The Epigenetics of Non-coding RNA

Lesley J. Collins1, Barbara Schönfeld2,3, and Xiaowei Sylvia Chen4
1Massey Genome Service, Massey University, Palmerston North, New Zealand
2Allan Wilson Centre for Molecular Ecology and Evolution, Massey University, Palmerston North, New Zealand
3Institute of Molecular BioSciences, Massey University, Palmerston North, New Zealand
4Department of Biochemistry, University of Otago, Dunedin 9054, New Zealand

INTRODUCTION

Non-protein-coding RNAs (ncRNAs) are RNAs that are transcribed from DNA but are not translated into proteins. Many are functional and are involved in the processing and regulation of other RNAs such as mRNA, tRNA, and rRNA. Processing-type ncRNAs include small nuclear RNAs (snRNAs) involved in splicing, small nucleolar RNAs (snoRNAs) that modify nucleotides in rRNAs and other RNAs, and RNase P that cleaves pre-tRNAs. Other small ncRNAs such as microRNAs (miRNAs) and short interfering RNAs (siRNAs) are involved in the regulation of target mRNAs and chromatin. Although many of these latter ncRNA classes are grouped under the term RNA interference (RNAi), it has become clear that there are many different ways that ncRNAs can interact with genes to up-regulate or down-regulate expression, to silence translation, or guide methylation [1–3]. Adding to these classes are long ncRNAs (typically >200 nt) that have also been implicated in gene regulation [4]. All of these ncRNAs form a network of processes, the RNA-infrastructure [2] that spans the cell not only spatially as RNAs move across the cell, but also temporally as the RNAs regulate gene processes during the cell cycle. Thus, the regulation of RNA processes may not only be transcriptional or translational, but also from their biogenesis and processing pathways [2]. However, when talking about gene regulation, it is RNAi that immediately comes to mind (especially in multicellular organisms) and it appears that RNAi-based ncRNAs and some longer ncRNAs have roles in epigenetic processes [5]. Some of these roles have been known for some time (e.g. X-chromosome inactivation [6] and gene imprinting [7]) but other roles in non-developmental mechanisms and cancer are only just coming to light.

We can cover only some of these mechanisms here but further reviews are available [5,7–10]. Although work in this area has clearly concentrated on mammalian examples there are many interesting mechanisms coming to light from non-mammalian species which we will cover to a small extent here. Presently we can divide the epigenetic-related classes of ncRNAs into two main groups; the long ncRNAs, and short ncRNAs including miRNAs, siRNAs, and Piwi-interacting RNAs (piRNAs). This chapter reviews both the long and short classes of ncRNAs involved in epigenetic regulation: those that generally act as cis-acting silencers, but also as trans-acting regulators of site specific modification and imprinted gene-silencing (Table 4.1). As the examples in the following sections will show, we are still very much in the early days...
of investigating how many characterized ncRNAs work to regulate processes such as RNA editing and methylation.

**SHORT ncRNAs AND EPIGENETICS**

RNAi is a mechanism by which short double-stranded RNAs (dsRNA) are used for sequence-specific regulation of gene expression, where some of the nucleotides on the ncRNA bind to either the coding or promoter region of an mRNA. This binding interferes with normal mRNA processing and consequently silences the expression of the mRNA. The three major classes are microRNA (miRNA), short interfering RNA (siRNA) and Piwi-interacting RNA (piRNA) which differ in their biogenesis and modes of target regulation ([11](#)) (Fig. 4.1).

Although best known for roles in regulating mRNA transcripts, these short ncRNAs are also directly involved in other cellular processes including chromatin-mediated gene silencing and DNA rearrangements ([2,12](#)). We will go through each class in turn highlighting how they are different, and review recent studies that indicate their use in epigenetics.

