Local and Genomic Determinants of siRNA-Mediated Heterochromatin Formation and Silencing in Fission Yeast

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters

<table>
<thead>
<tr>
<th>Citation</th>
<th>Yu, Ruby. 2016. Local and Genomic Determinants of siRNA-Mediated Heterochromatin Formation and Silencing in Fission Yeast. Doctoral dissertation, Harvard University, Graduate School of Arts &amp; Sciences.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Citable link</td>
<td><a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:33493406">http://nrs.harvard.edu/urn-3:HUL.InstRepos:33493406</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA">http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA</a></td>
</tr>
</tbody>
</table>
Local and genomic determinants of siRNA-mediated heterochromatin formation and silencing in fission yeast

A dissertation presented
by
Ruby Yu
to
The Division of Medical Sciences
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy
in the subject of
Biological and Biomedical Sciences

Harvard University
Cambridge, Massachusetts
January 2016
Local and genomic determinants of siRNA-mediated heterochromatin formation and silencing in fission yeast

Abstract

In eukaryotes, 20-30 nt small RNAs (sRNAs) regulate many cellular processes by a process called RNA interference (RNAi). sRNAs serve as adaptable specificity factors to direct silencing at either the post-transcriptional or transcriptional level. On one hand, the ability to target genome-wide is a valuable feature; however, this can potentially be a double-edged sword, as spurious expression of sRNAs could lead to inappropriate silencing. Fission yeast have developed mechanisms to combat inappropriate sRNA-mediated silencing. In fission yeast, small interfering siRNAs (siRNAs) and silenced heterochromatin are mutually dependent at pericentromeric repeat sequences, and are required for proper centromere function. Euchromatic regions of the genome typically do not generate siRNAs and are refractory to RNAi-mediated silencing and heterochromatin formation. This dissertation aims to identify mechanisms by which the cell protects itself against inappropriate RNAi-induced silencing or heterochromatin formation. Specifically, we seek to define 1) factors that determine whether a locus can produce siRNAs and 2) factors that determine whether a locus can be targeted by siRNAs for silencing or heterochromatin formation.

Using high throughput sequencing of small RNAs in cells overexpressing Dicer, I show that sites of overlapping transcription including convergent genes and centromeric repeats are potential substrates for Dicer activity, but at endogenous levels of Dicer only centromeric repeats generate siRNAs. RT-PCR and ChIP-seq experiments reveal that, at euchromatic loci, generation of siRNAs in cis does not correlate with reduced transcript levels or with methylation
of H3K9, a conserved marker for heterochromatin. Thus, there are features of euchromatic loci that protect them from siRNA-mediated silencing.

Genetic studies involving deletion or mutation of the 3’ UTR of an endogenous reporter gene in cells expressing complementary siRNAs identify transcriptional cleavage and polyadenylation signals as one of these protective factors. Direct sequencing of polyadenylated transcripts reveals a divergence in cleavage patterns between centromeric and mRNA-coding loci, supporting the idea that the ability of a nascent transcript to be targeted for silencing by siRNAs correlates with inefficient 3’ end processing.

Finally, I show that Mlo3, a gene associated with 3’ ends of ORFs and involved in mRNA export, antagonizes siRNA-mediated heterochromatin formation genome-wide. ChIP experiments and growth assays show that mlo3Δ cells are capable of generating siRNA-mediated H3K9 dimethylation at a locus in trans, but surprisingly, this rarely correlates with silencing of the targeted locus. Enrichment in H3K9 dimethylation that is not associated with silencing is not heritable through meiosis. However, once siRNA-induced silencing is established, heterochromatin is stable and transmissible through meiosis even in the absence of the driver siRNA-producing locus. Silencing is better correlated with H3K9me3 than H3K9me2, and we propose that H3K9me3 is required for formation of a silenced and heritable heterochromatic state.

Together, these results demonstrate that the cell has evolved several mechanisms to protect itself from spurious siRNA-mediated silencing or heterochromatin formation. First, limiting concentrations of Dicer restrict siRNA generation to repetitive sequences surrounding the centromeres. However, if siRNAs mapping to mRNAs are expressed, factors involved in efficient 3’ end processing antagonize silencing and H3K9 methylation directed by siRNAs. These factors include transcript cleavage and polyadenylation sequences and mRNA export factor mlo3+.
Table of Contents

Title Page ..........................................................................................................................i
Copyright Page ................................................................................................................ii
Abstract ...........................................................................................................................iii
Table of Contents ..............................................................................................................v
List of Figures and Tables ..................................................................................................viii
Acknowledgements .........................................................................................................ix

Chapter 1 - Introduction ..................................................................................................1

I. Heterochromatin and epigenetic regulation ..................................................................3

II. RNA interference .......................................................................................................7
   A. Core RNAi protein components .............................................................................8
      i. Argonautes ...........................................................................................................8
      ii. Dicer and Drosha ...............................................................................................9
   B. Classes of small RNAs ..........................................................................................10
      i. miRNA ...............................................................................................................11
      ii. siRNA ...............................................................................................................12
      iii. piRNA .............................................................................................................13

III. RNAi-mediated heterochromatin formation in S. pombe .........................................15
   A. Core complexes involved in the nucleation of RNAi-mediated heterochromatin ...16
      i. RITS, a nuclear RNAi effector complex ............................................................16
      ii. RDRC, an RNA-directed RNA polymerase complex ........................................17
      iii. CLRC mediates methylation of H3K9 ................................................................18
      iv. The ARC complex and siRNA loading onto Ago1 ............................................20
      v. Summary of the nascent transcript model .........................................................22
   B. Transcriptional and co-transcriptional gene silencing .........................................23
      i. Role of RNA Polymerase II ..............................................................................23
      ii. Role of splicing factors .....................................................................................24
      iii. HP1 and histone deacetylases (HDACs) mediate transcriptional gene silencing and co-transcriptional gene silencing .................................................................25
      iv. TRAMP and the exosome contribute to co-transcriptional gene silencing ..........27
   C. Biogenesis of early siRNAs and establishment of centromeric heterochromatin ......28
      i. Dcr1-independent primal RNAs (priRNAs) .........................................................30
      ii. Dicer-dependent generation of siRNAs from intramolecular hairpins ..............31
      iii. Dicer-based generation of siRNAs from overlapping transcripts ....................32
   D. Heterochromatin-inhibiting factors and the role of histone turnover ....................33
      i. Epe1, a putative demethylase ..........................................................................35
      ii. Mst2, a histone acetyltransferase .....................................................................37
      iii. Mlo3 and Cid14: transcript processing factors ................................................38
      iv. Paf1, a transcription elongation and 3' processing complex ..............................39

IV. Small RNA silencing of transposable elements in multicellular organisms ............41
   A. Arabidopsis thaliana .............................................................................................42
   B. Caenorhabditis elegans .........................................................................................43
   C. Drosophila melanogaster .......................................................................................44
   D. Mus musculus ......................................................................................................45

V. Regulation of gene expression and chromatin state by long non-coding RNA ..........46
Chapter 2 - Overexpression of Dcr1 reveals genome-wide overlapping transcription and bypasses the requirement for Rdp1 in RNAi-mediated silencing and heterochromatin... 65

Abstract ................................................................................................................. 67
Introduction ........................................................................................................... 68
Results ...................................................................................................................... 70
  Detection of Rdp1-independent primary siRNAs .................................................. 70
  Dicer overexpression boosts the levels of primary siRNAs and restores silencing in the absence of RDRC ................................................................. 72
  Dicer overexpression results in genome-wide generation of primary siRNAs from convergent or overlapping transcription units ........................................ 82
Discussion ............................................................................................................... 83
  Biogenesis of primary siRNA .............................................................................. 83
  Catalytic and structural functions of Dicer in centromeric silencing ...................... 87
  Silencing of euchromatic loci by small RNAs ...................................................... 87
References ............................................................................................................ 89

Chapter 3 - Transcriptional cleavage and polyadenylation signals inhibit silencing by RNAi and ectopic siRNAs ................................................................. 92

Abstract ............................................................................................................... 94
Introduction .......................................................................................................... 95
Results .................................................................................................................... 97
  Presence of mRNA 3’ processing signals inhibits the ability of hairpin siRNAs to promote silencing and secondary siRNA synthesis .............................................. 97
  A transcript lacking 3’ processing signals is targeted by RNAi ............................. 109
Discussion ............................................................................................................ 114
  Inhibition of siRNA-mediated silencing by 3’ processing signals ....................... 114
  Dcr1 directly targets the ura4-UTR^D locus for silencing in the absence of ectopic siRNAs .................................................. 116
  Intron-derived siRNAs are processed and can functionally target an intron-containing transcript ................................................................. 117
References ......................................................................................................... 119

Chapter 4 - Deletion of mRNA export factor mlo3^+ permits siRNA-mediated H3K9 dimethylation but not silencing ......................................................... 122

Abstract .............................................................................................................. 124
Introduction ......................................................................................................... 125
Results ................................................................................................................... 127
  Deletion of a factor linking 3’ end processing and RNA export permits high accumulation of ectopic siRNA-directed H3K9me2 at reporter genes ....................... 127
  Trans-acting siRNAs can induce ectopic establishment of meiotically heritable silenced heterochromatin ............................................................. 138
  Cells lacking Paf1 component leo1^+ induce siRNA-mediated silencing that correlates with lower H3K9me2 than in non-silencing mlo3^Δ cells ............................. 143
Discussion ......................................................................................................... 145

VI. Topics to be addressed in this thesis ................................................................ 49
VII. References ..................................................................................................... 52
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generation of high levels of H3K9 dimethylation is not sufficient for transcriptional gene silencing</td>
<td>146</td>
</tr>
<tr>
<td>References</td>
<td>150</td>
</tr>
<tr>
<td>Chapter 5 - Perspectives</td>
<td>154</td>
</tr>
<tr>
<td>Safeguarding gene expression: balancing silencing and safety mechanisms</td>
<td>155</td>
</tr>
<tr>
<td>Mechanisms by which factors antagonize siRNA-mediated targeting</td>
<td>155</td>
</tr>
<tr>
<td>Mechanisms that govern the differential requirements for siRNA-mediated H3K9me2 and H3K9me3</td>
<td>159</td>
</tr>
<tr>
<td>References</td>
<td>163</td>
</tr>
<tr>
<td>Appendix 1. Experimental Procedures</td>
<td>165</td>
</tr>
<tr>
<td>Appendix 2. Extended protocol for preparation of small RNA libraries</td>
<td>173</td>
</tr>
<tr>
<td>Appendix 3. Lists of strains, plasmids, and oligos used</td>
<td>180</td>
</tr>
</tbody>
</table>
List of Figures and Tables

Chapter 1
Figure 1.1: The nascent transcript model.................................................................22

Chapter 2
Figure 2.1: Fission yeast pericentromeric repeats produce primary siRNAs that increase in abundance with overexpression of Dcr1.................................................................71
Figure 2.2 Overexpression of Dcr1 partially restores centromeric silencing in the absence of the RNA-dependent RNA Polymerase Complex (RDRC)..............................................................73
Figure 2.3 Dcr1 plays a structural role in promoting H3K9 methylation independently of its catalytic activity ..............................................................................................................75
Figure 2.4. Dcr1 physically interacts with CLRC components .............................................78
Figure 2.5 The effect of overexpression of RNAi factors on silencing and pericentromeric H3K9 methylation ..................................................................................................................79
Figure 2.6. Dcr1 overexpression results in the generation of sense and antisense primary siRNAs at a subset of euchromatic transcription units ..........................................................80
Figure 2.7 Pathways that mediate siRNA generation and control siRNA-mediated heterochromatin formation ..................................................................................................................85

Chapter 3
Figure 3.1 The *ura4* 3' UTR inhibits hairpin siRNA-induced silencing..........................99
Figure 3.2 The *ura4* 3' UTR inhibits hairpin siRNA-induced heterochromatin formation ........101
Figure 3.3 The polyadenylation and cleavage signals in the *ura4* 3' UTR inhibit hairpin siRNA-induced silencing ..................................................................................................................102
Figure 3.4 The 3' UTR of *ura4* inhibits the generation of secondary siRNAs ......................104
Figure 3.5. Deletion of RDRC, CLRC, or TRAMP subunits affects spreading of secondary siRNAs .....................................................................................................................................107
Figure 3.6. The *ura4*-UTR*Delta* transcript is silenced in an RNAi- and TRAMP-dependent manner ...............................................................................................................................110
Figure 3.7. Overexpression of Dcr1 induces strong hairpin-independent silencing and heterochromatin formation at *ura4*-UTR*Delta* ...........................................................................111
Figure 3.8 RNAi-mediated silencing correlates with low or unusual polyadenylation signals .................................................................................................................................113
Figure 3.9 Control of siRNA biogenesis and function in *S. pombe*.....................................115

Chapter 4
Figure 4.1. Deletion of certain factors involved in transcriptional termination or mRNA export permits *trans* siRNA-mediated H3K9me2 ..............................................................................128
Table 4.1. List of deleted genes and their functions ............................................................131
Figure 4.2. Deletion of mRNA export factor *mlo3* permits *trans* siRNA-mediated H3K9me2 at multiple euchromatic loci. .....................................................................................................133
Figure 4.3. Enrichment of ectopic siRNA-mediated H3K9me2 in *mlo3Delta* cells is lost upon reintroduction of *mlo3* or passage through meiosis ..............................................................................136
Figure 4.4. *Trans* siRNA-mediated silencing in *mlo3Delta* is a rare event ..........................139
Figure 4.5. Deletion of Paf1 complex component *leo1* permits siRNA-mediated silencing but less spreading of H3K9me2 than *mlo3Delta* ..................................................................................144
Acknowledgements

I would like to sincerely thank the members of my dissertation examination committee, Drs. Fred Winston, Ting Wu, Carl Novina, and Richard Young, for their generous time in reading and critiquing this dissertation.

I would like to thank the members of my DAC, Drs. Fred Winston, Stirling Churchman, and Stephen Buratowski for their guidance and mentorship over the years.

I would like to thank the co-authors of my paper, Gloria Jih and Nahid Iglesias, for their hands and minds in collaborating with me.

I would like to thank graduate students past and present, particularly fellow S. pombe students Gloria Jih and Danny Holoch, for their companionship and for assisting me with protocols, literature, and graduate school in general. Additionally, I would like to thank all other past and present members of the Moazed lab for their support, advice, ears, and patience, particularly my baymate Keith and our lab technician Cece.

Thank you to Danesh for building a wonderful lab community, for mentoring me in the lab for 8 years, for always having patience and a positive attitude, and especially for always being excited about science.

I’d like to thank my friends, who have at times encouraged me to defect from science and join the circus.

Last but not least, I’d like to thank my parents, whose PhDs in Statistics and Physics made not having an advanced degree seem foreign. They were tirelessly dedicated to my education and well-being and successfully ensured that I would have opportunities they never dreamed of. Despite my lack of enthusiasm at the time, I now appreciate their heroic patience in personally teaching me extra math every weekend and Chinese every night until I was a teenager.
Chapter 1

Introduction
In all organisms, proper regulation of gene expression is crucial, from determining cell identity to responding to environmental changes. Changes in gene expression can occur at one of several different levels: at the chromatin level, transcription of a locus can be activated or repressed by transcription factors or epigenetic events triggered by histone modifications; at the transcript level, RNAs can be stabilized or degraded, either co-transcriptionally or post-transcriptionally; finally, protein levels can be fine-tuned by mechanisms that regulate translation, protein half-life, or both. Specificity factors- adapters that guide gene silencing or activating machinery to the appropriate target- play an integral role in these processes. Transcription factors, proteins that directly recognize short DNA sequences and mediate up- or down-regulation of transcription, represent one well-studied sub-class of specificity factors. More recently, non-coding RNAs, in the form of long non-coding RNAs (lncRNAs) or small RNAs (sRNAs), constitute a large class of molecules increasingly recognized for their widespread roles in gene regulation. RNAs can respond rapidly to changes in environmental conditions, moderating gene expression at the level of transcription (by transcriptional interference or serving as a scaffold for chromatin association), or by base-pairing with complementary nucleic acid sequences and lending specificity to otherwise non-specific silencing or activating machineries. In the introductory chapter to this dissertation, I will describe the characteristics and mechanisms of RNAs known to mediate gene silencing, focusing on classes that govern nucleation of silent chromatin or heterochromatin. First, I will review the importance of heterochromatin and mechanisms of epigenetic regulation. Second, I will describe RNA interference (RNAi) and the various mechanisms by which sRNAs regulate gene expression. In particular, I will focus on the roles that two classes of sRNAs play in heterochromatin formation and silencing of transposable elements, which is critical for maintenance of genome integrity. Of note is the integral role that lncRNAs play in sRNA-mediated transcriptional gene silencing, by alternately serving as precursors to sRNA production and as RNA scaffolds for sRNA-mediated association of heterochromatin machinery.
In the third section, I will describe in detail what is currently known about RNAi-mediated heterochromatin formation and silencing in the fission yeast *Schizosaccharomyces pombe*. The details of this fission yeast pathway inform models for the role of RNAi in transcriptional silencing of transposable elements in higher eukaryotes, which will be discussed in the fourth section. In the last section, I will briefly discuss Xist, a lncRNA required for heterochromatin formation during X-chromosome inactivation, and the possible mechanisms by which lncRNA can influence heterochromatin formation. I will close with open questions remaining in the field of siRNA-mediated heterochromatin formation in *S. pombe* that motivated the work presented in this dissertation.

I. Heterochromatin and epigenetic regulation

DNA in eukaryotic cells is packaged into a higher-order structure known as chromatin, which can exist in one of two basic states: euchromatin, which is associated with gene expression, and heterochromatin, which is associated with gene silencing. The basic unit of chromatin is the nucleosome, which consists of an octamer of four core histone proteins, H2A, H2B, H3, and H4, around which DNA is wound (Luger et al., 1997). The protruding N-terminal tails of these histones can undergo a plethora of post-translational modifications (including acetylation or methylation) that influence the resultant chromatin state, and the catalogue of different histone modifications is referred to as the “histone code” (Jenuwein and Allis, 2001; Strahl and Allis, 2000). Specific modifications distinguish heterochromatin and euchromatin and recruit reader proteins, ultimately leading to structural changes resulting in a silenced or transcriptionally active state.

Heterochromatin was originally identified under the microscope by its denser staining pattern and condensed structure; it can be further classified into one of two sub-types, constitutive and facultative. Constitutive heterochromatin is maintained at the centromeres and telomeres, which are low in genes and high in repetitive DNA sequences. Formation of
facultative heterochromatin, such as X-chromosome inactivation, is a cell-selected, heritable event resulting in condensation of chromatin and silencing of the genes contained within (Elgin and Grewal, 2003). The cytologically “condensed” appearance of heterochromatin was later observed to correlate with increased resistance to nucleases and meiotic recombination, and with more regularly spaced nucleosomes (Grewal and Elgin, 2007). We now know that heterochromatin is also marked by conserved histone and DNA modifications, including histone hypoacetylation, methylation of histone H3 lysine 9 (H3K9me), and 5-methylcytosine (5mC) DNA methylation. Studies in multiple organisms have shown that these modifications can be self-reinforcing and interdependent, demonstrating that cells employ multiple mechanisms to ensure the integrity of heterochromatin (Elgin and Grewal, 2003; Holoch and Moazed, 2015a).

H3K9 methylation is carried out by the Suv(var)3-9 class of histone methyltransferases (HMTs), which are conserved from fission yeast to humans (Nakayama et al., 2001; Rea et al., 2000; Schotta et al., 2002). H3K9me is recognized and bound by the chromodomain of heterochromatin protein 1 (HP1), which is required for the integrity of heterochromatin. The chromoshadow domain mediates HP1 dimerization and interactions with the H3K9 HMT or other proteins, contributing to spreading of heterochromatin (Canzio et al., 2011; Lachner et al., 2001; Lorentz et al., 1994). Collectively, these histone and DNA modifications, and the proteins that associate with them, contribute to one of the key characteristics of heterochromatin, which is the ability to be epigenetically (independent of DNA sequence) inherited.

Insertion of euchromatic genes within or adjacent to heterochromatin can result in stable phenotypes that are epigenetically inherited. The earliest example of this was documented in 1930 by H.J. Muller, who observed red-white mosaicism of the white\(^*\) gene in Drosophila. This was later attributed to rearrangements in chromosome structure that moved the white\(^*\) gene from its normally euchromatic locus to a locus adjacent to heterochromatin (Wakimoto, 1998). This effect is referred to as position-effect variegation (PEV) and has also been described in budding and fission yeast. In budding yeast, insertion of \textit{ADE2} adjacent to telomeres results in
red- and white- sectored colonies, reflecting genetically identical cells that have and have not silenced \textit{ADE2}, respectively (Gottschling et al., 1990). In fission yeast, insertion of \textit{ade6}+ within centromeric heterochromatin results in white-pink-red sectoring of colonies, reflecting progressive levels of silencing (Allshire et al., 1994). Furthermore, insertion of a 1.6kb fragment of the fission yeast outer centromeric repeat within euchromatin results in local establishment of H3K9me and HP1 association, which encroach upon adjacent genes to induce silencing (Partridge et al., 2002). These examples demonstrate two more important characteristics of heterochromatin: the existence of \textit{cis}-acting elements that can initiate heterochromatin outside of their native environments, and the ability of heterochromatin to spread in a sequence-independent manner into regions that lack those \textit{cis} elements (Moazed, 2009).

Because of the ability of heterochromatin to spread and silence loci in a sequence independent manner, the cell employs multiple mechanisms to restrict the domains of heterochromatin. In one example, deletion of boundary elements flanking the centromeres of fission yeast \textit{S. pombe} results in spreading of heterochromatin into euchromatic regions (Scott et al., 2006). More recently, deletion of two genes that negatively regulate heterochromatin was shown to result in uncontrolled spreading and ectopic formation of heterochromatin, silencing of essential genes, and severe growth defects (Wang et al., 2015). Cells rapidly adapted to these problems by forming facultative, silenced heterochromatin at the genes required for H3K9 methylation, demonstrating the ability of cells to respond to shifting environments by epigenetic adaptations.

Heterochromatin is required for many vital functions of life, including genome integrity and cell differentiation. Disruption of heterochromatin can therefore result in disease and abnormal cell phenotypes. Among the most conserved roles of heterochromatin is maintenance of genome integrity by silencing transposons, parasitic DNA elements that can insert themselves into the genome and wreak havoc (Malone and Hannon, 2009). Indeed, constitutive heterochromatin domains like the centromeres and telomeres of many eukaryotes are poor in
genes but enriched for repetitive DNA sequences, which appear to represent ancient transposable elements that have since lost mobility. Furthermore, Piwi-interacting RNA (piRNA) clusters, which initiate generation of piRNAs resulting in post-transcriptional or heterochromatic silencing of active transposons, are enriched in transposons or transposon fragments, and located either within or adjacent to heterochromatin (Iwasaki et al., 2015). Research into piRNA-mediated transcriptional gene silencing (TGS), governed by formation of silenced heterochromatin, as well as siRNA-mediated heterochromatin in fission yeast and plants, has challenged the view that heterochromatin is transcriptionally inert. In fact, both of these systems require the generation of transcripts in order to establish and maintain silencing of the locus—however, the transcripts are degraded and fail to accumulate (discussed in detail in following sections). Centromeric heterochromatin plays another fundamental role in genome stability by mediating sister chromatid cohesion during cell division, and disruption of centromeric heterochromatin results in chromosome segregation defects (Bernard et al., 2001). In development, progression of cells to a more differentiated state correlates with accumulation of histone modifications associated with heterochromatin and decline of histone modifications associated with active transcription (Bhaumik et al., 2007). Finally, misregulation of histone modifications has been associated with many diseases, including cancer. In mice, deletion of Suv39 H3K9 methyltransferases correlates with onset of B-cell lymphomas and other cancers (Bhaumik et al., 2007). These examples all underscore the critical role that heterochromatin plays in maintenance of cellular integrity.

In the remainder of this introduction, I will expand on mechanisms by which long or short RNAs regulate gene expression and chromatin state. First, in Section II, I will introduce RNA interference and the mechanisms by which small RNAs (sRNAs) mediate gene silencing. While most studies of sRNAs have focused on their role in post-transcriptional gene silencing, extensive studies in fungi and plants, and more recent studies in worms, flies, and mammals, have implicated sRNAs in a conserved nuclear pathway associated with heterochromatin
formation and transcriptional gene silencing. Strikingly, the main targets of nuclear sRNA pathways are transposable elements, selfish genetic elements that can mobilize to insert themselves throughout the genome and potentially cause catastrophic disruptions. A critical requirement for nuclear RNAi then is to accurately distinguish “native” sequences from “foreign” transposon sequences, the latter of which RNAi subsequently inactivates. In Chapters 2 through 4 of this dissertation, I will present largely published data regarding several redundant mechanisms by which this is achieved in fission yeast. First, in Section III of this introduction, I will discuss in detail what is known regarding the mechanism of RNAi-mediated heterochromatin formation and silencing in fission yeast. The well-studied fission yeast pathway serves as a model for nuclear mechanisms of RNAi that govern germline development and fertility in higher eukaryotes, which will be discussed in Section IV. Notably, a common feature of nuclear RNAi pathways is their reliance on long non-coding nascent transcripts, which serve as scaffolds for chromatin interactions resulting in formation of heterochromatin. In the final section of this introduction, I will discuss the various mechanisms by which long non-coding RNAs could potentially contribute to chromatin-based, RNAi-independent mechanisms of gene silencing, an emerging field in which much remains to be discovered.

II. RNA interference

RNA interference (RNAi) was originally described in the nematode C. elegans (Fire et al., 1998). In the course of investigating the hypothesis that antisense RNAs induce silencing of an endogenous gene, Fire, Mello, and colleagues discovered that double-stranded RNA (dsRNA) induces much more powerful silencing than single-stranded RNA (ssRNA). Injection of only several molecules of dsRNA was sufficient, implying an amplification mechanism that was later explained by the action of RNA-dependent RNA polymerases (RdRPs) (Fire et al., 1998; Sijen et al., 2001). Members of the conserved Argonaute family of proteins were implicated in this silencing mechanism (Catalanotto et al., 2000; Tabara et al., 1999) and later identified as
the core protein component of the RNA-induced silencing complex (RISC), which uses a single-stranded small RNA (sRNA) guide to target complementary transcripts for post-transcriptional gene silencing (PTGS) (Hammond et al., 2000; Hammond et al., 2001). RNaseIII-containing Dicer proteins were identified as the ribonucleases responsible for processing long dsRNA into duplex sRNAs (Bernstein et al., 2001). Dissection of sRNA biogenesis and mechanisms of silencing have revealed several important classes, including small interfering RNAs (siRNAs), microRNAs (miRNAs), and PIWI-interacting RNAs (piRNAs) (Moazed, 2009). In addition to roles in PTGS, two classes of sRNAs, siRNAs and piRNAs, have also been found to act in the nucleus, where they regulate gene expression at the level of chromatin. Below, I will describe each component of RNAi in more detail.

A. Core RNAi protein components

i. Argonautes

Argonaute family proteins are common to all RNAi pathways, and, together with the sRNA specificity factors they bind, constitute the core component of RNAi effector complexes like RISC. They are divided mainly into two clades, the AGO clade, which associates with ~21-25nt miRNAs and siRNAs, and the PIWI clade, which associates with ~24-31nt piRNAs (Carmell et al., 2002; Moazed, 2009). Many organisms have multiple Argonaute proteins, each associating with a specific set of sRNAs and mediating activity through specialized mechanisms (Ipsaro and Joshua-Tor, 2015). Argonautes contain 3 conserved domains, the PAZ, the MID, and the PIWI domains. The MID domain interacts with the 5’ phosphate of sRNAs, with contributions by the PIWI domain. Interactions between the nucleotide specificity loop of the MID domain and the first nucleotide of the sRNA tend to favor certain bases over the others, resulting in a 5’-nucleotide bias in sRNAs associated with a particular Argonaute, and contributing to appropriate sorting of sRNAs in organisms with multiple sRNA silencing pathways (Frank et al., 2012). The PAZ domain of Argonaute interacts with the 3’ end of
sRNAs, and arrangement of the PAZ domain relative to the MID domain also contributes to sRNA-binding specificity (Holoch and Moazed, 2015a; Ipsaro and Joshua-Tor, 2015). The PIWI domain of catalytically active Argonautes contains an RNaseH-like fold that is responsible for ‘slicer’ activity (Jinek and Doudna, 2009). Upon association of sRNA-bound Argonaute with a target mRNA, the endonuclease domain of catalytically active Argonautes can mediate cleavage (‘slicing’) of the target within the region of complementarity with the sRNA (Zamore et al., 2000), leading to subsequent transcript degradation. Slicing requires perfect complementarity between the sRNA and target, and thus is mainly mediated by siRNAs rather than miRNAs (discussed below). Slicer activity can also be required for ejection of the non-guide (passenger) strand of duplex siRNAs (Matranga et al., 2005; Miyoshi et al., 2005; Rand et al., 2005). Non-slicing activities of Argonaute involve association with GW182 proteins, which recruit downstream silencing proteins including poly(A)-binding protein and deadenylase complexes PAN2-PAN3 and CCR4-NOT (Jonas and Izaurralde, 2015).

**ii. Dicer and Drosha**

Dicer is responsible for processing of longer dsRNA into mature double-stranded siRNAs and, in combination with Drosha, miRNAs. Duplex Dicer products are ~21-25nt in length and characterized by 5’ monophosphates and 3’ 2-nt overhangs. Dicer contains two RNaseIII endonuclease domains that act as an intramolecular dimer and are each responsible for cleavage of one strand of the dsRNA. In addition, Dicer possesses a C-terminal dsRNA-binding domain (dsRBD), a PAZ domain sharing the same structure and 3’-overhang-binding residues as Argonaute PAZ domains, a DUF283 domain of unknown function, and in some organisms an N-terminal DEXD/H-box domain. Nonspecific cleavage generates the 3’ dinucleotide overhang of the substrate dsRNA, which associates with the PAZ domain. Following this, the RNaseIII domains mediate cleavage of the 5’ end at a fixed distance that is determined by the spacing
between the PAZ-proximal RNaseIII domain and the 3’ overhang-binding pocket of the PAZ domain (Ipsaro and Joshua-Tor, 2015; Jinek and Doudna, 2009).

Drosha is a nuclear RNaseIII-family protein that is required for processing of primary precursors of miRNAs (pri-miRNAs) into precursor miRNA (pre-miRNA), the first step of miRNA biogenesis. Drosha products are hairpins with phosphorylated 5’ ends and 3’ dinucleotide overhangs. Drosha consists of an N-terminal proline rich region, an arginine- and serine- rich region, two RNaseIII domains, and a dsRBD. Drosha itself appears to cleave nonspecifically, but associates with a trans-acting specificity factor that recognizes appropriate substrates (DGCR8 or, in invertebrates, PASHA). Subsequently, Drosha trims pri-miRNAs to ~60-70nt hairpins, which become substrates for Dicer. Together, Drosha and DGCR8/PASHA are known as the microprocessor (Ameres and Zamore, 2013; Ipsaro and Joshua-Tor, 2015; Jinek and Doudna, 2009).

B. Classes of small RNAs

The earliest evidence that RNA-mediated silencing was associated with small RNAs came from plants, where studies of multiple systems of post-transcriptional gene silencing (PTGS) consistently detected 25-nt RNA molecules that were antisense to the target mRNA (Hamilton and Baulcombe, 1999). 25-nt sRNAs were also found to be associated with RISC in D. melanogaster extracts (Hammond et al., 2000). Further work in D. melanogaster extracts showed that during RNAi, dsRNA is processed into 21-23 nt sRNAs, and that mRNAs are only cleaved in 21-23nt intervals at regions of complementarity with the dsRNA (Zamore et al., 2000). We now know that sRNAs serve as the specificity signals for RNAi-mediated silencing, and that there exist multiple classes of sRNAs. Below, I describe the roles and biogenesis of three of the most important classes of sRNAs, which are classified based on their mechanisms of generation and action: microRNAs (miRNAs), small interfering RNAs (siRNAs), and PIWI-interacting RNAs (piRNAs).
i. miRNA

microRNAs are ~21-25 nt sRNAs conserved in plants and animals that generally regulate gene expression post-transcriptionally. The first miRNA-coding locus and its target were discovered and characterized in *C. elegans* five years before the discovery of RNAi (Lee et al., 1993; Wightman et al., 1993). The precursors to miRNAs are endogenous hairpin RNAs, called primary precursors of miRNAs (pri-miRNAs). Pri-miRNAs are transcribed by RNA Polymerase II (Pol II) from either independent pri-miRNA genes or from introns of protein-coding genes: several miRNAs can be processed from one transcript (Ameres and Zamore, 2013). These pri-miRNAs are cleaved by RNaseIII-containing proteins, either Dicer alone or Dicer and Drosha. In plants, Dicer-like protein 1 (DCL-1) produces miRNA duplexes in the nucleus (Siomi and Siomi, 2009). In animals, miRNA production is a two-step process, taking place in both the nucleus and the cytoplasm. Drosha and a dsRNA binding partner (Pasha in flies and DGCR8 in mammals) are located in the nucleus, and convert one pri-miRNA into one or more ~60-70nt stem-loop pre-miRNAs. Pre-miRNAs are then exported to the cytoplasm, where Dicer mediates cleavage resulting in ~22nt miRNA duplexes, in combination with one or more dsRNA binding proteins that govern binding affinity, specificity, and cleavage site (Ameres and Zamore, 2013). In animals, miRNAs typically form imperfect duplexes with their targets, with the most important region of specificity being nucleotides 2-8 in the 5’ end of the miRNA (termed the ‘seed region.’) Complementarity to the ‘seed region’ is sufficient to direct Argonaute binding to a target, resulting in decreases of mRNA expression particularly when directed to the 3’ UTR of an mRNA. Because the seed sequence is so short, it is estimated that half or more than half of all miRNAs in mammals are regulated by miRNAs (Ameres and Zamore, 2013). In plants, where miRNAs can exhibit high complementarity to their target, miRNA-bound Argonautes have been shown to mediate slicing of the target mRNA, resulting in transcript degradation. However, miRNA-directed cleavage appears to be rare in mammals, due to limited regions of perfect
complementarity. miRNAs have also been shown to mediate translational repression and non-slicer dependent mRNA degradation. The mechanism of miRNA-mediated translational repression is still an active area of research but seems to occur at the step of translational initiation. To mediate transcript degradation, miRNA-bound Argonaute and its GW182 protein partner associate with deadenylases PAN2-PAN3 and CCR4-NOT, resulting in deadenylation of the target mRNA and subsequent decapping and 5’-3’ degradation by an exonuclease. Recent studies suggest that translational repression and transcript degradation may be coupled processes, and this coupling may involve the ability of deadenylase complex CCR4-NOT to contribute to translational repression independently of deadenylation (reviewed in Jonas and Izaurralde, 2015).

ii. siRNA

siRNAs are also ~21-25nt each, with 5’ monophosphate and 3’ OH groups. siRNAs are cleaved from long endogenous or exogenous dsRNA by Dicer to generate duplex siRNAs with 2-nt 3’ overhangs. Duplex siRNAs are loaded onto Argonaute, where one strand is discarded and the single stranded guide siRNA is retained. Argonaute and its associated effector complex are directed by the guide siRNA to a target transcript and mediate either post-transcriptional gene silencing (PTGS) or transcriptional gene silencing (TGS). In classical siRNA-mediated PTGS, Argonaute cleaves the target RNA at a site of complementarity that corresponds to the linkage between the 10th and 11th nucleotide of the siRNA, triggering subsequent degradation. In TGS pathways in S. pombe, A. thaliana, or C. elegans, siRNAs mediate DNA- or histone modifications resulting in formation of transcriptionally silenced heterochromatin (Holoch and Moazed, 2015a). In fungi, plants, and C. elegans, siRNA populations can be amplified by the activity of an RNA-dependent RNA polymerase (RdRP) recruited to a transcript bound by Argonaute (Dalmay et al., 2000; Sijen et al., 2001). In fungi and plants, RdRP uses the target RNA as a template to generate long dsRNA that can be cleaved by Dicer to generate more
siRNAs (Holoch and Moazed, 2015a; Motamedi et al., 2004). In C. elegans, RdRP directly generates secondary siRNAs in a Dicer-independent manner, using the target mRNA as a template; indicative of this, secondary siRNAs have 5’ triphosphates, rather than monophosphates, and are always antisense to the target RNA (Sijen et al., 2007). However, secondary siRNAs in C. elegans are not capable of contributing to a feed-forward loop that would result in uncontrolled production of siRNAs (Pak et al., 2012). Current models propose that primary siRNAs in C. elegans are responsible for amplifying the siRNA signal by directing generation of secondary siRNAs, and secondary siRNAs in turn are responsible for mediating silencing and RNAi activity, but not further siRNA amplification (Grishok, 2013).

iii. piRNA

piRNAs are only conserved in animals and bear marked differences from siRNAs and miRNAs. First, they are not Dicer products, being processed from single-stranded precursors. Because of this, piRNAs are on average larger than the aforementioned classes of sRNAs, with a wider distribution of sizes (~24-31 nt). piRNAs are also 2’-O-methylated at their 3’ ends. Second, piRNAs exclusively associate with the PIWI clade of Argonaute proteins, for which they are named. Last, piRNAs seem to operate solely in silencing of transposable elements, particularly in germline cells, and are thus integral for genome maintenance and reproduction (Iwasaki et al., 2015).

