Epigenetics: heterochromatin meets RNAi

Ingela Djupedal[1], Karl Ekwall[1]

[1]*Department of Biosciences and Medical Nutrition, Karolinska Institutet, Sweden & School of Life Sciences, University College Södertörn, NOVUM, 14157 Huddinge, Sweden*

The term epigenetics refers to heritable changes not encoded by DNA. The organization of DNA into chromatin fibers affects gene expression in a heritable manner and is therefore one mechanism of epigenetic inheritance. Large parts of eukaryotic genomes consist of constitutively highly condensed heterochromatin, important for maintaining genome integrity but also for silencing of genes within. Small RNA, together with factors typically associated with RNA interference (RNAi) targets homologous DNA sequences and recruits factors that modify the chromatin, commonly resulting in formation of heterochromatin and silencing of target genes. The scope of this review is to provide an overview of the roles of small RNA and the RNAi components, Dicer, Argonaute and RNA dependent polymerases in epigenetic inheritance via heterochromatin formation, exemplified with pathways from unicellular eukaryotes, plants and animals.

Keywords: epigenetics, heterochromatin, RNAi, siRNA, Argonaute, Dicer, RNA dependent polymerase
Cell Research (2009) **19**:282-295. doi: 10.1038/cr.2009.13; published online 3 February 2009

Introduction

In eukaryotes, the DNA is wrapped around histone proteins, forming chromatin fibers that coil up on themselves for further compaction. This arrangement affects the accessibility and activity of the DNA, which needs to be transcribed or at least replicated prior to each cell division. Moreover, the chromatin fibers need to be reconstituted at each cell division and different chromatin states may then be passed on to the daughter cells.

Histones are globular proteins with protruding N-terminal tails. There are four canonical types of histones, H2A, H2B, H3 and H4. Two copies of each histone together with 146 bp of DNA from a nucleosome – the building block of chromatin fibers. Chromatin fibers are more or less compacted depending on several factors, for example covalent modifications of the histone tails that create specific binding sites for various chromatin-binding proteins. Two forms of chromatin are generally distinguished, the accessible, early replicating and gene dense euchromatin and the highly condensed, inaccessible and late replicating heterochromatin. DNA organized into heterochromatin mostly consists of non-coding repeat sequences, and euchromatic genes that happen to be translocated into heterochromatin are epigenetically silenced. These genes were presumed to be silent due to the inaccessible structure of the heterochromatin that would hinder passage by RNA polymerase II (pol II).

Our understanding of the functions of heterochromatin and the mechanisms governing heterochromatin formation have been greatly improved over the last decades, much thanks to experiments conducted in model organisms such as the yeasts *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*, the fruit fly *Drosophila melanogaster* and the weed *Arabidopsis thaliana*. Many of the molecular mechanisms are highly conserved throughout evolution, pointing at their fundamental role in regulating and maintaining the genomes of living organisms. The most surprising discovery in the last few years was the realization that small RNA and factors from the RNAi machinery are involved in targeting protein factors that set up heterochromatin at homologous DNA sequences. The pathways of RNAi in heterochromatin establishment and maintenance have since been an exponentially developing field of research.

Hallmarks of heterochromatin are highly conserved

The repressive effect of heterochromatin on gene ex-

Correspondence: Karl Ekwall
Tel: +4686089133; Fax: +4687743358
E-mail: karl.ekwall@ki.se

pression was first investigated in *D. melanogaster* where translocation of genes into the vicinity of heterochromatic domains results in a variegated phenotype, called position effect variegation (PEV) (reviewed in [1]). In some clones of cells, heterochromatin will spread and cover the gene and thereby turn off its expression, whereas in other clones, the gene will remain euchromatic and expressed [2]. However, not all genes are inactivated by heterochromatin. There are hundreds of genes in *D. melanogaster* that are naturally residing within heterochromatin and display typical heterochromatic marks (H3K9me) while remaining actively transcribed [3]. Hence, the effects of heterochromatin on gene expression are variable, pointing at the complex nature of gene regulation.

The phenomenon of PEV has been exploited as a successful experimental setup to identify factors involved in the formation of heterochromatin. These factors are, in many cases, remarkably well conserved in distantly related species [4-10]. Histone Lysine Methyl Transferases (KMTs), enzymes that add one, two or three methyl groups to lysine residues within the tails of histones, as well as proteins with chromodomains that bind specifically to motifs generated by KMTs, are the best characterized factors required for setting up heterochromatin. Di- or trimethylation of lysine 9 of histone H3 (H3K9me) is a typical mark of heterochromatin, conserved in fungi, plants and animals. Other common modifications in heterochromatin are methylation of lysine 20 of histone H4 (H4K20me) and methylation of lysine 27 of histone H3 (H3K27me). These modifications are combined with a total hypoacetylation of the residues in histone tails as compared to euchromatin and therefore histone deacetylases (HDACs) are important enzymes for heterochromatin formation and propagation.

In euchromatin, histones H3 and H4 are generally highly acetylated and enriched for methylation of lysine 4 of histone H3 (H3K4me). Genome-wide chromatin-immunoprecipitation experiments (ChIP on chip) in *S. pombe* have demonstrated that H3K4me of euchromatin and H3K9me of heterochromatin are mutually exclusive [11].

The family of KMTs responsible for H3K9me contains a SET domain, which confers methyl transferase activity, and a chromodomain [12, 13]. Methylation of H3K9 creates a binding site for Heterochromatin Protein 1 (HP1), another protein originally found in screens for suppressors of PEV [4]. HP1 has one chromodomain joined by a hinge to a second chromoshadow domain [14, 15]. Chromodomains are found in several protein families, all involved in the transmission of stable chromatin states. The chromodomain recognizes and binds specifically to histone modifications, or in the case of the

chromoshadow domain, promotes self-assembly [16]. Intriguingly, some chromodomains i.e. MOF also interact with non-coding RNA [17]. KMTs specific for H3K9 have been found in complexes with HP1. Methylation of H3K9 by KMTs thus provides a binding site for HP1 that, in turn, can recruit more KMTs and HP1 leading to propagation and spreading of heterochromatin.

Not only histones are covalently modified in heterochromatin. Methylation of cytosine residues of the DNA is a common feature of silenced genes in heterochromatin in many eukaryotes. However, DNA methylation has not been found in the common model organisms, the yeasts *S. cerevisiae* and *S. pombe* or the nematode *Caenorhabditis elegans*, and only low levels are detected in the fruit fly *D. melanogaster* [18] (Table 1). *De novo* DNA methylation of cytosines can be targeted by methylation of H3K9 or by homologous RNA [19]. It is easy to visualize how DNA methylation is inherited from mother cell to daughter cell considering that DNA replication is semi-conservative and the newly synthesized strand may be methylated with the "old" strand as template [20, 21]. Likewise, half of the histones are partitioned to each DNA helix during S-phase, and may thereby guide histone modifications to newly incorporated histones. This would provide means for maintenance of the chromatin setting over cell divisions. However, histone deposition to the DNA is a dynamic process that occurs at all stages of the cell cycle [22].

In most eukaryotes, centromeres, telomeres and rDNA are typical heterochromatic regions. The sequence of centromeric and telomeric DNA is generally not conserved but consists of repetitive elements, remnants of transposons, and very few genes. The structure of both centromeres and telomeres is vital for ensuring proper segregation of sister chromatids during cell divisions and thus for keeping the genome intact. Failure to maintain a silenced state is often detrimental to higher organisms, and causes diseases such as cancer or deregulation of development resulting in embryonic death. For example, epigenetic activation of oncogenes and/or suppression of tumor-suppressor genes are common causes of cancer. Heterochromatin also plays a pivotal role in protecting the genome from parasitic DNA elements, such as transposons, which are "shut off" or silenced by heterochromatin. Heterochromatin is associated with silencing of genes and, in the case of X inactivation in female mammals, entire chromosomes (for review [23]).