### TABLE 4.1 ncRNAs Discussed in This Chapter and Their Abbreviations

<table>
<thead>
<tr>
<th>ncRNA</th>
<th>Length</th>
<th>Short Description</th>
<th>Suggested Reviews and Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>miRNA</td>
<td>21–23 nt</td>
<td>ssRNA folds into dsRNA structure; after processing and binding to RISC complex they target mRNAs to regulate translation.</td>
<td>Reviews [12,13,15,25] Figure 4.1A Rt1 [16–19] Figure 4.2 miR-290 [20,21]</td>
</tr>
<tr>
<td>siRNA</td>
<td>20–25 nt</td>
<td>Regulate a specific gene using complementary sequence. Post-Transcriptional Gene Silencing (PTGS) and Transcriptional Gene Silencing (TGS) pathways. Plants also use RNA-directed DNA methylation (RdDM).</td>
<td>Reviews [9,25–28] Figure 4.1B FLC gene [9,31,32]</td>
</tr>
<tr>
<td>piRNA</td>
<td>27–30 nt</td>
<td>Interact with PIWI proteins for chromatin regulation and transposon silencing. Scan RNAs (scnRNAs) are a type of piRNA.</td>
<td>Reviews [33–36] Figure 4.1C</td>
</tr>
<tr>
<td>XiRNAs</td>
<td>24–42 nt</td>
<td>Produced from Xist and Tsix long ncRNAs, required for controlling methylation of the future inactive X chromosome and of the Xist promoter region on the future active X chromosome.</td>
<td>XiRNAs in XCI [45,51,52] Figure 4.3</td>
</tr>
<tr>
<td>Long ncRNAs</td>
<td>&gt;200 nt</td>
<td>Many have specific targets and are critical for X chromosome inactivation in mammals (XCI), meiotic sex chromosome inactivation (MSCI), RoX (RNA on X) system in insects, and Hox gene regulation.</td>
<td>XCI [40,41,45,49,52] Figure 4.3 MSCI [54,55] RoX [56–58] HOX [63,65] Figure 4.4</td>
</tr>
</tbody>
</table>
The Epigenetics of Non-coding RNA

miRNAs are perhaps the most well known of the regulatory ncRNA classes, and the general miRNA processing pathway is shown in Figure 4.1. Yet we cannot assume that all miRNAs within one species will regulate their genes in the same way in another species. Studies have shown that not only can a single miRNA down-regulate expression of hundreds of its target genes [13], but some miRNAs use alternative methods of down-regulation, such as accelerated deadenylation of the polyA tail [14]. Other studies revealed that animal miRNAs can induce translational up-regulation, and that some plant miRNAs can function as translational inhibitors contrary to their original functional descriptions [reviewed in Ref. 15]. However, miRNAs are not merely regulating mRNA targets, but are also involved in intricate mechanisms that involve feedback, self-regulation and in some cases methylation.

An example (Fig. 4.2) comes from the mouse Dlk1-Dio3 region in which three protein genes, i.e. delta-like 1 (Dlk1), retrotransposon gene (Rtl1), and Dio3, are expressed exclusively from the paternal chromosome [16]. On the maternal chromosome these protein-coding genes are normally repressed, and several other transcripts are produced including one antisense to the Rtl1 gene. Regional imprinting of Rtl1 is predetermined by the methylation status of the nearby intergenic differentially methylated region (IG-DMR), which is methylated in the paternal chromosome, but not in the maternal. The maternally inherited unmethylated
Additional Epigenetic Processes

**FIGURE 4.2**
miRNAs and imprinting.
Methylation of the IG-DMR region on the paternal chromosome represses the expression of the Rtl1a antisense transcript, allowing expression of the Rtl1 transcript. Without this methylation, as on the maternal strand, the Rtl1a is processed to produce miRNAs including miR-136 and miR-127, which complementarily bind to the Rtl1 transcript and induce degradation.

State of IG-DMR is essential for maintaining the repression of the protein-coding genes and for the expression of the antisense transcript [16,17]. The Rtl1as (antisense) transcript (also known as antiPeg11) forms hairpin structures from which after processing, miRNAs are released including miR-127 and miR-136 [18]. These miRNAs are located near 2 CpG islands in the Rtl1 transcript, and regulate the expression of Rtl1 in trans by guiding RISC-mediated cleavage of any maternal transcript. Aberrant epigenetic reprogramming of miR-127, miR-136, or Rtl1 result in late-fetal and/or neonatal lethality [19].

