Transcriptional regulation in the context of DNA replication by RNA interference in fission yeast

Stephane E. Castel

Watson School of Biological Sciences

Cold Spring Harbor Laboratory

A thesis submitted for the degree of

PhilosophiæDoctor (Ph.D.)

June 19th, 2014

Thesis Committee

Thesis Advisor – Robert Martienssen

Committee Chair – Jim Hicks (CSHL)

Academic Advisor – Lloyd Trotman (CHSL)

External Member – Matthew O'Connell (Icahn School of Medicine at Mount Sinai, NY, USA)

External Examiner – Benoit Arcangioli (Institut Pasteur, Paris, France)

Defense Date: June 19th, 2014

Thesis Abstract

RNAi is a highly conserved mechanism of both post transcriptional and transcriptional silencing. More recently a role for RNAi in transcriptional termination has emerged. We have found that this function is of the utmost importance at the centromere, where it occurs in the context of DNA replication to remove stalled Pol II and facilitate epigenetic modification. Outside of the centromere Pol II termination by Dicer occurs at a variety of genomic loci that have in common transcription and replication collision. One striking example of this regulation occurs within the repetitive rDNA repeats. We used this locus as a model to study the interplay between transcription, DNA replication, and Dicer. In the absence of Dicer DNA damage increased and genomic integrity at rDNA repeats was compromised, resulting in a loss of rDNA copy number. This loss was enhanced during meiosis, suggesting it occurs via homologous recombination. We suggest that transcriptional termination by RNAi pathways in *S. pombe* occurs specifically at sites where replication and transcription compete. RNA Pol II release by RNAi is conserved in eukaryotes, so it's possible that the unique context of its action is also conserved.

Table of Contents

1. Introduction: nuclear RNAi pathways play a diverse role in transcriptional regulation across eukaryotes

This chapter has been reproduced with modifications from:

> Castel SE, Martienssen RA. RNA interference in the nucleus: roles for small RNAs in transcription, epigenetics and beyond. *Nature Reviews Genetics* 14, 100–112 (2013).

Since the discovery that double stranded RNA (dsRNA) can robustly silence genes in *C. elegans* and plants, RNAi has become a new paradigm for understanding gene regulation. The mechanism is well conserved across model organisms and utilizes short antisense RNA to inhibit translation, or to degrade cytoplasmic mRNA by post-transcriptional gene silencing (PTGS). PTGS protects against viral infection, prevents transposon mobilization, and regulates endogenous genes. Three classes of small RNA can regulate genes by targeting transcripts in the cytoplasm. These are miRNA that are hairpin-derived with imperfect complementarity to target transcripts and cause translational repression, small interfering RNA (siRNA) with perfect complementarity to targets and cause transcript degradation, and piwi RNA (piRNA) which target transposon transcripts in animal germlines. Traditionally the term RNAi has been used to describe siRNA pathways, however the mechanistic details of diverse small RNA pathways are converging, so in this review we use RNAi as an umbrella term to describe silencing that is dependent on small RNA.

In plants and fungi, RNAi pathways in the nucleus can repress target genes at the transcriptional level by guiding epigenetic modification of chromatin, for example via histone and DNA methyltransferases. At first these pathways were thought to be absent from metazoans, but recently, a parallel mechanism has been found in the germline. These findings have revealed a conserved nuclear role for RNAi in transcriptional gene silencing (TGS), and because it occurs in the germline, TGS can lead to transgenerational inheritance in absence of the initiating RNA, but dependent on endogenously produced small RNA. Such epigenetic inheritance is familiar in plants, but only recently described in metazoans.

1.1 Biogenesis of Nuclear Small RNA

The siRNA and piRNA pathways differ in the source of the primary RNA that elicits a response and the mechanism by which small RNA is subsequently produced and amplified. The Argonaute family of proteins are the effectors of RNAi and this family consists of two subclades: Ago proteins which are ubiquitously expressed and bind miRNA and siRNA, and Piwi proteins which were originally discovered in the germline and bind piRNA (Aravin et al., 2006; Cox et al., 1998; Kim et al., 2009).

1.1.1 siRNA Biogenesis

Double stranded RNA (dsRNA) is thought to be the trigger for most if not all siRNA biogenesis and can be generated by several means (Figure 1.1). Once dsRNA is available the biogenesis of siRNA requires action of the RNAse III like Dicer family of enzymes. Dicer cleaves dsRNA into 20-25nt siRNA duplexes with 2nt 3' OH overhangs (Colmenares et al., 2007) and 5' monophosphates (Colmenares et al., 2007; Pak and Fire, 2007). Dicer independent mechanisms of siRNA production have also been proposed in *Neurospora* (Lee et al., 2010), *S. pombe* (Halic and Moazed, 2010) and *C. elegans* (Aoki et al., 2007). The cellular location in which dsRNA processing occurs has implications for how siRNA biogenesis and nuclear effects are regulated. In *S. pombe* transcription, processing, RdRP amplification, and Ago mediated target cleavage are all intimately linked in the nucleus (Figure 1.1a) (Barraud et al., 2011; Colmenares et al., 2007; Emmerth et al., 2010; Irvine et al., 2006; Kato et al., 2005). In animals siRNA processing was originally thought to occur in the cytoplasm (Kim et al., 2009) however recent studies in *Drosophila* have shown that DCR2 is found predominantly in the nucleus challenging this view (Cernilogar et al., 2011). This is in contrast to *C. elegans* where in depth studies have validated the cytoplasmic processing of many siRNA pathways (Dalzell et al., 2011) (Figure 1.1b).

Once generated the siRNA duplexes are loaded into an appropriate effector Argonaute protein. The subcellular location where Argonaute loading takes place is not yet fully understood across model organisms. In Arabidopsis nuclear AGO4 loading is cytoplasmic and mediated by the heat shock protein HSP90, after which it is then imported into the nucleus (Ye et al., 2012). A requirement for

HSP90 in Ago loading has also been observed in *Drosophila* however where this process occurs is not known (Miyoshi et al., 2010). Like *Arabidopsis*, siRNA processing is nuclear in *S. pombe*, however it is not known where Ago1 loading occurs. The *C. elegans* nuclear Argonaute NRDE-3 is imported into the nucleus only when loaded with secondary siRNA that is produced in the cytoplasm (Guang et al., 2008). If cytoplasmic loading of Argonaute proteins is conserved across species this would have important implications for the regulation of nuclear RNAi.

1.1.2 piRNA biogenesis

The biogenesis of piRNA primarily occurs via a process known as the ping-pong cycle (Figure 1.1c) initially described in the *Drosophila* germline (Brennecke et al., 2007; Gunawardane et al., 2007). First, piRNA genomic clusters are transcribed to produce primary piRNA precursors. In the cytoplasm an unknown mechanism processes primary piRNA precursors into short 23-29nt antisense piRNA with a strong 5' uridine bias. These short ssRNA are loaded into the Piwi family Argonaute proteins Aub and Piwi. In the cytoplasm, the loaded Aub/Piwi then targets mRNA of active transposons for cleavage; this produces sense piRNA, which have a strong adenine bias at position 10. The sense piRNA is loaded into the Piwi family member Ago3, which then directs cleavage of primary piRNA precursors and the subsequent production of more antisense piRNA, completing the ping-pong cycle (Gunawardane et al., 2007). In the female germline Aub protein is restricted to the cytoplasm whereas Piwi is predominantly nuclear, indicating that Aub plays a larger role in the ping-pong cycle (Li et al., 2009). The nuclear localization of Piwi is lost in Ago3 mutants, suggesting that once loaded with piRNA produced by the ping-pong cycle Piwi is imported into the nucleus (Li et al., 2009). A less understood ping-pong independent piRNA biogenesis pathway operates in the somatic follicle cells that surround the female oocytes that is Piwi-dependent and Aub/Ago3 independent (Figure 1.4c) (Li et al., 2009; Malone et al., 2009).

Stability and turnover play an important role in the regulation of both the siRNA and piRNA pathways. Methylation of small RNA is a major determinant of their stability. Both piRNA and siRNA are 2'-O-methylated by the enzyme Hen1 across organisms (Ji and Chen, 2012). This methylation protects small RNA from both 3' uridylation and 3' truncation, which cause small RNA degradation and turnover.

The specificity of Hen1 could therefor contribute to cell-type specific small RNA profiles, and thus determine targets of RNAi, however such a mechanism has yet been uncovered.

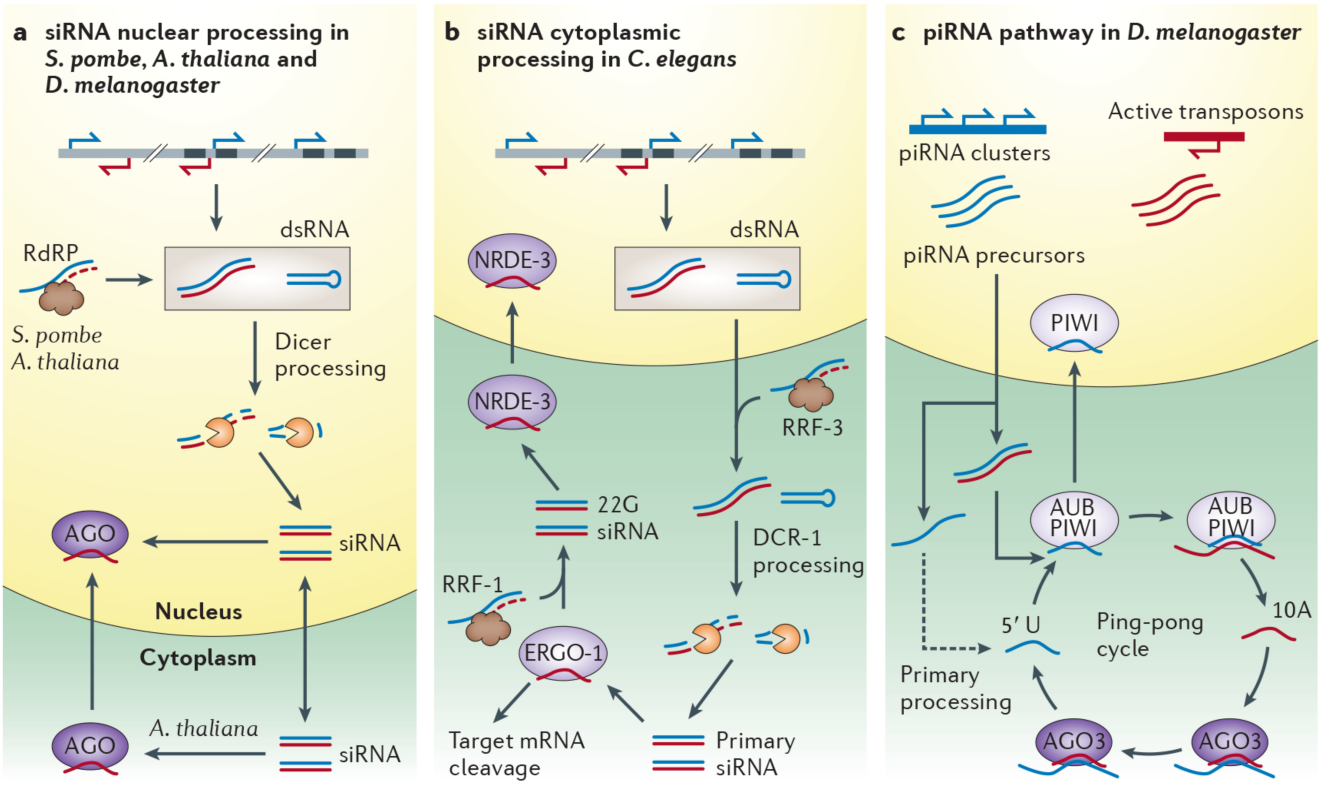

Figure 1.1 | Generalized pathways depicting the biogenesis of nuclear small RNA. A) siRNA processing takes place in the nucleus in *S. pombe* and *Drosophila* and the nucleoulus in *Arabidopsis*. dsRNA can be produced by convergent transcription, complementary transcripts, structured loci, or by RdRP activity in *Arabidopsis* and *S. pombe*. Dicer proteins generate siRNA that is loaded into an Argnoaute protein. In *Arabidopsis* siRNA are transported to the cytoplasm where Argonaute is loaded and then imported into the nucleus. B) In *C. elegans* siRNA processing occurs in the cytoplasm in a two-step fashion. Primary trigger dsRNA arises from nuclear transcription or the RdRP activity of RRF-3, which acts on transcripts in the cytoplasm. Primary processing by DCR-1 produces primary 26 nt siRNA which are loaded into ERGO-1. Loaded ERGO-1 can both facilitate PTGS in the cytoplasm and with RRF-1 generate secondary 22G siRNA siRNA. Secondary 22G siRNA is loaded into the nuclear Argonaute NRDE-3 in the cytoplasm that is then transported into the nucleus. C) piRNA biogenesis via the ping-pong cycle in the *Drosophila* female germline. Primary precursor piRNA antisense to active transposons (blue) is transcribed from heterochromatic piRNA clusters, and sense mRNA from active transposons (pink). In the cytoplasm primary processing generates antisense piRNA from primary precursor that is then

loaded into Aub or Piwi and cleaves sense transposon mRNA to produce sense piRNA. Additional antisense piRNA is produced by Ago3 mediated cleavage of antisense primary piRNA transcripts, completing the cycle. Only loaded Piwi is imported into the nucleus.

1.2 Mechanisms of Nuclear RNAi

Transcriptional Gene Silencing (TGS) was the first function of nuclear RNAi to be discovered, and refers to the process by which RNAi can reduce transcription by guiding localized heterochromatin formation at target genomic loci. A question that arises from this mechanism is how sequence specific targeting of chromatin modifications is achieved? As in the cytoplasm the substrate for nuclear RNAi has been shown to be RNA molecules, but these must be in close proximity to the locus they arose from so that epigenetic modification can be specific. This has lead to a model of co-transcriptional gene silencing (CTGS), whereby nuclear small RNA target nascent RNA molecules from RNA polymerases, and the effector complexes themselves interact with and regulate transcriptional machinery. The two examples of nuclear RNAi described here in mechanistic detail reveal that positive feedback loops are involved in chromatin modification. The nuclear RNAi complexes themselves are both attracted to repressive epigenetics marks, and deposit them, creating robust silencing at target loci.

1.2.1 Nuclear RNAi in S. pombe: TGS

A role for RNAi in TGS was identified in *S. pombe* where it is required for the formation of constitutive heterochromatin at pericentromeres. They are highly enriched for H3K9 methylation (H3K9me) and are composed of varying numbers of repeat units that are bi-directionally transcribed to form dsRNA that is then processed by Dcr1 into siRNA (Volpe et al., 2002). The RNA dependent RNA polymerase complex (RDRC) interacts with both Dcr1 (Colmenares et al., 2007) and Ago1 (Motamedi et al., 2004) to produce dsRNA and siRNA from Ago1 targeted transcripts and amplify the siRNA response. siRNA are loaded into Ago1, the principle member of the RNA Induced Transcriptional Silencing Complex (RITSC), and guide the RITSC to nascent pericentromeric ncRNA transcripts (Figure 1.2). The chromodomain protein Chp1 is also a member of the RITSC and contributes to its localization to heterochromatin by binding H3K9me (Verdel, 2004). Once the RITSC is localized to

repeat loci it facilitates H3K9 methylation by recruiting the cryptic loci regulator complex (CLRC) which contains Clr4, the sole H3K9 methyltransferase in *S. pombe* (Zhang et al., 2008). Interestingly the catalytic slicing activity of Ago1 is required for the deposition and spreading of H3K9me, particularly in reporter genes (Irvine et al., 2006). Catalytic activity is required for passenger strand release from Ago1 bound dsRNA, and thus is required to facilitate base pairing between loaded siRNA and their targets, explaining this observation (Buker et al., 2007). This suggests that nuclear RNAi, specifically siRNA-target base pairing, is required for the spreading of heterochromatin, a phenomenon originally described as position effect variegation. These interactions place the RITSC in a central role, integrating transcription and chromatin modification. They also create a positive feedback loop between siRNA generation, RITSC localization and H3K9 methylation. A fascinating consequence of this is that H3K9 methylation itself is required for siRNA generation. The coupling of transcription, siRNA production, and silencing in *S. pombe* suggests that TGS occurs in *cis*.

1.2.2 Nuclear RNAi in S. pombe: CTGS

The dependency of RITSC localization on base pairing with ncRNA transcripts presents an interesting paradox in that loci targeted by RNAi for TGS must be transcribed in order to be silenced. Supporting this idea, genetic screens for loss of silencing in *S. pombe* have identified two point mutations in RNA Pol II subunits that decouple transcription and the RITSC at the pericentromeres (Djupedal et al., 2005; Kato et al., 2005). A model linking transcription, RNAi and heterochromatin formation can be formed when these observations are taken in the context of the cell-cycle. Studies have shown that transcription of pericentromeric repeats targeted by RNAi occurs during S-phase, the same time at which DNA is replicating and chromatin modifications must be re-established (Chen et al., 2008; Kloc et al., 2008). DNA replication and transcription must also be coordinated to prevent collision of the two processes and subsequent replication fork stalling. We found that RNAi is required to facilitate the release of RNA Pol II and prevent read-through transcription into replicating DNA (Zaratiegui et al., 2011). This suggests that RNAi once recruited to an actively transcribing Pol II may be able to inhibit transcription during the later elongation phase, resulting in the release of Pol II, as shown in Figure 1.2c. These observations support a model of co-transcriptional gene silencing (CTGS) in *S. pombe* (Fig. 2d) that was first termed by Bühler *et al*. (Bühler et al., 2006).

The CTGS model explains the paradox behind TGS. A nascent RNA transcript is required for the initial targeting of RNAi to a locus, once this occurs the nuclear RNAi complex can promote both transcriptional silencing at the chromatin level, and can co-transcriptionally silence by releasing RNA Pol via an unknown mechanism. It will be interesting to understand how transcription is initiated in what has previously been thought of as a restrictive heterochromatic environment and the mechanism by which the RITSC can promote Pol II release.

There is growing evidence that nuclear RNAi may co-transcriptionally regulate loci outside of constitutive heterochromatin in *S. pombe*. It has been shown to play a role in preventing read-through transcription at convergently transcribed genes, presumably though RNA Pol II release (Gullerova et al., 2011; Gullerova and Proudfoot, 2008; Zofall et al., 2009). Additionally Dcr1 physically interacts with chromatin at euchromatic genes suggesting a role in gene regulation without histone modification (Woolcock et al., 2011). Indeed, nuclear Dcr1 plays a role in regulating heat stress responsive genes through a "thermoswitch" (Woolcock et al., 2012). In unstressed cells Dcr1 is nuclear localized and negatively regulates stress response genes, however under heat stress it is exported out of the nucleus and stress response genes are activated.

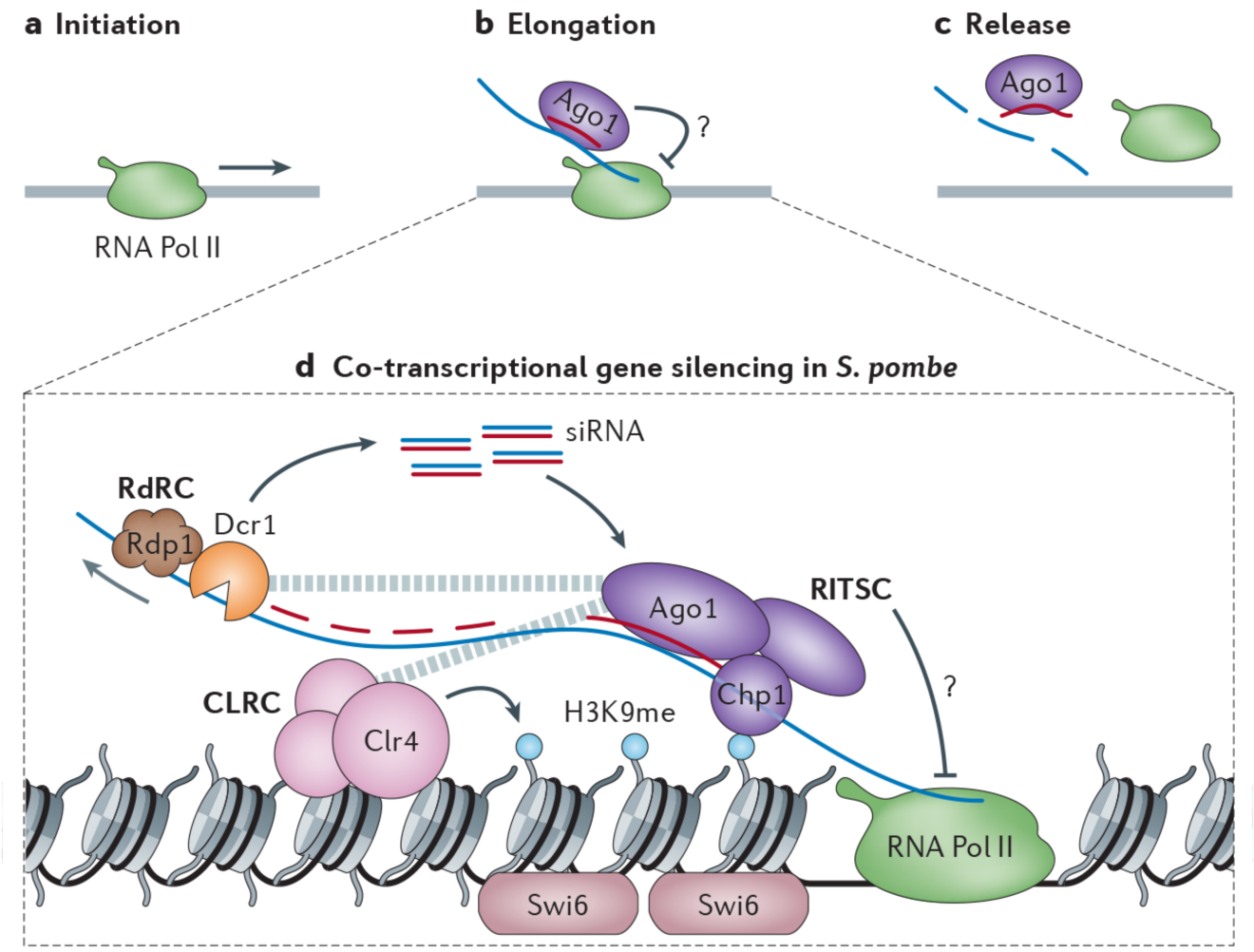

Figure 1.2 | Co-Transcriptional Gene Silencing in *S. pombe*. A) RNA Pol II initiates transcription at loci targeted by RNAi. B) During the elongation phase of transcription a Ago1 is guided to the nascent transcript and inhibits RNA Pol II transcription via an unknown mechanism. C) RNAi can lead to a release of RNA Pol II. d | A mechanistic model of RNAi acting during the elongation phase (2b) in *S. pombe*. The RITSC is localized through siRNA base-pairing with the nascent transcript, and chromatin interaction mediated by the chromodomain of Chp1. The RDRC couples dsRNA production by Rdp1 and siRNA cleavage by Dcr1 and is also associated with the nascent RNA Pol II transcript. The RITSC interacts with the CLRC that catalyzes H3K9 methylation at target loci. The RITSC promotes RNA Pol II release via an unknown mechanism. The dashed grey lines indicate interactions between complexes.

1.2.3 RNA Directed DNA Methylation in Arabidopsis

Transgene DNA methylation directed by viral RNA was discovered in plants long before a role for RNAi was known (Wassenegger et al., 1994), and later the involvement of small RNA and RNAi pathways in mediating TGS through cytosine methylation was first proposed in *Arabidopsis* (Aufsatz et al., 2002; Mette et al., 2000). There are many parallels between RNA Directed DNA Methylation (RdDM) in *Arabidopsis* and CTGS in *S. pombe*. For example, the requirement of transcription for silencing is common to both (Djupedal et al., 2005; Kato et al., 2005; Onodera et al., 2005), and both direct silencing at repetitive heterochromatic loci. RdDM differs from CTGS in *S. pombe* in that stepwise transcription by two RNA polymerases (Pol IV and Pol V) is required. Transcripts from Pol IV serve as substrates for siRNA generation, while nascent transcripts from Pol V are targeted by RNAi (Figure 1.3, reviewed extensively in Haag *et al*. 2011) (Haag and Pikaard, 2011). The initial template for Pol IV is not known, however it would presumably be loci that will be subject to RdDM. Pol IV physically interacts with the RNA-dependent polymerase 2 (RDR2) which produces dsRNA from transcripts (Law et al., 2011), that is subsequently processed into 24nt siRNA by Dicer Like 3 (DCL3) (Kasschau et al., 2007). These 24nt siRNA are exported into the cytoplasm where they are loaded into an Argonaute complex (Ye et al., 2012).

At least 3 of the 10 Argonautes found in *Arabidopsis* are involved in RdDM, but AGO4 was the first to be identified (Zilberman et al., 2003). Once loaded with 24nt siRNA in the cytoplasm AGO4 is imported into the nucleus and guided to complementary Pol V intergenic non-coding transcripts through siRNA target base pairing (Wierzbicki et al., 2008; 2009), and likely aided by direct protein-protein interaction with the Pol V subunit NRPE1 (El-Shami et al., 2007) and the Pol V associated GW/WG protein KTF1 (He et al., 2009; Rowley et al., 2011).

This co-transcriptional silencing by RNAi ultimately leads to the deposition of repressive cytosine methylation at loci transcribed by Pol V. In *Arabidopsis de novo* cytosine methylation is catalyzed by the enzyme DRM2 at loci targeted by RdDM (Cao et al., 2003). It might thus be expected to be a member of the RdDM effector complex alongside an Argonaute protein. Biochemical studies of a new complex member, RDM1 support this notion, as it interacts with both AGO4 and DRM2, and is

required for RdDM, bridging RNAi and cytosine methylation (Gao et al., 2010). The presence of a catalytically inactive DRM2 paralog DRM3 is also required for RdDM however its role is not known (Henderson et al., 2010). Once targeted DRM2 directs cytosine methylation in all cytosine contexts including at asymmetric CHH sites, to facilitate heterochromatin formation and TGS (Pélissier et al., 1999). Perhaps analogous to the role of Chp1 in localizing the RITSC to heterochromatin in *S. pombe*, the AGO4 associated protein RDM1 binds single stranded methylated DNA (Gao et al., 2010), and thus localizes AGO4 to methylated regions, creating a re-enforcing positive feedback loop.

Variations on the canonical RdDM pathway have been observed. AGO6 plays a partially redundant role with AGO4 (Zheng et al., 2007), and AGO9 is loaded with 24nt siRNA in the female germline, where its activity is required for transposon silencing, but a direct role in DNA methylation has not yet been established (Olmedo-Monfil et al., 2010). There is also evidence that transcripts from RNA Pol II (which chiefly transcribes euchromatic genes) as opposed to Pol V, are targeted by RdDM however the significance of this remains unclear (Gao et al., 2010; Zheng et al., 2009).

The RdDM pathway may be involved in H3K9 methylation, although it is uncertain if nuclear RNAi plays a direct role as in *S. pombe*. There is significant cross-talk between the two pathways as DNA methylation is required for the recruitment of the H3K9 methyltransferase SUVH4 / KYP (Johnson et al., 2007). At least two SUVH homologs are required for RdDM (Johnson et al., 2008) and small RNA from inverted repeats has been shown to influence H3K9 methylation to a greater extent than cytosine methylation suggesting a direct role (Enke et al., 2011).

RdDM may not be the only example of nuclear RNAi in *Arabidopsis*. There is evidence that another nuclear RNAi pathway involving DCL4 plays a co-transcriptional role in transcriptional termination. DCL4 was found to interact directly with chromatin in the 3' region of a Pol II transcribed endogenous gene to promote cleavage of the nascent transcript and transcription termination (Liu et al., 2012). Further study is needed to identify novel nuclear roles for other RNAi pathways.

A few examples outside of *Arabidopsis* indicate that siRNA may not be the only small RNA to direct DNA methylation in plants. In rice 24nt small RNA that arise from miRNA precursors termed long miRNA (lmiRNA) are RDR2-independent, processed by DCL3, and loaded into Ago4, which is normally associated with RdDM in *Arabidopsis* (Wu et al., 2010). These lmiRNA are able to direct highly sequence specific cytosine methylation at their own locus (in *cis*) and at complementary loci (in *trans*). Some lmiRNA have been identified in *Arabidopsis* however they have not been shown to direct DNA methylation (Vazquez et al., 2008). Similarly, in the moss *Physcomitrella patens* several 21nt miRNA have been shown to direct cytosine methylation at their targets (Khraiwesh et al., 2010). While both examples show that other classes of small RNA can direct DNA methylation neither uncover a novel effector pathway outside of RdDM.

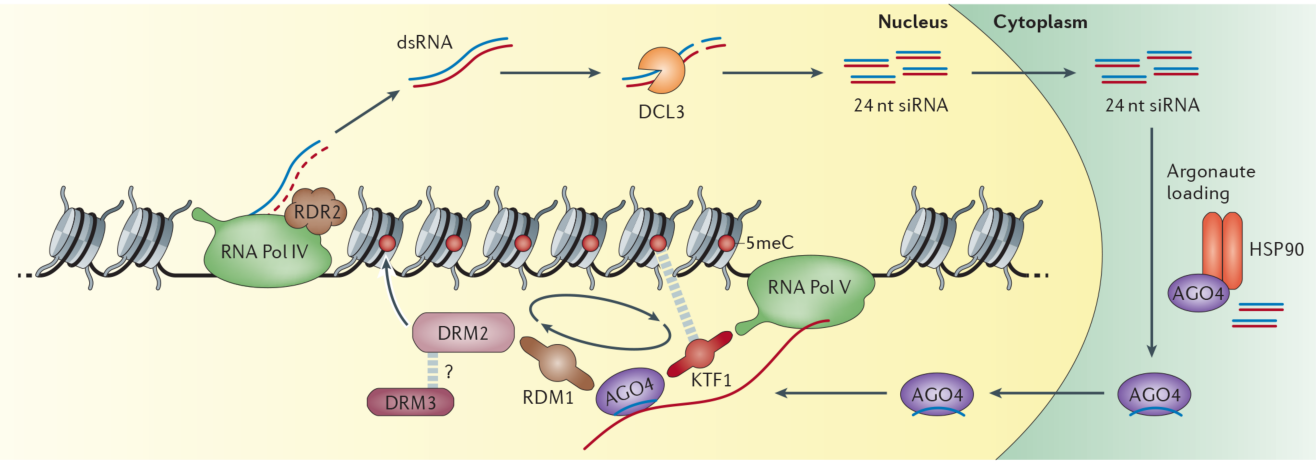

Figure 1.3 | The RNA-directed DNA Methylation pathway in *Arabidopsis*. RNA Pol IV transcribes ssRNA from repetitive heterochromatic loci. RDR2 physically associates with RNA Pol IV to produce dsRNA. DCL3 cleaves dsRNA to produce siRNA that are transported to the cytoplasm for AGO4 loading, facilitated by HSP90, which is then imported back into the nucleus. In the nucleus AGO4 targets nascent RNA Pol V transcripts through complementarity to siRNA and forms the RdDM complex presumably containing the catalytically active *de novo* DNA methyltransferase DRM2. The Pol V associated GW/WG protein KTF1 may act as an organizer by interacting with AGO4 and 5meC. Similarly, the AGO4 associated protein RDM1 can bind single stranded methylated DNA and interacts with DRM2. Both could contribute to a positive feedback loop between AGO4 localization and DNA methylation (circular arrows). DRM3, a catalytically inactive paralog of DRM2 is required for RdDM however its role is unknown. Once localized, DRM2 catalyzes methylation of cytosine in all sequence contexts. The dashed grey lines indicate interactions.