Biogenesis of piRNAs occurs in two steps, and is still a subject of intense research. Most of our information on piRNA biogenesis comes from Drosophila. Primary piRNA biogenesis generates primary piRNAs from long noncoding RNAs that are transcribed from one or more loci, called piRNA cluster loci. These primary piRNAs are mostly oriented antisense to transposon RNAs and are strongly enriched in 5’ uridines. The primary piRNA pathway is followed by the ‘ping pong’ cycle, which allows amplification of piRNA populations, resulting in secondary piRNAs that are complementary in their 5’ ends to the 5’-most 10 nucleotides of
primary piRNAs, and exhibit a bias for adenosine at nucleotide 10 (Brennecke et al., 2007; Gunawardane et al., 2007). piRNA clusters are enriched in ancient transposons whose sequences may vary from active transposons, but bear sufficient similarity to allow piRNAs to target active transposons. As previously mentioned, piRNA clusters are located near or within heterochromatin and can be unidirectionally or bidirectionally transcribed, making the mechanism by which primary piRNAs initiate an antisense bias unclear. It is currently thought that piRNA clusters serve as ‘traps’ that exploit the mobility of transposons; insertion of a new transposon into a piRNA cluster results in generation of complementary piRNAs and silencing of active transposon units (Malone and Hannon, 2009). Long piRNA cluster-derived transcripts are exported into the cytoplasm and processed into primary piRNAs by a mechanism that is still unknown. Primary piRNA production requires a number of mitochondria-associated factors, including endonuclease Zucchini (Zuc), Minotaur, and GasZ, but their precise role in processing of cluster transcripts is still unclear. However a model has emerged in outline suggesting that piRNA transcripts are cleaved to generate primary piRNA precursors, which are trimmed by a 3’-5’ exonuclease, 2’-O-methylated at their 3’ ends, and then loaded onto one of two PIWI proteins that associate with primary piRNAs: Aubergine (Aub), which remains cytoplasmic, or Piwi, which is then imported into the nucleus (Iwasaki et al., 2015).

Following primary piRNA production, the ping-pong cycle takes place in the cytoplasm and results in generation of secondary piRNAs. Aub, which is loaded with primary piRNAs containing a 5’ U and oriented antisense to transposons, associates with a sense transposon transcript and slices it between the 10th and 11th nucleotides of the primary piRNA, generating the 5’ end of the secondary piRNA. This secondary piRNA is sense to transposons, favors an A at the 10th nucleotide, and is complementary to the primary piRNA for the 5’-most 10 nucleotides. 3’ ends are cleaved in by an unknown mechanism. Secondary piRNAs are loaded onto cytoplasmic AGO3, the last PIWI protein in Drosophila. AGO3 then associates with transcripts oriented antisense to transposons and mediates cleavage resulting in generation of
more antisense-oriented piRNAs that are then loaded onto Aub. This cycle is known as the ping-pong cycle and accomplishes two important goals: 1) amplification of transposon-mapping piRNAs and in the course of this, 2) consumption and post-transcriptional silencing of transposon-derived transcripts (Iwasaki et al., 2015; Malone and Hannon, 2009). The mechanisms of piRNA-mediated silencing in worms, flies, and mice will be discussed in more detail below, following a discussion of the detailed mechanism of siRNA-mediated transcriptional gene silencing in the fission yeast Schizosaccharomyces pombe.

III. RNAi-mediated heterochromatin formation in S. pombe

Yeast are single-celled eukaryotes that are widely beloved for their genetic tractability and ability to stably maintain a haploid state. In particular, fission yeast Schizosaccharomyces pombe has become an attractive model system in which to study chromatin structure and function. While S. pombe possesses a compact genome with only 3 chromosomes, chromosome structure resembles that of higher eukaryotes, and many factors involved in silencing and heterochromatin formation are conserved between S. pombe and multicellular eukaryotes (Wood et al., 2002). Similar to higher eukaryotes, heterochromatin in fission yeast is associated with H3K9 methylation; there are often only one or two homologs of conserved genes important for heterochromatin formation, including a single copy of the histone H3K9 methyltransferase (clr4+), and two copies of heterochromatin protein HP1 (swi6+ and chp2+); and loci of constitutive heterochromatin (telomeres, the mating type locus, and centromeres) and their boundaries are well-defined (Bailis and Forsburg, 2002). Studies of heterochromatin in fission yeast have focused largely on the centromeres, which are regulated by the RNAi machinery, and characterized by bidirectionally transcribed dg and dh repeat sequences, which may be remnants of ancient transposons (Partridge et al., 2002; Volpe, 2002).
A. Core complexes involved in the nucleation of RNAi-mediated heterochromatin

Sequencing of the fission yeast genome uncovered only one ortholog each of RNAi proteins Argonaute (ago1\(^+\)), Dicer (dcr1\(^+\)), and RNA-dependent RNA polymerase (rdp1\(^+\)) (Wood et al., 2002). Shortly afterward, it was discovered that deletion of any of these components results in accumulation of centromeric transcripts mapping to both strands, loss of centromeric H3K9 methylation, and derepression of silencing at pericentromeric reporter genes (Volpe et al., 2002). Similarly, RNAi machinery was also found to regulate a centromere-homologous sequence (cenH) present at the mating type region, though at this locus redundant pathways render RNAi inessential for maintenance of heterochromatin (Hall et al., 2002). Purification and cloning revealed siRNAs complementary to pericentromeric repeats (Reinhart and Bartel, 2002). Later studies have largely defined the machinery and mechanism by which RNAi regulates heterochromatic silencing, though new insights continue to emerge.

i. RITS, a nuclear RNAi effector complex

Little more than a year after the discovery that RNAi regulates heterochromatin in fission yeast, the RITS (RNA-induced initiation of transcriptional gene silencing) complex, the primary RNAi effector complex involved in heterochromatic silencing, was identified (Verdel et al., 2004). RITS is composed of Ago1, chromodomain-containing protein Chp1, and GW-repeat protein Tas3 (reminiscent of Argonaute associations with GW-repeat proteins in miRNA-mediated silencing) (Verdel et al., 2004). Chp1 had been previously identified as necessary for proper chromosome segregation and silencing at centromeric but not mating type regions, in contrast to HP1 protein Swi6 and H3K9 methyltransferase Clr4, which are required for silencing at both loci (Doe et al., 1998; Thon and Verhein-Hansen, 2000). Subsequent work showed that the chromodomain of Chp1 binds methylated H3K9, and is required for localization of Chp1 to heterochromatin, proper chromosome segregation, and centromeric silencing and H3K9 methylation, but not for association with Ago1 and Tas3 (Partridge et al., 2002; Petrie et al.,
While RITS was found to localize to all heterochromatic loci, it is dispensable for maintenance of silencing at regions other than the centromere (Cam et al., 2005; Noma et al., 2004; Verdel et al., 2004). RITS was found to associate with centromeric siRNAs, and deletion of any of its components results in loss of centromeric silencing and heterochromatin formation (Noma et al., 2004; Sadaie et al., 2004). Interestingly, deletion of RITS and RNAi factors results in a significant, but not complete, loss of centromeric H3K9me, in contrast to complete loss in \(clr4\Delta\) cells (Sadaie et al., 2004). Tethering of RITS to a euchromatic reporter transcript (\(ura4\)-BoxB) via a Tas3-\(\lambda\)N fusion protein (\(\lambda\)N binds RNA BoxB sequences) results in \(ura4\) silencing, H3K9 methylation, and Swi6 association in a manner that requires RNAi and the heterochromatin machinery, and results in generation of \(ura4\) siRNAs (Bühler et al., 2006). Importantly, cells in which Tas3 was tethered to the endogenous BoxB-containing copy of \(ura4\) silenced the BoxB-containing copy of \(ura4\) on chromosome 3, but not a separate and identifiably distinct copy of \(ura4\) inserted on chromosome 2, providing an early example of the refractory nature of euchromatic genes to siRNA-mediated silencing in \textit{trans}.

\textbf{ii. RDRC, an RNA-directed RNA polymerase complex}

Once RITS stabilizes on chromatin, it recruits two other protein complexes. The first of these is the RNA-directed RNA polymerase complex, RDRC. RDRC is composed of conserved proteins RNA-dependent RNA polymerase Rdp1, RNA helicase Hrr1, and a member of the polyA polymerase family, Cid12 (Motamedi et al., 2004). RDRC associates with RITS in a Dcr1- and Clr4-dependent manner. Like RITS, RDRC associates with centromeric transcripts, and deletion of any of its components results in loss of centromeric heterochromatin and silencing, and loss of centromeric RITS localization (Motamedi et al, 2004). Furthermore, RDRC association with heterochromatin is RITS and Clr4-dependent, and cells containing a mutant form of \(rdp1\) lacking RNA-dependent RNA polymerase activity are defective in heterochromatin
formation and transcriptional silencing at centromeres (Sugiyama et al., 2005). Combined, this data supports a model in which RITS and RDRC associate with nascent centromeric transcripts.

RDRC components have been shown to interact with Dcr1 in a manner independent of Clr4 but dependent on RDRC and RITS integrity; this interaction was unaffected by nuclease treatment and did not require the catalytic activity of Dcr1 (Colmenares et al., 2007). In vivo, Rdpl may associate with Dcr1 via interactions with Dsh1, as Dsh1 co-immunoprecipitates with Rdpl and Dcr1, and deletion of dsh1 results in loss of association between Rdpl and Dcr1 as well as between RITS and RDRC (Kawakami et al., 2012). Dcr1 association with Rdpl in vitro stimulates the generation of dsRNA in a manner that does not require Dcr1 cleavage activity, and in vitro association of catalytically active but not catalytically inactive Dcr1 with RDRC results in the generation of siRNAs, indicating that Dcr1 stimulates the catalytic activity of Rdpl and can use its dsRNA product to generate siRNAs. As expected, mutation of the catalytic sites of Dcr1 results in loss of centromeric silencing and siRNAs (Colmenares et al., 2007).

iii. CLRC mediates methylation of H3K9

The second protein complex recruited to chromatin by RITS is the H3K9 methyltransferase-containing complex, termed CLRC (for Clr4-Rik1-Cul4). Clr4, a homologue of D. melanogaster Su(var)3-9 and mammalian Suv39h, possesses a methyltransferase SET domain as well a H3K9me-binding chromodomain, and is the sole H3K9 methyltransferase in fission yeast (Cam et al., 2005; Ivanova et al., 1998; Nakayama et al., 2001; Rea et al., 2000; Zhang et al., 2008). Interestingly, the H3K9me-binding affinity of the Clr4 chromodomain is required for proper localization of Clr4 to heterochromatic regions, and for spreading of CLRC away from heterochromatin nucleation sites (Al-Sady et al., 2013; Zhang et al., 2008). As previously mentioned, the chromodomain of Chp1 associates with H3K9me to stabilize RITS on chromatin (Partridge et al., 2002; Petrie et al., 2005). Related to this observation, deletion of clr4 (the sole H3K9 methyltransferase) results in loss of siRNAs mapping to centromeres
(Motamedi et al., 2004; Noma et al., 2004), though to a lesser extant than in \textit{dcr1}\Delta cells (Motamedi et al., 2004). In contrast, deletion of HP1 protein Swi6, which binds H3K9me and is required for centromeric silencing, has little effect on siRNA generation (Motamedi et al., 2004). Combined with previous data showing that deletion of genes involved in siRNA generation results in loss of H3K9me (Volpe et al., 2002), this data highlights the interdependent relationship between siRNAs and H3K9me.

\textbf{Clr4} associates with WD40 repeat-containing proteins Rik1 and Raf1 (Dos1/Cmc1/Clr8), as well as Raf2 (Dos2/Cmc2/Clr7) and cullin E3 ubiquitin ligase Cul4, and all components are required for heterochromatin formation and proper centromere silencing and function (Hong et al., 2005; Horn et al., 2005; Jia et al., 2005; Li et al., 2005; Nakayama et al., 2001; Thon et al., 2005). In a mechanism that is still not understood, the ubiquitin ligase activity of Cul4 is also required for heterochromatin formation. The ubiquitin ligase activity of cullin proteins requires modification of cullins by ubiquitin-like protein Nedd8 (Pan et al., 2004). A mutant form of Cul4 that cannot be neddylated is defective in centromeric silencing and heterochromatin formation and results in mislocalization of Clr4, and defects in centromeric silencing resulting from defective \textit{cul4} could not be overcome by overexpression of Clr4 (Jia et al., 2005). In contrast, localization of Rik1 to centromeric repeats does not require \textit{cul4}+, suggesting that Rik1 may serve as a targeting platform to recruit the ubiquitination and methyltransferase activity of remaining CLRC components (Jia et al., 2005; Zhang et al., 2008). Interestingly, in cells expressing a mutant form of \textit{cul4} and in cells lacking the Clr3 histone deacetylase (which is required for transcriptional silencing), Rik1 association with heterochromatic repeats correlates with increasing heterochromatic transcript and siRNA expression; this association requires functional RNAi machinery (Zhang et al., 2008).

Rik1 and Clr4 have both been shown to physically interact with RITS subunits in a manner that requires Dcr1 and RITS complex integrity, suggesting that intact RITS is capable of directly recruiting CLRC (Gerace et al., 2010; Zhang et al., 2008). In addition to direct
association between RITS and CLRC, LIM domain protein Stc1 also seems to bridge the RNAi and methyltransferase machineries. Stc1 physically associates with CLRC components in an RNAi-independent manner and physically associates with Ago1 in a CLRC- and Dcr1-dependent manner (Bayne et al., 2010). Together, this data highlights multiple mechanisms by which CLRC can be recruited to chromatin in an siRNA-dependent manner.

iv. The ARC complex and siRNA loading onto Ago1

In addition to its effector role in the RITS complex, Ago1 associates with two other proteins, Arb1 and Arb2, in a functionally distinct complex termed ARC, integrity of which does not require dcr1+ or tas3+ (Buker et al., 2007; Holoch and Moazed, 2015b). Arb1 is a fungally-conserved protein containing a C-terminal domain similar to those found in organellar maturases, which are involved in intron self-splicing. Arb2 is widely conserved among eukaryotes but has no identifiable defined domains (Buker et al., 2007). Both Arb1 and Arb2 are required for centromeric silencing, siRNA generation, and heterochromatin formation, but neither localizes to centromeric regions. Immunofluorescence showed that unlike Chp1, which localizes to a limited number of nuclear foci, Arb1, and Arb2 localize to the cytoplasm and nucleoplasm. In another contrast with RITS, experiments show that Tas3 preferentially associates with single stranded siRNAs, which is consistent with the role of RITS as an effector RNAi complex, whereas Arb1 preferentially associates with double stranded siRNA (Buker et al., 2007; Holoch and Moazed, 2015b). This data, in combination with the data that ARC components do not co-localize with RITS, suggest that ARC contributes to centromeric silencing, heterochromatin formation, and siRNA generation in a manner distinct from that of RITS.

Several lines of evidence have come together to suggest that ARC acts upstream of RITS, overseeing loading of duplex siRNAs onto Ago1. First, while association of Ago1 with Tas3 is largely lost in arb1Δ and arb2Δ cells, deletion of tas3+ does not reduce Ago1-Arb1
association (Holoch and Moazed, 2015b). This suggests that ARC acts upstream of RITS: loss of RITS components has no effect on ARC integrity, but loss of ARC components compromises RITS integrity. Holoch et al hypothesized that ARC acts upstream of RITS by governing the loading of siRNAs onto Ago1 (2015b). In support of this, they showed that Ago1 unassociated with siRNAs can enter ARC but not RITS: mutant forms of Ago1 unable to bind siRNAs failed to associate with Tas3, but retained Arb1 association. Furthermore, two lines of evidence show that ARC but not RITS can associate with duplex siRNAs. First, sRNA-binding assays showed that while ARC and RITS both associate with single-stranded siRNAs, double-stranded siRNAs, whose loading presumably precedes association of single-stranded siRNAs, only associate with ARC. Finally, the ability of Ago1 to associate with duplex siRNAs is lost in \( \text{arb}1\Delta \) but not \( \text{tas}3\Delta \) cells (Holoch and Moazed, 2015b). In summary, Holoch et al showed that ARC integrity is required for RITS integrity but not vice versa; ARC but not RITS can associate with Ago1 unbound to siRNAs; and ARC but not RITS can associate with duplex siRNAs (2015b).

The above evidence strongly supports a role for ARC in loading of siRNAs onto Ago1, upstream of RITS. Further evidence suggests that ARC also inhibits the slicer activity of Ago1, which is required for release of the passenger siRNA. The slicing activity of Argonautes has been implicated in the release of the passenger strand siRNAs in multiple organisms (Matranga et al., 2005; Miyoshi et al., 2005; Rand et al., 2005). Fission yeast Argonaute possesses slicing ability, which is critical for centromeric silencing and siRNA generation (Buker et al., 2007; Irvine et al., 2006). In cells expressing a slicing-defective form of Ago1, Chp1, Rd1 and H3K9me spreading at centromeres is reduced, while Ago1 enrichment is slightly increased (Buker et al., 2007; Irvine et al., 2006). In support of a role for ARC in passenger strand release, Arb1 was shown to inhibit the slicing activity of fission yeast but not \( D. \ melanogaster \) Ago1 (Buker et al., 2007). Generation of an Ago1 mutant defective in splicing showed that whereas wildtype Ago1 associates primarily with single stranded siRNAs, slicer-defective Ago1 associates primarily with dsRNA, similar to Arb1 (Ago1 mutants were expressed in cells that also possessed wildtype
Ago1 so that siRNA generation was preserved) (Buker et al., 2007). Thus, ARC serves as a checkpoint for Ago1 association with RITS, mediating loading of duplex siRNAs onto Ago1 and regulating slicer-dependent ejection of the passenger strand. The end result ensures that only Ago1 associated with a single-stranded guide siRNA enters the RITS effector complex.

Figure 1.1: The nascent transcript model.

v. Summary of the nascent transcript model

The above observations have all collectively contributed to the nascent transcript model of RNAi-mediated transcriptional silencing and heterochromatin formation in fission yeast. In this model, RITS, bound by a single stranded guide siRNA, associates with chromatin via RNA-RNA base pairing interactions between the guide siRNA and a nascent transcript. Importantly, RITS stabilization with heterochromatin is mediated by binding of the Chp1 chromodomain to
methylated H3K9. Next, RITS recruits two other protein complexes, RDRC and CLRC. RDRC uses the nascent transcript as a template to generate long dsRNA, which is cleaved by Dcr1, recruited by some combination of RDRC, RITS, and Dsh1, to generate double stranded siRNAs, amplifying the population of siRNAs. These in turn are loaded onto Ago1 in an ARC-dependent manner and the passenger strand of siRNA is ejected in a slicer-dependent manner. Ago1 bound to a single guide siRNA is then free to associate with the RITS complex once more. CLRC recruitment to the locus is mediated by interactions with RITS and Stc1, and may also involve association of CLRC component Rik1 with nascent transcripts. At chromatin, CLRC mediates methylation of H3K9, which stabilizes the association of Chp1 and remaining chromodomain-containing proteins. These include HP1 proteins, resulting in formation of heterochromatin, and importantly, Clr4 itself, allowing spreading of the H3K9me2 signal along chromatin beyond siRNA producing regions. The mechanisms by which subsequent silencing occurs, whether at the level of transcription or downstream, will be discussed below.

B. Transcriptional and co-transcriptional gene silencing

i. Role of RNA Polymerase II

Central to the seemingly contradictory nature of centromeric gene silencing in fission yeast is its requirement for transcription of loci that are to be silenced. However, this facet has been confirmed by multiple studies involving RNA Polymerase II. Pol II has been shown to ChIP to centromeric repeats in RNAi and heterochromatin mutants, and centromeric transcripts are polyadenylated, supporting the idea that Pol II is responsible for generation of centromeric transcripts (Chen et al., 2008; Djupedal et al., 2005; Kato et al., 2005). Cells expressing a temperature-sensitive mutation in a subunit of Pol II, rpb7-G150D, were impaired in chromosome segregation and centromeric silencing, siRNA generation, and heterochromatin formation, while few genes were down-regulated genome-wide, and these did not include factors known to be involved in centromeric silencing (Djupedal et al., 2005). Cells expressing
rpβ-G150D7 mutant showed much lower accumulation of centromeric transcripts than dcr1Δ cells, though still higher than in wt. This reduced transcript level was maintained in rpβ7-G150D dcr1Δ double mutants, indicating that rpβ7-G150D is defective at a step upstream of dcr1+, which is likely to involve the nature of the RNA that is transcribed (Djupedal et al., 2005).

A mutation in the second largest subunit of Pol II, rpβ2-m203, also showed phenotypes very similar to those of RNAi mutants: impaired chromosome segregation, and loss of silencing and heterochromatin formation at centromeres but not mating type loci, where RNAi is not required for maintenance (Kato et al., 2005). Importantly, expression of euchromatic gene act1+ and heterochromatin factors was not affected in rpβ2-m203 cells. Expression of centromeric transcripts was elevated in rpβ2-m203 cells to about half the level of dcr1Δ. In contrast to rpβ7-G150D cells, expression of centromeric transcripts in rpβ2-m203 dcr1Δ cells matched that of dcr1Δ cells, indicating that the phenotypes observed in the Pol II mutant were the result of impaired processing of centromeric transcripts rather than impaired transcription (Kato et al., 2005). Thus, Pol II contributes to centromeric silencing and heterochromatin formation by generating the nascent RNA template for siRNA-programmed RITS and possibly influencing whether the nascent RNA is a suitable template by coordinating the co-transcriptional RNA processing.

ii. Role of splicing factors

Intriguingly, and consistent with a co-transcriptional model of heterochromatic silencing, several splicing factors have been shown to play a role in centromeric silencing and siRNA generation, but not heterochromatin formation. Previously, several splicing factors were found to physically associate with RDRC component Cid12 (Motamedi et al., 2004). Furthermore, mutant forms of a number of fission yeast splicing factors resulted in alleviation of centromeric silencing of a reporter gene and endogenous centromeric transcripts, concomitant with a decrease in levels of centromeric siRNAs (Bayne et al., 2008; Chinen et al., 2010). Effects on centromeric
H3K9 methylation were weak in spliceosome mutants. This effect was downstream of transcription, as double mutants with \textit{dcr1}\textsuperscript{\Delta} showed similar transcript lengths and levels to the single \textit{dcr1}\textsuperscript{\Delta} strain. Silencing defects due to splicing mutants are also downstream of RITS recruitment, as introduction of mutations disrupted silencing when RITS was tethered to \textit{ura4-BoxB}, an intron-less transcript, via Tas3-\textit{\lambda}. ChIP experiments showed that certain splicing factors that were shown to be required for silencing of centromeric regions indeed physically associate with those loci in a Dcr1-dependent manner (Bayne et al., 2008). Initially, it was thought that splicing contributed to centromeric silencing in a manner beyond splicing and regulation of RNAi factors themselves, since replacement of known intron-containing RNAi factors with cDNA copies did not rescue silencing (Bayne et al., 2008). However, a more recent study, which examined the splicing of a larger number of factors required for RNAi-mediated heterochromatin formation, found several to be misregulated in a splicing mutant with centromeric silencing defects. Introduction of cDNAs corresponding to the mis-regulated RNAi factors largely alleviated silencing defects associated with the splicing mutant (Kallgren et al., 2014), leaving the role of splicing in RNAi-mediated silencing unclear.

\textbf{iii. HP1 and histone deacetylases (HDACs) mediate transcriptional gene silencing and co-transcriptional gene silencing}

Methylation of H3K9 in fission yeast is not sufficient to form fully functional heterochromatin and induce transcriptional gene silencing. Methylated H3K9 serves as an important binding site for the chromodomains of multiple proteins in fission yeast, including Chp1 and Clr4 (Al-Sady et al., 2013; Partridge et al., 2002; Petrie et al., 2005; Zhang et al., 2008). The two remaining chromodomain-containing proteins in fission yeast are HP1 homologs Swi6 and Chp2, which also possess chromo-shadow domains thought to be important for protein-protein interactions that allow self-association and spreading (Motamedi et al., 2008). Swi6 and Chp2 specifically bind H3K9me at all heterochromatic loci and are required for
silencing and proper centromere function, but not establishment of H3K9me domains (Bannister et al., 2001; Cam et al., 2005; Ekwall et al., 1995; Nakayama et al., 2001; Thon and Verhein-Hansen, 2000). Both Swi6 and Chp2 are required for localization and spreading of downstream proteins important for heterochromatin silencing, including SHREC, a complex that contains histone deacetylase (HDAC) Clr3 and mediates H3K14 deacetylation and transcriptional gene silencing (Motamedi et al., 2008; Sugiyama et al., 2007; Yamada et al., 2005). While deletion of chp2Δ or any SHREC components results in alleviation of silencing at all sites of constitutive heterochromatin, transcript accumulation occurs to a much lower extent than in clr4Δ cells (Motamedi et al., 2008). However, increases in H3K14Ac and Pol II occupancy in chp2Δ or SHREC mutants mirror those of clr4Δ cells, indicating that Pol II activity does not fully account for the silencing at constitutively heterochromatic loci (Motamedi et al., 2008).

Mass spectrometry analysis of Swi6-associated proteins revealed a number of nuclear proteins, including chromatin remodelers such as CHD proteins Hrp1 and Hrp3, FACT complex components Spt16 and Pob3; centromere-binding protein (CENP-B) homologs; and the SHREC and Clr6 HDAC complexes (Fischer et al., 2009; Motamedi et al., 2008). Many of these, including FACT and Hrp1, are themselves required for efficient silencing of heterochromatic regions (Lejeune et al., 2007; Walfridsson et al., 2005). In addition, Swi6 was found to associate with CENP-A, a histone variant involved in chromosome segregation, and Mis4, a cohesin loading factor. Localization of Mis4 and cohesin subunit Rad21 within centromeric heterochromatin was dependent on presence of swi6 (Fischer et al., 2009). Like chp2Δ, deletion of swi6Δ results in only a partial deprepression of silencing at heterochromatic loci; deletion of both HP1 homologs results in increased transcript levels and Pol II association, providing evidence that chp2 and swi6 function in non-overlapping pathways (Fischer et al., 2009; Motamedi et al., 2008). While both chp2Δ and swi6Δ have been shown to be required for centromeric silencing, only Swi6 binds centromeric transcripts (in a Clr4-dependent manner), and only Swi6 is required for generation of siRNAs (Bühler et al., 2006; Motamedi et al., 2008;
Motamedi et al., 2004). Swi6 was also shown to be required for Chp1 or Rdpl association with centromeric transcripts, also in a Clr4-dependent manner (Motamedi et al., 2008). These data illustrate the manners in which HP1 homologs Swi6 and Chp2 differentially contribute to silencing, heterochromatin formation, and subsequently, centromere function. Chp2 mainly contributes to recruitment of the SHREC HDAC complex, which mediates transcriptional gene silencing. Swi6 also contributes to SHREC association, but additionally contributes to siRNA generation, cohesin association, and recruitment of various other chromatin-associated factors.

iv. TRAMP and the exosome contribute to co-transcriptional gene silencing

I have mentioned above the evidence suggesting that while nucleation of heterochromatin negatively correlates with reduced Pol II occupancy, this reduction does not sufficiently explain the extent of silencing at fission yeast centromeres. For example, strains lacking HP1 protein chp2+ show the same level of Pol II occupancy at centromeres as strains lacking clr4+, but accumulate centromeric transcripts to a much lower degree (Motamedi et al., 2008). Furthermore, tethering of CLRC component Rik1 to a ura4 transcript results in silencing that requires the catalytic activities of RNAi machinery (RTS, RDRC, and Dcr1), but not CLRC, and does not correlate with H3K9me enrichment (Gerace et al., 2010). Therefore, there must be additional non-transcriptional mechanisms of RNA silencing at work. One of these involves RNAi-dependent co-transcriptional RNA degradation; another involves the S. pombe homolog of the TRAMP complex. In S. cerevisiae, the TRAMP complex, composed of RNA helicase Mtr4, poly(A) polymerase Trf4, and zinc knuckle protein Air2, is a nuclear partner to the exosome complex of 3’-5’ exonucleases, and mediates polyadenylation of transcripts targeted for degradation by the exosome (LaCava et al., 2005; Wyers et al., 2005). The fission yeast homolog for Trf4 is encoded by cid14+ (Win et al., 2006). The polyadenylation activity of Cid14 was subsequently found to be required for silencing at all heterochromatic loci, though its contribution to silencing was generally smaller than that of Clr4. Cid14 does not contribute to
proper localization of Chp1 or Swi6, and \textit{cid14}\textsuperscript{\textDelta} cells show no defect in H3K9 methylation at centromeres (Buhler et al., 2007). However, loss of Cid14 polyadenylation activity results in near complete loss of siRNAs. Deletion of exosome component \textit{rrp6}\textsuperscript{+} also results in loss of silencing at all heterochromatic loci, also to a lesser extent than \textit{clr4}\textsuperscript{\textDelta} (Buhler et al., 2007). Furthermore, while single \textit{rrp6}\textsuperscript{\Delta} deletion mutants have only a mild silencing defect, deletion of \textit{rrp6}\textsuperscript{+} exacerbated the phenotype of \textit{ago1}\textsuperscript{\Delta} cells, causing even higher levels of centromeric transcript accumulation and lower levels of H3K9 methylation (Reyes-Turcu et al., 2011a).

Together, this data demonstrates that the cell employs multiple pathways to ensure silencing of centromeric transcripts, including heterochromatin-based transcriptional silencing, and RNAi and TRAMP/exosome contributions to co-transcriptional silencing.

Yamanaka and colleagues found that in cells lacking \textit{rrp6}\textsuperscript{+}, many normally euchromatic loci form RNAi- and RITS-dependent heterochromatin domains (HOODs) to which both sense and antisense siRNAs map (2013). Interestingly, HOODs are not restricted to convergent genes or genes overlapping antisense ncRNAs. sRNA size distributions and first nucleotide distribution were consistent with those of siRNAs, and generation of siRNAs and heterochromatin required \textit{clr4}\textsuperscript{+}, \textit{rdp1}\textsuperscript{+}, \textit{dcr1}\textsuperscript{+} and \textit{ago1}\textsuperscript{+}. Several developmentally associated HOODs exhibited RNAi- and Clr4-dependent silencing (Yamanaka et al., 2013). This study further demonstrates that RNAi and the exosome cooperate to regulate many overlapping targets.

\textbf{C. Biogenesis of early siRNAs and establishment of centromeric heterochromatin}

While the maintenance of RNAi-mediated centromeric heterochromatin formation is now relatively well understood, its initiation is still a mystery. Because of the interdependence of siRNA generation and heterochromatin formation, it is unclear how the first siRNAs might be generated so that RITS can target centromeres for H3K9me and heterochromatin formation. Several studies have attempted to identify differential requirements distinguishing heterochromatin establishment and maintenance. In one such study, a mutant form of Tas3
(Tas3-WG) incapable of associating with Ago1 has little effect on maintenance of centromeric heterochromatin, but shows considerable defects in establishment of centromeric heterochromatin, silencing, and generation of siRNAs when clr4 is reintroduced to clr4Δ tas3-wg cells (Partridge et al., 2007); similar results were obtained in cells containing a mutant form of Chp1 that has reduced binding affinity to methylated H3K9 (Schalch et al., 2009). In contrast, tas3-wg cells show no defect in reconstitution of centromeric heterochromatin when siRNAs are depleted and reintroduced by deleting then reintroducing dcr1⁺, or upon deletion and reintroduction of RDRC components or ago1⁺. Low levels of centromeric H3K9me persist in cells lacking RNAi, and the authors argued that this suggests that H3K9 methylation is upstream of siRNA generation with regards to establishment of centromeric heterochromatin, and that RNAi is initially recruited to heterochromatin via intact RITS (Partridge et al., 2007; Shanker et al., 2010). While the data suggests that low levels of H3K9 methylation are sufficient to re-establish full RNAi-dependent silencing and heterochromatin in an RNAi mutant, another interpretation is that in the absence of fully functional RNAi, de novo establishment of heterochromatin at centromeres is not possible, and as such, RNAi activity must be required for heterochromatin establishment. Consistent with the latter hypothesis, other experiments have revealed no re-establishment of heterochromatin by reintroduction of clr4 into clr4Δ ago1Δ or clr4Δ dcr1Δ cells (Ragunathan et al., 2015). Furthermore, other results suggest a CLRC-independent pathway for recruitment of RNAi machinery to centromeric sequences. DNA adenine methyltransferase identification (DamID) experiments, which are more sensitive than ChIP and allow detection of indirect and transient protein-chromatin interactions, followed by hybridization of DNA sequences to tiling microarrays, revealed that Dcr1 associates with centromeric repeats in a Clr4-independent manner, and that Rdpl association, while greatly reduced in clr4Δ cells, was also not abolished (Woolcock et al., 2011). Moreover, Dcr1 and to a varying extent Rdpl were found to associate with euchromatic loci lacking H3K9 methylation, also in a Clr4-independent manner. These results indicate that RNAi factors can be recruited to
chromatin independently of initial H3K9me, consistent with the slightly higher levels of siRNAs present in clr4Δ cells compared to dcr1Δ cells (Halic and Moazed, 2010). Furthermore, previously published data (Iida et al., 2008; Simmer et al., 2010) along with data I will present in Chapter 3 show that ectopic siRNAs are capable of establishing de novo H3K9 methylation at a euchromatic locus. Together, these data refute the idea that H3K9 methylation must precede siRNA-mediated targeting of a locus and support an RNAi-dependent role for de novo heterochromatin establishment at certain loci.

