NATURE|Vol 457|22 January 2009|doi:10.1038/nature07754

On the road to reading the RNA-interference code

Haruhiko Siomi[1] & Mikiko C. Siomi[1,2]

The finding that sequence-specific gene silencing occurs in response to the presence of double-stranded RNAs has had an enormous impact on biology, uncovering an unsuspected level of regulation of gene expression. This process, known as RNA interference (RNAi) or RNA silencing, involves small non-coding RNAs, which associate with nuclease-containing regulatory complexes and then pair with complementary messenger RNA targets, thereby preventing the expression of these mRNAs. Remarkable progress has been made towards understanding the underlying mechanisms of RNAi, raising the prospect of deciphering the 'RNAi code' that, like transcription factors, allows the fine-tuning and networking of complex suites of gene activity, thereby specifying cellular physiology and development.

The discovery of RNA interference (RNAi)[1] heralded a revolution in RNA biology. Researchers uncovered 'hidden' layers of regulation of gene expression, in which many previously unidentified families of small RNAs (consisting of ~20–30 nucleotides) mediate gene silencing. A diverse set of gene-regulatory mechanisms were found to use key steps in the RNAi process, including mechanisms that silence endogenous genes and mechanisms that restrain the expression of parasitic and pathogenic invaders such as transposons and viruses[2–5].

The basic RNAi process can be divided into three steps[6,7]. First, a long double-stranded RNA (dsRNA) that is expressed in, or introduced into, the cell (for example, as a result of the base-pairing of sense and antisense transcripts or the formation of stem–loop structures) is processed into small RNA duplexes by a ribonuclease III (RNaseIII) enzyme known as Dicer. Second, these duplexes are unwound, and one strand is preferentially loaded into a protein complex known as the RNA-induced silencing complex (RISC). Third, this complex effectively searches the transcriptome and finds potential target RNAs. The loaded single-stranded RNA (ssRNA), called the guide strand, then directs an endonuclease that is present in the RISC (sometimes called the 'slicer' and now known to be an Argonaute protein[8–11]) to cleave messenger RNAs that contain sequence homologous to the ssRNA, over many rounds. In this way, the guide strand determines the sequence specificity of the RNAi response.

In different organisms, the RNAi pathways comprise different proteins and mechanisms, but they operate by strikingly convergent strategies. In all organisms that have been studied, RNAi involves two main components: small RNAs, which determine the specificity of the response; and Argonaute proteins, which carry out the repression. Depending on both the nature of the Argonaute in the RISC and the degree of complementarity between the small RNA and the target sequence in the mRNA, the association of the RISC with target mRNAs has been shown to have different outcomes: it can control protein synthesis and mRNA stability, maintain genome integrity or produce a specific set of small RNAs[8,12]. Analyses of the biogenesis of small RNAs and their targeting mechanisms have benefited from the advent of high-throughput sequencing technologies and sophisticated bioinformatics[13]. The picture emerging from these studies is that RNAi systems in different organisms have been refined in many ways, and such modifications include built-in molecular 'rulers' that define the size of small RNAs, structures that determine which strand of a small RNA is selected, mechanisms that direct further rounds of small RNA amplification, or safeguards against off-target (unrestricted and unrelated) silencing.

Another emerging finding in the field is that the activity of RNAi pathways is subject to intense regulation at various levels, from the level of biogenesis of small RNAs to the silencing mode of the RISC. In this Review, we describe the biogenesis of the guide strand of small RNAs and the formation and actions of the RISC, and we discuss the current understanding of the molecular mechanisms of RNAi in the light of recent insights into how silencing pathways are specified and regulated.

Biogenesis of small RNAs

A hallmark of RNAi is that short (~20–30 nucleotide) dsRNAs — known as small RNAs — are generated by the activity of RNaseIII enzymes (either Dicer alone or Drosha and Dicer). Two main categories of small RNAs have been defined on the basis of their precursors. The cleavage of exogenous long dsRNA precursors in response to viral infection or after artificial introduction generates short interfering RNAs (siRNAs), whereas the processing of genome-encoded stem–loop structures generates microRNAs (miRNAs). Using high-throughput sequencing technology, several new classes of endogenous small RNA species have recently been uncovered, and these include PIWI-interacting RNAs (piRNAs) and endogenous siRNAs (endo-siRNAs or esiRNAs).

A common feature of all of these small RNAs is that they are loaded onto Argonaute proteins to effect their targeting function (discussed further in the section 'Loading and sorting of small RNAs by the RISC'). An overview of the generation of small RNAs is presented in Fig. 1.

siRNA biogenesis

Dicer (Table 1) processes long RNA duplexes and generates siRNAs. These small RNAs are ~21–25-nucleotide duplexes with a phosphate group at both 5′ ends, and hydroxyl groups and two-nucleotide overhangs at both 3′ ends, all hallmarks of RNaseIII-mediated cleavage. The Dicer protein contains a PAZ domain, which binds to the 3′ end of an siRNA, and two RNaseIII domains, which have the catalytic activity. It functions as a monomer[14], but the RNaseIII domains associate with each other to

[1]Department of Molecular Biology, Keio University School of Medicine, 35 Shinanomachi, Shinjuku-ku, Tokyo 160-8582, Japan. [2]Japan Science and Technology Agency (JST), Core Research for Evolutional Science and Technology (CREST), 4-1-8 Hon-chou, Kawaguchi, Saitama 332-0012, Japan.

form an 'internal dimer' (see page 405). The distance between the PAZ domain and the two RNaseIII domains is the length spanned by 25 base pairs (bp) of RNA[15]. Thus, Dicer itself is a molecular ruler.

miRNA biogenesis

Similarly, miRNAs are short (~21–25-nucleotide) RNA molecules[16]; however, their biogenesis differs markedly from that of siRNAs. The primary precursors of miRNAs (pri-miRNAs) are encoded in the genome, and the relevant genomic regions are mostly transcribed by RNA polymerase II (ref. 17). The pri-miRNAs contain stem–loop structures that harbour the miRNA in the 5′ or 3′ half of the stem. During miRNA production in plants, one type of RNaseIII, Dicer-like protein 1 (DCL1), generates the miRNA–miRNA* duplex in the nucleus (miRNA* being the sequence in the stem–loop that pairs with the miRNA, equivalent to the passenger strand of siRNA duplexes; discussed later). By contrast, in animals, miRNAs are derived in a two-step process, in which the nuclear-localized RNaseIII Drosha defines one end of the miRNA–miRNA* duplex and releases a precursor miRNA (pre-miRNA) of ~65–70 nucleotides. The pre-miRNA hairpin is then exported to the cytoplasm, where Dicer completes the processing.