RNAi and epigenetics

The 1998 discovery that double stranded RNA mediates gene silencing (RNAi) [24] was awarded the 2007

Table 1 Compilation of features of heterochromatin and RNAi from yeast, animals and plants

	Fungi		Ciliates	Animals			Plants
	S. pombe	*N. Crassa*	*T. thermophila*	*C. elegans*	*D. melanogaster*	*H. sapiens*	*A. thaliana*
H3K9me	Yes	Yes	Yes	Yes	Yes	Yes	Yes
HP1	Swi6, Chp1 & Chp2	HP1	Pdd1p &Pdd3p	HPL1 &HPL2	HP1, HP1b & HP1c	HP1α, β & γ	TFL2/LHP1
DNA methylation	–	Yes	Yes	–	Low levels	Yes	Yes
short RNA	siRNA	siRNA	scnRNA	miRNA & siRNA	miRNA, siRNA, piRNA	miRNA, siRNA & piRNA	miRNA & siRNA
Dicer	Dcr1	DCL1 & DCL2	DCL1	Dcr1	Dcr1 & Dcr2	DCR1	DCL1, DCL2, DCL3 & DCL4
Argonaute	Ago1	QDE2 & SMS2	TWI1	27 AGO genes	Ago1-3, Piwi1-4	AGO1-4, PIWI1-4	AGO1-10
RDP	Rdp1	QDE1, SAD1 & RRP3	–	Ego1, Rrf1-3	–	–	RDR1-6
ds siRNA ribonuclease	Eri1	n.d.	–	Eri1	CG6393	THEX1	–
pol IV	–	–	–	–	–	–	NRPD1a & b

n.d.: Not determined

Nobel Prize in Medicine or Physiology. Today, RNAi has become a standard laboratory technique for investigating gene function by knockdown of the corresponding mRNAs. The molecular mechanisms of post-transcriptional gene silencing by endogenous microRNA (miRNA) genes through the degradation and/or interference with the translation of their mRNA targets have to a great extent been elucidated (for review [25]). Interestingly, many of the factors that mediate posttranscriptional silencing via RNAi also have functions in transcriptional gene silencing.

Ever since the late 1980's, there have been several reports of phenomena that later turned out to fall into this category. Introduction of antisense RNA was believed to mediate suppression of gene expression via specific binding to matching mRNAs and thereby interfering with their translation into proteins. Unexpectedly, the negative control in the form of the corresponding sense RNA also caused a certain degree of silencing [26]. In plants, RNA viruses were reported to direct DNA methylation of homologous genes [19] and introduction of multiple transgene copies resulted in silencing, also termed cosuppression, via DNA methylation of the endogenous gene copy [27, 28]. In the yeast *S. pombe*, small RNA and components of the RNAi pathway were demonstrated to be required for heterochromatin formation at homologous

DNA at centromeres [29, 30]. In the ciliate *Tetrahymena thermophila*, small RNA together with an Argonaute family protein were reported to be involved in generating H3K9me at histones at homologous DNA regions with subsequent elimination of DNA [31, 32]. Small RNA and factors required for RNAi, such as Dicer and Argonaute have since been implicated in the process of heterochromatin formation in various organisms, indicating that their role in the targeting and recruitment of factors such as KMTs and HP1 are generally conserved. In *D. melanogaster*, mutations of proteins belonging to the Argonaute family have been shown to reduce H3K9me as well as HP1 binding within heterochromatin [33] and Ago2 mutants show defects in centromeric heterochromatin formation in embryonic development [34]. Dicer mutants have been reported to display chromosome segregation defects in a human-chicken hybrid cell line [35] as well as loss of silencing at centromeres in mouse cell lines [36]. Furthermore, short interfering (siRNA) and RNAi have been demonstrated to trigger long-term, heritable gene silencing in *C. elegans* [37]. Silenced gene expression was inherited in about one third of the progeny, in some cases indefinitely, in a dominant fashion. Importantly, four genes encoding chromatin remodeling factors were required for maintenance of the silent state whereas the HDAC inhibitor, trichostatin A, reversed silencing [37].

Hence, in many eukaryotes siRNA/RNAi can induce epigenetic changes in gene expression via modifications of chromatin.

Small RNA – mediators of RNAi

There are several classes of 20-30 nucleotide long, short RNA, with different names depending on their genomic origin, their route of synthesis and their mode of action (for review [38]). The miRNAs typically direct post-transcriptional gene silencing (PTGS) via translational repression and differ from siRNA in that they originate from miRNA genes which code for precursors with potential to form local hairpin structures, with one distinct miRNA accumulating from each hairpin [39]. In addition, the miRNA genes are often conserved in between species [39]. Endogenous short interfering RNA (siRNA) originates from repetitive sequences, such as centromeric repeats, transposable elements or rDNA. Several different siRNA molecules along the length of the double-strand are detectable [39] and make up a considerable fraction of the short RNAs that have been cloned [29, 40-44]. The length of the siRNA varies between species, usually in between 19-22 nucleotides up to 26 nucleotides, and depends on the specificity of the enzymes that produce them. In *S. pombe*, centromeric siRNAs correspond to both strands of the DNA and do not show any strand bias. Genome wide protein binding maps in *S. pombe* have shown a strong correlation of binding of the RNAi components Ago1 and Rdp1 to the loci which generate siRNA [11]. Deep sequencing of seedlings and inflorescence of *A. thaliana* has also identified siRNA matching genes and intergenic regions with the genes represented showing a good distribution to many cellular processes and functions [41]. siRNA may silence genes at the post transcriptional level as well as at the transcriptional level and are known to associate with other Argonaute isoforms than the miRNA. However, in *D. melanogaster*, structural features of the double-stranded small RNA precursor, i.e. presence or absence of central mismatches, have been demonstrated to govern which Argonaute protein the small RNA will bind [45]. Likewise, in *A. thaliana*, miRNAs were demonstrated to preferentially start with a 5′-terminal uridine and bind to AGO1 whereas siRNAs associating with AGO2 and AGO4 generally started with a 5′-terminal adenosine [46]. Furthermore, substitution of the 5′-terminal nucleotide caused predictable changes in small RNA Argonaute association [46]. Thus, structure and sequence of small RNA are important determinants for their biological activity.

The amplification model, suggested by Yigit *et al.* [47,

48], makes a distinction between rare primary siRNA "triggers", products of Dicer processing of double stranded precursors, and abundant secondary siRNAs, synthesized by RNA dependent polymerases (RDPs). Secondary siRNAs promote the major effects on silencing of mRNAs but also, likely, on transcriptional silencing and heterochromatin formation [48]. Primary and secondary siRNAs have been demonstrated in *C. elegans*, having different modifications at the 5′ends [49, 50]. Primary siRNAs, which are the products of Dicer activity, have 5′-monophosphates and 3′-hydroxy groups whereas secondary RNAs are *de novo* synthesized by RDPs and have triphosphates at their 5′ end [49, 50]. In plants, as well as the single-cell algae *Chlamydomonas reinhardtii*, the 3′-end of miRNA/siRNA is methylated and thereby protected from degradation [43, 50, 51]. This feature is not conserved in *C. elegans* or *D. melanogaster* where the 3′-ends investigated were unmodified and homologues of the responsible methyl-transferase do not seem to be conserved outside the kingdom of plants [52].

In addition to miRNA acting on the post-transcriptional level and siRNA acting on chromatin, there is a third major class of small RNA – the Piwi-interacting RNA (piRNA). piRNAs are typically longer than miRNA and siRNA (25-29 nucleotides in *D. melanogaster*, 25-31 nucleotides in rat [53]), preferentially have a uridine at their 5′-terminus, originate from transposable elements or heterochromatic repeats and do not depend on Dicer for their synthesis [54, 55]. Instead, piRNAs are proposed to be generated via sequential cleavage by distinct Piwi-family proteins [56]. The piRNA pathway is required in the germline of animals for silencing of transposable elements. A parallel mechanism in the seedlings of plants, which do not have Piwi homologues, is the apparent use of transcriptional silencing via siRNA and Argonautes to control transposons. In plants, deep sequencing of small RNAs has revealed greater complexity in the ranks of siRNA in the germline and inflorescence tissue [41].

Argonautes, Dicers and RNA dependent polymerases

The major constituents of RNAi are, aside from the small RNA species, Argonaute and Dicer proteins (Table 1). Common features of Dicer proteins are the two RNAse III domains, which specifically cleave double stranded RNA, an N-terminal helicase domain and a PAZ domain [57]. The PAZ domain was first found among proteins of the Argonaute family [58] and is not present in Dicer proteins from fungi and protozoa [59]. The function of the PAZ domain is to bind the double stranded 5′-end of the siRNA precursor. The distance between the PAZ domain and the two RNAse III domains of Dicer

determines the length of the siRNAs [60]. Each RNAse III domain will mediate the endonucleolytic cleavage of one strand of the RNA and thus generate a new 3′-end. The 2-nucleotide 3′-end overhang typical of Dicer products are a consequence of the alignment of the two RNAse III domains on the double stranded RNA [61]. The cleavage reaction thus produces a new 5′-end to which the PAZ domain may bind and Dicer may processively generate new double stranded siRNA from the same siRNA precursor.