It remains unclear how initial recruitment of RNAi and initial generation of siRNAs takes place. As mentioned above, it has been suggested that the low levels of RNAi-independent H3K9 methylation at centromeres can recruit RITS and RDRC in an siRNA-independent manner (Noma et al., 2004; Partridge et al., 2007; Sadaie et al., 2004; Shanker et al., 2010), but there are also several alternative models that are not mutually exclusive. In the first model, intrinsic features of single stranded centromeric transcripts are recognized by RDRC, which then generates long dsRNA substrates for Dcr1. In the second model, overlapping transcripts resulting from bidirectional transcription at centromeres are recognized by Dcr1 and processed into siRNAs. As a corollary to this, Djupedal et al posited that due to the repetitive nature of centromeric sequences, individual centromeric transcripts could base-pair intramolecularly to form hairpins that could also be processed into siRNAs by Dcr1 (2009). So far, little data has been presented to support a model for Rdpl-based recruitment of downstream factors; on the contrary, Clr4-independent association of Dcr1 to centromeric repeats and euchromatic loci was found to be much stronger than Rdpl association to the same loci (Woolcock et al., 2011). In the following section, I will mainly report evidence supporting a model for Dcr1-based recognition of dsRNAs, resulting in subsequent generation of siRNAs and recruitment of downstream silencing machinery.

i. Dcr1-independent primal RNAs (priRNAs)
Because Dcr1 is the sole ribonuclease responsible for generation of siRNAs, it was widely presumed that cells lacking \( \text{dcr1}^+ \) also lack sRNAs. Indeed, initially, no difference in siRNA levels was observed between strains lacking various RNAi components. However, the development of more sensitive sRNA detection assays allowed Halic et al to distinguish higher levels of sRNAs associated with Flag-Ago1 than Flag alone in RNAi mutants, including, surprisingly, in \( \text{dcr1}^\Delta \) and \( \text{rdp1}^\Delta \) cells (2010). Deep sequencing of Flag-Ago1 associated sRNAs in these mutants revealed a significantly reduced, but still extant, population of sRNAs mapping to centromeric repeats. Centromeric sRNAs were lowest in \( \text{dcr1}^\Delta \) and \( \text{rdp1}^\Delta \) cells, and were significantly higher in \( \text{clr4}^\Delta \) cells (though still <1/10 of wildtype levels). The authors proposed that these \( \text{dcr1}^\Delta \) sRNAs, termed primal RNAs (priRNAs), could serve as the trigger siRNAs for establishment of centromeric heterochromatin. Furthermore, priRNAs might serve as an important model for studying piRNAs, a much larger and conserved class of Dicer-independent sRNAs. To this end, a follow-up study identified Triman as a 3'-5' exonuclease responsible for trimming of 3' ends of priRNAs and siRNAs (Marasovic et al., 2013), which may inform efforts to identify factors responsible for generation of 3' ends of piRNAs.

ii. Dicer-dependent generation of siRNAs from intramolecular hairpins

Deep sequencing of total sRNAs in fission yeast revealed that while centromere-derived sRNAs map to both strands of centromeric repeats, there was a slight preference for sRNAs mapping to the reverse strands of \( \text{dh} \) and \( \text{dg} \) repeats (Djupedal et al., 2009b). Furthermore, sRNAs map to distinct clusters, or “hot spots”, with significant cold spots lacking any sRNAs in between. This led the authors to hypothesize that “hot spot” regions correspond to sections of the RNA that fold back upon itself to form a hairpin structure, whereas “cold spots” remain single stranded. Based on secondary structure predictions, a 432nt sequence (termed \( \text{RevCen} \)) immediately downstream of a well-characterized centromeric promoter was selected for further study. Enzymatic and chemical cleavage analysis of the \textit{in vitro} transcribed fragment was
consistent with regions of duplex RNA, to which abundant siRNAs were found to map. Incubation of the in vitro transcribed RevCen with recombinant Dicer resulted in generation of appropriately sized sRNAs. sRNAs processed from a hairpin construct would not require activity by Rdpi; to test whether RevCen-originating sRNAs require Rdpi for generation, the authors sequenced sRNAs from rdpiΔ cells and found that, while centromeric siRNA levels were drastically reduced, some sRNAs did still map to RevCen. However, the authors did not sequence sRNAs in dcr1Δ cells and as such, the rdpi-independent RevCen sRNAs they observe may reflect dcr1-independent priRNAs. Because of this caveat, it is not possible to conclude from this study that endogenous centromeric transcripts produce hairpin structures that are processed by Dcr1 into siRNAs. Another model, which is not mutually exclusive with hairpin generation of some transcripts or at some frequency, is that Dcr1 processes overlapping convergent centromeric transcripts into siRNAs- I will present evidence supporting this model in Chapter 2.

iii. Dicer-based generation of siRNAs from overlapping transcripts

Analogous to, and not mutually exclusive with, the hypothesis that centromeric transcripts form intramolecular hairpin structures that can be processed into siRNAs by Dcr1 in the absence of rdpi+ is the alternative possibility that bidirectional transcription at centromeres results in overlapping transcripts that associate to generate long dsRNA that can be processed by Dcr1 into siRNAs in an rdpi-independent manner. I will present evidence supporting this hypothesis in Chapter 2 of this dissertation; however, some data has already been published regarding the hypothesis that the RNAi machinery can directly recognize and target overlapping transcripts. Gullerova et al have published numerous reports suggesting that RNAi can be recruited to convergent genes that transiently generate overlapping transcripts (2008; 2011; 2012). Using synchronized cell populations, the authors reported that convergent genes
generate overlapping dsRNA in G1 but not G2 (the longest) phase of the cycle, and G1-specific
dsRNA correlates with transient H3K9 methylation and Swi6 association in a Dcr1- and Ago1-
dependent but Rdpl-independent manner. Deletion of Dcr1 or Ago1 results in continued
presence of dsRNA during G2 (Gullerova et al., 2011; Gullerova and Proudfoot, 2008).

Expanding on this work, the authors examined the genes encoding RNAi and heterochromatin
factors, many of which are convergent themselves. They found that convergent RNAi genes are
downregulated during the G1-S transition, while tandem RNAi- and non RNAi- genes show no
change. Deletion of dcr1+, ago1+, or clr4+ results in upregulation of convergent RNAi transcripts
during the G1-S transition, while rdplΔ had no effect (Gullerova et al., 2011). Finally, it was
suggested that inserting a ura4+ gene in between two (strong) nmt1 promoters in convergent
orientation resulted in Ago1- and Dcr1- dependent but Rdpl-independent silencing of both
sense and antisense ura4 transcripts and generation of siRNAs (Gullerova and Proudfoot,
2012). In a strain also carrying ura4 at its endogenous locus, the convergent allele of ura4
(CTura4) was able to induce transcriptional gene silencing and H3K9 methylation of the
endogenous ura4+ locus in trans. However, attempts in our own lab to reproduce these results
have been unsuccessful; moreover, in past experiments I found that strong antisense
transcription at the ura4+ endogenous locus resulted in silencing by transcriptional interference
that was not associated with H3K9 methylation and was RNAi-independent (unpublished data).

Nevertheless, other studies (Iida et al., 2008) as well as results I will present in Chapters 2 and
3 of this dissertation indicate that, under certain conditions, convergent loci are capable of
generating siRNAs and being targeted by siRNAs for heterochromatin formation and silencing.

D. Heterochromatin-inhibiting factors and the role of histone turnover

In 2006, it was noted by Buhler et al that despite the generation of ura4 siRNAs when
tethered RITS mediates silencing at ura4, no trans-silencing or H3K9 methylation was detected
at a second copy copy of ura4+ located on another chromosome, demonstrating a cis restriction
to chromatin-based RNAi activity in fission yeast. Indeed, while there have been many examples of cis RNAi activity, for example the silencing of reporter genes when inserted into heterochromatic regions, or when adjacent to a 1.6kb fragment of centromeric repeat sequence inserted into euchromatin (Allshire et al., 1994; Partridge et al., 2002), reports of trans siRNA-mediated silencing are rare and inconsistent.

Several examples have been published of trans siRNA-mediated silencing and heterochromatin establishment at a euchromatic locus in wildtype cells, with no clear consensus on locus features that affect silencing and heterochromatin formation. As previously mentioned, Gullerova et al found that strong convergent transcription of a ura4 sequence resulted in RNAi-dependent silencing of both the convergent locus and the endogenous ura4 locus (2012). However, Iida et al found that expression a ura4 hairpin generating high levels of siRNAs could not silencing the endogenous ura4+ locus but did result in hairpin-dependent silencing, H3K9 methylation, and Swi6 association at the trp1::ura4 locus, which correlated with antisense transcription (2008). Silencing at the trp1::ura4 locus, once established, required all RNAi and heterochromatin machinery, and was not maintained in the absence of hairpin expression, indicating that constant targeting/re-establishment was required. Sigova et al found that expression of a gfp hairpin construct failed to induce transcriptional silencing or changes in chromatin at a copy of gfp inserted into the genome (2004). In contrast, Simmer et al showed that expression of a gfp hairpin results in silencing and heterochromatin formation at a ura4-gfp fusion construct inserted at the arg3 locus and silencing of an ade6-GFP locus; silencing did not correlate with antisense transcription (2010). Thus, it remained unclear what distinguishes a locus that can be targeted for RNAi-mediated silencing or heterochromatin formation from one that cannot.

A number of proteins have been shown to actively inhibit silencing and heterochromatin genome-wide. Deletion of these genes rescues the loss-of-silencing phenotype observed at centromeres in RNAi mutants. Their study provides insight into the myriad mechanisms by
which cells balance spreading and inhibition of heterochromatin formation, as well as non-RNAi mechanisms of recruiting silencing and heterochromatin machinery.

i. Epe1, a putative demethylase

*epe1*<sup>+</sup> was first identified in a screen for mutations that permit spreading of silencing beyond the defined boundaries of heterochromatin and silencing at the *mat* locus (Ayoub et al., 2003). Deletion of *epe1*<sup>+</sup> was found to promote spreading of H3K9me and Swi6, but has no effect on establishment of heterochromatin (Ayoub et al., 2003; Ragunathan et al., 2015). Overexpression of Epe1 results in disruption of heterochromatin and silencing, increased acetylation of H3K9 and H3K14, and increased methylation of H3K4, all of which are associated with euchromatin and expression (Ayoub et al., 2003; Zofall and Grewal, 2006). Localization of Epe1 mirrors that of H3K9me and Swi6, and is dependent on both Clr4 and Swi6, but not RNAi, suggesting a possible Swi6-dependent recruitment mechanism (Isaac et al., 2007; Zofall and Grewal, 2006). Consistent with this, Epe1 was shown to physically interact with Swi6 in a manner that did not require its JmjC domain (Zofall and Grewal, 2006). Deletion of *epe1*<sup>+</sup> rescued the loss of silencing and heterochromatin phenotype in both RNAi and SHREC (Clr3 HDAC) deletion strains; in *epe1*<sup>−</sup> RNAi<sup>−</sup> or *epe1*<sup>−</sup> *clr3*<sup>−</sup> double mutants, heterochromatic transcript accumulation and Pol II association with centromeric repeats was decreased, but Swi6 association was increased, compared to single RNAi or Clr3 mutants (Trewick et al., 2007; Zofall and Grewal, 2006). It is important to note that low levels of H3K9 methylation are still present in both RNAi and Clr3 mutants, and *epe1*<sup>−</sup> had no effect on the silencing phenotype of *clr4*<sup>−</sup> cells (Zofall and Grewal, 2006). Interestingly, overexpression of Epe1 resulted in a Swi6-dependent decrease of dimethylated H3K9, but appeared to have no effect on trimethylated H3K9 (Zofall and Grewal, 2006), which may reflect a more stable or protected state when H3K9 is tri- rather than di-methylated. Together, this data shows clearly that Epe1 counteracts and
restricts heterochromatin formation, likely a protective measure employed by the cell to balance mechanisms involved in spreading.

Epe1 is a jmjC-domain containing protein, and the jmjC domain is required for activity but not localization (Ayoub et al., 2003; Zofall and Grewal, 2006). Other jmjC-domain containing proteins have been shown to have demethylase activity, and data consistent with Epe1 demethylase activity has been published. Tethering Clr4 to a euchromatic reporter locus using the induceable tetracycline repressor (TetR) system results in establishment of strong heterochromatin and silencing of the gene. In wildtype cells, H3K9me and silencing are rapidly lost upon tetracycline-induced release of TetR-Clr4 (Audergon et al., 2015; Ragunathan et al., 2015). However, in *epe1*Δ cells, H3K9me and silencing were maintained through many generations and through meiosis (Audergon et al., 2015; Ragunathan et al., 2015). However, thus far no demethylase activity has been detected in Epe1, and its mechanism of action remains unknown.

The role of Epe1 in nucleosome turnover, which may reflect its putative role in H3K9 methylation, has also been explored. Aygun et al devised a system in which H3-Flag expression could be induced by growth on sucrose-containing medium after cell cycle arrest by growth on hydroxyurea (HU) (2013) By immunoprecipitating Flag-tagged H3 after micrococcal nuclease (MNase) digestion of crosslinked chromatin (MNase-ChIP), the authors could compare differences in incorporation of Flag-H3 and make inferences regarding nucleosome turnover. Heterochromatic regions were found to have much lower rates of nucleosome exchange than euchromatin, which was not a result of lower nucleosome occupancy. At pericentromeric heterochromatin, deletion of *clr4*+ or both HP1 homologs (*swi6Δ chp2Δ*), resulted in significant increases in nucleosome exchange; no changes were noted at euchromatic loci. At the *mat* locus, where RNAi is not required for silencing or heterochromatin, *dcr1Δ* had no effect on nucleosome exchange, and surprisingly, neither did *clr4Δ* or *swi6Δ chp2Δ*. However, deletion of the histone deactylase *clr3*, which can be recruited to *mat* by separate factors, resulted in
significant increase in nucleosome turnover at both the mating type locus and pericentromeric heterochromatin, in a manner that did not correlate with Pol II occupancy. To test the role of Epe1 on histone turnover, Epe1 was overexpressed, resulting in an increase in Flag-H3 incorporation across the pericentromeric region. Additionally, the authors found that deletion of RNAi factors results in increased histone turnover at the centromeres, but the additional deletion of *epe1* suppresses this effect (Aygun et al., 2013; Wang et al., 2015). This is unsurprising since deletion of *epe1* has been shown to suppress the H3K9me and silencing defect of RNAi mutants (Trewick et al., 2007; Zofall and Grewal, 2006).

ii. **Mst2, a histone acetyltransferase**

Mst2 mediates histone H3K14 acetylation, which has been shown to be upregulated at centromeres in heterochromatin deficient cells (Ayoub et al., 2003; Motamedi et al., 2008; Sugiyama et al., 2007). Reddy et al showed that deletion or catalytic inactivation of Mst2 results in rescue of centromeric H3K9 methylation and silencing in cells lacking RNAi machinery, but no effect on cells lacking CLRC, HP1, or SHREC (2011). However, in *mst2Δ dcr1Δ* cells in which heterochromatin was erased by various methods, pericentromeric heterochromatin could not be re-established. This effect was specific to Mst2, as deletion of another acetyltransferase, Gcn5, did not result in similar phenotypes. Interestingly, deletion of a Mediator component, *pmc2*+, also rescued loss of centromeric silencing in RNAiΔ cells, though to a much lesser extent, which together with the *mst2Δ* data, may suggest a role for reduced Pol II activity in permitting RNAi-independent heterochromatin, since H3K14ac correlates with active transcription (Reddy et al., 2011). Wang and colleagues found that disruption of centromeric heterochromatin by deletion of *dcr1* resulted in increased histone turnover, but this was reduced back to wildtype levels in *mst2Δ dcr1Δ* cells (2015).

How does deletion of (putative) chromatin-modifying enzymes enable establishment of RNAi-independent heterochromatin formation at centromeres? One possible mechanism is that
loss of epe1\(^+\) or mst2\(^+\) permits alternative mechanisms of CLRC recruitment to nascent centromeric transcripts, possibly involving direct association of Rik1 with RNA. However, recent experiments support an alternative model, in which deletion of epe1\(^+\) or mst2\(^+\), which negatively regulate H3K9 methylation either by direct demethylation or by promoting nucleosome turnover, results in increased levels of H3K9me and improved epigenetic maintenance (Audergon et al., 2015; Ragunathan et al., 2015).

iii. Mlo3 and Cid14: transcript processing factors

In addition to Mst2 and Epe1, both of which operate on histones, various RNA- and transcription- associated factors also negatively regulate heterochromatin formation. Mlo3 was originally identified in a screen for proteins that disrupt mitotic chromosome segregation when overexpressed. Strong overexpression of Mlo3 causes lethality and complete chromosome segregation failure, with a high percentage of daughter cells with either two nuclei or no nuclei (Javerzat et al., 1996). Mlo3 is a homolog of yeast YRA1 or human Aly/Ref, which link RNAPII and 3′ end processing factors to mRNA export (Johnson et al., 2009; Johnson et al., 2011; MacKellar and Greenleaf, 2011). mlo3\(^-\) cells accumulate transcripts in the nucleus (Reyes-Turcu et al., 2011). Mlo3 has been reported to associate with Clr4, Rik1, and Swi6, and while mlo3\(^-\) has no effect on centromeric H3K9me or Pol II occupancy, it shows a slight increase in centromeric transcript levels, with a corresponding decrease in centromeric siRNAs (Fischer et al., 2009; Zhang et al., 2011). In wildtype cells, Mlo3 associates mainly with euchromatic loci, particularly 3′ ends of ORFs, but in clr4\(^-\) cells, Mlo3 association with centromeric repeats is increased, perhaps reflecting increased transcription. Interestingly, antisense transcription, particularly at convergent genes, is upregulated in mlo3\(^-\) cells (Zhang et al., 2011). mlo3\(^-\) ago1\(^-\) and mlo3\(^-\) dcr1\(^-\) showed restored centromeric silencing and heterochromatin function with no effect on siRNA levels (Reyes-Turcu et al., 2011). Mlo3 was shown to associate with centromeric transcripts, particularly in ago1\(^-\) cells. In contrast to wildtype cells, mlo3\(^-\) cells
expressing a \textit{trp1} hairpin construct were able to establish H3K9 methylation at the endogenous \textit{trp1} locus in \textit{trans} (Reyes-Turcu et al., 2011). Finally, genome-wide analysis showed that in \textit{mlo3}\textsuperscript{\Delta} cells, H3K9 methylation forms in an RNAi-independent manner at some euchromatic genes, and this result correlates with \textit{mlo3}\textsuperscript{\Delta}-dependent increases in antisense transcript at those loci (Reyes-Turcu et al., 2011). Loci with \textit{mlo3}\textsuperscript{\Delta}-independent sense and antisense transcripts showed no increase in H3K9 methylation. In Chapter 4, I will present evidence that \textit{mlo3}\textsuperscript{\Delta} cells also display an improved ability for siRNAs to mediate dimethylation of H3K9 in \textit{trans}, indicating that the genome-wide defects in mRNA export, which might cause prolonged transcript association with chromatin, may positively affect siRNA-targeting.

In addition to its association with Clr4 and Rik1, Mlo3 also physically associates with all TRAMP components in a Clr4-, RNaseA/H-, and DNase- independent manner (Zhang et al., 2011). As previously mentioned, Cid14 is a component of the fission yeast TRAMP complex, and mediates polyadenylation of transcripts to stimulate degradation by the exosome. Deletion of \textit{cid14} results in phenotypes similar to \textit{mlo3}\textsuperscript{\Delta}, including suppression of centromeric heterochromatin and silencing defects in RNAi deletion cells (Reyes-Turcu et al., 2011). This may reflect mechanisms similar to that observed at HOODs, in which deletion of exosome component \textit{rrp6}\textsuperscript{+} results in stabilization of certain transcripts that are then targeted by RNAi for silencing and heterochromatin formation (Yamanaka et al., 2013). Perhaps deletion of transcript processing factors \textit{mlo3}\textsuperscript{+} or \textit{cid14}\textsuperscript{+} results in increased stabilization and association of centromeric transcripts with chromatin, allowing RNAi-independent mechanisms of CLRC recruitment. Alternatively, though not mutually exclusive, it is possible that deletion of \textit{mlo3}\textsuperscript{+} or \textit{cid14}\textsuperscript{+} also alters Pol II behavior and/or results in decreased nucleosome turnover, which would reduce removal of nucleosomes containing methylated H3K9.

\textbf{iv. Paf1, a transcription elongation and 3’ processing complex}
Deletion of a known Pol II elongation factor, *tfs1* (the fission yeast homolog of TFIIIS), has been shown to suppress silencing and heterochromatin defects in RNAi mutants in a Clr4-dependent manner, but suppression is variegated and weaker than in *mlo3Δ* (Reyes-Turcu et al., 2011). However, a screen for mutants that permit hairpin-mediated silencing of a reporter gene in *trans* revealed a significant role for another transcription elongation factor, the Paf1 complex (Kowalik et al., 2015). The Paf1 complex has been implicated in Pol II transcription elongation, 3’ end processing, and transcription-coupled histone modifications, among other roles (Tomson and Arndt, 2013). Mutants in all Paf1 components showed strong hairpin-mediated silencing and H3K9 methylation at an ade6 reporter gene, in an RNAi, CLRC, HP1, and HDAC-dependent manner. Silencing was accompanied by generation of secondary siRNAs along the ade6* reporter gene that did not map to the hairpin construct. The phenotype was not limited to hairpin constructs or the ade6 locus, as similar results were obtained using an euchromatic *ura4* reporter gene and a heterochromatic driver siRNA-producing copy of *ura4* in Paf1 mutants. The silenced epigenetic states were stable through mitosis and meiosis in cells that retained the Paf1 mutations, including cells that lost the driver siRNA-producing locus, indicating that the secondary siRNAs produced are sufficient to maintain the epigenetic state.

Insertion of a self-cleaving ribozyme sequence immediately downstream of the ade6 ORF resulted in loss of hairpin-dependent silencing in Paf1 mutants, indicating that it is likely the 3’ processing role of Paf1 that inhibits *trans* siRNA-mediated heterochromatin formation (Kowalik et al., 2015).

In a separate study, Paf1 was also found to affect spreading of heterochromatin past defined boundary elements, similar to Epe1 (Sadeghi et al., 2015). Mutations in Paf1 components *leo1* and *paf1*, but not other Paf1 components, resulted in silencing of an ade6* reporter gene inserted outside the heterochromatin boundaries of the *mat* locus. Additionally, mutations in *paf1* or *leo1* rescued the loss of centromeric silencing and chromosome segregation defects in RNAi mutant cells, but not in SHREC or HP1 mutants. While centromeric
heterochromatin showed no defects in leo1Δ cells, centromeric siRNAs were largely reduced, though still present, reflecting decreased entry of centromeric transcripts into the siRNA-generation pathway. A recombination-induced tag exchange (RITE) assay, in which expression of a hormone switches an epitope tag on histone H3, showed that in unsynchronized cells, deletion of leo1 resulted in a decrease of H3 incorporation at pericentromeric regions and mat, as well as the euchromatic and Pol II-transcribed regions tested, which could explain its heterochromatin-stabilizing effects (Sadeghi et al., 2015). This likely at least partially contributes to the ability of siRNAs to establish de novo heterochromatin in Paf1 mutants: decreased nucleosome turnover would result in prolonged association of H3K9me-containing nucleosomes with the locus and facilitated maintenance. However, this is not mutually exclusive with the model that mutation of Paf1 contributes to siRNA-mediated heterochromatin formation by delaying transcript cleavage and it is very possible that the two observations are related. Perhaps decreased Pol II activity at the 3’ ends of protein-coding genes in Paf1 mutant cells results in delayed transcript cleavage as well as reduced nucleosome turnover, contributing to siRNA-mediated heterochromatin formation and maintenance by distinct mechanisms.

IV. Small RNA silencing of transposable elements in multicellular organisms

Transposons are selfish nucleic acid elements that can insert themselves into the genome and “jump” from one locus to another. In the event that a parasitic transposon disrupts an important gene or regulatory network, this can be catastrophic for the cell, as evidenced by the devastating effects on fertility and germline development upon loss of transposon inactivation. For this reason, suppression of transposon mobilization is regarded as a critical aspect of genome maintenance. An important ability for cells seeking to silence “foreign” genetic elements but retain expression of “self” elements is the ability to distinguish between “self” and “non-self” sequences. To achieve this, many eukaryotes have evolved sRNA-based mechanisms of silencing that identify targets by base-pairing interactions and operate at both
the post-transcriptional and transcriptional levels. These systems exploit the identifying feature of mobilized transposons- transcription- to identify targets for silencing, using the transcripts both to amplify sRNA populations and as RNA landmarks that sRNAs can recognize by base-pairing interactions to recruit the heterochromatin machinery. Subsequently, cells propagate the silenced state with epigenetic markers (DNA methylation in plants and mice, H3K9 methylation in flies, worms, and fission yeast), ultimately resulting in formation of transcriptionally silenced heterochromatin. Here, I will discuss what is known of sRNA-based mechanisms of silencing and heterochromatin formation at transposons in multicellular eukaryotes.

**A. Arabidopsis thaliana**

The earliest discovery of RNA-directed DNA methylation (RdDM) was made in plants well before the discovery of RNAi (Wassenegger et al., 1994). The characterization of the conserved sRNA-mediated PTGS pathway led to the discovery in *Arabidopsis* that small RNAs can direct RdDM, resulting in transcriptional gene silencing (Mette et al., 2000). Since then, extensive research has clarified this mechanism. RdDM requires the activity of two RNA Polymerase II-related RNA polymerases, Pol IV and Pol V, which are involved in siRNA biogenesis and targeting, respectively. By mechanisms that are still unclear, but may involve histone modifications such as H3K9 methylation, Pol IV is recruited mainly to transposable elements and repeat sequences, where it produces a single-stranded transcript. RNA-dependent RNA polymerase 2 (RDR2) associates with Pol IV and uses the ssRNA as a template to generate dsRNA. DICER-LIKE 3 processes the dsRNA into 24-nt siRNAs, which are exported to the cytoplasm and loaded onto Argonaute AGO4. siRNA-bound AGO4 is then imported into the nucleus, where it is guided to a complementary nascent Pol V transcript and directly associates with Pol V, ultimately resulting in de novo methylation of cytosines and transcriptional silencing. It is unknown why a second polymerase is required, or what factors recruit Pol V, which associates mainly with euchromatic transposons and repeat sequences.
(Matzke and Mosher, 2014). However, this division, which is distinct from the cis regulation of centromeric transcripts in fission yeast, is reminiscent of animal piRNA systems, in which often heterochromatic piRNA clusters (which generate piRNAs) drive silencing of active transposons. About 70% of RdDM sites are also marked by H3K9 methylation, which can form a positive feedback loop with DNA methylation. Three histone methyltransferases, SUVH4, SUVH5, and SUVH6, mediate histone H3K9 methylation, which is closely associated with CMT3-mediated DNA methylation. SUVH proteins contain an SRA domain that binds methylated DNA, and correspondingly, CMT3 possesses a chromodomain that binds H3K9me (Matzke and Mosher, 2014), providing a self-reinforcing feedback loop that further stabilizes heterochromatin. Thus, certain aspects of siRNA-mediated transcriptional gene silencing, including the separation of siRNA-producing loci from siRNA-targeting loci, and the interplay between H3K9 and DNA methylation, are characteristic of heterochromatin formation in animals, making A. thaliana an important model system for sRNA-mediated transcriptional gene silencing.

B. Caenorhabditis elegans

The nematode C. elegans has proven to be an extremely valuable model for studying RNAi, being the organism where the first sRNA-target pair were identified five years before the discovery of RNAi (Lee et al., 1993; Wightman et al., 1993), and the organism in which RNAi was discovered (Fire et al., 1998). The roles of endogenous miRNAs in regulating development, and dsRNAs in triggering antiviral defense and transposon silencing have been well characterized (reviewed in Grishock, 2013). dsRNA in C. elegans mediates silencing of targets by a combination of post-transcriptional and transcriptional mechanisms, and a role for nuclear RNAi in reducing Pol II transcription has been clearly established (Guang et al., 2010). In somatic cells, siRNAs associate with nuclear Argonaute NRDE-3 to direct H3K9 methylation, resulting in heritable expression of siRNAs and heritable H3K9 methylation in F1 progeny (Burkhart et al., 2011; Burton et al., 2011). Exposure to dsRNA can also induce H3K9me3 that
persists 2 generations in the absence of the initial dsRNA trigger, spreads several kilobases beyond the initial triggering site, and requires the production of secondary siRNAs (Gu et al., 2012b). Important for the multigenerational inheritance of dsRNA-induced silencing is nuclear and germline-specific Argonaute HRDE-1, which associates with secondary siRNAs in the germ cells of the progeny of nematodes that were exposed to dsRNA (Ashe et al., 2012; Buckley et al., 2012). In germ cells, NRDE-3 is dispensable and HRDE-1 directs H3K9 methylation at targeted loci. Deficiencies in hrde-1 ultimately lead to sterile animals, consistent with HRDE-1 association with endogenous siRNAs. HRDE-1 is also important for piRNA silencing, and is thought to act downstream of piRNAs in germ-line silencing (Ashe et al., 2012). It is currently thought that piRNAs, or 21U-RNAs, associate with PIWI-protein PRG-1 initiate transposon silencing in the germline, and generation of RdRP-dependent secondary sRNAs allow HRDE-1 and various siRNA-associated and heterochromatin machinery to contribute to maintenance (Ashe et al., 2012; Lee et al., 2012; Shirayama et al., 2012). In contrast, nuclear Argonaute CSR-1 associates with sRNAs mapping antisense to nearly all germline-expressed genes and protects them from potential piRNA-mediated silencing, providing a beautiful example of “self” versus “nonself” recognition (Seth et al., 2013; Wedeles et al., 2013).

C. Drosophila melanogaster

Flies have also proven a valuable model system for both genetic and biochemical studies of RNAi mechanisms. The earliest biochemical studies of the RNAi machinery were performed in D. melanogaster cultured cells and extracts, resulting in the identification of RISC and Ago2 as a component of RISC (Hammond et al., 2000; Hammond et al., 2001). Piwi clade proteins were named after the founding Drosophila gene, which is essential for germline development (Malone and Hannon, 2009). In addition to Piwi, Drosophila possesses two other PIWI proteins, AGO3 and Aubergine (Aub). All are required for female fertility and transposon silencing, and all but Ago3 are required for male fertility. Aub and AGO3 act post-
transcriptionally via the ping-pong cycle described above, mediating target transcript cleavage in the cytoplasm. Piwi mediates transcriptional gene silencing in the nucleus by directing trimethylation of histone H3K9 at target transposons (Sienski et al., 2012), resulting in association of heterochromatin protein HP1a, decreased Pol II activity, and formation of silenced heterochromatin. Piwi-mediated silencing requires Maelstrom, but Maelstrom is not required for establishment of H3K9me3 (Iwasaki et al., 2015; Sienski et al., 2012). Interestingly, H3K9me3, coupled with association of a complex containing HP1a homolog Rhino, defines dual-stranded piRNA clusters; Piwi is required for Rhino association at certain loci, and generation of piRNAs requires H3K9me3 (Le Thomas et al., 2014; Mohn et al., 2014), which is similar to requirements for Ago1 and Swi6 in centromeric siRNA generation and H3K9 methylation in fission yeast. Interestingly, inheritance of maternally-derived piRNAs in the embryo is required for piRNA-mediated silencing in the progeny (Brennecke et al., 2008). Maternally-deposited piRNAs are responsible for initiating the ping-pong cycle and establishing deposition of H3K9me3 and Rhino at piRNA clusters, which permits transcription of the primary piRNA precursors (Le Thomas et al., 2014), and provides insight into how piRNA mechanisms are initiated in each new embryo.

D. *Mus musculus*

piRNAs were originally discovered in mice (Aravin et al., 2006; Girard et al., 2006), where they have been shown to silence transposons and repeat elements, both by PTGS and DNA methylation of CpG islands by mechanisms that are still largely unclear (Aravin et al., 2008; Aravin et al., 2007; Carmell et al., 2007; Kato et al., 2007). Mice express three PIWI proteins mainly in the germline, MIWI, MIWI2, and MILI, all of which affect spermatogenesis (Iwasaki et al., 2015). piRNA clusters in mammals are conserved in location (suggesting that location is functionally relevant, and may be related to heterochromatin) but not sequence (which is consistent with the idea of piRNA clusters being transposon ‘traps’), and include
unidirectional and bidirectional transcription. In contrast to *Drosophila*, piRNA biogenesis in mice may involve only one round of ping-pong. Reminiscent of ping-pong, Mili associates with primary piRNAs (which possess 5'Us and are oriented sense to transposons), whereas MIWI2 interacts with secondary piRNAs (possessing 10A and are antisense to transposons) (Aravin et al., 2006; Aravin et al., 2008). However, rather than continuing with subsequent ping-pong cycles, MIWI2 localizes to the nucleus in a Mili-dependent manner to direct DNA methylation at transposons (Aravin et al., 2008; Kuramochi-Miyagawa et al., 2008). Many details of the piRNA mechanism remain to be elucidated in mice, and work in the previously described systems, including fission yeast, will certainly help inform future studies.

**V. Regulation of gene expression and chromatin state by long non-coding RNA**

Whereas most non-protein-coding regions of the genome were previously thought to be non-transcribed, we now recognize that primary transcripts map to 75% of the human genome (Batista and Chang, 2013), including to regions of heterochromatin. Indeed, in previous sections I have discussed a requirement for transcription at heterochromatic loci, resulting in lncRNAs that are processed into sRNAs to target transposable elements for transcriptional gene silencing. Recent studies have also identified telomeric repeat-containing RNA (TERRA), a class of transcripts derived from constitutively heterochromatic telomeres that may play a role in telomere function (Azzalin et al., 2007; Schoeftner and Blasco, 2008). A more classic and better-studied lncRNA is X-inactivation-specific transcript, or Xist, which is required in *cis* for X-chromosome inactivation in mammals. Much remains to be discovered about the manner in which Xist and other lncRNAs contribute to mechanisms of silencing or activation, but the integral role of Xist in a fundamental process of the cell is unambiguous, and lncRNAs are likely to emerge as wide-spread and important regulators of gene expression and chromatin state, as sRNAs already have.
X-chromosome inactivation (XCI) is an example of dosage compensation, which is a necessary mechanism for ensuring proper gene expression levels from sex chromosomes. Because females have two copies of the X chromosome, and males only have one, one entire copy of the X chromosome in females is silenced by formation of facultative heterochromatin. The process by which this occurs is a subject of active research, and many details are still unknown. What has become clear, however, is the central role that various lncRNAs play in this mechanism, the star of which is Xist (Brockdorff et al., 1992; Brown et al., 1992; Lee and Bartolomei, 2013). Xist is a 17-20 kb spliced and polyadenylated transcript encoded within the 100-500 kb X-inactivation center (Xic), a region of the X chromosome that is necessary and sufficient for silencing in cis (Lee and Bartolomei, 2013). Within Xist are several conserved repeat sequences, including repeat-A (RepA, also an independent transcript), which is required for silencing of the X chromosome but not for localization of Xist, and repeat-C, which is important for association of Xist with the inactivated X chromosome (Xi). Deletion of Xist or disruption of its association with Xi results in loss of silencing (Beletskii et al., 2001; Marahrens et al., 1997; Penny et al., 1996).

At the initiation of XCI, Xist is expressed only from the chromosome to be inactivated. Within the Xist coding region is a “nucleation center” that consists of three Yin Yang 1 (YY1) transcription factor binding sites (Jeon and Lee, 2011). YY1 binds only Xi and is thought to co-transcriptionally tether Xist to chromatin. As inactivation progresses, Xist molecules spread and coat the entire inactivated X chromosome. Spreading of Xist correlates with increased association of Polycomb repressive complex 2 (PRC2), which mediates trimethylation of H3K27. However, it is unclear whether Xist directly contributes to PRC2 association, as Xist and PRC2 do not co-localize in a study using structured illumination microscopy (Cerase et al., 2014). Indeed, while some studies have reported direct interactions between the RepA region of Xist and PRC2, others have failed to identify PRC2 in purifications of Xist-interacting partners, and PRC2 association with the inactivated X chromosome is still observed when an Xist allele
lacking the RepA region is expressed (Galupa and Heard, 2015; Holoch and Moazed, 2015a). While Xist is required for PRC2 recruitment, a direct role in recruitment is unconfirmed. An intermediate step in PRC2 recruitment involves PRC2 cofactor Jarid2, which has been shown to interact with the B- and F- repeats of Xist (da Rocha et al., 2014).

Xist is in turn positively and negatively regulated by a number of other IncRNA-encoding loci. Expression of transcript Tsix negatively regulates Xist (Lee et al., 1999), likely through the act of antisense transcription itself (Stavropoulos et al., 2001). Jpx and Ftx, two IncRNA-encoding loci upstream of Xist, have been shown to positively regulate Xist expression, but data regarding the mechanisms by which they do so is contradictory and the mechanism remains unclear (Galupa and Heard, 2015).

As illustrated by Xist studies, the precise mechanisms by which IncRNAs moderate gene expression remain unclear. However, within the body of IncRNA research several themes emerge. First, transcription resulting in IncRNA can influence gene expression of overlapping or adjacent loci in cis by transcriptional interference, without a functional role for the IncRNA itself. This appears to be the case with antisense transcript Tsix, which negatively correlates with Xist expression. Second, IncRNAs that associate with chromatin could potentially serve as scaffolds for the recruitment of other regulating factors, including histone-modifying proteins. This is well-illustrated by the RNAi-independent mechanism of meiotic gene silencing in S. pombe, in which mRNAs encoding meiotic genes recruit silencing and heterochromatin machinery to their genomic loci.

