA in *piwi* mutants. Surprisingly then, we found increased Tc3 transposase mRNA in two independent *piwi* mutants and three independent *prg-1* mutants using primer pairs specific for 20 Tc3 loci (Figure 4D). Next, we assayed the germline excision rate of Tc1 and Tc3 transposons directly by the phenotypic reversion of *unc-22* transposon insertion alleles (Ketting et al., 1999). As shown in Table 1, *unc-22* reversion rates of Tc1, Tc3, or Tc4 insertion alleles were less than 10^{-6} in an otherwise wild-type background. In *mut-7* mutant animals the reversion rate was increased 100-fold, as reported previously (Ketting et al., 1999). Strikingly, in *prg-1* and *piwi* mutants Tc1 and Tc4 excision rates were not affected, but reversion of the Tc3 allele was increased 100- and 1000-fold, respectively (Table 1). These

expression pattern for piRNAs and *piwi* mRNA is consistent with expression in the germline. As Piwi proteins and piRNAs are thought to be exclusively expressed in germ cells in vertebrates (Carmell et al., 2007; Deng and Lin, 2002; Kuramochi-Miyagawa et al., 2004; Kuramochi-Miyagawa et al., 2001), but not in *Drosophila* (Brennecke et al., 2007; Cox et al., 1998; Gunawardane et al., 2007; Saito et al., 2006), we decided to test their germline restriction in *C. elegans* using a set of temperature-sensitive mutants (Figure 3D). 21UR-1 was absent from *glp-4(bn2*ts*)* and *glp-1(e2144*lf*)* mutant animals at the restrictive temperature, which are devoid of germ cells (Beanan and Strome, 1992). However, 21UR-1 levels were unchanged in *glp-1(ar202*gf,ts*)* mutants at the restrictive temperature, which are highly enriched in germ cells that have not yet entered meiosis (Pepper et al., 2003). 21UR-1 RNA was also detected in RNA from *fem-1(hc17*ts*)* (Kimble et al., 1984; Nelson et al., 1978) and *fem-3(q22*sd,ts*)* (Barton et al., 1987) mutants at the restrictive temperature that are devoid of sperm and oocytes, respectively. The same restricted pattern was also observed for *piwi* mRNA (Figure 3E and Figure S2). Together these data suggested that *C. elegans* Piwi and piRNAs are restricted to the male and female germline.
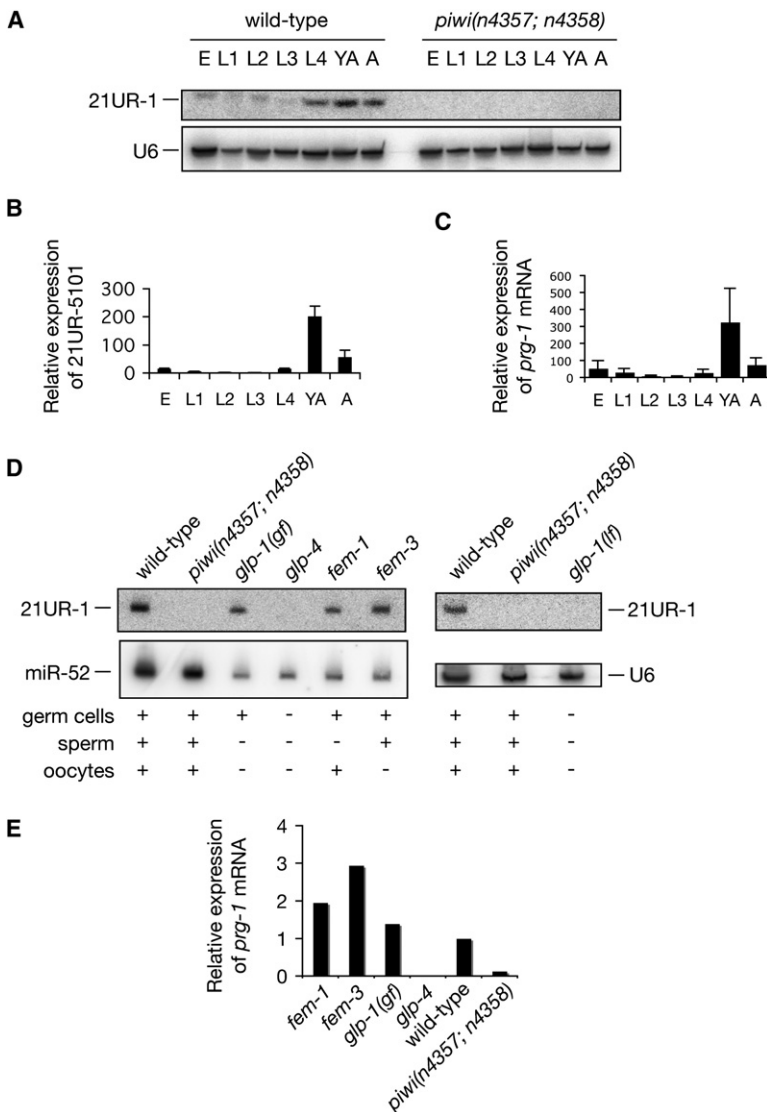
Figure 3. Piwi and piRNAs Are Restricted to the Male and Female Germline

(A) Profile of 21U-R1 expression during development. E, embryo. L1–L4, larval stages 1–4. YA, 12 hr adult. A, 24–48 hr adult. A U6 northern blot is shown as loading control.

(B) Quantitative RT-PCR of 21UR-5101. miR-52 expression was used as an internal control. Data are from three independent biological replicates. Error bars represent standard error of the mean.

(C) Quantitative RT-PCR of *prg-1* mRNA. Actin mRNA was used as an internal control. Data are from three independent biological replicates. Error bars represent standard error of the mean.

(D) Northern blot showing that piRNA expression is restricted to the male and female germline. *glp-1(gf)*, *glp-1(lf)*, *glp-4*, *fem-1*, and *fem-3* L1 larvae were grown to 12 hr adult stage at 25°C. Total RNA (20 μg) was loaded in each lane. U6 and miR-52 northern blots are shown as loading controls.

(E) Quantitative RT-PCR of *prg-1* mRNA. Actin mRNA was used as an internal control.

data suggest that PRG-1 and Piwi are powerful and specific suppressors of Tc3 transposition in the germline of *C. elegans*. As we found that Piwi and MUT-7 were both involved in germline transposon silencing, we wondered if these proteins shared additional functions. We therefore tested if Piwi, like MUT-7, is required for germline transgene silencing (Kim et al., 2005). However, we found that germline transgene silencing is intact in *piwi* mutants (Figure S3). These data suggest that Piwi and MUT-7 have overlapping and distinct roles in the *C. elegans* germline.