Multicellular organisms commonly have several isoforms of Dicer proteins which function in different RNAi pathways. Another class of enzyme containing an RNAse III motif is represented by Drosha, which is involved in the miRNA biogenesis of animals. There are no homologues to Drosha in plants, in which DCL1, one of the four Dicer proteins mediates pri-miRNA cleavage [62]. The cellular localization of Dicer has, in *S. pombe,* been demonstrated to be mainly cytoplasmic [63]. Others have, in the same organism, demonstrated a physical coupling between Dcr1 and RDP complexes at the centromeric heterochromatin within the nucleus [59] and the coupling of siRNA synthesis and transcriptional silencing has important implications for restriction of silencing in *cis*. In *A. thaliana*, Dicer-like 1-3 (DCL1, DCL2 and DCL3) are primarily localized in the nucleus [62].

The following step in the RNAi pathway is loading of the double stranded siRNA into protein complexes of which members of the Argonaute family are constitutive parts. Examples are RISC (RNA Induced Silencing Complex) or RITS (RNA-induced Initiation of Transcriptional Gene Silencing) [59, 64, 65]. The first member of the Argonaute family was identified in *A. thaliana* as a protein required for stem cell maintenance and development [66]. The protein family consists of three groups; the Argonaute group represented in all three kingdoms of life, the Piwi group with representatives in metazoans, and a third group exclusive to *C. elegans* which has as many as 27 different Argonaute encoding genes. The Argonaute proteins of this small worm have different patterns of expression in time and tissue, as well as distinct functions in chromosome segregation, development, embryo viability and fertility [47].

Argonaute proteins contain a PAZ domain that binds the 3′-end of the siRNA duplex, a MID domain, which binds the 5′-phosphate, and an RNAse H domain, which may direct endonucleolytic cleavage, or slicing, of RNA, like Dicer generating a 5′-phosphate and a 3′-hydroxy group [38, 61]. Once the double stranded siRNA is bound, Argonaute unwinds it and selectively retains the strand with the weakest base pair at the 5′-end [67]. The siRNA is then used as a guide for targeting homologous

RNA (or DNA) that may or may not be cleaved. The outcome varies depending on the type of Argonaute and protein complex involved. The RNAse H motif of many Argonautes appears to lack catalytic activity [47].

A third major type of RNAi protein is the RNA dependent RNA polymerases (RDPs). They are present and required for RNAi in fungi, nematodes and plants, but have not been identified in insects or vertebrates. It cannot be excluded, however, that some other, yet unidentified, enzyme fulfills the same function. For example, labeled siRNAs in extracts from *D. melanogaster* are, together with complementary mRNAs, incorporated into double stranded RNAs, which is subsequently processed into new small RNAs [68]. In mouse oocytes, siRNAs are generated from naturally occurring double-stranded RNAs, such as inverted repeats and antisense transcripts and have been shown to affect the expression of both mRNA and retrotransposons [69].

Biochemical investigation of an RDP from the fungus *Neurospora crassa* has demonstrated this enzyme to be capable of both primer-dependent and primer-independent synthesis of RNA from a single stranded RNA template, with primer-independent activity being more efficient. The enzyme could not support synthesis from a double stranded RNA template. Full-length copies complementary to the template were recovered, but the enzyme seemed to be most effective in generating 9-21 nucleotide long siRNA along the length of the template [70]. Full-length double-stranded RNA would be substrates for Dicer whereas the short RDP products proposedly would be directly incorporated into Argonaute complexes, guiding them towards complementary transcripts. Indeed, secondary siRNAs that are RDP products represent the most abundant class of small RNAs in *C. elegans* [49, 50].

The single RDP of *S. pombe* has been reported to mainly localize to the nucleus, at the centromeric heterochromatin, where it is proposed to use nascent centromeric transcripts as templates [30]. A small pool of the enzyme can also be detected in the cytoplasm where it seems to be associated with Argonaute and Dicer proteins [63]. Other organisms have several genes for RDP enzymes which function in different RNAi pathways, such as viral defense, PTGS and heterochromatin formation, with some isoforms localized to the cytoplasm and others within the nucleus [71].

In summary, representatives of the Dicer and Argonaute families are present in all organisms displaying RNAi whereas the RNA dependent polymerases are not required for RNAi in insects and vertebrates. The large number of Dicer and Argonaute isoforms in more complex organisms illustrates the many different pathways

and molecular mechanisms in which they are involved. Several Argonaute isoforms appear to lack endonucleolytic activity but are reported to increase the levels of small RNA, proposedly via binding and thereby stabilization of the small RNA. Transfection of constructs expressing the four human argonaute proteins (AGO1-4) has been reported to increase levels of mature miRNA, although only AGO2 has a complete RNAse H domain [72]. Likewise, in *C. elegans*, overexpression of Argonaute proteins predicted to be catalytically inactive are reported to enhance silencing [47]. The four Dicer isoforms of *A. thaliana* have been, when exposed to high doses of siRNA, reported to partially substitute for each other [73]. Hence, there is some functional redundancy between members of the same protein family.

S. pombe as a model for RNAi directed heterochromatin formation

The yeast *S. pombe* has one copy each of Dicer (Dcr1), Argonaute (Ago1) and an RDP (Rdp1) and the role of RNAi in heterochromatin formation was first discovered when knockout strains of Dcr1, Ago1 and Rdp1

displayed phenotypes typical for mutations in proteins involved in heterochromatin formation such as the KMT Clr4^{KMT1}, responsible for H3K9me, and the HP1 homologue Swi6. The RNAi knockout strains had decreased levels at the centromeres of H3K9me as well as Swi6 binding [30]. In addition, these strains had chromosome segregation defects [74]. The mutants also showed accumulation of transcripts derived from centromeres.

The current model for the role of RNAi in heterochromatin formation in *S. pombe* states that centromeric repeats are transcribed by pol II at the S-phase of the cell cycle [75-78] (Figure 1). Double-stranded RNAs are generated either via bidirectional transcription of the repeated sequences of the centromeres, the RNA molecule folding back on itself or by the activity of RNA dependent RNA polymerase on centromere-derived RNA templates (as shown in Figure 1). The double stranded precursor RNA is recognized and processed into small, ~25 nucleotide long [79], RNAs which are incorporated into protein complexes containing Argonaute proteins [65, 80]. The double stranded RNA is then sliced into an siRNA and a passenger strand that is degraded [67]. The Argonaute complex loaded with siRNA then targets
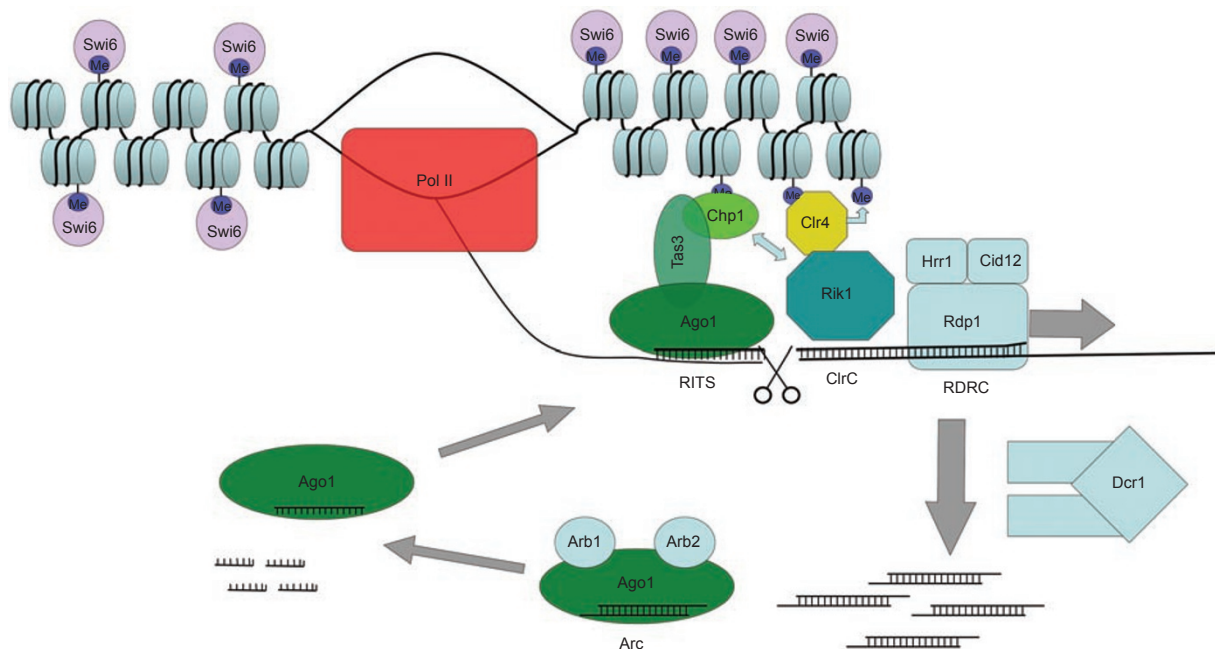