miRNAs have also been shown to be important in stem-cell self-renewal and differentiation (reviewed in Ref. 12). There are two types of stem cell, tissue stem cells (which include somatic and germline cells which develop, maintain, and repair tissues in developing and adult organisms), and embryonic stem cells (ES) which develop from an embryo to give rise to the fetus. Self-renewal (or self-replication) in tissue cells results in asymmetrical division, whereby one daughter cell retains the stem-cell properties, and the other daughter cell is committed to a differentiated function. This behavior is controlled inter-cellularly (between cells by cell signalling), as well as intra-cellularly through epigenetic, transcriptional, translational, and post-translational mechanisms. Recently, miRNAs have been found to be important players in controlling stem-cell fate and behavior. One example is the mouse miR-290–295 miRNA cluster, a group of miRNAs that share a 5’ proximal AAGUGC motif [20]. The expression of this cluster increases during pre-implantation development and remains high in undifferentiated ES cells, but then decreases after ES cell differentiation [20]. The miR-290 miRNAs act as post-transcriptional regulators of retinoblastoma-like 2 (Rbl2), which in turn acts as a transcriptional repressor of DNA methyl transferases (DNMTs), Dnmt3a and Dnmt3b. DNMTs epigenetically silence OCT4, a key transcription factor of ES cell renewal and differentiation [20,21]. Repression of Dnmt3a and Dnmt3b results in hypomethylation of the genome and especially the telomeres, leading to the appearance of long telomeres and increased telomere recombination. Alternatively, if Dicer is knocked out, miRNAs are depleted and the methylation of the Oct4 promoter is severely impaired during differentiation [20]. Many other candidate targets of the AAGUGC seed-containing miRNAs have been identified as well as many indirectly regulated targets [20], but it remains to be seen how other aspects of self-renewal and differentiation are affected by the miR-290 cluster.
This is only one example of many that show how miRNAs are directly or indirectly regulating key self-renewal or differentiating genes by either directly or indirectly affecting methylation processes. It is also possible for a miRNA to regulate another miRNA. An example of this action is miR-184, which negatively regulates miR-205 in human epithelial cells. Interfering with miR-205 dampens the Akt signaling pathway and is associated with a marked increase in keratinocyte apoptosis and cell death [22]. Current research (e.g. 22–24) is finding that more and more of such miRNAs are subjected to feedback from their target genes, and serve as a warning that what may appear at first to be “simple” regulation of an mRNA by an miRNA, may in fact have hidden features only revealed upon a detailed investigation of a mechanism.

As with miRNAs there are many subclasses of siRNAs that can be processed either as sense–antisense pairs (e.g. bidirectional promoter produced; Fig. 4.1 – siRNA pathway A), or as double-stranded transcripts which are subsequently cleaved by Dicer (Fig. 4.1 – siRNA pathway B) [25]. siRNA-based mechanisms have been well-studied in plants and fission yeast [9,26]. Although at first siRNAs in animals were mostly considered to be from foreign DNA or RNA (i.e. viral-induced exo-siRNAs), recent studies have characterized many more endogenously encoded siRNAs (endo-siRNAs) that appear to have a role in transposon control [reviewed in Ref. 27]. However, concentrating on the more well-known mechanisms in plants, one group of endo-siRNAs are the RDR2-dependent siRNAs which are preferentially associated with transposons, retroelements, and repetitive DNA, but also appear to guide methylation of specific DNA regions [26]. In plants, fission yeast, and to a small extent in mammals [28], both the transcriptional (TGS) as well as post-transcriptional gene silencing (PTGS) pathways are activated by dsRNAs. With the PTGS pathway, siRNAs direct mRNA degradation in the cytoplasm with no epigenetic incidence. However, TGS acts in the nucleus and is associated with chromatin modifications that silence transcription, and are maintained throughout the phases of the cell cycle [9].

In fission yeast more complicated models have been characterized. During TGS, the RITS (RNAi-Induced Transcriptional Silencing) complex is similar to RISC in containing Argonaute, but (unlike RISC) RITS localizes exclusively to the nucleus and contains at least one chromatin-binding module called a chromodomain [9]. Bound to a siRNA it mediates sequence specific heterochromatin formation and histone methylation. Recent models propose that RITS and RDRC (RNA-directed RNA polymerase Complex) are recruited to the site of intended heterochromatin formation when their associated siRNAs bind to a nascent RNA being transcribed at that site [29]. Thus, the binding of RITS to chromatin initiates heterochromatin formation which in turn results in TGS. Assembly of heterochromatin at a given genomic site comes with a heritable silencing of transcription. In fission yeast this mechanism is widely used to regulate heterochromatin formation, and a positive-feedback loop involving RDRP couples siRNA production to chromatin modifications [9].