Piwi Acts Upstream of an Endogenous siRNA Pathway

Previous work demonstrated a role for MUT-7 (Ketting et al., 1999) and endogenous siRNAs (Sijen and Plasterk, 2003) in germline transposon silencing. Our findings suggested a role for Piwi and piRNAs in the same process. We therefore wanted to determine how these two pathways related to each other. First, we tested if MUT-7 and endogenous siRNAs might act

upstream of piRNA expression. However, we found that piRNA levels were not affected in *mut-7* mutants by northern blotting and high-throughput sequencing (Figure 2C; Figure 5A). Next, we tested if Piwi and piRNAs might act upstream of MUT-7 and endogenous siRNAs. We searched for piRNAs mapping to Tc/mariner transposons and found that no known piRNAs mapped directly to Tc1 or Tc4. However, a single piRNA, 21UR-139, mapped to the sense strand of the transposase gene of 20 out of the 22 Tc3 insertions in the N2 genome (Table S2; Figure 5B). A second 21U-RNA in the Tc3 locus has also been identified (Batista et al., 2008; see Discussion). To identify endogenous siRNAs mapping to Tc transposons, we generated 5′ independent small-RNA libraries from wild-type and *piwi* mutants, as we expected these RNAs to carry 5′ di- or tri-phosphates (Pak and Fire, 2007; Ruby et al., 2006; Sijen et al., 2007). High-throughput sequencing identified a large number of endogenous small RNAs that map to Tc1 and Tc3 loci (Figure 5B; Supplemental Data). Interestingly, small RNAs mapping to the terminal inverted repeat (TIR) and the transposase open reading frame (ORF) of Tc3 were nearly absent in *piwi* mutants, but those mapping to Tc1 were unaffected. We confirmed these observations for small RNAs antisense to the transposase transcripts using an RNase protection assay (Figure 5C). While Tc1 and Tc3 antisense siRNAs were dependent on MUT-7, only Tc3 siRNAs were also dependent on Piwi. Unfortunately, we were unable to detect endogenous small RNAs from the TIRs using the same assay (data not shown). These data suggest that Piwi and piRNAs act upstream of a MUT-7-dependent endogenous siRNA pathway for Tc3 silencing. This finding raised the possibility that additional piRNAs are functionally linked to endogenous siRNAs. To test this hypothesis we examined piRNA loci for nearby endogenous siRNAs. Interestingly, we observed a local positional
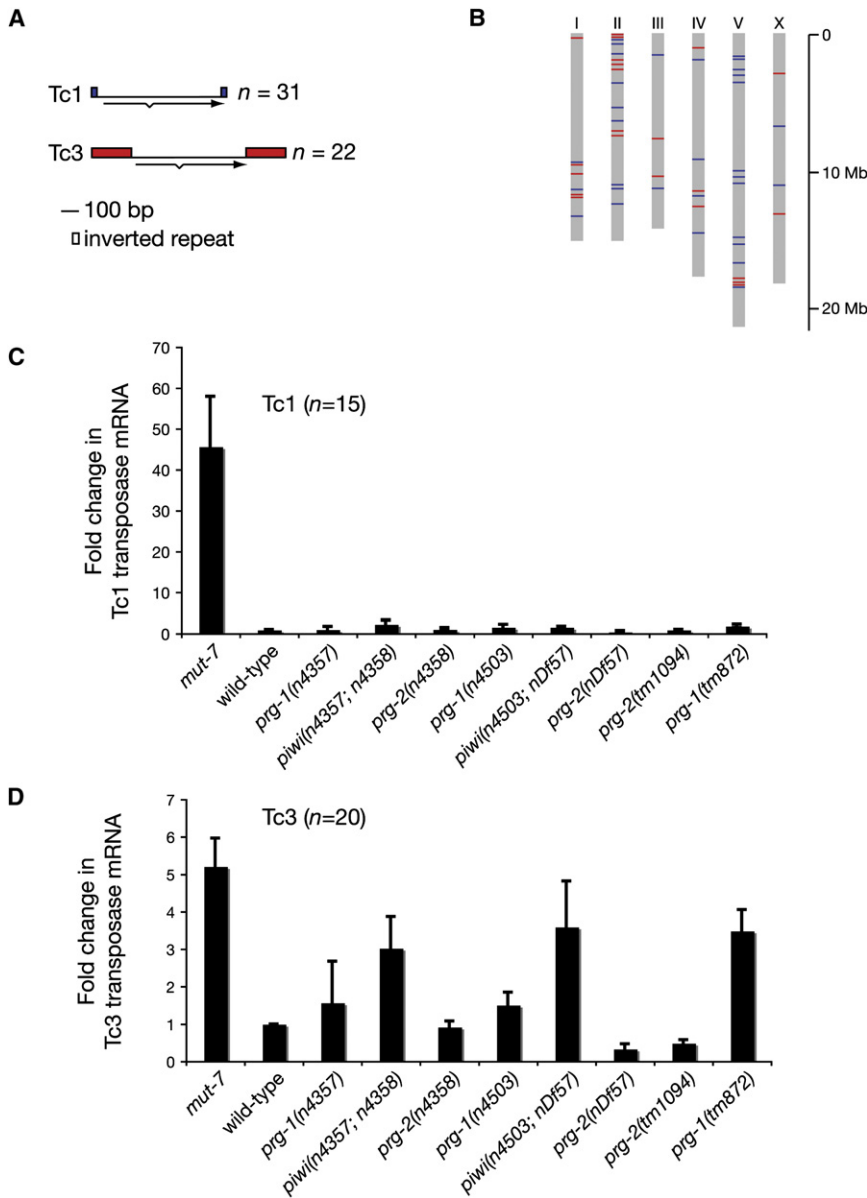
Figure 4. Piwi Is Required to Inhibit Tc3 Transposase Expression

(A) Diagram of the genomic structure of the two most common DNA transposons in *C. elegans*, Tc1, and Tc3. Tc1 and Tc3 are flanked by inverted repeats and encode a single spliced transcript for transposase. bp, base pairs. *n*, number of copies in the wild-type strain N2.

(B) Distribution of Tc1 and Tc3 transposons in the *C. elegans* genome.