Figure 1 Simplified model of the RNAi and heterochromatin feedback loop of *S. pombe*. Pol II transcription of the centromeric repeats leads to RITS recruitment via complementary siRNA and recognition of H3K9me. Ago1 cleaves the nascent RNA and recruits ClrC and RDRC. Clr4 of ClrC both binds and produces H3K9me (blue circles of histones), Rik1 interacts with Chp1 and potentially with nascent RNA. RDRC synthesizes dsRNA templated by the nascent RNA, producing either full-length transcripts which are Dcr1 substrates, or secondary siRNA that is potentially loaded directly into RITS. Primary, Dcr1-generated duplex siRNAs are incorporated in the Arc complex and subsequently, Ago exits Arc, slices the siRNA and together with Chp1 and Tas3 forms another RITS complex, ready for targeting homologous RNA.

KMTs and HP1 to homologous DNA sequences [30].

Genome-wide protein binding maps have demonstrated that Ago1 and Rdp1 preferentially bind to the same DNA regions that carry the H3K9me mark, i.e., the centromeres, telomeres and silent mating type loci [11]. In the same study Ago1 was chromatin-immunoprecipitated at rDNA loci where RNAi components also are involved in heterochromatin silencing. The investigation of RNAi in *S. pombe* has mainly focused on the molecular mechanisms of heterochromatin formation at the centromeres. Maintenance of heterochromatin at telomeres and the silent mating type loci is achieved by redundant, RNAi-independent mechanisms whereas disruption of RNAi affects both establishment and maintenance of the heterochromatin at centromeres [81].

The single copy RNAi genes of *S. pombe* function not only in setting up heterochromatin, but also in post-transcriptional gene silencing. Introduction of an siRNA hairpin construct directed to a GFP reporter gene caused a reduction in mRNA levels but not in the rate of transcription and the effect did not depend on Swi6 [82]. However, gene expression profiling experiments with RNAi knockout strains have revealed that relatively few genes depend on the RNAi pathway for their expression or repression, as compared to mutants or knock-outs of the classical heterochromatin proteins such as Clr4^{KMT1} [83]. The levels of eight proteins have been shown to be controlled by Dcr1 in *S. pombe* [84]. A novel function of RNAi has recently been discovered by its requirement for *de novo* establishment of CENP-A chromatin at the centromere [85]. CENP-A is a specialized Histone H3 variant, found at active centromeres and functioning as a platform for kinetochore assembly, and is highly conserved in eukaryotes [86]. Interestingly, RNAi and heterochromatin are not required for maintaining CENP-A chromatin once it has been established.

The combination of being an established model organism for formation of heterochromatin with a small and easily manipulated genome, centromeres akin to those of multicellular eukaryotes and single copies of the RNAi components has made *S. pombe* one of today's most well-known systems for studying how RNAi directs heterochromatin formation. Several components of the system have been identified and the molecular mechanisms are being dissected.

Silencing requires transcription

In *S. pombe*, there is an absolute requirement for H3K9me and the responsible KMT Clr4^{KMT1} in silencing [87, 88]. Transcription by pol II at centromeric repeats has been demonstrated to lead to recruitment of factors that modify chromatin, resulting in generation of the H3K9me mark [77]. It has recently been demonstrated that Clr4^{KMT1} not only creates the H3K9me mark but also recognizes and binds H3K9me via its chromodomain [89]. The ClrC complex, of which Clr4^{KMT1} is a part, is recruited to nascent transcripts in an RNAi-dependent manner via its putative RNA-binding component Rik1. The same protein also mediates a direct interaction with RITS. Hence, there is extensive interdependency between the generation of H3K9me marks and the RNAi system, making it difficult to distinguish what comes first.

The requirement for transcription has been elegantly demonstrated by tethering of the RITS subunit Tas3 to a reporter gene transcript. This was sufficient for heterochromatin formation and silencing in *cis* [90]. A homologous gene copy at another locus was not silenced, however. In order for this to occur, Eri1, a ribonuclease homologous to a protein originally identified as a negative regulator of RNAi in *C. elegans* [91], had to be knocked out. Only 1 of 10 000 *eri1Δ* cells achieved silencing of the reporter gene in *trans* but the silencing was stably propagated once initiated [90].

Ago1 of *S. pombe* has been identified in two complexes, the Argonaute siRNA Chaperone Complex (Arc) [92] and the RITS complex [64]. Arc consists of Ago1 together with Ago1-binding proteins 1 and 2 (Arb1 and Arb2) and incorporates only double stranded siRNA, presumably direct products of Dcr1. The RITS complex consists of Ago1 and single stranded centromeric siRNA together with the chromodomain protein Chp1, which recognizes and binds H3K9me, and Tas3 which in turn recognizes and binds Ago1 [93]. Chp1 interacts with Tas3 independently of Ago1 or siRNA [94]. Mutations in Tas3 have different consequences depending on which protein-binding domain they affect. A Tas3 mutation that abolishes binding of Ago1 can support maintenance but not re-establishement of centromeric heterochromatin. The protein will remain localized to heterochromatin due to the Chp1 interaction with H3K9me and Ago1 can independently localize to cognate loci guided by siRNA [95]. A Tas3 mutation that interferes with Chp1 binding, however, has a more severe phenotype. No Tas3 or Ago1 is recruited to the centromeres with loss of all subsequent steps in the RNAi pathway, resulting in failure to form heterochromatin and chromosome missegregation [88]. The recruitment of RITS complex therefore depends, primarily, on the H3K9me mark as well as guidance by siRNA. This co-dependency on siRNA and H3K9me has been suggested to stringently control heterochromatin formation at proper locations, as both criteria have to be fulfilled in order for heterochromatin to spread and silencing to be achieved.

The RITS complex recruits the Rdp1-containing complex RDRC (RNA-Directed RNA polymerase Complex) [87]. RDRC consists of Rdp1, Hrr1, a helicase that interacts with Ago1 in a Clr4-dependent manner, and Cid12 which is a polyA polymerase. The RDRC is capable of primer independent *in vitro* synthesis of full-length transcripts from a single stranded RNA template. RDRC recruited by RITS to loci with ongoing transcription may then use the nascent RNA as template for generating double stranded RNA and/or new siRNA. Although Dcr1 has not been chromatin-immunoprecipitated at the centromeres, it has been co-immunoprecipitated with RDRC. Furthermore, Dcr1 activity stimulated dsRNA synthesis by RDRC in a reconstituted system [59], implying that generation of siRNA might occur in *cis*.

The entire RNAi pathway thus seems to be assembled at the sites of heterochromatin formation in *S. pombe*. It may seem contradictory that the bulk mass of Ago1 and Dcr1 is localized to the cytoplasm considering the relatively limited role of PTGS in this organism [63]. However, the cytoplasmic distribution of these proteins may have a regulatory function in limiting the levels of RNAi enzymes in the nucleus and thus restricting the formation of heterochromatin to the proper sites of the genome.

RNAi and chromatin in plants

Like in *S. pombe*, there is transcription of heterochromatic, centromeric retrotransposons in rice combined with generation of corresponding small RNAs [96]. In addition, methylation of H3K9 and DNA methylation at centromeres and transposons are affected by mutations in RNAi components [97].

Plants have an extensive array of RNAi pathways, including Virus-Induced Gene Silencing (VIGS) and a vast number of endogenous miRNA genes crucial for control of plant physiology and development. Silencing can also spread in between cells. The genome of *A. thaliana* has four Dicer genes, ten genes encoding Argonaute proteins and six paralogues for RNA-dependent RNA polymerases (Table 1) [47, 98]. The various RNAi pathways in plants are interrelated and partially overlapping ([62] and references therein), and accumulation of double stranded precursor RNA above a threshold value can result in Dicer enzymes from parallel pathways substituting for each other [73]. Although plants were among the earliest systems investigated for RNA-mediated silencing of transcription, the large number of RNAi enzymes combined with the high level of interconnection between different RNAi pathways have made it difficult to produce a comprehensive picture of the phenomenon.

