CHAPTER 8

RNAi and Heterochromatin Assembly

Robert Martienssen\(^1\) and Danesh Moazed\(^2\)

\(^1\)Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724
\(^2\)Department of Cell Biology, Harvard Medical School, Boston, Massachusetts 02115-5730

CONTENTS

1. Overview of the RNAi Pathway, 153
2. Early Evidence Implicating RNA as an Intermediate in Transcriptional Gene Silencing, 154
3. RNAi and Heterochromatin Assembly in S. pombe, 155
4. Small RNAs Initiate Heterochromatin Assembly in Association with an RNAi Effector Complex, 156
5. dsRNA Synthesis and siRNA Generation, 157

6. RNA–RNA Versus RNA–DNA Recognition Models, 159
7. How Does RNAi Recruit Chromatin-modifying Enzymes?, 159
8. RNAi-mediated Chromatin and DNA Modifications in Arabidopsis, 160
9. Conservation of RNAi-mediated Chromatin Modifications in Animals, 162
10. Concluding Remarks, 164

References, 164
GENERAL SUMMARY

The intersection of RNA interference (RNAi) and heterochromatin formation brought together two areas of gene regulation that had previously been thought to operate by different, perhaps even unrelated, mechanisms. Using cytological staining methods, heterochromatin was originally defined nearly 80 years ago as those chromosome regions that retained a condensed appearance throughout the cell cycle. Early investigators studying the relationship between chromosome structure and gene expression noticed that certain chromosome rearrangements resulted in the spreading of heterochromatin into adjacent genes, which then became silent. But the stochastic nature of spreading gave rise to genetically identical populations of cells that had different phenotypes, providing a striking example of epigenetic regulation. The term RNA interference was first used to describe gene silencing when homologous antisense or double-stranded RNA (dsRNA) is introduced into the nematode Caenorhabditis elegans. It was soon recognized that a related mechanism accounted for posttranscriptional transgene silencing (PTGS) described earlier in petunia and other plants. In contrast, heterochromatin was widely believed to operate directly at the chromatin level to cause transcriptional repression, by a mechanism referred to as transcriptional gene silencing (TGS). This chapter focuses on the relationship between the RNAi pathway and the formation of epigenetically heritable heterochromatin at specific chromosome regions. It draws on recent examples that demonstrate this relationship in the fission yeast Schizosaccharomyces pombe and the mustard plant Arabidopsis thaliana.

The fission yeast nuclear genome is composed of three chromosomes that range from 3.5 Mb to 5.7 Mb in size. Each chromosome contains large blocks of repetitive DNA, particularly at centromeres, which are packaged into heterochromatin. In addition, the mating-type loci (which control cell type) and subtelomeric DNA regions also contain repetitive sequences that are packaged into heterochromatin. We now know that the assembly of DNA into heterochromatin plays both regulatory and structural roles. In the case of the mating-type loci in yeast, regulation of gene transcription by heterochromatin is important for cell-type identity. In the case of telomeres and centromeres, heterochromatin plays a structural role that is important for proper chromosome segregation during cell division. Moreover, repetitive DNA sequences and transposable elements account for a large fraction, in some cases more than half, of the genomes of many eukaryotic cells. Heterochromatin and associated mechanisms play a critical role in maintaining genome stability by regulating the activity of repeated sequences. Recent studies have uncovered a surprising requirement for components of the RNAi pathway in the process of heterochromatin formation in fission yeast and have provided insight into how these two pathways can work together at the chromatin level. Briefly, small interfering RNA (siRNA) molecules, which are a signature of RNAi and other dsRNA silencing mechanisms, assemble into the RNA-Induced Transcriptional Silencing (RITS) complex and direct epigenetic chromatin modifications and heterochromatin formation at complementary chromosome regions. RITS uses siRNA-dependent base-pairing to guide association with either DNA or nascent RNA sequences at the target locus destined to be silenced, an association that is stabilized by direct binding to methylated histone H3. The presence of these two activities in RITS triggers heterochromatin formation in concert with well-known heterochromatin-associated factors and directly links RNA silencing to heterochromatin modification.

In A. thaliana and other eukaryotes (with the exception of Saccharomyces cerevisiae), centromeric DNA regions are also composed of repetitive elements. These and other repeat sequences, such as retroelements and other transposons, are the source of siRNAs, attracting histone H3K9 and DNA methylation. Here again, several components of the RNAi pathway are required for the initiation and maintenance of these repressive methylation events. In this chapter, we discuss how heterochromatic siRNAs are produced and mediate DNA and/or chromatin modifications in fission yeast and A. thaliana.
1 Overview of the RNAi Pathway

Although the term RNAi was originally used to describe silencing that is mediated by exogenous dsRNA in *C. elegans* (Fire et al. 1998), it now broadly refers to gene silencing that is triggered by some kind of dsRNA. The steps involved in RNAi include the generation of dsRNA (which can be endogenous or exogenous such as viral RNA), processing into siRNA, and targeting of these molecules to either mRNAs (PTGS) or chromatin regions (TGS) to effect silencing. Therefore, before introducing the components of the RNAi machinery specific to TGS, we discuss the source of dsRNA that harnesses the RNAi machinery into action.

dsRNA may originate from bidirectional transcription of repetitive DNA elements, or transcription of RNA molecules that can base-pair internally to form dsRNA segments (see Fig. 1, a and b, respectively). For example, transcription through inverted repeat regions produces RNA molecules that fold back on themselves to produce hairpin structures. dsRNAs are then cleaved by Dicer, an RNase III class ribonuclease, which generates siRNAs. These are complementary duplexes, 21–27 nucleotides (nt) in size, that have a characteristic 2-nt overhang at each 3’ end of the duplex (Hamilton and Baulcombe 1999; Zamore et al. 2000; Bernstein et al. 2001; Elbashir et al. 2001; Hannon 2002; Zamore 2002; Bartel 2004; Baulcombe 2004). These duplexes are unwound into single-stranded siRNA to act as guides, through base-pairing interactions with complementary target sequences. They are therefore specificity factors and play a central role in all RNAi-mediated silencing mechanisms.