(C and D) Quantitative RT-PCR of Tc1 or Tc3 transposase mRNA. As the genomic copies of Tc1 and Tc3 have minor sequence variations, the number of transposon loci amplified by each qRT-PCR primer pair are shown (*n*). Actin mRNA was used as an internal control. Expression levels shown are relative to levels from wild-type RNA. Data are from three independent biological replicates. Error bars represent standard error of the mean.

bias of antisense siRNAs downstream of piRNA loci, and their accumulation appears to be dependent on Piwi (Figure 6A; see Discussion).

DISCUSSION

21U-RNAs Are piRNAs with Surprising Features

We identify the piRNAs of *C. elegans* as 21 nucleotide RNAs (Figure 1). Our observations are strongly supported by the finding that PRG-1 immunoprecipitates are 100-fold enriched for 21U-RNAs (Batista et al., 2008). The *C. elegans* 21U-RNAs/piRNAs share several features with piRNAs of other species. First, *C. elegans* piRNAs have 5′ uridines. This bias is common to piRNAs from *Drosophila* and vertebrates (Klattenhoff and Theurkauf, 2008). Second, several *C. elegans* piRNAs have a 5′ monophosphate and a 3′ modification blocking the 2′ or 3′ oxygen, likely a 2′-O-methyl group (data not shown; Ruby et al., 2006). This modification is shared by piRNAs from *Drosophila* (Horwich et al., 2007; Saito et al., 2007; Vagin et al., 2006), zebrafish (Houwing et al., 2007), and mice (Kirino and Mourelatos, 2007; Ohara et al., 2007). However, 3′ ends with 2′-O-methyl groups are not a unique feature of piRNAs, as plant miRNAs and siRNAs and *Drosophila* siRNAs also carry this modification (Ghildiyal et al., 2008; Li et al., 2005).

The piRNAs of *Drosophila*, zebrafish, mice, and rats are 24–30 nucleotides in length, and as such are distinct from miRNAs or endogenous siRNAs (Klattenhoff and Theurkauf, 2008). The reason for the size difference between 21U-RNAs and piRNAs of other organisms is unclear. Another distinction of *C. elegans* piRNAs is their genomic location and organization: while

Table 1. PIWI Is Required to Inhibit Tc3 Transposition in the Germline

	unc-22::Tc1	unc-22::Tc3	unc-22::Tc4
Wild-type	<10^{-6} [a]	<10^{-6} [a]	<10^{-6} [a]
prg-1(n4357)	10^{-6}	10^{-4}	10^{-6}
prg-1(tm872)	n.d.	10^{-3}	n.d.
piwi(n4357; n4358)	10^{-6}	10^{-3}	10^{-6}
piwi(n4503; nDf57)	n.d.	10^{-3}	n.d.
mut-7	10^{-4}	10^{-4}	10^{-4}

Transposition rates were estimated by scoring *unc-22* reversion rates. Germline excision was verified by scoring the progeny of revertants. Animals were grown at 20°C.

[a] No *unc-22* reversions were observed. n.d., not determined.
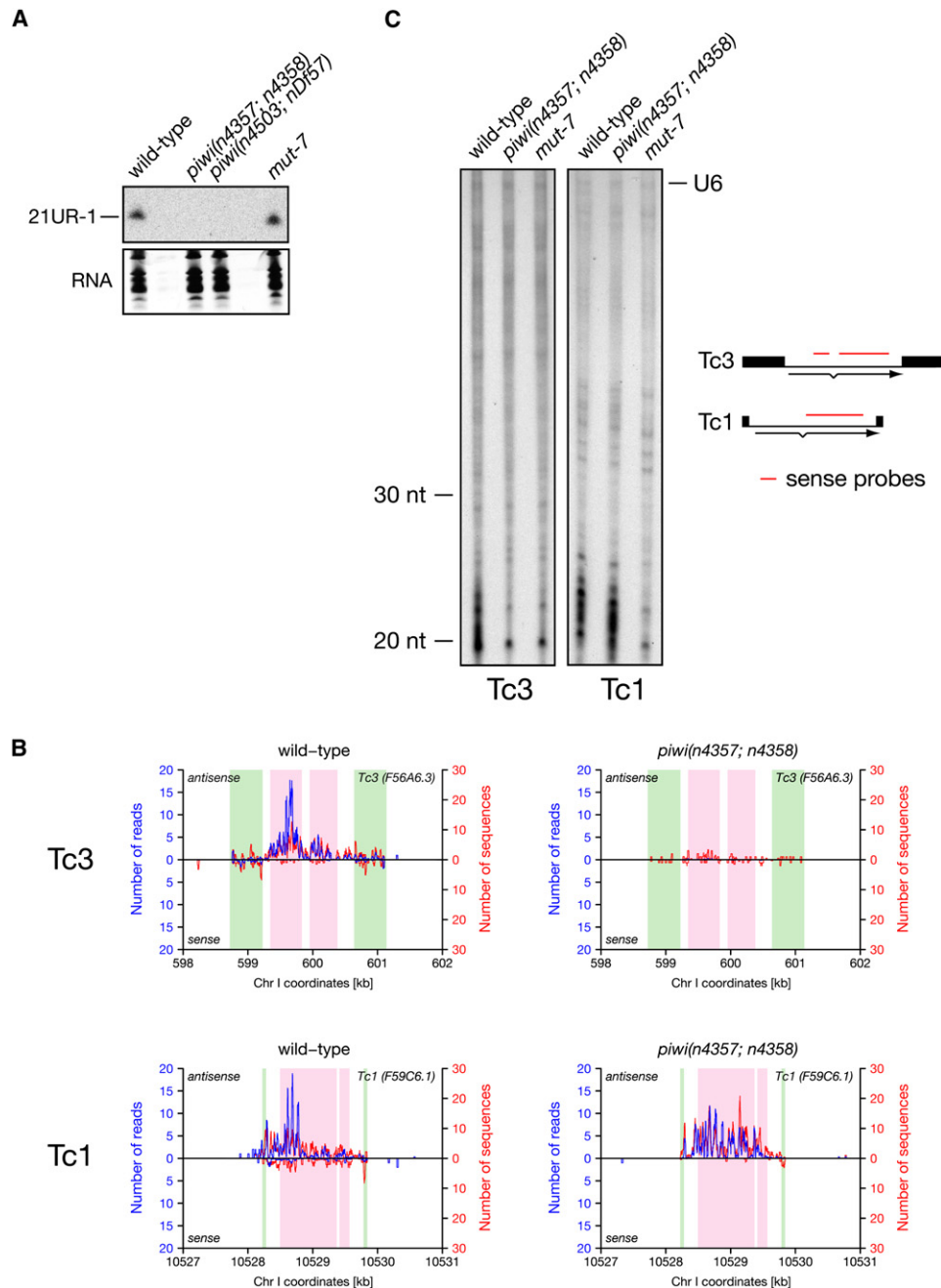
Figure 5. Piwi Acts Upstream of Endogenous siRNAs in Tc3 Silencing

(A) piRNA expression is independent of MUT-7. 21UR-1 northern blotting of total RNA of wild-type and mutant young adult *C. elegans*. Total RNA is shown as loading control (GelStar).