1.2.4 Metazoan Somatic Nuclear RNAi

While the germlines of metazoans have a clear role for nuclear RNAi (see *Nuclear RNAi in the Germline*), some evidence suggests that TGS also occurs in somatic cells, however the subject is controversial. Feeding *C. elegans* with dsRNA targeting an endogenous gene triggers H3K9 methylation at the target locus in somatic cells that is dependent on the nuclear RNAi pathway (NRDE) (Burton et al., 2011; Gu et al., 2012; Guang et al., 2008; 2010) and on the RDRP RRF-1 (Burkhart et al., 2011). There are many genes targetted by endogenous siRNA, and some but not all show a reduction of H3K9me in *nrde* mutants (Burkhart et al., 2011). In *Drosophila* somatic cells, mutations in siRNA pathway members *dcr2* or *ago2* affect expression of a centromeric reporter and result in a marked reduction of centromeric H3K9 methylation (Deshpande, 2005; Fagegaltier et al., 2009; Peng and Karpen, 2007).

As in fission yeast, proteins required for nuclear RNAi interact with the transcriptional machinery in metazoan somatic cells, suggesting that CTGS may be conserved. In human and *Drosophila* cells, Ago1 interacts directly with RNA Pol II by co-immunoprecipitation (Kavi and Birchler, 2009; Kim et al., 2006). In *Drosophila* S2 cells Ago2 and Dcr2 associate directly with both chromatin and RNA Pol II, and are required to inhibit the expression of heat-shock response genes under non-stress conditions by maintaining paused Pol II and preventing elongation (Cernilogar et al., 2011). In *C. elegans*, loci targetted by RNAi show a downstream decrease in RNA Pol II occupency that is dependent on the nuclear RNAi factor NRDE-2 and Argonaute NRDE-3, suggesting that siRNA may facilitate transcription termination (Guang et al., 2010). Overall current evidence suggests a conserved interaction of nuclear RNAi and the transcriptional machinery fitting a co-transcriptional model, however the role of these interactions needs further exploration.

1.3 Nuclear RNAi in the Germline

The germline is the battlefield on which evolutionary wars between selfish DNA elements and their hosts are played out because transposable element (TE) mobilization here would be inherited by future generations. Nuclear RNAi – the piRNA pathway in animals and various siRNA pathways in plants - is a front line defense.

1.3.1 Germline Nuclear RNAi in Arabidopsis

In plants, germline cells arise late in development from somatic stem cells (unlike in animals, in which the germline is specified early in development), and so transposons must be silenced extensively throughout development. Generally, chromatin marks that are present during somatic development must be reset in the germline. How this occurs selectively is a question that is actively being pursued. In somatic cells both the RdDM pathway and maintenance DNA methyltransferases keep transposons silent; however, this changes in the companion cells of the germline that will not contribute genetically to the next generation. The heterochromatin remodeler *ddm1* is a master regulator of transposons (Lippman et al., 2004) and is down-regulated in the supportive vegetative nucleus (VN), leading to transposon mobilization and the production of 21nt sRNA antisense to transposons (Slotkin et al., 2009) (Figure 1.4a). These 21nt sRNA can silence reporters expressed in sperm cells so they appear to act non-cell-autonomously. With regards to DNA methylation, unlike mammals which undergo whole genome demethylation during spermatogenesis (Popp et al., 2010), the *Arabidopsis* male germline retains symmetric methylation at levels similar to somatic cells (Calarco et al., 2012; Ibarra et al., 2012), but shows a reduction in the levels of asymmetric methylation specifically at a subset of retrotransposons that are later re-methylated in the developing embryo (Jullien et al., 2012).

In the female gametophyte the maintenance DNA methyltransferase *met1* is repressed (Jullien et al., 2008) and the DNA glycosylase *demeter,* which removes cytosine methylation, is expressed (Choi et al., 2002) in the diploid central cell (CC), that will later become the "extra-embryonic" endosperm (Figure 1.4b). This leads to global cytosine demethylation in the endosperm, accompanied by increased production of 24nt siRNA leading to non-CG hypermethylation at target sites, which are primarily

retroelements (Hsieh et al., 2009). These 24nt siRNA are bound by AGO9 in the central cell, and act non-cell-autonomously to control specification of gametic egg cells (Olmedo-Monfil et al., 2010). Currently there is no direct experimental evidence showing the movement of either 24nt siRNA or AGO9 from the central cell to the egg cell, however in *ago9* mutants transposable elements are activated in the egg cell where *ago9* is not expressed supporting this hypothesis. These observations suggest a hypothetical model by which transposons are revealed in companion cells, and are then used to generate small RNA that enforces transposon silencing in the germ cells (Martienssen, 2010), however it is not known if they can also direct TGS through nuclear RNAi.

1.3.2 The Drosophila piRNA Pathway

In animals, the role of the piRNA pathway in TE silencing has been best described in *Drosophila* ovaries. Within the ovaries piRNA silence transposons in somatic follicle cells surrounding the oocyte, germline nurse cells and the oocyte itself (Li et al., 2009; Malone et al., 2009) (Figure 1.4c). The somatic follicle cells produce only antisense piRNA from the *flamenco* locus that do not participate in the ping-pong cycle and are instead processed and loaded solely into Piwi. These piRNA predominantly target elements from the *gypsy* family of long terminal repeat (LTR) retroviruses. Gypsy family elements are able to propagate by producing viral packages in follicle cells that can infect germline cells, thus the *flamenco* derived piRNA pathway is thought to be an evolutionary counter to this class of transposons (Li et al., 2009). In nurse cells and ovaries the ping-pong cycle defends against a wide variety of TEs using input from all piRNA clusters and mRNA of active transposons (Brennecke et al., 2007; Malone et al., 2009). Here the piRNA pathway degrades transposon transcripts, and directs H3K9 methylation to transcriptionally silence transposons and prevent their mobilization (Wang and Elgin, 2011). Piwi has been shown to specifically interact with heterochromatin protein 1 (HP1a), a defining component of heterochromatin, and is chromatin associated itself (Brower-Toland et al., 2007). Furthermore, the nuclear localization of Piwi is required for chromatin-mediated repression of a subset of transposons suggesting a direct role (Klenov et al., 2011). Silencing of piRNA clusters themselves would be detrimental, as this would prevent primary piRNA from entering the cycle. This is solved by the HP1 variant Rhino that is restricted to germline nuclei and specifically localizes to piRNA clusters and promotes transcription of the heterochromatic clusters (Klattenhoff et al., 2009). How Rhino is localized to piRNA clusters and not active transposons remains unexplained.

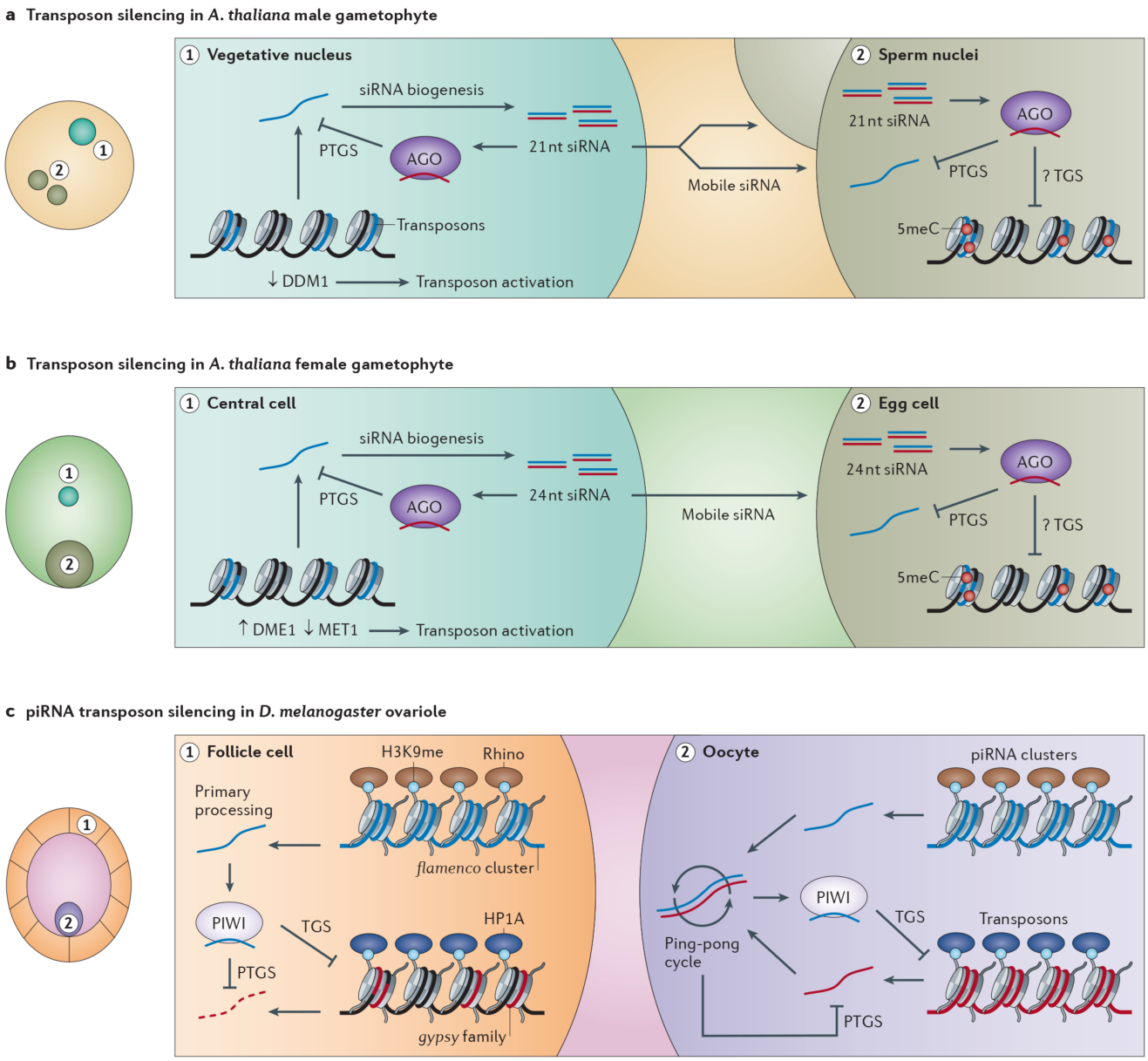

Figure 1.4 | RNAi mediated transposon silencing in the germline. A) In the supportive vegetative nucleus of the *Arabidopsis* male gametophyte *ddm1* expression is repressed which leads to the loss of cytosine methylation and reveals transposons. Transposons are processed into 21nt siRNA that are mobile and can direct PTGS in the sperm nuclei. They may also impact transposons transcriptionally by directing or inhibiting epigenetic modification. Red lollipops represent 5meC. B) The supportive central cell of the *Arabidopsis* female gametophyte reveals transposons for transcription by downregulating the maintenance DNA methyltransferase MET1 and expressing the DNA glycosylase DEMETER causing a loss of cytosine methylation. This activates the RdDM pathway and produces 24nt siRNA that may be transported to the egg cell to enforce transcriptional silencing through AGO9. Red lollipops represent 5meC. C) In the *Drosophila* ovariole the

flamenco cluster is expressed in somatic follicle cells, and generates piRNA independently of the ping-pong cycle. Loaded Piwi silences the *gypsy* family of retrotransposons which could otherwise form infectious particles. In oocytes and surrounding nurse cells all piRNA clusters are expressed and the primary transcripts enter the ping-pong cycle to produce piRNA. Active transposons are post-transcriptionally silenced, and nuclear Piwi promotes transcriptional silencing via H3K9 methylation, and HP1a localization. The HP1a homolog Rhino binds to heterochromatic piRNA clusters in place of HP1a and promotes transcription.

1.3.3 The Mouse piRNA Pathway

The Piwi pathway is highly conserved in animals and plays a similar role in the mouse germline. In mouse two Piwi homologs, MILI and MIWI2 are required for transposon silencing in the male germline. Loss of either causes transposon mobilization and sterility (Aravin et al., 2007; Carmell et al., 2007). The piRNA pathway however operates differently from *Drosophila*. In the mouse male germline transposons are globally derepressed by cytosine demethylation during early development. The piRNA pathway is then primed with individual transposons and re-establishes methylation patterns during development (Aravin et al., 2008; Watanabe et al., 2011). As MIWI2 is found in the nucleus it is likely to be the effector Argonaute of RNA directed DNA methylation in mouse (Aravin et al., 2009). The role of MIWI2 in establishing DNA methylation in the germline may not be direct as with AGO4 in *Arabidopsis (*or Ago1 for H3K9me in *S. pombe*) since co-immunoprecipitation experiments have failed to show interaction between MIWI2 and the *de novo* methyltransferases Dnmt3a and Dnmt3b. The role of nuclear RNAi in directing DNA methylation in mammals is nicely demonstrated at the imprinted *rasgfr1* locus where the piRNA pathway is required for *de novo* methylation in the male germline (Watanabe et al., 2011). Upstream of the differentially methylated region is an LTR that matches piRNAs with a typical pong-pong signature; probably these piRNAs can be generated due to the presence of another copy of the LTR in a piRNA cluster. The LTR is contained within a non-coding RNA that is transcribed specifically during spermatogenesis when *de novo* methylation occurs. This nascent ncRNA is targeted by piRNA and co-transcriptionally silenced by the deposition of DNA methylation. This may facilitate the spread of targeted silencing into the nearby *rasgfr1* locus, leading to imprinting, similarly to CTGS in *S. pombe*. Again the authors do not rule out the possibility that silencing by piRNA may be indirect, and this is a single locus example. The *rasgfr1* locus is however

unlikely to be the only example of RNAi directing imprinting or silencing of an endogenous gene, and hints that nuclear RNAi and transposon acquisition play a role in imprinting across organisms. Further genetic and biochemical dissection is needed to discern if the piRNA pathway plays a direct role in DNA methylation, and if so what the mechanistic details are. Specifically interactions between piRNA effectors and cytosine methyltransferases, and the use of exogenous reporters containing sequences complementary to known piRNA would provide convincing evidence.

1.3.4 Germline Nuclear RNAi in C. elegans

A class of small RNA termed 21U has been proposed to be the piRNA of *C. elegans* (Das et al., 2008; Ruby et al., 2006; Wang and Reinke, 2008). They associate with the Piwi-family protein PRG-1 which is required to silence Tc3 mariner transposons in the germline and for fertility (Batista et al., 2008; Das et al., 2008).

C. elegans 21U RNA originate from over 5700 loci dispersed over two broad clusters on chromosome IV (Ruby et al., 2006), however no evidence of a ping-pong cycle has been observed. The 21U pathway has been suggested to function by determining the specificity of the 22G siRNA and *nrde* pathways (see *Metazoan Somatic Nuclear RNAi* and Figure 1.1b) that direct TGS in the form of H3K9me at piRNA targets (Figure 1.5). Two avenues of study have validated this model. In *C. elegans* single copy transgenes with long exogenous DNA sequences, such as GFP, are stably silenced at a high frequency. This silencing correlates with H3K9me3 accumulation and is dependent on PRG-1 and 21U RNA accumulation for its establishment, and the germline specific nuclear Argonautes WAGO-9 and WAGO-10 that bind 22G RNA for its maintenance (Shirayama et al., 2012). Studies with reporter transgenes that contain sequences complementary to known 21U small RNAs (piRNA sensors) have revealed identical requirements for silencing and additionally implicated the HP1 ortholog HPL-2 and putative methyltransferases SET-25 and SET-32 in establishing H3K9me3 at loci targeted by piRNA (Ashe et al., 2012). Outside of transgenes, silencing at endogenous loci mediated by piRNA likely functions by the same mechanism. Indeed many endogenous loci that are targeted by 21U small RNA and silenced exhibit increased mRNA expression and a loss of corresponding 22G RNA in a *prg1* mutant background (Lee et al., 2012). RNAi also acts to establish repressive heterochromatin during meiosis at unpaired chromosomal regions in *C. elegans*. Specifically the RdRP EGO-1 and the Piwi family Argonaute protein CSR-1 are required for this process (Maine et al., 2005; She et al., 2009).

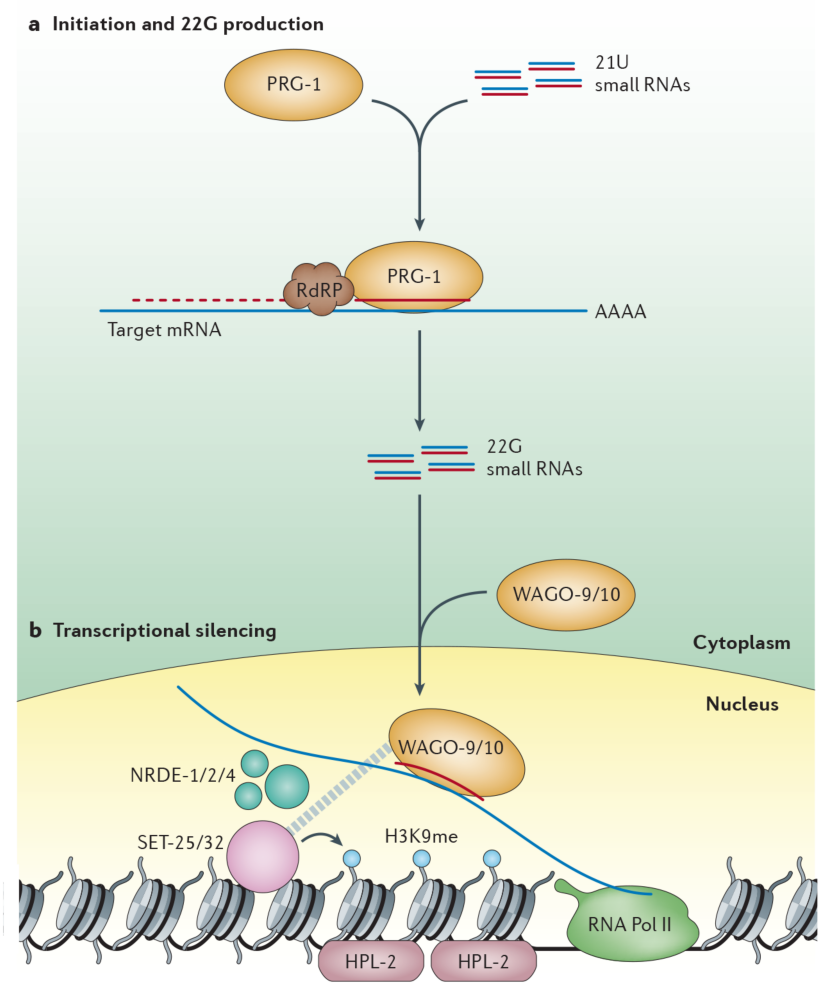

Figure 1.5 | piRNA (21U) Pathway in the *C. elegans* germline. A) The "21U" piRNA of *C. elegans* originate from two broad clusters on chromosome IV, however little is known about their biogenesis. They act with the Piwi family Argonaut PRG-1 to target mRNA in the cytoplasm. Targeting of PRG-1 to mRNA recruits a RdRP to produce abundant 22G siRNA. B) 22G siRNA is loaded into the germline specific nuclear Argonautes WAGO-9/10, which are closely related to NRDE-3, the nuclear Argonaute involved in somatic TGS. Loaded WAGO-9/10 is transported into the nucleus where it targets nascent transcripts of RNA Pol II and directs H3K9me that is dependent on the nuclear RNAi components NRDE-1/2/4. H3K9 methylation is catalyzed by two putative histone methyltransferases SET-25/32. The HP1 ortholog HPL-2 binds H3K9me and is required for multi-generational silencing.

1.3.5 Systemic TGS and Epigenetic Inheritance

The hypothesis that siRNA can move into *Arabidopsis* germ cells (see *Germline Nuclear RNAi in Arabidopsis*) has implications for epigenetic inheritance. Outside of the gametophytes, grafting experiments have shown that nuclear silencing signals can be transmitted from the root to shoot (Brosnan et al., 2007) and vice versa (Melnyk et al., 2011). Mobile 21 to 24nt siRNA are the effectors of this systemic silencing and can guide epigenetic modification through RdDM in recipient cells (Dunoyer et al., 2010; Molnar et al., 2010). These 24nt siRNA have been demonstrated to direct DNA methylation in meristematic root stem cells (Melnyk et al., 2011), and it is therefore tempting to speculate that they may act similarly in the shoot meristems (where germ cells are produced) to direct heritable epigenetic modification.

Systemic RNAi is well known in *C. elegans*, and there is recent evidence for small RNA mediated epigenetic inheritance. The progeny of animals exposed to dsRNA show H3K9 methylation of target loci and generate complementary small RNA for multiple generations (Burton et al., 2011). The appearance of siRNA precedes H3K9me in progeny so it's likely that this inheritance is indirect and is instead re-established by inherited siRNA in each generation. This process is dependent on the nuclear RNAi pathway, including the Argonaute NRDE-3. Furthermore small RNA produced against viral RNA can be transgenerationally inherited, and continue to persist even in the absence of the viral template itself (Rechavi et al., 2011). These studies both point to small RNA as an epigenetic vector, which can be inherited and through nuclear RNAi direct chromatin modifications in progeny. Once established, these chromatin modifications can be maintained and transmitted across generations even in the absence of the dsRNA trigger (Gu et al., 2012).

In *Drosophila* the makeup of cellular piRNA is epigenetically inherited. Reciprocal crosses have shown that progeny inherit the maternal piRNA composition, and this composition persists into adulthood (Brennecke et al., 2008). The maternally deposited piRNA may prime the ping-pong cycle and determine its specificity, or could potentially direct epigenetic modification to enforce a specific piRNA transcription program. A similar situation is seen in the *Arabidopsis* endosperm where

maternally deposited 24nt siRNA silence TEs and TE-associated genes during its development (Lu et al., 2012; Mosher et al., 2011).

1.4 Nuclear RNAi in Genome Maintenance and Repair

Nuclear RNAi plays a critical role in maintaining genome integrity by preventing transposon mobilization, however more direct roles in genome maintenance and DNA repair are emerging.

1.4.1 Chromosome structure and function

Proper chromosome condensation is required for segregation during mitosis. In *S. pombe* the loss of RNAi causes a high incidence of lagging chromosomes and sensitivity to a microtubule inhibiting drug (Volpe et al., 2003). Also, in the *Drosophila* germline the piRNA DEAD-box RNA helicase Vasa facilitates condensin I localization, which promotes chromosome condensation and is dependent on the piRNA components *aub* and *spindle-E* (Pek and Kai, 2011a). A *vasa* paralog *belle* acts analogously in somatic cells and requires the siRNA components *ago2* and *dcr2* (Pek and Kai, 2011b). Interestingly, a role for RNAi in cohesin localization has also been proposed in *S. pombe* (Gullerova and Proudfoot, 2008), suggesting a conserved role for RNAi in facilitating cohesin/condensin localization ensuring proper chromosome condensation.

The telomeres of *Drosophila* are unique in that they rely on a transposon based elongation mechanism (Shpiz and Kalmykova, 2011). The piRNA pathway has been found to regulate these telomeric transposons in the germline, and thus can regulate telomere length (Savitsky, 2006). Specifically *ago3* mutant embryos show an increase in telomeric transposition, and a subsequent increase in telomere length (Khurana et al., 2010). Additionally, *aub* and the RNA helicase *armitage* are involved in the production of telomere specific piRNA and their loss results in increased telomere fusion, suggesting another role for the piRNA pathway in telomere cap assembly (Khurana et al., 2010). Nuclear RNAi is also required for proper telomere function in *S. pombe*. Subtelomeric regions contain a region that is homologous to the pericentromeric repeats and this region facilitates RNAi-dependent heterochromatin

formation (Kanoh et al., 2005). It's possible that nuclear RNAi may have a conserved role in telomere maintenance across organisms.

Eukaryotic genomes contain extensive regions of repetitive DNA which if engaged in recombination can cause detrimental changes to chromosome structure. There is evidence that RNAi pathways may act to repress recombination in repetitive regions. The loss of RNAi in *S. pombe* cells leads to both an increase in meiotic recombination (Ellermeier et al., 2010), and a dependence on mitotic recombination in repetitive pericentromeric regions, as double mutants between RNAi components and the master regulator of homologous recombination *rad51* are synthetic lethal (Zaratiegui et al., 2011). This observation has also been made in *Drosophila* where RNAi mediated suppression of recombination is required to maintain stability of repetitive DNA (Peng and Karpen, 2007).

1.4.2 DNA damage response

A role for small RNA in the DNA damage response was first observed in *Neurospora*, where small RNA is generated from rDNA repeats when cells are treated with DNA damaging agents (Lee et al., 2009). More recently, RNAi has been shown to directly mediate DNA repair in *Arabidopsis*. Double strand breaks (DSBs) were found to induce a population of 21nt small RNA (Wei et al., 2012). These small RNA originate from the vicinity of the double strand breaks and their biogenesis requires the siRNA biogenesis factors RNA Pol IV and dicer-like proteins. They are recruited to DSBs by AGO2 and mediate repair, as mutants in *ago2* or biogenesis factors cause a reduction in DSB repair efficiency. The authors suggest that AGO2 recruits the DSB repair complex to damaged loci, analogously to the localization of DNA methylation complexes in RdDM. Importantly the results were validated in human cell lines pointing to a conserved role for RNAi in DSB repair. A similar finding has been reported in *Drosophila* cells where double strand breaks induce a localized production of siRNA that is dependent on Ago2 and Dcr2, members of the endo-siRNA pathway (Michalik et al., 2012). Upon DSB formation the DNA-damage response (DDR) pathway is activated and can arrest cell proliferation. Focus on this pathway has revealed that DICER and DROSHA-dependent small RNA are required for DDR activation in human, mouse, and zebrafish (Francia et al., 2012). It is therefore likely that the DDR

pathway may link RNAi and DNA repair, although the specific function of the small RNAs themselves remains a mystery.

1.4.3 Targeted genome elimination

Perhaps the most extreme role for nuclear RNAi in genome stability is in targeted genome elimination in *Tetrahymena*. *Tetrahymena* species retain two nuclei, a germline micro-nucleus (Mic) and a somatic macro-nucleus (Mac). After zygote formation a new Mac develops by the deletion of ~6000 internal eliminated sequences (IES). These IES are enriched for H3K9 methylation before deletion (Taverna et al., 2002) and produce a population of 28nt scan RNA (scnRNA) that associate with the Argonaute Twi1p (Mochizuki et al., 2002). A RNA helicase Ema1p facilitates the interaction between loaded Twi1p and chromatin by promoting base-pairing with nascent transcripts, fitting a co-transcriptional model (Aronica et al., 2008). It is hypothesized that this leads to the deposition of H3K9 methylation by a mechanism similar to *S. pombe* that then serves as a mark for DNA elimination in the Mac.

These examples show that in addition to silencing transposons nuclear RNAi has a conserved role in maintaining genome stability by participating in a variety of pathways across different organisms. In particular the link to DSB repair shows that Argonaute effector complexes can be directly involved in DNA repair. In other examples it is not clear if RNAi plays a direct role or if it simply maintains genome integrity through H3K9 methylation. Higher eukaryotes have numerous Argonaute proteins many of which are uncharacterized. Further investigation of these Argonautes may reveal novel roles in genome maintenance outside of classical RNAi.

1.5 Summary

Although a role for RNAi in the nucleus was first described in *Arabidopsis* and *S. pombe*, observations in key model organisms suggest that it is evolutionarily conserved. RNAi mediated transcriptional gene silencing has now been observed in plants, fungi, and metazoans, and evidence is mounting that it operates co-transcriptionally as in *S. pombe*. Across organisms nuclear RNAi operates predominantly at

heterochromatic loci where it facilitates sequence specific silencing through the direction of histone H3K9 methylation and/or cytosine methylation. Differences are however seen in small RNA biogenesis particularly in the subcellular localization of small RNA processing, and loading of Argonaute proteins, and could represent alternative approaches to regulating nuclear RNAi. Mechanistically it is still unclear in the context of the co-transcriptional model how nuclear RNAi complexes regulate transcriptional machinery. Outside of constitutive heterochromatin RNAi co-transcriptionally regulates some genes, and experiments are underway to determine if this is a widespread phenomenon across organisms.

The role played by nuclear RNAi in the germline to prevent the propagation of selfish DNA elements in future generations is significant and highly conserved. There is often a link between imprinted genes and nearby transposons, in mammals as well as in plants, which may be important in the evolution of some aspects of imprinting from germline transposon control. This field of study will be particularly fruitful in parallel with work on co-transcriptional models that could explain the spreading of silencing at transposon targets into nearby genes associated with non-coding RNA and RNAi. Outside of imprinting it is likely that small RNA themselves play a conserved role in epigenetic inheritance. As the ability to profile germline cells improves these question will be addressed.

Finally, the participation of nuclear RNAi in genome maintenance and DNA repair shows that there are other roles that nuclear small RNA and their effectors play outside of those involved in classical transcriptional silencing. Biochemical purification of novel Argonaute effectors in the context of DNA repair will help to identify the players.

2. RNAi promotes heterochromatic silencing through replication-coupled release of RNA Pol II

Abstract, figures and figure legends reproduced with modifications from:

> Zaratiegui M, Castel SE, Irvine DV, Kloc A, Ren J, Li F, de Castro E, Marín L, Chang AY, Goto D, Cande WZ, Antequera F, Arcangioli B, Martienssen RA. RNAi promotes heterochromatic silencing through replication-coupled release of RNA Pol II. *Nature* 479(7371), 135-138 (2011).

2.1 Abstract

Heterochromatin comprises tightly compacted repetitive regions of eukaryotic chromosomes. The inheritance of heterochromatin through mitosis requires RNA interference (RNAi), which guides histone modification during the DNA replication phase of the cell cycle. Here we show that the alternating arrangement of origins of replication and non-coding RNA in pericentromeric heterochromatin results in competition between transcription and replication in *Schizosaccharomyces pombe*. Co-transcriptional RNAi releases RNA polymerase II (Pol II), allowing completion of DNA replication by the leading strand DNA polymerase, and associated histone modifying enzyme that spread heterochromatin with the replication fork. In the absence of RNAi, stalled forks are repaired by homologous recombination without histone modification.

2.2 Introduction

The hallmark of epigenetic information is that it can be inherited through cell division. In fission yeast constitutive heterochromatic domains, defined by histone H3K9 methylation, are epigenetically inherited. The phenomenon of position effect variegation (PEV) aptly illustrates the inheritance of histone states. Originally described in *Drosophila*, PEV is the stochastic spreading of heterochromatin into a nearby gene, causing it to become silent (Elgin and Reuter, 2013). This silencing is inherited

through cell division, and if the gene has a phenotype associated with it, leads to a variegated expression pattern. In fission yeast reporter genes inserted nearby to either centromeric (Allshire et al., 1994) or mating type region (Ayoub et al., 1999) heterochromatin exhibit variegated expression patterns. RNAi only partially contributes to centromeric heterochromatin maintenance, however it is completely essential for PEV (Irvine et al., 2006), suggesting that RNAi is required for the inheritance and spreading of H3K9 methylation in fission yeast. Indeed, RNAi operates specifically during S-phase, the time at which histones marks are diluted and must be re-established to ensure faithful transmission of epigenetic information (Chen et al., 2008; Kloc et al., 2008). In S-phase histone H3 is phosphorylated at serine 10, which disrupts repressive heterochromatin by evicting Swi6, the fission yeast homolog of the metazoan HP1 (heterochromatin protein 1). Once repression is relieved, pericentromeric repeats are transcribed by RNA Pol II (Pol II) and processed into siRNA that are fed into the RNAi pathway. These observations suggest that DNA replication and transcription compete for DNA template within pericentromeric heterochromatin, something that is generally prevented elsewhere in the genome.

Head on collision between the replisome and transcription complexes causes replication fork stalling, and in order for replication to proceed the transcription complex must be removed. This can be accomplished through both recombination and transcription coupled repair pathways (Hanawalt and Spivak, 2008). If left unresolved the stalling can result in DNA damage and inappropriate recombination events that lead to genomic instability and a change in surrounding chromatin states (Bermejo et al., 2012). Other consequences of stalled forks are double strand breaks, which lead to the phosphorylation of histone H2A (γH2A). Surprisingly, genome wide profiling of γH2A localization in fission yeast revealed centromeric heterochromatin as a major site of γH2A enrichment (Rozenzhak et al., 2010). This localization was dependent on Clr4 (H3K9me2) and Swi6, suggesting that DNA repair pathways might play some role in heterochromatic domains.