To date, two related complexes have been identified that incorporate siRNA: RISC and RITS. In the RNA-Induced Silencing Complex (RISC), siRNAs recognize target mRNAs and initiate their degradation by endonucleolytic cleavage within the mRNA region that is base-paired to the siRNA (Hannon 2002; Bartel 2004). The RNase H domain of the Argonaute/PIWI family protein (a subunit of RISC) carries out this initial mRNA cleavage event. In the nuclear RNA-Induced Transcriptional Silencing (RITS) complex (similar to the RISC), siRNAs target the complex to chromosome regions for chromatin modification (Verdel et al. 2004; Buhler et al. 2006). It is the RITS-mediated RNA pathway that is the focus of this chapter.

The central Argonaute and Dicer proteins are required for an additional type of RNA silencing mechanism involving microRNAs (miRNA). RNA, transcribed from endogenous noncoding genes that initially form hairpin RNA structures, due to extended dsRNA regions, is processed into miRNA through a series of steps (Bartel 2004; Filipowicz et al. 2005). Like siRNAs, miRNAs are 21–24 nt in size and form part of the RISC via the Argonaute proteins, to target specific mRNAs. This targeting can result in mRNA cleavage via the PIWI/RNase H domain and translational repression involving interactions with the 7meG cap at the 5’-end of the mRNA. This may be coupled to sequestration of the miRNA to cytoplasmic RNA-processing organelles known as P bodies (Processing bodies). Thus, at least two different dsRNA-processing pathways result in the generation of siRNA or miRNA, yet these RNAs use a similar machinery to inactivate cognate mRNAs. The miRNA pathway distinguishes itself because miRNAs are all produced by endogenous noncoding genes that are largely developmentally regulated and, in turn, generally target and developmentally regulate the silencing of homologous genes.

Although dsRNAs can form by the annealing of forward and reverse RNAs that result from bidirectional

---

Figure 1. Sources of dsRNA, Which Act as a Substrate for Generation of siRNAs by the Dicer Ribonuclease, and Are the Trigger for RNA Silencing

(a) Bidirectional transcription has been observed at the *S. pombe* centromeric repeats and the *centR* region of the silent mating-type locus. (b) Transcription through inverted repeats found in many plant and animal cells can potentially produce dsRNA. (c) Transcription of aberrant RNAs that may lack proper processing signals may trigger dsRNA synthesis by RdRPs.
transcription or are present in hairpin structures, in some cells, RNAi requires an additional enzyme to make dsRNA. This is the RNA-directed RNA polymerase (RdRP) found in plants and C. elegans (Dalmau et al. 2000; Sijen et al. 2001). It uses siRNAs as primers to generate more dsRNA, which can then be processed into additional siRNA by Dicer. The primary function of RdRP is thus thought to be in amplification of the RNAi response, but, as discussed later, RdRPs may have more specific roles in initiating dsRNA synthesis (see Section 5). Indeed, it seems to be involved in a process adapted for producing a better host defense response to the introduction of exogenous dsRNA. This idea is strengthened by the fact that RdRPs are not involved in the miRNA silencing pathways (Sijen et al. 2001). Interestingly, insects (including Drosophila) and vertebrates (including mammals) lack recognizable RdRP-like sequences in their genomes, but it remains possible that other polymerases carry out dsRNA synthesis in these organisms.

What then is the function of the various RNA silencing mechanisms? They are widely conserved in organisms ranging from fission yeast to plants to human, and they play central roles in the regulation of gene expression and genome stability (through stable heterochromatin formation at centromeres and telomeres). In addition, these silencing mechanisms are involved in defense against transposons and RNA viruses through degradation of their RNA transcripts (Plasterk 2002; Li and Ding 2005). Finally, transcription from some transposons generates aberrant RNAs that trigger RNAi by a mechanism thought to involve the conversion of aberrant transcripts to dsRNA by RdRPs (Fig. 1) (Baulcombe 2004).

2. Early Evidence Implicating RNA as an Intermediate in Transcriptional Gene Silencing

Before discussing the better-understood examples of RNAi-based chromatin modifications in fission yeast and Arabidopsis, we briefly discuss early experiments that suggested a role for RNA in mediating chromatin and DNA modifications. The earliest evidence for the role of an RNA intermediate in TGS came from studies of plant viroids. The potato spindle tuber viroid (PSTV) consists of a 359-nt RNA genome and replicates via an RNA–RNA pathway. The introduction of PSTV into the tobacco genome results in the DNA methylation of homologous nuclear sequences, albeit transgenic in origin (Wasseneeger et al. 1994). However, these and integrated copies of PSTV DNA only become methylated in plants that support viroid RNA transcription, suggesting the involvement of an RNA intermediate that directs DNA methylation (Wasseneeger et al. 1994). Furthermore, in Arabidopsis, the production of aberrant transcripts somehow results in the DNA methylation of all homologous promoter regions and transcriptional gene silencing (Mette et al. 1999). This, together with the finding that the replication of viral genomes in plants leads to the production of small RNAs that are 22 nt in size, suggests that RNAi-related mechanisms mediate DNA methylation (Mette et al. 2000). These observations, as well as repeat-induced silencing by transgenes, which was first discovered in petunia and in tobacco, are now widely recognized as the earliest examples of silencing by RNAi (Napoli et al. 1990; discussed in Chapter 9).

Further evidence for a link between RNAi and TGS comes from studies of repeat-induced gene silencing in Drosophila (see Chapter 5). The introduction of multiple tandem copies of a transgene results in the silencing of both the transgene and the endogenous copies (Pall-Bhadra et al. 1999). This silencing requires the chromodomain protein Polycomb, which is also involved in the packaging of homeotic regulatory genes into heterochromatin-like structure outside of their proper domains of action (Francis and Kingston 2001). In addition, this repeat-induced gene silencing requires Piwi, a Drosophila Argonaute family member required for RNAi (Pall-Bhadra et al. 2002). In Tetrahymena, another Piwi protein family member, Twi1, is required for small RNA accumulation and the massive DNA elimination that is observed in the somatic macronucleus of the protozoa (see Chapter 7). These and more recent results discussed in Section 8 suggest that the RNAi pathway is involved in the assembly of repressive chromatin structures in flies.