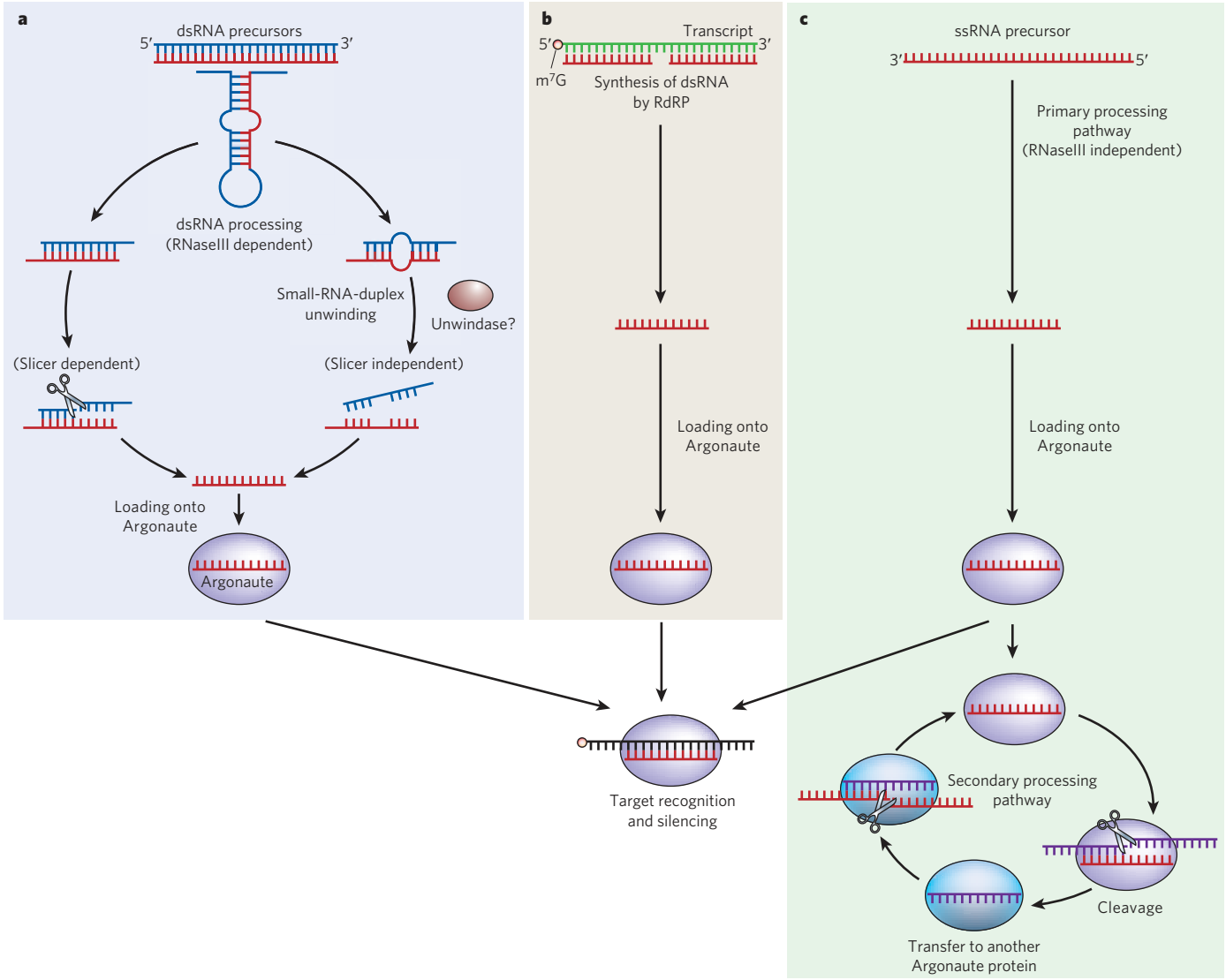

Figure 1 | Small RNA production and RNA silencing. a, Natural transcripts that form dsRNAs and hairpin-shaped structures can be sources of small RNAs. These precursors are processed by an RNaseIII enzyme (such as Drosha or Dicer), yielding small RNA duplexes. Duplexes with a perfect match (left pathway) are further processed by an enzyme with slicer activity (an Argonaute protein) into single-stranded small RNAs. By contrast, small RNA duplexes with a mismatch or bulge in the centre (right pathway) are not substrates for the slicer and thus become single-stranded in a cleavage-independent manner. The identity of the protein that carries out this unwinding is unknown. Single-stranded small RNAs are then loaded onto Argonaute proteins. The particular strand that is selected (sense or antisense) depends on thermodynamic stability. The loaded Argonaute proteins are guided to target mRNAs containing complementary sequence, and the expression of the corresponding genes is silenced. The mode of action of this silencing — whether the mRNA is cleaved or whether translation is just repressed — largely depends on the degree of complementarity between the target mRNAs and the Argonaute-associated small RNAs. **b**, Some small RNAs found in *Caenorhabditis elegans* and plants are known to be produced in an RNA-dependent RNA polymerase (RdRP)-dependent manner. Natural transcripts (often aberrant RNAs) can be substrates for this type of small RNA synthesis. This does not occur in organisms that lack RdRP activity, such as mammals and *Drosophila melanogaster*. Single-stranded small RNAs generated in this way can then be loaded onto Argonaute proteins and silence gene expression. **c**, The PIWI subfamily of Argonaute proteins, which are germline specific, are loaded with piRNAs. These complexes function to silence transposons. Single-stranded precursors give rise to piRNAs, through a mechanism called the primary processing pathway. The proteins required for this pathway are unknown. The silencing of transposons by PIWI proteins simultaneously amplifies piRNAs in germ cells. This pathway is known as the secondary processing pathway (or the ping-pong amplification loop) and is conserved in a variety of organisms, including mice and zebrafish. In this pathway, the slicer activity of the PIWI proteins reciprocally forms the 5′ ends of piRNAs by cleaving transposon transcripts (piRNA precursors). Proteins required to form the 3′ end of piRNAs remain unidentified.