(B) Tc3-associated small RNAs are absent in *piwi* mutants. Small RNAs mapping to the loci of Tc3 (top) and Tc1 (bottom) on chromosome I as identified by high-throughput sequencing of 5′ independent wild-type (left) and *piwi* mutant (right) libraries. Inverted repeat and exon sequences are indicated in green and pink, respectively. The number of aligned sequence reads (blue) and number of aligned unique sequences (red) were plotted for each base pair position, with the top and bottom graph in each panel corresponding to the antisense and sense strand relative to the transposase transcript. Read and sequence counts were corrected for multiple alignments to the genome. The total number of reads from wild-type and *piwi* mutant libraries were comparable (2,963,895 and 3,017,027 of reads with perfect matches to the reference genome, respectively).

(C) Tc3 transposase antisense siRNAs are dramatically reduced in *piwi* mutants. RNase protection assay using sense fragments of Tc1 and Tc3 transposase. Sense siRNAs were not detected above background levels using this assay (see Figure S4 in the Supplemental Data). U6 was used as an internal control, but its concentration had to be titrated down due to some interference with small-RNA detection (Figure S5).
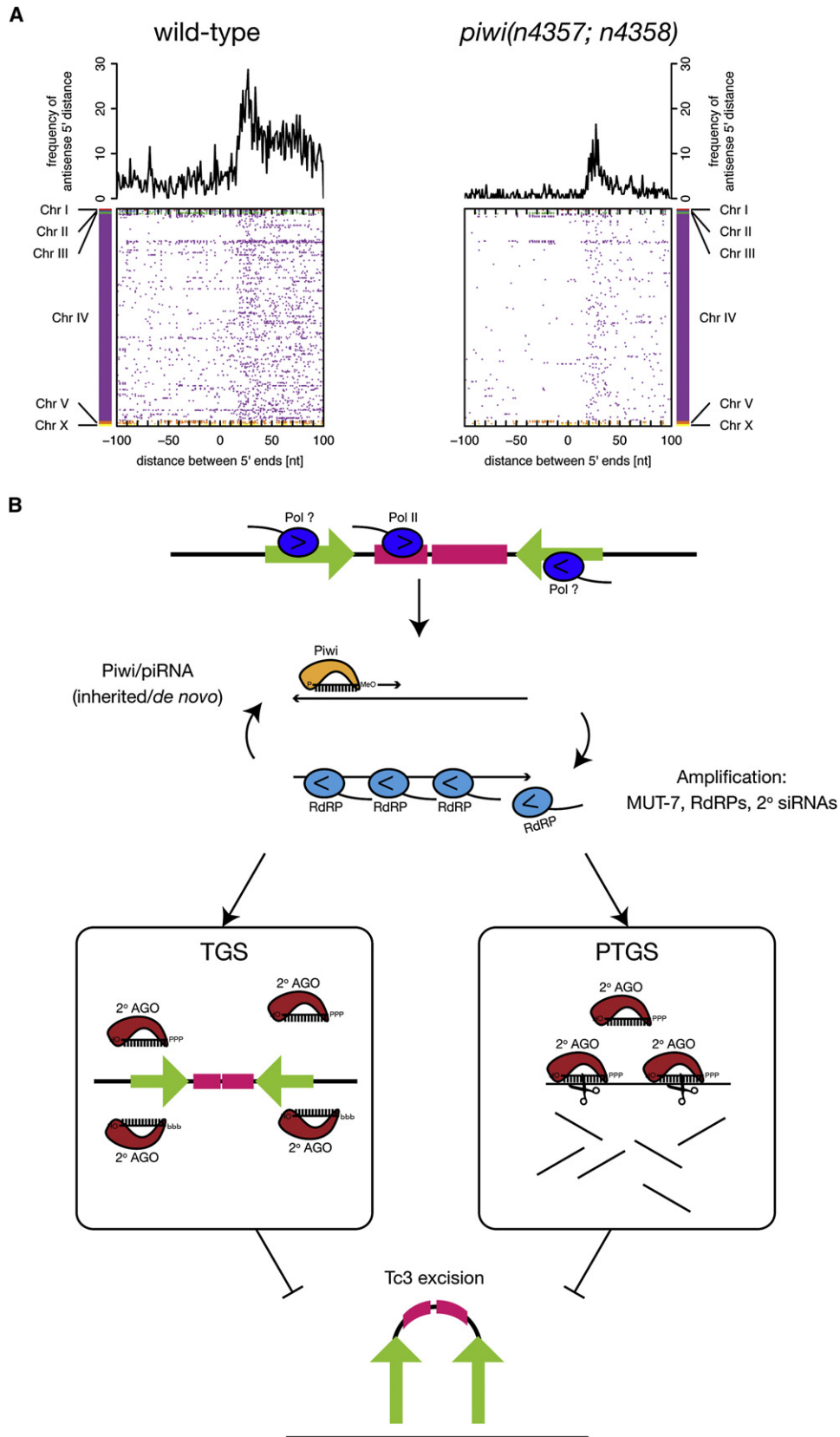
Drosophila and vertebrate piRNAs map to regions devoid of protein-coding genes (Klattenhoff and Theurkauf, 2008), some *C. elegans* piRNAs are interspersed with protein-coding genes (Ruby et al., 2006).

Biogenesis of piRNAs

The striking difference in genomic organization of *C. elegans* piRNA clusters compared to vertebrates and *Drosophila* might suggest a divergent mechanism of biogenesis. We found that *C. elegans* piRNAs are Dicer independent (Figure 2B), as are *Drosophila* and zebrafish piRNAs (Houwing et al., 2007; Vagin et al., 2006). Taken together with the strong local piRNA strand bias in all three species (Brennecke et al., 2007; Houwing et al., 2007; Ruby et al., 2006; Vagin et al., 2006), it is unlikely that piRNAs are generated through a double-stranded RNA intermediate. Instead, piRNAs might be generated directly as primary transcripts. In *C. elegans*, this could be achieved by an RdRP. However, two observations argue against this hypothesis. First, we found no evidence for the dependence of piRNA expression on RdRPs (Table 1), although we cannot exclude redundancy. Second, piRNAs lack 5′ triphosphates, a hallmark of other RdRP-derived small RNAs. We therefore favor a model in which piRNAs are derived from longer primary transcripts as is the case for miRNAs (Bartel, 2004). Interestingly, we show that Piwi is not essential for piRNA biogenesis, but that piRNA levels are dramatically reduced in *piwi* mutants (Figure 1E). This is analogous to *C. elegans* miRNAs, whose levels are reduced but are not absent in *alg-1; alg-2* Argonaute mutants (Grishok et al., 2001). We therefore postulate that Piwi is not an essential component of the piRNA biogenesis pathway but is required for piRNA accumulation or stability.