Other repeat-induced silencing mechanisms have been described in filamentous fungi, including Repeat-Induced Point mutation (RIP) in Neurospora crassa and Methylation Induced Pre-meiotically (MIP) in Ascolobus immersus, that do not appear to involve an RNA intermediate since they occur independently of the transcriptional state of the locus (Galagan and Selker 2004). Instead, RIP and MIP involve paired loci, where (for example) two out of three gene copies are silenced, suggesting some kind of DNA-DNA interaction mechanism involving homologous loci to induce silencing. Conversely, silencing of unpaired DNA in meiosis (MSUD), which also occurs in Neurospora, requires the RNAi pathway (Shiu et al. 2001; discussed in Chapter 6) and may have parallels in other organisms, including C. elegans (Maine et al. 2005; see Chapter 15).
3 RNAi and Heterochromatin Assembly in S. pombe

*S. pombe* chromosomes contain extensive heterochromatic DNA regions that are associated with underlying repetitive DNA elements at the centromeres and the silent mating-type loci (mat2/3) (Grewal 2000; Pidoux and Allshire 2004). Each fission yeast centromere contains a unique central core region (cnr) that is flanked by two types of repeats, called the innermost (imr) and outermost (otr) repeats (Fig. 2). The *otr* region itself is composed of *dh* and *dg* repeats.

Heterochromatin formation in *S. pombe* involves the concerted action of a number of trans-acting factors. These include histone deacetylases (HDACs), Clr4, a histone H3 lysine 9 methyltransferase (HKMT), and the histone H3K9-methyl binding proteins, Swi6 (an HP1 homolog) and Chp1. The initial recruitment of Swi6 and Clr4 to chromatin has been proposed to result in the spreading of H3K9 methylation and heterochromatin formation through sequential cycles of Clr4-catalyzed H3K9 methylation coupled to Swi6-mediated spreading to adjacent nucleosomes through its self-association (Grewal and Moazed 2003).

Mutation in components of the RNAi pathway surprisingly results in a loss of centromeric heterochromatin and the accumulation of noncoding forward and reverse transcripts from bidirectional promoters within each *dg* and *dh* repeat (Fig. 2) (Volpe et al. 2002). Fission yeast contains a single gene for each of the RNAi proteins, Dicer, Argonaute, and Rdp1 (*dcr1*, *ago1*, and *rdp1*, respectively). Deleting any of these genes results in the loss of histone H3K9 methylation, and mutants display defects in chromosome segregation, which are generally associated with defects in heterochromatin assembly (Provost et al. 2002; Volpe et al. 2003). Moreover, sequencing of a library of fission yeast small RNAs identified ~22-nt RNAs that mapped exclusively to centromeric repeat regions and ribosomal DNA repeats, suggesting that *cen* RNAs can produce dsRNAs that are processed into siRNAs (Reinhart and Bartel 2002). Thus, it was suggested that the RNAi pathway

---

**Figure 2. Organization of Heterochromatic Chromosome Regions in S. pombe and A. thaliana**

The centromere of *S. pombe* chromosome 1 is shown as an example. The unique central core (*cnr*) region is flanked by innermost (*imr* and *imrR*) and outermost (*otr* and *otrR*) repeats. The *otr* region is transcribed in both directions, giving rise to forward (blue) and reverse (red) transcripts. The region between the *mat2* and *mat3* genes contains a domain that is homologous to the centromeric *dg* and *dh* repeats (*cenH*) and is also bidirectionally transcribed. Atf1 and Pcr1 are DNA-binding proteins that act in parallel with RNAi in mating-type silencing. *Arabidopsis* centromeres are composed of 180-bp repeats (green) interspersed with retrotransposable elements (yellow). Forward transcripts initiating within the long terminal repeat (LTR) of the retroelement and reverse transcripts initiating within the 180-bp repeats are indicated.
could recruit Swi6 and Clr4 to chromatin to initiate and/or maintain heterochromatin formation at each of the above loci (Fig. 3) (Hall et al. 2002; Volpe et al. 2002).

Interestingly, both TGS and PTGS mechanisms appear to contribute to the down-regulation of cen RNAs. The forward strand transcript is primarily silenced at the transcriptional level, as demonstrated in RNAi mutants (Volpe et al. 2002). The reverse strand of cen transcripts, however, is not affected by Swi6 mutants (Volpe et al. 2002), and silencing of this cen-reverse transcript occurs primarily at the posttranscriptional level.

RNAi also plays a role in silencing the mating-type locus (mat2/3) (Hall et al. 2002). mat2/3 is interrupted by a region of DNA that is highly homologous to centromeric repeats (called cenH, cenHomology)(Fig. 2). Like the cen repeats, the cenH region is divergently transcribed to produce forward and reverse RNA (Noma et al. 2004). These cenH transcripts accumulate to high levels in RNAi mutants. In contrast, RNAi is not necessary for silencing of a reporter transgene inserted at mat2/3 if silencing is not first somehow compromised. The reason for this difference is that a partially redundant silencing mechanism, involving two DNA-binding proteins, Pcr1 and Atf1, which bind to mat2/3, recruits the heterochromatin machinery independently of RNAi (Jia et al. 2004). This is sufficient for silencing the reporter gene in the absence of RNAi but not for preventing the accumulation of non-coding cenH transcripts (Fig. 2).

4 Small RNAs Initiate Heterochromatin Assembly in Association with an RNAi Effector Complex

The discovery that the RNAi pathway is involved in heterochromatin formation in fission yeast and in transcriptional gene silencing in other systems raised the question of how it could directly regulate chromatin structure. Purification of Chp1, a chromodomain protein that is a structural component of heterochromatin, led to the identification of the RITS complex (Verdel et al. 2004). RITS contains the fission yeast Ago1 protein and Tas3, a protein of unknown function, in addition to Chp1. It also contains centromeric siRNAs, which are produced by the Dicer ribonuclease, and importantly, RITS associates with centromeric repeat regions in a siRNA-dependent fashion. RITS has therefore been proposed to use centromeric siRNAs to target specific chromosome regions for inactivation, and this provides a direct link between RNAi and heterochromatin assembly (Fig. 4).