Table 1 | Key proteins in RNA silencing in various organisms

Protein	Yeast (*Schizosaccharomyces pombe*)	Plant (*Arabidopsis thaliana*)	Nematode (*Caenorhabditis elegans*)	Fruitfly (*Drosophila melanogaster*)	Mammal	
					Mouse	Human
RNaseIII	Dcr1	DCL1	DCR-1	DCR-1	DICER1	DICER1
		DCL2	DRSH-1	DCR-2	DROSHA	DROSHA
		DCL3		DROSHA		
		DCL4				
Argonaute: AGO subfamily	Ago1	AGO1	ALG-1	AGO1	AGO1	AGO1
		AGO2	ALG-2	AGO2	AGO2	AGO2
		AGO4	3 others		AGO3	AGO3
		AGO5			AGO4	AGO4
		AGO6			AGO5 (possibly a pseudogene)	
		AGO7 (ZIPPY)				
		AGO10 (ZLL, PNH)				
		3 others				
Argonaute: PIWI subfamily	None	None	ERGO-1	AGO3	MILI (PIWIL2)	HILI (PIWIL2)
			PRG-1	PIWI	MIWI (PIWIL1)	HIWI (PIWIL1)
			PRG-2	AUB	MIWI2 (PIWIL4)	HIWI2 (PIWIL4)
						PIWIL3 (HIWI3)
Argonaute: WAGO subfamily	None	None	RDE-1	None	None	None
			SAGO-1			
			SAGO-2			
			PPW-1			
			PPW-2			
			CSR-1			
			NRDE-3			
			11 others			
Double-stranded-RNA-binding domain (dsRBD)-containing cofactor of RNaseIII	None	HYL1	PASH-1	PASHA	DGCR8	DGCR8
			RDE-4	R2D2	TRBP (TARBP2)	TRBP (TARBP2)
				LOQS	PACT (PRKRA)	PACT (PRKRA)
RNA-dependent RNA polymerase (RdRP)	Rdp1	RDR1	EGO-1	None	None	None
		RDR2 (SMD1)	RRF-1			
		RDR6 (SDE1, SGS2)	RRF-3			
		3 others	1 other			

Molecules that belong to these categories but have unknown functions are not listed but are indicated as 'others'. Common synonyms are indicated in parentheses. Data were taken from refs 8, 12, 29, 50 and 98.

Drosha is present in a large complex, known as the microprocessor complex, which functions like a molecular ruler to determine the cleavage site in the pri-miRNAs[18,19]. In this complex, Drosha interacts with its cofactor, known as DGCR8 or Pasha (depending on the species), which also binds to dsRNA (through its dsRNA-binding domain; dsRBD)[20,21]. A typical metazoan pri-miRNA consists of a 33-bp stem, a terminal loop and ssRNA flanking segments. The flanking segments are crucial for binding to DGCR8, and the 33-bp stem is also required for efficient binding. Drosha can interact transiently with the stem of this 'pre-cleavage' complex, and the processing centre of the enzyme, located at ~11 bp from the ssRNA–dsRNA junction, makes a staggered pair of breaks in the RNA to create the ~65–70-nucleotide pre-miRNA. Thus, DGCR8 might function as the molecular anchor that measures the distance from the ssRNA–dsRNA junction. It is possible that the microprocessor complex could recognize the terminal loop as ssRNA and bind to the stem–loop structure in the opposite orientation. In this case, abortive cleavage can occur at an alternative site ~11 bp from the terminal loop. However, most pri-miRNAs contain internal bulges or weakly paired bases ~11 bp from the terminal loop that mitigate processing from this direction[18].

Although many of the sequences encoding miRNAs are located within introns, clusters encoding miRNAs that are processed directly by the spliceosome, instead of Drosha, were recently identified[22,23]. The 3′ end of the stem–loop precursor of these intronic miRNAs (known as mirtrons) coincides with the 3′ splice site of a small annotated intron and is cleaved in the same splicing pathway as pre-mRNA in the nucleus instead of by Drosha. Subsequently, the mirtron precursors, which are released by the spliceosome in the shape of a lariat (lasso), are linearized by a de-branching enzyme. They then enter the miRNA-processing pathway directly (by mimicking the structural features of pre-miRNA hairpins) and are therefore exported to the cytoplasm and processed by a Dicer protein, bypassing Drosha-mediated cleavage.

The imprecision of Drosha or Dicer cleavage could result in the production of a set of miRNA–miRNA* duplexes with a variety of 5′ and 3′ ends. Most miRNAs in animals form imperfect hybrids with sequences in the target mRNA, with most of the pairing specificity being provided by the 5′-proximal region of the miRNA (that is, positions 2–8; also known as the seed region)[24,25]. Imprecise cleavage either alters the seed sequence or inverts the relative stabilities of the 5′ and 3′ ends of the duplex (see the section 'Loading and sorting of small RNAs by the RISC'). The results of recent deep-sequencing studies of small RNAs, however, indicate that human cells might take advantage of such imprecise cleavage, because the generation of a diverse set of miRNAs from a single precursor could be a way of broadening the network of factors and processes that are regulated by miRNAs[26–28].

RNaseIII-independent pathways of small RNA biogenesis

In some systems, small RNAs do not seem to be produced in response to dsRNA, but silencing signals are still amplified. Because these small RNAs do not arise from dsRNA precursors, RNaseIII enzymes cannot be involved in their generation. These findings therefore call into question the definition of RNAi. In this subsection, we describe the known RNaseIII-independent pathways of small RNA production, including those that generate piRNAs, 21U-RNAs, and secondary siRNAs in *Caenorhabditis elegans*.