There are similarities between constituents of the histone methyltransferase complex (CLRC) and DNA repair complexes. For example, Rik1, an essential member of the complex shares homology with DDB1 (DNA damage binding protein 1) family of proteins, which are involved in wide variety of processes, including nucleotide excision repair (NER), and the regulated degradation of proteins involved in DNA repair, transcription, and replication (Iovine et al., 2011). In fission yeast Rik1 acts

analogously to Ddb1, working with Raf1 (Rik1 associated factor) to facilitate the ubiquitination of targets by an E3 ubiquitin ligase in the CLRC, Cul4 (Buscaino et al., 2012). This activity is required for siRNA generation and H3K9 methylation, however its targets are unknown. In addition to similarities to DNA repair complexes, CLRC components also interact directly with DNA replication machinery, specifically Rik1 and a component of the leading strand DNA polymerase complex (Cdc20) (Li et al., 2011). Cdc20 is required for heterochromatin assembly, for the first time linking epigenetic modification to DNA replication in fission yeast. Another member of the Cdc20-Rik1 complex is Mms19, a TFIIH transcription factor with a well established role in DNA repair (Kou et al., 2008; Lauder et al., 1996).

These observations lead us to hypothesize that collisions between transcription and replication occur frequently within heterochromatin, and because of similarities between repair pathways and heterochromatin factors, could be resolved uniquely in the context of heterochromatin. We therefore sought to study the well-established role of RNAi in regulating transcription, specifically in the context of DNA replication. We found that transcription and replication are arranged in an alternating pattern within pericentromeric arms, and that replication fork stalling occurs within regions of transcription-replication conflict. This conflict produces stalled Pol II that is released by RNAi to facilitate DNA replication. Without release by RNAi, DNA damage occurs and fork restart by homologous recombination becomes necessary. We therefore propose a model by which RNAi plays a critical role in heterochromatic silencing through replication-coupled release of Pol II.

2.3 Results

2.3.1 Transcription and replication compete at the centromere

We first sought to produce a comprehensive profile of centromeric replication and transcription. Replication origins have been mapped both genome wide (Segurado et al., 2003) and at the centromeres (Smith et al., 1995), allowing us to place origins with respect to *dg* and *dh* repeat units (Figure 2.1a). We characterized the transcriptional landscape of centromeres by performing H3K9me2 ChIP-on-chip, small RNA sequencing, and ChIP-seq of both poised (pS5) and elongating (pS2) Pol II in dividing cells

(Figure 2.1b). Pol II ChIP-seq was performed in both wild-type (WT) and *dcr1*Δ cells and visualized as the increase in *dcr1*Δ ChIP read depth. There was a striking pattern of alternating transcription units and replication origins. In dividing cells regions transcribed by Pol II in the absence of RNAi corresponded to siRNA clusters, and peaked at polyA sites (Figure 2.1c), suggesting a defect in transcriptional termination. Supporting this, we have detected read-through centromeric transcripts up to 4.5kb in length in RNAi mutants (Zaratiegui et al., 2011).

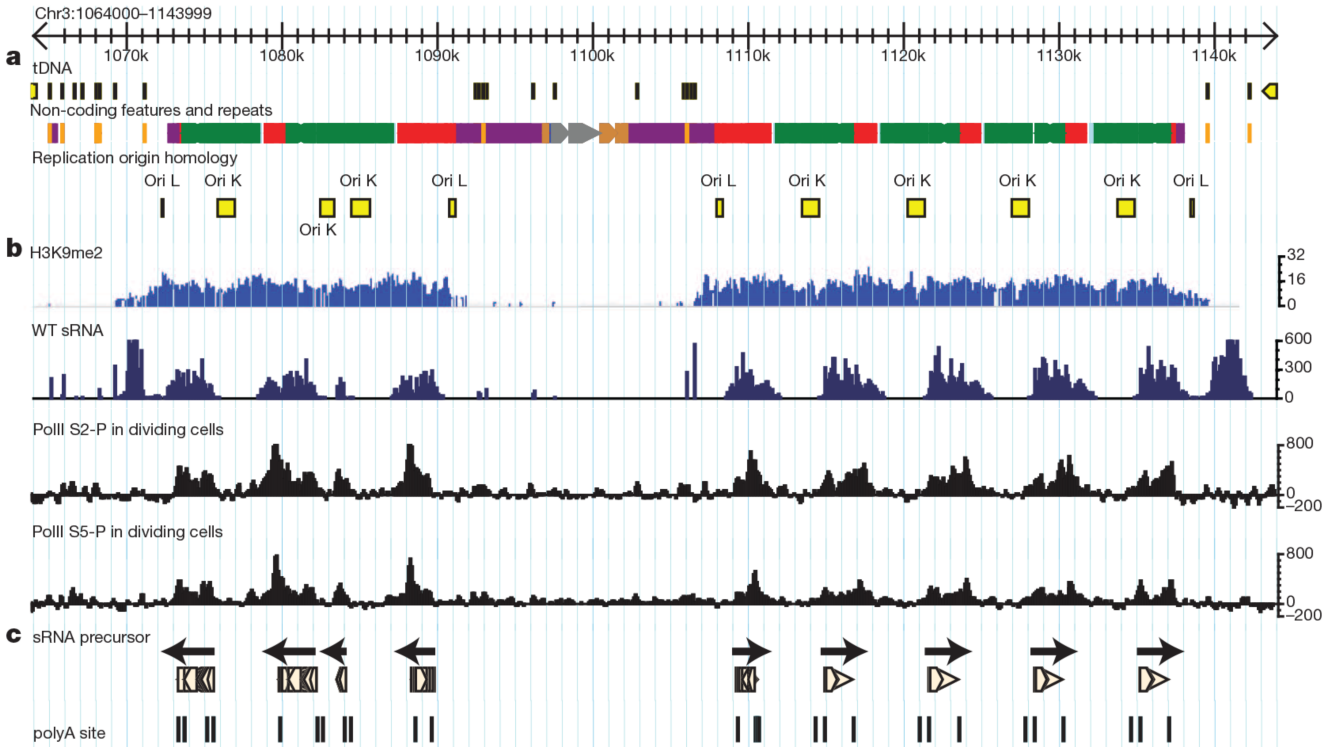

Figure 2.1 | Transcription and replication of pericentromeric heterochromatin in fission yeast. A) Pericentromeric heterochromatin on centromere 3. dh (red), dg (green) and imr (magenta) repeats are shown, bordered by tRNA genes (brown). Replication origins (yellow) are found in each repeat. B) Tiling microarrays of K9me2 ChIP (light blue) and clusters of small RNA sequences (dark blue) from wild-type cells. ChIP-seq reads corresponding to poised (S5-P) and elongating (S2-P) RNA polymerase II enriched in *dcr1*Δ cells relative to wild-type (WT) cells are in black. C) cDNA clones (beige) from *dcr1*Δ cells. PolyA sites are indicated as vertical lines and correspond to peaks of Pol II. Arrows indicate the direction of 'Forward' transcription.

2.3.2 RNAi is required to release Pol II and prevent DNA damage

Asynchronous *S. pombe* cultures contain predominantly G2 cells. To examine the effect of DNA replication on Pol II transcription we treated cells with hydroxyurea (HU), which arrests cells in S-phase by depleting dNTPs. Pol II ChIP was performed using arrested cells and again compared between WT and *dcr1*Δ (Figure 2.3a). Exemplified at a replication origin in centromere 1 (Figure 2.3b), there was a significant increase in *dcr1*Δ Pol II that extended into the replication origin in arrested cells that was absent in dividing cells. This indicates that during S-phase Pol II is released from transcription into replication origins by RNAi. Pol II ChIP-qPCR revealed similar enrichment within replication origins (Figure 2.2c).

Collision between transcription and replication result in stalled replication forks and DNA damage signaling. Rad52 binds single stranded DNA at such stalled replication forks, and replaces RPA with Rad51, a master regulator of homologous recombination (HR) (Lambert et al., 2010). We performed Rad52 ChIP-qPCR in synchronized cells using centromeric probes and found a dramatic increase in Rad52 within *dh* repeats, that was greatest during S-phase (Figure 2.2c). HR mediated by Rad51 can restart stalled replication forks (Lambert et al., 2010), and the presence of Rad52 at centromeric sites indicated that HR was engaged at the centromere in RNAi mutants. We found that double mutants between *rad51* and RNAi genes were inviable or sick (Figure 2.3d), revealing that fork restart by HR is essential in the absence of Pol II release by RNAi.

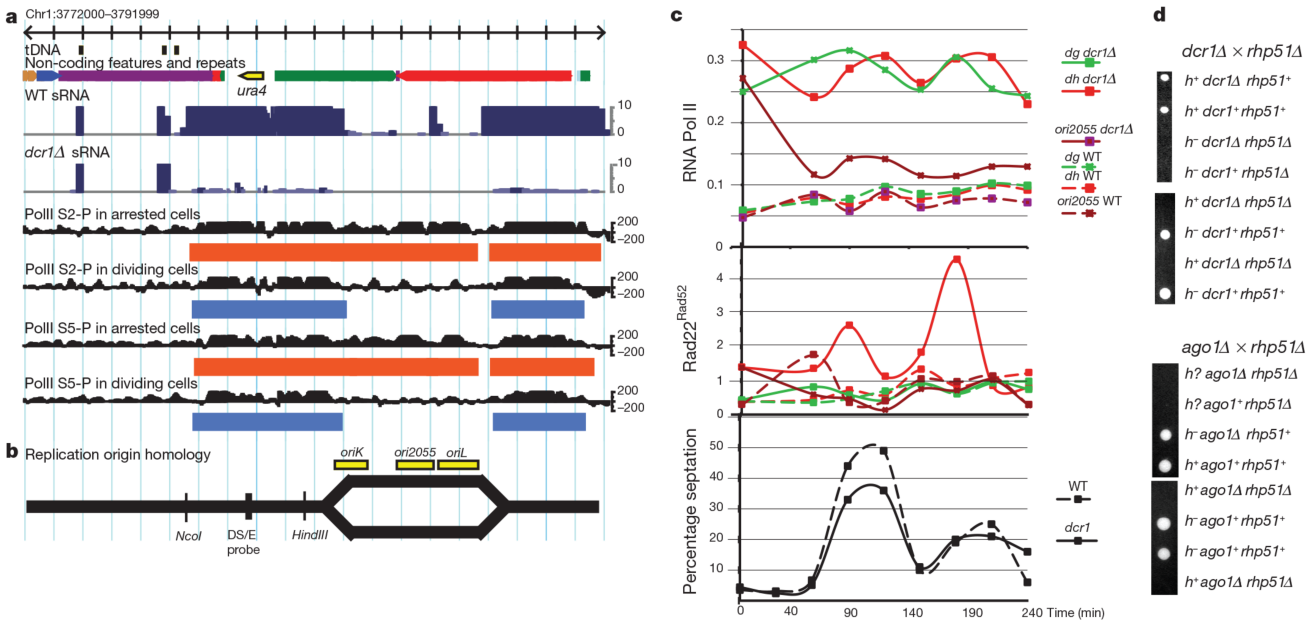

Figure 2.2 | RNA interference and DNA replication restrict RNA polymerase II accumulation and prevent DNA damage. A) Small RNA (blue) and Pol II ChIP-seq reads (black) and regions of significant enrichment (blue and red rectangles) from wild type and *dcr1*Δ on the right arm of centromere 1. B) A replication bubble is shown, initiated at one of the three origin homology regions at centromere 1 (yellow boxes). C) Chromatin 32mmunoprecipitation for RNA Pol II and Rad52 from hydroxyurea-arrested and released wild-type (dashed lines) and *dcr1*Δ (solid lines). Cell cycle progression after release from hydroxyurea block is monitored by septation index, which peaks coincident with S-phase. D) Representative parental and non-parental di-type tetrads from crosses between *rad51*Δ cells, defective in homologous recombination, and *dcr1*Δ or *ago1*Δ.

2.3.3 Replication fork stalling occurs in centromeric heterochromatin

Stalled replication forks can be detected at the molecular level using 2D gel electrophoresis. Signal intensity in the Y-arc can indicate a stalled replication fork, and an X-spike indicates replication intermediates, for example those formed during HR (Figure 2.3a). We performed 2D gel electrophoresis with probes against an inserted centromeric reporter gene (*ura4*$^{+}$, Figure 2.2b). In WT cells we saw both strong pausing in the Y-arc and an X-spike (Figure 2.3b). The intensity of the pausing and X-spike was correlated with centromeric transcription levels. Intensity increased in *dcr1*Δ cells (Figure 2.3c), with increased transcription, while it decreased in *mms19*Δ cells (Figure 2.3d), with

decreased transcription. In the absence of heterochromatin the X-spike (*swi6*Δ, Figure 2.3e) or both pausing and X-spike were lost (*clr4*Δ, Figure 2.3f). This may result from changes in replication timing in the absence of heterochromatin (Li et al., 2013) that would presumably de-synchronize transcription and DNA replication.

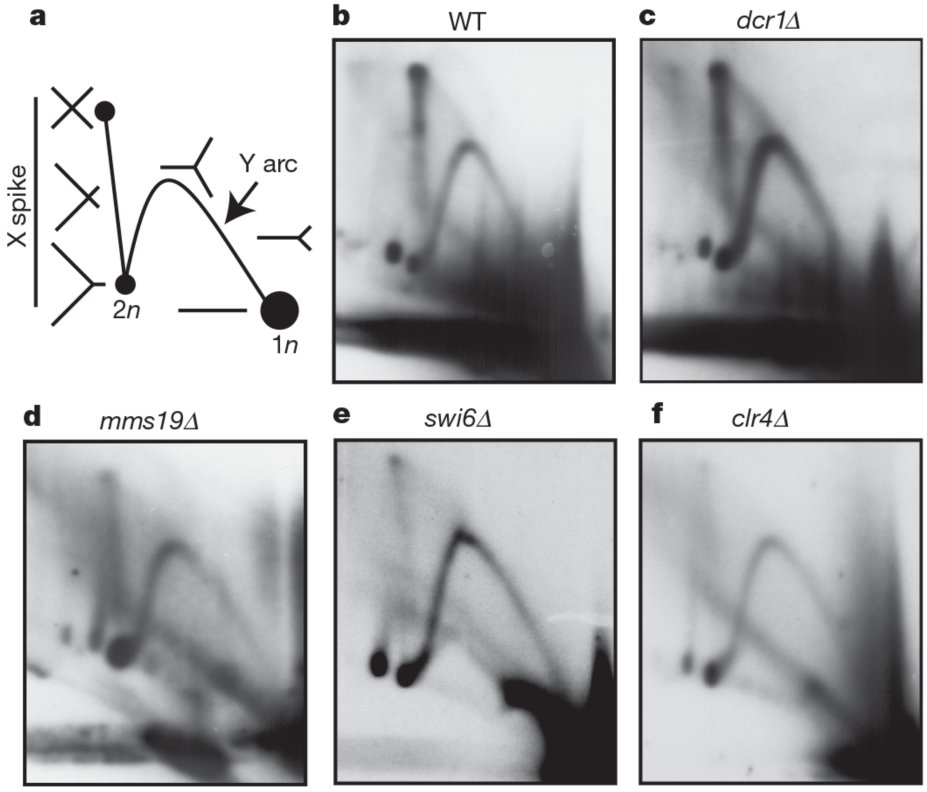

Figure 2.3 | Replication fork stalling during heterochromatin replication. Replication intermediates in wild-type and mutant cells resolved by 2D gel electrophoresis and probed with the unique DS/E probe from the ura4 transgene within the dg repeat on chromosome 1 (Fig. 2.2a). A) Schematic of replication intermediates in 2D gels indicates joint molecules (X), and forks (Y). B-F) Junction molecules indicate fork stalling in WT (B) and *dcr1*Δ mutant cells (C), and are reduced in *mms19*Δ (d), *swi6*Δ (e) and *clr4*Δ (f).

2.4 Discussion

The organization and dynamics of fission yeast centromeres produces an environment where transcription and DNA replication collide. Within the centromeres, RNAi has a well-described co-transcriptional role in regulating Pol II transcription. We found that this roll is of particular importance to resolve transcription-replication collisions resulting from S-phase transcription of centromeric repeats (Figure 2.4a). Without Pol II release by RNAi replication forks are restarted by homologous recombination, which allows replication to proceed, but results in epigenetic changes at stall sites. Such fork restart might disrupt interactions between the CLRC and DNA replication machinery, and prevent the spreading of H3K9me2, explaining why PEV is dependent on RNAi (Figure 2.4b). This presents a new paradigm whereby RNAi promotes heterochromatic silencing through replication-coupled release of RNA Pol II.

It is well documented that head on Pol II transcription causes replication fork pausing (Prado and Aguilera, 2005), that can be restarted by the activity of homologous recombination (Carr and Lambert, 2013; Lambert et al., 2010). Such a process must also involve the removal of Pol II if replication is to proceed, however the mechanism behind this is unknown. Chromatin remodeling plays an integral role in HR (Lans et al., 2012), so it's possible that this remodeling contributes to Pol II removal. Rad54, a key component of HR, is a swi2/snf2 family chromatin remodeler that translocates approximately 10 kb from Rad51 nucleation sites (Amitani et al., 2006) and possesses nucleosome remodeling activities (Alexeev et al., 2003), which could remove Pol II and other chromatin bound proteins.

Whether RNAi factors are recruited to stalled RNA Pol II, or simply act during the elongation phase of transcription remains unknown. There are a few possibilities for the release of stalled Pol II. In one case the CLRC, with its similarities to transcription coupled repair pathways, could recognize stalled Pol II and recruit RNAi components. A second possibility is that RNAi recognizes stalled Pol II, and recruits the CLRC, which could be involved in releasing Pol II, perhaps through ubiquitination and degradation by Cul4 and Rik1. Finally, the cleavage of nascent RNA molecules from transcribing Pol II by either Dicer or Argonaute would produce a free 5' end. This aberrant 5' end could be a substrate for a 5' to 3' exonuclease, which could release RNA Pol II through a "torpedo" model. A similar model of Pol II

release has been proposed at some genes in *S. cerivisiae*, mediated by hairpin formation and RNAse III cleavage (Ghazal et al., 2009).

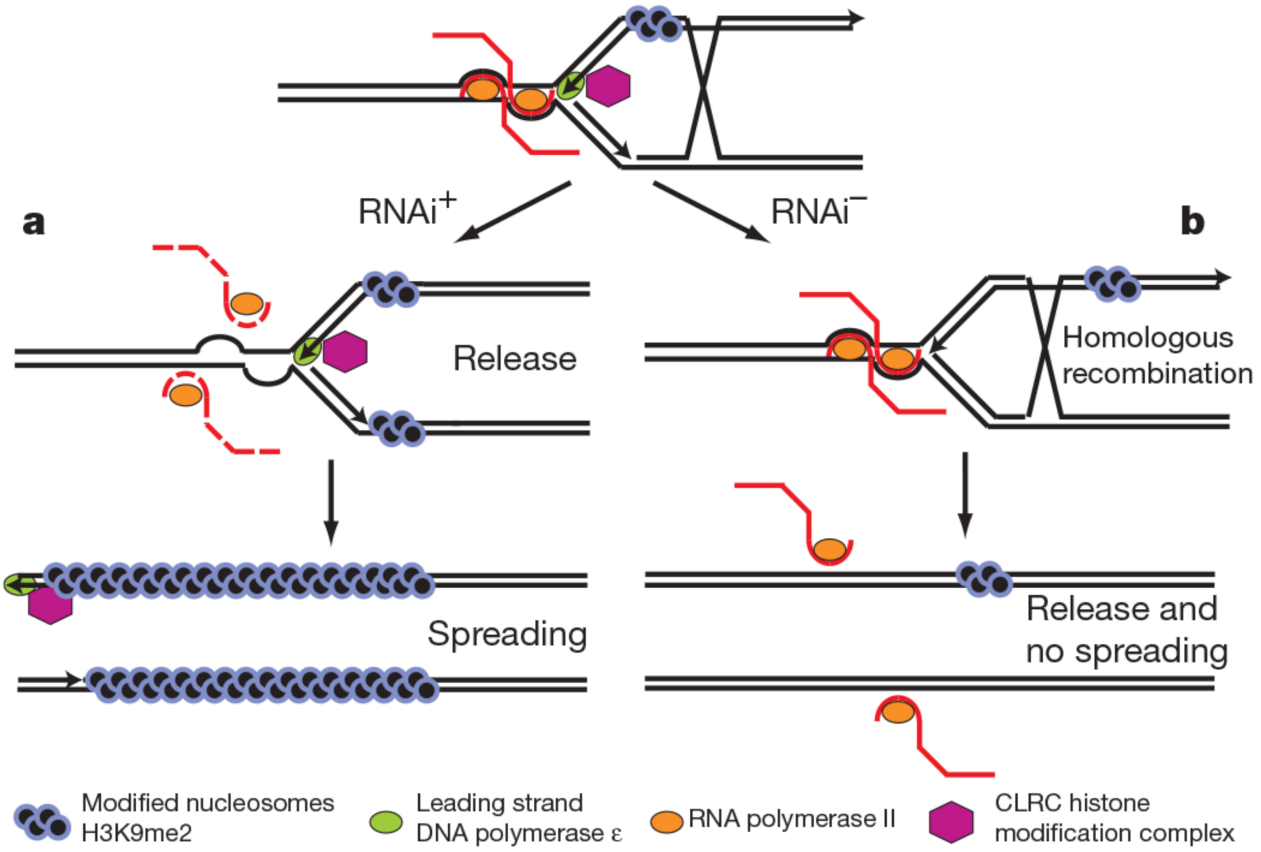

Figure 2.4 | Replication-coupled transcriptional silencing through histone modification and RNAi. A) The Rik1 complex (red hexagon) is recruited to heterochromatic replication forks by interactions with methylated histone H3K9me2 and with the leading strand DNA polymerase (Pol e, green). Swi6 induces origin firing, but collision with RNA polymerase II (orange) stalls replication forks. RNAi releases Pol II by processing of pre-siRNA transcripts (red lines), allowing leading strand DNA polymerase to complete DNA replication and the associated Rik1 histone modification complex (red hexagon) to spread histone modification (black circles). B) In the absence of RNAi, origins fire but Pol II is not released, stalling replication forks. Stalled Pol II signals repair via homologous recombination instead. Recombination could in principle occur with sister chromatids (shown here) or with other copies of the same repeat (not shown). DNA polymerase e and the associated Rik1 complex are lost along with the replisome, and fail to spread histone modification into neighboring reporter genes.

The observation that transcription and heterochromatin replication timing can be uncoupled in *swi6*Δ and *clr4*Δ mutants begs the question of why collision need happen in the first place. Hypothetically, a pool of centromeric siRNA could be generated earlier in the cell cycle that could direct H3K9me2 during S-phase, preventing collision. Without transcription during S-phase however, siRNA and target nascent RNA base pairing could not occur, rendering RNAi unable to direct heterochromatin. It's also possible that histone modifications can only be established during DNA replication, perhaps because of the coupling between the CLRC and replisome. Therefore to satisfy both the requirement of targeting and histone modification S-phase transcription may be unavoidable, and thus in organisms with RNAi the pathway has also evolved to resolve this conflict.

Silencing by small RNA pathways indicates a necessity to silence elements in *trans*. A good example in metazoans is the piRNA pathway in *Drosophila* described in the introduction. In this system heterochromatic piRNA clusters are transcribed to produce piRNA that can silence transposons genome wide in *trans*. Some transposons are expressed during S-phase (Zhang et al., 2014), mimicking the scenario seen within fission yeast centromeres, which themselves have been hypothesized to be remnants of ancient transposons (Wong and Choo, 2004). Retrotransposons block replication forks as part of mechanism to maintain their copy number, and CENP-B counteracts this stimulating fork progression (Zaratiegui et al., 2010). Like transposons, centromeric repeats are bound by CENP-B (Nakagawa et al., 2002) and block replication forks, supporting a transposon origin for centromeric repeats. These observations suggest that the control of DNA replication may have been an important factor during the co-evolution of RNAi and the transposable elements it controls.

2.5 Contributions

Much of this work (Chapter 2) was done in collaboration with both lab members and other research groups. Collaborators performed the following experiments:

- sRNA Sequencing (Figure 2.1b) – experiment and analysis performed by Mikel Zaratiegui.
- Cell Cycle Pol II and Rad52 ChIP-qPCR (Figure 2.2c) – experiment and analysis performed by Danielle Irvine.
- *rad51* and RNAi synthetic lethality (Figure 2.2d) – experiment performed by Jie Ren.
- 2D gels of centromeric heterochromatin replication (Figure 2.3) – experiment performed by Paco Antequera's lab.

3. Transcriptional termination by Dicer at sites of replication stress maintains genomic integrity

This chapter has been reproduced with modifications from:

> Castel SE, Ren J, Bhattacharjee S, Chang AY, Sanchez M, Valbuena A, Antequera F, Martienssen RA. Transcriptional termination by Dicer at sites of replication stress maintains genomic stability.

3.1 Abstract

Nuclear RNA interference is an important regulator of transcription and epigenetic modification, but the underlying mechanisms remain elusive. Using a genome-wide approach in the fission yeast *S. pombe* we have found that Dicer, but not other components of the canonical RNAi pathway, releases Pol II from the 3' end of highly transcribed genes, and, surprisingly, from antisense transcription of rRNA and tRNA genes, which are normally transcribed by Pol I and Pol III. These Dicer-terminated loci correspond to sites of replication stress and DNA damage, likely resulting from transcription-replication collisions. At the rDNA loci, release of Pol II facilitates DNA replication and prevents homologous recombination, which would otherwise leads to loss of rDNA repeats especially during meiosis. Our results reveal a novel role for Dicer-mediated transcription termination in genome maintenance and may account for widespread regulation of genome stability by nuclear RNAi in higher eukaryotes.

3.2 Introduction

Nuclear RNA interference (RNAi) has emerged as an important regulator of gene expression and epigenetic inheritance in eukaryotes, and studies of fission yeast centromeres have provided mechanistic insight into the process by which RNAi directs epigenetic modification (Bühler and Gasser, 2009; Castel and Martienssen, 2013; Goto and Nakayama, 2012; Grewal, 2010; Lejeune et al., 2011). In *S. pombe,* RNAi is required to direct H3 lysine 9 methylation (H3K9me2) and H3K4

demethylation within the heterochromatic repeats flanking each centromere (Volpe et al., 2002). Tightly regulated transcription within these repeats (Djupedal et al., 2005) leads to the production of double stranded RNA (dsRNA), in part via RNA dependent RNA polymerase (Rdp1), that is processed into small interfering RNA (sRNA) by the sole Dicer in *S. pombe,* Dcr1 (Colmenares et al., 2007). sRNA are loaded into Argonaute (Ago1), guiding it back to complementary nascent RNA (from transcribing RNA Polymerase II) where it directs the deposition of H3K9me2 through the histone methyltransferase Clr4 via "co-transcriptional gene silencing"(Bühler et al., 2006; Irvine et al., 2006).

Similar mechanisms of RNAi based silencing have been discovered in higher eukaryotes. In *Arabidopsis* the RNA-directed DNA methylation (RdDM) pathway co-transcriptionally directs *de novo* cytosine methylation at loci transcribed by RNA Pol V (Law and Jacobsen, 2010). In the *C. elegans* germline the 21U small RNA pathway directs H3K9 methylation through 22G-loaded Argonautes in the nucleus, closely resembling *S. pombe* (Shirayama et al., 2012). Classically, these silencing pathways have been thought to act on heterochromatic repetitive elements, such as transposons, but more recently a broader role at euchromatic genes has been discovered. Studies in *Arabidopsis* (Liu et al., 2012), *Drosophila* (Cernilogar et al., 2011), *C. elegans* (Guang et al., 2010), and *S. pombe* (Gullerova et al., 2011; Gullerova and Proudfoot, 2008; Woolcock et al., 2012) have implicated nuclear small RNA pathways in the regulation of Pol II at individual euchromatic genes. In fission yeast, this conserved function of RNAi (Pol II release) is particularly important in the context of DNA replication. Centromeric repeat units in *S. pombe* are transcribed during S-phase, the time at which DNA replication occurs and epigenetic marks must be re-established (Chen et al., 2008; Kloc et al., 2008). This leads to a collision between Pol II and the replisome that is resolved by RNAi through the release of Pol II (Zaratiegui et al., 2011). In the absence of RNAi stalled replication forks are restarted through homologous recombination (HR), and this results in the loss of epigenetic modifications (Zaratiegui et al., 2011).

We have found that Dicer coordinates transcription and replication outside of pericentromeres, identifying a novel role for Dicer in transcription termination and maintaining genomic stability. Pol II accumulation is a hallmark of polymerase collision, and we found an increase in Pol II enrichment in *dcr1*Δ cells at previously uncharacterized loci including protein coding genes, tDNA, and rDNA. Dicer-dependent sRNAs were detected at these loci, but transcriptional termination was not dependent on

other RNAi pathway components, demonstrating for the first time a Dicer-specific role in Pol II release. These loci are strongly correlated with sites of replication pausing, and thus likely represent collisions between transcription and replication (Bermejo et al., 2012). We focused on one particularly striking and unexpected site of Pol II regulation, the rDNA repeats, where we found that Dicer is required for rDNA copy number maintenance. Our findings suggest that in *S. pombe* Dicer has a genome wide role in terminating transcription by releasing Pol II at sites of collision between transcription and replication, and thus maintains genome stability.

3.3 Results

3.3.1 Dicer has a genome wide role in Pol II regulation

To identify sites transcriptionally regulated by Dicer we profiled Pol II accumulation in WT and *dcr1*Δ mitotic cells by ChIP-seq using antibodies raised against both "poised" (S5 phosphorylated, pS5) and "elongating" (S2 phosphorylated, pS2) forms of Pol II. We observed genome wide effects on Pol II enrichment in the absence of Dicer that were not limited to the centromeric repeats (Figure 3.1a). Indeed, striking regions of enrichment were visible not only at the centromeres, but also within the subtelomeric rDNA repeats on chromosome III, which are normally transcribed by Pol I. A total of 224 high-confidence (FDR <= 0.01) regions of increased Pol II enrichment in *dcr1*Δ (as compared to WT) were identified using both antibodies and replicates. Features found within these regions largely contained protein coding genes, non-coding RNA (ncRNA), centromeric repeats, tDNA, and rDNA (Figure 3.1b).

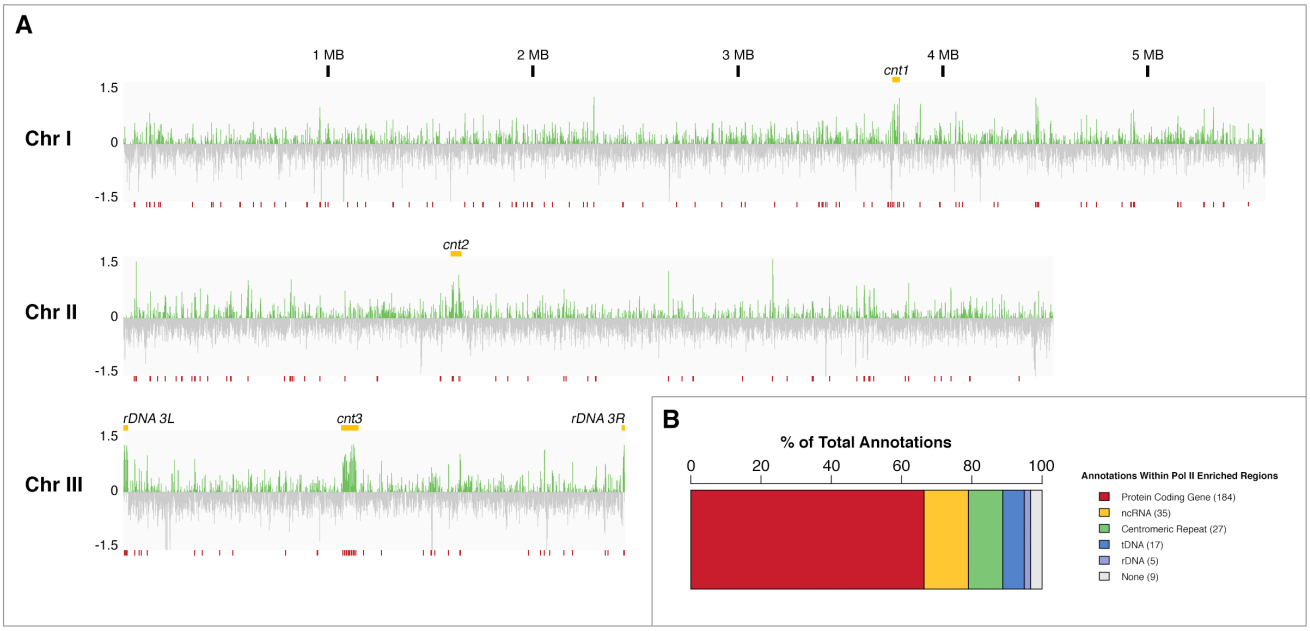

Figure 3.1 | Pol II transcriptional profile by ChIP-seq reveals novel Dicer-regulated loci. A) Chromosome wide view of $\log_2$(ratio) between *dcr1*Δ and WT Pol II enrichment. Repeat features including centromeres and rDNA clusters are indicated (yellow). Regions of statistically significant (FDR <= 0.01) enrichment across combined initiating and elongating Pol II replicates are indicated (red). B) Count of annotated features contained within regions of increased Pol II enrichment.