The piRNA genomic organization in *Drosophila* suggests that whole piRNA clusters could be transcribed as single primary transcripts and then processed into many mature piRNAs (Brennecke et al., 2007). Our finding that 21U-RNAs are piRNAs now suggests an alternative model. Many 21U-RNA loci are associated with a conserved upstream motif containing an eight nucleotide core consensus sequence CTGTTTCA (Ruby et al., 2006; see Supplemental Data). This motif is located approximately 38 bases upstream of the base corresponding to the 5′ uridine of the 21U-RNA. Ruby et al. have postulated that this motif is part of a 21U-RNA promoter. We conclude that piRNAs might be transcribed as individual transcripts in *C. elegans* and perhaps other species. Alternatively, these motifs might also be involved posttranscriptionally in piRNA biogenesis. Having linked a conserved motif to piRNAs in *C. elegans*, it will be important to search for such motifs in other species. Perhaps such a motif might be found only for a subset of "primary" piRNAs in *Drosophila* or vertebrates.

No Evidence for Ping-Pong in *C. elegans*

Recently, a compelling model for piRNA production in *Drosophila* named the Ping-Pong amplification loop has been put forward (Brennecke et al., 2007; Gunawardane et al., 2007). In this model Piwi/Aub is bound by a primary piRNA with a 5′ uridine. Target slicing by this complex followed by 3′ end processing results in the generation of a secondary piRNA that is antisense to the primary piRNA, overlaps by ten complementary nucleotides, and has an adenine at position 10. The loop is completed by AGO3, the third Piwi protein of *Drosophila*, which binds the secondary piRNA and, given a suitable template, in turn generates piRNAs with 5′ uridines. This model is supported by the finding that Piwi/Aub-bound piRNAs have a strong bias for 5′ uridine and AGO3-bound piRNAs have a strong bias for adenine at position 10 (Brennecke et al., 2007; Gunawardane et al., 2007). To test the Ping-Pong hypothesis in *C. elegans*, we analyzed the overlap among 21U-RNAs and between 21U-RNAs and other *C. elegans* small RNAs (Figure 6A). We found few cases of overlap and no adenine peak at position 10 (Figure S6). We conclude that there is currently no evidence for a role of Ping-Pong to generate piRNAs or to couple piRNAs to endogenous siRNA pathways in *C. elegans*.

Piwi and piRNAs Act Upstream of Secondary siRNA Pathways in Tc3 Silencing

We suggest that *C. elegans* and *Drosophila* share primary piRNAs, but while amplification of small RNAs in *Drosophila* might be achieved through a Ping-Pong amplification loop, in *C. elegans* amplification might be achieved through secondary siRNAs and RdRPs (Figure 6B). We demonstrate that Tc3 transposon silencing and associated siRNAs depend on Piwi and MUT-7. However, we cannot directly implicate a specific piRNA in Tc3 transposon silencing. One attractive candidate might be 21UR-139, which maps to the Tc3 transposase ORF and might act in *cis*. A second 21U-RNA also maps to the Tc3 locus (Batista et al., 2008). Both are sense relative to the transposase gene. We propose a speculative model to explain how sense piRNAs might be required for antisense siRNA production at the Tc3 locus. Piwi loaded with sense piRNAs might stimulate RdRP activity using Tc3 antisense transcripts as templates, in a manner analogous to siRNAs in RNAi amplification (Baulcombe, 2007). The sense transcripts would be targeted by antisense siRNAs and act as a template for a second round of small RNA-stimulated RdRP activity (Figure 6B). Such a loop would maintain high levels of

---

Figure 6. Piwi and piRNAs May Act Upstream of siRNAs

(A) Small RNAs mapping to the opposite strand of nearby 21U-RNA loci show a preference for locations downstream of the 21U-RNA locus and were reduced in the *piwi* mutant (right) as compared to wild-type (left). Proximate small RNAs on the same strand as the 21U-RNA were also reduced in the *piwi* mutant (Figure S7). (Bottom) Rows correspond to 6021 21U-RNA loci, ordered by genomic position with colors representing different chromosomes. For a given row (21U-RNA locus), dots correspond to the relative position of nearby antisense small RNAs as defined by the distance of the 5′ end of the cloned small RNA relative to the 21U-RNA 5′ end. (Top) Shown is the frequency of distances between 5′ ends of 21U-RNAs and antisense small RNAs. Frequencies were based on the normalized number of loci for 5′ unique sequences.

(B) A speculative model of the role of Piwi in Tc3 silencing. TGS, transcriptional gene silencing. PTGS, posttranscriptional gene silencing. See Discussion for an explanation.

Tc3 siRNAs. In the case of the Tc3 transposase, this might lead to PTGS against the mRNA. This is in agreement with the low number of siRNAs mapping to the intron as compared to the two exons of the Tc3 transposase (Figure 5B). In the case of the Tc3 TIR, this might involve chromatin-mediated TGS. If the TIR-associated siRNAs are functionally distinct from the siRNAs mapping to the Tc3 transposase ORF, these two pathways might be separable. Indeed, Batista et al. (2008) observed that *prg-1* mutants lack only TIR-associated siRNAs, but not siRNAs associated with the transposase ORF (Batista et al., 2008). Together, these data suggest that PRG-1 and PRG-2 might have functionally distinct roles in Tc3 silencing. This is particularly intriguing as PRG-1, but not PRG-2, is required for piRNA accumulation (Figure 2A).

The amplification loop we propose is sequence independent, given an initiating 21U-RNA. Therefore the link between piRNA and siRNA pathways might not be restricted to Tc3. Indeed, we observe a positional bias for antisense siRNAs close to 21U-RNAs to map downstream of the 21U-RNA locus (Figure 6A). piRNA and siRNA pathways might be linked in other species too. *Drosophila* and vertebrates have lost secondary siRNAs and RdRPs, which are found in plants, yeast, and *C. elegans*. However, recently endogenous siRNAs in mouse oocytes were shown to map to the same loci as piRNAs (Tam et al., 2008).