The small RNAs known as piRNAs were named for their ability to bind to a group of Argonaute proteins known as PIWI proteins. As noted earlier, members of the Argonaute family bind directly to small guide RNAs and lie at the core of all known RISCs[8]. Argonaute proteins consist of a variable amino-terminal domain and three conserved domains (the PAZ, middle (MID) and PIWI domains)[8,29,30]. The 3′ end of a small RNA interacts with the PAZ domain, whereas the phosphate group at the 5′ end of small RNAs binds to a cleft bridging the MID domain and the PIWI domain[29,30] (see page 405). The PIWI domain has an RNaseH-like folded structure[10] and slicer activity (although some Argonaute proteins seem to have no slicer activity). There are three phylogenetic groups of Argonaute proteins[29]: the AGO subfamily (or AGO clade), named after the founding member *Arabidopsis thaliana* ARGONAUTE 1 (AGO1); the PIWI subfamily, named after *D. melanogaster* PIWI (*P*-element-induced wimpy testis); and the WAGO (worm-specific Argonaute) subfamily of *C. elegans*-specific proteins. PIWI-subfamily proteins bind to piRNAs[31–37] (Table 1). These small RNAs have been found only in germ cells, and they are important for germline development and suppress transposon activity in the germline cells of mammals, fish and *D. melanogaster*. They are ~24–31 nucleotides (slightly longer than miRNAs), usually have a uridine at the 5′ end and carry a 5′ monophosphate. Unlike mammalian miRNAs, but similarly to plant miRNAs, piRNAs have a 2′-*O*-methyl (2′-*O*-Me) modification on the nucleotide at the 3′ end, a modification that is carried out by a HEN1-like methyltransferase[38–42]. If Dicer is mutated, the production of piRNAs is not affected, indicating that their biogenesis is distinct from that of miRNAs and siRNAs and does not involve dsRNA precursors[31,42].

The sequencing of small RNAs associated with *D. melanogaster* PIWI-subfamily proteins (PIWI, Aubergine (AUB) and AGO3)[43,44] showed that piRNAs associated with AUB and PIWI are derived mainly from the antisense strand of retrotransposons, whereas AGO3-associated piRNAs arise mainly from the sense strand. AUB- and PIWI-associated piRNAs show a strong preference for uridine at their 5′ ends, whereas AGO3-associated piRNAs show a preference for adenosine at nucleotide 10. The first ten nucleotides of AUB-associated piRNAs can be complementary to the first ten nucleotides of AGO3-associated piRNAs. In addition, PIWI-subfamily proteins have slicer activity that allows them to cleave an RNA substrate opposite position 10 of their bound piRNA[32,44]. These observations suggest that piRNAs have a self-amplifying loop (Fig. 1), in which sense piRNAs associated with AGO3 cleave long antisense transcripts and guide the formation of the 5′ end of antisense piRNAs bound to AUB or PIWI, and vice versa. Thus, in this amplification loop, which is called the ping-pong cycle[43], transposons are both a source of piRNAs and a target of piRNA-mediated silencing. After the resultant cleavage products have been loaded onto another member of the PIWI subfamily, further (as yet unidentified) nuclease activity generates the 3′ end of the piRNA, with the specific size of the piRNA determined by the footprint of the PIWI-subfamily protein on the RNA, a step that seems to precede 2′-*O*-Me modification[38]. In each PIWI-subfamily protein, the PAZ domain might be positioned at a distance from the MID domain that corresponds to the length of each piRNA. Thus, the PAZ domain might function as part of a molecular ruler for processing piRNAs of a defined size. Signatures of this amplification cycle are also apparent in zebrafish (*Danio rerio*) germ cells and in mammalian germ cells before the pachytene stage of meiosis during spermatogenesis[42,45].

PIWI-subfamily proteins and, presumably, their associated piRNAs are loaded into embryos from the ova[8], implying that the piRNAs that initiate an amplification cycle of piRNA biogenesis (which generates secondary piRNAs) could be supplied by germline transmission. But several findings indicate that there must be mechanisms of piRNA biogenesis other than amplification induced by maternal piRNAs. First, the amplification cycle in *D. melanogaster* engages mainly AGO3 and AUB[43,44], but piRNAs are still loaded onto PIWI, which is spatially separated from these proteins at the subcellular and cell-type levels[32,43,44]. Second, piRNAs derived from a particular piRNA cluster in the genome (the *flamenco* locus) associate almost exclusively with PIWI[43]. These findings indicate that *flamenco*-derived piRNAs are produced by a pathway independent of the amplification loop. Whether such a piRNA-biogenesis pathway exists remains to be determined.

What at first seemed to be another type of small RNA, 21U-RNA, is found in *C. elegans*. These small RNAs are precisely 21 nucleotides and have a bias towards uridine at the 5′ end (but not in the remaining 20 nucleotides), and the genetic regions that encode them contain a characteristic sequence motif ~42 bp upstream of the first nucleotide of the small RNA[46]. It is possible that these RNAs are derived from thousands of separate, autonomously expressed, loci that are broadly scattered in two large regions of one chromosome. They are expressed solely in the germ line and interact with the PIWI-subfamily protein PRG-1 (refs 47, 48); therefore, 21U-RNAs are the *C. elegans* equivalent of piRNAs by definition. Like piRNAs, they depend on PRG-1 activity for their accumulation and are independent of DCR-1 (the *C. elegans* Dicer protein) for their production. *C. elegans* with mutations in *prg-1* have a smaller brood and a temperature-sensitive sterile phenotype, which is consistent with the idea that PIWI-subfamily proteins are involved in germline maintenance. Like the piRNAs found in mammalian germ cells in pachytene[33,34], 21U-RNAs have remarkable sequence diversity but lack obvious targets.

Small RNAs with a similar role to piRNAs have also been found in the ciliate *Tetrahymena thermophila*. These scan RNAs (scnRNAs) direct the elimination of transposon-like DNA sequences and associate with a PIWI-subfamily protein, TWI1 (ref. 8) but, in contrast to piRNAs and 21U-RNAs, are produced by a Dicer-dependent pathway[49].

These three examples (piRNAs, 21U-RNAs and scnRNAs) indicate that the core PIWI and piRNA machinery might have evolved to produce small RNAs and silence targets by different strategies.