Like RISC, which mediates PTGS, RITS uses siRNAs for target recognition. Unlike RISC, however, RITS associates with chromatin and initiates heterochromatin formation as opposed to mRNA inactivation. How can siRNAs target specific chromosome regions? Two possible mechanisms have been proposed. In the first model, siRNAs bound to Ago1 in the RITS complex must somehow base-pair with an unwound DNA double helix. In the second model, RITS-associated siRNAs base-pair with noncoding RNA transcripts at the target locus (Fig. 4).

According to either model, the association of RITS with chromatin via siRNA results in the recruitment of the Clr4 HKMT and subsequent histone H3K9 methylation. This is followed by Swi6 binding and the spreading of H3K9 methylation and heterochromatin. However, Clr4 is also required for the association of RITS with chromatin, suggesting that it provides methylated H3K9 to which the RITS complex can bind, thereby stabilizing its association with chromatin. The chromodomain of Chp1 was already known to bind specifically to methylated H3K9 residues (Partridge et al. 2002), and mutations in Clr4 or the chromodomain of Chp1 that are involved in this interaction result in a loss of RITS binding to chromatin (Partridge et al. 2002; Noma et al. 2004). Moreover, RITS can also bind to chromatin domains that are coated with methylated H3K9 through the chromo-
through base-pairing of siRNA with either DNA or RNA transcripts.

Recent evidence strongly supports a role for RITS and siRNAs in the initiation of heterochromatin assembly. Buhler et al. (2006) used a site-specific RNA-binding protein to artificially tether the RITS complex to the RNA transcript of the normally active urad4' gene. Remarkably, this tethering results in the generation of urad4' siRNAs and silencing of the urad4' gene in a manner that requires both RNAi and heterochromatin components. In addition, this system allowed a direct evaluation of the ability of newly generated siRNAs to initiate H3K9 methylation and Swi6 binding, which are molecular markers for heterochromatin formation. Interestingly, the newly generated urad4' siRNAs were found to be under negative control by the conserved siRNA ribonuclease, Eri1, which restricts them to the locus where they are produced. However, when the gene encoding Eri1 is deleted, urad4' siRNAs are able to act in trans to silence a second copy of the urad4' gene, which is inserted on a different chromosome in the same cell. This experiment therefore demonstrates that siRNAs can act as specificity factors that direct RITS and heterochromatin assembly to a previously active region of the genome.

The ability of siRNAs to initiate silencing in *S. pombe* has also been examined using a different method, which relies on the expression of a hairpin RNA to produce siRNAs homologous to a GFP transgene (Sigova et al. 2004). In this system, hairpin siRNAs promoted silencing of the GFP reporter gene at the PTGS, but not TGS, level (Sigova et al. 2004). It is unclear why the hairpin siRNAs cannot induce TGS and heterochromatin assembly at the chromosomal copy of GFP. One possible explanation is that heterochromatin is assembled at specific subnuclear locations, and assembly outside these locations occurs inefficiently (Gasser et al. 2004; Chapter 4).

5 dsRNA Synthesis and siRNA Generation

Bidirectional transcription of centromeric DNA repeats could in principle provide the initial source of dsRNA in fission yeast (Volpe et al. 2002). dsRNA resulting from the annealing of forward and reverse transcripts could then be a substrate for the Dicer ribonuclease. However, RNA-directed RNA polymerase (Rdp1) and its associated cofactors, as well as the Ctr4 HKMT, are also required for siRNA production by Dicer (Hong et al. 2005; Li et al. 2005; Buhler et al. 2006). These observations indicate that the generation of heterochromatic siRNAs by Dicer is coupled to chromatin and Rdpl-dependent events (Fig. 4).
The Rdpl enzyme resides in a multiprotein complex that also contains Hrr1, an RNA helicase, and Cid12, a member of the β family of DNA polymerases which includes poly(A) polymerase enzymes (Motamedi et al. 2004). This complex has been termed RNA-directed RNA polymerase complex (RDRC), and all of its subunits are required for heterochromatin formation at centromeric DNA regions (Motamedi et al. 2004). As expected from the presence of Rdpl, RDRC has RNA-directed RNA polymerase activity in vitro, and mutations that abolish this activity also abolish RNAi-dependent silencing in vivo (Motamedi et al. 2004; Sugiyama et al. 2005). The in vitro RNA synthesis activity of RDRC does not require an siRNA primer (Motamedi et al. 2004). RITS may therefore provide in vivo specificity by recruiting RDRC to selected RNA templates via siRNA. Consistent with this hypothesis, subunits of the RDRC are required for siRNA generation, and RITS complexes purified from cells that lack any subunit of the RDRC are devoid of siRNAs (Motamedi et al. 2004; Li et al. 2005; Sugiyama et al. 2005; Buhler et al. 2006).

The presence of Cid12 in the RDRC is intriguing and raises the possibility that another polymerase activity participates in chromosome-associated RNA silencing. Because some members of this family have poly(A) polymerase activity, one possibility is that adenylation of Rdpl-produced dsRNA may be important for their further processing. Interestingly, Cid12-like proteins are conserved throughout eukaryotes (Table 1); mutations in Rde-3, a C. elegans member of this family, result in defective RNAi (Chen et al. 2005), corroborating a conserved role for these enzymes in the RNAi pathway.

There is evidence for dsRNA synthesis and processing associated with the generation of heterochromatic siRNAs occurring on the chromosome, at sites of transcription of noncoding centromeric RNAs (Fig. 4). Evidence includes, first, that Rdpl can be cross-linked to centromeric DNA repeats (Volpe et al. 2002; Sugiyama et al. 2005), and to the forward and reverse RNA transcripts that originate from these regions (Motamedi et al. 2004). As is the case with cross-linking to DNA, cross-linking to centromeric RNAs requires Dicer and Clr4, and is therefore siRNA- and chromatin-dependent. Second, siRNA generation requires chromatin components, including Clr4, Swi6, and the HDAC Sir2 (Hong et al. 2005; Li et al. 2005; Buhler et al. 2006). Finally, the association of RDRC with RITS is dependent on siRNAs as well as Clr4, suggesting that it occurs on chromatin (Motamedi et al.