In fission yeast, transcripts mapping to a handful of meiotic genes (mei4, ssm4, rec8, spo5) accumulate at the onset of meiosis but are undetectable in vegetative cells due to the activity of an RNA elimination pathway. YTH family protein Mmi1 mediates this silencing by associating with hexanucleotide sequences that exist within a section of each meiotic mRNA termed 'determinant of selective removal' (DSR) (Harigaya et al., 2006; Yamashita et al., 2012). Mmi1-mediated silencing requires the conserved exosome degradation complex, a poly(A)
polymerase and poly(A) binding protein, and zinc-finger protein Red1 (Harigaya et al., 2006; Sugiyama and Sugioka-Sugiyama, 2011; Yamanaka et al., 2010). Red1 physically interacts with H3K9 methyltransferase Clr4, resulting in enrichment of H3K9me2 at mei4 and ssm4 (but not rec8 or spo5) that is lost when cells undergo nitrogen starvation (inducing meiosis) (Zofall et al., 2012). Surprisingly, deletion of clr4+ does not affect transcript levels or Pol II occupancy at mei4 or ssm4, suggesting that the H3K9me2 does not truly reflect “heterochromatin” and may simply be vestigial (Egan et al., 2014): however, it is possible that H3K9me2 affects expression in certain conditions. Nevertheless, these studies clearly illustrate a case in which a transcript can act as a targeted scaffold for chromatin modifications. Sequences within meiotic mRNAs are recognized by adaptor protein Mmi1, resulting in recruitment of both transcript degradation machinery and histone modifier Clr4.

In conclusion, long non-coding RNAs represent an exciting new subclass of regulatory molecules. Previous studies in sRNA-mediated heterochromatin formation in fission yeast, plants, worms, flies, and mice have already shown that nascent IncRNAs serve critical roles in heterochromatin formation by serving as targeted scaffolds that are recognized as “nonself” by Argonaute/PIWI, facilitating recruitment of silencing and heterochromatin machinery to chromatin. It is likely that IncRNAs serve similar roles outside the context of RNAi, in which recognition by sequence-specific RNA-binding adaptor proteins is followed by recruitment of non-specific machinery. However, it is also possible that some IncRNAs can directly recruit factors involved in heterochromatin formation and gene expression. In this regard, various IncRNAs, including Xist and HOTAIR, have been shown to associate with PRC2, though the specificity of these interactions is unclear (Holoch and Moazed, 2015a).

**VI. Topics to be addressed in this thesis**

In this dissertation, I will address the issue of “self” versus “nonself” recognition as it relates to the mechanism of siRNA-mediated heterochromatin formation in fission yeast. As has
been demonstrated in many model systems, the ability of sRNA-based silencing systems to distinguish loci that are appropriate for silencing from loci that are inappropriate is critical to proper regulatory functions, and different systems have evolved different strategies. *C. elegans* contain a population of sRNAs mapping antisense to all germline-expressed genes that associates with a specific Argonaute and protects those genes from silencing. In plants, flies, and mice, siRNA-producing loci are clearly defined by association of certain factors and separated from targeted active loci. In fission yeast, we do not observe “protection” sRNAs mapping to antisense expressed genes, and in contrast to plants, flies, and mice, we observe a significant barrier to siRNA-mediated silencing in *trans*. This barrier is likely to reflect protective “self” recognition mechanisms in fission yeast, which I seek to further dissect by asking two questions. First, what factors govern the ability of a locus to produce siRNAs? Second, what features render mRNA-coding loci refractory to siRNA-mediated silencing and heterochromatin formation in *trans*? In Chapter 2, I will present evidence indicating that genome-wide convergent transcription results in generation of dsRNA that is capable of serving as an Rdsp1-independent Dcr1-substrate for siRNA generation. However, the cell enforces at least two levels of control to ensure that genome-wide siRNA generation does not occur or induce misregulated generation of silenced heterochromatin regions. At the first level, cells maintain low levels of Dcr1, so that siRNA generation only occurs at appropriate loci. At the second level, cells enforce various pathways to make euchromatic loci refractory to *cis*- and *trans*-regulated RNAi activity. In Chapter 3, I will show that 3’ end processing signals in the 3’ UTR of a transcript inhibit both *cis* and *trans* siRNA-mediated silencing and heterochromatin formation of the locus. In Chapter 4, I will present data demonstrating that deletion of the mRNA export factor Mlo3 permits formation of large domains of *trans* siRNA-mediated H3K9me2 that rarely correlate with silencing. In cells that do not silence the targeted locus, H3K9me2 is not heritable through meiosis. Establishment of silencing correlates with increased H3K9 di- and tri-methylation at the immediately targeted locus and the ability to inherit the silenced state through meiosis. Together, this data indicates
that cells have evolved many redundant mechanisms to distinguish “self” from “non-self” loci, thereby restricting siRNA activity to appropriate targets and protecting the protein-coding portion of the genome.
VII. References


recruitment of the RNA elimination machinery to silent meiotic genes. Open Biology 2, 120014-120014.


Chapter 2

Overexpression of Dcr1 reveals genome-wide overlapping transcription and bypasses the requirement for Rdp1 in RNAi-mediated silencing and heterochromatin

Ruby Yu and Danesh Moazed
This chapter consists primarily of a portion of a published report:


All experiments and analyses were performed by Ruby Yu with the exception of ChIP-qPCR assays in Figure 2.3, which were performed by Gloria Jih, and co-immunoprecipitation experiments in Figure 2.4, which were performed by Nahid Iglesias. In addition, ChIP-seq libraries, appearing in Figures 2.3, 2.5, and 2.6, were made by Gloria Jih, then processed and analyzed by Ruby Yu. The original manuscript was written by Ruby Yu and Danesh Moazed. The introduction, results and discussion were modified for this dissertation.
Abstract

Endogenous small interfering RNAs (siRNAs) and other classes of small RNA provide the specificity signals for silencing of transposons and repeated DNA elements at the post-transcriptional and transcriptional levels. However, the determinants that define an siRNA-producing region or control the silencing function of siRNAs are poorly understood. Here we show that convergent antisense transcription and availability of the Dicer ribonuclease are the key determinants for primary siRNA generation. Euchromatic convergent transcription units can generate overlapping transcript that are processed by Dicer into siRNAs, but heterochromatin formation is restricted to pericentromeric DNA repeats, due in part to mechanisms that will be elucidated in Chapter 3. Surprisingly, Dicer makes dual contributions to heterochromatin formation, promoting histone H3 lysine 9 methylation independently of its catalytic activity, in addition to its well-known role in catalyzing siRNA generation. Our results reveal distinct mechanisms that limit siRNA generation to centromeric DNA repeats and prevent spurious siRNA-mediated silencing at euchromatic loci.
Introduction

RNA-based mechanisms silence the expression of transposons and foreign DNA sequences in eukaryotic organisms ranging from yeast to human (Ghildiyal and Zamore, 2009; Moazed, 2009; Sabin et al., 2013). This silencing is associated with the generation of small RNA (sRNA) molecules that act at both the post-transcriptional and transcriptional levels. The generation of the sRNA silencing trigger is associated with specific transcription events but the mechanisms that distinguish between foreign or aberrant transcription events and normal cellular transcription are poorly understood. These mechanisms are fundamentally important because inappropriate generation of sRNA could lead to spurious silencing of essential genes. Cells must therefore have evolved mechanisms that tightly regulate sRNA biogenesis and function.

In the fission yeast *Schizosaccharomyces pombe*, centromeres are surrounded by DNA repeats that are thought to be transposon remnants (Ekwall, 2007; Rhind et al., 2011). These repeats are assembled into heterochromatin, which contributes to de novo centromere assembly and proper chromosome segregation (Allshire et al., 1994; Allshire et al., 1995). Transcription of the repeats gives rise to long noncoding RNAs, termed cenRNAs, which are processed into small interfering RNAs (siRNAs) by the RNAi machinery (Reinhart and Bartel, 2002; Verdel et al., 2004; Volpe et al., 2002). The siRNAs load onto the Ago1 subunit of RNA-Induced Transcriptional Silencing (RITS) complex, which also contains the GW-domain protein Tas3 and the chromodomain protein Chp1 (Verdel et al., 2004). RITS then targets the centromeric repeats via base pairing between its siRNAs and nascent cenRNAs, leading to recruitment of the Clr4 methyltransferase complex (CLRC), methylation of histone H3 lysine 9 (H3K9), and recruitment of HP1 proteins and histone deacetylases (Bayne et al., 2010b; Gerace et al., 2010; Motamedi et al., 2004; Noma et al., 2004). H3K9 methylation also creates a binding pocket for Chp1 and stabilizes RITS association with chromatin (Noma et al., 2004). This step is critical for the recruitment of the RNA-Dependent RNA polymerase Complex (RDRC) and
synthesis of the double stranded RNA (dsRNA), which is subsequently processed into siRNA by the Dicer (Dcr1) ribonuclease (Motamedi et al., 2004; Sugiyama et al., 2005). Consistent with their role in mediating heterochromatin formation, the RITS and RDRC complexes localize to foci in the nucleus that overlap heterochromatin proteins such as Swi6 (Colmenares et al., 2007; Noma et al., 2004; Sadaie et al., 2004). Moreover, although endogenous Dcr1 has not been detected by immunofluorescence, Dcr1 is physically associated with RDRC and when moderately overexpressed localizes to the nuclear periphery (Colmenares et al., 2007; Emmerth et al., 2010). It has remained unclear how localization of Dcr1 to the nuclear periphery contributes to its siRNA generation function.

Initiation of the cascade of events leading to RDRC-dependent siRNA amplification and heterochromatin formation requires a primary small RNA trigger. The analysis of centromeric transcription patterns and high-throughput sequencing of small RNA libraries has provided some clues into the origin of the primary small RNA. The outermost centromeric repeats are transcribed in both directions to produce overlapping complementary cenRNAs (Volpe et al., 2002). These transcripts can potentially base pair to provide a source of dsRNA that can be processed into primary siRNA by Dcr1. In addition, cenRNAs may form stem-loop structures that are cleaved by Dcr1 (Djupedal et al., 2009). If such primary siRNAs were produced by Dcr1, the levels of sRNAs should be higher in rdp1Δ than in dcr1Δ cells, since the fission yeast Rdpl is only required for siRNA amplification or spreading into neighboring sequences but does not appear to produce small RNAs. Previous high-throughput analysis of sRNA libraries have failed to detect any difference in sRNA levels between rdp1Δ and dcr1Δ cells, suggesting that in fission yeast Dcr1 does not produce primary Rdpl-independent siRNAs (Halic and Moazed, 2010). On the other hand, analysis of the sRNA sequencing data revealed a population of Dcr1- and Rdpl-independent sRNA, called primal RNA (priRNA) (Halic and Moazed, 2010). priRNAs have the same size and 5’ nucleotide preference as siRNAs and associate with Ago1, suggesting that they may perform the sRNA trigger function. However, priRNAs map to the
entire transcriptome, making it unlikely that their generation alone triggers the downstream events of siRNA amplification and heterochromatin formation.

In this study we use a combination of approaches to address how primary siRNAs are generated and what limits their ability to silence target sequences. We show that a class of Dicer-dependent primary siRNAs can be detected in *S. pombe* and provide evidence that convergent overlapping transcription gives rise to genome-wide dsRNA that can be processed into primary siRNA by Dicer. The biogenesis and function of primary siRNAs is controlled at multiple levels. First, Dicer availability controls the processing of dsRNA to primary siRNA as genome-wide siRNA production is observed when Dicer is overexpressed. Second, the ability of endogenous primary siRNAs to induce heterochromatin formation is restricted to pericentromeric DNA regions, suggesting that features of cenRNA or centromeric DNA repeats contribute to the ability of primary siRNAs to induce heterochromatin. Finally, we show that Dcr1 physically associates with the CLRC methyltransferase complex and plays a structural role in centromeric H3K9 methylation, distinct from its more commonly known functions in siRNA generation and silencing. Together, these results provide new insight into the overlapping pathways that regulate siRNA biogenesis and siRNA-mediated silencing, and demonstrate roles for Dicer in heterochromatin formation beyond siRNA generation.

**Results**

**Detection of Rdpl-independent primary siRNAs**

Previous analysis of high-throughput small RNA (sRNA) sequence reads did not reveal a convincing difference in the distribution of Ago1-associated centromeric sRNAs between *rdp1Δ* and *dcr1Δ* cells (Halic and Moazed, 2010). However, increases in the depth of Illumina sequencing technology have allowed us to greatly increase genome coverage in small RNA libraries, so that whereas previous libraries contained between 1 and 6 million aligned reads, our current libraries contained between 20 and 45 million aligned reads. With this increased
Figure 2.1: Fission yeast pericentromeric repeats produce primary siRNAs that increase in abundance with overexpression of Dcr1.

(A) Ago1-associated small RNA (sRNA) reads obtained by high-throughput sequencing of small RNA libraries using the Illumina platform, mapped to the left half of the centromere of chromosome 1 (cen1) and normalized by total number of reads (in reads per million, y axis). Note the 100-fold difference in scale of y-axis between wildtype and mutant libraries. Chromosome coordinates and the location of the centromere (central core, cnt), the innermost (imr) and outermost dg and dh repeats are indicated above and below the sRNA reads, respectively.

(B) Sum of normalized sRNA reads per library mapping to the dg or dh centromeric repeats of all three chromosomes.

(C) Western blot showing that when grown in medium lacking thiamine, Flag-Dcr1 controlled by the nmt1 promoter is overexpressed about 100-fold relative to endogenously tagged Flag-Dcr1. When grown under repressive conditions (15μM thiamine), overexpression is reduced to less than 20-fold.

(D, E) sRNA reads from sequencing of total small RNA libraries prepared from rdp1Δ cells transformed with a control (rdp1Δ+v) or with a Dcr1 overexpression plasmid (rdp1Δ+Dcr1OE) mapping to dg1L and dh1L, respectively. Total genome-wide proportion of sRNAs mapping to dg and dh, respectively, are on the right.
coverage, we found that Ago1-associated centromeric sRNA were 4-fold more abundant in
rdp1Δ cells than dcr1Δ cells (Figure 2.1A-B). We therefore conclude that a population of Dcr1-
dependent but Rdp1-independent primary siRNAs is present in fission yeast.

**Dicer overexpression boosts the levels of primary siRNAs and restores silencing in the absence of RDRC**

Previous studies have indicated that Dcr1 is a limiting factor in siRNA generation. Dcr1 overexpression boosts the levels of hairpin- and centromere-derived siRNAs and partially rescues silencing and H3K9 methylation in rdp1Δ cells (Iida et al., 2008; Kawakami et al., 2012). To understand how the dsRNA substrate for overexpressed Dcr1 is generated, we performed genome-wide deep sequencing of small RNA libraries from these cells. Dicer was expressed on a plasmid under the control of the strong nmt1 promoter, which is maximally induced in the absence of thiamine (Maundrell, 1990). When grown under non-repressive conditions, overexpressed Dcr1 levels were about 100-fold higher than endogenous Dcr1 (Figure 2.1C). Overexpression of Dcr1 in rdp1Δ cells resulted in a 20-fold increase in centromeric dg and dh primary siRNA levels (Figure 2.1D-E). These observations suggest that Dcr1 availability limits the production of primary siRNAs from dsRNA, which in turn is generated by base pairing between forward and reverse centromeric dg and dh transcripts.

Consistent with its effect on siRNA levels and previous findings (Kawakami et al., 2012), overexpression of wildtype Dcr1 partially rescued the loss of silencing phenotype at a centromeric *ura4*+ transgene (*otr1R::ura4*) in cells lacking RDRC subunits *rdp1*+, *cid12*+, or *hrr1*+ (Figure 2.2A-C). Silencing was assessed by growth on drug 5FOA, which is toxic to cells expressing *ura4*+. To test whether this phenomenon requires Dcr1 catalytic activity, we also overexpressed Dcr1 with amino acid substitutions at either one (Dcr1-D937A, hereafter referred to as Dcr1-1A) or both (Dcr1-D937A-D1127A, hereafter referred to as Dcr1-2A) RNase III catalytic sites (Figure 2.2A) (Colmenares et al., 2007). While Dcr1-1A maintains *in vitro* dsRNA
Figure 2.2 Overexpression of Dcr1 partially restores centromeric silencing in the absence of the RNA-dependent RNA Polymerase Complex (RDRC).

(A) Diagram of the Dcr1 protein and the location of functional domains and catalytic site mutations in the RNase III domains used in this study.

(B) Ten-fold serial dilutions of cells spotted on non-selective, -URA, or 5FOA-containing medium showing that Dcr1 overexpression restores silencing in *rdp1Δ* cells. Mutations in one or both RNase III domains (*Dcr1-1A* or *Dcr1-2A*, respectively) diminish or abolish the rescuing activity.

(C) Silencing assays showing that Dcr1 overexpression restores silencing in RDRC component *hrr1Δ* and *cid12Δ* cells.

(D, E) qRT-PCR assays showing that the overexpression of wildtype but not catalytically inactive Dcr1 silenced the expression of the *dg* or *dh* transcripts, respectively. Each sample was normalized to *rdp1Δ*+vector (which has a value of 1.0). P-values are based on a one sample Student’s t-test, testing that the mean is less than 1. Error bars reflect standard deviation.

cleavage activity, no cleavage activity is detected in Dcr1-2A (Colmenares et al., 2007). In contrast to wildtype Dcr1, overexpression of the mutant Dcr1 enzymes in *rdp1Δ* either restored
very weak silencing or did not restore silencing at all, respectively (Figure 2.2B). Quantitative RT-PCR (qRT-PCR) showed a partial restoration of silencing for endogenous centromeric dg and dh transcription in rdp1Δ+Dcr1OE compared to rdp1Δ cells (Figure 2.2D-E). With the exception of very weak silencing of dg transcripts in rdp1Δ+Dcr1-1A-OE cells, neither of the Dcr1 catalytic mutants restored noticeable dg or dh silencing (Figure 2.2D-E), suggesting that silencing was coupled to the increase in primary siRNA levels generated by Dcr1-mediated dsRNA cleavage.

We next performed ChIP-seq and ChIP-qPCR to determine whether H3K9 methylation was also restored by Dicer overexpression. At the pericentromeric DNA repeats, Dcr1 overexpression in rdp1Δ cells (rdp1Δ+Dcr1OE) restored H3K9 dimethylation (H3K9me2) levels to near wildtype levels, but had only a modest, if any, effect on H3K9me2 levels at subtelomeric DNA regions (Figure 2.3A-B). As previously shown, when grown in the absence of thiamine, Dcr1 under the control of the nmt1 promoter was over 100-fold overexpressed (Figure 2.1C). However, even when Dcr1 overexpression was reduced to less than 20-fold over endogenous levels by growth in thiamine, there was still considerable restoration of centromeric H3K9me2 in rdp1Δ cells (Figures 2.1C and 2.3C). Interestingly, the overexpression of Dcr1-1A in rdp1Δ cells also resulted in increased levels of H3K9me2 despite the fact that the single catalytic mutant form of Dcr1 did not support efficient silencing (Figure 2.3D and 2.2D-E). Since this Dcr1 mutant has been shown to retain in vitro dsRNA cleavage activity (Colmenares et al., 2007), we tested whether overexpression of a Dcr1 with mutations in both RNase III domains (Dcr1-2A) could induce H3K9me2 in rdp1Δ cells. Surprisingly, overexpression of the catalytically dead Dcr1 also induced a ~4-fold increase in centromeric H3K9me2 in rdp1Δ cells, close to the levels induced by overexpression of the wildtype protein (~5.7 fold increase) (Figure 2.3D). Notably, while wildtype Dcr1 was consistently able to induce strong H3K9me2 when overexpressed in rdp1Δ cells, there was considerable clonal variation for either of the Dcr1 catalytic mutants, likely due to low levels of centromeric siRNAs in these backgrounds. This variation was also reflected in
Figure 2.3 Dcr1 plays a structural role in promoting H3K9 methylation independently of its catalytic activity.

(A, B) ChIP-seq experiments showing H3K9 dimethylation (H3K9me2) levels at the pericentromeric regions and the left telomere of chromosome 1 for the indicated wildtype and mutant cells, respectively. H3K9me2 ChIP libraries were sequenced on the Illumina platform and normalized to reads per million (y axis). Chromosome coordinates are indicated above the reads. See Figure 2.1 legend for centromere abbreviations. tlh1, telomeric RecQ type helicase.

(C) ChIP experiment showing that reduced Dcr1 overexpression for cells grown in +thiamine media is still sufficient for restoration of centromeric H3K9me2 in rdp1Δ cells.

(D) ChIP-qPCR assays showing that overexpression of wildtype and the indicated mutant Dcr1 enzymes restores H3K9 methylation at the dg repeats in rdp1Δ cells.

(E) Overexpression of wildtype or catalytically inactive Dcr1 proteins suppresses the thiaabendazole (TBZ) sensitivity of rdp1Δ cells. Note variable rescue of TBZ sensitivity in cells overexpressing mutant Dcr1. Cells were spotted in 5-fold serial dilutions on medium lacking leucine, with or without 17mg/L TBZ.

(F) ChIP assays showing that endogenous wildtype Dcr1 is required for rescue of H3K9me2 by overexpression of catalytically inactive Dcr1 in rdp1Δ cells.

(G) qRT-PCR assays showing that overexpression of Dcr1 or Rdp1 in ago1Δ cells fails to induce silencing of dg transcripts.

(H) ChIP-seq experiments showing pericentromeric H3K9 dimethylation (H3K9me2) levels for ago1Δ cells with overexpression of RNAi proteins. H3K9me2 ChIP libraries were sequenced on the Illumina platform and normalized to reads per million (y axis). Chromosome coordinates are indicated above the reads. See Figure 2.1 legends for centromere abbreviations. tlh1, telomeric RecQ type helicase.
Figure 2.3 (Continued)
assays testing for sensitivity of cells to microtubule-inhibiting drug thiobendazole (TBZ) (Figure 2.3E). While overexpression of wildtype Dcr1 regularly rescued sensitivity to TBZ in rdp1Δ cells, overexpression of Dcr1 mutants resulted in no rescue, slight rescue, or very strong rescue to TBZ sensitivity in rdp1Δ cells (Figure 2.3E). These observations suggest that Dcr1 plays a non-catalytic structural role in mediating H3K9 methylation and functional heterochromatin, which under conditions of overexpression can bypass the requirement for Rdp1. Importantly, overexpression of either Dcr1 catalytic mutant did not restore centromeric H3K9me2 in rdp1Δ dcr1Δ double mutant cells, indicating that the low levels of siRNA produced by endogenous Dcr1 were required for the restoration of Rdp1-independent H3K9me2 (Figure 2.3F). Consistent with this observation, overexpression of Dcr1 in ago1Δ cells did not result in restoration of silencing or increased centromeric H3K9me2 (Figure 2.3G-H). Together, these results demonstrate that the structural role of Dcr1 still depends on Dcr1-generated small RNAs and Ago1.

To further explore the non-catalytic role of Dcr1 in heterochromatin formation, we performed co-immunoprecipitation (Co-IP) experiments testing for interaction between Dcr1 and components of the Clr4-containing CLRC complex. We found that overexpressed Dcr1 co-immunoprecipitated with CLRC subunits Clr4 and Raf1, and the interaction between overexpressed Dcr1 and Raf1 did not require Rdp1 (Figures 2.4A-B). Endogenous Dcr1 also co-immunoprecipitated with Raf1, demonstrating that the physical association between Dcr1 and CLRC occurred without Dcr1 overexpression (Figure 2.4B, right side). Finally, we found that overexpressed Dcr1 and Dcr1-2A, the catalytically inactive Dcr1 mutant, both co-immunoprecipitated with the CLRC subunit Rik1 (Figure 2.4C). These results support a model in which overexpressed Dcr1 localizes to centromeres in an siRNA- and Ago1-dependent manner and contributes to CLRC recruitment and H3K9 methylation independently of its dsRNA cleavage activity.
Figure 2.4. Dcr1 physically interacts with CLRC components.

(A) Co-immunoprecipitation assays showing that overexpressed TAP-Dcr1 associates with Flag-tagged CLRC subunits Raf1 and Clr4.

(B) Co-immunoprecipitation assays showing that overexpressed TAP-Dcr1 interacts with Flag-Raf1 in the absence of Rdp1 (left) and that endogenously expressed TAP-Dcr1 interacts with Flag-Raf1 (right).

(C) Co-immunoprecipitation assays showing that overexpressed Flag-tagged Dcr1 and Dcr1-2A associate with myc-tagged CLRC subunit Rik1.

To determine whether the restoration of silencing upon Dcr1 overexpression in rdp1Δ cells still required the RITS complex and whether restoration of silencing could occur by overexpression of other RNAi enzymes, we individually overexpressed each of Dcr1, Ago1, or Rdp1 in rdp1Δ, dcr1Δ, and ago1Δ cells. No silencing of centromeric transcripts was observed except for overexpression of Dcr1 in rdp1Δ cells (Figure 2.3G and 2.5A-B). In contrast, ChIP-seq analysis of H3K9 methylation showed that overexpression of RNAi proteins in different RNAi mutant backgrounds commonly affected centromeric H3K9me2 levels to some degree (less than 2-fold increase), suggesting that the requirements for H3K9me2 levels are not as tightly restricted as they are for transcriptional silencing (Figure 2.3H and 2.5C-E).
Figure 2.5 The effect of overexpression of RNAi factors on silencing and pericentromeric H3K9 methylation.

(A-B) qRT-PCR assays showing that upon overexpression of Dcr1, Rdpl, or Ago1 in dcr1Δ or rdplΔ backgrounds, only Dcr1 overexpression in rdplΔ induces silencing of dg transcripts.

(C) Summation of all reads mapping to centromeres in indicated methylated H3K9 ChIP-seq libraries, in reads per million.

(D-E) ChIP-seq experiments showing pericentromeric H3K9 dimethylation (H3K9me2) levels for wildtype and the indicated mutant cells with overexpression of different RNAi proteins. H3K9me2 ChIP libraries were sequenced on the Illumina platform and normalized to reads per million (y axis). Chromosome coordinates are indicated above the reads. See Figure 2.1 legends for centromere abbreviations. tlh1, telomeric RecQ type helicase.
Figure 2.6. Dcr1 overexpression results in the generation of sense and antisense primary siRNAs at a subset of euchromatic transcription units.

(A) Summary of sRNA sequencing results showing the effect of Dcr1 overexpression in rdp1Δ cells on the percentage of sRNA reads mapping complementary to protein-coding RNA (mRNA), noncoding RNA (ncRNA), ribosomal RNA (rRNA), transfer RNA (tRNA), and small nuclear RNA (snRNA).

(B) Summary of sRNA sequencing results showing the effect of Dcr1 overexpression on the percentage of convergent or non-convergent transcription units with antisense sRNAs. Convergent units are classified as genes convergent with genes or genes convergent with ncRNAs.

(C-F) Examples of transcription units with elevated sense and antisense primary siRNA levels upon Dcr1 overexpression. Blue peaks indicate siRNA sequencing reads, green peaks indicate H3K9me2 ChIP-seq reads.
Figure 2.6 (Continued)
Dicer overexpression results in genome-wide generation of primary siRNAs from convergent or overlapping transcription units

In our analysis of the sRNA sequencing data in Dcr1-overexpression libraries, we observed a striking global upregulation of siRNAs mapping antisense to mRNA-coding genes, and structural RNAs, rRNA, tRNA, and snRNA (Figure 2.6A). Upon Dcr1 overexpression in rdp1Δ cells, about 47% of convergent mRNA-coding gene pairs and 43% of mRNA-coding genes convergent with noncoding RNAs (ncRNAs) produced some antisense siRNAs, compared to 3.3% and 1.9%, respectively, in the absence of Dcr1 overexpression (Figure 2.6B). This association between siRNAs and convergent transcription suggests that overlapping sense and antisense transcripts from these loci produce dsRNA that is cleaved by overexpressed Dcr1. Interestingly, when Dcr1 was overexpressed, antisense siRNAs also mapped to 21% of non-convergent genes in rdp1Δ cells, compared to 0.7% when Dcr1 was not overexpressed. However, one quarter (102) of non-convergent genes that produced antisense siRNAs were tandem to a convergent gene that also expressed antisense siRNAs, suggesting the possibility that a long antisense read-through transcript produced dsRNA. This increase in genome-wide siRNAs may be due to increased Dcr1 availability upon overexpression. Fission yeast centromeres cluster near the spindle pole body, and moderately expressed Dcr1 associates mainly with the nuclear periphery (Funabiki et al., 1993). We propose that this co-localization contributes to the preferential centromere-specific generation of primary siRNAs. However, when Dcr1 is overexpressed, it is present throughout the nucleus and cytoplasm and is likely to gain access to dsRNAs that are produced genome-wide.

Upon inspecting the distribution of reads at specific genomic loci, we found numerous examples of convergent or overlapping transcription units to which sense and antisense siRNAs mapped in approximately 1:1 ratios (Figure 2.6C-F). Analysis of RNA levels and H3K9 methylation for genes expressing sense and antisense sRNAs did not reveal any significant increase in H3K9 methylation or reduction in RNA levels upon Dcr1 overexpression (Figure
2.6C-F, data not shown). Therefore, unlike centromeric primary siRNAs, euchromatic siRNAs, which are present at levels comparable to centromeric primary siRNAs, are unable to induce silencing or heterochromatin formation.

Discussion

The findings presented here provide insight into the mechanisms that determine which genomic loci can produce primary siRNAs and the factors that determine whether the primary siRNA signal is amplified and mediates heterochromatic gene silencing. Our results indicate that convergent and overlapping transcription at the pericentromeric repeats and throughout the genome provides a source of dsRNA that can give rise to primary siRNAs. The availability of Dicer determines the levels of primary siRNAs within the centromeric repeats and limits siRNA production at euchromatic loci. Surprisingly, Dicer also makes contributions to histone H3K9 methylation independently of its siRNA-producing enzymatic activity, likely as a structural component that contributes to the recruitment of the CLRC methyltransferase complex. Based on these findings and the previously established localization of Dicer and centromeres to the nuclear periphery, we propose a model for the specific generation of primary siRNAs from pericentromeric transcripts based on co-localization of Dicer and its RNA substrate to the nuclear periphery.

Biogenesis of primary siRNA

The mechanisms that determine which transcripts give rise to siRNA play a fundamental role in defining how RNAi regulates genome organization and function. In S. pombe, transcripts from the centromeric dg and dh repeats, and related sequences at the mating type locus and subtelomeric DNA regions, are precursors for siRNA synthesis (Buhler et al., 2008; Cam et al., 2005; Halic and Moazed, 2010; Reinhart and Bartel, 2002; Verdel et al., 2004). In addition to Dicer, siRNA generation from these loci requires several complexes that include the RNA-
dependent RNA polymerase complex (RDRC), the Ago1-containing RITS complex, and the Clr4 methyltransferase complex (CLRC) (Buhler et al., 2006; Motamedi et al., 2004; Noma et al., 2004; Verdel et al., 2004). Models of siRNA amplification propose that the RITS complex targets the noncoding dg and dh RNAs and recruits RDRC to initiate siRNA amplification. However, this targeting requires the presence of a trigger sRNA bound to Ago1 in the RITS complex that must base pair with the target RNA (Halic and Moazed, 2010). We previously described a population of Dcr1-independent small RNAs, called priRNAs, which could potentially perform this trigger function (Halic and Moazed, 2010). In this study, we achieved a deeper coverage of Ago1-associated small RNAs and were able to detect a population of Dcr1-dependent but Rdpl-independent sRNAs. We conclude that these sRNAs define a population of primary siRNAs that result from Dcr1-mediated cleavage of dsRNAs produced from the inter-molecular base pairing of forward and reverse centromeric dg and dh RNAs. Although an alternative model, involving Dcr1-mediated cleavage of intra-molecular dg or dh stem-loop structures (Djupedal et al., 2009; Kawakami et al., 2012), cannot be ruled out, the even distribution of primary siRNAs along both strands of the transcribed regions of dg and dh and the similarity of this distribution to that of secondary Rdpl-amplified siRNAs in wildtype cells leads us to favor the former model (Figure 2.1). Thus S. pombe contains two distinct populations of Dcr1-dependent and Dcr1-independent small RNAs that can act as triggers for further siRNA amplification and heterochromatin formation (Figure 2.7).

We previously showed that Dcr1 overexpression results in increased siRNA generation from a ura4+ hairpin construct, suggesting that Dcr1 is limiting for siRNA generation even when an abundant dsRNA substrate is present (Iida et al., 2008). The results presented here show that Dcr1 overexpression boosts primary siRNA levels by about 20-fold and partially suppresses the requirement for all three subunits of the RDRC complex in centromeric silencing. Consistent with these results, a recent report showed that Dcr1 overexpression restored silencing in rdplΔ
and $dsh1\Delta$ cells (Kawakami et al., 2012). Dsh1 is a putative inner nuclear membrane protein that may help localize Dcr1 to centromeric transcripts by a mechanism that remains to be

**Figure 2.7 Pathways that mediate siRNA generation and control siRNA-mediated heterochromatin formation.**
Overlapping or antisense transcription at the pericentromeric DNA repeats results in dsRNA generation, which is preferentially processed into siRNA because the centromeres localize to the nuclear periphery, where Dicer (Dcr1) concentrations are high. In addition to primary siRNAs generated by this pathway, primal RNAs (priRNA) are generated from single stranded centromeric RNAs and loaded onto Ago1. Dcr1 physically associates with the CLRC H3K9 methyltransferase complex and also plays a structural role in heterochromatin formation by facilitating CLRC recruitment (denoted as Dcr1*). RDRC, RNA-Dependent RNA polymerase Complex; ARC, Argonaute Chaperone; RITS, RNA-Induced Transcriptional Silencing; CLRC, Clr4 -Rik1-Cul4 complex; filled red circles, H3K9me2/3.
defined. These observations suggest that the transcription of centromeric repeats provides a source of dsRNA that can be directly processed into primary siRNAs by Dcr1, providing a trigger for secondary siRNA amplification (Figure 2.7). However, as discussed below, our results also suggest that Dcr1 plays distinct catalytic and non-catalytic roles in promoting RNAi-dependent heterochromatin formation.

Further support for the idea that overlapping transcription provides a source of dsRNA for Dicer comes from the observation that Dcr1 overexpression results in transcriptome-wide generation of siRNAs from convergent or overlapping transcription units. In Dcr1 overexpressing cells, in which Dcr1 is not restricted to the nuclear periphery (Emmerth et al., 2010), we observe antisense primary siRNAs that map to approximately half of all protein-coding genes that are convergent either with other protein coding genes or contain annotated overlapping antisense RNAs (Figure 2.6B). These results strongly suggest that pervasive dsRNA, resulting from base pairing overlapping complementary transcripts, is formed in *S. pombe* but is normally not processed into siRNA because Dcr1 levels and localization are strictly regulated. The *S. pombe* Dcr1 activity is regulated at multiple levels via its association with heterochromatin through interactions with the RITS and RDRC complexes (Colmenares et al., 2007), its localization to the nucleus and nuclear periphery (Emmerth et al., 2010), and its direct or indirect interaction with the nuclear pore complex and the Dsh1 protein (Kawakami et al., 2012). However, the significance of these interactions for siRNA generation has been unclear. Since endogenous fission yeast Dcr1 preferentially processes centromeric RNAs into primary siRNAs (Figure 2.1), we propose that the localization of both Dicer and centromeric repeats to the nuclear periphery limits siRNA generation and amplification to the subset of overlapping RNAs that are transcribed from centromeric repeats. The mechanisms that mediate the localization of centromeres to the nuclear periphery remain to be determined but are likely to involve interactions with inner nuclear membrane proteins.
We did not observe a significant overlap between the genes that produce antisense
primary siRNAs in our experiments and the recently reported euchromatic genes that generated
siRNAs in rrp6D cells (Yamanaka et al., 2013), the convergent gene pairs that establish
transient heterochromatin in G1 phase (Gullerova et al., 2011), or the euchromatic Dcr1-
associated loci (Woolcock et al., 2011). Therefore, additional locus-specific mechanisms are
likely involved in controlling siRNA generation.

Catalytic and structural functions of Dicer in centromeric silencing

Our analysis of the effects of Dcr1 overexpression on histone H3K9 methylation and
silencing reveal unexpected catalytic and structural roles for Dcr1 in promoting silencing and
H3K9 methylation. Although only wildtype Dcr1 overexpression restores silencing in rdp1Δ cells
(Figures 2.1, 2.2), the overexpression of a catalytically dead Dcr1 in rdp1Δ cells results in a
substantial increase in H3K9 methylation levels (Figure 2.3D). Remarkably, H3K9 methylation
levels induced by the catalytically dead Dcr1 are only 20% lower than the methylation that is
induced by overexpression of wildtype Dcr1. These results give rise to two important
conclusions. First, they suggest that Dcr1 can play a structural role in promoting H3K9
methylation independent of its siRNA-producing activity and Rdp1. We note that the ability of
the catalytically dead Dcr1 to promote H3K9 methylation in rdp1Δ cells relies on Ago1 and
endogenous Dcr1, therefore Dcr1 is likely to contribute to the recruitment of the CLRC complex
to nascent cenRNAs in steps downstream of siRNA generation (Figure 2.7). Second, they
suggest that the catalytic activity of Dcr1 plays an important role in silencing beyond promoting
H3K9 methylation. We propose that this role involves the co-transcriptional degradation of RNA
that is transcribed from heterochromatic regions, the byproduct of which may be siRNAs.