RNA silencing pathways include mechanisms that downregulate endogenous genes and restrain the expression of selfish or exogenous genetic material, and these pathways often share common components such as Dicer. Therefore, there should be competition between different silencing pathways for particular components. Ways to overcome such competition should also exist; for example, by amplifying a weak silencing signal. In *C. elegans*, distinct Argonaute proteins operate at different stages of RNAi, directing gene silencing in a sequential manner[50] — the second stage of which involves RNaseIII-independent biogenesis of small RNAs. First, a primary Argonaute protein (such as RDE-1 for exogenous siRNAs (exo-siRNAs) and ERGO-1 for endo-siRNAs) is guided by 'primary' siRNAs (that is, a first round of siRNAs), which have been generated from long dsRNAs by DCR-1. Second, the silencing signal is amplified by the production of 'secondary' siRNAs by the action of RNA-dependent RNA polymerases (RdRPs) (Fig. 1). These secondary siRNAs then bind differentially to secondary Argonaute proteins (SAGOs, members of the WAGO subfamily), which mediate downstream silencing. In plants, RNAs with aberrant features, including lack of a poly(A) tail and lack of a 5′ cap, are copied into double-stranded forms by RdRPs and become substrates for Dicer, which converts them into siRNA duplexes[12]. By contrast, the *C. elegans* somatic RdRP mostly produces 21-nucleotide, single-stranded, 5′-triphosphorylated small RNAs directly from the target mRNA in a primer-independent manner without the need for Dicer-mediated cleavage of dsRNA[51–53]. Such recruitment of an RdRP directly to the target mRNA allows dsRNA synthesis without consuming the siRNAs generated in response to the original trigger, although it is unclear how the 3′ end of these secondary siRNAs is formed and what the molecular ruler is that determines their size.

Blurring of the boundaries between small RNA types

As described above, the three main classes of small RNA — siRNAs, miRNAs and piRNAs — are distinct in their biogenesis and cellular roles. However, recent findings blur these distinctions and show that there are even more-complex interactions between factors involved in small RNA biogenesis. Deep sequencing of small RNAs from somatic tissues and cultured somatic cells in *D. melanogaster* has uncovered another class of small RNA, consisting of 3′-methylated, 21-nucleotide RNAs derived from the *D. melanogaster* genome. These endogenous RNAs are derived from transposons and from several loci, including

loci that encode *cis*-natural antisense transcript pairs, and long stem–loop structures containing many mismatched pairs in their stems[54–57]. In *D. melanogaster*, distinct Dicer-containing complexes produce exo-siRNAs and miRNAs[58,59]. DCR-1 generates miRNAs, acting with its dsRNA-binding protein partner, Loquacious (LOQS)[60,61], and the miRNAs are loaded onto AGO1. By contrast, DCR-2, together with its dsRNA-binding protein partner, R2D2 (ref. 62), generates exo-siRNAs, which are loaded onto AGO2. Like exo-siRNAs, the recently discovered endogenous small RNAs are produced by the DCR-2-dependent pathway and are loaded onto AGO2, and they are therefore called endo-siRNAs. However, the generation of many endo-siRNAs requires LOQS[54,56], the dsRBD-containing partner of DCR-1 in the miRNA pathway[60,61], but not R2D2, the partner of DCR-2 (ref. 62). In *D. melanogaster* deficient in DCR-2 or AGO2, the expression of transposons increases, so endo-siRNAs might be the main mechanism for silencing 'selfish' genetic elements in somatic cells, which lack the piRNA pathway. Therefore, endo-siRNAs and piRNAs are fundamentally similar in that they defend organisms against nucleic-acid-based 'parasites'. This finding also shows that *D. melanogaster* has two RNAi pathways that repress transposon expression. Mouse oocytes have also been shown to contain endo-siRNAs. These RNAs are derived from various sources, including transposons[63,64]; however, some are processed from overlapping regions of functional genes and their cognate pseudogenes. This finding suggests that pseudogenes, which have been thought to be non-functional protein 'fossils', might regulate the expression of their founder genes.

Although siRNAs and miRNAs are categorized in terms of their origin rather than their size or function[7,12], the discovery of endo-siRNAs makes it difficult to distinguish between siRNAs and miRNAs. This blurring of the boundaries between the different types of small RNA has interesting evolutionary implications. The long stem–loop structures that are processed to form endo-siRNAs are reminiscent of the pre-miRNAs in plants. One hypothesis for the evolutionary origin of plant miRNAs is that new plant miRNA loci might evolve from the inverted duplication of founder loci, which when transcribed would result in hairpin RNAs[12]. These hairpin RNAs would have almost perfect self-complementarity and might be processed by Dicer-like enzymes other than DCL1, the main miRNA-processing enzyme in plants, because DCL1 has limited activity against such substrates. Subsequent acquisition of mutations as a result of genetic drift would produce a hairpin with imperfect complementarity, which could then be processed by DCL1. Thus, the stem–loop structures from which endo-siRNAs are derived could be evolutionary intermediates that are gradually transformed into miRNA precursors. It is possible that such an adaptive switch could also occur during the evolution of miRNA-encoding genes in *D. melanogaster*, in which DCR-1 would then generate miRNAs instead of endo-siRNAs being generated by DCR-2.

Loading and sorting of small RNAs by the RISC

In gene silencing pathways initiated by dsRNA precursors, Dicer-mediated cleavage yields small dsRNA intermediates (small RNA duplexes). These small RNA duplexes must be dissociated into 'competent' single strands in order to function as guides for RISCs. For each small RNA duplex, only one strand, the guide strand, is loaded onto a specific Argonaute protein and assembled into the active RISC; the other strand, the passenger strand, is destroyed. Many eukaryotes express more than one Argonaute protein, and these proteins bind to small RNAs in a sequence-independent manner. So how are small RNAs sorted and loaded onto a specific Argonaute protein?