In plants, siRNAs are involved in RNA-directed DNA Methylation (RdDM), which was first observed in viroid infected tobacco plants where sequences similar in sequence to the viral genes became methylated [30]. The exact mechanism for RdDM has not yet been characterized but in a general model [9], the plant specific RNA Polymerase IV is (somehow) recruited to a target genomic site; once there it synthesizes an ssRNA which RDR2 uses as a template to construct dsRNA that is processed by DCL3 (plant dicer) into siRNAs that bind AGO4 proteins. An AGO4 protein bound to an siRNA is thought to form a complex with PolVVb and DRM2 to guide DNA and chromatin methylation at the target genomic region [9,26]. One example is the FLC gene (Flowering locus gene C), a key MADS box transcription factor with key cell differentiation roles similar to that of HOX in animals. FLC gene expression is low during flowering in *Arabidopsis thaliana*, maintained by the Polycomb group of silencing proteins. A transposon in an intron of FLC is believed to nucleate formation of silent chromatin by attracting DNA and H3K9 methylation [9,31]. siRNAs complementary to
the 3’ end of the FLC gene have now been detected and their accumulation requires DCL2, RDR2, and PolIVa [32]. However, the siRNAs do not depend on the transposon but instead on antisense transcription of FLC 3’ UTR by a mechanism that is not yet clear [9].

The use of piRNAs (PIWI-interacting RNAs) in epigenetic processes is (like the siRNAs) only just coming under detailed investigation. Although also found in mammals and some ciliates piRNAs have been studied in greater detail in *Drosophila melanogaster*, in both germline and, more recently, soma cells [33,34], where they play critical roles in transposon “control” (i.e. preventing transposon activation and hence keeping the levels of transposons interrupting genes to a minimum) [35]. *Drosophila* piRNAs reside in clusters usually within heterochromatin or at heterochromatin–euchromatin boundaries. These piRNA clusters are repeat-rich regions composed of ancient fragmented transposon copies representing all major classes and element families [35]. Unlike miRNAs and siRNAs, piRNAs are not produced by “Dicing” (Fig. 4.1), but mainly by bi-directional promoters and what is known as the “ping pong” cycle of biogenesis and amplification. This cycle is initiated by primary piRNAs arising from piRNA clusters. Those piRNAs that are antisense to expressed transposons identify and cleave their targets, resulting in a set of new sense piRNAs in an AGO3 complex termed secondary piRNA. The AGO3-bound piRNA targets any transposon target that contains antisense transposon sequences. This cleavage then generates additional antisense piRNAs and the cycle can continue. This forms an effective small ncRNA-based transposon immune system.

piRNAs are now being proposed as possible vectors for carrying epigenetic inheritance [36]. An example comes from *Drosophila* strains that differ in the presence of a specific transposon, where crosses produce sterile progeny (hybrid dysgenesis), but only if the transposon is paternally inherited. Maternally inherited piRNAs are thought to play a role in this transposon silencing [36]. Both PIWI and Aubergine (Aub) proteins are deposited into developing oocytes and accumulate in the pole plasm suggesting a mechanism of transfer of maternal piRNAs into the germ lines of their progeny [36]. piRNA clusters alone have been shown to be insufficient to inactivate some transposons within a single generation. Instead maternally inherited siRNAs appear to prime the “resistance”-type control system at each generation to achieve full immunity. It is also thought that, since environment can influence the content of maternal small RNA populations, these RNAs could epigenetically alter the phenotype of progeny [35].

In mammals, transposon control by TGS occurs using PIWI-type proteins Milli and Miwi2 (Line-1 non-LTR), and IAP (LTR) retrotransposons, along with DNA methylation during embryogenesis in male germ cells (prospermatogonia) [35]. Like AGO3, Milli binds preferentially to piRNAs corresponding to transposon sense strands while Miwi2 contains mainly antisense piRNAs. piRNAs in prospermatogonia are derived from transposon rich piRNA clusters. There is evidence for a ping-pong amplification cycle as seen in *Drosophila*, but as yet its involvement in epigenetic inheritance is not characterized. Although the transmission of phenotype via piRNAs has only been demonstrated to date in *Drosophila*, the accumulation of small RNAs in the oocytes of other species is known and opens the way for this phenomenon to be more widespread [35].