Still a large number of enzymes involved in RNAi

in plants have been identified and the complexity of the system is beginning to be dissected [73]. A common consequence of all plant RNAi pathways is RNA-dependent DNA methylation of cytosines and thereby silencing of genes at the transcriptional level [99]. The methylation is confined along the length of the transcribed regions. Introduction of inverted repeat sequences that fold back on themselves produces double stranded RNAs, which are processed into siRNAs and cause DNA methylation of homologous promoters resulting in epigenetic silencing of the downstream gene [100, 101]. In two recent studies genome-wide DNA methylation maps of *A. thaliana* were compared to the occurrence of small RNAs [102, 103]. siRNA clusters showed an 80% overlap with CG methylation and there was good correlation between siRNA abundance and CG methylation at the transcribed regions of inverted and tandem repeats [102]. Analysis of DNA methyl transferase mutants revealed that decreases of DNA methylation generally were accompanied by a reduction in the associated small RNA and vice versa [103]. Yet, only one third of the genomic DNA methylation was correlated with the incidence of small RNA, indicating that other mechanisms are of great importance in targeting DNA for methylation [103]. The reported requirement for a putative HDAC suggests that heterochromatic marks may help propagate some forms of DNA methylation [104].

The chromatin modifying RNAi pathway in plants depends on a fourth "silencing" RNA polymerase – only conserved among plant species – for its integrity [105-107]. Duplicate gene copies with homology to the two largest subunits of the canonical RNA polymerases I, II and III, have been identified. Both copies of the largest subunit, NRPD1a and NRPD1b, are expressed but non-essential for viability. The level of conservation as compared to RNA polymerases I, II and III, is rather low including the active site [105]. Only one of the genes encoding the second largest subunit, named NRPD2a, is expressed and it has better resemblance to the other polymerases [105]. NRPD2 probably associates with both NRPD1a and NRPD1b. All three expressed pol IV subunits are involved in the synthesis of siRNA. The phenotypes of mutations in NRPD1a and NRPD1b differ, with that of NRPD1a being more severe and affecting a larger number of siRNA clusters [106]. Therefore NRPD1b, which only affects siRNA from a subset of the NRPD1a-dependent clusters, has been supposed to act downstream [40, 106]. NRPD1a and b differ mainly in their C-terminal domains. The cognate domain of pol II functions in coordinating protein interactions with various forms of RNA processing and it was therefore suggested that the different functions of NRPD1a and NRPD1b may be due

to interactions with different protein factors [107].

Deep sequencing of small RNAs from NRPD1a and NRPD1b mutants revealed transposons and centromeres as key sources of NRPD1a-dependent siRNAs [40]. Surprisingly, the effect of mutant NRPD1b on DNA methylation is separable from its effect on siRNA accumulation, indicating more complex mechanisms of pol IV action than previously anticipated. Another unexpected finding was an activating role for NRPD1a and b, where siRNA directs demethylation and thus increased expression of the genes in question [40]. The functions of RNAi in epigenetic changes of chromatin in plants therefore seem to be even more divergent than what is recognized today with potentially opposite outcomes, i.e. repression or activation, depending on as yet unknown factors in the local context.

RNAi and chromatin in mammals

There are several known forms of RNA-mediated regulation of animal genomes. For example, long noncoding RNAs are part of the mechanisms that set up imprinting of genes, thus creating epigenetic asymmetry between parental genomes, and regulate X-chromosome inactivation (reviewed in [23, 108]).

The large genomes of multicellular organisms have a low percentage of coding genes in combination with long arrays of non-coding nucleotides and repeat elements. The emergence of powerful sequencing techniques has revealed an unanticipated complexity in the transcriptome of animals. Extensive analysis of 1% of the human genome has demonstrated that the majority of base pairs are transcribed [109]. Transcripts are derived from non-coding and silent regions and may overlap with protein-coding transcripts. In addition, most transcription units seem to have several transcription start sites. Transcription of both strands of the DNA is widespread, with 72% of transcription units having sense-antisense pairs with effects on each other's regulation [110]. This regulatory effect may be achieved either via collision of RNA polymerase II complexes at the DNA template or by siRNA production from the antisense strand.

A role for RNA molecules together with histone modifications in generating a higher-order heterochromatin structure at centromeres in mammalian cells has been demonstrated [111]. Knockout of Dicer in mice causes early embryonic death due to failure to maintain stem cells [112, 113]. Dicer has also been reported to be essential for heterochromatin formation. Dicer depletion causes accumulation of centromere-derived transcripts, and a reduction in the levels of siRNA as well as di- and trimethylation of H3K9, with chromatid segregation

defects as a consequence [35]. Thus, centromeres are transcribed with concomitant siRNA production and subsequent heterochromatin formation via methylation of histones and DNA. Like in *S. pombe*, there are also additional RNAi-independent mechanisms participating in the establishment of heterochromatin in mammals [114].

Elucidation of the mechanism of RNAi-mediated heterochromatin targeting by introduction of siRNAs homologous to promoters of reporter genes has demonstrated that two of the four human Argonaute proteins, AGO1 and AGO2, mediate transcriptional gene silencing via interactions with pol II and KMTs specific for H3K9me and H3K27me [113, 115]. There is a requirement for active transcription by pol II as evidenced by the fact that addition of alpha-amanitin inhibits silencing [116]. A number of human genes have been shown to give rise to low amounts of transcripts from the upstream promoter region and these low-abundance transcripts are necessary for siRNA-induced transcriptional silencing, thus providing evidence that siRNA acts via base pairing to nascent RNA [117].

Unexpectedly, in some cases introduction of certain exogenous siRNA species causes activation of transcription rather than repression [118, 119]. Similarly to gene silencing, gene activation is achieved with duplex siRNAs homologous to the promoter regions but is associated with loss of heterochromatin marks. The exact mechanism for activation rather than repression is poorly understood. A requirement for AGO2 has been reported [118] whereas others claim that gene activation is not associated with increased AGO1 or AGO2 recruitment [119]. Exogenous siRNAs are thus capable of stable and specific epigenetic regulation of target genes, raising the possibility that similar mechanism may be employed by endogenous siRNAs.

Indeed, the previously mentioned sense-antisense pairs provide an example of endogenous RNA-directed epigenetic silencing. Investigation of a naturally occurring sense/antisense pair of a tumor suppressor gene frequently deregulated in several types of cancers revealed that an antisense construct could silence corresponding sense transcription in *cis* and *trans* [120]. Expression of antisense RNA induced silencing via heterochromatin formation, but not DNA methylation, independently of Dicer function [120].

Other examples of endogenous RNAs regulating gene expression include naturally occurring RNA species derived from introns and spacer regions between the rDNA genes, and transcripts containing mutations generating premature translation termination signals. Short DNA repeats within introns are capable of transcriptional suppression of their gene of origin, thus generating a nega-

tive feed-back loop [121]. High levels of transcription will lead to an abundance of siRNA that prevents accumulation of the cognate mRNA. Likewise, transcripts derived from the intergenic spacer that separates rDNA genes have been shown to promote heterochromatin marks, for example H3K9me and HP1 association at the rDNA locus [122]. Finally, RNAi has interconnections to the nonsense-mediated decay pathway; premature translation-termination codons can induce changes in chromatin and transcriptional gene silencing which apparently can be inhibited by the mammalian homologue of the siRNAse Eri1 [123]. However, the underlying mechanisms behind these phenomena have not been elucidated in detail and it is not known whether they share all the features of typical RNAi-mediated transcriptional gene silencing as described above.

Perspective

Epigenetic modulation of transcription via the RNAi machinery can result in either repression or activation of gene expression in both plants and animals. Further investigations are needed to reveal the mechanisms underlying the different outcomes. Adding more complexity, expression and processing of miRNAs in mice have been demonstrated to affect DNA methylation via control of genes that repress DNA methyltransferase expression [124, 125]. Hence, regulation of genes and gene products is intertwined in complex patterns with different layers of regulatory components affecting each other. We are only beginning to realize the impact that these mechanisms may have on gene regulation in general and on propagation of epigenetic states.