Loading

A small RNA generated from dsRNA precursors is converted from a duplex into a single-stranded form as it is loaded into the RISC. The key steps in converting the RISC from its precursor form (the pre-RISC), which contains the small RNA duplex, to its mature form (the holo-RISC), which contains the guide strand, are small RNA strand unwinding and preferential strand selection. The prevalent view of RISC loading is that thermodynamic asymmetry along small RNA duplex determines which RNA strand is retained and which is discarded. More specifically, the strand that has its 5′ end at the thermodynamically less stable end of the small RNA duplex is preferentially loaded into the RISC as the guide strand, a phenomenon referred to as the asymmetry rule[65,66].

For siRNAs, the known interactions between Dicer and the Argonaute proteins[8] indicate that the production of the small RNA and the assembly of the RISC might be physically coupled. For example, in *D. melanogaster*, DCR-2 does not simply transfer siRNAs to a distinct RISC but, instead, forms part of the RISC together with the siRNAs, indicating that the role of DCR-2 extends beyond the initiation phase. The loading of siRNA duplexes onto AGO2 is facilitated by the RISC-loading complex, which contains DCR-2 and its dsRBD-containing partner, R2D2 (refs 62, 67). The particular strand of the siRNA duplex that is loaded onto AGO2 seems to be determined by the orientation of the DCR-2–R2D2 heterodimer on the siRNA duplex[68]. R2D2 is thought to sense the thermodynamic stability of the siRNA duplexes and bind to the more stable end of the siRNA, whereas DCR-2 is recruited to the less stable end. The heterodimer probably recruits AGO2 through an interaction between DCR-2 and AGO2. Previous models have proposed that the transition from a double-stranded silencing trigger to a single-stranded one is mediated by an unidentified ATP-dependent RNA helicase. However, the unwinding of the siRNA duplex and the loading of a single strand into the RISC are facilitated by the slicing of the unincorporated (passenger) strand by AGO2, a process that does not require ATP[69–71] (Fig. 1). Cleavage in the middle of the passenger strand, as though the passenger strand were an mRNA target, would be expected to reduce the annealing temperature and the free energy of duplex formation, which in turn facilitates the separation of the siRNA strands. These data support a model in which siRNAs are initially loaded as duplexes onto an AGO2-containing pre-RISC (Fig. 2).

By contrast, in humans, pre-miRNAs are known to bind to a preformed trimeric complex of AGO2, DICER1 and DICER1's dsRBD-containing partner, TRBP[72]. This complex can cleave target RNAs using pre-miRNA and can distinguish miRNA from miRNA*, in the absence of ATP hydrolysis[72,73], suggesting that DICER1-mediated cleavage and sensing of thermodynamic stability occur in series in the AGO2–DICER1–TRBP complex.

This process by which a pre-RISC is converted to a holo-RISC can also occur by a slicer-independent mechanism. Three of the four Argonaute proteins in humans (AGO1, AGO3 and AGO4) lack slicer activity but are nonetheless loaded with single-stranded guide siRNAs[9,11,28]. Similarly, single-stranded miRNAs are found associated with AGO2 in humans, despite the expectation that mismatches in the unwound pre-miRNA should block the passenger-strand cleavage activity of AGO2. Thus, a cleavage-independent (bypass) mechanism for RISC assembly must exist. RNA helicase A has been identified as a candidate for unwinding the duplex in this process[74].

Sorting

Once assembled, RISCs mediate a range of the effector steps in all RNA silencing mechanisms, from repressing translation to maintaining genome stability. The specialized functions of RISCs are likely to result from the particular proteins that associate with each Argonaute protein. In other words, the different RISC variants are distinguished by their constituent Argonaute protein. Thus, it is crucial that a specific set of small guide RNAs is directed to a specific Argonaute protein. Analyses of how different types of small RNA are channelled to different Argonaute proteins show that there are multiple mechanisms: the determinants for small RNA sorting vary from the structure of the small RNA duplex to the identity of the 5′ nucleotide and the presence and extent of modifications to this nucleotide.

In *D. melanogaster*, pre-miRNAs are processed by DCR-1, whereas exo-siRNA duplexes are produced by DCR-2 from long dsRNAs[58] (Fig. 2a). Small RNAs then seem to be loaded onto either AGO1 or AGO2, depending on the structure of a small intermediate RNA duplex[75]. If the duplex has a bulge in the middle (frequently observed in miRNA