3.3.2 Dicer releases Pol II at the 3' end of highly transcribed genes

Many sites of significant Pol II enrichment fell within protein coding genes, and within these genes enrichment was most often found within the open reading frame (ORF) and at the 3' end, rather than in the promoter region (Figure 3.2a). We then calculated Pol II enrichment specifically at all protein coding genes and found that 235 genes showed a significant (FDR <= 0.01) increase in *dcr1*Δ as compared to WT (Table 3.1). Importantly, the few previously experimentally validated Dicer-regulated protein coding genes (*hsp16*, *hsp104*, *hsp9*) (Woolcock et al., 2012) were present in our list, validating the approach.

We noticed a striking pattern of Pol II enrichment at the 3' end of Dicer-regulated genes accompanied by sense Dicer-dependent small RNA (sRNA), exemplified in Figure 3.2b. Meta-analysis of Dicer-

regulated protein coding genes showed an increase in elongating Pol II occupancy peaking sharply at the 3' end (Figure 3.2c), suggesting a defect in transcriptional termination. This peak was accompanied by a peak in Dicer-dependent sRNA sense to protein coding transcripts indicative of direct Dicer activity. Dicer-dependent sense sRNA at the 3' end of Dicer targets was significant compared to all other genes (~3.75x, $p < 0.0001$) (Figure 3.3a) and matched the expected size distribution for Dicer products (Figure 3.3b). The presence of sense only sRNA suggested that Dicer might be acting on hairpins at the 3' end of genes. We used mfold to predict RNA structure at 3' end of the example gene shown in Figure 2b and found a hairpin candidate that matched both sRNA peaks (Figure 3.3c).

We performed RNA-seq to distinguish termination defects at Dicer-regulated genes from transcriptional increase. Failure to release Pol II during termination can result in a decrease in RNA transcript levels (Padmanabhan et al., 2012; West and Proudfoot, 2009). In *dcr1*Δ cells we saw a decrease of at least 20% in transcript level at 75% of Dicer-regulated genes, supporting a defect in termination (Figure 3.3d). Importantly, the decrease in expression of Dicer-regulated genes was significant when compared to all other protein coding genes ($p < 0.05$, Figure 3.3e). We did not see any evidence of run-on transcription at Dicer-regulated genes (Figure 3.3f), suggesting that Dicer is involved in the release but not pausing of Pol II during termination.

A GO annotation analysis of these genes revealed enrichment in many core cellular processes, most substantially in translation (Table 3.2). These categories contain many highly transcribed genes, and suggested that this may be a common feature. Indeed Dicer-regulated genes were highly transcribed in WT cells as compared to global gene transcription levels ($p < 0.001$, Figure 3.2d).

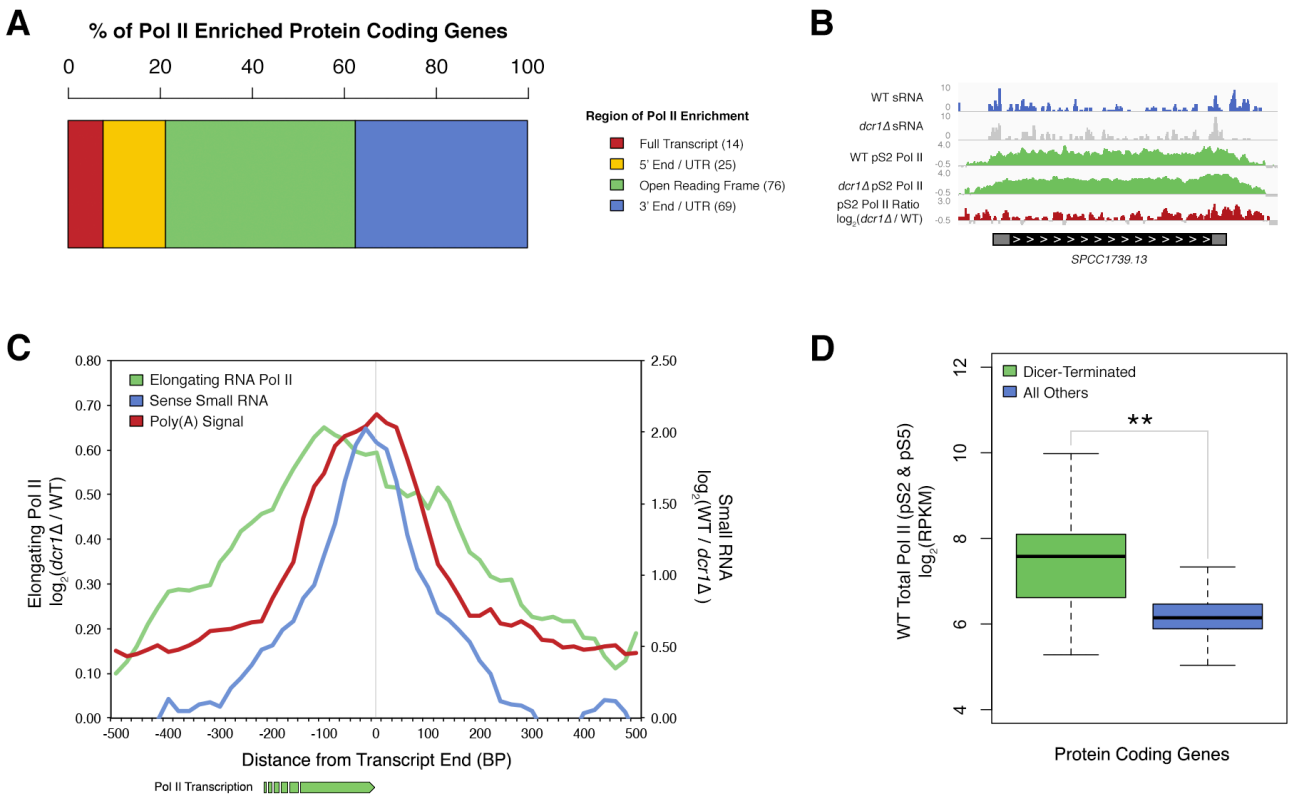

Figure 3.2 | Dicer releases Pol II at the 3' end of highly transcribed genes. A) Region of Pol II enrichment within protein coding genes identified in Figure 1b. B) Example of a Dicer-terminated gene (*SPCC1739.13*). Sense sRNA reads in WT (blue) and *dcr1*Δ (grey), pS2 Pol II enrichment in WT and *dcr1*Δ (green), $\log_2$(ratio) between *dcr1*Δ and WT pS2 (red). C) Average elongating Pol II enrichment (green, $\log_2$(*dcr1*Δ / WT)), sRNA (blue, $\log_2$(WT / dcr1Δ)), and Poly(A) signal at the 3' end of all Dicer-regulated protein coding genes (listed in Table 3.1). Enrichment averaged over a 100bp sliding window for ease of viewing. D) WT Pol II enrichment (average of initiating (pS5) and elongating (pS2)) at all Dicer-regulated protein coding genes (blue) vs all other mRNA annotations (green) (** = $p < 0.01$). See Table 3.2 for GO Term enrichment of Dicer-terminated genes.

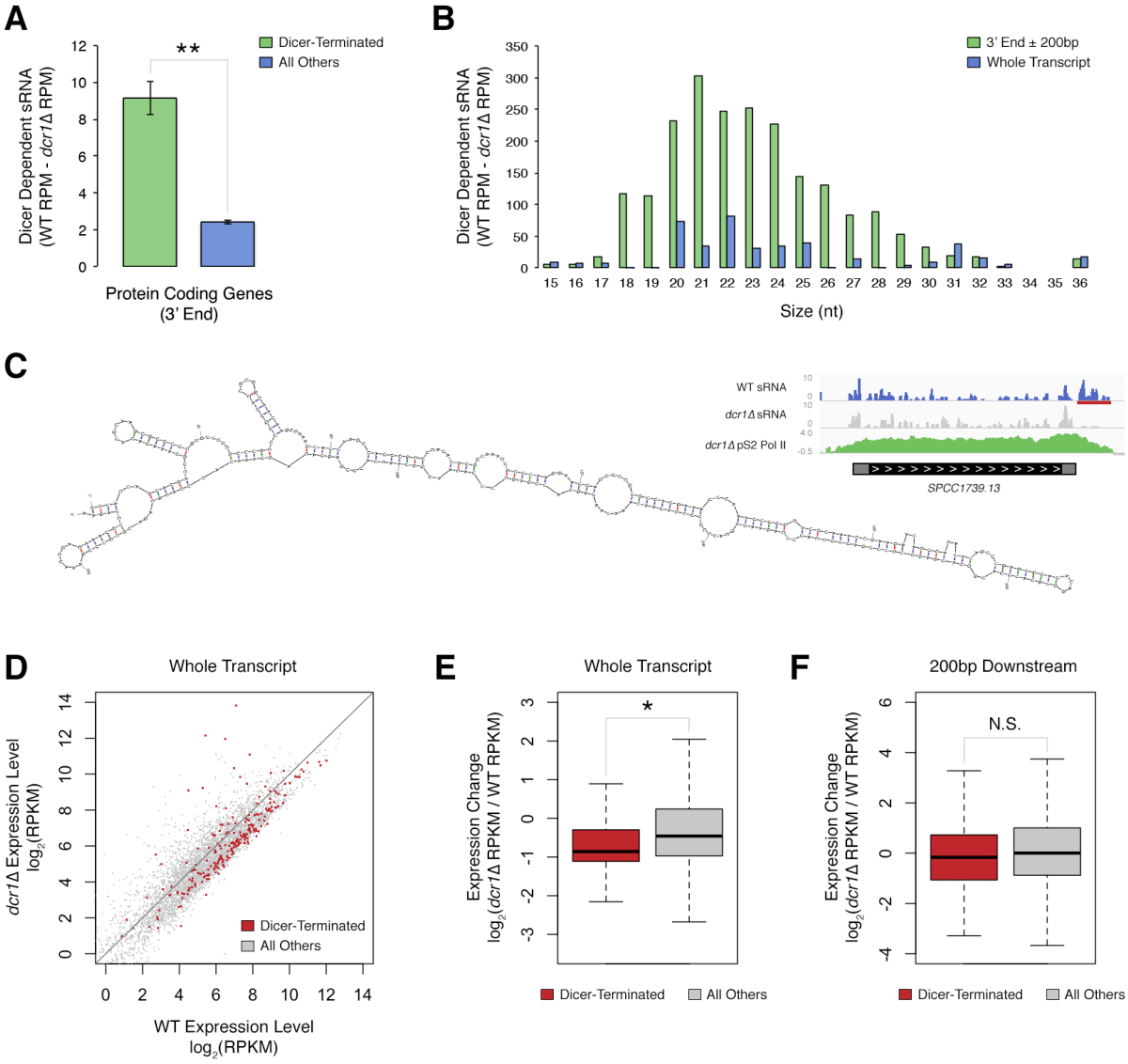

Figure 3.3 | Dicer activity at protein coding genes. A) Average sense Dicer dependent small RNA at the 3' end of Dicer-terminated genes compared to all other protein coding genes. 3' end was defined as the transcription termination site (TTS) ± 200bp, corresponding to the sRNA peak in Figure 2C. Dicer dependent sRNA level calculated as WT RPM - *dcr1*Δ RPM. Data are represented as mean ± SEM. The significance of differences between groups is indicated (** = $p < 0.00001$). B) Size distribution of sense Dicer dependent sRNA at the 3' end compared to the whole transcript of Dicer-terminated genes. 3' end was defined as TSS ± 200bp. Dicer dependent sRNA level in each size bin calculated as WT RPM - *dcr1*Δ RPM. C) mfold prediction of RNA secondary structure at an example Dicer-terminated protein coding gene (*SPCC1739.13*), also shown in Figure 2C. Regions corresponding to Dicer dependent sRNA are highlighted (red). D) Normalized expression level (RPKM) of Dicer-terminated (red) or all other (grey) protein coding genes in WT and *dcr1*Δ

cells determined by RNA-seq. E) Expression change calculated as $\log_2$(*dcr1*Δ RPKM / WT RPKM) at either Dicer-terminated (red) or all other (grey) protein coding genes. The significance of differences between groups is indicated (* = $p < 0.05$). F) Expression change calculated as $\log_2$(*dcr1*Δ RPKM / WT RPKM) in 200bp regions downstream of either Dicer-terminated (red) or all other (grey) protein coding genes. The significance of differences between groups is indicated. For RNA-seq analysis only regions containing at least 1 read in both WT and *dcr1*Δ libraries were considered.

3.3.3 Dicer releases Pol II from antisense tDNA and rDNA transcription

Surprisingly, many tRNA genes (tDNA), which are normally transcribed by RNA Pol III, were found within regions of Pol II accumulation. Because single tDNA are very short (<100bp) we assessed Pol II occupancy at all chromosomal tDNA and found an increase in *dcr1*Δ vs WT (Figure 3.4a). Elongating Pol II peaked at the 5' end of tRNA genes, and there was an accompanying peak of antisense Dicer-dependent sRNA, suggesting antisense Pol II transcription. When quantified individually, 108 of 171 tDNA showed an increase of at least 20%, whereas only 4 showed a decrease (Figure 3.4b, Table 3.3). The increase was statistically significant across replicates for 37 of the 108, and there was no bias towards pericentromeric tDNA. sRNA that peaked at the site of Pol II accumulation was antisense to tRNA, Dicer-dependent (Figure 3.4c), and fell within the expected size range for Dicer products when compared to sense tRNA processing fragments (Figure 3.4d).

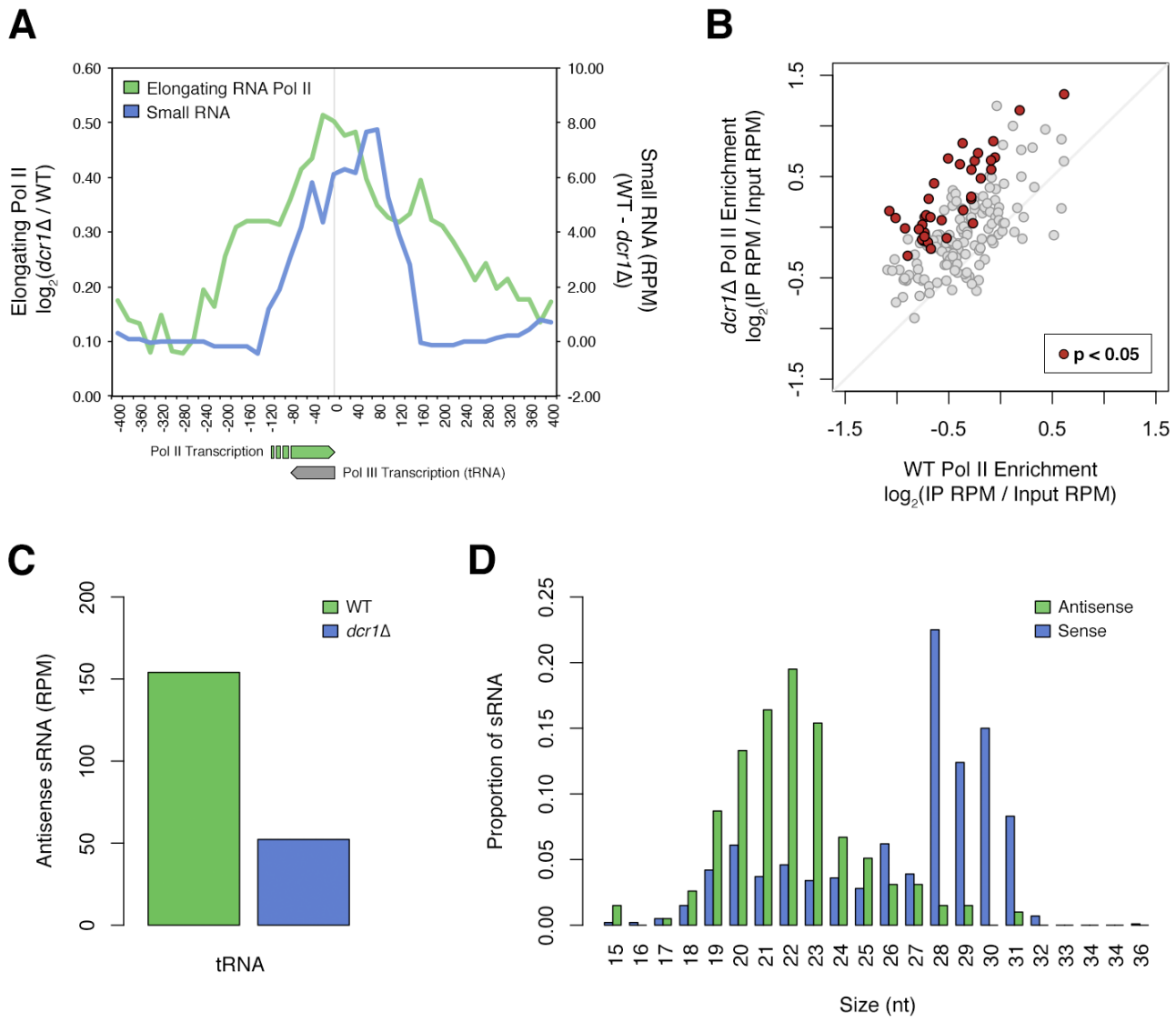

Figure 3.4 | Dicer releases Pol II from antisense Pol II transcription at tDNA. A) Average elongating Pol II enrichment (green, $\log_2(dcr1\Delta$ / WT)), and sRNA level (RPM, WT – *dcr1*Δ) at nuclear tRNA genes. Enrichment averaged over a 100bp sliding window for ease of viewing. Direction of Pol III tRNA transcription and antisense Pol II transcription are indicated. B) Average of pS5 and pS2 Pol II enrichment at each of the *S. pombe* genomic tDNA, for both *dcr1*Δ and WT. tDNA with statistically significant ($p < 0.05$) Pol II enrichment are indicated in red (listed in Table 3.3). C) Normalized read counts of antisense sRNA mapping to tDNA in WT (green) and *dcr1*Δ (blue). D) Size distribution of antisense (green) and sense (purple) sRNA mapping to tDNA in WT cells.

Similar to tDNA, but more striking at the genome-wide scale was Pol II accumulation within the subtelomeric rDNA repeats on chromosome 3 (Figure 3.1a and Figure 3.5a). Enrichment of both poised and elongating Pol II at rDNA repeats was significantly increased in *dcr1*Δ vs WT when quantified across replicates ($p < 0.05$) (Figure 3.5b). Poised Pol II peaked at the 3' end of Pol I transcripts while elongating Pol II peaked at the 5' end, again suggesting antisense transcription. We also identified a population of Dicer-dependent sRNA antisense to 35S rRNA that peaked at the 3' end of Pol II transcription (Figure 3.5 a and c). These sRNA fell within the expected size range for Dicer products, unlike sense rRNA fragments, which were more evenly distributed and most likely degradation products (Figure 3.5d). Consistently, overexpression of *dcr1* even in the absence of the RNA-dependent RNA polymerase (Rdp1) results in a dramatic increase in sRNA levels antisense to both rDNA (~38.5 fold) and tDNA (~4.5 fold) while a comparable increase in sense sRNA is not seen (Yu et al., 2013).

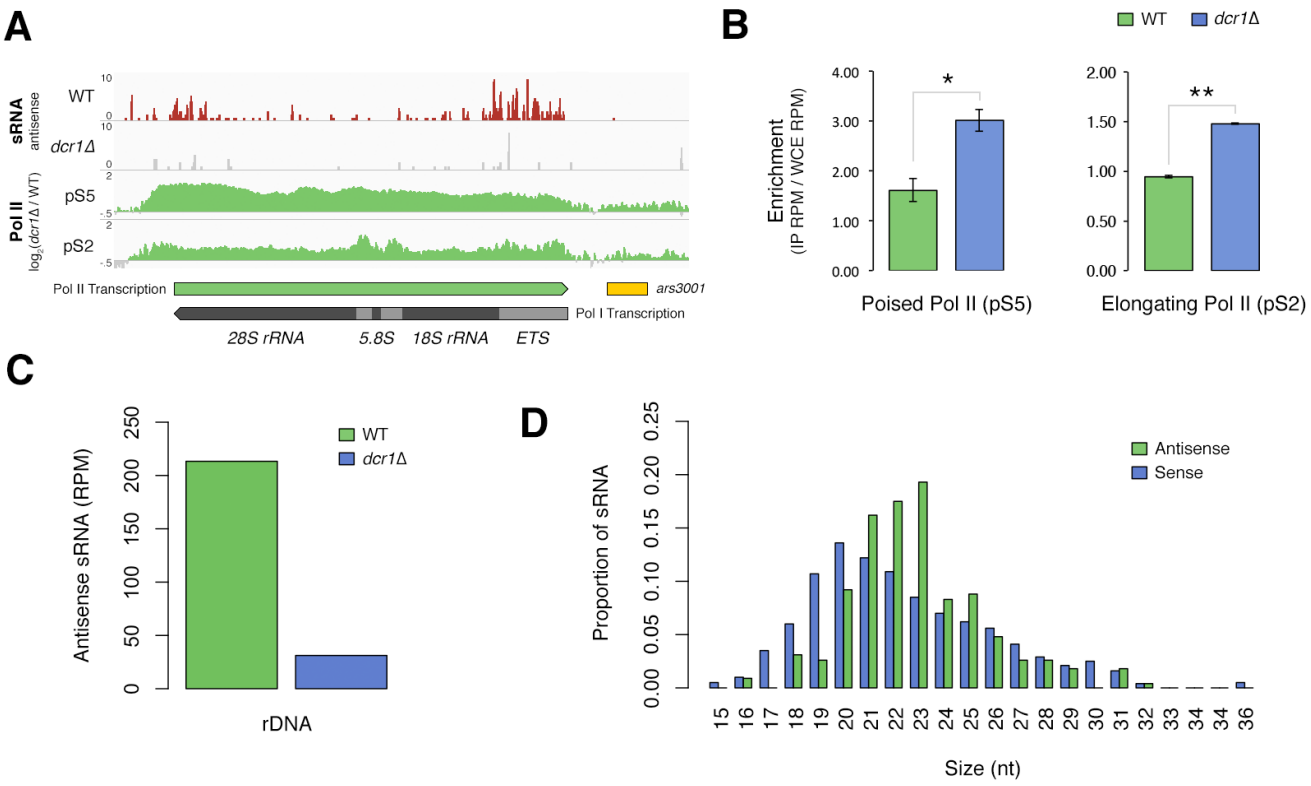

Figure 3.5 | Dicer releases Pol II from antisense transcription at subtelomeric rDNA. A) Distribution of sRNA reads mapping antisense to repetitive subtelomeric rRNA genes in WT (red) and Dicer (grey) and both poised and elongating Pol II accumulation in *dcr1*Δ vs WT (green, $\log_2$(fold change)). Direction of Pol I rRNA transcription and antisense Pol II transcription are indicated. Annotations for 18S, 5.8S, 28S rRNA genes, externally transcribed spacer (ETS) and replication origin containing region (ars3001) are shown. B) Quantification of poised (pS5) and

elongating (pS2) Pol II enrichment within subtelomeric rDNA repeat regions containing sRNA for WT (green) and *dcr1*Δ (blue). Data are represented as mean ± SEM. The significance of differences across replicates is indicated (** = $p < 0.01$, * = $p < 0.05$). C) Normalized read counts of antisense sRNA mapping to rDNA in WT (green) and *dcr1*Δ (blue). D) Size distribution of antisense (green) and sense (purple) sRNA mapping to rDNA in WT cells.

3.3.4 The canonical RNAi pathway is not involved in Pol II release at novel Dicer-terminated loci

Transcriptional gene silencing (TGS) in *S. pombe* occurs when sRNA generated by Dicer are loaded into Argonaute (Ago1) and direct H3K9me2 deposition at target loci (Castel and Martienssen, 2013). We tested the involvement of the RNAi pathway in transcriptional regulation at tDNA and rDNA by performing Pol II ChIP-qPCR in *ago1*Δ cells alongside *dcr1*Δ and WT. We saw no increase in Pol II enrichment at tDNA or rDNA between *ago1*Δ and WT, unlike in *dcr1*Δ (Figure 3.6a), while as expected both *ago1*Δ and *dcr1*Δ showed a similar increase in Pol II enrichment at the centromeric *dh* repeat (Figure 3.6a).

An indicator of RNAi mediated chromatin silencing is repressive H3K9 methylation at target sites, and both tDNA and rDNA are enriched for this mark in *S. pombe* (Figure 3.6b). We assessed the contribution of H3K9 methylation to transcriptional regulation by performing H3K9me2 ChIP-seq in WT and *dcr1*Δ cells. There was no decrease in H3K9me2 at either protein coding genes, tDNA or rDNA in *dcr1*Δ, while a decrease at centromeric repeats was seen as expected (Figure 3.6b). In fact, there was a slight (~ 10%, $p < 0.05$) increase of H3K9me2 at novel Dicer targets that likely represents higher background levels in *dcr1*Δ samples due to the absence of centromeric heterochromatin.

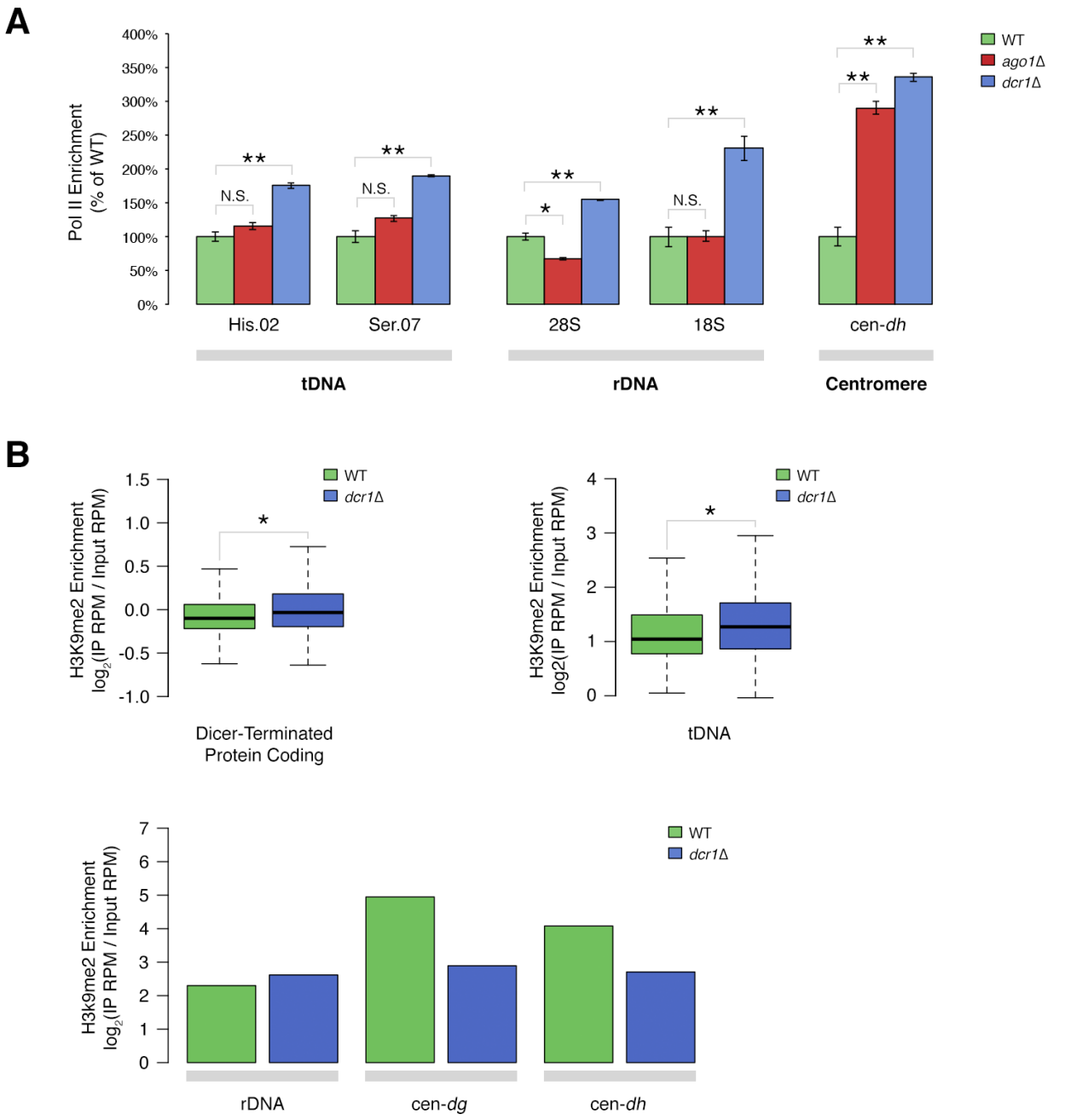

Figure 3.6 | The canonical RNAi pathway is not involved in Pol II transcriptional regulation at tDNA and rDNA. A) pS2 Pol II enrichment by ChIP-qPCR at representative euchromatic tDNA and rDNA loci with centromeric *dh* repeat included for comparison. ChIP enrichment quantified relative to input and displayed as % of WT. Data are represented as mean ± SEM. B) H3K9me2 ChIP-seq enrichment at Dicer-regulated protein coding genes, tDNA, rDNA, and centromeric repeats in WT (green) and *dcr1*Δ (blue). The significance of differences is indicated (** = $p < 0.01$, * = $p < 0.05$).

3.3.5 Dicer terminates transcription at sites of replication stress

The novel Dicer-regulated loci (actively transcribed genes, tDNA, and rDNA) that we have identified all represent 'difficult to replicate' regions because frequent passage of transcription complexes creates a barrier to DNA replication that can stall forks (Alzu et al., 2012; Sabouri et al., 2012). The homologous recombination (HR) protein Rad52 is recruited to stalled replication forks, which may eventually collapse leading to DNA damage and checkpoint (Chk1) activation via Crb2 (Nakamura et al., 2004). We performed both Rad52 and Crb2 ChIP-Seq to identify stalled and collapsed forks and to correlate these with Dicer-terminated genes. We found a strong correlation between both Rad52 and Crb2 enriched genes and Dicer-terminated genes in WT cells (Figure 3.7a), with over 55% of Dicer-terminated genes enriched for at least one protein. This correlation was also seen in *dcr1*Δ cells, however there was a large decrease in Rad52, while Crb2 remained similar (Figure 3.7b). We observed a similar correlation at tRNA genes, but unlike protein coding genes the number of Crb2-enriched tRNA genes increased by 30% in *dcr1*Δ. We validated the correlation between Rad52 enrichment and replication pausing by 2D gel electrophoresis at a protein-coding gene (*hsp90*) and a tDNA cluster, both of which showed significant Rad52 enrichment and Pol II accumulation (Figure 3.7c). We found that replication pausing occurs at these loci, indicating that sites of Rad52 enrichment detected by ChIP-seq are bona-fide difficult-to-replicate regions (Figure 3.7c).