Table 1. Conservation of RNAi and heterochromatin proteins

<table>
<thead>
<tr>
<th>S. pombe</th>
<th>A. thaliana</th>
<th>C. elegans</th>
<th>Drosophila</th>
<th>H. sapiens</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dcr1</td>
<td>DCL1 to 4</td>
<td>Dcr-1</td>
<td>Dcr1 and 2</td>
<td>DCR-1</td>
</tr>
<tr>
<td>Ago1</td>
<td>AGO1 to 10</td>
<td>Rde-1, Alg-1 and -2</td>
<td>Ago1 to 3, Piwi</td>
<td>AGO-1 to AGO-4</td>
</tr>
<tr>
<td>Chp1*</td>
<td>CMT3</td>
<td>Aubergine/Sting</td>
<td>Piwi-1 to Piwi-4</td>
<td>PIWI-1 to PIWI-4</td>
</tr>
<tr>
<td>Tas3*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rdpl</td>
<td>RDR1 to 6</td>
<td>Ego-1, Rrf-1 to -3</td>
<td>GH20028p</td>
<td>KIAA1404</td>
</tr>
<tr>
<td>Hrr1</td>
<td>SGS2/SDE3</td>
<td>ZK1067.2</td>
<td>CG11265*</td>
<td>POLS*</td>
</tr>
<tr>
<td>Cid12</td>
<td></td>
<td>Rde-3, Trf-4*</td>
<td>HP1</td>
<td>HP1α, β, γ</td>
</tr>
<tr>
<td>Swi6</td>
<td>LHP1 (TFL2)</td>
<td>Hpl-1, Hpl-2, F32E10.6</td>
<td>Su(var)3-9</td>
<td>SUV39H1 and 2</td>
</tr>
<tr>
<td>Clr4</td>
<td>SUVH2 to 6</td>
<td></td>
<td>Ddb1</td>
<td>DDB1</td>
</tr>
<tr>
<td>Rik1*</td>
<td>DDB1</td>
<td>M18.5</td>
<td>Cu4</td>
<td>Cu4</td>
</tr>
<tr>
<td>Cid12</td>
<td></td>
<td></td>
<td>Sir2</td>
<td>SIRT1</td>
</tr>
<tr>
<td>Sir2</td>
<td>SIR2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eri1</td>
<td>ERI1</td>
<td></td>
<td>CG6393</td>
<td>THEX1</td>
</tr>
</tbody>
</table>

* An obvious ortholog of the chromodomian protein, Chp1, has not been identified in the other model organisms listed here, but most eukaryotic cells contain multiple chromodomian proteins. CMT3 in Arabidopsis is a chromodomian DNA methyltransferase, which acts in the same pathway as AGO4 and may be analogous to Chp1.

* No obvious orthologs of Tas3 have been identified, but it shares weak sequence similarity with a mouse ovary testis specific protein (NP_031552).

* Cid12 belongs to a large family of conserved proteins that share sequence similarity with the classic poly(A) polymerase as well as 2'-5'-oligoadenylate enzymes.

* C. elegans have about 20 SET domain proteins, but an H3K9 HMT has not yet been identified in this organism.

* S. pombe contains another Rik1-like protein, Ddb1, which is involved in DNA damage repair. Metazaoans and plants appear to contain only a single Rik1-like gene, called Ddb1, which has been shown to be involved in DNA damage repair, but it is unknown whether it also participates in heterochromatin formation.
Thus, the generation of dsRNA and heterochromatic siRNAs may involve the recruitment of RDRC to chromatin-associated nascent pre-mRNA transcripts as illustrated in Figure 5 (Martienssen et al. 2005; Verdel and Moazed 2005). The fact that transcription and siRNA generation are likely to occur simultaneously reinforces the difference between RNA silencing mechanisms that mediate chromatin modifications and PTGS. However, this distinction is unlikely to be absolute. For example, in _C. elegans_, mutations in several chromatin components, similar to _S. pombe_, result in defects in RNAi and transposon-induced RNA silencing (see Table 1) (Sijen and Plasterk 2003; Grishok et al. 2005; Kim et al. 2005), raising the possibility that in some cases dsRNA synthesis and processing may occur on the chromosome regardless of whether silencing occurs at the TGS or PTGS level.

6 RNA–RNA Versus RNA–DNA Recognition Models

An outstanding question in working out the role of RNAi in heterochromatin assembly is whether RITS/RDRC associates with DNA or nascent RNA. The observation that tethering components of the RNAi machinery to a gene transcript can induce heterochromatin-dependent gene silencing in _cis_ clearly demonstrates that this process can be promoted via initial interactions with nascent RNA transcripts (Buhler et al. 2006). Importantly, _cis_-restriction rules out the possibility that the initial events of dsRNA synthesis and siRNA generation occur on mature transcripts where mRNA products from different alleles cannot be distinguished. Furthermore, a direct prediction of the RNA–RNA interaction model is that transcription at the target locus should be required for RNAi-mediated heterochromatin assembly. Although the requirement for transcription has not been directly tested, mutations in two different subunits of RNA polymerase II (RNA pol II), denoted Rpb2 and Rpb7, have specific defects in siRNA generation and heterochromatin assembly, but not on general transcription (Djupedal et al. 2005; Kato et al. 2005). This is reminiscent of Rbp1 mutants, which have defects in histone modifications (i.e., H3K4 methylation and H2B ubiquitination) coupled to transcriptional elongation (Hampsey and Reinberg 2003), and provides a precedent for the hypothesis that RNAi-mediated H3K9 methylation and heterochromatin formation could be coupled to transcriptional elongation via the association of RNAi complexes with RNA pol II. In fact, contrary to the widely held view that heterochromatin is an inaccessible structure that inhibits transcription, RNAi-mediated heterochromatin assembly has little or no effect on the association of RNA pol II with _S. pombe_ centromeric repeats (Volpe et al. 2002; Djupedal et al. 2005; Kato et al. 2005; Buhler et al. 2006).

Therefore, nascent RNA transcripts, which act as templates for RITS in the RNA–RNA recognition model, are present in heterochromatic domains (Fig. 4).