Silencing of euchromatic loci by small RNAs
Previous studies have linked convergent transcription to siRNA-mediated silencing and heterochromatin formation. Gullerova et al. showed that certain convergent gene pairs establish transient heterochromatin during G1-S phase, while Iida et al. showed that silencing of a euchromatic locus by hairpin-generated siRNAs correlates with antisense transcription (Gullerova and Proudfoot, 2008; Iida et al., 2008). However, euchromatic primary siRNAs, in cells that overexpress Dcr1, do not appear to induce silencing of their corresponding loci. In rdp1Δ cells overexpressing Dcr1, primary siRNA densities at many euchromatic convergent transcription units were comparable to siRNA levels at centromeric repeats, but these loci (in unsynchronized cells) lack H3K9 methylation and are not silenced (Figure 2.6). Therefore, while convergent transcription units have the potential to generate primary siRNAs, these loci are refractory to siRNA-mediated silencing when compared to pericentromeric repeats. This observation suggests that euchromatic loci possess features that inhibit siRNA-mediated silencing or lack features that permit siRNA-mediated silencing. As discussed in the next chapter, one inhibitory mechanism involves the presence of 3’ end processing signals in euchromatic transcripts.
References


Chapter 3

Transcriptional cleavage and polyadenylation signals inhibit silencing by RNAi and ectopic siRNAs

Ruby Yu and Danesh Moazed
Chapter 3

This chapter consists primarily of a portion of a published report:


All experiments and analyses were performed by Ruby Yu. The original manuscript was written by Ruby Yu and Danesh Moazed. The introduction, results and discussion were modified for this dissertation.
Abstract

Small RNAs serve as important and adaptive specificity factors both in regulation of gene expression and in genome defense. In fission yeast, short interfering RNAs (siRNAs) play a critical role in silencing at the constitutively heterochromatic pericentromeric repeats; siRNAs and methylated histone H3K9, a conserved marker for heterochromatin, exhibit a mutual interdependence at these loci. In a mechanism that is not yet fully understood, the ability of siRNAs to efficiently nucleate heterochromatin and silence a target locus is restricted to sequences bearing similarity to those of the pericentromeric repeats, and attempts to induce silencing or heterochromatin formation at a euchromatic locus have met with limited success. In this chapter, we show that transcriptional polyadenylation and cleavage sequences within the 3’ UTR of reporter gene ura4+ inhibit the ability of the RNAi machinery to target the locus for silencing and heterochromatin formation. Furthermore, direct sequencing of polyadenylated transcripts reveals that centromeric RNAs are polyadenylated, but centromeric sites of cleavage and polyadenylation are widely distributed and poorly used, in contrast with those at mRNAs, which are limited in number and highly used. Together, these data suggest that the efficiency of cleavage and polyadenylation of nascent transcripts may be a key factor in distinguishing loci that can be silenced by siRNAs from those that cannot, and implicate transcript residence time as a likely determinant of successful targeting by RNAi.
Introduction

In eukaryotes, RNAi is a central RNA regulatory mechanism that has been identified in plants, fungi, worms, flies, and mammals. While RNAi-based silencing occurs via a variety of mechanisms, they are all defined by the association of 20-30 nucleotide small RNAs with a member of the PAZ- and PIWI-domain-containing Argonaute family of proteins (Hutvagner and Simard, 2008). These small RNAs, which include microRNAs (miRNAs), short interfering RNAs (siRNAs), and Piwi-interacting RNAs (piRNAs), bind the Argonaute protein and guide it to complementary sequences targeted for silencing.

In classical post-transcriptional gene silencing (PTGS), long double stranded RNAs (dsRNAs) are cleaved by a member of the RNaseIII family called Dicer to produce 21-25 nucleotide dsRNAs. One strand of the dsRNA product is loaded onto Argonaute at the core of the RNAi- induced silencing complex (RISC) (Bernstein et al., 2001; Hutvagner and Simard, 2008). RISC is guided to mRNAs complementary to the guide strand by sequence-specific RNA-RNA interactions and mediates silencing by transcript degradation or translational repression. Silencing can also require an siRNA-amplification cycle involving an RNA-dependent RNA polymerase (RdRP) (Sijen et al., 2001). In some organisms, including flies and worms, proteins of the RNAi pathway have also been implicated in chromatin-dependent transcriptional gene silencing. Studies in D. melanogaster ovarian somatic cells show that H3K9 methylation of transposable elements are Piwi-dependent, and targeting Piwi to an ectopic site containing piRNA-complementary sequences results in increased H3K9 methylation and silencing of the locus (Huang et al., 2013; Le Thomas et al., 2013).

In fission yeast, pericentromeric repeats are constitutively maintained as heterochromatin and silenced in an RNAi-dependent manner, and centromeric transcripts give rise to an endogenous population of siRNAs, which are required for maintenance of silencing and heterochromatin (Verdel et al., 2004; Volpe et al., 2002). Fission yeast have one gene each for RNAi components Dicer (dcr1*), Argonaute (ago1*), and RdRP (rdp1*). These have also
been shown to interact in two RNAi silencing complexes known as the RNA-induced transcriptional silencing complex (RITS) and the RNA-directed RNA polymerase complex (RDRC) (Motamedi et al., 2004; Verdel et al., 2004).

The RITS complex is composed of the Argonaut protein Ago1, a GW-repeat protein Tas3, a chromodomain protein Chp1, and siRNA (Verdel et al., 2004). RITS is guided by siRNAs to complementary nascent transcripts in centromeric regions, where Chp1 associates with H3K9me and stabilizes RITS (Buhler et al., 2006; Motamedi et al., 2004; Verdel et al., 2004). RDRC then associates with this complex in a Dcr1- and Clr4-dependent manner and uses the nascent transcript as a template to synthesize double stranded RNA (dsRNA), which is then processed by Dcr1 into siRNAs (Colmenares et al., 2007; Motamedi et al., 2004; Sugiyama et al., 2005). Double-stranded siRNAs are loaded onto another Ago1 complex, termed ARC, and passed on to RITS, completing an siRNA amplification loop (Buker et al., 2007; Holoch and Moazed, 2015). The chromatin-associated RITS also recruits the CLRC complex containing H3K9 histone methyltransferase Clr4, which mediates H3K9 methylation and subsequent Swi6 association (Bayne et al., 2010; Gerace et al., 2010; Hong et al., 2005).

As previously mentioned, ectopic targeting of Piwi to a reporter locus in *D. melanogaster* results in H3K9 methylation and transcriptional silencing (Huang et al., 2013; Le Thomas et al., 2013). In *C. elegans*, exposure to exogenous dsRNAs can trigger H3K9me3 enrichment and transcriptional silencing at complementary genomic reporter loci (Gu et al., 2012; Guang et al., 2010). However, studies involving long hairpin constructs in fission yeast show that while Dcr1 efficiently processes the hairpin dsRNA into siRNAs, which load onto the RITS complex, hairpin siRNAs can only induce heterochromatin formation at a subset of targets (Iida et al., 2008). For example, hairpin generated siRNAs fail to induce heterochromatin at a GFP reporter gene or the endogenous *ura4*+ locus, but have variable silencing activity at the *trp1*+::*ura4*+, *arg3*::*ura4*+-GFP, and *ade6*+::GFP loci (Iida et al., 2008; Sigova et al., 2004; Simmer et al., 2010). Although at the *trp1*+::*ura4*+ locus hairpin-mediated silencing correlates with antisense transcription, at the
arg3::ura4*-GFP locus hairpin-mediated silencing occurs in the apparent absence of antisense transcription. Furthermore, data presented in chapter 2 showed that although overexpression of Dcr1 in rdp1Δ cells results in generation of sense and antisense siRNAs at both euchromatic and centromeric overlapping transcription units, silencing and heterochromatin formation only occur at centromeric loci, and not mRNAs. The mechanisms that determine whether primary siRNAs can trigger the amplification of secondary siRNAs and heterochromatin formation therefore remain unclear.

In this chapter, we show that transcriptional polyadenylation and cleavage signals in the 3' untranslated region (3' UTR) of a euchromatic gene inhibit the ability of exogenous siRNAs to target the locus for silencing or heterochromatin formation, and a transcript lacking the 3’ UTR is directly silenced by RNAi machinery in the absence of exogenous siRNA expression. Furthermore, using direct sequencing of polyadenylated transcripts, we show that unlike mRNAs, which generally have a low number of well-used sites of cleavage and polyadenylation, centromeric transcripts have many poorly-used sites of cleavage, suggesting a different and weaker pattern of cleavage and polyadenylation than at mRNAs and euchromatic loci. Together, these results suggest that 3’ end processing signals may protect mRNA-coding genes from RNAi-mediated heterochromatin assembly.

Results

Presence of mRNA 3’ processing signals inhibits the ability of hairpin siRNAs to promote silencing and secondary siRNA synthesis

In Chapter 2, we showed that Dcr1 overexpression promotes primary siRNA generation from convergent overlapping centromeric transcripts and euchromatic transcription units. However, silencing and Rdpl-dependent siRNA amplification (secondary siRNA synthesis) only occur within the centromeric repeats. This suggests that some feature of mRNA coding genes protects them from RNAi silencing. In this regard, we have previously shown that ura4* at its
endogenous locus is refractory to silencing by abundant hairpin-produced siRNAs, but that ura4+ could be silenced by hairpin siRNAs at a different locus where we observed antisense transcription (Iida et al., 2008). However, Simmer et al. showed that hairpin-mediated silencing of a different reporter gene did not correlate with antisense transcription, suggesting that antisense transcription is not a required feature for a locus to be silenced by RNAi (Simmer et al., 2010). We speculated that antisense transcription might make a locus more susceptible to siRNA-mediated silencing because antisense transcription through the 3’ UTR of a gene could disrupt proper 3’ end processing and termination of the sense transcript. mRNA-coding genes contain specific sequences at their 3’-untranslated regions that may inhibit the ability of siRNAs to use them as co-transcriptional scaffolds for the recruitment of H3K9 methylation and siRNA amplification machineries. We therefore asked whether disrupting the ura4+ 3’ UTR sequence, which would likely interfere with 3’ end processing of the transcript, would make the endogenous ura4+ locus more amenable to hairpin-mediated silencing.

We generated strains with different modifications to the ura4+ endogenous locus: the ura4-UTR\textsuperscript{DIS} locus, where the 3’ UTR was disrupted by insertion of the kanMX gene between the ura4+ stop codon and its 3’ UTR; the ura4-UTR\textsuperscript{D} locus, where the 3’ UTR was deleted and replaced it with the kanMX gene; and the ura4-UTR-kanMX locus, where the kanMX antibiotic resistance marker was inserted downstream of the 3’ UTR (Figure 3.1A). Consistent with previous results (Iida et al., 2008), the ura4+ hairpin did not induce detectable silencing at the unmodified ura4+ or the ura4-UTR-kanMX loci, as measured by growth on 5FOA, which is toxic to cells expressing ura4+ (Figure 3.1B-C). However, the hairpin did induce weak silencing of the ura4-UTR\textsuperscript{DIS} allele and strong silencing of the ura4-UTR\textsuperscript{D} allele (Figure 3.1B). Not surprisingly, ura4-UTR\textsuperscript{D} showed weak growth on 5-FOA medium even in the absence of hairpin expression. Possibly, this is because the heterologous 3’ UTR from the distantly related budding yeast Ashbya gossypii at the kanMX gene is not fully functional in fission yeast (Goldstein and
Figure 3.1 The *ura4* 3' UTR inhibits hairpin siRNA-induced silencing

(A) Diagrams of the wildtype *ura4* locus and *ura4* alleles with modified 3' UTRs. *ura4*+, wildtype *ura4*; *ura4*-UTR-kanMX, *ura4* allele with the *kanMX* antibiotic resistance marker inserted downstream of the 3' UTR; *ura4*-UTR<sup>DIS</sup>, *ura4* allele in which the 3' UTR is disrupted by insertion of *kanMX*; *ura4*-UTR<sup>D</sup>, *ura4* allele in which the 3' UTR is replaced with *kanMX*. Grey boxes indicate regions of complementarity with the *ura4* hp.

(B) Silencing assays showing the effect of 3' UTR displacement or deletion on hairpin-induced silencing. Ten-fold serial dilutions of the indicated cells were plated on EMMC non-selective, -URA, or 5FOA-containing medium. *otr1R::ura4* serves as a positive control. hp, *ura4*+ hairpin.

(C) Ten-fold serial dilutions of control *ura4*+ or *ura4*-UTR-kan strains with or without *ura4*+ hairpin expression, grown on non-selective EMMC, -URA, or 5FOA-containing medium.

(D) Strand-specific qRT-PCR showing the effect of the *ura4* hairpin on *ura4* sense (top) and antisense (bottom) transcripts.

(E) Northern blot showing the expression of *ura4* RNA from wildtype or the indicated mutant alleles.
Figure 3.1 (Continued)
McCusker, 1999), or because there is transcriptional interference from antisense transcription at the \textit{ura4-UTR}\textsuperscript{D} locus (Figure 3.1A,D). Hairpin-induced silencing in \textit{ura4-UTR}\textsuperscript{DIS} cells was accompanied by a reduction in the levels of \textit{ura4}\textsuperscript{+} sense RNA (Figure 3.1D). Despite the considerable hairpin-dependent increase in silencing of the \textit{ura4-UTR}\textsuperscript{D} cells on 5-FOA-containing medium, we did not observe a corresponding decrease in \textit{ura4}\textsuperscript{+} sense transcript levels. This may be due to the low levels of \textit{ura4-UTR}\textsuperscript{D} RNA even in the absence of the hairpin and the fact that cells were grown in non-selective medium for RNA isolation. However, high levels of antisense transcripts in \textit{ura4-UTR}\textsuperscript{D} were reduced more than 5-fold upon expression of the hairpin, suggesting that the hairpin does mediate transcriptional silencing at this locus (Figure 3.1D). Interestingly, all modified \textit{ura4} alleles showed much higher levels of antisense transcription than the endogenous unmodified locus, likely due to a cryptic reverse promoter in the \textit{kanMX} gene (Figure 3.1A,D). Northern analysis showed that the \textit{ura4-UTR}\textsuperscript{DIS} and \textit{ura4-UTR}\textsuperscript{D} transcripts were longer than wildtype \textit{ura4}, likely due to fusion with the \textit{kanMX} transcript (Figure 3.1E). Finally, ChIP experiments showed that hairpin-induced silencing of the \textit{ura4-UTR}\textsuperscript{D} locus is accompanied by a reduction in the levels of antisense transcripts.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure3.png}
\caption{The \textit{ura4} 3\textsuperscript{'} UTR inhibits hairpin siRNA-induced heterochromatin formation. (A) ChIP experiments showing that the \textit{ura4}\textsuperscript{+} hairpin preferentially induces H3K9 methylation at the \textit{ura4-UTR}\textsuperscript{D} locus. H3K9 methylation was lost upon the deletion of either \textit{clr4} (\textit{clr4}\textsuperscript{Δ}) or the \textit{ura4}\textsuperscript{+} hairpin (\textit{hp}\textsuperscript{Δ}). (B) ChIP assay showing that the hairpin-dependent H3K9me2 observed at \textit{ura4-UTR}\textsuperscript{D} requires RNAi (\textit{ago1}, \textit{dcr1}, \textit{rdp1}), CLRC (\textit{clr4}) and TRAMP (\textit{cid14}) components.}
\end{figure}
UTR^D allele was accompanied by H3K9 methylation, which was lost upon deletion of the hairpin or RNAi factors (Figure 3.2A-B). These results indicate that sequences within the 3' UTR of ura4^+, possibly cleavage or polyadenylation signals, interfere with RNAi-mediated transcriptional gene silencing and provide an explanation for the inability of hairpin-produced siRNAs to silence euchromatic reporter genes at the transcriptional level. Furthermore, consistent with previous findings, the continuous dependence of ectopic silencing and H3K9 methylation on the presence of the siRNA-producing hairpin locus shows that the silent state cannot be maintained in the absence of an siRNA-producing driver locus.

![Diagram of the elements of the ura4^+ 3' UTR. Four downstream polyadenylation and termination signals are indicated: two Site Determining Elements (SDE1 and SDE2), an efficiency element (EE), and a downstream sequence element (DSE). These sites were mutated as in Birse et al., 1997 to generate strains ura4-UTR^123, ura4-UTR^4, and ura4-UTR^1234, all of which have a downstream KanMX gene as in ura4-UTR-kanMX. Superscript numbers and black boxes indicate elements that have been mutated.](image)

![Northern blot showing the expression of ura4 RNA from indicated wildtype or mutant alleles.](image)

![Silencing assays showing the effect of specific polyadenylation or termination signal mutations on hairpin-induced silencing. Ten-fold serial dilutions of the indicated cells were plated on EMMC non-selective, -URA, or 5FOA-containing medium. ura4-UTR^D serves as a positive control. hp, ura4^+ hairpin.](image)
To more specifically test the idea that polyadenylation and cleavage signals in the 3’ UTR inhibit siRNA-mediated silencing, we generated a series of directed mutations or deletions in four termination signals downstream of the ura4 open reading frame: two site-determinant elements (SDE1 and SDE2), one efficiency element (EE), and one downstream sequence element (DSE), as identified in previous studies (Birse et al., 1997; Humphrey et al., 1994) (Figure 3.3A). The elements were numbered 1 to 4, respectively, with the following mutations introduced at each: in ura4-UTR\textsuperscript{123}, a 34-base pair (bp) deletion in SDE1, a 4-bp substitution in SDE2, and a substitution of 11 bp of the efficiency element with an 8-nucleotide linker sequence; in ura4-UTR\textsuperscript{4}, deletion of the 31-bp DSE; and in ura4-UTR\textsuperscript{1234}, a combination of all the above mutations and deletions. All strains included a KanMX cassette inserted downstream of the DSE, as in construct ura4-UTR-kanMX (Figures 3.1A and 3.3A). Northern analysis showed a slightly longer transcript in ura4-UTR\textsuperscript{123} and a much longer run-on transcript in ura4-UTR\textsuperscript{1234}, comparable to that seen in ura4-UTR\textsuperscript{DIS} (Figure 3.3B). Growth assays on 5-FOA showed hairpin mediated silencing at both ura4-UTR\textsuperscript{123} and ura4-UTR\textsuperscript{1234} but not ura4-UTR\textsuperscript{4}, demonstrating that trans silencing by ectopic siRNAs correlates with loss of proper 3’ end processing and read-through transcription (Figure 3.3C).

The above results suggest that signals within the 3’ UTR region of ura4\textsuperscript{+} interfere with siRNA-mediated steps that may involve the recruitment of RDRC/Dcr1 and secondary siRNA generation, or the recruitment of the CLRC complex and H3K9 methylation, or both. In order to distinguish between these possibilities, we tested whether 3’ UTR sequences affected secondary siRNA generation by sequencing siRNA libraries of either wildtype ura4\textsuperscript{+}, ura4-UTR\textsuperscript{DIS}, or ura4-UTR\textsuperscript{D} cells that expressed the ura4\textsuperscript{+} hairpin (Figure 3.4A-D). In these experiments, secondary siRNA synthesis is expected to result from Rdg1-dependent dsRNA synthesis at the targeted locus and spreading of siRNAs to regions flanking the siRNA targeted domain. As shown in Figure 3.4B-D and summarized in Figure 3.4E, hairpin expression induced a low amount of secondary siRNA synthesis, indicated by spreading of siRNA reads along the
Figure 3.4 The 3' UTR of ura4 inhibits the generation of secondary siRNAs.

(A) Diagram of the ura4 hairpin RNA containing the cox4 intron. Blue arrows, double stranded region complementary to ura4*; gold loop, cox4 intron.

(B-D) siRNA-seq of Ago1-associated siRNAs at the wildtype ura4*, ura4-UTR^{DIS}, or ura4-UTR^{D} loci, shown in log scale. Note the spreading of siRNAs to the tam14 gene at the ura4-UTR^{D} locus.

(E) Sum of normalized siRNA-seq read densities mapping to the indicated loci, in reads per million per kilobase (rpmkpb). Total reads mapping to ura4-UTR^{D} are approximately 30-40% of those in ura4+ or ura4-UTR^{DIS} due to an unusually high number of reads mapping to one siRNA, located in rDNA, in the ura4-UTR^{D} sample.

(F) Zoomed out view of hairpin-induced secondary siRNA reads mapping to the cox4 locus (left). The levels of intronic cox4 siRNAs are comparable to the levels of siRNAs mapping to the ura4 hairpin shown in B-D.

(G) Northern blot showing hairpin-dependent isoforms containing the cox4 intron.
Figure 3.4 (Continued)
ura4 transcript. However, secondary siRNAs were more abundant at the ura4-UTR\textsuperscript{DIS} locus, and by far the most abundant at the ura4-UTR\textsuperscript{D} locus. These results indicate that signals within the ura4\textsuperscript{+} 3’ UTR region interfere with primary siRNA-mediated recruitment of the machinery that promotes secondary siRNA synthesis.

The ura4\textsuperscript{+} hairpin construct contains a 355 nucleotide intron from the cox4\textsuperscript{+} gene that forms the loop in the stem-loop hairpin structure (Figure 3.4A) (Iida et al., 2008). We observed a high proportion of reads mapping sense and antisense to the cox4 intron loop of the ura4 hairpin, at a comparable density in reads per million per kilobase (rpmPKb) to the siRNAs from the dsRNA stem of the ura4 hairpin (Figure 3.4E-F). The cox4 intron may be inefficiently spliced out from the long hairpin construct, thus becoming a substrate for RDRC. Consistent with this idea, Northern analysis showed the presence of hairpin-dependent RNAs containing the cox4 intron, suggesting that the intron was not efficiently spliced (Fig 3.4G).

We next tested the effect of mutations in RNA processing or heterochromatin proteins on siRNA synthesis at the ura4-UTR\textsuperscript{D} and cox4\textsuperscript{+} loci. At the ura4-UTR\textsuperscript{D} locus, deletion of rdp1\textsuperscript{+} resulted in loss of secondary siRNAs (Figure 3.5A-B). In contrast, deletion of clr4\textsuperscript{+} resulted in only an ~2-fold decrease in secondary siRNA synthesis (Figure 3.5A-B). Consistent with previous northern blot analysis (Iida et al., 2008), deletion of either rdp1\textsuperscript{+} or clr4\textsuperscript{+} did not diminish primary hairpin siRNA levels (Figure 3.5A). Furthermore, deletion of TRAMP component cid14\textsuperscript{+}, which results in a decrease in the levels of centromeric siRNAs (Buhler et al., 2007), resulted in a modest decrease in the levels of hairpin-derived and secondary siRNAs (Figure 3.5A-B).

At the cox4\textsuperscript{+} locus, we detected low levels of antisense siRNAs that mapped to the mature spliced cox4 RNA regions flanking the intron (Figure 3.5A,C). These tertiary siRNAs occurred with very low efficiency, probably due to two factors. First, the cox4 intron is co-transcriptionally spliced out before it can become a target for RNAi. Second, the cox4 3’UTR was intact, and 3’ end processing proceeded efficiently. As expected, the cox4 tertiary siRNAs
Figure 3.5. Deletion of RDRC, CLRC, or TRAMP subunits affects spreading of secondary siRNAs.

(A) Table summarizing primary or secondary siRNA read densities in rpmkb mapping to the indicated loci. Reads mapping genome-wide in ura4-UTR$^D$ are approximately 30-40% of those in ura4+ or ura4-UTR$^{DIS}$ due to an unusually high number of reads mapping to one siRNA located in rDNA.

(B-C) Zoomed-in view of siRNA-seq reads mapping to ura4-UTR$^D$ or cox4, respectively. Note increase in cox4 tertiary siRNAs in clr4Δ cells.
Figure 3.5 (Continued)
were lost in \textit{rdp1}\Delta cells and decreased in \textit{cid14}\Delta cells, but surprisingly, became more abundant in \textit{clr4}\Delta cells (Figure 3.5A,C). One likely explanation for this observation is that in cells lacking \textit{clr4}^*, the RNAi machinery is released from the centromere and becomes more available at the \textit{cox4} locus. The detection of siRNAs that map to the \textit{cox4} mature mRNA indicates that abundant intronic \textit{cox4} secondary siRNAs can target the \textit{cox4} pre-mRNA, most likely co-
transcriptionally, and initiate tertiary siRNA synthesis.

\textbf{A transcript lacking 3' processing signals is targeted by RNAi}

As noted above, cells carrying a \textit{ura4}^+ gene lacking a 3' UTR grow on medium containing 5-FOA, indicating that even without being targeted by hairpin \textit{ura4} siRNAs the aberrant \textit{ura4-UTR}^D RNAs are at least partially silenced. We wished to test whether the RNAi pathway contributed to this silencing, perhaps along with the well-known TRAMP/exosome pathway. Towards this goal, we first made deletions of RNAi genes \textit{ago1}^+, \textit{rdp1}^+, and \textit{dcr1}^+, as well as CLRC component \textit{clr4}^*, and \textit{cid14}^*, the non-canonical poly(A) polymerase in the TRAMP complex that targets aberrant transcripts for degradation by the exosome. Deletion of any of these genes resulted in an immediate loss of silencing at the \textit{ura4-UTR}^D locus, suggesting that both RNAi and the TRAMP/exosome pathways were required for silencing of the defective \textit{ura4-UTR}^D RNA (Figure 3.6). However, we noted that extended passaging of RNAi\Delta or \textit{clr4}\Delta (but not \textit{cid14}\Delta) cells resulted in a regaining of \textit{ura4-UTR}^D silencing, suggesting that the TRAMP/exosome pathway compensated for the loss of the RNAi-mediated mechanism of silencing.
Figure 3.6. The ura4-UTR0 transcript is silenced in an RNAi- and TRAMP-dependent manner.
Silencing assays showing that RNAi is required for silencing of a ura4 allele that lacks 3’ end processing signals (ura4-UTR0). ura4-UTR0 silencing is lost upon deletion of RNAi components (dcr1, rdp1, ago1), the H3K9 methyltransferase clr4, or the TRAMP component cid14.

We next tested whether Dcr1 overexpression could potentiate the RNAi-dependent, but ura4+ hairpin independent, silencing of ura4 alleles with defective 3’ UTRs. We observed strong silencing at ura4-UTR0 but none of the other alleles, and this silencing was maintained even when overexpression of the thiamin repressible nmt1-dcr1+ was reduced by growth in thiamine-containing medium (Figures 3.7A-C; see Figure 2.1C for Dcr1 protein levels when nmt1-dcr1+ cells were grown in thiamine-containing medium). In addition, Dcr1 overexpression resulted in strong Ago1- and Rdp1-dependent H3K9me2 at the ura4-UTR0 locus, comparable to H3K9me2 levels at centromeric reporter gene otr1R::ura4 and much stronger than that observed with the ura4 hairpin (Figure 3.7D-E). Weak H3K9me2 at the ura4-UTR0 locus was also observed when the catalytically inactive Dcr1-2A was overexpressed (Figure 3.7E). Together, these results suggest that a transcript lacking proper 3’ end processing signals is targeted for silencing and heterochromatin formation by the RNAi machinery.

We propose that silencing of the ura4-UTR0 locus, described above (Figure 3.6 and 3.7A), and silencing of the outer centromeric repeats occur via similar RNAi-mediated mechanisms. Both the ura4-UTR0 locus and centromeric repeats possess sense and antisense
Figure 3.7. Overexpression of Dcr1 induces strong hairpin-independent silencing and heterochromatin formation at ura4-UTR\(^\delta\).

(A) Silencing assay showing that Dcr1 overexpression induces stronger silencing at \textit{ura4-UTR}\(^\delta\) than overexpression of the siRNA-producing \textit{ura4} hairpin.

(B) Silencing assay showing that Dcr1 overexpression does not induce silencing at the wildtype \textit{ura4}\(^+\) or \textit{ura4-UTR-kan} loci.

(C) Silencing assay on growth medium containing 30 uM thiamine showing that the reduced expression of Dcr1 under repressive conditions is sufficient for strong silencing at \textit{ura4-UTR}\(^\delta\).

(D) ChIP experiments showing that Dcr1 overexpression induces 15-fold higher H3K9 methylation at the \textit{ura4-UTR}\(^\delta\) locus than a \textit{ura4} hairpin. Note that H3K9me at \textit{ura4-UTR}\(^\delta\) +\textit{Dcr1OE} was comparable to that of a \textit{ura4} gene inserted at an outer centromeric repeat.

(E) ChIP assays showing that overexpression of a catalytically dead Dcr1 mutant induces low levels of H3K9me2 at \textit{ura4-UTR}\(^\delta\) and that the H3K9me2 observed when wildtype Dcr1 is overexpressed requires \textit{ago1}\(^+\) and \textit{rdp1}\(^+\).
Figure 3.7 (Continued)
transcripts, and, as described below, both possess weak or abnormal polyadenylation and termination signals. We performed direct RNA sequencing of total polyadenylated transcripts in either wildtype or rdp1Δ cells and used a strategy that mapped the last nucleotide before the polyA tail. The polyA sequencing (polyA-seq) results revealed notable differences in the polyadenylation pattern of transcripts arising from the outer centromeric repeats compared to
mRNAs. We summed all reads mapping either downstream of annotated transcription units or within the dg/dh repeats of each chromosome in a wildtype strain. In general, we observed a good correlation between transcript abundance, based on microarray expression data (Dutrow et al., 2008), and the number of polyA-seq reads (Figure 3.8A). As expected, the centromeric dg and dh transcripts ranked low compared to genic transcription units and relative to their expression levels, although they contain higher polyA levels than structural RNAs, such as rRNAs and tRNAs (Figure 3.8A). Moreover, in contrast to mRNAs, which contained one or two major polyA peaks, the centromeric dg and dh transcripts contained multiple weak polyA peaks that generally mapped throughout the transcribed region that gives rise to siRNAs (Figures 3.8B-C). These observations suggest that centromeric transcripts, like the \textit{ura4-UTR}^D locus, possess weak or abnormal 3’ end processing signals, which may contribute to their susceptibility to RNAi-mediated silencing.

\textbf{Discussion}

The results in this chapter have elucidated one feature by which fission yeast protects its protein-coding genome from inappropriate RNAi-mediated heterochromatin formation. Our data suggest that mRNA 3’ end processing signals located in the 3’ UTR limit the ability of ectopically produced siRNAs and endogenous RNAi to silence a target gene, thus providing an important layer of protection against spurious siRNA-mediated silencing of euchromatic genes.

\textbf{Inhibition of siRNA-mediated silencing by 3’ processing signals}

Most protein-coding genes contain signals in their 3’ untranslated regions (3’ \textit{UTRs}) that mediate efficient cleavage, polyadenylation, and export of their mRNA (Proudfoot, 2011; Richard and Manley, 2009). Our findings suggest that in addition to these critical roles in mRNA metabolism, signals in 3’ \textit{UTRs} protect protein-coding genes from siRNA-mediated silencing (Figure 3.9A). The deletion or displacement of the 3’ \textit{UTR} makes the euchromatic \textit{ura4}^+ gene,
which is normally refractory to siRNA-mediated silencing, a target for siRNA-mediated silencing. This silencing is coupled to histone H3K9 methylation and the generation of secondary siRNAs at the targeted locus. We propose that proximity to polyadenylation signals prevents siRNA-mediated silencing by promoting the release and export of the RNA before it can be targeted by the siRNA-containing RITS complex (Figure 3.9). This targeting is required for both the recruitment of RDRC, which synthesizes the dsRNA precursor for the generation of secondary siRNA by Dcr1, and recruitment of the CLRC complex, which methylates histone H3. Indeed,

Figure 3.9 Control of siRNA biogenesis and function in S. pombe.
(A) Convergent transcription at euchromatic loci results in dsRNA generation but because the concentration of Dicer (Dcr1) is limiting, this dsRNA is not processed into siRNA unless Dcr1 is overexpressed (discussed in Chapter 2). Signals in the 3’ UTR inhibit siRNA-mediated heterochromatin formation by promoting cleavage and polyadenylation, which is coupled to mRNA release from the site of transcription and export from the nucleus.
(B) Centromeric RNAs contain weak polyadenylation/transcription termination signals and their slower release from the site of transcription may allow them to act as scaffolds for siRNA-mediated heterochromatin formation. Dcr1* denotes a non-catalytic role for Dcr1 in promoting heterochromatin formation at steps downstream of siRNA generation, discussed in Chapter 2.
sequencing of polyadenylated transcripts showed striking differences in patterns of cleavage and polyadenylation between mRNAs and cenRNAs (Figures 3.8, 3.9). At each mRNA, the number of distinct polyadenylation sites was low, with a high number of reads mapping to each site (Figure 3.8B). In contrast, the pattern at pericentromeric repeats was markedly different, with many peaks of polyadenylation distributed across both strands of the transcribed region, each with a low number of reads (Figure 3.8C). This is consistent with the fact that centromeres are constitutively maintained as heterochromatin via endogenous siRNAs, whereas euchromatic loci are refractory to silencing by endogenous or exogenous siRNAs (Figure 3.1B-C, 2.6). In addition, a recent study has reported that mutation of components in the Paf1 complex, which plays a role in transcriptional elongation and 3’ end processing, permits siRNA-mediated heterochromatin formation and silencing at a euchromatic reporter gene in trans (Kowalik et al., 2015). However, siRNA-mediated silencing was largely lost when a self-cleaving ribozyme sequence was inserted between the ORF and the 3’ UTR, indicating that transcript residence time is critical for targeting of a locus by RNAi.

**Dcr1 directly targets the *ura4-UTR<sup>0</sup>* locus for silencing in the absence of ectopic siRNAs**

Another striking observation was that a *ura4* allele lacking the 3’ processing signals is directly targeted for silencing by RNAi. First, we found that the partial silencing of *ura4-UTR<sup>0</sup>* requires both the RNAi and the TRAMP/exosome pathways, suggesting that RNAi may play a global role in silencing aberrant RNAs without proper 3’ end processing signals (Figure 3.6). Second, overexpression of Dcr1 in the absence of ectopic siRNA expression induced strong heterochromatic silencing at the *ura4-UTR<sup>0</sup>* locus, comparable to that of a copy of *ura4* inserted at the centromeric repeats (*otr1R::ura4*) (Figure 3.7A,D). This is almost certainly dependent on the presence of an antisense overlapping transcript produced by a cryptic promoter in the *KanMX* gene inserted in place of the 3’UTR, and is consistent with data presented in Chapter 2 that overexpressed Dcr1 is able to process dsRNA generated from overlapping convergent
euchromatic transcripts into siRNAs (Figure 2.6). At endogenous euchromatic convergent transcription units, where normal 3’ end processing occurs, siRNA generation did not correlate with silencing or heterochromatin formation (Figure 2.6). In contrast, transcriptional gene silencing and heterochromatin assembly were strongly induced at the 3’-end processing impaired cenRNAs and the aberrant ura4-UTR^D transcript, further highlighting the importance of impaired transcriptional cleavage and polyadenylation in RNAi-mediated silencing.

The fact that high expression of Dcr1 levels correlates with much stronger silencing at ura4-UTR^D than expression of high ectopic siRNAs is likely a result of altered Dcr1 localization when Dcr1 is overexpressed. Whereas Dcr1 expressed at wildtype levels localizes mainly to the nuclear periphery, elevated Dcr1 levels likely allow Dcr1 to directly localize to the overlapping transcripts at the ura4-UTR^D, where it plays the dual role of generating siRNAs, which subsequently recruit the RITS complex, and directly recruiting the CLRC complex (see Chapter 2), rather than relying solely on RITS-based recruitment. Consistent with this hypothesis, overexpression of a catalytically inactive Dcr1 that lacks ribonuclease activity but still co-immunoprecipitates with a CLRC component (Figure 2.4C) leads to a ~6x enrichment in H3K9me2 over WT at the targeted ura4^+ locus (compared to a ~400x increase when wildtype Dcr1 is overexpressed) (Figure 3.7D). Together, this data reinforces conclusions from Chapter 2 and previously published results (Emmerth et al., 2010) that controlling the level of Dcr1 protein is critical for ensuring proper Dcr1 localization and appropriate siRNA targeting of the genome.