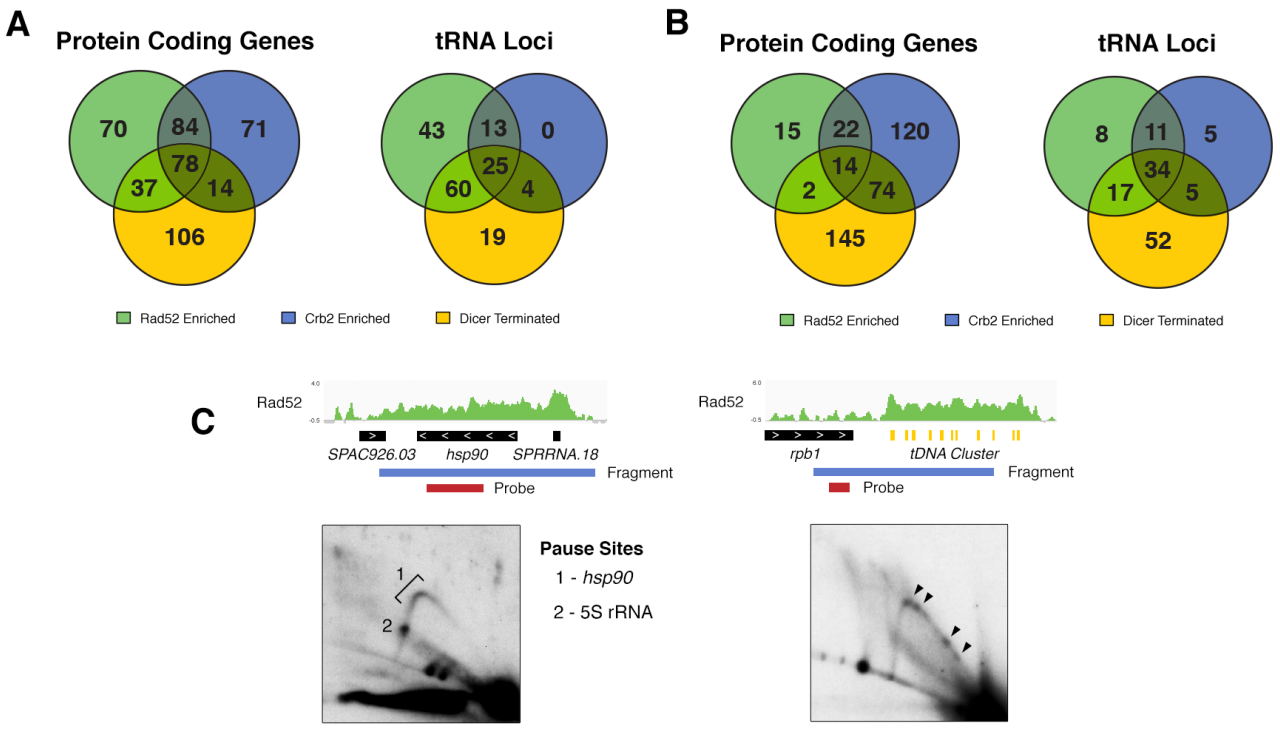

Figure 3.7 | Dicer releases Pol II at sites of replication stress. A) Overlap between Rad52 enrichment, Crb2 enrichment and Dicer termination at either protein coding or tRNA genes in WT cells. B) Overlap in *dcr1*Δ cells. C) Rad52 enrichment at *hsp90* and adjacent 5S rRNA gene determined by ChIP-seq, and accompanying 2D gel of fragment containing both features. Sites of replication pausing, within the 5S rRNA gene and *hsp90* are indicated. Rad52 enrichment in WT cells at the tDNA cluster in the left outer arm of centromere 2, and accompanying 2D gel of fragment containing the cluster. Arrowheads indicate sites of major pausing.

Somewhat surprisingly Rad52 is lost from many of the Dicer-terminated genes in *dcr1*Δ. Analysis of sequencing read distribution within Rad52 peaks clearly shows that in *dcr1*Δ the bulk (72%) of Rad52 is localized to the rDNA repeats (Figure 3.8a). Indeed we have previously observed a substantial increase in Rad52 foci in *dcr1*Δ cells as compared to WT (Zaratiegui et al., 2011), however the total level of Rad52 remains unchanged (data not shown), indicating that the Rad52 pool is limited. Rad52 nucleation occurs at sites of DNA damage during S-phase and subsequently spreads from the stall site (Zhou et al., 2013). To determine the precise location of replication stalling within rDNA we synchronized cells and performed Rad52 ChIP in S-phase. During S-phase Rad52 enrichment in *dcr1*Δ cells peaked over programmed replication pause sites (Sanchez, 1998) and replication origins within

rDNA repeats (Figure 3.8b). We also saw overlapping peaks of Crb2 enrichment in *dcr1*Δ vs WT, indicating fork collapse at these loci (Figure 3.8b).

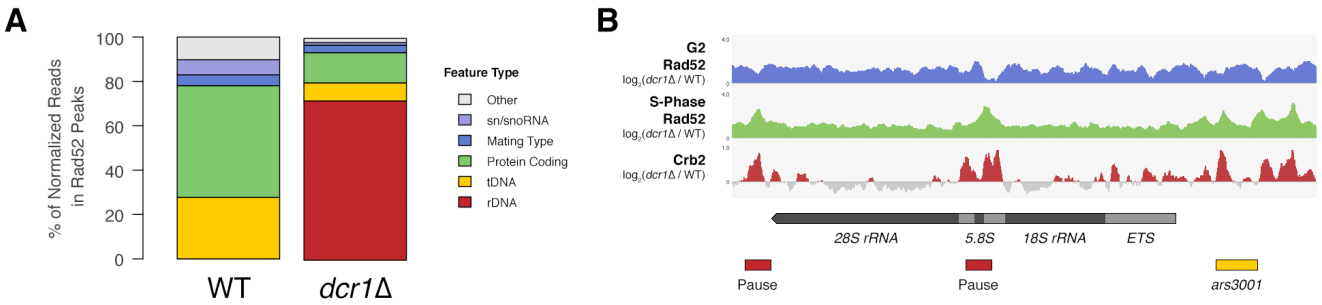

Figure 3.8 | DNA damage accumulates within rDNA repeats in *dcr1*Δ. A) Distribution of Rad52 by feature type in WT and *dcr1*Δ determined by normalized Rad52 ChIP-seq read counts in peaks. B) Rad52 enrichment at rDNA in both unsynchronized (G2) and S-phase cells determined by ChIP-seq in *dcr1*Δ vs WT. Crb2 enrichment at rDNA in unsynchronized cells determined by ChIP-seq in *dcr1*Δ vs WT. Enrichment is shown as log2(*dcr1*Δ / WT). rDNA annotations (black), programmed pause sites (red), and replication origin (yellow) are indicated.

RNA:DNA hybrids can occur at stall sites resulting from transcription and replication collision (Aguilera and García-Muse, 2012; Alzu et al., 2012; Bermejo et al., 2012). These hybrids are themselves highly recombinogenic, and recruit Rad52 (Wahba et al., 2013). We hypothesized that hybrids might form within rDNA repeats due to failed Pol II release in *dcr1*Δ. We performed DNA:RNA Immunoprecipitation (DRIP) (Ginno et al., 2012) to assess hybrid formation at rDNA and found increased hybrid levels in *dcr1*Δ (Figure 3.9a) that was statistically significant across replicates in the 18S and 28S regions (Figure 3.9b). Conversely, hybrids that are Dicer-dependent have been reported at centromeric repeats (Nakama et al., 2012) and our DRIP-seq results support this (Figure 3.9c). We did however observe an increase in hybrids with regions containing replication origins, again suggesting hybrid formation as a consequence of replication and transcription collision.

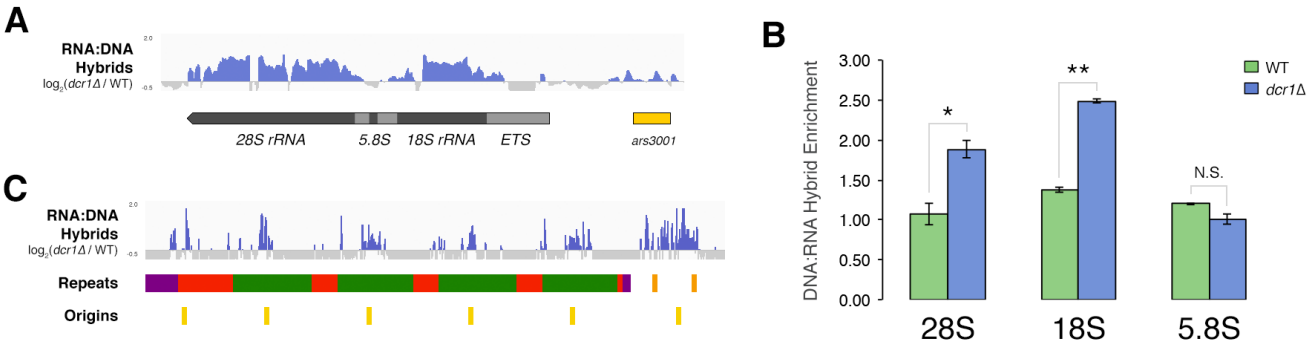

Figure 3.9 | RNA:DNA Hybrids form at sites of transcription-replication collision in *dcr1*Δ. A) RNA:DNA hybrid enrichment detected by DRIP-Seq at rDNA, shown as the $\log_2$(ratio) between *dcr1*Δ and WT. rRNA genes (black) and replication origin (yellow) are indicated. B) Quantification of RNA:DNA hybrids at rDNA in WT (green) and *dcr1*Δ (blue) across replicates and at different regions. Data are represented as mean ± SEM. The significance of differences is indicated (** = $p < 0.01$, * = $p < 0.05$). C) RNA:DNA hybrid enrichment within pericentromeric repeats, shown as the $\log_2$(ratio) between *dcr1*Δ and WT. Centromeric annotations are indicated, *dh* (red), *dg* (green), *imr* (purple), tRNA genes (orange), and replication origins (yellow).

3.3.6 Dicer is required for copy number maintenance of rDNA repeats

The dramatic increase in Rad52 enrichment, and the presence of RNA:DNA hybrids at rDNA in *dcr1*Δ suggested that recombination within the repeats could result in genomic instability. We isolated whole chromosomes from individual WT and *dcr1*Δ colonies of varying growth rates using pulsed-field gel electrophoresis (PFGE). Consistently *dcr1*Δ cells showed a significant reduction in chromosome III size suggesting a loss of subtelomeric rDNA repeats (Figure 3.10a). To understand the dynamics of rDNA loss we created de novo *dcr1* deletion strains from WT cells and assessed rDNA copy number using quantitative polymerase chain reaction (qPCR). Freshly transformed *dcr1*Δ cells showed a 30% reduction in rDNA copy number, which remained stable through 72h of continuous mitotic division (Figure 3.10b). As expected this loss was not seen in *ago1*Δ cells, again indicating a Dicer-specific function. We then tested rDNA stability through meiosis, by assessing copy number through four generations of progeny. Strikingly, rDNA repeats were progressively lost at each meiosis in *dcr1*Δ by approximately 7%, while being maintained in WT and in *ago1*Δ cells (Figure 3.10c).

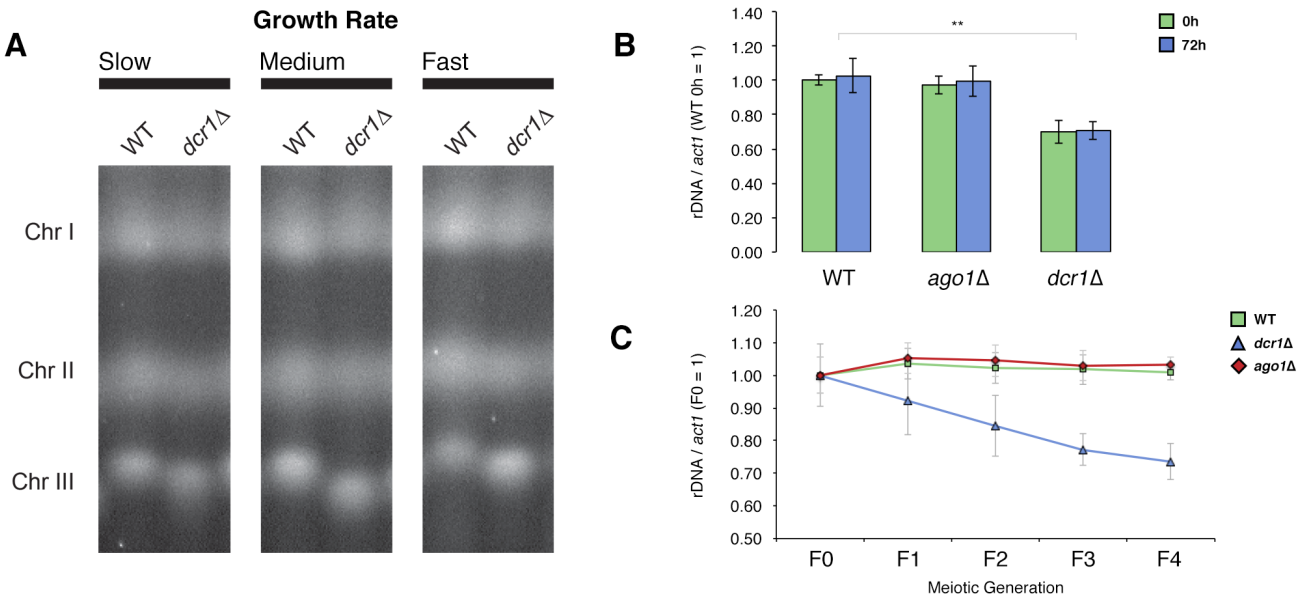

Figure 3.10 | Dicer is required for copy number maintenance of rDNA repeats. A) Whole chromosomes isolated by Contour Clamped Homogenous Electric Field Pulsed Field Gel Electrophoresis (CHEF-PFGE) from individual WT (WT) and *dcr1*Δ colonies of varying growth rates (slow, medium, fast) run side-by-side for comparison. B) rDNA copy number determined by qPCR of genomic DNA from 6 colonies of WT and freshly transformed *dcr1* or *ago1* knockout cells, and cells after 72h (~25-30 generations) of mitotic division. C) rDNA copy number of WT and freshly transformed *dcr1* and *ago1* knockout cells (F0) and subsequent meiotic generations (F*n*). Copy number is normalized to F0 of each strain. Data are represented as mean ± SD. The significance of differences is indicated (** = $p < 0.01$, * = $p < 0.05$). See also Figure S3.

3.3.7 Dicer is required in the face of replicative stress at rDNA

Programmed replication fork pausing facilitates the directional replication of DNA, preventing the collision of replication forks with transcription complexes. The histone demethylase Lsd1 is required for replication fork pausing within rDNA (Holmes et al., 2012), and is enriched at tDNA (Lan et al., 2007) where it may also play the same role, as H3K9me2 spreads across tRNA boundaries in *lsd1* single mutants (Lan et al., 2007) and depends on association of CLRC with the replisome (Li et al.,

2011; Zaratiegui et al., 2011). We hypothesized that in the absence of programmed fork pausing, collisions between Pol II and replication forks would increase, and that Dicer would be required to resolve these. *lsd1* single mutants are slow growing but viable, however we found that *dcr1* and *lsd1* are synthetically lethal, supporting our hypothesis (Figure 3.11a). Increased activity of Dicer in the face of replicative stress at rDNA should result in higher sRNA levels. The helicase Pfh1 is required for replication fork progression through rDNA, and in its absence stalling occurs (Sabouri et al., 2012). We sequenced sRNA from temperature sensitive *pfh1-R23* cells (Tanaka et al., 2002) at a semi-permissive temperature, to induce replication stress in rDNA without arresting growth. As predicted, there was a Dicer-dependent increase in antisense sRNA originating from rDNA in *pfh1* cells relative to WT, supporting increased Dicer activity in the face of replicative stress (Figure 3.11b). Importantly this increase was not due to rDNA repeat expansion, since rDNA copy number in *pfh1-R23* was not significantly different from WT (Figure 3.11c). Intriguingly, the *pfh1-R23 dcr1*Δ double mutant showed a further reduction in rDNA copy number as compared to both parents ($p < 0.05$) suggesting that increased replication stress in the absence of Dicer results in an enhancement of rDNA loss.

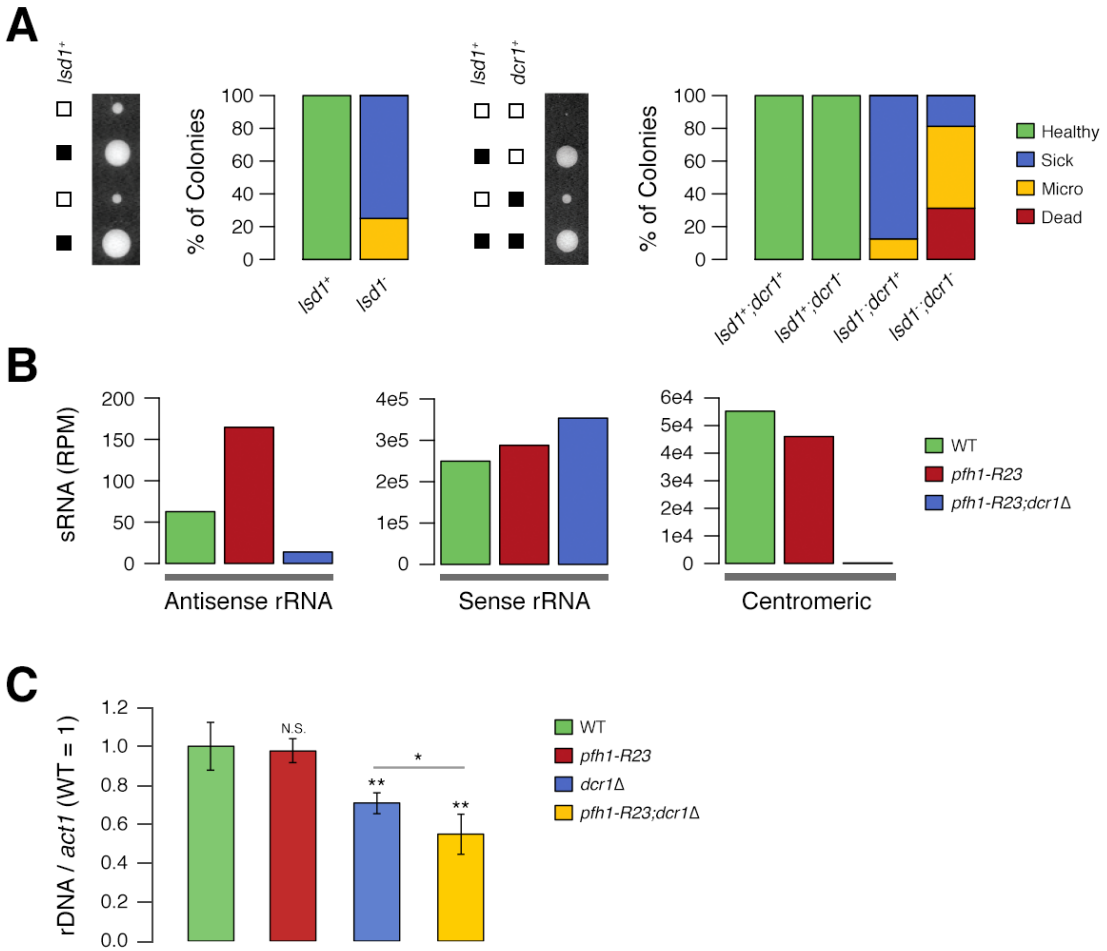

Figure 3.11 | Dicer release is essential in the face of replication stress. A) Representative colonies from tetrad dissection of *lsd1* (N=6 tetrads) and *lsd1;dcr1* heterozygous diploids (N=12 tetrads) with genotypes indicated, and breakdown of colonies by phenotypic class (healthy, sick, micro, dead). B) sRNA reads (RPM) mapping to antisense rRNA, sense rRNA, and centromeric repeats from WT (green), partial loss of function allele *pfh1-R23* (red), and *pfh1-R23;dcr1*Δ double mutant cells (purple) grown at a semi-permissive temperature (30°C). C) Quantification of rDNA copy numbers was performed using real-time qPCR. Four colonies each of WT (WT, green), *pfh1-R23* (red), *dcr1*Δ (blue), and *pfh1-R23;dcr1*Δ (yellow) were used. Data are represented as mean ± SD. rDNA signal was normalized to a single copy gene (*act1*) and divided by the average WT enrichment. The significance of groups compared to WT is shown directly above bars, while other comparisons are indicated (** = $p < 0.01$, * = $p < 0.05$).

3.4 Discussion

3.4.1 A Dicer-specific role in transcriptional termination

Previous genome-wide studies aimed at identifying targets of RNAi in *S. pombe* have focused on RNA transcript levels and histone modification, but have failed to identify a consensus group of targets outside of heterochromatin (Gullerova et al., 2011; Hansen et al., 2005; Woolcock et al., 2012; 2011; Yamanaka et al., 2012). Because of the well-established role of RNAi in transcriptional silencing we interrogated Pol II directly at the chromatin level using ChIP-seq in WT and *dcr1*Δ cells. With this robust approach we identified a comprehensive set of loci that showed a statistically significant increase of Pol II in *dcr1*Δ, suggesting transcriptional regulation by Dicer. These diverse loci included highly transcribed protein coding genes, tDNA, and rDNA in addition to pericentromeric repeats.

At these loci Pol II accumulation was most striking at the 3' end of the transcription unit, suggesting a termination defect in *dcr1*Δ cells. We present several lines of evidence indicating that Dicer promotes transcriptional termination. Canonical termination involves two steps, the first being Pol II pausing, and the second being Pol II release (Park et al., 2004; Yang and Roberts, 1989). We saw no evidence of run-on transcription, which is indicative of a pausing defect, by either Pol II ChIP-seq or RNA-seq. Pol II peaked just upstream of the transcription stop site, and RNA-seq showed reduced transcript levels of Dicer-terminated genes in *dcr1*Δ cells, consistent with a release defect. In some instances we observed an increase in Pol II occupancy that extended upstream of the 3' end. A failure to remove stalled Pol II at the 3' end has been shown to result in an upstream Pol II "pile-up" (Hanawalt and Spivak, 2008; Trautinger et al., 2005), and could explain this observation.

We found strong evidence for direct Dicer activity at regions of Pol II accumulation in the form of Dicer-dependent sRNA that matched the expected size distribution. Because of their putative role in transcriptional termination we termed them termination-associated small RNA (tsRNA). As an RNase III enzyme Dicer's substrate, an RNA duplex, can be generated in two ways. First, the presence of antisense tsRNA from tDNA and rDNA, and their persistence in *rdp1*Δ cells, strongly suggests that antisense transcription by Pol II is occurring at these loci, providing the potential for dsRNA.

Alternatively, secondary structures in the Pol II transcribed RNA molecules themselves might produce hairpins. This latter pathway is more likely at protein coding genes since tsRNA arise almost exclusively from the sense strand at the 3' end, where Pol II accumulated consistent with a failure to release. Predicting secondary RNA structure at an example Dicer-terminated gene revealed a hairpin corresponding to sRNA peaks, consistent with this model.

Our findings suggest that tsRNA do not function through Ago1. Importantly, we did not see Pol II accumulation in *ago1*Δ cells at tDNA or rDNA, and they did not show a reduction in rDNA copy number. Ago1 has strong binding specificity for sRNA with a 5' uridine (Bühler et al., 2008) that tsRNA do not exhibit, while *in vitro* studies have demonstrated that Dicer has no 5' preference in production of sRNA (Weinberg et al., 2011). The TRAMP complex containing the poly(A) polymerase Cid14 targets aberrant small RNA molecules arising from tDNA and rDNA, and prevents their loading into Ago1 (Bühler et al., 2008). In the absence of Cid14, Ago1 binds antisense sRNA arising from rDNA, and this association is dependent on both Dcr1 and Rdp1. Furthermore, overexpression of Dicer results in an increase of sRNA mapping to tDNA and rDNA (Yu et al., 2013). Taken in light of our results these sRNA are likely products of Dicer's role in transcriptional termination. Since Ago1 is not required for this process, the TRAMP complex ensures that these tsRNA do not enter the RNAi pathway. Since tsRNA are not loaded into Ago1 they would be unstable, which explains their relatively low abundance compared to siRNA. Whether these tsRNA themselves play a direct role in termination, or are simply a by-product of Dicer activity remains unanswered.

Transcriptional termination mediated by a 3' hairpin and cleavage by RNAse III (Rnt1p) is conserved in *S. cerivisiae* (Ghazal et al., 2009). Cleavage by Rnt1p results in both transcript degradation at the 3' end by the 3'-5' exoribonuclease Rrp6, and transcriptional termination at the 5' end by the 5'-3' exoribonuclease Rat1p through the "torpedo" model (Figure 3.12b). A similar model of transcriptional termination has been put forward in human cells, involving the RNAse III Drosha and Rrp6 (Wagschal et al., 2012). Much like the tsRNA we identified, sRNA termed "termination-associated sRNA" (TASRs) are generated at the 3' end of protein coding genes in human cells, and could potentially be related to termination (Kapranov et al., 2007).

If this were a conserved function of RNAse III enzymes it would explain Dicer's role in transcriptional termination in *S. pombe* (Figure S6). The only other non-ribosomal RNase III enzyme in *S. pombe*, Pac1, has to date been implicated in post-transcriptional regulation of meiotic genes (Iino et al., 1991) and processing of snRNA and rRNA (Rotondo and Frendewey, 2001). Supporting this model, the nuclear exosome containing Rrp6 in *S. pombe* processes pericentromeric transcripts that are also cleaved by Dicer (Reyes-Turcu et al., 2011). Pol II termination at Dicer-cleaved transcripts in *S. pombe* could be carried out by the Rat1p homolog Dhp1. Interestingly a *ts* allele of Dhp1 shows chromosome segregation defects (Shobuike et al., 2001) similar to those seen in *dcr1*Δ mutants (Volpe et al., 2003).

3.4.2 Dicer acts in the unique context of transcription and replication collision

Why should Dicer promote transcriptional termination at some loci and not others? Our results suggest that in *S. pombe* this regulation occurs specifically at sites where collision between transcription and replication occurs. Head on collisions first occur at the 3' end of transcribed regions and result in stalled Pol II, correlating to sites of Dicer activity. A similar function has been proposed for the RNA/DNA helicase Sen1, which terminates transcription of non-polyadenylated transcripts, and has other functions in replication fork progression (Bermejo et al., 2012; Mischo et al., 2011). Such collision also results in stalled replication forks (Azvolinsky et al., 2009) and the recruitment of Rad52 (Lambert et al., 2010). If stalled forks are not resolved they will collapse, leading to γH2A deposition and Crb2 recruitment (Rozenzhak et al., 2010). We performed both Rad52 and Crb2 ChIP-seq and found a strong correlation between Dicer-terminated loci and peaks of both proteins in WT cells. This suggests that these loci are "natural" sites of replication stress and pausing. Indeed highly transcribed RNA Pol II genes, tDNA, and rDNA all constitute "difficult-to-replicate" regions in *S. pombe* (Sabouri et al., 2012). The presence of Rad52 in WT cells at Dicer-terminated loci suggests that collision, fork stalling, and Rad52 localization occur upstream of Dicer termination.

These findings suggest a model whereby Dicer terminates transcription by releasing stalled Pol II specifically at transcription-replication collisions (Figure 3.12a). This explains why run-on transcription is not observed at Dicer targets, as collision with replisome would presumably prevent further transcription. In this model Dicer does not prevent collisions from occurring in the first place, but

through termination it does resolve them. Without Dicer termination, Rad52 and Crb2 persist at stall sites, which ultimately must be restarted by homologous recombination for replication to proceed (Lambert et al., 2010). It's possible that Dicer is specifically recruited to stalled forks through a pathway not yet understood. Supporting this model, in *Neurospora* Rad52 is required for the generation of aberrant RNA (aRNA) from rDNA repeats by HU-induced replication fork stalling, which are then processed by Dicer into qiRNA (QDE-2 interacting) (Zhang et al., 2013). Of all the loci we detected, only the pericentromeric repeats were enriched for H3K9me2 in a Dicer-dependent manner, consistent with the idea that CLRC is recruited to heterochromatin, but not to euchromatin, for spreading via the replisome (Chen et al., 2008; Li et al., 2011; Zaratiegui et al., 2011).

3.4.3 Dicer is required for genome stability at rDNA

The subtelomeric rDNA repeats are a suitable locus to study the necessity of Dicer termination at collision sites because of their well-known replication dynamics, and tolerance of copy number change. In Dicer's absence we observed an increase in Pol II accumulation and a dramatic increase of Rad52 within rDNA repeats. This was accompanied by a reduction in rDNA copy number likely occurring through homologous recombination. After an initial loss, copy number subsequently remained stable for 72h of mitotic division, however further loss occurred in subsequent meiotic generations. Recombination pathways are hyper activated as part of the normal meiotic progression, and it is possible that without Dicer this leads to an enhancement of rDNA loss. Similarly, RNAi prevents detrimental recombination at the centromeres during meiosis (Ellermeier et al., 2010).

The direction of DNA replication within rDNA repeats is tightly controlled to prevent collisions with transcribing Pol I that would result in stalled forks. However, the presence of antisense tsRNA at rDNA and patterns of poised and elongating Pol II enrichment suggest that Pol II transcription occurs antisense to Pol I and would therefore collide with replication. Our results show that Dicer-termination is required at these collision sites to prevent recombination and thus maintain genomic stability. We demonstrated this by increasing replication stress within rDNA using a partial loss-of-function *pfh1* allele, and both observing an increase in Dicer-dependent sRNA in the *pfh1-R23* single mutant, and an enhancement of rDNA loss in the double *pfh1-R23 dcr1*Δ mutant. Furthermore, in the absence of

programmed fork pausing, Pol II release by Dicer is essential, revealed by the synthetic lethality of *lsd1* and *dcr1*. We also detected increased RNA:DNA hybrids in *dcr1*Δ at rDNA that form when ssDNA is exposed as a result of fork stalling outside of programed pause sites, providing further evidence of collision. Similar hybrids occur within *S. cerivisiae* rDNA in the absence of a master repressor of transcription, *sin3,* which leads to Rad52 recruitment and genomic instability (Gottlieb and Esposito, 1989; Wahba et al., 2011).

The presence of antisense sRNA and enrichment of Pol II suggest that antisense transcription of rDNA occurs even in WT cells. In *S. cerevisiae* Pol II transcription of the intergenic spacer region stimulates recombination and copy number change, thought to be mediated by loss of cohesin localization (Kobayashi and Ganley, 2005). Similar to what we describe, within rDNA Pol II is released by the exosome in budding yeast (Vasiljeva et al., 2008). Pol II transcription is negatively regulated by the silencing protein Sir2 (Smith and Boeke, 1997), and a balance between transcription and silencing therefore regulates copy number. A similar mechanism may exist in *S. pombe*, whereby some level of Pol II transcription is required to promote basal recombination that maintains copy number. This would presumably lead to transcription-replication collisions that are resolved by Dicer. Similarly to Sir2, Dicer is required for cohesin localization at some loci (Gullerova and Proudfoot, 2008), and may also suppress recombination in rDNA by this mechanism.

Dicer regulation of rDNA appears to be conserved across eukaryotes with RNAi. *Drosophila* DCR-2 is required to maintain K9 methylation at rDNA repeats, and thus their stability (Peng and Karpen, 2007). In *Neurospora* Dicer produces sRNA from rDNA repeats and is similarly required for their stability (Cecere and Cogoni, 2009). Dicer physically localizes to rDNA repeats in mouse ES cells (Sinkkonen *et al*., 2010). Thus studies across eukaryotes support a conserved role for Dicer at rDNA, and our research suggests a specific effect tied to Pol II regulation.