Small RNAs are, as we have described here, integral parts of a great variety of cellular processes. The role of RNA as a regulator is not restricted to eukaryotes. Non-coding RNA regulating gene expression seems to be a universal theme. Various classes of non-coding or small RNA species with regulatory functions have been described in bacteria [126] and *Archaea* [127]. A study in *S. cerevisiae*, which lacks the classical RNAi components, has demonstrated antisense transcription as being responsible for promoting specific histone deacetylation that lead to repression of homologous genes, indicating that even though there are no homologues of either Dicer or Argonaute, RNA may still govern epigenetic inheritance of chromatin states [128]. This alternative RNA-based regulatory mechanism has a requirement for components of the exosome, a protein complex involved in degradation of aberrant RNAs. Components of the exosome have also been reported to be involved in heterochromatin formation in *S. pombe* where they work in parallel with the RNAi pathway [129]. Considering that the exosome functions in heterochromatin formation in these distantly related yeast species, it seems probable that similar mechanisms are present in other eukaryotes.

Why then, is RNA universally employed as a gene regulatory agent? As proposed by Susan Gottesman [126], RNA has the advantages of rapid synthesis and degradation – no intermediate protein regulator needs to be produced. A single small RNA can potentially modulate the expression of several genes/mRNAs as long as they share common regulatory motifs. It is likely that additional, novel RNA regulatory pathways will be discovered. Annotation of the newly discovered plethora of non-coding transcripts in organisms ranging from yeasts to human will help identify new classes of regulatory RNAs as well as implicate their mechanisms of actions. The next few years of research in this field holds great promise for new, exciting discoveries.

Acknowledgments

References

1   Ebert A, Lein S, Schotta G, Reuter G. Histone modification and the control of heterochromatic gene silencing in *Drosophila*. *Chromosome Res* 2006; **14**:377-392.
2   Zhimulev IF, Belyaeva ES, Bgatov AV, Baricheva EM, Vlassova IE. Cytogenetic and molecular aspects of position effect variegation in *Drosophila melanogaster*. *Chromosoma* 1988; **96**:255-261.
3   Yasuhara JC, Wakimoto BT. Molecular landscape of modified histones in *Drosophila* heterochromatic genes and euchromatin-heterochromatin transition zones. *PLoS genetics* 2008; **4**:e16.
4   James TC, Elgin SC. Identification of a nonhistone chromosomal protein associated with heterochromatin in *Drosophila melanogaster* and its gene. *Mol Biol Cell* 1986; **6**:3862-3872.
5   Ma J, Hwang KK, Worman HJ, Courvalin JC, Eissenberg JC. Expression and functional analysis of three isoforms of human heterochromatin-associated protein HP1 in *Drosophila*. *Chromosoma* 2001; **109**:536-544.
6   Schotta G, Ebert A, Krauss V, *et al.* Central role of *Drosophila* SU(VAR)3-9 in histone H3-K9 methylation and heterochromatic gene silencing. *EMBO J* 2002; **21**:1121-1131.
7   Baumbusch LO, Thorstensen T, Krauss V, *et al.* The *Arabidopsis thaliana* genome contains at least 29 active genes encoding SET domain proteins that can be assigned to four evolutionarily conserved classes. *Nucleic Acids Res* 2001; **29**:4319-4333.

8   Jackson JP, Lindroth AM, Cao X, Jacobsen SE. Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. *Nature* 2002; **416**:556-560.

9   Naumann K, Fischer A, Hofmann I, *et al*. Pivotal role of AtSUVH2 in heterochromatic histone methylation and gene silencing in *Arabidopsis*. *EMBO J* 2005; **24**:1418-1429.

10 Tschiersch B, Hofmann A, Krauss V, *et al*. The protein encoded by the *Drosophila* position-effect variegation suppressor gene Su(var)3-9 combines domains of antagonistic regulators of homeotic gene complexes. *EMBO J* 1994; **13**:3822-3831.

11 Cam HP, Sugiyama T, Chen ES, *et al*. Comprehensive analysis of heterochromatin- and RNAi-mediated epigenetic control of the fission yeast genome. *Nat Genet* 2005; **37**:809-819.

12 Aagaard L, Laible G, Selenko P, *et al*. Functional mammalian homologues of the *Drosophila* PEV-modifier Su(var)3-9 encode centromere-associated proteins which complex with the heterochromatin component M31. *EMBO J* 1999; **18**:1923-1938.

13 Rea S, Eisenhaber F, O'Carroll D, *et al*. Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* 2000; **406**:593-599.

14 Paro R, Hogness DS. The Polycomb protein shares a homologous domain with a heterochromatin-associated protein of *Drosophila*. *Proc Natl Acad Sci U S A* 1991; **88**:263-267.

15 Aasland R, Stewart AF. The chromo shadow domain, a second chromo domain in heterochromatin-binding protein 1, HP1. *Nucleic Acids Res* 1995; **23**:3168-3173.

16 Eissenberg JC. Molecular biology of the chromo domain: an ancient chromatin module comes of age. *Gene* 2001; **275**:19-29.

17 Akhtar A, Zink D, Becker PB. Chromodomains are protein-RNA interaction modules. *Nature* 2000; **407**:405-409.

18 Gowher H, Leismann O, Jeltsch A. DNA of *Drosophila melanogaster* contains 5-methylcytosine. *EMBO J* 2000; **19**:6918-6923.

19 Wassenegger M, Heimes S, Riedel L, Sanger HL. RNA-directed *de novo* methylation of genomic sequences in plants. *Cell* 1994; **76**:567-576.

20 Holliday R, Pugh JE. DNA modification mechanisms and gene activity during development. *Science* 1975; **187**:226-232.

21 Riggs AD. X inactivation, differentiation, and DNA methylation. *Cytogenet Cell Genet* 1975; **14**:9-25.

22 Polo SE, Almouzni G. Chromatin assembly: a basic recipe with various flavours. *Curr Opin Genet Dev* 2006; **16**:104-111.

23 Heard E, Disteche CM. Dosage compensation in mammals: fine-tuning the expression of the X chromosome. *Genes Dev* 2006; **20**:1848-1867.

24 Fire A, Xu S, Montgomery MK, *et al*. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 1998; **391**:806-811.

25 Jackson RJ, Standart N. How do microRNAs regulate gene expression? *Sci STKE* 2007; **2007**:re1.

26 Smith CJ, Watson CF, Bird CR, *et al*. Expression of a truncated tomato polygalacturonase gene inhibits expression of the endogenous gene in transgenic plants. *Mol Gen Genet* 1990; **224**:477-481.

27 van der Krol AR, Mur LA, Beld M, Mol JN, Stuitje AR. Flavonoid genes in petunia: addition of a limited number of gene copies may lead to a suppression of gene expression. *Plant Cell* 1990; **2**:291-299.

28 Napoli C, Lemieux C, Jorgensen R. Introduction of a Chimeric Chalcone Synthase Gene into Petunia Results in Reversible Co-Suppression of Homologous Genes in trans. *Plant Cell* 1990; **2**:279-289.

29 Reinhart BJ, Bartel DP. Small RNAs correspond to centromere heterochromatic repeats. *Science* 2002; **297**:1831.

30 Volpe TA, Kidner C, Hall IM, *et al*. Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science* 2002; **297**:1833-1837.

31 Mochizuki K, Fine NA, Fujisawa T, Gorovsky MA. Analysis of a piwi-related gene implicates small RNAs in genome rearrangement in *tetrahymena*. *Cell* 2002; **110**:689-699.

32 Taverna SD, Coyne RS, Allis CD. Methylation of histone h3 at lysine 9 targets programmed DNA elimination in *tetrahymena*. *Cell* 2002; **110**:701-711.

33 Pal-Bhadra M, Leibovitch BA, Gandhi SG, *et al*. Heterochromatic silencing and HP1 localization in *Drosophila* are dependent on the RNAi machinery. *Science* 2004; **303**:669-672.

34 Deshpande G, Calhoun G, Schedl P. *Drosophila* argonaute-2 is required early in embryogenesis for the assembly of centric/centromeric heterochromatin, nuclear division, nuclear migration, and germ-cell formation. *Genes Dev* 2005; **19**:1680-1685.

35 Fukagawa T, Nogami M, Yoshikawa M, *et al*. Dicer is essential for formation of the heterochromatin structure in vertebrate cells. *Nat cell biol* 2004; **6**:784-791.

36 Kanellopoulou C, Muljo SA, Kung AL, *et al*. Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. *Genes Dev* 2005; **19**:489-501.

37 Vastenhouw NL, Brunschwig K, Okihara KL, *et al*. Gene expression: long-term gene silencing by RNAi. *Nature* 2006; **442**:882.

38 Farazi TA, Juranek SA, Tuschl T. The growing catalog of small RNAs and their association with distinct Argonaute/Piwi family members. *Development* 2008; **135**:1201-1214.