**Intron-derived siRNAs are processed and can functionally target an intron-containing transcript**

Our analysis of siRNAs in cells that express the ura4 hairpin (ura4-hp) uncovered new targets for ura4 hairpin siRNAs that provide further support for the co-transcriptional targeting of nascent transcripts by RNAi. First, the hairpin itself becomes a target for Rdpl-dependent dsRNA synthesis resulting in high levels of siRNAs that correspond to the cox4 intron in the
hairpin loop (Figure 3.4F). The presence of these intronic cox4 secondary siRNAs suggests that the hairpin RNA is targeted co-transcriptionally prior to splicing and removal of the intron. However, in this case, the cox4 intron is located within a hairpin structure flanked by inverted repeats that can form a 250 bp dsRNA stem. Splicing may therefore be inefficient (Figure 3.4G) allowing dsRNA synthesis by RDRC prior to intron removal. Second, we observe a low level of tertiary siRNAs at the endogenous cox4 locus at regions flanking its intron. Here again, Rdp1 association must occur prior to splicing of the cox4 intron at its endogenous transcript. The low levels of cox4 tertiary siRNAs may reflect rapid removal of the intron at the endogenous cox4 gene before it can be targeted by abundant intronic hairpin-produced siRNAs, or rapid processing and export of the transcript away from the locus due to the presence of endogenous 3’ processing factors; these explanations are not mutually exclusive.

In the case of the intronic cox4 siRNAs, the density of reads originating from the cox4 intron is comparable to the density of reads mapping to the double stranded region in the ura4 hairpin, implying that a substantial portion of the ura4 reads originating from the hairpin are actually Rdp1 derived (indeed, ura4 antisense hairpin siRNAs decrease dramatically in rdp1Δ). Furthermore, the hairpin itself is a much better target for secondary siRNA generation than the ura4 mRNAs, even when the 3’ processing signals are deleted (e.g. ura4-UTR0). A possible explanation for this observation is that the presence of an intron in the hairpin RNA and its targeting by the splicing machinery may improve the ability of RNAi to target the transcript. In this regard, previous work has uncovered links between splicing and RNAi in S. pombe (Bayne et al., 2008; Motamedi et al., 2004), and in C. neoformans transcripts with suboptimal splicing signals become RNAi targets (Dumesic et al., 2013).
References


Chapter 4

Deletion of mRNA export factor \( mlo3^+ \) permits siRNA-mediated H3K9 dimethylation but not silencing

Ruby Yu and Danesh Moazed
Chapter 4

This chapter contains unpublished work. All experiments were performed by Ruby Yu. The introduction, results and discussion were written by Ruby Yu.
Abstract

small RNAs (sRNAs) regulate gene expression throughout eukaryotes. A widely conserved role of sRNAs is to silence transposable elements that could cause catastrophic disruptions within the genome. In fission yeast, this is role is critical, and short interfering RNAs (siRNAs) are required for establishment and maintenance of silenced heterochromatin at pericentromeric regions. Importantly, siRNA-mediated heterochromatin formation and silencing are tightly linked, and factors that influence one typically influence the other. Here we show that mRNA export factor Mlo3 inhibits siRNA-mediated heterochromatin formation at euchromatic loci. Cells lacking mlo3\(^+\), but not wildtype cells, exhibit very high levels of H3K9 dimethylation (H3K9me2, a conserved marker for heterochromatin) at a locus targeted by ectopic siRNAs. Surprisingly, this rarely results in silencing of the complementary transcript, and in cells that do not show silencing, H3K9me2 is not transmitted through meiosis. In contrast, mlo3\(^-\) cells that silence a reporter gene targeted by siRNAs contain high levels of H3K9me2 and H3K9me3, and furthermore can epigenetically transmit the silenced state. Together, these results show that siRNA-mediated H3K9 methylation can be functionally separated from silencing. Thus, the formation of transcriptionally silenced heterochromatin cannot be inferred from high levels of H3K9me2, and the formation of epigenetically heritable silenced heterochromatin requires a switch in chromatin state beyond H3K9me2. We propose that this switch is a transition to H3K9me3.
Introduction

RNAs have been linked to gene silencing and heterochromatin formation in many organisms. In one of the most well-studied mechanisms, classical RNA interference (RNAi), a ribonuclease-containing Dicer protein cleaves long double-stranded RNAs (dsRNAs) to generate 21-25 nt small RNA (sRNA) duplexes. One strand of sRNA is loaded onto an Argonaute effector protein, which together with the RNA-induced silencing complex (RISC), mediates post-transcriptional gene silencing (PTGS) by transcript degradation or translational repression (Bernstein et al., 2001; Fire et al., 1998; Hamilton and Baulcombe, 1999; Hammond et al., 2000; Hammond et al., 2001). In many organisms, including plants, fungi, worms, flies, and possibly mammals, the RNAi machinery has also been shown to operate on chromatin, mediating DNA or histone methylation resulting in transcriptional gene silencing (TGS) of either protein-coding genes or, largely, parasitic elements of the genome (Castel and Martienssen, 2013). In C. elegans, a class of small RNAs termed piRNAs (for PIWI-interacting RNAs) has been implicated along with the RNAi machinery in mediating methylation of histone modifications and likely inducing formation of heterochromatin (Ashe et al., 2012; Gu et al., 2012; Guang et al., 2010; Lee et al., 2012; Shirayama et al., 2012).

In fission yeast, another class of small RNAs, siRNAs, mediates silencing and heterochromatin formation at centromeres (Reinhart and Bartel, 2002; Verdel et al., 2004; Volpe et al., 2002). In this system, the fission yeast Argonaute homolog, Ago1, associated with two other proteins in the RNA-induced initiation of transcriptional gene silencing (RITS) complex, is guided to chromatin by RNA-RNA interactions between an Ago1-bound guide siRNA and a nascent transcript (Motamedi et al., 2004; Verdel et al., 2004). RITS subsequently recruits RDRC, an RNA-dependent RNA polymerase-containing complex, which uses the nascent template as a guide to generate long dsRNA, which is cleaved by the S. pombe Dicer homolog, Dcr1, to generate and amplify siRNAs (Colmenares et al., 2007; Motamedi et al., 2004; Sugiyama et al., 2005). RITS also recruits the histone H3K9 methyltransferase Clr4 in the CLRC
(Clr4-Rik1-Cul4) complex, which mediate methylation of histone H3K9, resulting in association of HP1 proteins Swi6 and Chp2, various histone deacetylases, and eventual formation of transcriptionally silenced heterochromatin (Bayne et al., 2010; Gerace et al., 2011; Hong et al., 2005; Jia et al., 2005; Motamedi et al., 2008; Sugiyama et al., 2007; Thon et al., 2005; Zhang et al., 2008). In an interesting paradox, the mechanism of transcriptional gene silencing in fission yeast requires transcription by Pol II and the presence of a nascent transcript on chromatin.

Attempts to induce silencing and heterochromatin at euchromatic reporter genes have met with varying success, and it is thought that some feature or combination of features at euchromatic loci and protein-coding genes renders them generally refractory to siRNA-mediated silencing. Several studies involving hairpin-derived siRNAs have induced silencing and H3K9 methylation of a complementary locus in trans (Iida et al., 2008; Simmer et al., 2010), but it is unclear what distinguishes loci that can be silenced from those that cannot. In Chapter 3, I presented data showing that 3’ end processing signals in the 3’UTR of protein-coding genes inhibit siRNA-mediated targeting, suggesting that delayed dissociation of the nascent transcript from chromatin may be important for allowing the RNAi machinery to target a locus for H3K9 methylation and silencing. A recently published study supports this hypothesis, showing that deletion of components in transcriptional elongation and 3’ end processing complex Paf1 permits heritable siRNA-mediated silencing and heterochromatin formation at a target locus, but that this silencing is lost when a self-cleaving ribozyme sequence is inserted immediately downstream of the ORF (Kowalik et al., 2015).

The Paf1 complex has also been shown to inhibit RNAi-mediated heterochromatin formation. Several reports have characterized genes that suppress the requirement for RNAi in centromeric heterochromatin formation. Deletion of one of these genes, including putative H3K9 demethylase Epe1, H3K14 acetyltransferase Mst2, Paf1 complex components, nuclear exosome chaperone TRAMP component Cid14, and RNA export protein Mlo3, results in the rescue of centromeric silencing and heterochromatin defects observed in cells lacking the RNAi
machinery (Kowalik et al., 2015; Reddy et al., 2011; Reyes-Turcu et al., 2011; Trewick et al., 2007; Zofall and Grewal, 2006). Importantly, deletion of any of these genes does not rescue the loss of silencing in clr4Δ cells. Interestingly, combined deletion of two of these genes, epe1Δ mst2Δ, results in uncontrolled spreading of heterochromatin and severe growth defects (Wang et al., 2015). Mst2, Epe1, and Paf1 have also been shown to promote nucleosome turnover (Aygün et al., 2013; Sadeghi et al., 2015; Wang et al., 2015).

In this chapter we further characterize the phenotype of cells lacking the RNA export protein Mlo3. We confirm previous results showing that deletion of mlo3+ permits siRNA-mediated methylation of H3K9 at the targeted locus in trans (Reyes-Turcu et al., 2011). Enrichment and spreading of H3K9 methylation at the target locus is highly variable but its levels can be comparable to what is found at native heterochromatic loci. Surprisingly, despite levels of H3K9me2 enrichment reaching up to 20% of those at pericentromeric repeats, silencing of the targeted locus in mlo3Δ cells occurs in less than 0.5% of cells. This silencing correlates with high levels of di- and tri-methylated H3K9, comparable to levels at pericentromic heterochromatin. Interestingly, the progression of a locus from an unsilenced state with high H3K9me2 to a silenced heterochromatic state seems to be required for epigenetic inheritance of heterochromatin through meiosis. Together, these results show that factors that contribute to H3K9 methylation can be separated from those that contribute to silencing and that high levels of H3K9 dimethylation cannot be equated with transcriptional gene silencing.

**Results**

**Deletion of a factor linking 3’ end processing and RNA export permits high accumulation of ectopic siRNA-directed H3K9me2 at reporter genes**

Based on our results in Chapter 3 demonstrating that transcriptional polyadenylation and cleavage signals in the 3’ UTR of ura4+ inhibit siRNA-mediated silencing and heterochromatin formation, we hypothesized that transcript residence time on chromatin is important for the
Figure 4.1. Deletion of certain factors involved in transcriptional termination or mRNA export permits trans siRNA-mediated H3K9me2.

(A) Diagram of ura4+ endogenous locus. Shaded grey box indicates regions complementary to the ura4hp construct.

(B) ChIP-qPCR assays showing enrichment of H3K9me2 in individual hrp1Δ or ctf1Δ clones. Samples are normalized to WT cells expressing ura4hp.

(C) Schematic of the cen::ade6+ insertion (top) or the endogenous ade6+ locus.

(D) ChIP-qPCR assays showing that deletion of hrp1+ or ctf1+ results in enrichment of H3K9me2 at the endogenous ade6+ locus in strains expressing cen::ade6+. Samples are normalized to WT cells expressing cen::ade6+. Error bars indicate standard deviation.

(E) ChIP-qPCR assays showing enrichment (or lack thereof) of H3K9me2 at the endogenous ade6+ locus in various transcriptional termination, mRNA export, and 3’ end processing deletion mutants carrying cen::ade6+.
Figure 4.1 (Continued)
ability of the locus to be targeted by the RNAi machinery, and that deletion of factors involved in 3’ end processing might permit more widespread siRNA targeting of euchromatic genes. To test this, we made deletions of a number of non-essential genes with annotated 3’ end processing activity and tested them for ura4+ silencing in the presence of a ura4 hairpin construct. We did not observe silencing in any of the deletion strains tested (data not shown). Surprisingly, when we performed ChIP for H3K9 dimethylation (H3K9me2), we observed occasional weak enrichment (1- to 3-fold higher than WT) for H3K9me2 at the ura4+ locus in two of the strains, ctf1Δ and hrp1Δ, but this enrichment was not consistent from one clone to another (Figure 4.1A,B). Ctf1 encodes the S. pombe homolog of cleavage/polyadenylation cleavage-stimulation factor 64 (CstF-64) in mammals and Rna15p in S. cerevisiae, and has been shown to play a role in transcriptional termination (Aranda and Proudfoot, 2001). Hrp1 is a CHD chrom-helicase/ATPase DNA binding chromatin remodeling factor, a homologue of S. cerevisiae Chd1p. In fission yeast, Hrp1 has been demonstrated to have RNA Polymerase II transcriptional termination activity (Alén et al., 2002) and interacts with centromeres, where it has been shown to be required for efficient pericentromeric silencing (Walfridsson et al., 2005).

Since the phenotype at the targeted ura4+ locus was so weak, we decided to also test these deletions in a strain where a centromeric insertion of ade6 (cen::ade6+) served as the driver siRNA-producing locus, which would be expected to target the endogenous euchromatic copy of ade6+. We again failed to observe silencing of the endogenous ade6+ locus in any of the strains tested. However, when we performed ChIP, we observed enrichment of H3K9me2 at ade6+- proximal loci. Because the two copies of ade6+ are indistinguishable in these strains, PCR oligos were used to probe genes that were not included in the cen::ade6+ fragment: either vtc4+, which is convergent with endogenous ade6+, or srk1+, which is 14kb upstream of ade6+ (Figure 4.1C). At vtc4+ we observed 3- or 2-fold enrichment in H3K9me2 in the hrp1Δ and ctf1Δ mutants, respectively (Figure 4.1D). While this result was reproducible and suggested that deleting genes involved in 3’ end processing could affect trans siRNA-directed heterochromatin
formation, the effect was very weak. This may reflect the fact that \( hrp1^+ \) and \( ctf1^+ \), as well as all of the subsequent genes we deleted, are necessarily non-essential, and thus likely do not play critical roles in 3’ end processing or cell viability. Nevertheless, we decided to expand our deletion set and test other factors.

### Table 4.1. List of deleted genes and their functions.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ctf1</td>
<td>mRNA cleavage and polyadenylation specificity factor complex subunit</td>
</tr>
<tr>
<td>( hrp1 )</td>
<td>helicase, regulation of chromatin silencing at centromere, termination of RNA Polymerase II transcription</td>
</tr>
<tr>
<td>( mex67 )</td>
<td>mRNA export</td>
</tr>
<tr>
<td>( dss1/rpn15 )</td>
<td>RNA export, proteasome regulation</td>
</tr>
<tr>
<td>( puf6 )</td>
<td>mRNA 3’ UTR binding, translation inhibition</td>
</tr>
<tr>
<td>( SPCC126.11c )</td>
<td>RNA binding, associates with Cid14 and Mlo3</td>
</tr>
<tr>
<td>( nrc1/bdf2 )</td>
<td>inhibits heterochromatin spreading</td>
</tr>
<tr>
<td>( nab2 )</td>
<td>polyA binding protein, mRNA export</td>
</tr>
<tr>
<td>( tho1 )</td>
<td>mRNA export, transcription elongation from RNA Pol II promoter</td>
</tr>
<tr>
<td>( rhn1 )</td>
<td>termination of RNA Polymerase II transcription, poly(A)-coupled</td>
</tr>
<tr>
<td>( nup132 )</td>
<td>nuclear pore, mRNA export</td>
</tr>
<tr>
<td>( nxt1/mtr2 )</td>
<td>mRNA export, associates with poly(A)-binding protein Pab2</td>
</tr>
</tbody>
</table>

Because factors involved in 3’ end processing had such weak effects, we decided to expand our deletions to factors involved in RNA export. After cleavage and polyadenylation of the protein-coding transcript, mRNAs are guided by RNA export factors to nuclear pores, where they are exported from the nucleus. We reasoned that deletion of certain RNA export factors may result in delayed disassociation of mRNA from its genomic locus, and deleted 10 more factors involved in RNA export, transcriptional termination, or 3’ end processing in \( cen::ade6^+ \)-
containing cells (Table 4.1). ChIP experiments showed that only deletion of dss1⁺, a conserved RNA export factor, resulted in enrichment for H3K9me2 at the endogenous ade6⁺ locus, and that despite high variability, the enrichment and spreading of H3K9me2 was much stronger than in ctf1Δ or hrp1Δ. Indeed, the average enrichment at vtc4⁺, a gene convergent with ade6⁺, was 15-fold higher than in WT, and H3K9me2 spread at least 10 kb upstream of ade6⁺ to srk1⁺, where H3K9me2 enrichment was 5-fold higher than in WT (Figure 4.1E).

Since dss1Δ produced such a strong phenotype, we decided to examine its known interacting partners and again test for H3K9me2 in cen::ade6⁺ strains. Dss1 is known to associate with chromatin and a nuclear pore protein, Nup146, and forms a complex with two other proteins involved in RNA export, Uap56 and Mlo3, linking mRNAs to the nuclear pore complex (Thakurta et al., 2005; Thakurta et al., 2007). Cells lacking Uap56 or Nup146 are inviable, but mlo3⁺ is non-essential. Mlo3 is a homolog of S. cerevisiae YRA1, which has been shown to link RNA Pol II and 3’ end processing factors to the RNA export machinery (Johnson et al., 2009; MacKellar and Greenleaf, 2011). In fission yeast, Mlo3 has been shown to suppress antisense transcription genome-wide and ChIPs mainly to euchromatic loci, specifically 3’ ends of ORFs (Zhang et al., 2011). Deletion of dss1⁺ or mlo3⁺ results in nuclear accumulation of polyadenylated RNAs. We made several cen::ade6⁺ mlo3Δ strains and again tested endogenous ade6⁺-proximal levels of H3K9me2. We observed considerable clone-to-clone variability in the strength and extent of H3K9me2, with enrichment at vtc4⁺ ranging from 2- to 32- fold, and enrichment at srk1⁺ ranging from 1- to 17- fold over WT cen::ade6⁺ (Figure 4.2A). Notably, enrichment of H3K9me2 at vtc4⁺ in two of the clones reached levels as high as 1/5 of centromeric (dg) H3K9me2, which is unusually high for siRNA-based targeting of a euchromatic gene in trans. We also performed ChIP in several mlo3Δ strains expressing a ura4 hairpin construct (ura4hp) and observed consistent but much lower (2- to 4.5- fold) enrichment in H3K9me2 at ura4⁺ (Figure 4.2C). Importantly, H3K9me2 nucleation at the target locus always depended on the presence of a driver siRNA-producing locus (Figure 4.2A,C). We then tested
Figure 4.2. Deletion of mRNA export factor mlo3\(^+\) permits trans siRNA-mediated H3K9me2 at multiple euchromatic loci.

(A) ChIP-qPCR assays showing that deletion of mlo3\(^+\) results in variable enrichment of H3K9me2 at the endogenous ade6\(^+\) locus only in strains expressing cen::ade6\(^+\).

(B) Same ChIP-qPCR experiment as in (A), showing H3K9me2 enrichment at centromeric dg repeats and ade6\(^+\). Note high levels of enrichment at ade6\(^+\) only in strains containing cen::ade6, and that the highest levels of enrichment at vtc4\(^+\) shown in (A) are \(~1/5\) of the H3K9me2 levels at centromeres.

(C) ChIP-qPCR assays showing that deletion of mlo3\(^+\) results in enrichment of H3K9me2 at the endogenous ura4\(^+\) locus only in strains expressing ura4hp.

(D) Reverse transcription of total RNA followed by qPCR (qRT-PCR) showing a lack of silencing of the ade6\(^+\) endogenous locus in mlo3\(\Delta\) cells carrying cen::ade6\(^+\).

(E) qRT-PCR showing a lack of silencing of the ura4\(^+\) endogenous locus in mlo3\(\Delta\) cells expressing ura4hp.
Figure 4.2 (Continued)
ura4hp and cen::ade6+-dependent silencing in mlo3Δ backgrounds by growth assays on 5FOA-containing medium or low adenine medium, respectively, and failed to observe any silencing (data not shown). This result was confirmed with qRT-PCR (Figure 4.2D,E).

We next tested the ability of the targeted ade6+ locus to maintain H3K9me2 upon either reintroduction of Mlo3 or removal of the driver cen::ade6+ locus. Mlo3 has been shown to be methylated by Clr4 on residue K167, but it is not clear whether this modification is functionally important (Zhang et al., 2011). We first reintroduced either Mlo3 or a mutant form of Mlo3 that cannot be methylated (K167A) into two of the cen::ade6+ mlo3Δ strains with the strongest trans H3K9me2 and found that all H3K9me2 was lost, showing that continued absence of Mlo3 is required to stabilize the domain of H3K9me2 and that methylation of Mlo3 does not affect its role in destabilizing H3K9me2 (Figure 4.3A).

In order to test whether H3K9me could be epigenetically inherited in the absence of siRNAs, we attempted to remove the driver cen::ade6+ locus. This proved difficult to achieve by homologous recombination in a standard transformation, since the centromeres are maintained as heterochromatin, which is less accessible to the recombination machinery, and centromeric sequences are repetitive. Therefore, we decided to cross cen::ade6+ mlo3Δ strains with strong H3K9me2 to ade6-M210 mlo3Δ strains. Because each parent has a different allele of ade6, we could test whether cells with endogenous ade6+ inherited the copy that originally had H3K9me2, and whether cells with ade6-M210, which never had H3K9me2, remained free of H3K9me2 (Figure 4.3B). After crossing, we tested H3K9me2 in mlo3Δ ade6+ cells and cen::ade6+ mlo3Δ ade6+ as a control. We found that mlo3Δ ade6+ cells lacking the cen::ade6+ driver lacked H3K9me2 at the ade6+-proximal loci, indicating that the methylation state of the ade6+ locus was not inherited in the absence of siRNAs produced from the cen::ade6+copy. More surprisingly, we found that even cen::ade6+ mlo3Δ ade6+ progeny completely lacked H3K9me2 at the endogenous ade6+ locus (Figure 4.3C). Thus, passage through meiosis results in erasure of siRNA- and mlo3Δ-dependent trans H3K9me2. This may be due to upregulation of epe1+,
putative demethylase that has no effect on establishment but inhibits maintenance of a silenced and heterochromatic state (Audergon et al., 2015; Ragunathan et al., 2015). Interestingly, H3K9me2 was reduced ~2x and ~3x at dg and cen::ade6+ as well.

Figure 4.3. Enrichment of ectopic siRNA-mediated H3K9me2 in mlo3Δ cells is lost upon reintroduction of mlo3Δ or passage through meiosis
(A) ChIP assay showing that the enrichment of H3K9me2 observed in mlo3Δ cen::ade6+ cells is lost when either mlo3Δ or mlo3-K167A are reintroduced.
(B) Schematic of the cross in (C). Red genotypes indicate progeny of interest, carrying the copy of endogenous ade6+ that was originally enriched in H3K9me2.
(C) ChIP assay showing that all progeny (obtained by random spore analysis) containing the parentally H3K9me2-enriched copy of endogenous ade6+ lose H3K9me2 enrichment, even those carrying cen::ade6+. Note that dg and cen::ade6+ also show reduced H3K9me2.
(D) Random spore analysis of a cross between cen::ade6+ cells and mlo3Δ cells.
Because mlo3Δ cells exhibit high enrichment and spreading of H3K9 methylation but no silencing, and H3K9 methylation is erased when cen::ade6+ mlo3Δ cells are crossed to cen+ mlo3Δ cells, we suspected that silencing of the H3K9me-enriched domain at the endogenous ade6 locus might be detrimental to the cell. Closer inspection revealed that the gene encoding ribosomal protein L23, rpl2302+, located within the H3K9 methylation domain between ade6+ and srk1+, is essential (Figure 4.1C). ChIP assays show that rpl2302+ is significantly enriched in H3K9 dimethylation (Figure 4.2A). A recent study reported that construction of a strain containing deletions of both epe1+, a putative demethylase that inhibits spreading of H3K9 methylation, and mst2+, a histone deacetylase also involved in inhibiting heterochromatin, resulted in uncontrolled spreading of heterochromatin resulting in severe growth defects upon initial strain construction, likely due to silencing of essential genes. Cells were able to adapt to this by epigenetically silencing CLRC components, reducing the cell's ability to methylate H3K9 and form heterochromatin (Wang et al., 2015). We had casually observed that mlo3Δ cells have a slow growth phenotype and that mlo3Δ cells expressing cen::ade6+ seemed to be particularly sick. We hypothesized that the lack of silencing in mlo3Δ cells may be due to adaptations to avoid silencing the essential rpl2302+ gene, despite high levels of H3K9me2. To test this, we performed tetrad and random spore analysis after crossing wildtype cen::ade6+ cells to mlo3Δ cells. The tetrad analysis was problematic, as there was a very low frequency of spore survival that did not seem to correlate with genotype, and this was likely due to RNA export defects. Random spore analysis of the same cross, followed by screening of progeny, showed that of 24 progeny randomly screened, 4 were cen::ade6+ mlo3Δ, 8 were cen+ mlo3Δ, 2 were cen+ wt, and 2 were cen::ade6+ wt (Figure 4.3D). The remaining 8 cells had uncertain mlo3 status. Cells were plated onto rich medium; since there was no specific selection for survival of cen::ade6+ mlo3Δ cells, it seems clear this combination is not lethal, and it is unlikely to predominantly result in synthetic growth defects. However, since this analysis was performed with random spore analysis rather than tetrad dissection, we cannot exclude the possibility that a subset of
cen::ade6+ mlo3Δ cells were extremely sick and never grew, or that surviving cells have acquired an epigenetic adaption. To test for the latter possibility, we will perform ChIP-seq experiments to look for potential peaks of H3K9 methylation at genes encoding silencing factors.

*Trans-acting siRNAs can induce ectopic establishment of meiotically heritable silenced heterochromatin*

As previously mentioned, we did not observe silencing of the endogenous ade6+ or ura4+ loci in mlo3Δ cells upon targeting with cen::ade6+ or ura4hp-derived siRNAs, respectively, even in cells with very high levels of H3K9me2. However, we decided to test whether silencing could occur in mlo3Δ cells, possibly at very low frequency. To do this, we took a strain with high trans H3K9me2 at the ade6+ endogenous locus and plated ~1000 cells per plate on rich media with low adenine. While the vast majority of cells formed white colonies, indicating normal ade6+ expression, an average of 5 cells per plate formed red colonies, indicating strong silencing, or pink, indicating weaker silencing. We picked 12 red colonies and struck cells out to single colony. From each of these populations, we picked one white and one red colony and struck each of these to single colony. Colonies derived from a white cell were completely white in 8 lines, 10-20% light pink in 3 lines, and 50% light pink in 1 line. Populations derived from red cells were always a mix of red, white, and pink cells (average 50% white and 50% red or pink), ruling out the possibility that silencing was a result of a mutation in the adenine biosynthetic pathway, and strongly suggesting a heritable but unstable and epigenetic mechanism of silencing (Figure 4.4A). From the final purified population of silenced cells, we selected the three lines that showed the highest penetrance of silencing, with 60-70% colonies growing red or pink. From each of these three lines, we propagated cultures from one white and one red colony for ChIP-based analyses using anti-H3K9me2, anti-H3K9me3, or anti-Pol II antibodies. As expected, both red cells, and white cells, which have lost silencing, showed enrichment for
**Figure 4.4.** *Trans* siRNA-mediated silencing in *mlo3Δ* is a rare event.

(A) ~1000 cells from a strain with high H3K9me2 but no silencing of the endogenous ade6* locus were spread per plate. On each plate there were several red or pink cells. Red cells were selected and purified to single colonies on rich media containing low adenine twice. An example of growth after the second restreak is shown. Note the mixed colony colors. From the 3 sets of purified progeny with the highest proportions of red cells, one red and one white colony were selected and propagated for remaining experiments.

(B) ChIP assay showing increased H3K9 dimethylation in red cells than white cells, particularly at vtc4*, which is convergent with ade6*.

(C) ChIP assay showing increased H3K9 trimethylation at vtc4*, but not the more distal srk1*.

(D) ChIP assay showing Pol II occupancy at the ade6*-proximal loci. No significant differences were observed between red and white cen::ade6* mlo3Δ cells.

(E) Red cen::ade6* mlo3Δ cells were crossed to mlo3Δ lacking cen::ade6* and random spore analysis was performed. Number of progeny matching each genotype and phenotype are shown.

(F) Example plate growth on YE plate after spreading ~100 spores from the cross in (E).

(G) 9 red cen::ade6* mlo3Δ ade6* colonies were restruck to single colony and all showed a variegated silencing phenotype; a sample is shown.

(H) 3 red cen* mlo3Δ ade6* colonies were restruck to single colony. A line that showed variegation of white and light pink colonies is shown; the remaining two lines lost silencing.
Figure 4.4 (Continued)
H3K9me2 at ade6* because our primers do not discriminate between the endogenous and centromeric copy (Figure 4.4B). On the other hand, we found that red cells showed ~7x more enrichment for H3K9me2 at vtc4*, ~5x more enrichment at rpl2302*, and ~2x more enrichment at srk1* when compared to white cells, indicating that in silenced cells, H3K9me2 in regions immediate and distal to ade6* is enriched, but enrichment is higher in the immediate vicinity of ade6* and relative spreading is reduced compared to that observed in unsilenced cells (Figures 4.2A and 4.4B). Notably, H3K9me2 enrichment at vtc4* reached 60% of the levels of ade6 (which reflects H3K9me2 enrichment at both centromeric and endogenous copies of ade6*), which were previously shown to be slightly lower than but comparable to H3K9me2 levels at the centromeric dg repeats (Figures 4.2B, 4.4B). We note that since silencing in even the red cells is incomplete (silencing is lost in a proportion of cells with each doubling), the calculated enrichments reflect a mixed population of silenced and unsilenced cells, and are likely an underestimate of levels in only silenced cells. Recent evidence has suggested that H3K9 trimethylation (H3K9me3) may play a different role in silencing and heterochromatin formation than H3K9 dimethylation (GJ, unpublished data). In the previous experiments examining H3K9 methylation in unsilenced cells, we had only tested H3K9 dimethylation. ChIP assays for H3K9me3 showed that similar to H3K9me2, there was strong enrichment at vtc4* in silenced cells compared to cells that had lost silencing (~14-fold increase), and levels of enrichment were comparable to ade6* itself, but this enrichment did not spread to srk1*, where enrichment levels in all strains matched those in clr4Δ (Figure 4.4C). Because we saw increased silencing and enrichment for H3K9me2/3 in red cells, we suspected that we would be able to detect transcriptional gene silencing at the endogenous ade6* locus by assaying Pol II occupancy using ChIP. However, we detected no changes in Pol II occupancy between red or white mlo3Δ cells (Figure 4.4D). This may be due to the presence of unsilenced cells in the “red” population due to phenotypic variegation; changes in Pol II occupancy may be difficult to detect if only a
fraction of cells contain a silenced ade6\textsuperscript{+} locus, and even when clr4\textsuperscript{+} is deleted, the increase in Pol II occupancy at ade6 (endogenous and centromeric copies) is only \(~\)4.5-fold.

Previously, we had observed that the predecessor to these cells, which had high levels of H3K9me2 but failed to silence ade6\textsuperscript{+} expression, was unable to transmit H3K9 methylation through meiosis when crossed with mlo3∆ cells lacking the cen::ade6\textsuperscript{+} locus, even in cen::ade6\textsuperscript{+} mlo3∆ cells carrying the parental copy of ade6\textsuperscript{+} that originally had H3K9me2 (Figure 4.3B,C). We decided to test whether silenced progeny could propagate the silenced state through meiosis. We crossed red cen::ade6\textsuperscript{+} ade6\textsuperscript{+} mlo3∆ cells to ade6\textsuperscript{+} mlo3∆ cells and performed random spore analysis, plating the spores onto rich medium with low adenine (Figure 4.4E,F). Note that both parental strains had wildtype copies of endogenous ade6\textsuperscript{+}, so that it was not possible to distinguish which parental copy of endogenous ade6 each progeny inherited.

The resulting progeny formed mostly white colonies, with some red and more pink or sectored colonies (Figure 4.4F). We screened 12 red-containing progeny and 12 white progeny for presence of mlo3\textsuperscript{+} and cen::ade6\textsuperscript{+} and found that while most progeny matched parental genotypes to phenotypes, there were 4 white cen::ade6\textsuperscript{+} mlo3∆ progeny (and 9 red), which was expected since even the parental cells showed variegated silencing. Surprisingly, we obtained 3 red cen\textsuperscript{+} mlo3∆ cells (and 8 white), indicating that the silenced phenotype could be transmitted through meiosis in the absence of the siRNAs produced by cen::ade6\textsuperscript{+} (Figure 4.4E). This was in stark contrast to our earlier cross involving the non-silenced cen::ade6\textsuperscript{+} mlo3∆ predecessor, which lost H3K9 methylation at the endogenous ade6\textsuperscript{+} locus in all progeny (Figure 4.3B,C).

These results demonstrate that in mlo3∆ cells siRNAs can mediate the establishment of functional heterochromatin at a low frequency and that once established, the silent state can be epigenetically inherited through meiosis. One factor that may be critical for crossing the threshold to generation of silenced heterochromatin is trimethylation of H3K9, which was elevated to nearly centromeric levels in silenced cells, but reduced back to background levels in cells that had lost the ability to silence ade6\textsuperscript{+} in trans. Trimethylation of H3K9 could be important
for recruiting downstream factors involved in silencing and compaction of chromatin to a stable and more transcriptionally repressed state. Further testing is required to determine if this is the case. Based on the data shown so far, we cannot rule out that enrichment of H3K9 dimethylation, which also correlated with silencing in mlo3Δ cells, contributes to silencing, but data presented below will show that the ability to silence a locus targeted by siRNAs does not always correlate with increasing levels of H3K9me2.

Purification of red cells to single colony showed that cen::ade6+-containing cells maintained a phenotype of variegated silencing (Figure 4.4G). However, cells that lacked cen::ade6+ rapidly lost silencing and all following progeny were either white or barely pink, indicating that maintenance of the epigenetic phenotype in the absence of the driver siRNA locus is less stable (Figure 4.4H). Previously, we and others have found that a silenced phenotype can only be epigenetically inherited in the absence of putative demethylase epe1+, and another group has found that a silenced phenotype can be epigenetically inherited in the absence of the Paf1 transcriptional elongation and 3’ end processing complex (Audergon et al., 2015; Ragunathan et al., 2015; Kowalik et al., 2015). Our data indicates that mlo3+ also affects epigenetic inheritance, but it remains to be determined whether maintenance in mlo3Δ cells will require RNAi as is the case with deletions of the Paf1 complex.

**Cells lacking Paf1 component leo1+ induce siRNA-mediated silencing that correlates with lower H3K9me2 than in non-silencing mlo3Δ cells**

A recent study identified a set of mutants in the transcriptional elongation and 3’ end processing Paf1 complex that also permit trans siRNA-mediated H3K9me2 (Kowalik et al 2015). In Paf1 mutants, H3K9me2 is associated with silencing of the target locus, and silencing is maintained after loss of the driver siRNA locus through meiosis. The conserved Paf1 complex has been shown to be involved in transcriptional elongation by Pol II, transcriptional termination and 3’ end processing (Tomson and Arndt, 2013). We deleted Paf1 component leo1+ in cells
containing cen::ade6+ and compared the effects of siRNA-mediated targeting at the endogenous ade6+ locus to those in mlo3∆ cells that do not silence ade6+.