Other Roles of Piwi in *C. elegans*?

We demonstrate a role for Piwi in inhibiting Tc3 mobility. However, it seems unlikely that the defects in germline development observed in *piwi* mutants are solely due to Tc3 activity. It has been proposed that mobile element excision and excessive DNA breaks result in meiotic catastrophe in *piwi* mutants in *Drosophila* (Klattenhoff and Theurkauf, 2008). We cannot exclude the possibility that other mobile elements in addition to Tc3 are hyperactivated in *piwi* mutants in *C. elegans*, with similar consequences. Alternatively, Piwi/piRNAs might be involved in gene regulation during germline development in *C. elegans*. Surveying whole-animal gene expression in wild-type and *piwi* mutant adult *C. elegans* using Affymetrix mRNA profiling, we did not observe striking changes in gene expression (data not shown). However, changes might only be revealed by studying gene expression in germ cells. Indeed, a recent study suggests that the expression of a subset of mRNAs expressed during spermatogenesis is deregulated in male gonads isolated from *prg-1* mutants (Wang and Reinke, 2008). The striking clustering of piRNAs on chromosome IV (Ruby et al., 2006) raises a number of interesting questions. Do piRNAs act in *cis*, or do they act in *trans*, analogous to miRNA function? If piRNAs act in *trans*, what are the rules for piRNA:piRNA target interaction? Do piRNAs always act through siRNA pathways as we have demonstrated for Tc3? Having identified piRNAs in *C. elegans*, we hope to address these questions using the powerful genetic tools that this organism offers.

EXPERIMENTAL PROCEDURES

C. elegans Strains and Culture

See the Supplemental Data and Tables S3 and S4 for strain information and culture conditions.

High-Throughput Sequencing and Data Analysis

We generated 5′-dependent libraries for N2 (2,719,949 reads), *piwi(n4357; n4358)* (764,960 reads), *mut-7* (3,475,722 reads), and *dcr-1* (149,997 reads) and 5′-independent libraries for N2 (2,963,895 reads) and *piwi(n4357; n4358)* (3,017,027 reads) that were sequenced using the Illumina/Solexa platform. See the Supplemental Data for additional information.

RNAi Experiments

RNAi experiments were carried out as reported previously (Fire et al., 1998; Timmons et al., 2001). Bacterial strains carrying plasmids expressing double-stranded RNA for the gene of interest were obtained from the Ahringer Laboratory genome-wide RNA library (Fraser et al., 2000; Kamath et al., 2003). All RNAi library inserts were confirmed by sequencing.

Transposon Excision Experiments

Transposon excision assays were carried out as described previously (Ketting et al., 1999).

SUPPLEMENTAL DATA

The Supplemental Data include Supplemental Experimental Procedures, eight figures, and five tables and can be found with this article online at http://www.molecule.org/cgi/content/full/31/1/79/DC1/.

REFERENCES

Ambros, V., Lee, R.C., Lavanway, A., Williams, P.T., and Jewell, D. (2003). MicroRNAs and other tiny endogenous RNAs in *C. elegans*. Curr. Biol. *13*, 807–818.

Aravin, A.A., Naumova, N.M., Tulin, A.V., Vagin, V.V., Rozovsky, Y.M., and Gvozdev, V.A. (2001). Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the D. melanogaster germline. Curr. Biol. *11*, 1017–1027.

Aravin, A., Gaidatzis, D., Pfeffer, S., Lagos-Quintana, M., Landgraf, P., Iovino, N., Morris, P., Brownstein, M.J., Kuramochi-Miyagawa, S., Nakano, T., et al. (2006). A novel class of small RNAs bind to MILI protein in mouse testes. Nature *442*, 203–207.

Bartel, D.P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. Cell *116*, 281–297.

Barton, M.K., Schedl, T.B., and Kimble, J. (1987). Gain-of-function mutations of fem-3, a sex-determination gene in Caenorhabditis elegans. Genetics *115*, 107–119.

Batista, P.J., Ruby, J.G., Claycomb, J.M., Chiang, R., Fahlgren, N., Kasschau, K.D., Chaves, D.A., Gu, W., Vasale, J.J., Duan, S., et al. (2008). PRG-1 and 21U-RNAs interact to form the piRNA complex required for fertility in *C. elegans*. Mol. Cell *31*, this issue, 67–78.

Baulcombe, D.C. (2007). Molecular biology. Amplified silencing. Science *315*, 199–200.

Beanan, M.J., and Strome, S. (1992). Characterization of a germ-line proliferation mutation in *C. elegans*. Development *116*, 755–766.

Brennecke, J., Aravin, A.A., Stark, A., Dus, M., Kellis, M., Sachidanandam, R., and Hannon, G.J. (2007). Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. Cell *128*, 1089–1103.

Carmell, M.A., Xuan, Z., Zhang, M.Q., and Hannon, G.J. (2002). The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis. Genes Dev. *16*, 2733–2742.

Carmell, M.A., Girard, A., van de Kant, H.J., Bourc'his, D., Bestor, T.H., de Rooij, D.G., and Hannon, G.J. (2007). MIWI2 is essential for spermatogenesis and repression of transposons in the mouse male germline. Dev. Cell *12*, 503–514.

C. elegans Sequencing Consortium (1998). Genome sequence of the nematode C. elegans: a platform for investigating biology. Science *282*, 2012–2018.

Cerutti, L., Mian, N., and Bateman, A. (2000). Domains in gene silencing and cell differentiation proteins: the novel PAZ domain and redefinition of the Piwi domain. Trends Biochem. Sci. *25*, 481–482.

Chen, Y., Pane, A., and Schüpbach, T. (2007). Cutoff and aubergine mutations result in retrotransposon upregulation and checkpoint activation in Drosophila. Curr. Biol. *17*, 637–642.

Cox, D.N., Chao, A., Baker, J., Chang, L., Qiao, D., and Lin, H. (1998). A novel class of evolutionarily conserved genes defined by piwi are essential for stem cell self-renewal. Genes Dev. *12*, 3715–3727.

Cox, D.N., Chao, A., and Lin, H. (2000). piwi encodes a nucleoplasmic factor whose activity modulates the number and division rate of germline stem cells. Development *127*, 503–514.

Deng, W., and Lin, H. (2002). miwi, a murine homolog of piwi, encodes a cytoplasmic protein essential for spermatogenesis. Dev. Cell *2*, 819–830.