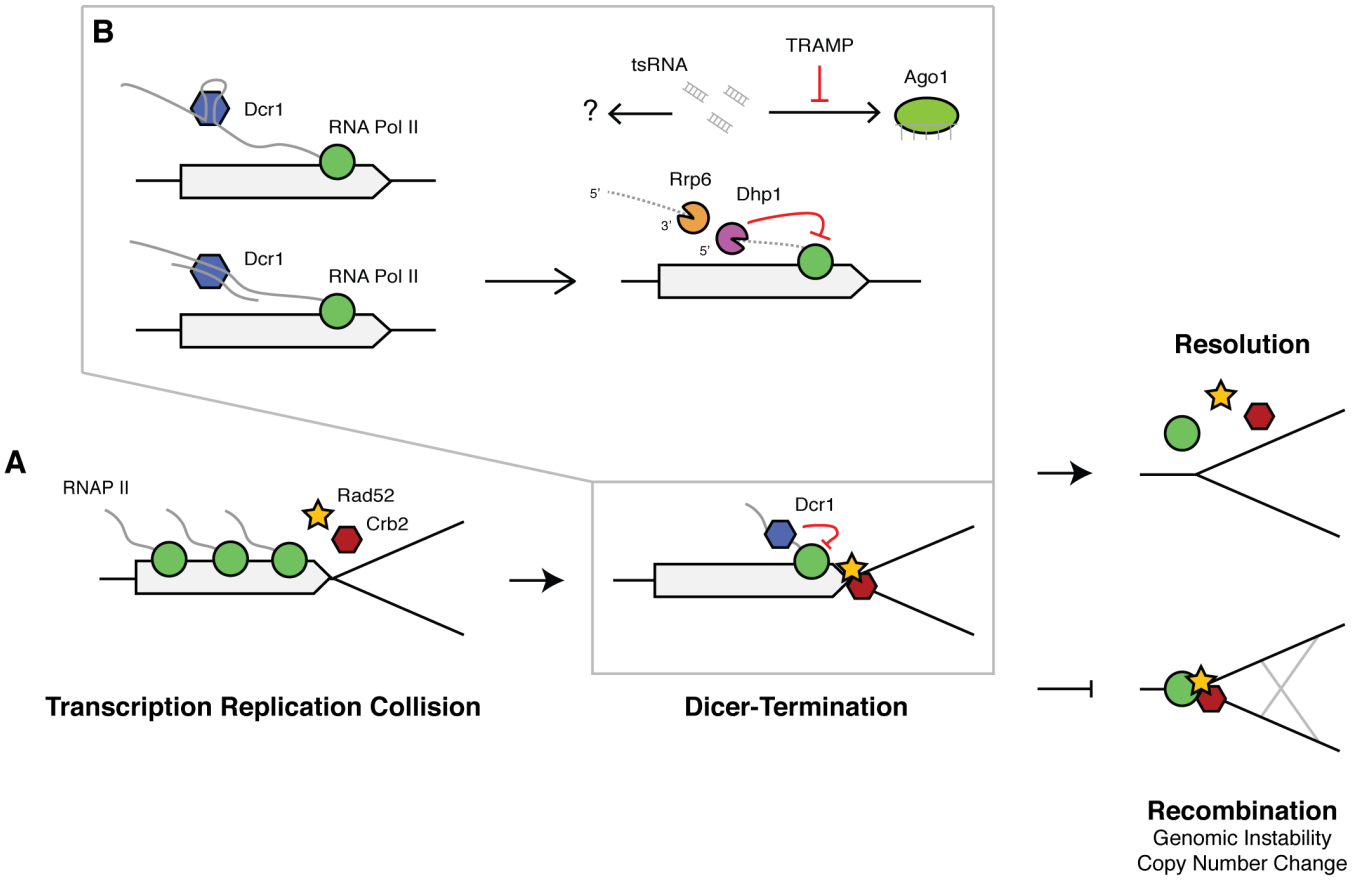

Figure 3.12 | Dicer-termination of Pol II transcription at stalled replication forks maintains genomic stability. A) Transcription by RNA Pol II (green circle) and DNA replication collide producing stalled replication forks that recruit Rad52 (yellow star) and Crb2 (red hexagon). Dicer (blue hexagon) acts at these sites to release Pol II and facilitate replication. Without Dicer homologous recombination is necessary to restart the replication fork and results in genomic instability and copy number changes. See Figure S5 for mechanism of Pol II release by Dicer. B) Dicer cleaves structured RNA (protein coding genes) or double stranded RNA (tDNA, rDNA) producing termination-associated sRNA (tsRNA) and creating free ends for the action of the exonucleases Rrp6 (orange) and Dhp1 (pink). Dhp1 releases Pol II through 5' to 3' exonuclease activity and the torpedo model. tsRNA are prevented from entering the canonical sRNA pathway by the TRAMP complex.

3.5 Contributions

Some of this work (Chapter 3) was done in collaboration with both lab members and other research groups. The following experiments were performed as a collaboration, with my contributions listed:

- H3K9me2 ChIP-seq – collaboration with An-Yun Chang. An-Yun performed the ChIP experiment, I analyzed the ChIP-seq data, and combined it with ChIP-seq data I had previously done myself.
- RNA-seq – collaboration with An-Yun Chang. An-Yun performed the experiment, and I analyzed the data.
- Rad52 ChIP-seq – collaboration with Jie Ren. Assisted her with ChIP experiment itself. I created libraries, sequenced all samples, and did all analysis.
- 2D gels – performed by Paco Antequera's lab. I designed the experiment, but did not perform the 2D gels themselves as we lack the equipment and expertise to do them at CSHL.
- rDNA copy number through meiosis – collaboration with Sonali Bhattacharjee. I designed the experiment, optimized rDNA quantification protocols, and worked closely with Sonali to complete it.

3.6 Supplementary Tables

Table 3.1. Protein coding genes with significant RNA Pol II enrichment in *dcr1*Δ vs wild type. List of all protein coding genes showing significant increase in Pol II enrichment in *dcr1*Δ vs WT calculated using MEDIPS and a cutoff of FDR < 0.01.

SPBC609.04	SPAC19G12.10c	SPBC11C11.09c	SPAC26F1.10c	SPBC1685.10
SPBC3E7.02c	SPBP8B7.06	SPBC2G2.05	SPAC1805.13	SPAC13C5.05c
SPCC1739.13	SPCC576.08c	SPBC1709.05	SPAPB17E12.05	SPAC22A12.04c
SPAC13G7.02c	SPCC645.14c	SPBC800.04c	SPAC15E1.03	SPBC56F2.08c
SPAP8A3.04c	SPAC3G9.03	SPAC3H5.10	SPCC757.13	SPBC1685.02c
SPCC663.06c	SPAC13D6.02c	SPAC821.09	SPAC3A12.10	SPAC25H1.08c
SPAC9E9.09c	SPBC27.08c	SPAPJ760.03c	SPAC24H6.07	SPBC28F2.03
SPCC330.06c	SPBC1711.14	SPCC1322.11	SPCC330.14c	SPAC23C11.02c

SPAC926.04c	SPAC3H5.04	SPBC29A3.12	SPBC1711.08	SPCP1E11.09c
SPBC16D10.08c	SPAP7G5.05	SPAPYUG7.03c	SPBC23G7.15c	SPCC962.04
SPAC1B3.03c	SPBP8B7.05c	SPBC21C3.13	SPAC3H5.07	SPBC776.11
SPBC21C3.19	SPAC9E9.13	SPCC1020.01c	SPAC1805.12c	SPBC1861.01c
SPCC757.07c	SPAC9G1.03c	SPAC1002.19	SPBC577.02	SPBC1778.01c
SPBC839.06	SPAPB17E12.13	SPCC1223.05c	SPAC6F6.07c	SPAC11D3.05
SPAPB8E5.06c	SPAC664.04c	SPBC30B4.09	SPBP8B7.15c	SPBC3E7.12c
SPCC663.08c	SPBC16G5.14c	SPAC110.04c	SPAC29A4.02c	SPBC16H5.08c
SPAC17A5.03	SPAC3C7.14c	SPBC1711.15c	SPAC1783.08c	SPBC17G9.10
SPBC215.05	SPCC576.09	SPCC74.05	SPBC30B4.04c	SPBC19F8.08
SPAC3H5.05c	SPAC977.12	SPAC806.03c	SPAC11G7.04	SPBC11C11.07
SPAC664.05	SPBC106.17c	SPAC13G6.07c	SPAC4F10.14c	SPCC970.05
SPAC1F7.13c	SPAC6G10.11c	SPAC1071.08	SPBC1271.10c	SPBC365.03c
SPBC106.02c	SPCC1682.14	SPAC6B12.15	SPBC685.07c	SPAC30D11.12
SPAC9E9.11	SPAC6G9.09c	SPAC11E3.15	SPAC750.01	SPCC1223.08c
SPAC2H10.01	SPAC222.09	SPBC16D10.11c	SPBC3B9.12	SPAC11D3.14c
SPAC23A1.10	SPAC31G5.03	SPBC713.11c	SPAC1B3.04c	SPAC869.11
SPAC343.12	SPAC17G6.06	SPAC24C9.12c	SPBC839.04	SPCC550.05
SPBC839.15c	SPAC19G12.16c	SPAC144.11	SPBC83.17	SPAC26F1.04c
SPAC343.20	SPBC15D4.05	SPBC17G9.07	SPAC26F1.14c	SPBC3B9.01
SPBC26H8.06	SPCC1259.01c	SPAC513.07	SPAC521.05	SPAC23C11.03
SPAC3H5.12c	SPAC644.15	SPBC3D6.15	SPBC4F6.09	SPBC1709.06
SPCC16C4.13c	SPBPB8B6.05c	SPAC19G12.09	SPAC22A12.15c	SPCC613.06
SPCC613.05c	SPAC328.03	SPAC1805.11c	SPAC4F10.20	SPAC15A10.04c
SPAC8C9.08	SPAC9E9.12c	SPBC12C2.04	SPAPB1E7.04c	SPBC16G5.13
SPBC887.17	SPAC18G6.14c	SPCC5E4.07	SPAC22G7.06c	SPBC27B12.03c
SPBC685.06	SPAC26A3.04	SPAC12G12.04	SPAC5D6.01	SPBC1685.09
SPBC29A3.04	SPAC1071.07c	SPAC694.05c	SPAC29E6.08	SPCC569.05c
SPAC11D3.13	SPAC1687.06c	SPBC2G5.06c	SPAC5D6.02c	SPAC4D7.09
SPBP8B7.03c	SPBC3B9.13c	SPAC1B3.05	SPBC18E5.06	SPAC16C9.03
SPAC23A1.11	SPAPB1E7.12	SPCC364.03	SPAC3G6.13c	SPAC17A5.04c
SPBP22H7.08	SPAC9E9.08	SPCC70.12c	SPAC14C4.09	SPAC31A2.12
SPCC622.18	SPCC18.14c	SPAC31G5.17c	SPAC13G6.02c	SPBC839.05c
SPAC22H12.04c	SPAC959.07	SPAC1805.10	SPBC2G2.06c	SPAC1F3.09
SPAC140.02	SPBC18H10.13	SPACUNK4.16c	SPCP31B10.08c	SPBC106.18

SPCC1183.08c	SPCC830.07c	SPAC328.10c	SPBC29B5.01	
SPBC1734.11	SPAC890.08	SPCC550.06c	SPAC227.13c	
SPCC736.15	SPAC2C4.16c	SPBC4F6.17c	SPBC83.02c	
SPAC17G6.13	SPBC21B10.10	SPAC1F12.02c	SPAC16C9.02c	
SPAPJ698.02c	SPAC589.10c	SPAC959.08	SPAC343.21	

Table 3.2. GO term enrichment analysis of protein coding genes with increased Pol II in *dcr1*Δ,

GO ID	Gene Ontology Term	Cluster Frequency	Genome Frequency	Corrected P-Value
GO:0002181	Cytoplasmic translation	50.40%	8.60%	5.37E-65
GO:0006412	Translation	53.00%	11.80%	1.38E-54
GO:0019538	Protein metabolic process	62.70%	28.80%	7.28E-26
GO:0044267	Cellular protein metabolic process	61.40%	28.00%	1.06E-25
GO:0034645	Cellular macromolecule biosynthetic process	58.10%	26.10%	6.19E-24
GO:0044249	Cellular biosynthetic process	64.80%	34.00%	1.08E-20
GO:0010467	Gene expression	61.90%	29.90%	3.20E-20
GO:1901576	Organic substance biosynthetic process	64.80%	34.30%	3.22E-20
GO:0009058	Biosynthetic process	65.30%	34.70%	6.03E-20
GO:0042254	Ribosome biogenesis	22.00%	6.10%	2.39E-14
GO:0022613	Ribonucleoprotein complex biogenesis	22.50%	7.10%	3.79E-12
GO:0043170	Macromolecule metabolic process	72.00%	49.50%	1.56E-10
GO:0044260	Cellular macromolecule metabolic process	70.30%	48.00%	3.38E-10
GO:0044238	Primary metabolic process	79.70%	60.10%	2.98E-08
GO:0044237	Cellular metabolic process	81.40%	62.10%	3.69E-08
GO:0008152	Metabolic process	82.60%	64.10%	1.13E-07

GO:0071704	Organic substance metabolic process	80.50%	62.20%	2.68E-07
GO:0009408	Response to heat	5.90%	0.80%	1.37E-06
GO:0034605	Cellular response to heat	5.90%	0.80%	1.37E-06
GO:0009266	Response to temperature stimulus	5.90%	0.80%	1.91E-06
GO:0006457	Protein folding	8.50%	2.10%	3.19E-05
GO:0009987	Cellular process	90.80%	78.40%	1.10E-04
GO:0042026	Protein refolding	2.10%	0.10%	5.20E-04
GO:0061077	Chaperone-mediated protein folding	2.50%	0.20%	1.49E-03
GO:0044085	Cellular component biogenesis	25.80%	14.90%	3.07E-03

Table 3.3. Nuclear tDNA with significant RNA Pol II enrichment in *dcr1*Δ vs wild type. List of all nuclear tRNA genes showing significant increase in Pol II enrichment *dcr1*Δ vs WT across replicates with a cutoff of $P < 0.05$.

SPCTRNAGLU.10	SPBTRNAPRO.08	SPCTRNALEU.12	SPBTRNAGLN.01	SPCTRNATHR.08
SPBTRNAMET.06	SPCTRNALYS.11	SPCTRNAASN.06	SPATRNAALA.04	SPATRNATYR.01
SPCTRNAMET.07	SPBTRNAMET.04	SPBTRNAGLY.09	SPBTRNAVAL.06	SPATRNAARG.03
SPCTRNAGLY.12	SPCTRNAGLN.06	SPATRNASER.04	SPBTRNAVAL.07	SPCTRNAASP.08
SPBTRNAILE.08	SPCTRNAARG.09	SPATRNAASP.01	SPCTRNAVAL.11	SPATRNAMET.03
SPBTRNAGLU.05	SPBTRNAPRO.06	SPBTRNATYR.04	SPBTRNAGLN.03	
SPCTRNAVAL.10	SPATRNAPRO.01	SPBTRNAASN.03	SPBTRNALYS.07	
SPCTRNASER.07	SPATRNAVAL.01	SPBTRNAHIS.02	SPCTRNAVAL.09	

4. Discussion and Future Direction

4.1 Transcriptional termination by RNAi

The co-transcriptional model of RNAi silencing first put forward in fission yeast implied that transcriptional termination was part of the process. Our results, which separate H3K9 methylation from Pol II release, in combination with many studies in *S. pombe* (Gullerova and Proudfoot, 2008; Woolcock et al., 2012; 2011) and the identification of co-transcriptional silencing in higher eukaryotes indicate that this is a conserved feature of nuclear RNAi. Broadly, transcriptional termination involves two stages, RNA Pol II pausing, and Pol II release. Our work suggests that RNAi is involved in the latter stage of termination. As opposed to constituting a completely novel termination pathway, it appears that nuclear RNAi pathways can feed targets into the exosome and exonuclease-based termination pathways. This could be achieved by producing a free uncapped 5' end on a nascent RNA molecule from transcribing RNA Pol II through either Dicing or Slicing that is a substrate for exonuclease digestion (Figure 4.1). At loci with antisense transcription (centromere, tDNA, rDNA), dsRNA is readily available and is a substrate for Dicing. At protein coding genes, where we observed only sense Dicer-dependent sRNA, it's possible that secondary structures, such as hairpins, could produce substrates for Dicer-cleavage. A meta-analysis of mRNA secondary structure in *Arabidopsis* found that RNA secondary structure peaks in the 3' UTR of protein coding genes, lending support to our hypothesis (Li et al., 2012). It's important to note that while we have put forward a model of exonuclease based termination at RNAi targets because of similarities to other systems, we have not yet directly tested it. Future studies should aim at unequivocally demonstrating release of Pol II by this torpedo model at RNAi targets. Preliminary results suggest this is the case, as a null mutant of a protein in the Dhp1 exonuclease complex (Din1) shows increased centromeric Pol II occupancy (data not shown). This however needs to be validated at other RNAi targets.

Exonuclease mediated termination is conserved from yeast (Ghazal et al., 2009) to humans (Wagschal et al., 2012; West et al., 2004) so it's possible that RNAi evolved alongside RNA surveillance pathways in an early eukaryotic ancestor. Indeed RNAi itself is highly conserved among eukaryotes, and was likely present in the last common ancestor of eukakaryotes (LECA) (Shabalina and Koonin, 2008). Supporting this link is the intimate evolutionary relationship between RNAi, RNA surveillance, and

splicing pathways that is apparent across eukaryotes (Tabach et al., 2013). In fission yeast some splicing factors are required for RNAi directed silencing (Bayne et al., 2008), and RNA surveillance pathways, such as the exosome, share common targets with RNAi, supporting a co-evolutionary relationship. In *mlo3*Δ mutants the exosome can take over silencing from RNAi at the centromere (Reyes-Turcu et al., 2011), and conversely in an *rrp6*Δ background RNAi takes over silencing at a diverse set of genes (Yamanaka et al., 2012). The promiscuous targeting of RNAi in *rrp6*Δ mutants could be explained by the accumulation of transcripts that would otherwise be degraded, and act as a substrate for sRNA generation. This seems to be the case at the centromere, where sRNA levels are highly elevated in *rrp6*Δ mutant cells. Similarly, at the rDNA, antisense sRNA levels increase in an *rrp6*Δ mutant, supporting the involvement of the exosome in transcript degradation (Marasovic et al., 2013). How transcription is terminated could be the key determinant of whether a transcript is silenced by the exosome or RNAi. The efficacy of hairpins to silence in *trans* is greatly increased if the target is missing a canonical 3' processing signal, suggesting that proper termination inhibits RNAi targeting (Yu et al., 2013).

The target overlap between RNA metabolism pathways and RNAi is not limited to fission yeast. Like exosome mutants in *S. pombe*, mutants of the 3' to 5' exonuclease, XRN4 in *Arabidopsis* generate promiscuous sRNA at a number of protein coding genes (Gregory et al., 2008). Furthermore, this exonuclease degrades transcripts that have been targeted by miRNA, lending support to the hypothesis that RNAi can feed transcripts into RNA metabolism pathways (Souret et al., 2004). The overlap between RNA metabolism and silencing has also been demonstrated by studies of proteins involved in regulating flowering time in *Arabidopsis*. FCA and FCP are required both for small RNA mediated transgene silencing (Bäurle et al., 2007), and 3' RNA processing and transcription termination (Sonmez et al., 2011).

At the novel Dicer-terminated loci that we identified in *S. pombe* there was either a lack of either H3K9 methylation or no change in methylation, meaning that not all targets of Dicer are targets for the full RNAi pathway. The TRAMP complex was previously thought to target and degrade aberrant small RNA by feeding them into the exosome, and thus prevent their entry into the RNAi pathway (Bühler et al., 2008). We suggest that the TRAMP complex acts as a gatekeeper to the chromatin based silencing aspect of the RNAi pathway. Supporting this, in the absence of the TRAMP component Cid14, levels

of Ago1 associated sRNA arising from Dicer-terminated loci increase. This could mean that the TRAMP complex is able to survey sRNA pools and selectively degrade some Dicer products. The increase in sRNA levels in exosome deficient cells also supports this model, as sRNA targeted by TRAMP can no longer be degraded. Without the exosome, Dicer-dependent sRNA generated from rDNA are able to direct H3K9me2, unlike in WT cells (Marasovic et al., 2013). How some sRNA are targeted for degradation while others are not is unknown. It's possible that 2' methylation has an effect on sRNA stability, analogous to other systems, since *S. pombe* has a homolog of the Hen1 methyltransferase whose function has not yet been described. It might also be that sRNA produced by RdRP amplification (such as centromeric siRNA) are protected from TRAMP degradation. In contrast, sRNA may be degraded non-discriminately by the exosome, resulting in high levels of turnover, ensuring that only those targets with highest sRNA levels are loaded into Ago1.

Why RNAi acts at some loci to remove Pol II and terminate transcription and not others remains a major outstanding question. Our research suggests that the context of transcription, rather than an intrinsic sequence-dependent signal, is likely the most important aspect of RNAi targeting. Specifically, we find that Dicer promotes termination at sites of stalled Pol II. It could be that aberrant Pol II pausing, accompanied a lack of canonical termination is a signal for Dicer-termination. Studies in *Arabidopsis* have shown that when termination is compromised in RNA processing mutants, RNA-silencing pathways are able to recognize the aberrant 3' end, generate sRNA, and elicit a silencing response (Herr et al., 2006).

It is clear that RNAi is not essential for transcriptional termination in eukaryotes, as some organisms, such as *S. cerivisiae* have lost it completely (Weinberg et al., 2011), and many RNAi null mutants are viable. This suggests that in the absence of RNAi, termination at targets is still able to occur, perhaps through a parallel pathway. At one of the loci we described here, the rDNA repeats, Pol II termination is carried out by the Nrd1/Sen1 exosome complex in *S. cerivisiae* (Vasiljeva et al., 2008), exemplifying the interchangeability of termination pathways. Co-transcriptional regulation by RNAi has been identified in other higher organisms including *C. elegans* (Guang et al., 2010), *Drosophila* (Cernilogar et al., 2011), and *Arabidopsis* (Liu et al., 2012), and while the mechanism behind this regulation is not yet understood, it could be similar to what we describe here. Of the higher eukaryotic systems *C.*

elegans would be an ideal system to verify the conservation of the mechanism behind RNAi based transcriptional termination.

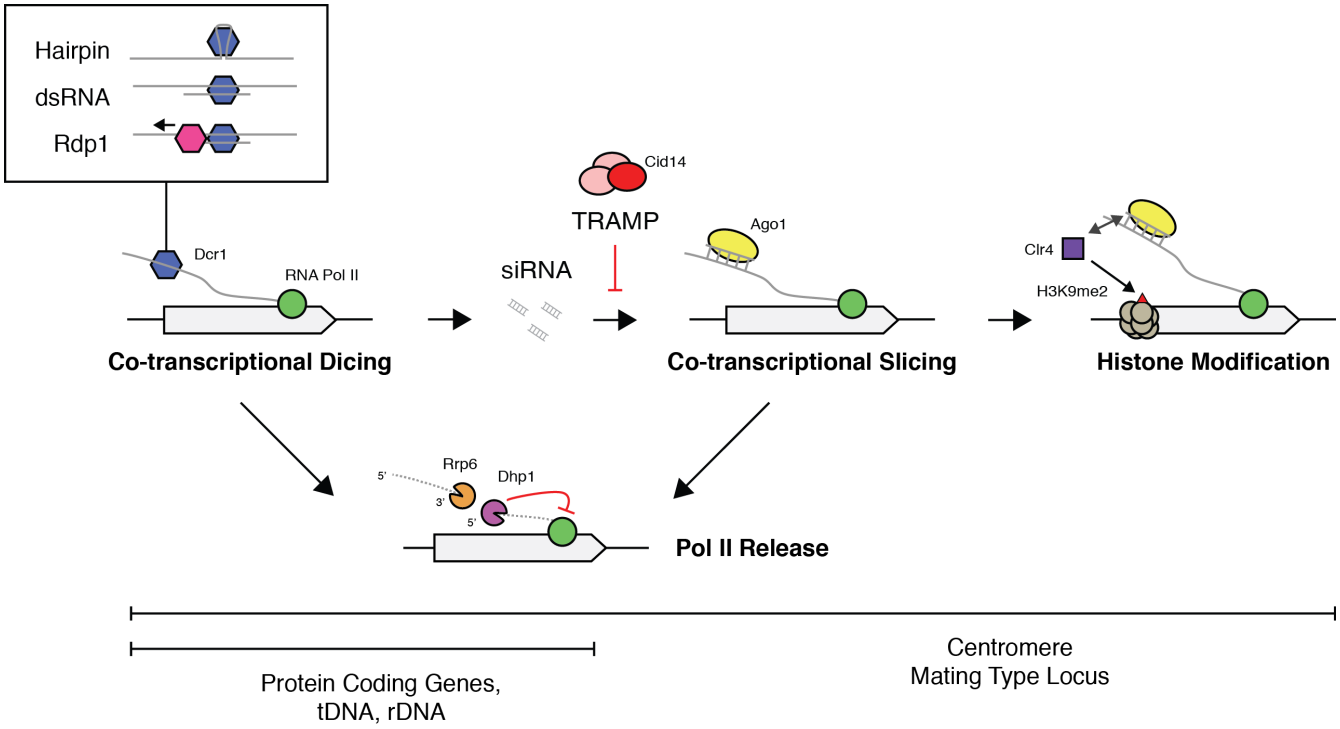

Figure 4.1 | Co-transcriptional silencing and epigenetic modification by RNAi at diverse loci in *S. pombe*. Dcr1 cleaves dsRNA arising from a hairpin, complementary RNA, or through Rdp1 from nascent Pol II transcripts to form 20-23nt siRNA. The TRAMP complex containing the poly(A) polymerase Cid14 acts as a gatekeeper to the RNAi pathway by selectively degrading some sRNA. Ago1 loaded with siRNA targets complementary nascent Pol II transcripts. Cleavage of nascent transcripts by both Dcr1 and Ago1 produces free 3' and 5' ends that are recognized by the exosome (Rrp6) at the free 3' end, and the 5' to 3' exonuclease Dhp1 at the free 5' end. Dhp1 exonuclease activity releases transcribing Pol II by the torpedo model. Ago1 targets are further silenced at the chromatin level through the deposition of H3K9me2 by the CLRC containing the histone methyltransferase Clr4.

4.2 RNAi and DNA replication

The context of both transcriptional termination and epigenetic modification by RNAi is that it occurs during S-phase while DNA replication proceeds. RNAi genes are expressed and proteins are detectable throughout the cell cycle (Gullerova et al., 2011), suggesting that they are not temporally limited to S-phase. This does not rule out the possibility of regulation by post-translational modification of RNAi proteins that could render them inactive outside of S-phase. What is clear however is that the targets of RNAi share in common S-phase transcription, providing a clear context for RNAi action. This could explain why in *S. pombe* constructs designed to silence genes in *trans* through RNAi, with for example hairpins, work at low efficiency (Simmer et al., 2010), since the targets might not be expressed during S-phase. When reporter genes are inserted into pericentromeric heterochromatin they come under the control of promoters that drive repeat expression during S-phase, potentially explaining their robust silencing. However this effect cannot be disentangled from heterochromatin spreading. Future studies should aim at determining if S-phase transcription is a requirement, or enhancer of RNAi based silencing. In the case of epigenetic modification, this requirement seems even more evident, because H3K9 methylation is coupled to DNA replication (Li et al., 2011).

Our research suggests that RNAi acts not only at sites of S-phase transcription, but also more specifically at transcription and replication collisions that produce stalled Pol II. This could be related to the evolution of RNAi as a transposon defense mechanism. Retrotransposons in *S. pombe* control the direction of their replication by directionally pausing replication forks, which may prevent their eviction through recombination pathways (Zaratiegui et al., 2010). While RNAi doesn't normally target transposons in *S. pombe*, this appears to be an exception. If this is a shared characteristic among transposons it's possible that RNAi evolved to target stalled replication forks. Similarly, the targets of transcriptional termination by Dicer that we identified are highly correlated with replication fork pausing and DNA damage, although at these sites transcription is likely the cause of stalling. In the case of repetitive elements (tDNA and rDNA), by resolving stalled forks Dicer prevents copy number changes, again drawing a parallel to the control of transposons.

Stalled Pol II itself could be the signal that recruits Dicer to transcription-replication collisions. Such a mechanism would be analogous to transcription-coupled repair pathways, that use stalled Pol II to detect DNA lesions (Hanawalt and Spivak, 2008). Similar to what we describe in *dcr1*Δ cells, a number of mutants in 3' end cleavage and polyadenylation factors, and RNA processing factors exhibit genomic instability, presumably related to their inability to correctly terminate Pol II transcription in the context of replication (Luna et al., 2005). This work was performed in budding yeast, where RNA processing factors may have taken over Pol II release in the absence of RNAi. For example, in budding yeast the RNA/DNA helicase Sen1 facilitates transcription termination, and by doing so maintains genomic stability, similar to the role we describe for Dicer (Bermejo et al., 2012; Mischo et al., 2011). Also similar to Dicer, Sen1 along with the NRD complex carry out Pol II termination within the rDNA, exemplifying the interchangeability of various termination pathways (Vasiljeva et al., 2008). Resembling RNAi genes in *S. pombe*, *sen1* is synthetic lethal with genes involved in HR, likely because of the dependence on HR to restart forks stalled by Pol II transcription (Mischo et al., 2011). Proper coordination of transcription and replication is a challenge that all organisms face, thus the removal of stalled Pol II is of particular importance. For example, in rapidly dividing human cancer cells, but not resting cells, the major source of γ-H2AX, and thus double strand breaks, is at highly transcribed genes, directly demonstrating the consequence of clashes between transcription and replication (Seo et al., 2012). In bacteria, which are among the most rapidly dividing organisms, the orientation of most genes is arranged to avoid head-on collision with replication, and thus prevent detrimental mutations and genomic instability (Paul et al., 2013). It is therefore unsurprising that cells employ many, often redundant pathways as fail-safes to address this, exemplified by the myriad of ways termination can occur. Understanding how collisions are resolved in model organisms will have implications for higher eukaryotes and may even help in the development of novel cancer treatments.

The more general link between RNAi pathways and processes involved in DNA replication is not limited to *S. pombe*. Recently RNAi has been implicated in DNA repair across higher eukaryotes. In *Arabidopsis* (Wei et al., 2012), *Drosophila* (Michalik et al., 2012) and vertebrates (Francia et al., 2012) a robust RNAi response occurs at double strand breaks (DSB). There are competing hypothesis as to the role of small RNA and RNAi at these breaks. One suggests that they guide chromatin modifications or recruit repair proteins to the break site (Wei et al., 2012). Another hypothesis that is supported by our work is that RNAi is required to prevent Pol II transcription at a break site (Michalik et al., 2012). Double strand breaks will stall any replication complex that encounters them, so it's possible that the

stall, and not the break, constitutes the signal for RNAi response. We found strong correlation between double strand breaks (Crb2) and Dicer-terminated protein coding genes, which supports this hypothesis. In vertebrates Dicer and Drosha are required to activate a repair response, but not other RNAi components, similar to our observations. However, unlike in *Arabidopsis* we did not see a reduction in double strand break repair efficiency in the absence of RNAi (data not shown). It's possible that plants, with their prolific expansion of RNAi pathways have evolved a mechanism for RNAi mediated DNA repair that is not conserved in other lineages.

Our observations suggest that Pol II release, at least in *S. pombe*, is the key role performed by RNAi at stalled replication forks that leads to their resolution, however we have not ruled out other possibilities. It's possible for example that sRNA could recruit repair complexes to the stalled fork. Preliminary studies in our lab have shown a physical interaction between Dicer and Rad51 suggesting that they could form a complex. These questions could be better addressed by moving from genome wide correlational studies, such as those presented here, to a more experimental system, where transcription and replication can be manipulated.

5. Experimental Procedures

5.1 Wet bench methods

Yeast Strains and Growth

Strains used in this work are listed in Table 5.1. Standard media (YEA) and genetic protocols for fission yeast were used. Cells for *pfh1-R23 ts* allele sRNA-seq experiment were grown at the semi-permissive temperature as determined by plate assay (Figure 5.1).

Cell Cycle Synchronization

Cells were arrested with 15mM HU for 4 hr 20 min, washed twice in HU free media, and released. Samples for ChIP and microscopy were taken at regular intervals. Synchrony was then measured using septation index. Samples for Rad52 ChIP-seq were taken from the first S-phase, occurring approximately 90 minutes after release.