**LONG ncRNAs AND EPIGENETICS**

During the last few years, evidence of complex, long ncRNA mediated epigenetic control systems has increased dramatically [3,37,38]. In a famous example, X chromosome inactivation (XCI) studied largely in mice, ensures only one of the two X chromosomes in XX females is expressed during development, and involves two long ncRNAs: Xist (17 kb) and its antisense transcript Tsix (40 kb) (reviewed in Refs 6,39,40). Xist RNA is expressed at a low level in both females and males before differentiation [41], but upon cell differentiation, Xist RNA coats the future inactive X chromosome (Xi) triggering extensive histone methylation
whereas Tsix appears to restrict Xist activity on the future active X chromosome (Xa) [43]. Recent studies, especially in mice, have revealed a more complex regulatory network of XCI which involves the interaction of long and short ncRNAs (Fig. 4.3).

To explain in more detail, in mice pre-XCI embryonic stem cells (ES) (Fig. 4.3A), Tsix is transcribed at a much higher level than Xist and triggers cytosine methylation within both Tsix and Xist genes, resulting in epigenetically equal competency for transcription and random X-inactivation [44]. The transcriptional level of Xist is elevated when the major pluripotency factors Nanog, Oct3/4, and Sox2 dissociate from intron 1 within Xist initiating XCI [44] (Fig. 4.3B and C). The coating of Xist on the future inactive X chromosome (Xi) (Fig. 4.3B) forms a silent chromatin compartment where X-linked genes become “localized” through Xist binding [45]. Xist RNA is required for chromosome-wide methylation in undifferentiated ES cells during the onset of X inactivation; however, once established, the maintenance of the heterochromatic state is independent of Xist RNA [46]. In contrast, the Polycomb repressive complex PCR2 is recruited by the RepA (a 1.6 kb ncRNA within Xist), and responsible for the maintenance of Xi [47,48]. On the future active X chromosome Xa (Fig. 4.3C), the level of Xist expression is largely controlled by its antisense transcript Tsix. The expression of Tsix is restricted to Xa [49] and associates with the DNA methyltransferase Dnmt3a to direct methylation on Xist promoter [50]. However, this methylation event is transient and does not play a role during the initiation of XCI [50]. Besides directing histone modification, the Tsix RNA can also down-regulate Xist expression through antisense binding. It has been shown that Tsix transcription across the Xist promoter is crucial for Xist regulation [44].
In addition small ncRNAs are also involved in XCI. Dicer-dependent XiRNAs are produced from both the Xist and Tsix ncRNA transcripts [44] and are required for methylation along the future inactive X chromosome, and methylation of the CpG island of the Xist promoter region in the future active X chromosome Xa [51]. Although XiRNAs are produced with Dicer, RNAi is not directly involved in X chromosome inactivation; instead it appears to maintain the steady-state level of the Xist RNA [52]. Adding to this increasingly complex network, RepA has been found to mediate the heterochromatic configuration of the Xist promoter through recruiting PRC2 [53].

Although most of the studies on XCI have been in placental mammals (and especially mice), recent work in marsupials has shown that a very similar mechanism exists although marsupials do not have the Xist RNA [54]. Here it is thought that male meiotic sex chromosome inactivation (MSCI) plays a greater role in dosage compensation. In mice, (reviewed in Ref. 55), MSCI silencing of the X chromosome genes is initiated during male meiosis, but unlike XCI, MSCI is transient, occurring during each round of spermatogenesis with some X-linked genes reactivating, then subsequently becoming silenced in the female. In marsupials, this can be demonstrated since XCI appears not to result from inheriting an X chromosome already inactivated by MSCI, but instead the inactivation takes place in the female (although the exact timing is not as yet known) [54]. It is also suggested [55] that some X-linked miRNAs escape MSCI and may contribute to the mechanisms regulating MSCI in an RNAi-like manner. Commonalities between the placental mammal and marsupial models (including enrichment of H3K27 trimethylation on the Xi and association of the Xi with the nucleolus [54]), indicate that aspects of the XCI system may be more conserved than originally thought [54].