We first examined spreading of H3K9me2 at the endogenous ade6+ locus in three different leo1∆ clones compared to an unsilenced mlo3∆ clone with strong H3K9me2. While there was significant enrichment and spreading of H3K9me2 in leo1∆ cells (~8x at vtc4 and ~5x

Figure 4.5. Deletion of Paf1 complex component leo1+ permits siRNA-mediated silencing but less spreading of H3K9me2 than mlo3∆.
(A) ChIP-qPCR assay showing less spreading of H3K9me2 at the endogenous ade6+ locus in cen::ade6+ -carrying leo1∆ cells compared to mlo3∆ cells.
(B) qRT-PCR showing 2-fold reduction of only ade6+ transcripts in leo1∆ cells.
(C) ChIP-qPCR assay showing similar or increased Pol II enrichment at ade6+ in leo1∆ cells but reduced Pol II enrichment in mlo3∆ cells.
(D) ChIP-qPCR assay showing reduced Pol II enrichment at all ade6-surrounding genes in mlo3∆ cells with or without cen::ade6.
enrichment was much lower than in the \textit{mlo3}\Delta clone (~30x at \textit{vtc4}, ~20x at \textit{rpl2302} and \textit{srk1}) (Figure 4.5A). However, qRT-PCR showed that in the \textit{leo1}\Delta cells, there was a ~2-fold reduction in \textit{ade6}\textsuperscript{+} transcript levels (this mainly reflects the endogenous \textit{ade6}\textsuperscript{+} transcript, since the \textit{cen::ade6}\textsuperscript{+} transcript is silenced) (Figure 4.5B). This silencing was unique to the \textit{ade6}\textsuperscript{+} transcript and did not spread to neighboring loci that also contained H3K9me2. As shown previously (Figure 4.2D), \textit{mlo3}\Delta cells did not exhibit detectable silencing at any of the genes. Importantly, and together with our previous data dissecting \textit{mlo3}\Delta cells that have gained or lost silencing, this data indicates that while enrichment of H3K9me2 is required for generation of silenced heterochromatin, increasing levels of H3K9me2 do not always correlate with stronger silencing.

Since \textit{mlo3}\textsuperscript{+} and \textit{leo1}\textsuperscript{+} are linked to transcription and 3' end processing, we performed ChIP to assess whether Pol II occupancy changes in either deletion. While deletion of \textit{leo1}\textsuperscript{+} did not affect Pol II occupancy at \textit{ade6}\textsuperscript{+} or the more distal genes \textit{rpl2302}\textsuperscript{+} and \textit{srk1}\textsuperscript{+}, Pol II enrichment increased almost 2-fold in the ORF of \textit{vtc4}\textsuperscript{+}, which is convergent to \textit{ade6}\textsuperscript{+} (Figure 4.5C). This increased enrichment of Pol II may reflect slower elongation or pausing due to loss of functional Paf1 complex, allowing the nascent transcripts more time to be targeted by RNAi in \textit{leo1}\Delta cells. On the other hand, in \textit{mlo3}\Delta cells, we observed a decrease in Pol II enrichment at all probed genes near the endogenous \textit{ade6}\textsuperscript{+} locus in all clones, even those lacking \textit{cen::ade6}\textsuperscript{+} (Figure 4.5C,D), indicating that the decreased Pol II occupancy was likely a genome-wide effect of \textit{mlo3}\Delta and not an effect of RNAi targeting of the locus. However, this reduced Pol II association was not associated with reduced transcript levels or functional silencing.

**Discussion**

In this chapter we presented data showing correlations, or lack thereof, between siRNA-mediated methylation of histone H3K9 and gene silencing. First, we showed that in cells lacking export factor \textit{mlo3}\textsuperscript{+}, ectopic siRNAs mediated high levels of H3K9me2 at a euchromatic locus in
the absence of associated silencing. At a low frequency, ectopic siRNAs established silencing in mlo3Δ cells, which correlated with local levels of H3K9me2 and H3K9me3 that were comparable to those of native heterochromatic dg repeats. Finally, we found that in contrast with non-silencing sites of siRNA-mediated H3K9me2, the silenced state was transmitted through meiosis, even in the absence of the driver siRNA-producing locus. We discuss the implications of these findings below.

**Generation of high levels of H3K9 dimethylation is not sufficient for transcriptional gene silencing**

Cells lacking mlo3+ showed an impressively large domain of enriched H3K9 dimethylation at the ade6+ locus targeted by siRNAs, with enrichment levels up to 32-fold higher than wildtype at the convergent gene vtc4+ and up to 17-fold higher at srk1+, which is 15 kb upstream of ade6+, but this H3K9 dimethylation did not result in silencing of the target ade6+ locus. Plating the progeny of a cell containing high levels of H3K9me2 at the target locus revealed that 0.5% of cells that did silence the endogenous ade6+ locus, and these exhibited levels of di- and tri-methylated H3K9 at vtc4+ that approached the levels of ade6+ (which reflects both the endogenous and centromeric copies), with no corresponding increase at the more distal srk1+. Importantly, we note that as the gain of a silencing phenotype correlated with increases in H3K9me2/3, correspondingly the loss of the silencing phenotype correlated with significant decreases in H3K9me2/3, below the levels observed in ancestral cells that never had silencing (Figures 4.2A,B and 4.4B). The lack of spreading in silenced cells may be related to the presence of the essential gene rpl2302+ within the domain of H3K9 dimethylation. From this data alone we would have concluded that extremely high levels of H3K9 dimethylation are required to transition to a silenced heterochromatic state. However, ChIP assays in cells lacking Paf1 complex leo1+, which consistently show siRNA-mediated silencing, showed that leo1Δ cells that exhibit siRNA-mediated silencing of ade6+ actually have a considerably lower level of
H3K9 dimethylation than non-silencing mlo3Δ cells at all tested loci, and much reduced spreading, with no measureable enrichment of H3K9 dimethylation at srk1+. Furthermore, in mlo3Δ cells that lack silencing, the very high levels of H3K9me2 at endogenous ade6* were erased when cells underwent meiosis, even though they were crossed to cells that also lacked mlo3*, and even in progeny that still carried the centromeric copy of ade6*. This may be due to activity of known heterochromatin antagonist epe1*. Progeny of these cells that silenced the endogenous copy of ade6* retained silencing after undergoing the same cross with cen* mlo3Δ cells, both in cells that still possessed cen::ade6* and in cells that lost the centromeric copy.

This data underscores two points. First, local enrichment of H3K9 methylation correlates with silencing and may be more important than wide spreading. In fact, uncontrolled spreading of H3K9 methylation may result in silencing of essential genes, causing lethality or extreme growth defects, which leads to adaptations that block silencing (Wang et al., 2015). Second, and more importantly, steps beyond H3K9 dimethylation are required to generate a silenced and heritable epigenetic state. H3K9 dimethylation, while necessary, is not a sufficient predictor of silencing. It is possible that H3K9 trimethylation is a better marker for a silenced heterochromatin state, or a combination of histone modifications or silencing factors may be required, with no one component sufficient. To test this hypothesis, we will perform ChIP-seq assaying for H3K9me3 in mlo3Δ cells that silence, have never silenced, or have lost silencing at an siRNA-targeted locus, and compare the results to those in silencing leo1Δ cells. Possibly, for reasons that are unclear, while the generation of siRNA-mediated dimethylation is relatively easy in mlo3Δ cells, the transition from dimethylation to trimethylation is difficult, and only rarely is there sufficient accumulation of H3K9me3 to switch to a silenced heterochromatric state. Even in the continued presence of ade6 siRNAs, the heterochromatic state is unstable, possibly due to the putative demethylase activity of Epe1, and at a variable rate, a proportion of cells lose the silenced state. Interestingly, cells that have lost the silenced state do not seem to retain any memory of ever having been silenced, as their levels of H3K9me2/3 match the levels in cells
that were never silenced, and similar to cells that have never been silenced, they exhibit a very low frequency of switching to a silenced state.

Regarding the low frequency of switching to a silenced and heterochromatic state, there are several possible explanations. The first, which we speculated above, is that switching from dimethylated to trimethylated H3K9 is an extremely inefficient process in mlo3Δ cells; why this would be the case is unclear, but the explanation may be related to the high extent of spreading of H3K9me2 at the endogenous ade6+ locus. Perhaps some feature of the chromatin or transcriptional landscape at the endogenous ade6+ locus in mlo3Δ cells causes Clr4 to favor spreading the dimethyl H3K9 mark rather than stabilization at the endogenous ade6+ gene to generate H3K9me3. In this regard, we note that all mlo3Δ cells, whether they carried cen::ade6+ or not, exhibited a decrease in Pol II occupancy (that was not associated with changes in transcript levels) at all measured loci.

Relevant to this discussion is the presence of the essential gene rpl2302+ within the domain of H3K9 dimethylation. It is possible that cells select against silencing the endogenous ade6+ locus due to the proximity or rpl2302+, and we note that a recent study reported epigenetic silencing of CLRC components themselves to alleviate growth defects due to uncontrolled spreading of heterochromatin (Wang et al., 2015). A similar mechanism, resulting in silencing of a heterochromatin factor other than CLRC, may be at work here.

These explanations are not mutually exclusive with the possibility that lack of silencing is an indirect effect of mRNA export defects in mlo3Δ cells resulting in low protein levels of some gene required for silencing. Occasionally, due to stochastic changes, sufficient levels may be generated to induce silencing and generate even higher levels of H3K9 methylation. Possibly, this factor assists in switching from H3K9 dimethylation to trimethylation. However, mlo3Δ cells do not exhibit silencing defects at the centromere, and in fact deletion of mlo3+ rescues the defects in silencing and heterochromatin at the centromere due to deletion of RNAi components. However, it’s possible that silencing factors, which may be limited in protein level,
would preferentially go to the centromeres rather than a euchromatic locus, and this preference may be exacerbated upon decreased export of transcripts.
References


Chapter 5

Perspectives
Safeguarding gene expression: balancing silencing and safety mechanisms

The experiments I have presented in this dissertation emphasize the redundancy of mechanisms the cell has evolved to limit siRNA activity outside of constitutively heterochromatic regions. While many euchromatic loci, particularly those with overlapping antisense transcripts, are capable of generating siRNAs and could thus potentially become targets for the silencing activity of the RNAi machinery, a number of features protect protein-coding loci. Among these, I have identified regulation of Dcr1 localization and expression as a way to restrict siRNA production, and efficient 3’ end processing as a way to restrict targeting by siRNAs. Others have described mechanisms for repression of antisense transcription, reducing possible Dicer substrates, and removal of silencing modifications on histones. Additionally, I have observed examples in which siRNA-mediated H3K9me2 accumulates at a locus in the absence of silencing, indicating that H3K9 methylation and silencing are separable events. In these instances, requirements for silencing that function downstream of H3K9 methylation must not have been met. Factors that govern the transition from H3K9me2 to H3K9me3 may be responsible for converting euchromatin into stable and heritably silenced heterochromatin, but without the proper chromosomal context, this conversion is infrequent. Below, I discuss the implications of these findings and possible experiments to further elucidate the mechanisms by which the cell protects itself from potentially harmful siRNAs.

Mechanisms by which factors antagonize siRNA-mediated targeting

In the published fission yeast literature, there have been two types of experiments studying factors that inhibit heterochromatin formation and heterochromatic silencing. The first involves studying RNAi-independent heterochromatin formation at centromeric repeats, in which deletion of some gene rescues the loss of silencing phenotype resulting from deletions of the RNAi machinery (but importantly, not heterochromatin machinery). So far, inhibitory roles in heterochromatin formation have been well characterized in a putative demethylase (Epe1), a
histone deacetylase (Mst2), and two chromatin-associated factors involved in 3’ end processing of Pol II products (Paf1 complex and Mlo3), and a handful of other genes involved in transcriptional elongation (tfs1 and pmc2) or RNA degradation (cid14) have also been shown to rescue centromeric silencing in RNAi mutants (Reddy et al., 2011; Reyes-Turcu et al., 2011; Sadeghi et al., 2015; Trewick et al., 2007; Wang et al., 2015; Zofall and Grewal, 2006). It is important to note that a low level of H3K9 methylation exists in RNAi mutants, which may be important for nucleating heterochromatin. Indeed, mst2Δ dcr1Δ cells exhibit increased levels of heterochromatin maintenance, but similar to dcr1Δ cells, cannot re-establish heterochromatin or silencing when clr4Δ is also deleted and then reintroduced (Reddy et al., 2011).

In the second set of experiments, siRNA-mediated nucleation of silencing and heterochromatin in trans is tested after local or genome-wide manipulations. The results presented in Chapters 3 and 4 of this dissertation have implicated 3’ end processing factors (cis cleavage and polyadenylation sequences as well as global export factor Mlo3) in inhibiting the ability of siRNAs to target a locus, and another study indicates an important role for transcriptional elongation and 3’ end processing complex Paf1 (Kowalik et al., 2015). Furthermore, we and others have also observed an increase in siRNA-mediated H3K9me2 upon deletion of transcriptional termination factors ctf1+, hrp1+, and res2+ (Kowalik et al., 2015). These factors are all associated with Pol II transcription, which is consistent with the nascent transcript model of RNAi-mediated heterochromatin formation, in which siRNA-bound RITS must recognize a nascent transcript and associate with chromatin to mediate silencing and chromatin modifications. I note that two Pol II-associated factors, Mlo3 and Paf1, contribute both to RNAi-dependent and RNAi-independent heterochromatin formation.

How do each of these factors individually contribute to heterochromatin inhibition? Do they all act through different mechanisms, or do some factors act in the same pathway (for example, 3’ end processing factors Mlo3 and Paf1?). To test this, epistasis analysis like that performed by Wang and colleagues (2015), showing that simultaneous deletion of epe1+ and
$mst2^*$ results in severe growth defects, would be useful. In this regard, I have unsuccessfully attempted to generate $mlo3^\Delta$ $epe1^\Delta$ cells by transformation, but have yet to perform tetrad dissection.

Our results showed that Mlo3 inhibits siRNA-mediated H3K9 methylation, but $mlo3^\Delta$ cells could rarely establish silencing or heterochromatin in trans. However, previous studies have shown that $mlo3^\Delta$ $dcr1^\Delta$ cells rescue the loss of centromeric silencing phenotype observed in $dcr1^\Delta$ cells, and that $mlo3^\Delta$ can form islands of H3K9 methylation at some euchromatic loci (Reyes-Turcu et al., 2011). Furthermore, the study which originally identified $mlo3^+$ (as ‘misseggregation and lethal when over expressed’ gene number 3) in fission yeast showed that overexpression of $mlo3^+$ is lethal, and sub-lethal levels of overexpression result in complete chromosome segregation failure, with a high percentage of daughter cells with either 2 or no nuclei (Javerzat, 1996). With this combined data, I hypothesize that Mlo3 inhibits the ability of a locus to generate H3K9 methylation more than it specifically inhibits RNAi activity. Perhaps, like Epe1 and Mst2, Mlo3 positively regulates nucleosome turnover. This could be tested by systems in which an epitope-tagged histone is inducibly expressed, similar to those described previously (Aygün et al., 2013; Javerzat, 1996; Wang et al., 2015).

Similarly, I wonder how 3’ end processing in general antagonizes siRNA-mediated targeting. One possible mechanism, which I put forth in Chapter 3, is that inefficient 3’ end processing causes longer residence time of the nascent transcript in chromatin, which gives the RNAi machinery more time to associate with the nascent transcript and stabilize on chromatin. In support of this model, a recent study showed that adding a self-cleaving ribozyme sequence upstream of the 3’ UTR of a transcript in Paf1 mutant cells abrogates the ability of siRNAs to target the locus for silencing (Kowalik et al., 2015). To further test this, one could artificially tether RNA to a non-complementary genomic locus using modified CRISPR. Modern CRISPR requires two only two components: the Cas9 RNA- and DNA binding protein, and a single guide RNA (sgRNA) that is bound by Cas9 and directs localization using a 20-nt target sequence.
Recent studies have shown that the 3’ end of the guide RNA can be extended and decorated with accessory RNA domains, which can subsequently recruit appropriate RNA-binding proteins (Shechner et al., 2015; Zalatan et al., 2015). To set up an artificial tethering system, the target sequence would be complementary to one reporter gene, for example ade6+, and a long (50-300nt) sequence complementary to another reporter gene, for example ura4+, would be appended to the 3’ end of the sgRNA. Co-expression with a Cas9 mutant lacking endonuclease activity (dCas9) would result in tethering of a long ura4- and ura4hp-complementary transcript to the ade6+ locus, and experiments to test siRNA-mediated silencing or H3K9 methylation could be performed. Ultimately this system would need to be coupled to dCas9-sgRNA modules with different residence times on chromatin, which are not yet developed, in order to determine the relationship between template RNA residence time of chromatin and siRNA-mediated histone H3K9 methylation.

Another possible mechanism by which 3’ end processing factors could inhibit siRNA-mediated targeting, which is not mutually exclusive with transcript retention, is by suppression of antisense transcripts. Previous studies have noted that mlo3Δ cells show elevated levels of antisense transcripts genome-wide (Zhang et al., 2011). Certain euchromatic loci generate RNAi-independent H3K9 methylation in mlo3Δ cells, in a manner that correlates with local generation of mlo3Δ-dependent antisense transcripts (Reyes-Turcu et al., 2011). Additionally, other studies have shown that sites of convergent transcription can be silenced in an RNAi-dependent manner (Gullerova and Proudfoot, 2008; Iida et al., 2008; Zhang et al., 2011), and our own data has shown that euchromatic sites of overlapping transcription are potential targets for Dicer activity (Chapter 2). Finally, we note that at the ura4-UTRΔ locus, which had antisense transcription, elevated Dcr1 levels generated levels of silencing and H3K9me2 that were comparable to those of an allele of ura4+ inserted at the centromere (Chapter 3). Thus, elevated antisense transcription at loci lacking proper 3’ end processing may facilitate targeting by siRNAs, possibly due to Dcr1 association resulting in both siRNA generation and CLRC
recruitment, or by simply providing another RNA scaffold to which RNA-binding RNAi (RITS or Dcr1) or CLRC machinery (Rik1) can associate. One way to test this would be by examining whether Dcr1 overexpression in \textit{mlo3\textDelta} cells promotes H3K9 methylation and chromatin association of RNAi-machinery genome-wide (by crossing \textit{mlo3\textDelta} cells to cells overexpressing Dcr1, performing tetrad dissection, and testing for silencing in the resulting \textit{mlo3\textDelta Dcr1OE} cells). Because the combination of \textit{mlo}\textDelta and \textit{Dcr1OE} is likely to be lethal or result in severe growth defects, overexpression of catalytically inactive Dcr1 mutant Dcr1-2A might be more appropriate. Results in \textit{rdp1\textDelta} and \textit{ura4-UTR_{D}} cells showed that overexpression of Dcr1-2A induces H3K9me2 but not silencing at loci where weak 3’ end processing signals correlates with overlapping transcription.

Other possible mechanisms by which failed 3’ end processing could enhance the ability of a locus to be targeted by siRNAs or generate heterochromatin in an RNAi-independent manner include changes in Pol II activity. Reduction of Pol II processivity in transcription could affect transcript residence time (a factor expanded upon above), decrease transcription-coupled nucleosome turnover thereby stabilizing existing H3K9 methylation, or be a signal to recruit downstream factors involved in silencing and heterochromatin formation. Alternatively, a failure in 3’ end processing could recruit an alternative RNA degradation pathway such as TRAMP-exosome, which is known to interact with the RNAi machinery.

**Mechanisms that govern the differential requirements for siRNA-mediated H3K9me2 and H3K9me3**

Data presented in this dissertation showed that it is possible for cells to generate high levels of H3K9me2 but not silencing. In Chapter 2, we showed that overexpression of catalytically inactive Dcr1 lacking ribonuclease activity resulted in consistent generation of H3K9me2 at centromeric repeats in cells lacking \textit{rdp1\textDelta}. Levels of H3K9me2 enrichment were variable but averaged around 50% of WT levels. However, this enrichment in H3K9me2 did not
correlate with silencing of centromeric transcripts and resulted in inconsistent rescue of chromosome segregation defects, as measured by growth on microtubule-destabilizing drug TBZ. H3K9me2 required the presence of endogenous dcr1*, indicating a continued requirement for siRNAs. Additionally, in mlo3Δ cells that failed to mediate siRNA-directed silencing of a target locus in trans, we saw high levels of H3K9me2 that could not be inherited through meiosis. Notably, in both of these systems, H3K9me2 that was not associated with silencing was subject to high variability. Cells that exhibited ectopic siRNA-mediated silencing showed levels of H3K9me2 and H3K9me3 that were comparable to levels at pericentromeric repeats; cells that had lost silencing showed enrichment in H3K9me2 at the target locus that were several fold higher than WT but several fold lower than in silenced cells, and only background levels of H3K9me3 enrichment. This data, combined with high levels of H3K9me2 that were also observed in cells that had no history of silencing, suggest that H3K9me2 is not sufficient for silencing, but that H3K9me3 may be. This requires further testing: we note that we have not yet assayed H3K9me3 in cells in which silencing was never established, which potentially could also have high levels of H3K9me3 that are completely lost when cells lose silencing. Nevertheless, we remain interested in the hypothesis that the switch from H3K9me2 to H3K9me3 is critical for maturation of heterochromatin, and this leads to two important questions. First, what governs why H3K9me2 is more readily established than heterochromatin? Related to this, what governs the transition from H3K9me2 to heterochromatin, and if H3K9me3 is critical for this transition, what governs the transition from H3K9me2 to H3K9me3?

Regarding the less stringent requirements for H3K9me2 than H3K9me3 establishment, perhaps, due to the kinetics of mediating the progression from unmethylated H3K9 to trimethylated H3K9, a more stable association between CLRC and chromatin (which could be mediated indirectly by CLRC-recruiting components of the RNAi pathway) is required for reaching a dimethylated state than for a trimethylated state. We know from recent studies that directly tethering Clr4-TetR to a euchromatic reporter gene by tetO sites results in strong
establishment of H3K9me2/3 and silencing at the locus (Audergon et al., 2015; Ragunathan et al., 2015). This indicates that stable and prolonged association of Clr4 is sufficient to mediate H3K9me3 and silencing. However, in the absence of direct Clr4 tethering, it is likely that multiple recruitment methods must cooperate to allow a sufficiently stable association of Clr4 with chromatin to mediate H3K9me3. Possibly, H3K9me2 requires less stable association or less residence time of Clr4 on chromatin. For example, perhaps RITS association with an RNA scaffold is sufficient for generation of H3K9me2, but other recruitment factors are required to stabilize CLRC association for trimethylation, or to stabilize the other protein complexes involved in mediating TGS. This could be tested using the transcript tethering system described above. If siRNA targeting of a tethered transcript results in H3K9me2 but not silencing, it would suggest that an RNA scaffold is sufficient for siRNA-mediated H3K9me2, but additional factors are required for generation of heterochromatin. Alternatively, the idea that de-stabilizing RITS/CLRC-chromatin association negatively affects silencing more than H3K9me2 could be tested by further probing cells that express a self-cleaving transcript. In paf1 mutant cells expressing hairpin siRNAs, transcripts with active self-cleaving ribozyme sequences at the target locus exhibited large abrogation of siRNA-mediated silencing, whereas transcripts with a mutant ribozyme sequence that could not mediate self-cleavage were very efficiently silenced (Kowalik et al., 2015). However, the authors did not test for heterochromatin formation, and I would be curious to know the state of H3K9me2 and me3 in the cells that could not be silenced. I note that deletion of Paf1 is likely to affect local chromatin structure and the receptiveness of euchromatin to establishment of siRNA-mediated silencing or H3K9 methylation in a manner beyond transcript residence time, for example by altering nucleosome turnover, increased antisense transcription, or reduced association of other inhibitory complexes. Enrichment of H3K9me2 but not H3K9me3 in hairpin- and ribozyme-expressing Paf1 mutants would support the hypothesis that multiple factors must cooperate for stable chromatin association of CLRC
resulting in silencing, but that reduced associations are sufficient for establishment of only H3K9me2.

In conclusion, the studies in this dissertation further highlight the potential for sites of overlapping transcription to become targets of RNAi and identify 3’ end processing as a mechanism by which mRNA-coding loci combat RNAi-mediated silencing. In the course of this work, I also discovered that siRNA-mediated H3K9me2 occurs more readily than siRNA-mediated transcriptional gene silencing, and that H3K9me2 is not a reliable predictor for generation of heritable and silenced heterochromatin. Exciting work remains to be done to elucidate the mechanism by which 3’ end processing antagonizes RNAi-mediated H3K9 methylation and the requirements, beyond H3K9me2, for generating transcriptionally silenced heterochromatin.
References


Appendix 1. Experimental Procedures

Plasmid and Strain Constructions

Cells were transformed using electroporation. 10ml cultures of cells were grown overnight in rich media (YES) and harvested by centrifugation at 4400 rpm for 1 min at a concentration of $10^7$ cells/ml. Cell pellets were washed 3 times in ice cold 1.2M sorbitol, then resuspended in 100ul 1.2M sorbitol and transferred to an ice-cold electroporation cuvette [model]. Either 10ng of purified plasmid DNA or 5ul of unpurified PCR product was added to the cell suspension. Electroporation was performed using the S. pombe setting on a Biorad Micropulser. Immediately after electroporation, 400ul 1.2M sorbitol was added to the electroporation cuvette. Either 250ul or 25ul of cells were plated, and cells were replica plated onto drug-containing media the next day, if appropriate.

Plasmids used were previously published in Colmenares et al., 2007. S. pombe strains and plasmids used in this study are described in Tables S1 and S2, respectively.

Random Spore Analysis

For mating, fresh parental cells (within 3 days of growth on plate) were picked from plates with pipette tips and mixed in 100ul sterile water. Mixed cells were spotted onto malt extract (ME) plates. After 2-3 days growth at 30C, presence of tetrads was confirmed by microscopy. If tetrads were confirmed, one loopful of cells were resuspended in 1 ml of water + 5 ul glusalase and incubated rotating at 25C ON. Cells were counted and either 100 or 1000 spores were plated on YES and incubated at 32C. Genotype was tested by colony PCR.

Growth Assays

Cells were grown in 5 ml of complete minimal medium (EMMC, from Sunrise Science Products) at 30°C overnight, diluted into fresh EMMC medium, and harvested in log phase ($10^6$-$10^7$)
Cells were washed with water, then resuspended in water to a concentration of $2 \times 10^7$ cells/ml. For silencing assays, 5 ul serial 10-fold dilutions were then spotted on appropriate plates (EMMC, EMMC-URA, and EMMC+5FOA, with or without Leucine). For TBZ assays, 5 ul serial 5-fold dilutions were spotted on non-selective plates or plates containing 17mg/L TBZ. Plates were incubated at 30°C for 3-7 days and photographed. For ade6 growth assays, cells were struck onto low adenine rich media (YE).

**sRNA Libraries**

To purify total sRNAs, cells were grown in 100 ml YES (Moreno et al., 1991) to a concentration of ~$2 \times 10^7$ cells/ml. Pellets were processed using the mirVana™ miRNA Isolation kit (Ambion), and the resulting RNA used for library construction.

To purify Ago1-associated sRNAs, cells transformed with plasmid pREP1-N3X-FLAG-Ago1 (Buker et al., 2007) were grown in 3L EMMC–Leu to a concentration of 1-2x$10^7$ cells/ml, spun down, washed twice in water, and pellets flash-frozen in liquid nitrogen. Cells were resuspended in 1 ml lysis buffer (50mM HEPES pH 7.6, 5mM MgOAc, 0.1 mM EDTA, 0.1 mM EGTA, 300 mM NaOAc, 5% glycerol, 1mM PMSF, 0.25% NP-40, plus Roche Complete Protease Inhibitor Cocktail tablet) per 1g cells in 15 mL tubes. Glass beads (Biospec, >1vol/vol cell suspension, 0.5 mm) were added and cells were disrupted in a FastPrep®-24 homogenizer (MP Biomedicals) at 6.5 meters/second for 3 times 30 s. Tubes were punctured with a 21 gage needle and the flow-through was collected in 50 ml tubes by spinning at 1000 rpm for 2 min. The flow-through was transferred to fresh tubes and spun down in a microfuge (Eppendorf 5415R) at 4300 rpm for 15 min. The supernatant was then transferred to a fresh tube. For each 1 ml supernatant, 15 ul pre-washed M2 anti-Flag resin was added and samples were incubated at 4°C for 2h. The resin was then collected by spinning at 1400 rpm for 1 min (Eppendorf 5415R), washed 3 times with lysis buffer, and once with lysis buffer without NP-40. The washed resin was resuspended in diethyl pyrocarbonate-treated water, and extracted with an equal
volume of Phenol/Chloroform/Isoamylalcohol. The aqueous phase was transferred to phase
lock tubes and extracted again with Phenol/Chloroform/Isoamylalcohol, precipitated by the
addition of 2 volume EtOH, 1/10 volume 3M NaOAc pH 5.2, and 60 ug glycogen, and
resuspended in 10 ul H2O.

Total and Ago1-associated small RNA libraries were constructed as previously
described (Halic and Moazed, 2010). Briefly, 21-24nt RNA was size-selected on a 17.5%
polyacrylamide/7M urea gel and ligated to a 3’ adapter. The ligated species were size-selected
on a 17.5% polyacrylamide/7M urea gel and ligated to a 5’ adapter. RNA was then reverse
transcribed into cDNA and PCR-amplified in a two-step process. Amplified cDNA was gel-
purified and sequenced on an Illumina GAIIx or High-Seq platform. Please see the detailed
protocol in Appendix 2.

**Analysis of polyA and sRNA sequences**

Sequencing data was analyzed using customized Python scripts (R.Y., unpublished), which are
available upon request. For sRNA libraries, reads with maximum 1 nt mismatch were aligned to
the *S. pombe* genome using Maq (http://maq.sourceforge.net/), normalized for reads per
million, and visualized using IGV (http://www.broad.mit.edu/igv/). Reads mapping to more than
one location were randomly assigned. For wildtype Ago1-associated sRNA libraries,
approximately 30% of reads generally map to rDNA (as for *ura4*+ and *ura4-UTR*DIS with *ura4-hp*
libraries), though this proportion may vary. For unclear reasons, in the *ura4-UTR*DO +*ura4-hp*
library, 75% of reads mapped to rDNA, resulting in read values (in reads per million) that are 2.5
fold lower for the remaining reads mapping genome-wide, including to regions of interest such
as *dg* and *dh* repeats, *ura4*, and *cox4*. The raw and processed small RNA and polyA data are
publicly available at the NCBI Gene Expression Omnibus under accession number GSE52535.
The deposited processed data can be visualized using IGV software.
Total RNA Purification
Total RNA was purified using the hot phenol method (Leeds et al., 1991). Briefly, 25 ml of cells were grown in YES medium overnight to log phase. Cells were harvested by centrifugation at 4400 rpm for 1 min and resuspended in 300 ul 50mM NaOAc pH 5.2, 10mM EDTA, to which 50 µl of 10% SDS and 300 ul acidified phenol (pH 4.3) were added. Cells were then incubated at 65°C for 10 min, vortexed briefly every minute, and incubated on ice for 5 min. Entire samples were then transferred to 1.5 ml Phase Lock Gel tubes (Eppendorf), to which 250 ul chloroform was added. Tubes were shaken manually, then centrifuged at maximum speed for 5 min. The supernatant was transferred to a new tube, and 60 ug glycogen, 1/10 volume 5M NaCl, and 2 volumes of 100% EtOH were added. Samples were then centrifuged (Eppendorf 5415R) at 12000 rpm for 35 min. Pellets were washed with 70% EtOH and centrifuged again at 12000 rpm for 5 min, dried, and resuspended in 86 ul RNase-free water.

Total RNA preparations were further purified using RNeasy® Mini kits (Qiagen) following the "RNeasy Mini Protocol for RNA Cleanup" provided in the manufacturer handbook.

polyA Sequencing
Purified total RNA was submitted to Helicos for direct RNA sequencing of immobilized polyadenylated RNA on oligo dT chips (Ozsolak et al., 2010). Reads with minimum 20 nt length and maximum 1 mismatch were aligned using the basic pipeline of the Helisphere software. Reads mapping to more than one location were randomly assigned. Files were normalized for reads per million and converted to IGV-viewable formats.

Reverse Transcription
cDNA was prepared using transcript-specific oligonucleotide primers (Table S3) and Superscript III reverse transcriptase from Invitrogen.
Quantitative PCR

DNA or cDNA was amplified with the Taq polymerase using primers described in table S3 in the presence of SYBR Green. For ChIP-qPCR, reported values are % of input using the $\Delta C_T$ method. For qRT-PCR, relative RNA levels were quantified using the $\Delta C_T$ method and normalized to either act1 or ura4 levels. Enrichment relative to the indicated RNA was calculated after normalization. Error bars in all figures indicate standard deviation.

Northern blot

Northern blot was performed as described previously (Buhler et al., 2006).

Chromatin Immunoprecipitation (ChIP)

ChIP experiments were performed as previously described (Huang and Moazed, 2003). Briefly, cells were cultured overnight in YES medium, diluted into fresh YES medium to OD$_{600} = 0.4$, and harvested at late log phase (OD$_{600} = 1.3-1.7$). Cells were fixed with 1% formaldehyde for 15 min at room temperature (RT), then quenched with 130 mM glycine for 5 min at RT, harvested by centrifugation, washed twice with TBS (50 mM Tris, pH 7.6, 150 mM NaCl), and flash-frozen. Cell pellets were resuspended in 500 ul lysis buffer (50 mM Heps-KOH, pH 7.5, 500 mM NaCl, 1mM EDTA, 1% Triton X-100, 0.1% SDS, and protease inhibitors) and disrupted by bead beating (Mini Beadbeater, Biospec products) for 6x30 sec with 1 ml 0.5 mm glass beads. Tubes were punctured and the flow-through was collected in a new tube by centrifugation. After sonication for 3x20 sec at 40% amplitude (Branson Digital Sonifier), the extract was centrifuged (Eppendorf 5415R) for 15 min at 13000 rpm. The soluble chromatin was then transferred to a fresh tube and normalized for protein concentration by the Bradford assay. For each sample, 10 ul lysate was saved as input, to which 240 ul of 1xTE/1% SDS were added (TE: 50 mM Tris pH 8.0, 1 mM EDTA). For anti-H3K9me3 (Diagenode), antibody was pre-incubated with washed Dynabeads M280 streptavidin, and for each immunoprecipitation 2 ug antibody coupled to 73 ul
beads was used. For anti-H3K9me2 (Abcam, ab1220) or anti-Pol II (Covance, 8WG16), antibody was pre-incubated with washed Dynabeads Protein A, and for each immunoprecipitation, 2 ug antibody coupled to 30ul beads was used. Bead plus antibody mixtures were added to 400 ul soluble chromatin, and lysis buffer was added for a final volume of 600 ul. Samples were rotated for 2h at 4°C, the beads were collected on magnetic stands, and washed 3 times with 1 ml lysis buffer and once with 1 ml TE, and eluted with 100 ul pre-heated buffer (50 mM Tris pH 8.0, 10mM EDTA, 1% SDS) at 65°C for 15 min. The eluate was collected and 150 ul 1xTE/0.67% SDS was added. Input and immunoprecipitated samples were incubated overnight at 65°C to reverse crosslinks, diluted with 250 ul TE, and treated with 50 ug RNase A at 37°C for 1h. 60 ug glycogen and 5 ul proteinase K (Roche) was added and incubation was continued at 55°C for 1h. 22 ul of 10M LiCl was added and the samples were extracted with phenol/chloroform, and EtOH precipitated. DNA was resuspended in 100 ul of 10 mM Tris pH 7.5 and 50 mM NaCl. 5ul of immunoprecipitated DNA was used for qPCR (SYBR Green) using an Applied Biosystems 7900HT lightycler and normalized relative to input.

Sample preparation for multiplex ChIP-seq

Libraries for Illumina sequencing were constructed following the manufacturer’s protocols, starting with ∼5 ng of immune-precipitated DNA fragments. Each library was generated with custom-made adapters carrying unique barcode sequence at the ligating end (Wong and Struhl, 2011). Barcoded libraries were mixed and sequenced with Illumina HiSeq2000. Raw reads were separated according to their barcodes and mapped to the *S. pombe* genome using Bowtie. Mapped reads were normalized to reads per million and visualized in IGV.