Desset, S., Meignin, C., Dastugue, B., and Vaury, C. (2003). COM, a heterochromatic locus governing the control of independent endogenous retroviruses from Drosophila melanogaster. Genetics *164*, 501–509.

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

Fraser, A.G., Kamath, R.S., Zipperlen, P., Martinez-Campos, M., Sohrmann, M., and Ahringer, J. (2000). Functional genomic analysis of C. elegans chromosome I by systematic RNA interference. Nature *408*, 325–330.

Ghildiyal, M., Seitz, H., Horwich, M.D., Li, C., Du, T., Lee, S., Xu, J., Kittler, E.L., Zapp, M.L., Weng, Z., et al. (2008). Endogenous siRNAs derived from transposons and mRNAs in Drosophila somatic cells. Science *320*, 1077–1081.

Girard, A., Sachidanandam, R., Hannon, G.J., and Carmell, M.A. (2006). A germline-specific class of small RNAs binds mammalian Piwi proteins. Nature *442*, 199–202.

Grishok, A., Pasquinelli, A.E., Conte, D., Li, N., Parrish, S., Ha, I., Baillie, D.L., Fire, A., Ruvkun, G., and Mello, C.C. (2001). Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control *C. elegans* developmental timing. Cell *106*, 23–34.

Grivna, S.T., Beyret, E., Wang, Z., and Lin, H. (2006). A novel class of small RNAs in mouse spermatogenic cells. Genes Dev. *20*, 1709–1714.

Gunawardane, L.S., Saito, K., Nishida, K.M., Miyoshi, K., Kawamura, Y., Nagami, T., Siomi, H., and Siomi, M.C. (2007). A slicer-mediated mechanism for repeat-associated siRNA 5′ end formation in Drosophila. Science *315*, 1587–1590.

Horwich, M.D., Li, C., Matranga, C., Vagin, V., Farley, G., Wang, P., and Zamore, P.D. (2007). The *Drosophila* RNA methyltransferase, DmHen1, modifies germline piRNAs and single-stranded siRNAs in RISC. Curr. Biol. *17*, 1265–1272.

Houwing, S., Kamminga, L.M., Berezikov, E., Cronembold, D., Girard, A., van den Elst, H., Filippov, D.V., Blaser, H., Raz, E., Moens, C.B., et al. (2007). A role for Piwi and piRNAs in germ cell maintenance and transposon silencing in Zebrafish. Cell *129*, 69–82.

Joshua-Tor, L. (2006). The Argonautes. Cold Spring Harb. Symp. Quant. Biol. *71*, 67–72.

Kamath, R.S., Fraser, A.G., Dong, Y., Poulin, G., Durbin, R., Gotta, M., Kanapin, A., Le Bot, N., Moreno, S., Sohrmann, M., et al. (2003). Systematic functional analysis of the *Caenorhabditis elegans* genome using RNAi. Nature *421*, 231–237.

Ketting, R.F., Haverkamp, T.H., van Luenen, H.G., and Plasterk, R.H. (1999). Mut-7 of *C. elegans*, required for transposon silencing and RNA interference, is a homolog of Werner syndrome helicase and RNaseD. Cell *99*, 133–141.

Ketting, R.F., Fischer, S.E., Bernstein, E., Sijen, T., Hannon, G.J., and Plasterk, R.H. (2001). Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans. Genes Dev. *15*, 2654–2659.

Kim, J.K., Gabel, H.W., Kamath, R.S., Tewari, M., Pasquinelli, A., Rual, J.F., Kennedy, S., Dybbs, M., Bertin, N., Kaplan, J.M., et al. (2005). Functional genomic analysis of RNA interference in C. elegans. Science *308*, 1164–1167.

Kimble, J., Edgar, L., and Hirsh, D. (1984). Specification of male development in Caenorhabditis elegans: the fem genes. Dev. Biol. *105*, 234–239.

Kirino, Y., and Mourelatos, Z. (2007). Mouse Piwi-interacting RNAs are 2′-O-methylated at their 3′ termini. Nat. Struct. Mol. Biol. *14*, 347–348.

Klattenhoff, C., Bratu, D.P., McGinnis-Schultz, N., Koppetsch, B.S., Cook, H.A., and Theurkauf, W.E. (2007). *Drosophila* rasiRNA pathway mutations disrupt embryonic axis specification through activation of an ATR/Chk2 DNA damage response. Dev. Cell *12*, 45–55.

Klattenhoff, C., and Theurkauf, W. (2008). Biogenesis and germline functions of piRNAs. Development *135*, 3–9.

Knight, S.W., and Bass, B.L. (2001). A role for the RNase III enzyme DCR-1 in RNA interference and germ line development in *Caenorhabditis elegans*. Science *293*, 2269–2271.

Kuramochi-Miyagawa, S., Kimura, T., Yomogida, K., Kuroiwa, A., Tadokoro, Y., Fujita, Y., Sato, M., Matsuda, Y., and Nakano, T. (2001). Two mouse piwi-related genes: miwi and mili. Mech. Dev. *108*, 121–133.

Kuramochi-Miyagawa, S., Kimura, T., Ijiri, T.W., Isobe, T., Asada, N., Fujita, Y., Ikawa, M., Iwai, N., Okabe, M., Deng, W., et al. (2004). Mili, a mammalian member of piwi family gene, is essential for spermatogenesis. Development *131*, 839–849.

Lau, N.C., Lim, L.P., Weinstein, E.G., and Bartel, D.P. (2001). An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. Science *294*, 858–862.

Lau, N.C., Seto, A.G., Kim, J., Kuramochi-Miyagawa, S., Nakano, T., Bartel, D.P., and Kingston, R.E. (2006). Characterization of the piRNA complex from rat testes. Science *313*, 363–367.

Lee, R.C., and Ambros, V.R. (2001). An extensive class of small RNAs in Caenorhabditis elegans. Science *294*, 862–864.

Li, J., Yang, Z., Yu, B., Liu, J., and Chen, X. (2005). Methylation protects miRNAs and siRNAs from a 3′-end uridylation activity in Arabidopsis. Curr. Biol. *15*, 1501–1507.

Lim, L.P., Lau, N.C., Weinstein, E.G., Abdelhakim, A., Yekta, S., Rhoades, M.W., Burge, C.B., and Bartel, D.P. (2003). The microRNAs of *Caenorhabditis elegans*. Genes Dev. *17*, 991–1008.

Lin, H., and Spradling, A.C. (1997). A novel group of pumilio mutations affects the asymmetric division of germline stem cells in the Drosophila ovary. Development *124*, 2463–2476.