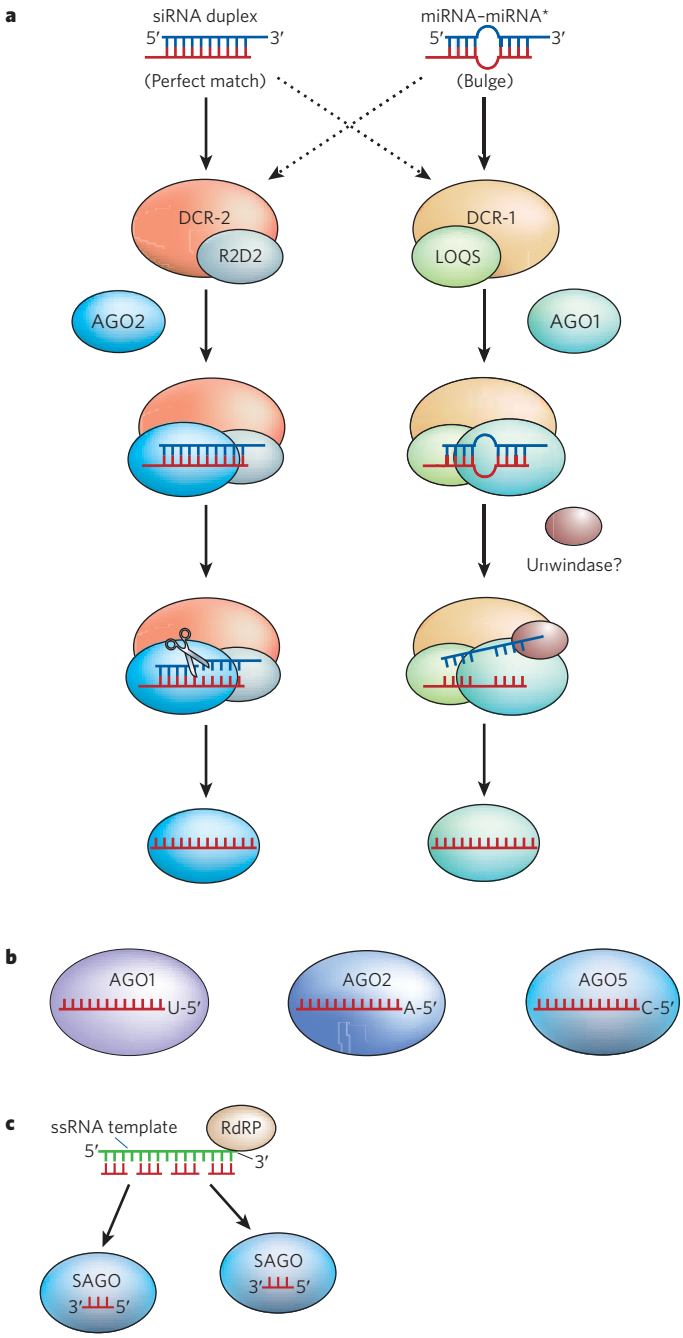

Figure 2 | Sorting of small RNAs onto distinct Argonaute proteins. Small RNAs are sorted onto specific Argonaute proteins, and this process occurs by several mechanisms. **a**, In *Drosophila melanogaster*, small RNAs originating from a duplex are loaded onto one of two Argonaute proteins (AGO1 or AGO2), on the basis of the structure of the small RNA duplex. If the duplex has a mismatch or a bulge in the centre (as miRNAs do), then the RNA is routed to AGO1. If the duplex is perfectly matched (as siRNAs are), then the small RNA is routed to AGO2. This selectivity occurs because the small RNAs are loaded onto Argonaute proteins from a Dicer-containing complex, and the two forms of Dicer, DCR-1 and DCR-2, associate with different RNA structures. DCR-2 pairs with R2D2, and this heterodimer binds to highly paired small RNA duplexes but recognizes small RNA duplexes with a central mismatch only poorly. AGO2 favours binding to DCR-2–R2D2 over binding to the other Dicer-containing complex, DCR-1–LOQS, which binds to small RNAs with bulges. Further processing into single-stranded small RNAs is described in Fig. 1. **b**, *Arabidopsis thaliana* miRNAs and *trans*-acting siRNAs (ta-siRNAs) have a 5′ uridine and preferentially associate with AGO1. By contrast, AGO2 and AGO5 show preferences for small RNAs containing 5′ adenosines and 5′ cytidines, respectively. However, it is unlikely that the 5′ nucleotide is the sole determinant of selective loading in *A. thaliana*. **c**, Secondary endo-siRNAs in *Caenorhabditis elegans*, as well in *Schizosaccharomyces pombe*, have a striking strand bias in which only the antisense siRNA is loaded onto Argonaute proteins. These siRNAs correspond to the RNA strand synthesized by RdRP. In *C. elegans*, RdRP produces small RNAs directly from the target mRNA in a primer-independent manner. Thus, these secondary small RNAs show negative polarity, and this mechanism reinforces the silencing carried out by the primary small RNAs.

precursors), the small RNA is routed to AGO1. If the duplex is perfectly matched, the small RNA is channelled to AGO2. This is because the DCR-2–R2D2 heterodimer, which recruits AGO2 to form the pre-RISC, binds well to highly paired small RNA duplexes but poorly to duplexes with central mismatches. Thus, the DCR-2–R2D2 heterodimer not only determines the polarity of siRNA loading on the basis of thermodynamic stability rules but also functions as a gatekeeper for AGO2-containing RISC assembly, promoting the incorporation of siRNAs over miRNAs. These observations suggest that each siRNA duplex dissociates from the active site of the Dicer protein after it is produced and is subsequently recaptured by the DCR-2–R2D2 heterodimer. However, although AGO1 favours binding to small RNA duplexes with central mismatches, a large proportion of miRNA–miRNA* duplexes with a base-paired central region still enter into AGO1-containing RISCs[55], suggesting that the AGO1-loading pathway is selective and not a default pathway for small RNAs rejected by the AGO2 pathway.

The identity of the nucleotide at the 5′ end and the extent to which this nucleotide is phosphorylated also influence which Argonaute protein the small RNA associates with. In contrast to what is observed in *D. melanogaster*, processing by Dicer may be uncoupled from association with Argonaute proteins in *A. thaliana* because, in this species, the miRNAs are all generated by one particular Dicer protein, DCL1, but are still sorted and loaded onto different Argonaute proteins. In *A. thaliana*, miRNAs and *trans*-acting siRNAs (ta-siRNAs), a class of small RNAs that regulate plant development[12], generally have a 5′ uridine and preferentially associate with AGO1 (ref. 76) (Fig. 2b). By contrast, AGO2 associates preferentially with small RNAs containing 5′ adenosines, and AGO5 prefers 5′ cytidines. Interestingly, if the opposite strand of a miRNA (that is, miRNA*) has a 5′ adenosine or a 5′ cytidine, it is bound to AGO2 or AGO5, respectively. These findings have led to the hypothesis that the binding affinity of Argonaute proteins for small RNAs is determined by the nucleotide at the 5′ end. Although these 5′-nucleotide preferences generally hold true for these Argonaute proteins in plants, exceptions have been observed: the *A. thaliana* miRNA known as miR-172 has a 5′ adenosine but preferentially associates with AGO1 (ref. 77); and AGO7 preferentially associates with

miR-390, which has a 5′ adenosine[77]. Therefore, the 5′ nucleotide does not seem to be the sole determinant of Argonaute association.

Another mechanism might operate for secondary siRNAs in *C. elegans*. These small RNAs are specifically loaded onto SAGOs[50]. Secondary siRNAs carry a 5′-triphosphate modification[51,52], the hallmark of RdRP products, which might function as a recognition element for SAGO binding while excluding binding by a primary Argonaute, such as RDE-1.