ChIP

ChIP was performed from chromatin fixed by 3% PFA for 30 minutes at room temperature. Fixation was stopped using 125 mM Glycine. Cell lysis and chromatin fragmentation were carried out using a Bioruptor on high setting with 8 cycles of 30s pulsing and 1m cooling. Chromatin was quantified using the Bradford assay, and generally 750ug of chromatin was used per IP. The amount of antibody used varied based on the IP being performed. The following antibodies were used: Pol II pS2 - Abcam ab5095, Pol II pS5 - Abcam ab5131, H3K9me2 – Milipore 07-441, Myc - Invitrogen R950, GFP (Rad52) - Abcam ab290, RNA:DNA hybrids - S9.6.

ChIP-qPCR

qPCR was performed using the primers listed in Table 5.2. Average CT was calculated across technical triplicates for each sample. IP enrichment was calculated as % of input (whole cell extract) and

presented relative to WT. Each IP was performed in triplicate. Significance was calculated using a two-tailed heteroscedastic T-test.

DRIP-Seq

Performed as per Ginno *et al*., 2012 modified to include a zymolase digestion, necessary to breakdown the fission yeast cell wall. Input nucleic acid was fragmented with DdeI digestion before the IP was performed. S9.6 antibody was purified from ATCC HB-8730 hybridoma cell supernatant using Pierce Chromatography Cartridges Protein G (89926). Mitochondrial reads were removed during sequence processing. IP enrichment was calculated relative to RNAse H treated IP using the formula: DRIP enrichment = (IP RPM (- RNAse H) / Input RPM (- RNAse H)) / (IP RPM (+ RNAse H) / Input RPM (+ RNAse H)).

Illumina Sequencing

Genomic DNA libraries were created using either the standard Illumina protocol (Pol II pS5, pS2 and H3K9me2), or with the Nugen Ovation Ultralow DR kit (0330, all others). Small RNA libraries were created using the NEBNext Small RNA kit (E7300). Sequencing was performed on Illumina GA II, Illumina HiSeq, or Illumina MiSeq platforms depending on the experiment. A full list of all libraries used in this work is listed in Table 5.3.

Clamped Homogenous Electric Field (CHEF) Pulsed Field Gel Electrophoresis (PFGE)

Mid-log cells were grown in standard media (YEA), harvested, and embedded in 1% low melting temperature agarose. A Biorad CHEF Mapper pulsed-field gel apparatus was used to run 1X TAE, 0.8% PFGE grade agarose at 14°C with the following settings: 48hr two state run with a gradient of 2.0V / cm, angle of 120°, and a period of 1800s. Gels were stained with Ethidium Bromide and visualized under UV light.

5.2 Bioinformatic analysis

Illumina Read Processing and Alignment

Illumina reads were quality filtered using Trimmomatic and aligned to the *S. pombe* genome assembly ASM294v2.21 using Bowtie v2.1.0 and local alignment, with multi-mappers randomly distributed. For genomic DNA libraries all duplicate reads were discarded. Read counts were normalized to reads per million (RPM), using total library size.

ChIP-Seq Analysis

ChIP peaks versus appropriate inputs (whole cell extract) were called using MACS v1.4. When replicates were performed only peaks found in all replicates were considered. MEDIPS v1.12.9 was used to compare Pol II enrichment genome wide between WT and *dcr1*Δ experiments. Differential coverage in MEDIPS was calculated using EdgeR and a cutoff of FDR < 0.01. Genome browser tracks and meta-analysis were created using enrichment (IP reads per million (RPM) / input rpm) of representative replicates. Enrichment at individual loci was calculated as IP RPM / input RPM within the genomic interval and significance was calculated using a two-tailed heteroscedastic T-test.

Metaplot Analysis

Software was developed to produce metaplots based around a defined set of genomic features. Briefly, reads were first converted into 1 BP intervals at either the 5' end, midpoint, or 3' end depending on the type of metaplot to be produced. Next the zero point of each genomic feature to be interrogated was defined as either the 5' end, midpoint, or 3' end. Bins of a specified size were produced around the zero point of each feature. Reads in the samples were then intersected against each of the bins to give a read sum, read standard error, and read average. In the case of ChIP samples enrichment was presented as IP read sum / input (WCE) read sum. If applicable fold change between samples was then calculated at each bin and plotted.

5.3 Figures and tables related to experimental procedures

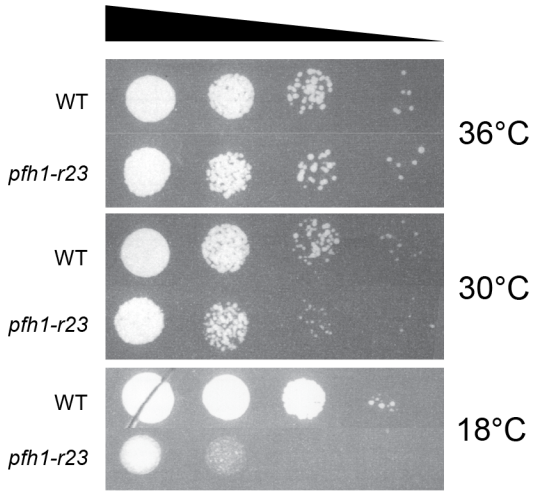

Figure 5.1 | The *pfh1-R23* allele is semi-permissive at 30°C, related to Experimental Procedures. Spot assay of WT (WT) and *pfh1-R23* mutant cells on YEA plates at 36°C, 30°C, and 18°C. The fully permissive temperature for the *ts* allele is 36°C, while the restrictive is 18°C. At the intermediate semi-permissive temperature growth is slowed.

Table 5.1 | *S. pombe* strains used in this work.

Strain Name	Genotype	Reference	Use
DG21	*h-, otr1R(SphI)::ura4+, ura4-DS/E, leu1-32, ade6-216, his7- 366*	(Li et al., 2005)	sRNA-seq, Pol II ChIP
DG690	*h-, delta-dcr1::kanMX6, otr1R(SphI)::ura4+, ura4-DS/E, leu1- 32, ade6-210, his7-366*	(Irvine et al., 2009)	sRNA-seq, Pol II ChIP
ZB20	*h-, delta-ago1::kanMX6, otr1R(SphI)::ura4+, ura4-DS/E, leu1-32, ade6-216, his7-366*	(Zaratiegui et al., 2011)	Pol II ChIP

ZB515	*h-, rad22-YFP::kanMX6, ade6+*	(Meister et al., 2003)	Rad52 ChIP
AK69	*rad22-YFP::kanMX6, delta-dcr1::kanMX6, ade6-M210*	(Zaratiegui et al., 2011)	Rad52 ChIP
MO2299	*crb2-18myc, ura4-D18, leu1-32, ade6-M216*	(Elzen and O'Connell, 2004)	Crb2 ChIP
SC035	*dcr1-kanMX, crb2-18myc, ura4-D18, leu1-32, ade6-M216*	this study	Crb2 ChIP
H276	*h+, pfh1-r23, leu1-32*	(Tanaka et al., 2002)	sRNA-seq
SC028	*pfh1-r23, dcr1-kanMX*	this study	sRNA-seq
pB317	*delta-lsd1::ura4;lsd1+, ade6-210;ade6-216*	(Holmes et al., 2012)	Synthetic Lethality
SC044	*delta-dcr1::KanMX6;dcr1+, ade-216;ade-210, delta-lsd1::ura4;lsd1+*	this study	Synthetic Lethality
SB13	*h+, dcr1Δ::kanMX6, otr1R(SphI)::ura4, ura4-DS/E, ade6-216, leu1-32*	this study	Meiotic rDNA F1
SB14	*h+, ago1Δ::hpMX4, otr1R(SphI)::ura4, ura4-DS/E, ade6-216, leu1-32*	this study	Meiotic rDNA F1
SB15	*h+, otr1R(SphI)::ura4, ura4-DS/E, ade6-216, leu1-32*	this study	Meiotic rDNA F1
SB16	*h+, otr1R(SphI)::ura4, ura4-DS/E, ade6-216, leu1-32*	this study	Meiotic rDNA F2
SB17	*h+, ago1Δ::hpMX4, otr1R(SphI)::ura4, ura4-DS/E, ade6-216, leu1-32*	this study	Meiotic rDNA F2
SB18	*h+, otr1R(SphI)::ura4, ura4-DS/E, ade6-216, leu1-*	this study	Meiotic

	32		rDNA F2
SB19	*h+, dcr1Δ::kanMX6, otr1R(SphI)::ura4, ura4-DS/E, ade6-216, leu1-32*	this study	Meiotic rDNA F3
SB20	*h+, ago1Δ::hpMX4, otr1R(SphI)::ura4, ura4-DS/E, ade6-216, leu1-32*	this study	Meiotic rDNA F3
SB21	*h+, otr1R(SphI)::ura4, ura4-DS/E, ade6-216, leu1-32*	this study	Meiotic rDNA F3
SB22	*h+, dcr1Δ::kanMX6, otr1R(SphI)::ura4, ura4-DS/E, ade6-216, leu1-32*	this study	Meiotic rDNA F4
SB23	*h+, ago1Δ::hpMX4, otr1R(SphI)::ura4, ura4-DS/E, ade6-216, leu1-32*	this study	Meiotic rDNA F4
SB24	*h+, otr1R(SphI)::ura4, ura4-DS/E, ade6-216, leu1-32*	this study	Meiotic rDNA F4

Table 5.2 | Oligonucleotides used in this work.

Name	**Sequence**
p30_qPCR_F (*dh*)	CCATATCAATTTCCCATGTTCC
p30_qPCR_R (*dh*)	CATCAAGCGAGTCGAGATGA
p33_qPCR_F (*dg*)	TATCCTGCGTCTCGGTATCC
p33_qPCR_R (*dg*)	CTGTTCGTGAATGCTGAGAAAG
p20_qPCR_F (*ori2055*)	CCGGCGATTGAGAAAGACTACAA
p20_qPCR_R (*ori2055*)	TCGAAAAGATACGGCCAATAACA
act1_qPCR_F	TGCACCTGCCTTTTATGTTG
act1_qPCR_R	TGGGAACAGTGTGGGTAACA
rDNA18S_qPCR_F	CCCTGCATTGTTATTTCTTG

rDNA18S_qPCR_R	TCAACTTTCGATGGTAGGAT
rDNA28S_qPCR_F	GCTTGGTTGAATTTCTTCAC
rDNA28S_qPCR_R	CCAACTTAGAACTGGTACGG
HIS.02_qPCR_F	CTGGTGTGGGCACTTACTAT
HIS.02_qPCR_R	ATGGATCTATTTGGGATGC
SER.07_qPCR_F	CCGCAACAGATTTCTAGTCT
SER.07_qPCR_R	AGCTTTTATAATTTTCGACTT

Table 5.3 | Illumina sequencing libraries and statistics.

Library	Run Type	Total Reads	Aligned Reads	Duplicates Removed
polII_dg21_1_s2	PE 50	3.78E+06	3.24E+06	1.77E+06
polII_dg21_1_s5	PE 50	8.96E+06	8.38E+06	3.45E+06
polII_dg21_2_s2	PE 50	1.07E+07	9.10E+06	2.26E+06
polII_dg21_2_s5	PE 50	7.68E+06	7.14E+06	2.48E+06
polII_dg21_wce	PE 50	2.23E+06	2.06E+06	2.00E+06
polII_dg690_1_s2	PE 50	1.04E+07	9.52E+06	2.07E+06
polII_dg690_1_s5	PE 50	8.64E+06	8.23E+06	2.71E+06
polII_dg690_2_s2	PE 50	2.85E+06	2.63E+06	1.07E+06
polII_dg690_2_s5	PE 50	4.03E+06	3.73E+06	2.21E+06
polII_dg690_wce	PE 50	4.68E+06	4.19E+06	3.91E+06
srna_dg21	SR 36	1.75E+06	1.27E+06	N/A
srna_dg690	SR 36	1.40E+06	1.09E+06	N/A

h3k9me2_fy648_ip	PE 150	2.38E+07	2.28E+07	9.65E+06
h3k9me2_fy648_wce	PE 150	1.76E+06	1.70E+06	1.63E+06
h3k9me2_dg690_ip	PE 150	1.86E+07	1.75E+07	9.55E+06
h3k9me2_dg690_wce	PE 150	6.48E+05	6.16E+05	6.04E+05
rad52_zb515_exp_ip	PE 100	1.77E+06	1.05E+06	9.02E+05
rad52_zb515_exp_wce	PE 100	1.72E+06	1.70E+06	1.69E+06
rad52_ak69_exp_ip	PE 100	2.50E+06	6.71E+05	5.40E+05
rad52_ak69_exp_wce	PE 100	9.13E+05	8.92E+05	8.88E+05
rad52_zb515_s_ip	PE 100	2.06E+06	7.63E+05	6.63E+05
rad52_zb515_s_wce	PE 100	1.56E+06	1.54E+06	1.53E+06
rad52_ak69_s_ip	PE 100	1.89E+06	8.15E+05	5.97E+05
rad52_ak69_s_wce	PE 100	2.13E+06	2.10E+06	2.08E+06
crb2_wt_ip1	PE 150	2.61E+06	2.34E+06	2.28E+06
crb2_wt_ip2	PE 150	1.75E+06	1.49E+06	1.43E+06
crb2_wt_ip3	PE 150	3.79E+06	2.78E+06	2.71E+06
crb2_wt_wce	PE 150	2.58E+06	2.55E+06	2.51E+06
crb2_dcr_ip1	PE 150	2.48E+06	2.31E+06	2.25E+06
crb2_dcr_ip2	PE 150	2.17E+06	1.92E+06	1.88E+06
crb2_dcr_ip3	PE 150	2.22E+06	2.02E+06	1.98E+06
crb2_dcr_wce	PE 150	2.89E+06	2.85E+06	2.81E+06
srna_pfh1_dg21	SR 36	1.05E+06	9.43E+05	N/A

srna_pfh1_h276	SR 36	1.16E+06	1.05E+06	N/A
srna_pfh1_sc028	SR 36	1.44E+06	1.30E+06	N/A
drip_dg21_ip_1_-h	PE 150	1.62E+07	4.78E+06	2.34E+06
drip_dg690_ip_1_-h	PE 150	1.21E+07	3.88E+06	1.94E+06
drip_dg690_ip_1_+h	PE 150	1.47E+07	7.24E+06	2.78E+06
drip_dg21_ip_2_-h	PE 150	1.26E+07	6.13E+06	2.88E+06
drip_dg21_ip_2_+h	PE 150	1.59E+07	6.64E+06	2.77E+06
drip_dg690_ip_2_-h	PE 150	9.93E+06	2.40E+06	1.49E+06
drip_dg690_ip_2_+h	PE 150	1.40E+07	6.56E+06	2.79E+06
drip_dg21_wce_-h	PE 150	3.70E+07	3.65E+07	1.97E+07
drip_dg21_wce_+h	PE 150	1.90E+07	1.84E+07	1.09E+07
drip_dg690_wce_-h	PE 150	3.10E+07	3.03E+07	1.82E+07
drip_dg690_wce_+h	PE 150	1.52E+07	1.48E+07	9.01E+06

6. References

Aguilera, A., and García-Muse, T. (2012). R loops: from transcription byproducts to threats to genome stability. Mol Cell *46*, 115–124.

Alexeev, A., Mazin, A., and Kowalczykowski, S.C. (2003). Rad54 protein possesses chromatin-remodeling activity stimulated by the Rad51-ssDNA nucleoprotein filament. Nat. Struct. Biol. *10*, 182–186.

Allshire, R.C., Javerzat, J.-P., Redhead, N.J., and Cranston, G. (1994). Position effect variegation at fission yeast centromeres. Cell *76*, 157–169.

Alzu, A., Bermejo, R., Begnis, M., Lucca, C., Piccini, D., Carotenuto, W., Saponaro, M., Brambati, A., Cocito, A., Foiani, M., et al. (2012). Senataxin associates with replication forks to protect fork integrity across RNA-polymerase-II-transcribed genes. Cell *151*, 835–846.

Amitani, I., Baskin, R.J., and Kowalczykowski, S.C. (2006). Visualization of Rad54, a chromatin remodeling protein, translocating on single DNA molecules. Mol Cell *23*, 143–148.

Aoki, K., Moriguchi, H., Yoshioka, T., Okawa, K., and Tabara, H. (2007). In vitro analyses of the production and activity of secondary small interfering RNAs in C. elegans. Embo J *26*, 5007–5019.

Aravin, A.A., Sachidanandam, R., Girard, A., Fejes-Toth, K., and Hannon, G.J. (2007). Developmentally Regulated piRNA Clusters Implicate MILI in Transposon Control. Science *316*, 744–747.

Aravin, A.A., Sachidanandam, R., Bourc'his, D., Schaefer, C., Pezic, D., Toth, K.F., Bestor, T., and Hannon, G.J. (2008). A piRNA Pathway Primed by Individual Transposons Is Linked to De Novo DNA Methylation in Mice. Mol Cell *31*, 785–799.

Aravin, A.A., van der Heijden, G.W., Castañeda, J., Vagin, V.V., Hannon, G.J., and Bortvin, A. (2009). Cytoplasmic Compartmentalization of the Fetal piRNA Pathway in Mice. PLoS Genet *5*, e1000764.

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.

Aronica, L., Bednenko, J., Noto, T., DeSouza, L.V., Siu, K.W.M., Loidl, J., Pearlman, R.E., Gorovsky, M.A., and Mochizuki, K. (2008). Study of an RNA helicase implicates small RNA-noncoding RNA interactions in programmed DNA elimination in Tetrahymena. *22*, 2228–2241.

Ashe, A., Sapetschnig, A., Weick, E.-M., Mitchell, J., Bagijn, M.P., Cording, A.C., Doebley, A.-L., Goldstein, L.D., Lehrbach, N.J., Le Pen, J., et al. (2012). piRNAs Can Trigger a Multigenerational Epigenetic Memory in the Germline of C. elegans. Cell *150*, 88–99.

Aufsatz, W., Mette, M.F., van der Winden, J., Matzke, A.J.M., and Matzke, M. (2002). RNA-directed DNA methylation in Arabidopsis. Proc Natl Acad Sci USA *99 Suppl 4*, 16499–16506.

Ayoub, N., Goldshmidt, I., and Cohen, A. (1999). Position effect variegation at the mating-type locus of fission yeast: a cis-acting element inhibits covariegated expression of genes in the silent and expressed domains. Genetics *152*, 495–508.

Azvolinsky, A., Giresi, P.G., Lieb, J.D., and Zakian, V.A. (2009). Highly Transcribed RNA Polymerase II Genes Are Impediments to Replication Fork Progression in Saccharomyces cerevisiae. Mol Cell *34*, 722–734.

Barraud, P., Emmerth, S., Shimada, Y., Hotz, H.-R., Allain, F.H.-T., and Bühler, M. (2011). An extended dsRBD with a novel zinc-binding motif mediates nuclear retention of fission yeast Dicer. Embo J *30*, 4223–4235.

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.

Bayne, E.H., Portoso, M., Kagansky, A., Kos-Braun, I.C., Urano, T., Ekwall, K., Alves, F., Rappsilber, J., and Allshire, R.C. (2008). Splicing factors facilitate RNAi-directed silencing in fission yeast. Science *322*, 602–606.

Bäurle, I., Smith, L., Baulcombe, D.C., and Dean, C. (2007). Widespread role for the flowering-time regulators FCA and FPA in RNA-mediated chromatin silencing. Science *318*, 109–112.

Bermejo, R., Lai, M.S., and Foiani, M. (2012). Preventing replication stress to maintain genome stability: resolving conflicts between replication and transcription. Mol Cell *45*, 710–718.

Brennecke, J., Malone, C.D., Aravin, A.A., Sachidanandam, R., Stark, A., and Hannon, G.J. (2008). An Epigenetic Role for Maternally Inherited piRNAs in Transposon Silencing. Science *322*, 1387–1392.

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.

Brosnan, C.A., Mitter, N., Christie, M., Smith, N.A., Waterhouse, P.M., and Carroll, B.J. (2007). Nuclear gene silencing directs reception of long-distance mRNA silencing in Arabidopsis. Proc Natl Acad Sci USA *104*, 14741–14746.

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

Buker, S.M., Iida, T., Bühler, M., Villén, J., Gygi, S.P., Nakayama, J.-I., and Moazed, D. (2007). Two different Argonaute complexes are required for siRNA generation and heterochromatin assembly in fission yeast. Nat Struct Mol Biol *14*, 200–207.

Burkhart, K.B., Guang, S., Buckley, B.A., Wong, L., Bochner, A.F., and Kennedy, S. (2011). A pre-mRNA-associating factor links endogenous siRNAs to chromatin regulation. PLoS Genet *7*, e1002249.

Burton, N.O., Burkhart, K.B., and Kennedy, S. (2011). Nuclear RNAi maintains heritable gene silencing in Caenorhabditis elegans. Proc Natl Acad Sci USA *108*, 19683–19688.

Buscaino, A., White, S.A., Houston, D.R., Lejeune, E., Simmer, F., de Lima Alves, F., Diyora, P.T., Urano, T., Bayne, E.H., Rappsilber, J., et al. (2012). Raf1 Is a DCAF for the Rik1 DDB1-Like Protein and Has Separable Roles in siRNA Generation and Chromatin Modification. PLoS Genet *8*, e1002499.

Bühler, M., and Gasser, S.M. (2009). Silent chromatin at the middle and ends: lessons from yeasts. Embo J *28*, 2149–2161.

Bühler, M., Spies, N., Bartel, D.P., and Moazed, D. (2008). TRAMP-mediated RNA surveillance prevents spurious entry of RNAs into the Schizosaccharomyces pombe siRNA pathway. Nat Struct Mol Biol *15*, 1015–1023.

Bühler, M., Verdel, A., and Moazed, D. (2006). Tethering RITS to a nascent transcript initiates RNAi- and heterochromatin-dependent gene silencing. Cell *125*, 873–886.

Calarco, J.P., Borges, F., Donoghue, M.T.A., Van Ex, F., Jullien, P.E., Lopes, T., Gardner, R., Berger, F., Feijó, J.A., Becker, J.D., et al. (2012). Reprogramming of DNA Methylation in Pollen Guides Epigenetic Inheritance via Small RNA. Cell *151*, 194–205.

Cao, X., Aufsatz, W., Zilberman, D., Mette, M.F., Huang, M.S., Matzke, M., and Jacobsen, S.E. (2003). Role of the DRM and CMT3 methyltransferases in RNA-directed DNA methylation. Curr Biol *13*, 2212–2217.

Carmell, M.A., Girard, A., van de Kant, H.J.G., 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.

Carr, A.M., and Lambert, S. (2013). Replication Stress-Induced Genome Instability: The Dark Side of Replication Maintenance by Homologous Recombination. Journal of Molecular Biology *425*, 4733–4744.

Castel, S.E., and Martienssen, R.A. (2013). RNA interference in the nucleus: roles for small RNAs in transcription, epigenetics and beyond. Nat Rev Genet *14*, 100–112.

Cecere, G., and Cogoni, C. (2009). Quelling targets the rDNA locus and functions in rDNA copy number control. BMC Microbiol *9*, 44.

Cernilogar, F.M., Onorati, M.C., Kothe, G.O., Burroughs, A.M., Parsi, K.M., Breiling, A., Sardo, Lo, F., Saxena, A., Miyoshi, K., Siomi, H., et al. (2011). Chromatin-associated RNA interference components contribute to transcriptional regulation in Drosophila. Nature *480*, 391–395.

Chen, E.S., Zhang, K., Nicolas, E., Cam, H.P., Zofall, M., and Grewal, S.I.S. (2008). Cell cycle control of centromeric repeat transcription and heterochromatin assembly. Nature *451*, 734–737.

Choi, Y., Gehring, M., Johnson, L., Hannon, M., Harada, J.J., Goldberg, R.B., Jacobsen, S.E., and Fischer, R.L. (2002). DEMETER, a DNA glycosylase domain protein, is required for endosperm gene imprinting and seed viability in arabidopsis. Cell *110*, 33–42.

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

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.

Dalzell, J.J., McVeigh, P., Warnock, N.D., Mitreva, M., Bird, D.M., Abad, P., Fleming, C.C., Day, T.A., Mousley, A., Marks, N.J., et al. (2011). RNAi effector diversity in nematodes. PLoS Negl Trop

Dis *5*, e1176.

Das, P.P., Bagijn, M.P., Goldstein, L.D., Woolford, J.R., Lehrbach, N.J., Sapetschnig, A., Buhecha, H.R., Gilchrist, M.J., Howe, K.L., Stark, R., et al. (2008). Piwi and piRNAs Act Upstream of an Endogenous siRNA Pathway to Suppress Tc3 Transposon Mobility in the Caenorhabditis elegans Germline. Mol Cell *31*, 79–90.

Deshpande, G. (2005). Drosophila argonaute-2 is required early in embryogenesis for the assembly of centric/centromeric heterochromatin, nuclear division, nuclear migration, and germ-cell formation. Genes Dev *19*, 1680–1685.

Djupedal, I., Portoso, M., Spåhr, H., Bonilla, C., Gustafsson, C.M., Allshire, R.C., and Ekwall, K. (2005). RNA Pol II subunit Rpb7 promotes centromeric transcription and RNAi-directed chromatin silencing. Genes Dev *19*, 2301–2306.

Dunoyer, P., Brosnan, C.A., Schott, G., Wang, Y., Jay, F., Alioua, A., Himber, C., and Voinnet, O. (2010). An endogenous, systemic RNAi pathway in plants. Embo J *29*, 1699–1712.

El-Shami, M., Pontier, D., Lahmy, S., Braun, L., Picart, C., Vega, D., Hakimi, M.-A., Jacobsen, S.E., Cooke, R., and Lagrange, T. (2007). Reiterated WG/GW motifs form functionally and evolutionarily conserved ARGONAUTE-binding platforms in RNAi-related components. Genes Dev *21*, 2539–2544.

Elgin, S.C.R., and Reuter, G. (2013). Position-effect variegation, heterochromatin formation, and gene silencing in Drosophila. Cold Spring Harb Perspect Biol *5*, a017780.

Ellermeier, C., Higuchi, E.C., Phadnis, N., Holm, L., Geelhood, J.L., Thon, G., and Smith, G.R. (2010). RNAi and heterochromatin repress centromeric meiotic recombination. Proc Natl Acad Sci USA *107*, 8701–8705.

Elzen, den, N.R., and O'Connell, M.J. (2004). Recovery from DNA damage checkpoint arrest by PP1-mediated inhibition of Chk1. Embo J *23*, 908–918.

Emmerth, S., Schober, H., Gaidatzis, D., Roloff, T., Jacobeit, K., and Bühler, M. (2010). Nuclear retention of fission yeast dicer is a prerequisite for RNAi-mediated heterochromatin assembly. Dev Cell *18*, 102–113.

Enke, R.A., Dong, Z., and Bender, J. (2011). Small RNAs Prevent Transcription-Coupled Loss of Histone H3 Lysine 9 Methylation in Arabidopsis thaliana. PLoS Genet *7*, e1002350.

Fagegaltier, D., Bouge, A.L., Berry, B., Poisot, E., Sismeiro, O., Coppee, J.Y., Theodore, L., Voinnet, O., and Antoniewski, C. (2009). The endogenous siRNA pathway is involved in heterochromatin formation in Drosophila. Proc Natl Acad Sci USA *106*, 21258–21263.

Francia, S., Michelini, F., Saxena, A., Tang, D., de Hoon, M., Anelli, V., Mione, M., Carninci, P., and d'Adda di Fagagna, F. (2012). Site-specific DICER and DROSHA RNA products control the DNA-damage response. Nature *488*, 231–235.

Gao, Z., Liu, H.-L., Daxinger, L., Pontes, O., He, X., Qian, W., Lin, H., Xie, M., Lorkovic, Z.J., Zhang, S., et al. (2010). An RNA polymerase II- and AGO4-associated protein acts in RNA-directed DNA methylation. Nature *465*, 106–109.

Ghazal, G., Gagnon, J., Jacques, P.-É., Landry, J.-R., Robert, F., and Abou Elela, S. (2009). Yeast RNase III Triggers Polyadenylation-Independent Transcription Termination. Mol Cell *36*, 99–109.

Ginno, P.A., Lott, P.L., Christensen, H.C., Korf, I., and Chedin, F. (2012). R-Loop Formation Is a Distinctive Characteristic of Unmethylated Human CpG Island Promoters. Mol Cell.

Goto, D.B., and Nakayama, J.-I. (2012). RNA and epigenetic silencing: insight from fission yeast. Dev. Growth Differ. *54*, 129–141.

Gottlieb, S., and Esposito, R.E. (1989). A new role for a yeast transcriptional silencer gene, SIR2, in regulation of recombination in ribosomal DNA. Cell *56*, 771–776.

Gregory, B.D., O'Malley, R.C., Lister, R., Urich, M.A., Tonti-Filippini, J., Chen, H., Millar, A.H., and Ecker, J.R. (2008). A link between RNA metabolism and silencing affecting Arabidopsis development. Dev Cell *14*, 854–866.

Grewal, S.I. (2010). RNAi-dependent formation of heterochromatin and its diverse functions. Curr Opin Genet Dev *20*, 134–141.

Gu, S.G., Pak, J., Guang, S., Maniar, J.M., Kennedy, S., and Fire, A. (2012). Amplification of siRNA in Caenorhabditis elegans generates a transgenerational sequence-targeted histone H3 lysine 9 methylation footprint. Nat Genet *44*, 157–164.

Guang, S., Bochner, A.F., Pavelec, D.M., Burkhart, K.B., Harding, S., Lachowiec, J., and Kennedy, S. (2008). An Argonaute Transports siRNAs from the Cytoplasm to the Nucleus. Science *321*, 537–541.

Guang, S., Bochner, A.F., Burkhart, K.B., Burton, N., Pavelec, D.M., and Kennedy, S. (2010). Small regulatory RNAs inhibit RNA polymerase II during the elongation phase of transcription. Nature *465*, 1097–1101.

Gullerova, M., Moazed, D., and Proudfoot, N.J. (2011). Autoregulation of convergent RNAi genes in fission yeast. Genes Dev *25*, 556–568.

Gullerova, M., and Proudfoot, N.J. (2008). Cohesin complex promotes transcriptional termination between convergent genes in S. pombe. Cell *132*, 983–995.

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.

Haag, J.R., and Pikaard, C.S. (2011). Multisubunit RNA polymerases IV and V: purveyors of non-coding RNA for plant gene silencing. Nat Rev Mol Cell Biol *12*, 483–492.

Halic, M., and Moazed, D. (2010). Dicer-independent primal RNAs trigger RNAi and heterochromatin formation. Cell *140*, 504–516.

Hanawalt, P.C., and Spivak, G. (2008). Transcription-coupled DNA repair: two decades of progress and surprises. Nat Rev Mol Cell Biol *9*, 958–970.

Hansen, K.R., Burns, G., Mata, J., Volpe, T.A., Martienssen, R.A., Bähler, J., and Thon, G. (2005). Global effects on gene expression in fission yeast by silencing and RNA interference machineries. Mol

Cell Biol *25*, 590–601.

He, X.-J., Hsu, Y.-F., Zhu, S., Wierzbicki, A.T., Pontes, O., Pikaard, C.S., Liu, H.-L., Wang, C.-S., Jin, H., and Zhu, J.-K. (2009). An Effector of RNA-Directed DNA Methylation in Arabidopsis Is an ARGONAUTE 4- and RNA-Binding Protein. Cell *137*, 498–508.

Henderson, I.R., Deleris, A., Wong, W., Zhong, X., Chin, H.G., Horwitz, G.A., Kelly, K.A., Pradhan, S., and Jacobsen, S.E. (2010). The De Novo Cytosine Methyltransferase DRM2 Requires Intact UBA Domains and a Catalytically Mutated Paralog DRM3 during RNA–Directed DNA Methylation in Arabidopsis thaliana. PLoS Genet *6*, e1001182.

Herr, A.J., Molnar, A., Jones, A., and Baulcombe, D.C. (2006). Defective RNA processing enhances RNA silencing and influences flowering of Arabidopsis. Proc Natl Acad Sci USA *103*, 14994–15001.

Holmes, A., Roseaulin, L., Schurra, C., Waxin, H., Lambert, S., Zaratiegui, M., Martienssen, R.A., and Arcangioli, B. (2012). Lsd1 and Lsd2 Control Programmed Replication Fork Pauses and Imprinting in Fission Yeast. Cell Reports *2*, 1513–1520.