Co-immunoprecipitation assay

Flag/TAP-tagged strains and non-tagged control strains were grown in 1L of EMMC-Leu or YES
media for O/E or endogenous Dcr1, respectively, harvested at OD$_{600}$ = 1.5-2.5, wash twice with 1x TBS, and frozen. Cell pellets were resuspended in 1 volume of ice-cold lysis buffer (20 mM Hepes–NaOH at pH 7.5, 100 mM NaCl, 5 mM MgCl$_2$, 1 mM EDTA pH 8.0, 0.5 mM DTT, 10% glycerol, 0.25% Triton X-100, 1 mM PMSF, complete protease inhibitor mix [Roche]) and disrupted by bead beating (MagNa Lyser, Roche) at 4°C for 4x 45 sec with 1 ml glass beads. Tubes were punctured and the flow-through was collected in a new tube by centrifugation. The flow-through was spun down at 13,000 rpm for 3 min at 4°C (Eppendorf 5415R), transferred to a fresh tube and centrifuged at 13,000 rpm for 15 min at 4°C. The supernatant was then transferred to a fresh tube and incubated with 30 µl (endogenous Dcr1) or 50 µl (O/E Dcr1) of Dynabeads Pan Mouse IgG (Invitrogen; for TAP co-IPs) or EZView Red ANTI-FLAG M2 Affinity Gel (Sigma; for Flag co-IPs) for 3h at 4°C. Beads were washed four times with ice-cold lysis buffer, and associated proteins were eluted from the beads by addition of 1x sample buffer (100 mM Tris-HCl pH6.8, 4% SDS, 20% glycerol, 200 mM DTT, 0.05% bromophenol blue) and incubated for 10 min at 65°C. For Western blotting, anti-myc-HRP (Invitrogen), anti-Flag-HRP (Sigma) and PAP antibody (Sigma) were used.

**Western blotting**

Immunoprecipitations were carried out as described in the Supplemental Experimental Procedures. For western blotting, 2 ml of cells grown to OD$_{600}$ = 1 were pelleted. Cells were resuspended in 450 ul cold 2 M NaOH + 7.5% β-Mercaptoethanol and incubated on ice for 15 minutes. 450 ul of 55% trichloroacetic acid was added, followed by incubation on ice for 15 minute, and centrifugation at 13,000 rpm for 10 minutes (Eppendorf 5415R). Pellets were washed once in acetone, dried, and resuspended in high urea sample buffer (8 M urea, 5% SDS, 2 M Tris pH 6.8, 1.5% DTT, bromophenol blue). For blotting, anti-Flag-HRP (Sigma) and anti-actin (Abcam 8224) were used.
References


Appendix 2. Extended protocol for preparation of small RNA libraries

Preparing cells

1. Transform strains of interest with plasmid pREP1-N3X-FLAG-Ago1 (PDM 817) and plate on EMMC-Leu.

2. Grow up a 50-100ml culture overnight in EMMC-Leu. Use this starting culture to grow up 2-3 liters of cells to a density of OD = 2 to OD = 5 on the Nanodrop, putting no more than 1L of cell suspension in 2L flasks.

   Note: Time this carefully, using a doubling time range of 4-6 hours. Aim for the small culture to reach OD = 2 (or rather, a 10x dilution has OD = 0.2, at the appropriate time and use this to inoculate the larger culture, so that the cells never leave exponential phase.) From a small culture with OD = 2, it usually takes about 24 hours for cells to grow to OD = 4.

3. Spin cells in 1L Beckman containers (wash them out first) at 6500 rpm for 7 minutes in the Beckman centrifuge. Wash with 1L MilliQ water and spin again.

4. Transfer to 15ml conical tubes (use about 6 ml of water so don't exceed volume) and spin for 2 mins at 4400 rpm.

5. Discard the water and flash freeze the cells in liquid nitrogen. Store at -80°C until ready to proceed with FLAG-Ago1 pulldown.

FLAG-Ago1 pulldown

Solutions:  
Lysis Buffer w/ NP-40 (make fresh, 30ml for 3-4 samples)  
Lysis Buffer w/o NP-40 (can store at RT)

2x lysis buffer w/o NP-40:  
1M HEPES pH 7.6  
1M MgOAc  
0.5M EDTA  
0.5M EGTA  
3M NaOAc (not pH'd)  
50% glycerol  
H₂O  
TOTAL:  
50 ml

lysis buffer:  
2x lysis buffer  
100mM PMSF  
H₂O  
(10% NP-40  
(Roche Complete Proteinase Inhibitor Cocktail  
TOTAL:  
30 ml

1. Resuspend the cell pellet in 1ml lysis buffer (w/ NP-40) per 1g cells and add 0.5mm glass beads to the top (make sure there are <8ml cell suspension/15ml tube).

2. Bead-beat in MP homogenizer at max speed (6.5m/s) for 3x30s (HG setting), putting cells on ice for 5 min between beats.
   - while waiting between beats, use a screwdriver and some kind of weight to poke 2 adjacent holes in the caps of 50ml Falcon tubes so 15 ml tubes can be hammered in later (you could reuse these caps if you want).
3. After bead-beating, wipe off the bottom of the 15ml tubes and use some kind of weight (like Tiff's hand weight) to hammer them into the caps of 50ml tubes, until about the 9ml mark on the 15ml tube is level with the cap (or whatever seems appropriate so that the tube won't touch the liquid that comes out.)

4. Unscrew the 50ml caps and poke holes in the bottom of the 15ml tubes with the BD 21G1 needles (green) then screw the cap back into the 50ml tube.

5. Spin 2 min at 1000rpm in the 4°C Sorvall to separate beads and lysates, then transfer the lysates and cell debris to a pre-chilled 15ml conical tube and spin 15 mins at 4300 in the Sorvall.

6. While cells are spinning, wash the M2 anti-Flag resin 3x with 1ml lysis buffer or water (make sure to thoroughly resuspend the beads in the bottle and to cut the tip off your pipette tip, using max 200ul beads per 50 ml lysates). Spin at no higher than 1400rpm at 4°C. Resuspend in lysis buffer for 50% slurry.

7. Collect 3ul of the cell lysates for a diagnostic protein gel later and transfer the rest of the lysates to a new pre-chilled 15ml conical tube. Add the anti-FLAG resin and incubate 2 hr at 4°C on the rotator.

8. Spin down (1400rpm) and wash 3x with lysis buffer, then 1x with lysis buffer w/o NP-40.

9. Add 415ul H2O(DEPC) and save 15ul for protein gel. Add 400ul Phenol/Chloroform/Isoamylalcohol, vortex, and spin 5 mins max speed.
   - During this time pre-spin 2ml phase-lock heavy tubes 5 mins at max speed.

10. Transfer the aqueous phase to phase lock tubes, avoiding the pink beads, add 400ul more Phenol/Chloroform/Isoamylalcohol, vortex, and spin 5 min at max speed.

11. Transfer aqueous phase to new eppendorf, add 3ul glycogen (20mg/ml), 1/10 volume 3M NaOAc pH 5.2 (RNAse free), at least 2 volumes of EtOH (fill to the top if you want), and leave at least 4-6 hrs (or ON) at -80°C.

12. Wash with 70% EtOH, dry (not too long) and resuspend in 10 ul H2O (DEPC).
   - Take 0.5 ul and add 0.5 ul water to take OD260, should get 2-20ug RNA.

13. Run a protein gel to check that the IP worked, checking the lysate and the post-IP resin. The post-IP sample should have the FLAG-Ago1 band between 90 and 100 kDa.
   - Pour 5 ml for the BioRad gel casting system (or until almost to where the bottom of the combs would reach) then fill to the top with 70% ethanol and let the gel set, about 20-30 mins.
   - Pour out the ethanol and pour in the stacking gel to the top, put in the comb.

<table>
<thead>
<tr>
<th>Separating gel (5ml)</th>
<th>final concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>30% acrylamide</td>
<td>1.67 ml</td>
</tr>
<tr>
<td>2M Tris-Cl pH 8.8</td>
<td>937.5 ul</td>
</tr>
<tr>
<td>10% SDS</td>
<td>50 ul</td>
</tr>
<tr>
<td>10% APS</td>
<td>50 ul</td>
</tr>
<tr>
<td>TEMED</td>
<td>5 ul</td>
</tr>
<tr>
<td>H2O</td>
<td>2.34 ml</td>
</tr>
</tbody>
</table>
### Stacking gel (5ml, only need 1-2ml)

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume</th>
<th>Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>30% acrylamide</td>
<td>833ul</td>
<td>5%</td>
</tr>
<tr>
<td>1M Tris-Cl pH 6.8</td>
<td>625ul</td>
<td>125mM</td>
</tr>
<tr>
<td>10% SDS</td>
<td>50ul</td>
<td>0.1%</td>
</tr>
<tr>
<td>10% APS</td>
<td>50ul</td>
<td></td>
</tr>
<tr>
<td>TEMED</td>
<td>5ul</td>
<td></td>
</tr>
<tr>
<td>H₂O</td>
<td>2.34ml</td>
<td></td>
</tr>
</tbody>
</table>

### Purification of ~21nt siRNAs

17.5% polyacrylamide/7M Urea Gel:

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume</th>
<th>Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Urea</td>
<td>8.4g</td>
<td></td>
</tr>
<tr>
<td>40% PAA (19:1)</td>
<td>8.8ml</td>
<td></td>
</tr>
<tr>
<td>10x TBE</td>
<td>2ml</td>
<td></td>
</tr>
<tr>
<td>H₂O (DEPC)</td>
<td>very little (add last)</td>
<td></td>
</tr>
<tr>
<td>TOTAL</td>
<td>20 ml</td>
<td></td>
</tr>
</tbody>
</table>

Note: The urea is hard to dissolve but heating the solution up will speed this up. To avoid having to make this up a lot you can make 50 ml up and store at 4°C. Only about 7ml are required for a mini gel. When ready to cast, use 45 ul of 10% APS and 10 ul of TEMED for a 7 ml gel.

* Use DEPC water for everything until you have cDNA and make sure to wipe down all the containers, glass, etc. with a strong detergent like RNAse ZAP to avoid contamination.

1. Run the Ago1-purified RNA out on a 17.5%PAA/7M Urea gel, skipping at least 2 lanes per sample. To run, make a stock solution of formamide + bromophenol blue and add at least 1 V to each sample. Boil at 95°C for 2 min then load onto gel. Also load a 20nt and 30nt marker on each gel.

2. Run the gel in 1x TBE (made with DEPC water) between 2-4 Watts, until bromophenol blue, which runs with the 10nt fragment, almost runs off the gel (this takes > 1 hr).

   NOTE: If you keep the voltage constant, the current will drop off as the gel heats, so you can keep increasing the voltage in 10 V increments to keep the gel running as fast as possible.

3. Stain the gel in a 1:10,000 solution of SYBR Green (for RNA) in DEPC water, between 20 and 40 minutes. SYBR Green is light sensitive so keep it covered with foil.

4. Rinse 3x quickly in DEPC water.

5. View the gel using the GFP setting of the Fuji and cut out the lower between the 20nt and 30 nt markers, closer to the 20nt marker.
6. Crush the band in an eppendorf using the side of a P1000 pipette tip, then add at least 2 volumes-worth of the gel slice of 0.4M NaOAc pH 5.2, or just at 500 ul. Rotate ON at 4°C.

7. Run the solution and gel slices through a spin column (PALL centrifugal device) to separate the solution from the gel slices. Transfer liquid to an eppendorf.

8. Add 3 ul glycogen (20 mg/ml) and ethanol precipitate, leaving at -80°C at least 6 hours (or just ON). Wash with 70% EtOH and redissolve in an appropriate amount for the 3’ ligation.

**Ligation to 3’ adaptor**

*Mario’s 3’ linker: miRNA Cloning Linker-1 from IDT, AppCTGTAGGCACCATCAAT/ddC/)

T4 RNA ligase buffer: 500 mM Tris-HCl (pH7.5), 100 mM MgCl₂, 100 mM DTT

**Ligation reaction:**

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>10x T4 RNA ligase buffer (w/o ATP)</td>
<td>1 ul</td>
</tr>
<tr>
<td>RNAse inhibitor (40U/ul)</td>
<td>0.5 ul</td>
</tr>
<tr>
<td>small RNA</td>
<td>7.3 ul</td>
</tr>
<tr>
<td>RNA 3’ linker (100uM)</td>
<td>0.2 ul</td>
</tr>
<tr>
<td>T4 RNA ligase (Takara 2050a)</td>
<td>1 ul</td>
</tr>
<tr>
<td><strong>TOTAL:</strong></td>
<td><strong>10 ul</strong></td>
</tr>
</tbody>
</table>

**Control reaction:**

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>10x T4 RNA ligase buffer (w/o ATP)</td>
<td>0.5 ul</td>
</tr>
<tr>
<td>RNAse inhibitor (20U/ul)</td>
<td>0.25 ul</td>
</tr>
<tr>
<td>20 nt RNA oligo (10uM)</td>
<td>1 ul</td>
</tr>
<tr>
<td>RNA 3’ linker (100uM)</td>
<td>0.1 ul</td>
</tr>
<tr>
<td>T4 RNA ligase (Takara 2050a)</td>
<td>0.5 ul</td>
</tr>
<tr>
<td>H₂O</td>
<td>2.65 ul</td>
</tr>
<tr>
<td><strong>TOTAL:</strong></td>
<td><strong>5 ul</strong></td>
</tr>
</tbody>
</table>

1. Incubate at 20°C for 2 hrs then 4°C ON (or at least 4 hours).

2. Run another denaturing 17.5% PAA/7M urea gel to separated out 3’-ligated siRNAs, use 20nt and 30nt markers and also compare to the control reaction: band should be >30nt, though if you had little RNA to start with it may be too faint to see.
3. Stain in a 1:10,000 SYBR Green solution for 30 mins, wash 3x with DEPC water, and view on Fuji. Cut out the appropriate bands, including for the control reaction, crush in an eppendorf, and add 500 ul or 2 gel volumes of 0.4M NaOAc. Tumble ON at 4°C.

4. Run the gel slices and solution through a spin column, transfer the liquid to a new eppendorf, add 3 ul glycogen (20mg/ml), and ethanol precipitate at -80°C, ON or at least 6 hours. Wash with 70% ethanol and resuspend in appropriate volume for 5' ligation reaction.

**Ligation to 5' adaptor**

Mario's 5' adaptor: GUU CAG AGU UCU ACA GUC CGA CGA UC

**5' ligation reaction:**

<table>
<thead>
<tr>
<th>Component</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>10x T4 RNA ligase buffer (10mM ATP)</td>
<td>1 ul</td>
</tr>
<tr>
<td>RNAse inhibitor (40U/ul)</td>
<td>0.5 ul</td>
</tr>
<tr>
<td>BSA (1mg/ml)</td>
<td>0.6 ul</td>
</tr>
<tr>
<td>5' linker (100uM)</td>
<td>0.2 ul</td>
</tr>
<tr>
<td>T4 RNA ligase (Takara)</td>
<td>0.5 ul</td>
</tr>
<tr>
<td>3'-ligated siRNA</td>
<td>7.2 ul</td>
</tr>
<tr>
<td>TOTAL:</td>
<td>10 ul</td>
</tr>
</tbody>
</table>

1. Run the 5' ligation reaction at 20°C for 2 hours, 4°C ON or at least 4 hours. Also run a 5 ul half-reaction with the control 3'-ligation.

**RT-PCR to make cDNA**

multiplex 3' RT oligo: TGGCACCCGAGAATTCCAATTGATGGTGCCTACAG (prYU 478)

**RT-PCR mastermix** (one reaction):

<table>
<thead>
<tr>
<th>Component</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>5x FS buffer</td>
<td>4.4 ul</td>
</tr>
<tr>
<td>50mM MgCl2</td>
<td>2.2 ul</td>
</tr>
<tr>
<td>0.1M DTT</td>
<td>2.2 ul</td>
</tr>
<tr>
<td>20mM dNTPs</td>
<td>0.55 ul</td>
</tr>
<tr>
<td>RNAse inhibitor (40U/ul)</td>
<td>1.1 ul</td>
</tr>
<tr>
<td>TOTAL:</td>
<td>10.45 ul</td>
</tr>
</tbody>
</table>

1. Incubate 5.5 pmol (0.55 ul of 10uM) of the 3' RT oligo with the 5' ligation reaction, 5 min at 65°C (destroys the ligase, anneals the oligo).

2. Make a mastermix of the RT solution (including a one-half control reaction) and add 10.45 ul to each reaction (5.25 ul to control). For the actual samples, there are now 21 ul: divide this into a 19 ul +RT reaction and 2 ul -RT reaction.

3. Add 0.5 ul Superscript RTIII to the +RT reactions and incubate all reactions at 44°C for 1 hr, 85°C for 5 mins.

4. Add RNAse H to all reactions (0.5 ul for +RT, 0.1 ul for -RT) and incubate at 37°C for 30 min.
Cloning cDNA amplifications (2-step PCR)

First round PCR: short oligos complementary to adaptors

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>5x Phusion buffer</td>
<td>10 ul</td>
</tr>
<tr>
<td>multiplex 3' RT oligo (10uM) (prYU 478)</td>
<td>0.25 ul</td>
</tr>
<tr>
<td>5' oligo (cMO 13279, 10uM) (prYU 231)</td>
<td>0.25 ul</td>
</tr>
<tr>
<td>2mM dNTPs</td>
<td>5 ul</td>
</tr>
<tr>
<td>Phusion</td>
<td>0.5 ul</td>
</tr>
<tr>
<td>H₂O</td>
<td>32 ul</td>
</tr>
<tr>
<td>cDNA</td>
<td>2 ul</td>
</tr>
<tr>
<td>TOTAL:</td>
<td>50 ul</td>
</tr>
</tbody>
</table>

Second round PCR: add 10x excess of longer Illumina primers

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>5x Phusion buffer</td>
<td>2 ul</td>
</tr>
<tr>
<td>cMO 13278 (P5, 10uM) (prYU 229)</td>
<td>2.5 ul</td>
</tr>
<tr>
<td>RPI#, 10uM (prYU 480-484, 497-500)</td>
<td>2.5 ul</td>
</tr>
<tr>
<td>H₂O</td>
<td>3 ul</td>
</tr>
<tr>
<td>TOTAL:</td>
<td>10 ul</td>
</tr>
</tbody>
</table>

Reaction condition cycles:

<table>
<thead>
<tr>
<th>Temperature</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>94°C</td>
<td>15s</td>
</tr>
<tr>
<td>50°C</td>
<td>15s</td>
</tr>
<tr>
<td>72°C</td>
<td>15s</td>
</tr>
</tbody>
</table>

1. For your +RT samples, run one 50 ul first-round PCR reaction each; run half-reactions for the control and -RT reactions. Repeat the reaction cycles 8 times total.

2. To each reaction, add 10 ul (or 5 ul for half-reactions) of the second-round PCR mix, for a total volume of 60 ul (or 30 ul). Run the same thermocycler conditions for 16 cycles more. For the +RT samples, take out 2 ul aliquots at 4, 6, 8, 10, 12, and 16 cycles to test the appropriate number of cycles. For the control +RT reaction, take aliquots at 4, 8, and 12 cycles only.

3. Run the samples out on a 1.5mm (Biorad system) 8%PAA(19:1)/TBE gel without urea (use DNA loading dye with bromophenol blue and xylene CF), including a 50bp DNA ladder on each gel. Run the gel until the xylene blue just runs off (at 4 Watts, this takes over an hour), stain with 1:10,000 SYBR Green for 15-20 mins, wash 3x with water, image, and pick the appropriate cycle number for each sample. If the control reaction worked you can compare your samples to the control, but this isn't crucial. The band of interest is 150bp long. The lower bands are the amplification of the two adaptors ligated to each other. After too many cycles bands and smearing >150bp appear.

* Bands with rectangles indicate proper cycle number
4. Pick the appropriate cycle number and run an optimized 2-step PCR, with one 50ul first-step reactions per sample.

5. For each sample there should be 60 ul of amplified cDNA. Add 40ul water to each sample and Ethanol precipitate at -80°C for 1 hr, wash with 70% ethanol, and resuspend in at least 15 ul water. If necessary, run each sample in several lanes on the gel.

   Note: if using the BioRad gel-casting system use the 1.5mm gels and run one reaction in 2-3 lanes. Otherwise too much DNA in one lane will result in awful smearing.

6. Run the samples out on a 8%PAA(19:1)/TBE gel, skipping at least 2 lanes per sample (include a 50bp ladder), until just after xylene blue runs off. Cut out the appropriate bands (right at 150bp), crush in an eppendorf, and elute ON with 350 ul or 2 gel volumes of 0.4M NaOAc pH 5.2.

7. Run the solution through a spin column and EtOH (w/ 5ug glycogen) at 80°C at least 1 hr. Wash with 75% ethanol and resuspend in 20ul H₂O total for each sample (so split between the 3 tubes containing bands for one library). Take an OD on the Nanodrop. Keep in mind that anything below 10ng/ul is not very reliable so you should also compare to what you see on the gel.

8. Run 1-4ul of each sample out on a 8%PAA(19:1)/TBE gel along with a DNA ladder dilution set to determine concentration and check that you have the appropriate fragment.

   Note: the NEB 50bp ladder is given at a concentration of 1ug/ul total, and the 150bp band specifically is at 46ng/ul. Dilute the ladder so that you have lanes of 20ng, 10ng, 5ng, 3ng, 2ng, and 1ng of the 150bp fragment. You want to be as accurate as possible with the concentrations so that each library is about equimolar. Tufts requires that you have a concentration of at least 2ng/ul (20nM). If you don’t have enough then run more PCRs.

   Another note: For some reason you will see a smear on this final gel that is between 150bp and 200bp. We don’t know why this occurs but hopefully you can also see a distinct band at 150bp. If you do see a strong smear then you might just want to rely on your Nanodrop concentration, since you won’t be able to tell how much DNA is in your smear. Danny observed a 150bp band and a smear and when Tufts ran that sample through Bioanalyzer they reported only one band at 158bp, so the smear is probably just a gel artifact.
### Table S1. List of S. pombe strains used

<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPY 120</td>
<td>h+ leu1-32 ura4-D18 imr1R(Ncol)::ura4+ oril ade6-216 ago1Δ::TAP-KanMX6</td>
</tr>
<tr>
<td>SPY 137</td>
<td>h+ otr1R(Sphl)::ura4+ ura4-DS/E leu1-32 ade6-M210</td>
</tr>
<tr>
<td>SPY 815</td>
<td>h+ otr1R(Sphl)::ura4+ ura4-DS/E leu1-32 ade6-M210 clr4Δ::kanMX</td>
</tr>
<tr>
<td>SPY 1491</td>
<td>h- ade6-M210 leu1-32</td>
</tr>
<tr>
<td>SPY 1647</td>
<td>h- ade6-M210 leu1-32 shura4-5/natR</td>
</tr>
<tr>
<td>SPY 1716</td>
<td>h+ leu1-32 ade6-M210 ura4+D18 otr1R(Sphl)::ade6+ 3xFlag-Clr4</td>
</tr>
<tr>
<td>SPY 1774</td>
<td>h+ leu1-32 ade6-216 Chp1-13xmyc::hphMX 3xFlag-Raf1</td>
</tr>
<tr>
<td>SPY 1776</td>
<td>h+ leu1-32 ade6-216 Chp1-13xmyc::hphMX 3xFlag-Raf1 rdp1Δ::natR</td>
</tr>
<tr>
<td>SPY 1830</td>
<td>h+ leu1-32 ura4-D18 imr1R(Ncol)::ura4+ oril ade6-216 Rik1-13xmyc::hphMX</td>
</tr>
<tr>
<td>SPY 2010</td>
<td>h- ade6-M210 leu1-32 ura4-UTR – kanMX</td>
</tr>
<tr>
<td>SPY 2011</td>
<td>h- ade6-M210 leu1-32 ura4-UTR^DIS/kanMX</td>
</tr>
<tr>
<td>SPY 2013</td>
<td>h- ade6-M210 leu1-32 ura4-UTR - kanMX shura4-5/natR</td>
</tr>
<tr>
<td>SPY 2014</td>
<td>h- ade6-M210 leu1-32 ura4-UTR^DIS/kanMX shura4-5/natR</td>
</tr>
<tr>
<td>SPY 2015</td>
<td>h- ade6-M210 leu1-32 ura4-UTR^D::kanMX shura4-5/natR</td>
</tr>
<tr>
<td>SPY 2020</td>
<td>h- ade6-M210 leu1-32 ura4-UTR^D::kanMX shura4-5/natR ago1Δ::hphMX</td>
</tr>
<tr>
<td>SPY 2021</td>
<td>h- ade6-M210 leu1-32 ura4-UTR^D::kanMX shura4-5/natR clr4Δ::hphMX</td>
</tr>
<tr>
<td>SPY 2022</td>
<td>h- ade6-M210 leu1-32 ura4-UTR^D::kanMX</td>
</tr>
<tr>
<td>SPY 2046</td>
<td>h- ade6-M210 leu1-32 ura4-UTR^D::kanMX cid14Δ::hphMX</td>
</tr>
<tr>
<td>SPY 2079</td>
<td>h- ade6-M210 leu1-32 ura4-UTR^D::kanMX shura4-5/natR rdp1Δ::hphMX</td>
</tr>
<tr>
<td>SPY 2081</td>
<td>h- ade6-M210 leu1-32 ura4-UTR^D::kanMX shura4-5/natR dcr1Δ::hphMX</td>
</tr>
<tr>
<td>SPY 2150, 2215-6</td>
<td>h- ade6-M210 leu1-32 ura4+ shura4-5/natMX hrp1Δ::kanMX #16</td>
</tr>
<tr>
<td>SPY 2164</td>
<td>h- ade6-M210 leu1-32 ura4-UTR^D::kanMX shura4-5/natR cid14Δ::hphMX</td>
</tr>
<tr>
<td>SPY 2171</td>
<td>h- ade6-M210 leu1-32 ura4-UTR^D::kanMX shura4-5/natR =&gt; shura4-5Δ/hphMX</td>
</tr>
<tr>
<td>SPY 2199</td>
<td>h- ade6-M210 leu1-32 ura4+ dcr1Δ::hphMX</td>
</tr>
<tr>
<td>SPY 2201</td>
<td>h- ade6-M210 leu1-32 ura4+ ago1Δ::hphMX</td>
</tr>
<tr>
<td>SPY 2431</td>
<td>h- ade6-M210 leu1-32 ura4-UTR^D::kanMX rdp1Δ::hphMX</td>
</tr>
<tr>
<td>SPY 2433</td>
<td>h- ade6-M210 leu1-32 ura4-UTR^D::kanMX dcr1Δ::hphMX</td>
</tr>
<tr>
<td>SPY 2435</td>
<td>h- ade6-M210 leu1-32 ura4-UTR^D::kanMX clr4Δ::hphMX</td>
</tr>
<tr>
<td>SPY 2441</td>
<td>h+ otr1R(Sphl)::ura4+ ura4-DS/E leu1-32 ade6-M210 rdp1Δ::TAP-KanMX6</td>
</tr>
<tr>
<td>SPY 2444</td>
<td>h+ otr1R(Sphl)::ura4+ ura4-DS/E leu1-32 ade6-M210 hrr1Δ::KanMX6</td>
</tr>
<tr>
<td>SPY 2447</td>
<td>h+ otr1R(Sphl)::ura4+ ura4-DS/E leu1-32 ade6-M210 cid12Δ::KanMX6</td>
</tr>
<tr>
<td>SPY 2467</td>
<td>h- ade6-M210 leu1-32 ura4-UTR^D::kanMX ago1Δ::hphMX</td>
</tr>
<tr>
<td>SPY 2468</td>
<td>h- ade6-M210 leu1-32 ura4+ rdp1Δ::hphMX</td>
</tr>
<tr>
<td>SPY 2703</td>
<td>h- ade6-M210 leu1-32 ura4+ rdp1Δ::hphMX dcr1Δ::kanMX</td>
</tr>
<tr>
<td>SPY 4079-80</td>
<td>h- leu1-32 ade6-M210 mlo3Δ::hphMX #1</td>
</tr>
<tr>
<td>SPY 4081</td>
<td>h+ otr1R(Sphl)::ade6+ ura4-D18 leu1-32 #1</td>
</tr>
<tr>
<td>SPY 4494</td>
<td>h+ leu1-32 ura4-D18 imr1R(Ncol)::ura4* ade6-216 Kan-TAP-Dcr1 Nat-Flag-Raf1</td>
</tr>
<tr>
<td>SPY 4524-35</td>
<td>h+ otr1R(Sphl)::ade6+ ura4-D18 leu1-32 mlo3Δ::kan</td>
</tr>
<tr>
<td>SPY 4536</td>
<td>h- ade6-M210 leu1-32 Nat-3xFlag-Dcr1</td>
</tr>
<tr>
<td>SPY 4555</td>
<td>h- ade6-M210 leu1-32 ura4-UTR^D::kanMX</td>
</tr>
</tbody>
</table>
Table S2. List of plasmids used in this study

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PDM 817</td>
<td>pREP1-3xFlag-Ago1</td>
</tr>
<tr>
<td>PDM 829</td>
<td>pREP-nmt1/LEU2</td>
</tr>
<tr>
<td>PDM 914</td>
<td>pREPFLAG-Dcr1</td>
</tr>
<tr>
<td>PDM 915</td>
<td>pREPFLAG-Dcr1-D837A</td>
</tr>
<tr>
<td>PDM 935</td>
<td>pREPFLAG-Dcr1∆C103, Ascl-Xmal</td>
</tr>
<tr>
<td>PDM 936</td>
<td>pREPFLAG-Rdp1, Ascl-Xmal</td>
</tr>
<tr>
<td>PDM 941</td>
<td>pREPFLAG-Dcr1-D837A,D1127A</td>
</tr>
</tbody>
</table>

SPY 4557  h- ade6-M210 leu1-32 ura4-UTR\textsuperscript{123}::kanMX shura4-5/natR
SPY 4559  h- ade6-M210 leu1-32 ura4-UTR\textsuperscript{4}::kanMX
SPY 4561  h- ade6-M210 leu1-32 ura4-UTR\textsuperscript{4}::kanMX shura4-5/natR
SPY 4563  h- ade6-M210 leu1-32 ura4-UTR\textsuperscript{123}::kanMX
SPY 4565  h- ade6-M210 leu1-32 ura4-UTR\textsuperscript{123}::kanMX shura4-5/natR
SPY 4749-51 h- ade6-M210 leu1-32 mlo3\Delta::hphMX
SPY 4752-3 h+ otr1R(Sphl)::ade6+ ura4-D18 leu1-32 mex67\Delta::hphMX
SPY 4754-6 h+ otr1R(Sphl)::ade6+ ura4-D18 leu1-32 dss1\Delta::hphMX
SPY 4757-8 h+ otr1R(Sphl)::ade6+ ura4-D18 leu1-32 puf6\Delta::hphMX
SPY 4759-60 h+ otr1R(Sphl)::ade6+ ura4-D18 leu1-32 SPCC126.11\Delta::hphMX
SPY 4761-2 h+ otr1R(Sphl)::ade6+ ura4-D18 leu1-32 nrc1\Delta::hphMX
SPY 4763-4 h+ otr1R(Sphl)::ade6+ ura4-D18 leu1-32 nab2\Delta::hphMX
SPY 4765-6 h+ otr1R(Sphl)::ade6+ ura4-D18 leu1-32 tho1\Delta::hphMX
SPY 4767-8 h+ otr1R(Sphl)::ade6+ ura4-D18 leu1-32 rnh1\Delta::hphMX
SPY 4769-70 h+ otr1R(Sphl)::ade6+ ura4-D18 leu1-32 nup132\Delta::hphMX
SPY 4771-2 h+ otr1R(Sphl)::ade6+ ura4-D18 leu1-32 nrt1\Delta::hphMX
SPY 4844-7 h- leu1-32 ade6-M210 ura4+ ura4h-5/natMX mlo3\Delta::kan
SPY 4910-4 h+ otr1R(Sphl)::ade6+ ura4-D18 leu1-32 hph1\Delta::kan
SPY 4940-3 h+ otr1R(Sphl)::ade6+ ura4-D18 leu1-32 mlo3\Delta::kanMX -> mlo3::Nat-mlo3-K167A
SPY 4944-53 h+ otr1R(Sphl)::ade6+ ura4-D18 leu1-32 mlo3\Delta::kanMX -> Nat-Mlo3
SPY 4952  h- leu1-32 ura4-D18 imr1R(Ncol)::ura4+ oril ade6-216 Nat-Flag-Raf1
SPY 4915-6 h+ otr1R(Sphl)::ade6+ ura4-D18 leu1-32 ctf1\Delta::kan
SPY 4917-9 h- leu1-32 ade6-M210 ura4+ ura4h-5/natMX ctf1\Delta::kan
SPY 5058  h+ otr1R(Sphl)::ade6+ ura4-D18 leu1-32 clr4\Delta::hphMX
SPY 5063  h- leu1-32 ade6-M210 ura4+ ura4h-5/natMX clr4\Delta::hphMX
SPY 6113-5 h+ otr1R(Sphl)::ade6+ ura4-D18 leu1-32 leo1\Delta::hphMX
Table S3. List of oligonucleotides used in this study

<table>
<thead>
<tr>
<th>Primer</th>
<th>Description</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>AS45</td>
<td>fbp1Fw (qPCR)</td>
<td>AATGACAATTCCCCACTAGCC</td>
</tr>
<tr>
<td>AS46</td>
<td>fbp1Rv (qPCR)</td>
<td>ACTTCAGCTAGGATTCACCTGG</td>
</tr>
<tr>
<td>AS112</td>
<td>ura4Fw (qPCR)</td>
<td>TACAGTGGCCAGGCAGGTTATTAT</td>
</tr>
<tr>
<td>AS113</td>
<td>ura4Rv (qPCR)</td>
<td>TAGGAAACCAGTAGCCAAAGAGCCT</td>
</tr>
<tr>
<td>AS131</td>
<td>dgFw (qPCR)</td>
<td>AAGGAATGTCATCGTCGAAATT</td>
</tr>
<tr>
<td>AS132</td>
<td>dgRv (qPCR)</td>
<td>TGCTTCACGGGTATTTTTTTGAAATC</td>
</tr>
<tr>
<td>AS133</td>
<td>dhFw (qPCR)</td>
<td>GTATTTGGATTCATCGGTACTATGG</td>
</tr>
<tr>
<td>AS134</td>
<td>dhRv (qPCR)</td>
<td>ACTACATCGACACAGAAAGAAAAACAA</td>
</tr>
<tr>
<td>GJ195</td>
<td>dgFw (qPCR)</td>
<td>GGTTAAAGCGGTTGGTTGGACTG</td>
</tr>
<tr>
<td>GJ196</td>
<td>dgRv (qPCR)</td>
<td>TGACGAGGACATTCCATTATACGC</td>
</tr>
<tr>
<td>mb 91</td>
<td>act1 (Northern)</td>
<td>TCCTTTTGCATACGATCGGCAATA</td>
</tr>
<tr>
<td>mb 134</td>
<td>ura4 (Northern)</td>
<td>TTAATGCTGAGAAAGTCTTTTGCTGATATGC</td>
</tr>
<tr>
<td>prYU 1</td>
<td>ura4Fw (qPCR)</td>
<td>GGTT CCTACACAGAGAAACCTTAG</td>
</tr>
<tr>
<td>pMO 284</td>
<td>ura4Rv (qPCR)</td>
<td>AGCTCCATAGACTCCACGACCAAC</td>
</tr>
<tr>
<td>ade6Fw</td>
<td>ade6Fw (qPCR)</td>
<td>GCAGGCCAACAGGTTTGGTTATC</td>
</tr>
<tr>
<td>ade6Rv</td>
<td>ade6Rv (qPCR)</td>
<td>CCAAGTGCTTTTGATGGCAATG</td>
</tr>
<tr>
<td>vtc4Fw</td>
<td>vtc4Fw (qPCR)</td>
<td>CTGCAGATGCATCCAGGACATTC</td>
</tr>
<tr>
<td>vtc4Rv</td>
<td>vtc4Rv (qPCR)</td>
<td>CCGGTACTAGTAGAGCCAAATA</td>
</tr>
<tr>
<td>srk1Fw</td>
<td>srk1Fw (qPCR)</td>
<td>GGGAGTTGATATGTTGGCTTTG</td>
</tr>
<tr>
<td>srk1Rv</td>
<td>srk1Rv (qPCR)</td>
<td>ATACTCCGCTCGTGAAGACTTT</td>
</tr>
<tr>
<td>prYU 955</td>
<td>rpl2302Fw (qPCR)</td>
<td>GGTTACTGTTGCCTCTTTACAC</td>
</tr>
<tr>
<td>prYU 956</td>
<td>rpl2302Rv (qPCR)</td>
<td>CGCTCAATTTAGGCCATACT</td>
</tr>
<tr>
<td>prYU 762</td>
<td>cox4Fw (qPCR)</td>
<td>TGATCCTACCGCATATGGGCTG</td>
</tr>
<tr>
<td>prYU 603</td>
<td>cox4Rv (qPCR)</td>
<td>TACAGATCCGACATTCCAGGACACCG</td>
</tr>
</tbody>
</table>