In insects however, dosage compensation is achieved not by silencing but by a 2-fold increase of X-linked genes in males, relative to females [56–58]. In this mechanism the male-specific lethal (MSL) complex (consisting of MSL1, MSL2, MSL3, MOF (males absent on first), and MLE (maleless)) binds to genes along the male X chromosome. Associated with this complex are two long ncRNAs, roX1 and roX2 (RNA on X), that direct activation, rather than silencing, of their target genes [56,59]. roX1 and roX2 transcripts spread along the X chromosome recruiting the histone deacetylation protein complex, which generates an open chromatin conformation to facilitate active transcription [60,61]. How roX RNA regulates changes in the localization and activity of the MSL complex, is still poorly understood [62], and likewise how the MSL complex achieves dosage compensation [58]. Studies are beginning to indicate that target genes are enriched at the 3’ end and not at promoter sites leading to a model that the MSL complex affects elongation, resulting perhaps in hyper-transcription of the targeted genes or chromatin looping [58]. What is clear is that the RoX RNAs are a key part of the insect dosage compensation mechanism, and more study is needed to uncover the finer details [58].

Recent studies have also revealed long non-coding RNAs regulating the Hox gene cluster in insects and vertebrates (reviewed in Ref. 63). First found in Drosophila, the Hox family of proteins are critical determinants of correct patterning of the axis during embryonic development [64]. A large number of non-coding transcripts have been identified within the Hox gene cluster [63], the majority of which are found as antisense transcripts from intergenic regions, and are coordinately induced with their 3’-end Hox genes [65]. In Drosophila, the Bithorax Hox gene cluster (BX-C) regulation is extremely complex, with the regulatory region containing enhancers, silencers, maintenance elements, boundary elements, and possibly other elements not yet characterized [63].

Included in this regulation are the long ncRNAs, bxd RNAs, and iab RNAs [64], involved in regulation of their downstream Hox genes, Ultrabithorax (Ubx), Abdominal-A (abd-A), and Abdominal-B (abd-B) [63,66]. Bxd RNAs are expressed in different cells and germ layers, consistent with each bxd ncRNA having a unique role [63]. This spatial regulation may account for the observed mosaic expression pattern of the Hox genes in early embryos.
The Epigenetics of Non-coding RNA

In particular there has been some attention focused on one of these ncRNAs, HOTAIR, identified as regulating chromatin silencing of the adjacent Hox locus [65]. Figure 4.4 illustrates the mechanism by which the HOTAIR RNA regulates expression of Hox gene clusters through epigenetic control. HOTAIR, a 2158 nt spliced and polyadenylated long ncRNA is transcribed as a single copy on the antisense strand of the HOXC gene [65]. siRNA knockdown of HOTAIR results in transcriptional activation of the HOXD gene locus spanning four genes on a different chromosome. The HOTAIR RNA is physically associated with the Polycomb Repressive Complex 2 (PCR2), and is required for H3K27me3 modification and transcriptional silencing at HoxD [48,65]. Interestingly, HOTAIR transcription is linked to Polycomb group protein deposition and HOXD silencing on a different chromosome, demonstrating the action of an ncRNA in trans [48,65]. This action, but in cis, is seen in other long ncRNAs such as RepA, Kcnq1ot1, and AIR [48].

A number of models have been studied to reveal the roles of long ncRNAs in imprinted gene clusters. Short ncRNAs clearly have a role in epigenetic imprinting. In mouse, for example, 80 genes are grouped into clusters [7,67] and in many cases, one or more ncRNAs expressed from within a gene cluster play a crucial role in regulating the expression of the gene cluster [68]. This regulation directs chromatin modification forming an “epigenetic memory” within the same cell lineage [69]. Expression of genes in an imprinted cluster is generally controlled by a cis-regulatory region, the Imprint Control Element (ICE), which carries parental information in the form of DNA methylation [70]. Several gene clusters controlled by ICE are insulin-like growth factor 2 (Igf2), insulin-like growth factor 2 receptor (Igf2r), potassium voltage-gated channel (Kcnq1), and guanine nucleotide binding protein α stimulating factor (Gnas). Each of these clusters carries one ncRNA gene on the parental chromosome with unmethylated ICE [70]. The paternally imprinted Igf2 cluster contains a 2.5 kb spliced long ncRNA H19, which correlates with the methylation silencing of Igf2 genes [71], despite not having a direct role in maintaining silencing of the Igf2 cluster [72]. Although not necessary for the silencing of the Igf2 cluster, transgenic studies have revealed that H19 expression is sufficient for acquiring paternal-allele-specific methylation of the Igf2 genes [73]. It is possible that some of these long imprinting ncRNAs are in fact miRNA precursors [7], but as with many of the ncRNAs discussed here, details will emerge as further investigations proceed.