The RNA–RNA targeting model is also supported by the observation that components of both the RITS and RDRC complexes can be localized to noncoding centromeric RNAs using in vivo cross-linking experiments (Motamedi et al. 2004). This localization is siRNA-dependent, which suggests that it involves base-pairing interactions with the noncoding RNA. In addition, it requires the Clr4 HKMT, suggesting that it is coupled to binding of RITS to methylated H3K9 and occurs on chromatin. Nonetheless, the possibility that siRNAs can also recognize DNA directly through base-pairing interactions cannot be ruled out. For example, in plants, siRNAs that are complementary to promoter regions that are (presumably) not transcribed can still direct DNA methylation, another modification which takes place during heterochromatin formation within these regions (see Chapter 9).

7 How Does RNAi Recruit Chromatin-modifying Enzymes?

The recruitment of Clr4 and Swi6 is a key step in initiating histone H3K9 methylation and heterochromatin assembly, through an autoregulatory modification-
binding model (Figs. 3 and 4) (Grewal and Moazed 2003). However, because RITS association to chromatin and Clr4-catalyzed histone H3K9 methylation are interdependent processes, it has been difficult to determine the event that provides the initial trigger for RNAi-dependent heterochromatin assembly. One solution to this chicken-and-egg problem is that siRNA-dependent base-pairing interactions could provide the initial signal for heterochromatin assembly (Fig. 4). Consistent with this hypothesis, de novo generation of urad' siRNAs promotes silencing of a previously active copy of the urad' gene that is coupled to the recruitment of RITS and Swi6 to chromatin (Buhl er et al. 2006). The initial binding of RITS may, however, be transient and difficult to detect; stable binding of RITS to chromatin would require dual interactions between (1) RITS-bound siRNAs and the nascent transcript and (2) the chromodomain binding of Clr1 to methylated H3K9. In this model, RITS itself directly recruits Clr4. Alternatively, Clr4 may be recruited by a parallel pathway that involves one or more DNA-binding proteins, as is the case at the silent mating type and telomeric regions (Jia et al. 2004; Kanoh et al. 2005). In either scenario, Clr4-mediated H3K9 methylation would be required to stabilize RITS association with chromatin, which then leads to the recruitment of RDRC, dsRNA synthesis, and siRNA generation (Fig. 5).

Clr4 has recently been found to be a component of a multiprotein complex that contains the heterochromatin protein Rik1, a Cullin E3 ubiquitin ligase, Cul4, and several other proteins (Hong et al. 2005; Horn et al. 2005; Jia et al. 2005; Li et al. 2005). These Clr4-associated proteins further strengthen the link between RNA and heterochromatin formation. The Rik1 protein is a member of a large family of β propeller WD repeat proteins that have been implicated in RNA or DNA binding. Members of this protein family include the Cleavage Polyadenylation Specificity Factor A (CPSF-A) involved in pre-mRNA splicing, and the DNA damage binding 1 (Ddb1) protein involved in binding UV-damaged DNA. CPSF-A is of particular interest because Rik1 shares sequence similarity with its putative RNA-binding domain involved in the recognition of mRNA polyadenylation sequences (Barbino et al. 2000). The Ddb1 protein, like Rik1, is a component of a Cul4 E3 ubiquitin ligase complex and is involved in the recognition and repair of UV-damaged DNA (Higa et al. 2003; Zhong et al. 2003). An exciting possibility is that Rik1 acts in a fashion that is similar to CPSF-A and Ddb1, binding to an RNAi-generated product during heterochromatin assembly (Fig. 5).

8 RNAi-mediated Chromatin and DNA Modifications in Arabidopsis

The mechanism by which RNAi guides heterochromatic modifications in plants is similar to the mechanism in fission yeast, but there are also many differences. The most important difference is that plants have methylated DNA at many repressive heterochromatin regions: In this respect they resemble vertebrates, but differ from worms and Drosophila (Lippman and Martienssen 2004). Four genetic screens for mutants that relieve RNA-mediated TGS have recovered mutants in H3K9-specific HKMTs, and in RNAi components, but they have also uncovered the required function of DNA methyltransferases, SWI/SNF remodeling complexes, and a novel RNA polymerase (Baulcombe 2004). These screens are described in detail in Chapter 9, but here we briefly compare the mechanism in fission yeast and plants.

Each of the silencing mutant screens used inverted repeats introduced in trans to induce the silencing of endogenous or transgenic reporter genes. Relief from silencing indicates a mutation that has arisen in a necessary component of the silencing pathway. The endogenous genes used were PAI2 (involved in amino acid biosynthesis) (Mathieu and Bender, 2004) and SUPERMAN (a transcription factor that regulates flower development) (Chan et al. 2004), and the reporter genes used were driven by either a strong viral promoter or a strong seed-specific promoter (Matzke et al. 2004). In each case, the promoter was targeted for silencing, in some cases along with the rest of the gene. A number of the genes found through these screens are illustrated in Figure 6 (see also Table 1 of Chapter 9). Only one RNAi mutant was identified, in only one of the screens, and this was the argonaute gene AGO4. However, three of the screens recovered mutants in DNA methyltransferases, including MET1 and CMT3. A third DNA methyltransferase related to the mammalian DNMT3 was identified by reverse genetics, as this activity is encoded by DRM1 and DRM2, two redundant genes unable to be determined in single mutant screens (see Chapter 9). Indeed, redundancy may account for the failure to recover additional components of the RNAi apparatus: for example, although DCL3 (DICER-LIKE 3) and RNA-DEPENDENT RNA POLYMERASE 2 (RDRP) are predominantly required for production of the 24-nt siRNA associated with transposons and repeats, at least two other DCL genes in Arabidopsis can substitute for DCL3 to some extent (Gasciolliri et al. 2005).
produce siRNA if they are transcribed (Martienssen 2003). In these cases, loss of siRNA is correlated with the loss of H3K9me2 (Cao et al. 2003; Lippman et al. 2003). Mutants in the SW12/SNF2 chromatin-remodeling ATPase DDM1 (decreased DNA methylation) also abolish siRNA and H3K9me2 accumulation from a wide range of transposons, although when siRNA is retained, so is H3K9me2 (Lippman et al. 2003). It is possible, therefore, that siRNA in plants is bound to the chromosome via methylated DNA instead of, or in addition to, binding via methylated histones as is the case in S. pombe (Fig. 7).