39 Lim LP, Lau NC, Weinstein EG, *et al*. The microRNAs of *Caenorhabditis elegans*. *Genes Dev* 2003; **17**:991-1008.

40 Mosher RA, Schwach F, Studholme D, Baulcombe DC. PolIVb influences RNA-directed DNA methylation independently of its role in siRNA biogenesis. *Proc Natl Acad Sci U S A* 2008; **105**:3145-3150.

41 Lu C, Tej SS, Luo S, *et al*. Elucidation of the small RNA component of the transcriptome. *Science* 2005; **309**:1567-1569.

42 Aravin AA, Lagos-Quintana M, Yalcin A, *et al*. The small RNA profile during *Drosophila melanogaster* development. *Dev Cell* 2003; **5**:337-350.

43 Molnar A, Schwach F, Studholme DJ, Thuenemann EC, Baulcombe DC. miRNAs control gene expression in the single-cell alga *Chlamydomonas reinhardtii*. *Nature* 2007; **447**:1126-1129.

44 Morin RD, Aksay G, Dolgosheina E, *et al*. Comparative analysis of the small RNA transcriptomes of *Pinus contorta* and *Oryza sativa*. *Genome Res* 2008; **18**:571-584.

45 Tomari Y, Du T, Zamore PD. Sorting of *Drosophila* small silencing RNAs. *Cell* 2007; **130**:299-308.

46 Mi S, Cai T, Hu Y, *et al*. Sorting of small RNAs into *Arabidopsis argonaute* complexes is directed by the 5′ terminal nucleotide. *Cell* 2008; **133**:116-127.

47 Yigit E, Batista PJ, Bei Y, *et al*. Analysis of the *C. elegans*

Argonaute family reveals that distinct *Argonautes* act sequentially during RNAi. *Cell* 2006; **127**:747-757.

48 Aoki K, Moriguchi H, Yoshioka T, Okawa K, Tabara H. *In vitro* analyses of the production and activity of secondary small interfering RNAs in *C. elegans*. *EMBO J* 2007; **26**:5007-5019.

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

50 Sijen T, Steiner FA, Thijssen KL, Plasterk RH. Secondary siR-NAs result from unprimed RNA synthesis and form a distinct class. *Science* 2007; **315**:244-247.

51 Li J, Yang Z, Yu B, Liu J, Chen X. Methylation protects miR-NAs and siRNAs from a 3′-end uridylation activity in *Arabidopsis*. *Curr Biol* 2005; **15**:1501-1507.

52 Yu B, Yang Z, Li J, *et al.* Methylation as a crucial step in plant microRNA biogenesis. *Science* 2005; **307**:932-935.

53 Lau NC, Seto AG, Kim J, *et al.* Characterization of the piRNA complex from rat testes. *Science* 2006; **313**:363-367.

54 Saito K, Nishida KM, Mori T, *et al.* Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the *Drosophila* genome. *Genes Dev* 2006; **20**:2214-2222.

55 Vagin VV, Sigova A, Li C, *et al.* A distinct small RNA pathway silences selfish genetic elements in the germline. *Science* 2006; **313**:320-324.

56 Brennecke J, Aravin AA, Stark A, *et al.* Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. *Cell* 2007; **128**:1089-1103.

57 Bernstein E, Caudy AA, Hammond SM, Hannon GJ. Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* 2001; **409**:363-366.

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

59 Colmenares SU, Buker SM, Buhler M, Dlakic M, Moazed D. Coupling of double-stranded RNA synthesis and siRNA generation in fission yeast RNAi. *Mol cell* 2007; **27**:449-461.

60 Macrae IJ, Li F, Zhou K, Cande WZ, Doudna JA. Structure of Dicer and mechanistic implications for RNAi. *Cold Spring Harb Symp Quant Biol* 2006; **71**:73-80.

61 Hammond SM. Dicing and slicing: the core machinery of the RNA interference pathway. *FEBS lett* 2005; **579**:5822-5829.

62 Meins F, Jr., Si-Ammour A, Blevins T. RNA silencing systems and their relevance to plant development. *Annu Rev Cell Dev Biol* 2005; **21**:297-318.

63 Carmichael JB, Stoica C, Parker H, *et al.* RNA interference effector proteins localize to mobile cytoplasmic puncta in *Schizosaccharomyces pombe*. *Traffic* 2006; **7**:1032-1044.

64 Verdel A, Jia S, Gerber S, *et al.* RNAi-mediated targeting of heterochromatin by the RITS complex. *Science* 2004; **303**:672-676.

65 Hammond SM, Bernstein E, Beach D, Hannon GJ. An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. *Nature* 2000; **404**:293-296.

66 Carmell MA, Xuan Z, Zhang MQ, Hannon GJ. The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis. *Genes Dev* 2002; **16**:2733-2742.

67 Schwarz DS, Hutvagner G, Du T, *et al.* Asymmetry in the assembly of the RNAi enzyme complex. *Cell* 2003; **115**:199-208.

68 Lipardi C, Wei Q, Paterson BM. RNAi as random degradative PCR: siRNA primers convert mRNA into dsRNAs that are degraded to generate new siRNAs. *Cell* 2001; **107**:297-307.

69 Watanabe T, Totoki Y, Toyoda A, *et al.* Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. *Nature* 2008; **453**:539-543.

70 Makeyev EV, Bamford DH. Cellular RNA-dependent RNA polymerase involved in posttranscriptional gene silencing has two distinct activity modes. *Mol cell* 2002; **10**:1417-1427.

71 Wassenegger M, Krczal G. Nomenclature and functions of RNA-directed RNA polymerases. *Trends Plant Sci* 2006; **11**:142-151.

72 Diederichs S, Haber DA. Dual role for argonautes in microRNA processing and posttranscriptional regulation of microRNA expression. *Cell* 2007; **131**:1097-1108.

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

74 Provost P, Silverstein RA, Dishart D, *et al.* Dicer is required for chromosome segregation and gene silencing in fission yeast cells. *Proc Natl Acad Sci U S A* 2002; **99**:16648-16653.

75 Kato H, Goto DB, Martienssen RA, *et al.* RNA polymerase II is required for RNAi-dependent heterochromatin assembly. *Science* 2005; **309**:467-469.

76 Djupedal I, Portoso M, Spahr H, *et al.* RNA Pol II subunit Rpb7 promotes centromeric transcription and RNAi-directed chromatin silencing. *Genes Dev* 2005; **19**:2301-2306.

77 Chen ES, Zhang K, Nicolas E, *et al.* Cell cycle control of centromeric repeat transcription and heterochromatin assembly. *Nature* 2008; **451**:734-737.

78 Kloc A, Zaratiegui M, Nora E, Martienssen R. RNA interference guides histone modification during the S phase of chromosomal replication. *Curr Biol* 2008; **18**:490-495.

79 Hamilton AJ, Baulcombe DC. A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science* 1999; **286**:950-952.

80 Hammond SM, Boettcher S, Caudy AA, Kobayashi R, Hannon GJ. Argonaute2, a link between genetic and biochemical analyses of RNAi. *Science* 2001; **293**:1146-1150.

81 Hall IM, Shankaranarayana GD, Noma K, *et al.* Establishment and maintenance of a heterochromatin domain. *Science* 2002; **297**:2232-2237.

82 Sigova A, Rhind N, Zamore PD. A single Argonaute protein mediates both transcriptional and posttranscriptional silencing in *Schizosaccharomyces pombe*. *Genes Dev* 2004; **18**:2359-2367.

83 Hansen KR, Burns G, Mata J, *et al.* Global effects on gene expression in fission yeast by silencing and RNA interference machineries. *Mol Cell Biol* 2005; **25**:590-601.

84 Gobeil LA, Plante P, Rohani M, Ouellette M, Provost P. Involvement of Dcr1 in post-transcriptional regulation of gene expression in *Schizosaccharomyces pombe*. *Front Biosci* 2008; **13**:2203-2215.

85 Folco HD, Pidoux AL, Urano T, Allshire RC. Heterochromatin and RNAi are required to establish CENP-A chromatin at cen-

tromeres. *Science* 2008; **319**:94-97.

86 Cleveland DW, Mao Y, Sullivan KF. Centromeres and kineto-chores: from epigenetics to mitotic checkpoint signaling. *Cell* 2003; **112**:407-421.

87 Motamedi MR, Verdel A, Colmenares SU, *et al.* Two RNAi complexes, RITS and RDRC, physically interact and localize to noncoding centromeric RNAs. *Cell* 2004; **119**:789-802.