Mochizuki, K., Fine, N.A., Fujisawa, T., and Gorovsky, M.A. (2002). Analysis of a piwi-related gene implicates small RNAs in genome rearrangement in tetrahymena. Cell *110*, 689–699.

Moerman, D.G., and Waterston, R.H. (1984). Spontaneous unstable unc-22 IV mutations in C. elegans var. Bergerac. Genetics *108*, 859–877.

Nelson, G.A., Lew, K.K., and Ward, S. (1978). Intersex, a temperature-sensitive mutant of the nematode Caenorhabditis elegans. Dev. Biol. *66*, 386–409.

Ohara, T., Sakaguchi, Y., Suzuki, T., Ueda, H., Miyauchi, K., and Suzuki, T. (2007). The 3′ termini of mouse Piwi-interacting RNAs are 2′-O-methylated. Nat. Struct. Mol. Biol. *14*, 349–350.

Pak, J., and Fire, A. (2007). Distinct populations of primary and secondary effectors during RNAi in *C. elegans*. Science *315*, 241–244.

Pane, A., Wehr, K., and Schüpbach, T. (2007). zucchini and squash encode two putative nucleases required for rasiRNA production in the *Drosophila* germline. Dev. Cell *12*, 851–862.

Pepper, A.S., Killian, D.J., and Hubbard, E.J. (2003). Genetic analysis of Caenorhabditis elegans glp-1 mutants suggests receptor interaction or competition. Genetics *163*, 115–132.

Prud'homme, N., Gans, M., Masson, M., Terzian, C., and Bucheton, A. (1995). Flamenco, a gene controlling the gypsy retrovirus of Drosophila melanogaster. Genetics *139*, 697–711.

Reddien, P.W., Oviedo, N.J., Jennings, J.R., Jenkin, J.C., and Sánchez Alvarado, A. (2005). SMEDWI-2 is a PIWI-like protein that regulates planarian stem cells. Science *310*, 1327–1330.

Ruby, J.G., Jan, C., Player, C., Axtell, M.J., Lee, W., Nusbaum, C., Ge, H., and Bartel, D.P. (2006). Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in *C. elegans*. Cell *127*, 1193–1207.

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

Saito, K., Sakaguchi, Y., Suzuki, T., Suzuki, T., Siomi, H., and Siomi, M.C. (2007). Pimet, the Drosophila homolog of HEN1, mediates 2′-O-methylation of Piwi- interacting RNAs at their 3′ ends. Genes Dev. *21*, 1603–1608.

Savitsky, M., Kwon, D., Georgiev, P., Kalmykova, A., and Gvozdev, V. (2006). Telomere elongation is under the control of the RNAi-based mechanism in the Drosophila germline. Genes Dev. *20*, 345–354.

Sijen, T., and Plasterk, R.H. (2003). Transposon silencing in the Caenorhabditis elegans germ line by natural RNAi. Nature *426*, 310–314.

Sijen, T., Steiner, F.A., Thijssen, K.L., and Plasterk, R.H. (2007). Secondary siRNAs result from unprimed RNA synthesis and form a distinct class. Science *315*, 244–247.

Simmer, F., Tijsterman, M., Parrish, S., Koushika, S.P., Nonet, M.L., Fire, A., Ahringer, J., and Plasterk, R.H. (2002). Loss of the putative RNA-directed RNA polymerase RRF-3 makes *C. elegans* hypersensitive to RNAi. Curr. Biol. *12*, 1317–1319.

Smardon, A., Spoerke, J.M., Stacey, S.C., Klein, M.E., Mackin, N., and Maine, E.M. (2000). EGO-1 is related to RNA-directed RNA polymerase and functions in germ-line development and RNA interference in *C. elegans*. Curr. Biol. *10*, 169–178.

Tabara, H., Sarkissian, M., Kelly, W.G., Fleenor, J., Grishok, A., Timmons, L., Fire, A., and Mello, C.C. (1999). The rde-1 gene, RNA interference, and transposon silencing in *C. elegans*. Cell *99*, 123–132.

Tam, O.H., Aravin, A.A., Stein, P., Girard, A., Murchison, E.P., Cheloufi, S., Hodges, E., Anger, M., Sachidanandam, R., Schultz, R.M., et al. (2008). Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. Nature *453*, 534–538.

Timmons, L., Court, D.L., and Fire, A. (2001). Ingestion of bacterially expressed dsRNAs can produce specific and potent genetic interference in Caenorhabditis elegans. Gene *263*, 103–112.

Vagin, V.V., Sigova, A., Li, C., Seitz, H., Gvozdev, V., and Zamore, P.D. (2006). A distinct small RNA pathway silences selfish genetic elements in the germline. Science *313*, 320–324.

van Luenen, H.G., Colloms, S.D., and Plasterk, R.H. (1993). Mobilization of quiet, endogenous Tc3 transposons of Caenorhabditis elegans by forced expression of Tc3 transposase. EMBO J. *12*, 2513–2520.

van Luenen, H.G., Colloms, S.D., and Plasterk, R.H. (1994). The mechanism of transposition of Tc3 in *C. elegans*. Cell *79*, 293–301.

Vos, J.C., van Luenen, H.G., and Plasterk, R.H. (1993). Characterization of the Caenorhabditis elegans Tc1 transposase in vivo and in vitro. Genes Dev. *7*, 1244–1253.

Vos, J.C., De Baere, I., and Plasterk, R.H. (1996). Transposase is the only nematode protein required for in vitro transposition of Tc1. Genes Dev. *10*, 755–761.

Wang, G., and Reinke, V. (2008). A *C. elegans* Piwi, PRG-1, regulates 21U-RNAs during spermatogenesis. Curr. Biol. *18*, in press. Published online May 22, 2008. 10.1016/j.cub.2008.05.009.

Watanabe, T., Takeda, A., Tsukiyama, T., Mise, K., Okuno, T., Sasaki, H., Minami, N., and Imai, H. (2006). Identification and characterization of two novel classes of small RNAs in the mouse germline: retrotransposon-derived siRNAs in oocytes and germline small RNAs in testes. Genes Dev. *20*, 1732–1743.

Yigit, E., Batista, P.J., Bei, Y., Pang, K.M., Chen, C.C., Tolia, N.H., Joshua-Tor, L., Mitani, S., Simard, M.J., and Mello, C.C. (2006). Analysis of the *C. elegans* Argonaute family reveals that distinct Argonautes act sequentially during RNAi. Cell *127*, 747–757.