DDM1 has an exquisite specificity for transposons and repeats, and must somehow recognize these as being different from genes. siRNA, perhaps bound to the chromosome by methyl-binding proteins, would have the required specificity to make this distinction. Transposons and repeats in Arabidopsis are a major source of 24-nt, and some 21-nt, siRNA, consistent with this idea (Lippman et al. 2004). Centromeric satellite repeats, which are arranged in tens of thousands of tandem copies on either side of each centromere, are also transcribed and processed by RNAi (Fig. 6). This processing depends on DCL3, RDR2, and DDM1. Silencing also depends on H3K9me2 and CMT3. However, silencing is more complex than in fission yeast, as retrotransposon insertions into the repeats can silence them, and this depends on other mechanisms including MET1, DDM1, and the histone deacetylase HDA6 (May et al. 2005).

As mentioned earlier, in fission yeast, subunits of RNA pol II are required for silencing and siRNA production, supporting the idea that the RNAi- and chromatin-modification apparatus is recruited to the chromosome by nascent transcripts (Fig. 5). In Arabidopsis, two subunits of a novel RNA polymerase (RNA pol IV) were recovered in one of the four screens mentioned above (Kanno et al. 2005) but were first isolated as weak mutants in PTGS, along with mutants in RNA-dependent RNA polymerase (Herr et al. 2005). It is not yet known what template is used by RNA pol IV, but both methylated DNA (Onodera et al. 2005) and double-stranded RNA have been proposed (Vaughn and Martienssen 2005). Only the largest subunits are unique to RNA pol IV, which presumably uses the same complement of small subunits as RNA pol II. Additional SW12/SNF2 chromatin remodelers that were also recovered in these screens may alter local chromatin structure to facilitate processivity of RNA polymerases. It is therefore likely that they facilitate transcription by RNA pol IV (Kanno et al. 2004). A similar role can be proposed for DDM1, although the require-
flowering. No such phenotype is observed in mutants of RNAi, even when siRNA is lost. Instead, RNAi may play a role in initiating FWA silencing, because FWA transgenes are rapidly silenced when first introduced into a plant, and this silencing depends on DCL3, RDR2, and AGO4 (Chan et al. 2004). Silencing might then be maintained by DNA methylation, regulated by MET1. Similarly, transposons that lose siRNA in met1 mutants cannot be re-silenced in backcrosses, but those that do not lose siRNA can be re-silenced, implicating siRNA in reestablishing silencing in cis rather than in trans (Lippman et al. 2003). Similarly, late-flowering FWA alleles are stably inherited in backcrosses after being removed from met1 or ddm1 mutant backgrounds because maintenance of epialleles is heritable (Soppe et al. 2000).

Finally, it is possible that miRNA may guide DNA methylation of genes in some circumstances. mRNA from the PHABULOSA gene is targeted for cleavage by miRNA 165 and 166 in Arabidopsis, and the gene itself is methylated downstream from the region that matches the miRNA. Interestingly, this match spans an exon junction, so that the spliced RNA must interact with the miRNA if this guides methylation (Bao et al. 2004). However, other members of the same gene family are not methylated in this way, and neither are most other miRNA target genes (Martienssen et al. 2004; Ronemus and Martienssen 2005). Conversely, several other genes are methylated in the Arabidopsis genome, and typically at their 3' end, in a mechanism that requires MET1 but not DDM1 (Lippman et al. 2004; Tran et al. 2005). It remains to be seen whether RNA is involved in these cases.

9 Conservation of RNAi-mediated Chromatin Modifications in Animals

Perhaps the most widely studied examples of epigenetic silencing are found in animals, including Drosophila and C. elegans, as well as the mouse. The role of RNA and RNA interference in transcriptional silencing and heterochromatic modifications appears to be conserved in some model animals as well as in protists and plants. In Drosophila, both PIWI and the PIWI class Argonaute homolog, Aubergine (Sting), are required for epigenetic and heterochromatic silencing (see also Chapter 5). Gypsy retrotransposons are the target of silencing in ovary follicle cells and female gonads by PIWI itself (Sarot et al. 2004). This is mediated by the heterochromatic gene Flamenco (with as-yet-unknown function), and requires the 5'UTR of the Gypsy polyprotein gene. The detection of 25–27-nt small RNAs from this region suggests it occurs
via an RNAi-mediated mechanism. Cut-and-paste DNA transposons are also affected by RNAi. For example, certain telomeric P elements (a type of DNA transposon) can suppress transposition to elsewhere in the genome when inherited through the female germ line, resulting in a strongly repressive "cytotype." This repression is completely dependent on the PIWI homolog, Aubergine, as well as the Swi6 homolog HP1 (Reiss et al. 2004). However, not all P-repressive cytotypes such as those mediated by other, nontelomeric P elements are dependent on Aubergine or HP1.

Unlinked transgenes in Drosophila are silenced post-transcriptionally when present in many copies (Pal-Bhadra et al. 1997, 2002). Silencing is associated with large amounts of 21-nt siRNA and depends on PIWI. Transgene fusions can also silence each other transcriptionally, in a manner that requires the Polycumb chromatin repressor. This silencing is not associated with increased levels of siRNA from the transgene transcript but is (largely) dependent on PIWI. Involvement of Polycumb in this example, and HP1 in other examples, of PIWI-dependent silencing, implicates the RNAi pathway and histone methylation in the silencing process. Tandem transgene arrays also exhibit position-effect variegation in Drosophila, and this variegation is strongly suppressed by mutants in HP1 as well as in piwi, aubergine, and the putative RNA helicase Spindle-E (homeless) (Pal-Bhadra et al. 2004). Transgenes inserted within centric heterochromatin are also affected, and heterochromatic levels of H3K9me2 are reduced in spindle-E mutant cells. These observations strongly support a role for both chromatin proteins and components of the RNAi pathway in gene silencing within Drosophila heterochromatin.