Hsieh, T.-F., Ibarra, C.A., Silva, P., Zemach, A., Eshed-Williams, L., Fischer, R.L., and Zilberman, D. (2009). Genome-wide demethylation of Arabidopsis endosperm. Science *324*, 1451–1454.

Ibarra, C.A., Feng, X., Schoft, V.K., Hsieh, T.F., Uzawa, R., Rodrigues, J.A., Zemach, A., Chumak, N., Machlicova, A., Nishimura, T., et al. (2012). Active DNA Demethylation in Plant Companion Cells Reinforces Transposon Methylation in Gametes. Science *337*, 1360–1364.

Iino, Y., Sugimoto, A., and Yamamoto, M. (1991). S. pombe pac1+, whose overexpression inhibits sexual development, encodes a ribonuclease III-like RNase. Embo J *10*, 221–226.

Iovine, B., Iannella, M.L., and Bevilacqua, M.A. (2011). Damage-specific DNA binding protein 1 (DDB1): a protein with a wide range of functions. Int J Biochem Cell Biol *43*, 1664–1667.

Irvine, D.V., Goto, D.B., Vaughn, M.W., Nakaseko, Y., McCombie, W.R., Yanagida, M., and Martienssen, R. (2009). Mapping epigenetic mutations in fission yeast using whole-genome next-generation sequencing. Genome Research *19*, 1077–1083.

Irvine, D.V., Zaratiegui, M., Tolia, N.H., Goto, D.B., Chitwood, D.H., Vaughn, M.W., Joshua-Tor, L., and Martienssen, R.A. (2006). Argonaute slicing is required for heterochromatic silencing and spreading. Science *313*, 1134–1137.

Ji, L., and Chen, X. (2012). Regulation of small RNA stability: methylation and beyond. Cell Research.

Johnson, L.M., Bostick, M., Zhang, X., Kraft, E., Henderson, I., Callis, J., and Jacobsen, S.E. (2007). The SRA methyl-cytosine-binding domain links DNA and histone methylation. Curr Biol *17*, 379–384.

Johnson, L.M., Law, J.A., Khattar, A., Henderson, I.R., and Jacobsen, S.E. (2008). SRA-Domain Proteins Required for DRM2-Mediated De Novo DNA Methylation. PLoS Genet *4*, e1000280.

Jullien, P.E., Mosquna, A., Ingouff, M., Sakata, T., Ohad, N., and Berger, F. (2008). Retinoblastoma and its binding partner MSI1 control imprinting in Arabidopsis. PLoS Biol *6*, e194.

Jullien, P.E., Susaki, D., Yelagandula, R., Higashiyama, T., and Berger, F. (2012). DNA Methylation

Dynamics during Sexual Reproduction in Arabidopsis thaliana. Current Biology *22*, 1825–1830.

Kanoh, J., Sadaie, M., Urano, T., and Ishikawa, F. (2005). Telomere Binding Protein Taz1 Establishes Swi6 Heterochromatin Independently of RNAi at Telomeres. Current Biology *15*, 1808–1819.

Kapranov, P., Cheng, J., Dike, S., Nix, D.A., Duttagupta, R., Willingham, A.T., Stadler, P.F., Hertel, J., Hackermüller, J., Hofacker, I.L., et al. (2007). RNA maps reveal new RNA classes and a possible function for pervasive transcription. Science *316*, 1484–1488.

Kasschau, K.D., Fahlgren, N., Chapman, E.J., Sullivan, C.M., Cumbie, J.S., Givan, S.A., and Carrington, J.C. (2007). Genome-Wide Profiling and Analysis of Arabidopsis siRNAs. PLoS Biol *5*, e57.

Kato, H., Goto, D.B., Martienssen, R.A., Urano, T., Furukawa, K., and Murakami, Y. (2005). RNA polymerase II is required for RNAi-dependent heterochromatin assembly. Science *309*, 467–469.

Kavi, H.H., and Birchler, J.A. (2009). Interaction of RNA polymerase II and the small RNA machinery affects heterochromatic silencing in Drosophila. Epigenetics Chromatin *2*, 15.

Khraiwesh, B., Arif, M.A., Seumel, G.I., Ossowski, S., Weigel, D., Reski, R., and Frank, W. (2010). Transcriptional Control of Gene Expression by MicroRNAs. Cell *140*, 111–122.

Khurana, J.S., Xu, J., Weng, Z., and Theurkauf, W.E. (2010). Distinct Functions for the Drosophila piRNA Pathway in Genome Maintenance and Telomere Protection. PLoS Genet *6*, e1001246.

Kim, D.H., Villeneuve, L.M., Morris, K.V., and Rossi, J.J. (2006). Argonaute-1 directs siRNA-mediated transcriptional gene silencing in human cells. Nat Struct Mol Biol *13*, 793–797.

Kim, V.N., Han, J., and Siomi, M.C. (2009). Biogenesis of small RNAs in animals. Nat Rev Mol Cell Biol *10*, 126–139.

Klattenhoff, C., Xi, H., Li, C., Lee, S., Xu, J., Khurana, J.S., Zhang, F., Schultz, N., Koppetsch, B.S., Nowosielska, A., et al. (2009). The Drosophila HP1 homolog Rhino is required for transposon silencing and piRNA production by dual-strand clusters. Cell *138*, 1137–1149.

Klenov, M.S., Sokolova, O.A., Yakushev, E.Y., Stolyarenko, A.D., Mikhaleva, E.A., Lavrov, S.A., and Gvozdev, V.A. (2011). Separation of stem cell maintenance and transposon silencing functions of Piwi protein. Proc Natl Acad Sci USA *108*, 18760–18765.

Kloc, A., Zaratiegui, M., Nora, E., and Martienssen, R. (2008). RNA interference guides histone modification during the S phase of chromosomal replication. Curr Biol *18*, 490–495.

Kobayashi, T., and Ganley, A.R.D. (2005). Recombination regulation by transcription-induced cohesin dissociation in rDNA repeats. Science *309*, 1581–1584.

Kou, H., Zhou, Y., Gorospe, R.M.C., and Wang, Z. (2008). Mms19 protein functions in nucleotide excision repair by sustaining an adequate cellular concentration of the TFIIH component Rad3. Proc Natl Acad Sci USA *105*, 15714–15719.

Lambert, S., Mizuno, K., Blaisonneau, J., Martineau, S., Chanet, R., Fréon, K., Murray, J.M., Carr, A.M., and Baldacci, G. (2010). Homologous recombination restarts blocked replication forks at the

expense of genome rearrangements by template exchange. Mol Cell *39*, 346–359.

Lan, F., Zaratiegui, M., Villén, J., Vaughn, M.W., Verdel, A., Huarte, M., Shi, Y., Gygi, S.P., Moazed, D., Martienssen, R.A., et al. (2007). S. pombe LSD1 Homologs Regulate Heterochromatin Propagation and Euchromatic Gene Transcription. Mol Cell *26*, 89–101.

Lans, H., Marteijn, J.A., and Vermeulen, W. (2012). ATP-dependent chromatin remodeling in the DNA-damage response. Epigenetics Chromatin *5*, 4.

Lauder, S., Bankmann, M., Guzder, S.N., Sung, P., Prakash, L., and Prakash, S. (1996). Dual requirement for the yeast MMS19 gene in DNA repair and RNA polymerase II transcription. Mol Cell Biol *16*, 6783–6793.

Law, J.A., and Jacobsen, S.E. (2010). Establishing, maintaining and modifying DNA methylation patterns in plants and animals. Nat Rev Genet *11*, 204–220.

Law, J.A., Vashisht, A.A., Wohlschlegel, J.A., and Jacobsen, S.E. (2011). SHH1, a Homeodomain Protein Required for DNA Methylation, As Well As RDR2, RDM4, and Chromatin Remodeling Factors, Associate with RNA Polymerase IV. PLoS Genet *7*, e1002195.

Lee, H.-C., Chang, S.-S., Choudhary, S., Aalto, A.P., Maiti, M., Bamford, D.H., and Liu, Y. (2009). qiRNA is a new type of small interfering RNA induced by DNA damage. Nature *459*, 274–277.

Lee, H.-C., Gu, W., Shirayama, M., Youngman, E., Conte, D., and Mello, C.C. (2012). C. elegans piRNAs Mediate the Genome-wide Surveillance of Germline Transcripts. Cell.

Lee, H.-C., Li, L., Gu, W., Xue, Z., Crosthwaite, S.K., Pertsemlidis, A., Lewis, Z.A., Freitag, M., Selker, E.U., and Mello, C.C. (2010). Diverse Pathways Generate MicroRNA-like RNAs and Dicer-Independent Small Interfering RNAs in Fungi. Mol Cell *38*, 803–814.

Lejeune, E., Bayne, E.H., and Allshire, R.C. (2011). On the Connection between RNAi and Heterochromatin at Centromeres. Cold Spring Harbor Symposia on Quantitative Biology.

Li, C., Vagin, V.V., Lee, S., Xu, J., Ma, S., Xi, H., Seitz, H., Horwich, M.D., Syrzycka, M., Honda, B.M., et al. (2009). Collapse of germline piRNAs in the absence of Argonaute3 reveals somatic piRNAs in flies. Cell *137*, 509–521.

Li, F., Zheng, Q., Vandivier, L.E., Willmann, M.R., Chen, Y., and Gregory, B.D. (2012). Regulatory impact of RNA secondary structure across the Arabidopsis transcriptome. The Plant Cell Online *24*, 4346–4359.

Li, F., Goto, D.B., Zaratiegui, M., Tang, X., Martienssen, R., and Cande, W.Z. (2005). Two novel proteins, dos1 and dos2, interact with rik1 to regulate heterochromatic RNA interference and histone modification. Curr Biol *15*, 1448–1457.

Li, F., Martienssen, R., and Cande, W.Z. (2011). Coordination of DNA replication and histone modification by the Rik1-Dos2 complex. Nature *475*, 244–248.

Li, P.-C., Green, M.D., and Forsburg, S.L. (2013). Mutations disrupting histone methylation have different effects on replication timing in S. pombe centromere. PLoS ONE *8*, e61464.

Lippman, Z., Gendrel, A.-V., Black, M., Vaughn, M.W., Dedhia, N., Richard McCombie, W., 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.

Liu, F., Bakht, S., and Dean, C. (2012). Cotranscriptional role for Arabidopsis DICER-LIKE 4 in transcription termination. Science *335*, 1621–1623.

Lu, J., Zhang, C., Baulcombe, D.C., and Chen, Z.J. (2012). Maternal siRNAs as regulators of parental genome imbalance and gene expression in endosperm of Arabidopsis seeds. Proc Natl Acad Sci USA.

Luna, R., Jimeno, S., Marín, M., Huertas, P., García-Rubio, M., and Aguilera, A. (2005). Interdependence between transcription and mRNP processing and export, and its impact on genetic stability. Mol Cell *18*, 711–722.

Maine, E.M., Hauth, J., Ratliff, T., Vought, V.E., She, X., and Kelly, W.G. (2005). EGO-1, a putative RNA-dependent RNA polymerase, is required for heterochromatin assembly on unpaired dna during C. elegans meiosis. Curr Biol *15*, 1972–1978.

Malone, C.D., Brennecke, J., Dus, M., Stark, A., McCombie, W.R., Sachidanandam, R., and Hannon, G.J. (2009). Specialized piRNA pathways act in germline and somatic tissues of the Drosophila ovary. Cell *137*, 522–535.

Marasovic, M., Zocco, M., and Halic, M. (2013). Argonaute and Triman generate dicer-independent priRNAs and mature siRNAs to initiate heterochromatin formation. Mol Cell *52*, 173–183.

Martienssen, R.A. (2010). Heterochromatin, small RNA and post-fertilization dysgenesis in allopolyploid and interploid hybrids of Arabidopsis. New Phytol. *186*, 46–53.

Meister, P., Poidevin, M., Francesconi, S., Tratner, I., Zarzov, P., and Baldacci, G. (2003). Nuclear factories for signalling and repairing DNA double strand breaks in living fission yeast. Nucleic Acids Res *31*, 5064–5073.

Melnyk, C.W., Molnar, A., Bassett, A., and Baulcombe, D.C. (2011). Mobile 24 nt small RNAs direct transcriptional gene silencing in the root meristems of Arabidopsis thaliana. Curr Biol *21*, 1678–1683.

Mette, M.F., Aufsatz, W., van der Winden, J., Matzke, M.A., and Matzke, A.J. (2000). Transcriptional silencing and promoter methylation triggered by double-stranded RNA. Embo J *19*, 5194–5201.

Michalik, K.M., Böttcher, R., and Förstemann, K. (2012). A small RNA response at DNA ends in Drosophila. Nucleic Acids Res.

Mischo, H.E., Gómez-González, B., Grzechnik, P., Rondón, A.G., Wei, W., Steinmetz, L., Aguilera, A., and Proudfoot, N.J. (2011). Yeast Sen1 helicase protects the genome from transcription-associated instability. Mol Cell *41*, 21–32.

Miyoshi, T., Takeuchi, A., Siomi, H., and Siomi, M.C. (2010). A direct role for Hsp90 in pre-RISC formation in Drosophila. Nat Struct Mol Biol *17*, 1024–1026.

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.

Molnar, A., Melnyk, C.W., Bassett, A., Hardcastle, T.J., Dunn, R., and Baulcombe, D.C. (2010). Small Silencing RNAs in Plants Are Mobile and Direct Epigenetic Modification in Recipient Cells. Science *328*, 872–875.

Mosher, R.A., Tan, E.H., Shin, J., Fischer, R.L., Pikaard, C.S., and Baulcombe, D.C. (2011). An Atypical Epigenetic Mechanism Affects Uniparental Expression of Pol IV-Dependent siRNAs. PLoS ONE *6*, e25756.

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.

Nakagawa, H., Lee, J.-K., Hurwitz, J., Allshire, R.C., Nakayama, J.-I., Grewal, S.I.S., Tanaka, K., and Murakami, Y. (2002). Fission yeast CENP-B homologs nucleate centromeric heterochromatin by promoting heterochromatin-specific histone tail modifications. Genes Dev *16*, 1766–1778.

Nakama, M., Kawakami, K., Kajitani, T., Urano, T., and Murakami, Y. (2012). DNA-RNA hybrid formation mediates RNAi-directed heterochromatin formation. Genes Cells *17*, 218–233.

Nakamura, T.M., Du, L.-L., Redon, C., and Russell, P. (2004). Histone H2A phosphorylation controls Crb2 recruitment at DNA breaks, maintains checkpoint arrest, and influences DNA repair in fission yeast. Mol Cell Biol *24*, 6215–6230.

Olmedo-Monfil, V., Durán-Figueroa, N., Arteaga-Vázquez, M., Demesa-Arévalo, E., Autran, D., Grimanelli, D., Slotkin, R.K., Martienssen, R.A., and Vielle-Calzada, J.-P. (2010). Control of female gamete formation by a small RNA pathway in Arabidopsis. Nature *464*, 628–632.

Onodera, Y., Haag, J.R., Ream, T., Costa Nunes, P., Pontes, O., and Pikaard, C.S. (2005). Plant nuclear RNA polymerase IV mediates siRNA and DNA methylation-dependent heterochromatin formation. Cell *120*, 613–622.

Padmanabhan, K., Robles, M.S., Westerling, T., and Weitz, C.J. (2012). Feedback regulation of transcriptional termination by the mammalian circadian clock PERIOD complex. Science *337*, 599–602.

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

Park, N.J., Tsao, D.C., and Martinson, H.G. (2004). The two steps of poly(A)-dependent termination, pausing and release, can be uncoupled by truncation of the RNA polymerase II carboxyl-terminal repeat domain. Mol Cell Biol *24*, 4092–4103.

Paul, S., Million-Weaver, S., Chattopadhyay, S., Sokurenko, E., and Merrikh, H. (2013). Accelerated gene evolution through replication-transcription conflicts. Nature *495*, 512–515.

Pek, J.W., and Kai, T. (2011a). A Role for Vasa in Regulating Mitotic Chromosome Condensation in Drosophila. Curr Biol *21*, 39–44.

Pek, J.W., and Kai, T. (2011b). DEAD-box RNA helicase Belle/DDX3 and the RNA interference pathway promote mitotic chromosome segregation. Proc Natl Acad Sci USA *108*, 12007–12012.

Peng, J.C., and Karpen, G.H. (2007). H3K9 methylation and RNA interference regulate nucleolar organization and repeated DNA stability. Nat Cell Biol *9*, 25–35.

Pélissier, T., Thalmeir, S., Kempe, D., Sänger, H.L., and Wassenegger, M. (1999). Heavy de novo methylation at symmetrical and non-symmetrical sites is a hallmark of RNA-directed DNA methylation. Nucleic Acids Res *27*, 1625–1634.

Popp, C., Dean, W., Feng, S., Cokus, S.J., Andrews, S., Pellegrini, M., Jacobsen, S.E., and Reik, W. (2010). Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. Nature *463*, 1101–1105.

Prado, F., and Aguilera, A. (2005). Impairment of replication fork progression mediates RNA polII transcription-associated recombination. Embo J *24*, 1267–1276.

Rechavi, O., Minevich, G., and Hobert, O. (2011). Transgenerational Inheritance of an Acquired Small RNA-Based Antiviral Response in C. elegans. Cell.

Reyes-Turcu, F.E., Zhang, K., Zofall, M., Chen, E., and Grewal, S.I.S. (2011). Defects in RNA quality control factors reveal RNAi-independent nucleation of heterochromatin. Nat Struct Mol Biol *18*, 1132–1138.

Rotondo, G., and Frendewey, D. (2001). Pac1 ribonuclease of Schizosaccharomyces pombe. Meth Enzymol *342*, 168–193.

Rowley, M.J., Avrutsky, M.I., Sifuentes, C.J., Pereira, L., and Wierzbicki, A.T. (2011). Independent Chromatin Binding of ARGONAUTE4 and SPT5L/KTF1 Mediates Transcriptional Gene Silencing. PLoS Genet *7*, e1002120.

Rozenzhak, S., Mejía-Ramírez, E., Williams, J.S., Schaffer, L., Hammond, J.A., Head, S.R., and Russell, P. (2010). Rad3 decorates critical chromosomal domains with gammaH2A to protect genome integrity during S-Phase in fission yeast. PLoS Genet *6*, e1001032.

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.

Sabouri, N., McDonald, K.R., Webb, C.J., Cristea, I.M., and Zakian, V.A. (2012). DNA replication through hard-to-replicate sites, including both highly transcribed RNA Pol II and Pol III genes, requires the S. pombe Pfh1 helicase. Genes Dev *26*, 581–593.

Sanchez, J. (1998). Ribosomal DNA Replication in the Fission Yeast,Schizosaccharomyces pombe. Exp. Cell Res. *238*, 220–230.

Savitsky, M. (2006). Telomere elongation is under the control of the RNAi-based mechanism in the Drosophila germline. Genes Dev *20*, 345–354.

Segurado, M., de Luis, A., and Antequera, F. (2003). Genome-wide distribution of DNA replication origins at A+T-rich islands in Schizosaccharomyces pombe. EMBO Rep *4*, 1048–1053.

Seo, J., Kim, S.C., Lee, H.-S., Kim, J.K., Shon, H.J., Salleh, N.L.M., Desai, K.V., Lee, J.H., Kang, E.-S., Kim, J.S., et al. (2012). Genome-wide profiles of H2AX and γ-H2AX differentiate endogenous and

exogenous DNA damage hotspots in human cells. Nucleic Acids Res *40*, 5965–5974.

Shabalina, S.A., and Koonin, E.V. (2008). Origins and evolution of eukaryotic RNA interference. Trends Ecol. Evol. (Amst.) *23*, 578–587.

She, X., Xu, X., Fedotov, A., Kelly, W.G., and Maine, E.M. (2009). Regulation of Heterochromatin Assembly on Unpaired Chromosomes during Caenorhabditis elegans Meiosis by Components of a Small RNA-Mediated Pathway. PLoS Genet *5*, e1000624.

Shirayama, M., Seth, M., Lee, H.-C., Gu, W., Ishidate, T., Conte, D., and Mello, C.C. (2012). piRNAs Initiate an Epigenetic Memory of Nonself RNA in the C. elegans Germline. Cell.

Shobuike, T., Tatebayashi, K., Tani, T., Sugano, S., and Ikeda, H. (2001). The dhp1(+) gene, encoding a putative nuclear 5“-->3” exoribonuclease, is required for proper chromosome segregation in fission yeast. Nucleic Acids Res *29*, 1326–1333.

Shpiz, S., and Kalmykova, A. (2011). Role of piRNAs in the Drosophila telomere homeostasis. Mge *1*, 274–278.

Simmer, F., Buscaino, A., Kos-Braun, I.C., Kagansky, A., Boukaba, A., Urano, T., Kerr, A.R.W., and Allshire, R.C. (2010). Hairpin RNA induces secondary small interfering RNA synthesis and silencing in trans in fission yeast. EMBO Rep *11*, 112–118.

Slotkin, R.K., Vaughn, M., Borges, F., Tanurdzić, M., Becker, J.D., Feijó, J.A., and Martienssen, R.A. (2009). Epigenetic reprogramming and small RNA silencing of transposable elements in pollen. Cell *136*, 461–472.

Smith, J.G., Caddle, M.S., Bulboaca, G.H., Wohlgemuth, J.G., Baum, M., Clarke, L., and Calos, M.P. (1995). Replication of centromere II of Schizosaccharomyces pombe. Mol Cell Biol *15*, 5165–5172.

Smith, J.S., and Boeke, J.D. (1997). An unusual form of transcriptional silencing in yeast ribosomal DNA. Genes Dev *11*, 241–254.

Sonmez, C., Bäurle, I., Magusin, A., Dreos, R., Laubinger, S., Weigel, D., and Dean, C. (2011). RNA 3' processing functions of Arabidopsis FCA and FPA limit intergenic transcription. Proc Natl Acad Sci USA *108*, 8508–8513.

Souret, F.F., Kastenmayer, J.P., and Green, P.J. (2004). AtXRN4 degrades mRNA in Arabidopsis and its substrates include selected miRNA targets. Mol Cell *15*, 173–183.

Tabach, Y., Billi, A.C., Hayes, G.D., Newman, M.A., Zuk, O., Gabel, H., Kamath, R., Yacoby, K., Chapman, B., Garcia, S.M., et al. (2013). Identification of small RNA pathway genes using patterns of phylogenetic conservation and divergence. Nature *493*, 694–698.

Tanaka, H., Ryu, G.-H., Seo, Y.-S., Tanaka, K., Okayama, H., MacNeill, S.A., and Yuasa, Y. (2002). The fission yeast pfh1(+) gene encodes an essential 5“ to 3” DNA helicase required for the completion of S-phase. Nucleic Acids Res *30*, 4728–4739.

Taverna, S.D., Coyne, R.S., and Allis, C.D. (2002). Methylation of histone h3 at lysine 9 targets programmed DNA elimination in tetrahymena. Cell *110*, 701–711.

Trautinger, B.W., Jaktaji, R.P., Rusakova, E., and Lloyd, R.G. (2005). RNA polymerase modulators and DNA repair activities resolve conflicts between DNA replication and transcription. Mol Cell *19*, 247–258.

Vasiljeva, L., Kim, M., Terzi, N., Soares, L.M., and Buratowski, S. (2008). Transcription termination and RNA degradation contribute to silencing of RNA polymerase II transcription within heterochromatin. Mol Cell *29*, 313–323.

Vazquez, F., Blevins, T., Ailhas, J., Boller, T., and Meins, F. (2008). Evolution of Arabidopsis MIR genes generates novel microRNA classes. Nucleic Acids Res *36*, 6429–6438.

Verdel, A. (2004). RNAi-Mediated Targeting of Heterochromatin by the RITS Complex. Science *303*, 672–676.

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

Volpe, T., Schramke, V., Hamilton, G.L., White, S.A., Teng, G., Martienssen, R.A., and Allshire, R.C. (2003). RNA interference is required for normal centromere function in fission yeast. Chromosome Res *11*, 137–146.

Wagschal, A., Rousset, E., Basavarajaiah, P., Contreras, X., Harwig, A., Laurent-Chabalier, S., Nakamura, M., Chen, X., Zhang, K., Meziane, O., et al. (2012). Microprocessor, Setx, Xrn2, and Rrp6 Co-operate to Induce Premature Termination of Transcription by RNAPII. Cell *150*, 1147–1157.

Wahba, L., Amon, J.D., Koshland, D., and Vuica-Ross, M. (2011). RNase H and multiple RNA biogenesis factors cooperate to prevent RNA:DNA hybrids from generating genome instability. Mol Cell *44*, 978–988.

Wahba, L., Gore, S.K., and Koshland, D. (2013). The homologous recombination machinery modulates the formation of RNA-DNA hybrids and associated chromosome instability. Elife *2*, e00505.

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

Wang, S.H., and Elgin, S.C.R. (2011). Drosophila Piwi functions downstream of piRNA production mediating a chromatin-based transposon silencing mechanism in female germ line. Proc Natl Acad Sci USA *108*, 21164–21169.

Wassenegger, M., Heimes, S., Riedel, L., and Sänger, H.-L. (1994). RNA-directed de novo methylation of genomic sequences in plants. Cell *76*, 567–576.

Watanabe, T., Tomizawa, S.I., Mitsuya, K., Totoki, Y., Yamamoto, Y., Kuramochi-Miyagawa, S., Iida, N., Hoki, Y., Murphy, P.J., Toyoda, A., et al. (2011). Role for piRNAs and Noncoding RNA in de Novo DNA Methylation of the Imprinted Mouse Rasgrf1 Locus. Science *332*, 848–852.

Wei, W., Ba, Z., Gao, M., Wu, Y., Ma, Y., Amiard, S., White, C.I., Danielsen, J.M.R., Yang, Y.-G., and Qi, Y. (2012). A Role for Small RNAs in DNA Double-Strand Break Repair. Cell *149*, 101–112.

Weinberg, D.E., Nakanishi, K., Patel, D.J., and Bartel, D.P. (2011). The inside-out mechanism of Dicers from budding yeasts. Cell *146*, 262–276.

West, S., and Proudfoot, N.J. (2009). Transcriptional termination enhances protein expression in human cells. Mol Cell *33*, 354–364.

West, S., Gromak, N., and Proudfoot, N.J. (2004). Human 5“ --> 3” exonuclease Xrn2 promotes transcription termination at co-transcriptional cleavage sites. Nature *432*, 522–525.

Wierzbicki, A.T., Haag, J.R., and Pikaard, C.S. (2008). Noncoding transcription by RNA polymerase Pol IVb/Pol V mediates transcriptional silencing of overlapping and adjacent genes. Cell *135*, 635–648.

Wierzbicki, A.T., Ream, T.S., Haag, J.R., and Pikaard, C.S. (2009). RNA polymerase V transcription guides ARGONAUTE4 to chromatin. Nat Genet *41*, 630–634.

Wong, L.H., and Choo, K.H.A. (2004). Evolutionary dynamics of transposable elements at the centromere. Trends Genet *20*, 611–616.

Woolcock, K.J., Stunnenberg, R., Gaidatzis, D., Hotz, H.R., Emmerth, S., Barraud, P., and Buhler, M. (2012). RNAi keeps Atf1-bound stress response genes in check at nuclear pores. Genes Dev.

Woolcock, K.J., Gaidatzis, D., Punga, T., and Bühler, M. (2011). Dicer associates with chromatin to repress genome activity in Schizosaccharomyces pombe. Nat Struct Mol Biol *18*, 94–99.

Wu, L., Zhou, H., Zhang, Q., Zhang, J., Ni, F., Liu, C., and Qi, Y. (2010). DNA methylation mediated by a microRNA pathway. Mol Cell *38*, 465–475.

Yamanaka, S., Mehta, S., Reyes-Turcu, F.E., Zhuang, F., Fuchs, R.T., Rong, Y., Robb, G.B., and Grewal, S.I.S. (2012). RNAi triggered by specialized machinery silences developmental genes and retrotransposons. Nature –.

Yang, X.J., and Roberts, J.W. (1989). Gene Q antiterminator proteins of Escherichia coli phages 82 and lambda suppress pausing by RNA polymerase at a rho-dependent terminator and at other sites. Proc Natl Acad Sci USA *86*, 5301–5305.

Ye, R., Wang, W., Iki, T., Liu, C., Wu, Y., Ishikawa, M., Zhou, X., and Qi, Y. (2012). Cytoplasmic Assembly and Selective Nuclear Import of Arabidopsis ARGONAUTE4/siRNA Complexes. Mol Cell 1–12.

Yu, R., Jih, G., Iglesias, N., and Moazed, D. (2013). Determinants of Heterochromatic siRNA Biogenesis and Function. Mol Cell.

Zaratiegui, M., Castel, S.E., Irvine, D.V., Kloc, A., Ren, J., Li, F., de Castro, E., Marín, L., Chang, A.-Y., Goto, D., et al. (2011). RNAi promotes heterochromatic silencing through replication-coupled release of RNA Pol II. Nature *479*, 135–138.

Zaratiegui, M., Vaughn, M.W., Irvine, D.V., Goto, D., Watt, S., Bähler, J., Arcangioli, B., and Martienssen, R.A. (2010). CENP-B preserves genome integrity at replication forks paused by retrotransposon LTR. Nature.

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

Zhang, L., Beaucher, M., Cheng, Y., and Rong, Y.S. (2014). Coordination of transposon expression with DNA replication in the targeting of telomeric retrotransposons in Drosophila. Embo J.

Zhang, Z., Chang, S.-S., Zhang, Z., Xue, Z., Zhang, H., Li, S., and Liu, Y. (2013). Homologous recombination as a mechanism to recognize repetitive DNA sequences in an RNAi pathway. Genes Dev *27*, 145–150.

Zheng, B., Wang, Z., Li, S., Yu, B., Liu, J.-Y., and Chen, X. (2009). Intergenic transcription by RNA polymerase II coordinates Pol IV and Pol V in siRNA-directed transcriptional gene silencing in Arabidopsis. Genes Dev *23*, 2850–2860.

Zheng, X., Zhu, J., Kapoor, A., and Zhu, J.-K. (2007). Role of Arabidopsis AGO6 in siRNA accumulation, DNA methylation and transcriptional gene silencing. Embo J *26*, 1691–1701.

Zhou, Z.-X., Zhang, M.-J., Peng, X., Takayama, Y., Xu, X.-Y., Huang, L.-Z., and Du, L.-L. (2013). Mapping genomic hotspots of DNA damage by a single-strand-DNA-compatible and strand-specific ChIP-seq method. Genome Research *23*, 705–715.

Zilberman, D., Cao, X., and Jacobsen, S.E. (2003). ARGONAUTE4 control of locus-specific siRNA accumulation and DNA and histone methylation. Science *299*, 716–719.

Zofall, M., Fischer, T., Zhang, K., Zhou, M., Cui, B., Veenstra, T.D., and Grewal, S.I.S. (2009). Histone H2A.Z cooperates with RNAi and heterochromatin factors to suppress antisense RNAs. Nature *461*, 419–422.

7. Acknowledgements

I thank M. O'Connell, and B. Arcangioli for strains, advice, discussion, and hosting me in their labs, and M.T.A. Donoghue for discussion and advice relating to statistical analysis. I acknowledge financial support by a fellowship from the Natural Sciences and Engineering Research Council of Canada (PGSD) and a Cashin Scholarship from the Watson School of Biological Sciences.

Thank you to my advisor Rob Martienseen for providing me both with the freedom to work independently, and a plethora of ideas to work on. I would also like to thank the Martienssen Lab members, past and present, who provided me with scientific help through countless conversations, and moral support when I needed it most. I'd especially like to thank those lab members who became my close friends.

Outside of the lab, at CSHL I'd like to thank members of my class year (WSBS 2009). Without your friendship I don't think I would have made it through the PhD gauntlet. Thank you to the Watson School of Biological Sciences for providing me with support throughout my PhD.

Thanks to my family who supported my decision to go to graduate school, specifically my dad, whose advice to "do what you love" has stuck with me (for better or for worse), and my step mother, who convinced me to come back to graduate school and finish my PhD.

Most importantly, thank you Blackford Bar, Wine and Cheese, and Cocktails.