**CONCLUSIONS**

The investigation of ncRNA-related epigenetic mechanisms is at this point relatively new, but the rise of new sequencing technologies has already revealed epigenetic regulation at the
Additional Epigenetic Processes

Deep-sequencing technologies (also known as Next Generation Sequencing or NGS) have not only enabled analysis of histone modifications and methylation sites across entire genomes [74], but are enabling the detection of ncRNAs important in the regulation of these modifications. An example is Wang et al. (2009) [75] where maize organ specific distributions of canonical miRNAs and endogenous siRNAs have been linked to epigenetic modifications, H3K27me3, and DNA methylation [75]. This provides an important link between the epigenome and the transcriptome.

An interesting titbit is that the role of ncRNA in epigenetics has even been investigated in space. Spaceflight is a unique environment comprising of cosmic irradiation, microgravity, and space magnetic fields. A study of rice plants germinated from seed subjected to spaceflight showed altered methylation patterns and gene expression in six transposable elements and 11 cellular genes including siRNA related proteins Ago1 and Ago4 [76]. All of the detected alterations in the cellular genes were hypermethylation events occurring at CNG sites. This is consistent with the idea that plant CNG methylation is more prone to perturbation by environmental stresses [76,77].

While most studies have been conducted on major model organisms, there is now some information on ncRNA-based epigenetic mechanisms in protist lineages. piRNA-type RNAs (scan RNAs or scnRNAs) from ciliates are produced during the reorganization of the macronucleus during sexual development when some exons can become “scrambled” [35]. In Tetrahymena thermophila ~6000 IES internal eliminated sequences consisting of transposon-like and other repeats are targeted for removal by RNA-directed heterochromatin marking by scnRNAs [78]. Although the actual molecular mechanism is not as yet known, scnRNAs pair with either DNA or RNA from the parental macronucleus to be sorted, and then “selected” transcripts are moved to the newly developed macronucleus where they induce heterochromatin formation on the IES prior to elimination.

Recently, long ncRNAs have been found in the malaria parasite Plasmodium falciparum, where sterile sense and antisense RNAs are transcribed from the var virulence gene family and coat chromatin in a similar way to the animal Xist RNA and the Drosophila roX RNAs [79]. The regulation of the VSP genes involved in antigenic switching in the Diplomonad Giardia lamblia is also thought to be epigenetically regulated [80]. Subsequently there has been the identification of key RNAi proteins [81,82], some miRNAs [83,84], and a little on the regulatory mechanism of the VSP genes [85]; however, nothing is known as yet about chromatin modifications in Giardia. Further studies on protists are essential if we are to understand how ncRNAs in general regulate epigenetics and to understand how such mechanisms evolved.

Despite the individual variances in these pathways, miRNAs, siRNAs, and piRNAs all share several key protein components including Argonaute, PIWI, RDRP, and Dicer. Many RNA-directed epigenetic regulation events thus appear to be sharing protein and RNA components with the RNAi pathway if not dependent on the latter. We also note that evolution of ncRNAs by duplication could allow epigenetic states (e.g. methylation and imprinting) between the two copies to differ [1,86]. Since there are instances where a single trans-acting siRNA may have ~2300 predicted gene targets [7], this type of duplication could possibly result in a significant change in phenotype [1]. There is no doubt that the next few years will see more a greater understanding of ncRNA-related epigenetic mechanisms and perhaps then we can move on to constructive evolutionary analysis.

ACKNOWLEDGEMENTS

Many thanks go to David Penny, Peter Lockhart, and all, at the Allan Wilson Centre for Molecular Ecology and Evolution in Palmerston North, for support and encouragement. Funding for this project was from the Allan Wilson Centre, with funding also supplied to LJC by the Health Research Council of New Zealand.
References

Additional Epigenetic Processes