In the Drosophila male germ line, the heterochromatic Suppressor of Stellate repeats (Su(stel)), located on the Y chromosome, are transcribed first on the antisense strand, and then on both strands during spermatocyte development, possibly following the insertion of a nearby transposon (Aravin et al. 2001). These nuclear transcripts are required to silence sense transcripts of the closely related X-linked Stellate gene, whose overexpression results in defects in spermatogenesis. Although heterochromatic sequences are involved, silencing in this case appears to be posttranscriptional, is associated with 25–27-nt siRNA, and depends on both Aubergine and Spindle-E.

In C. elegans, examples of TGS in somatic cells have been reported. This depends on the RNAi pathway genes rde-1, dcr-1, rde-4, and rrf-1, as well as HP1 homologs and the histone modification apparatus (Grishok et al. 2005).

Somatic heterochromatin is not widespread in C. elegans, but an example of naturally occurring RNAi-dependent heterochromatic silencing has been described in the germ line (Sijen and Plasterk 2003). During meiosis, unpaired sequences, such as the X chromosome in males, are silenced via H3K9me2, and this silencing depends on RNA-dependent RNA polymerase (Maine et al. 2005; see Chapter 15), reminiscent of meiotic silencing of unpaired DNA (MSUD) in Neurospora (see Shiu et al. 2001; Chapter 6). However, other components of the RNAi apparatus have not yet been implicated in this process, and it is not known whether it is related mechanistically to RNAi-mediated heterochromatin assembly in fission yeast.

Finally, like Drosophila, mammalian cells lack genes related to RNA-dependent RNA polymerases found in plants, worms, and fungi. Nonetheless, antisense RNA has been implicated in the most widely studied epigenetic phenomena of all, imprinting and X inactivation (see Chapters 19 and 17, respectively). In the case of X inactivation, a 17-kb spliced and polyadenylated noncoding RNA known as Xist is required to silence the inactive X chromosome from which it is expressed. Conversely, Xist itself is silenced on the active X chromosome, a process that depends in part on the antisense RNA Tsix. Silencing is accompanied by modification of histones associated with upstream chromatin regions, which are marked with H3K9me2 and H3K27me3 (see Chapter 17). Silencing of other imprinted loci in the mouse, including Igf2r and the Dlk1-Gtl2 region, is also maintained by antisense transcripts from the paternal or maternal allele, respectively. In the case of Dlk1-Gtl2, this noncoding RNA is specifically processed into miRNA that targets the antisense transcript from the paternal allele, encoding a sushi (gypsy) class retrotransposon (Davis et al. 2005).

Although the parallels with forward and reverse transcription from heterochromatic repeats in S. pombe are many, a role for RNAi itself in imprinting and X inactivation has so far proved elusive. Nonetheless, introduction of siRNA into cancer cell lines can result in chromatin being marked with H3K9me2 at homologous promoters (Ting et al. 2005). In some cases, it can also result in DNA methylation (Morris et al. 2004), perhaps mediated by direct binding of small RNA with DNA methyltransferases and DNA methylation binding proteins (Jeffery and Nakiely 2004). Finally, Dicer knockout vertebrate cell lines have chromosome segregation defects reminiscent of those found in fission yeast mutants, accompanied by changes in heterochromatic morphology, expression of satellite repeats, and mislocalization of cohesin (Fukagawa et al. 2004; Kanellopoulou et al. 2005).
10 Concluding Remarks

The possibility that genes may be regulated by small RNA molecules was suggested over 40 years ago (Jacob and Monod 1961), as well as the notion that “control RNA” might be related to repeats (Britten and Davidson 1969). Since the identification of the lambda and lac repressors as site-specific DNA-binding proteins in Escherichia coli and the infecting bacteriophage lambda (Gilbert and Muller-Hill 1966; Ptashne 1967), studies of gene regulation have focused almost exclusively on the role of nucleic-acid-binding proteins as specificity factors. The discovery of small RNA molecules as specificity agents in diverse RNA silencing mechanisms now clearly establishes a role for RNA as a sequence-specific regulator of genes and their RNA products. Studies in fission yeast, Arabidopsis, and other model organisms have revealed a surprisingly direct role for small RNAs in mediating epigenetic modifications of the genome that direct gene silencing and contribute to heterochromatic domains necessary for genome stability and nuclear division. Many important mechanistic questions remain at large, and future studies are likely to provide more surprises about how RNA regulates gene expression.

References


Lippman Z., May B., Yordan C., Singer T., and Martienssen R. 2003. Distinct mechanisms determine transposon inheritance and methyla-


Partridge J.F., Scott K.S., Bannister A.J., Kouzarides T., and Allshire R.C. 2002. cis-acting DNA from fission yeast centromeres mediates his-


Epigenetic Regulation in Plants

Marjori Matzke and Ortrun Mittelsten Scheid

Gregor Mendel Institute of Molecular Plant Biology, Austrian Academy of Sciences, A-1030 Vienna, Austria

CONTENTS

1. Benefits of Plants in Epigenetic Research, 169
   1.1 Plants and Mammals Are Similar in Terms of (Epi)Genome Organization, 169
   1.2 Plants Provide Additional Topics for Epigenetic Research, 169
   1.3 Plants Tolerate Methodological Approaches That Are Difficult in Mammals, 170
   1.4 Plants Have a Proven Record of Contributing to Epigenetic Research, 174

2. Molecular Components of Chromatin in Plants, 175
   2.1 Regulators of DNA Methylation in Plants, 175
   2.2 Histone-modifying Enzymes, 177
   2.3 Other Chromatin Proteins, 178

3. Molecular Components of RNAi-mediated Gene Silencing Pathways, 180

3.1 Elaboration of RNAi-mediated Silencing in Plants, 180
   3.2 Pathway 1: Transgene-related Posttranscriptional and Virus-induced Silencing (PTGS/VIGS), 181
   3.3 Pathway 2: Regulation of Plant Development by RNAs and Trans-acting siRNAs, 183
   3.4 Pathway 3: Transgene-related Transcriptional Silencing, RNA-directed DNA Methylation, and Heterochromatin Formation, 184

4. Epigenetic Regulation without RNA Involvement, 186

5. Outlook, 186

References, 187