88 Debeauchamp JL, Moses A, Noffsinger VJ, *et al.* Chp1-Tas3 interaction is required to recruit RITS to fission yeast centromeres and for maintenance of centromeric heterochromatin. *Mol Cell Biol* 2008; **28**:2154-2166.

89 Zhang K, Mosch K, Fischle W, Grewal SI. Roles of the Clr4 methyltransferase complex in nucleation, spreading and maintenance of heterochromatin. *Nat Struct Mol Biol* 2008; **15**:381-388

90 Buhler M, Verdel A, Moazed D. Tethering RITS to a nascent transcript initiates RNAi- and heterochromatin-dependent gene silencing. *Cell* 2006; **125**:873-886.

91 Kennedy S, Wang D, Ruvkun G. A conserved siRNA-degrading RNase negatively regulates RNA interference in *C. elegans*. *Nature* 2004; **427**:645-649.

92 Buker SM, Iida T, Buhler M, *et al.* Two different Argonaute complexes are required for siRNA generation and heterochromatin assembly in fission yeast. *Nat Struct Mol Biol* 2007; **14**:200-207.

93 Till S, Lejeune E, Thermann R, *et al.* A conserved motif in Argonaute-interacting proteins mediates functional interactions through the Argonaute PIWI domain. *Nat Struct Mol Biol* 2007; **14**:897-903.

94 Petrie VJ, Wuitschick JD, Givens CD, Kosinski AM, Partridge JF. RNA interference (RNAi)-dependent and RNAi-independent association of the Chp1 chromodomain protein with distinct heterochromatic loci in fission yeast. *Mol Cell Biol* 2005; **25**:2331-2346.

95 Partridge JF, DeBeauchamp JL, Kosinski AM, *et al.* Functional separation of the requirements for establishment and maintenance of centromeric heterochromatin. *Mol cell* 2007; **26**:593-602.

96 Neumann P, Yan H, Jiang J. The centromeric retrotransposons of rice are transcribed and differentially processed by RNA interference. *Genetics* 2007; **176**:749-761.

97 Zilberman D, Cao X, Jacobsen SE. ARGONAUTE4 control of locus-specific siRNA accumulation and DNA and histone methylation. *Science* 2003; **299**:716-719.

98 Schauer SE, Jacobsen SE, Meinke DW, Ray A. DICER-LIKE1: blind men and elephants in *Arabidopsis* development. *Trends Plant Sci* 2002; **7**:487-491.

99 Matzke MA, Birchler JA. RNAi-mediated pathways in the nucleus. *Nat Rev* 2005; **6**:24-35.

100 Mette MF, Aufsatz W, van der Winden J, Matzke MA, Matzke AJ. Transcriptional silencing and promoter methylation triggered by double-stranded RNA. *EMBO J* 2000; **19**:5194-5201.

101 Jones L, Ratcliff F, Baulcombe DC. RNA-directed transcriptional gene silencing in plants can be inherited independently of the RNA trigger and requires Met1 for maintenance. *Curr Biol* 2001; **11**:747-757.

102 Cokus SJ, Feng S, Zhang X, *et al.* Shotgun bisulphite sequencing of the *Arabidopsis* genome reveals DNA methylation patterning. *Nature* 2008; **452**:215-219.

103 Lister R, O'Malley RC, Tonti-Filippini J, *et al.* Highly integrated single-base resolution maps of the epigenome in *Arabidopsis*. *Cell* 2008; **133**:523-536.

104 Aufsatz W, Mette MF, van der Winden J, Matzke M, Matzke AJ. HDA6, a putative histone deacetylase needed to enhance DNA methylation induced by double-stranded RNA. *EMBO J* 2002; **21**:832-6841.

105 Onodera Y, Haag JR, Ream T, *et al.* Plant nuclear RNA polymerase IV mediates siRNA and DNA methylation-dependent heterochromatin formation. *Cell* 2005; **120**:613-622.

106 Herr AJ, Jensen MB, Dalmay T, Baulcombe DC. RNA polymerase IV directs silencing of endogenous DNA. *Science* 2005; **308**:118-120.

107 Kanno T, Huettel B, Mette MF, *et al.* Atypical RNA polymerase subunits required for RNA-directed DNA methylation. *Nat Genet* 2005; **37**:761-765.

108 Pauler FM, Koerner MV, Barlow DP. Silencing by imprinted noncoding RNAs: is transcription the answer? *Trends Genet* 2007; **23**:284-292.

109 Birney E, Stamatoyannopoulos JA, Dutta A, *et al.* Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* 2007; **447**:799-816.

110 Katayama S, Tomaru Y, Kasukawa T, *et al.* Antisense transcription in the mammalian transcriptome. *Science* 2005; **309**:1564-1566.

111 Maison C, Bailly D, Peters AH, *et al.* Higher-order structure in pericentric heterochromatin involves a distinct pattern of histone modification and an RNA component. *Nat Genet* 2002; **30**:329-334.

112 Bernstein E, Kim SY, Carmell MA, *et al.* Dicer is essential for mouse development. *Nat Genet* 2003; **35**:215-217.

113 Kim DH, Villeneuve LM, Morris KV, Rossi JJ. Argonaute-1 directs siRNA-mediated transcriptional gene silencing in human cells. *Nat Struct Mol Biol* 2006; **13**:793-797.

114 Wang F, Koyama N, Nishida H, *et al.* The assembly and maintenance of heterochromatin initiated by transgene repeats are independent of the RNA interference pathway in mammalian cells. *Mol Cell Biol* 2006; **26**:4028-4040.

115 Morris KV, Chan SW, Jacobsen SE, Looney DJ. Small interfering RNA-induced transcriptional gene silencing in human cells. *Science* 2004; **305**:1289-1292.

116 Weinberg MS, Villeneuve LM, Ehsani A, *et al.* The antisense strand of small interfering RNAs directs histone methylation and transcriptional gene silencing in human cells. *RNA* 2006; **12**:256-262.

117 Han J, Kim D, Morris KV. Promoter-associated RNA is required for RNA-directed transcriptional gene silencing in human cells. *Proc Natl Acad Sci U S A* 2007; **104**:12422-12427.

118 Li LC, Okino ST, Zhao H, *et al.* Small dsRNAs induce transcriptional activation in human cells. *Proc Natl Acad Sci U S A* 2006; **103**:17337-17342.

119 Janowski BA, Younger ST, Hardy DB, *et al.* Activating gene expression in mammalian cells with promoter-targeted duplex RNAs. *Nat Chem Biol* 2007; **3**:166-173.

120 Yu W, Gius D, Onyango P, *et al.* Epigenetic silencing of tumour suppressor gene p15 by its antisense RNA. *Nature* 2008; **451**:202-206.

121 Zhang MX, Ou H, Shen YH, *et al.* Regulation of endothelial

nitric oxide synthase by small RNA. *Proc Natl Acad Sci U S A* 2005; **102**:16967-16972.

122 Mayer C, Schmitz KM, Li J, Grummt I, Santoro R. Intergenic transcripts regulate the epigenetic state of rRNA genes. *Mol cell* 2006; **22**:351-361.

123 Buhler M, Mohn F, Stalder L, Muhlemann O. Transcriptional silencing of nonsense codon-containing immunoglobulin minigenes. *Mol cell* 2005; **18**:307-317.

124 Benetti R, Gonzalo S, Jaco I, *et al*. A mammalian microRNA cluster controls DNA methylation and telomere recombination via Rbl2-dependent regulation of DNA methyltransferases. *Nat Struct Mol Biol* 2008; **15**:268-279.

125 Sinkkonen L, Hugenschmidt T, Berninger P, *et al*. MicroR-NAs control *de novo* DNA methylation through regulation of

transcriptional repressors in mouse embryonic stem cells. *Nat Struct Mol Biol* 2008; **15**:259-267.

126 Gottesman S. The small RNA regulators of *Escherichia coli*: roles and mechanisms*. *Annu Rev Microbiol* 2004; **58**:303-328.

127 Dennis PP, Omer A. Small non-coding RNAs in Archaea. *Curr Opin Microbiol* 2005; **8**:685-694.

128 Camblong J, Iglesias N, Fickentscher C, Dieppois G, Stutz F. Antisense RNA stabilization induces transcriptional gene silencing via histone deacetylation in *S. cerevisiae*. *Cell* 2007; **131**:706-717.

129 Buhler M, Haas W, Gygi SP, Moazed D. RNAi-dependent and -independent RNA turnover mechanisms contribute to hetero-chromatic gene silencing. *Cell* 2007; **129**:707-721.