Endo-siRNAs in *C. elegans* (including the secondary siRNAs just mentioned) and *Schizosaccharomyces pombe* (fission yeast) have a striking strand bias in which only the antisense siRNA strand, corresponding to the RNA strand synthesized by RdRP, is loaded into Argonaute-containing complexes. Because *C. elegans* RdRPs produce small RNAs directly from the target mRNA, in a primer-independent manner (Fig. 2c), all secondary siRNAs have a negative polarity and function to reinforce the silencing of the target mRNA[50–52]. In *S. pombe*, the strand bias is probably the result of a different mechanism. The physical association of Dicer with an RdRP-containing complex known as RDRC and an Argonaute-containing complex known as the RNA-induced transcriptional silencing complex (RITS) (see page 413) may facilitate the loading of siRNAs onto Argonaute proteins in a directional manner as Dicer moves along and cleaves the dsRNA products of RdRP, giving rise to an antisense strand bias. This suggests that the polarity of Dicer processing defines the polarity of the siRNA strand loaded onto the Argonaute protein.

Argonaute proteins have diversified over evolutionary timescales, evolving a range of functions[8,12]. These findings about small RNA sorting imply that the diversification of the Argonaute proteins is a consequence of which small RNA they recruit. It is possible that the conformation of the Argonaute protein dictates which small RNAs it partners, but the structures of eukaryotic Argonaute proteins will need to be determined before this can be assessed.

Safeguards in silencing pathways

During RNA silencing, a single non-sequence-specific RNA-binding protein (Argonaute) is loaded with small guide RNAs with a variety of sequences, resulting in effector complexes (RISCs). Thus, this system requires gatekeepers to ensure that Argonaute can bind to small guide RNAs but not to degraded small RNAs, thereby avoiding 'off-target' silencing. Such gatekeeper systems seem to depend mainly on structural features specific for small guide RNAs.

As described earlier, Dicer helps to load siRNAs into the RISC, preventing siRNAs from diffusing freely in the cytoplasm after their production. This function of Dicer probably also aids in the discrimination of genuine siRNAs from various RNA-degradation products in the cell. Processing by RNaseIII enzymes (such as Dicer) characteristically yields small RNAs with 5′ monophosphates and 3′ two-nucleotide overhangs. The PAZ domain of Argonaute proteins might, as a first step, distinguish degraded RNAs (derived from unrelated pathways) from these small RNAs by binding to the characteristic 3′ overhangs of the small RNAs[8,12]. In addition, to become incorporated into the RISC and mediate cleavage of the target mRNA, the guide strand of an siRNA must have a phosphate group at the 5′ end[78]. In humans, the 5′ end of siRNAs is phosphorylated by the enzyme CLP1 (ref. 79), which also has roles in splicing transfer RNAs and forming the 3′ ends of mRNAs. Interestingly, both tRNA splicing and mRNA 3′-end formation occur in the nucleus[80,81], suggesting that siRNA duplexes with a 5′ hydroxyl group are transported to, or diffuse into, the nucleus and, after phosphorylation by CLP1, are exported to the cytoplasm and assembled into the RISC.

Amplification of the silencing signal needs to be balanced against the dangers of amplifying off-target silencing. For example, the slicer-mediated ping-pong mechanism for piRNA production does not lead to 'transitive' RNA silencing (in which RdRPs synthesize siRNAs complementary to sequences upstream or downstream of the initial trigger region in the target mRNA). Instead, it leads to conservative amplification of functional primary piRNA sequences (those inherited by germline transmission). However, it is conceivable that any off-target events mediated by RdRPs could lead to a chain reaction or transitive effect of silencing with deleterious consequences. Thus, there must be safeguards to prevent the pervasive use of RdRPs. A striking aspect of RdRP-based trigger amplification is that amplification occurs only when a target has been engaged, so amplification of the silencing signal is limited to cases in which there is a real target[51,52]. In *C. elegans*, the processing of the trigger dsRNA and the loading of primary siRNAs into the RDE-1-containing complex seem to be inherently inefficient, limiting the first round of target recognition by RDE-1-containing complexes and minimizing the risk of amplifying off-target silencing reactions[50]. In addition, each secondary siRNA seems to be generated by non-processive self-termination by RdRP, thereby restricting transitive effects[51–53]. Furthermore, secondary siRNAs associate with SAGOs, which lack catalytic residues for cleaving mRNAs, suggesting that these complexes cannot generate cleaved substrates for further amplification, which in turn would prevent them from inducing the exponential generation of secondary siRNAs[50] (but see also ref. 53 for a conflicting viewpoint). SAGOs are also present in limited supply and thus have a restricted capacity to support multiple simultaneous silencing reactions.

Another factor is that in *C. elegans* and *S. pombe* the RNAi machinery is negatively regulated by a conserved siRNA nuclease called enhanced RNAi (ERI-1 and Eri1, respectively)[82,83]. In *S. pombe*, transgene silencing is linked to a protein complex resembling the TRAMP complex of *Saccharomyces cerevisiae* (budding yeast), which carries out surveillance in the nucleus, targeting aberrant transcripts for degradation by the exosome[84]. Thus, RNAi in *S. pombe* is actively restricted from exerting its effects throughout the genome and seems to be subject to competition from RNA quality-control machinery.

Target-sensing modes and effector modes of the RISC

When the RISC is loaded with the guide strand of a small RNA, how does it find its target mRNA? Most of the binding energy that tethers a RISC to a target mRNA is from nucleotides in the seed region of the small RNA[85]. It seems that the accessibility of the target site can be sensed by the intrinsic, nonspecific affinity of RISC for ssRNA, which follows the initial specific association between the RISC and the target (through the 5′ seed region of the small RNA)[86]. But the accessibility of the target site correlates directly with the efficiency of cleavage, indicating that the RISC cannot unfold structured RNA.

Target mRNAs are present in the cell in complex with ribonucleoproteins (RNPs)[87], so target accessibility is also controlled by several RNA-binding proteins that either mask the target binding site or facilitate unfolding of the target. Therefore, the function of a RISC seems to be context-dependent, with its effector mode influenced not only by the structures of the small-RNA-binding sites on the target but also by the particular proteins associated with each Argonaute protein. For example, animal miRNAs silence gene expression by at least three independent mechanisms through binding sites that are mostly in the 3′ untranslated region of target mRNAs: by cleaving mRNAs, by repressing their translation and/or by promoting mRNA degradation[88,89]. However, the contribution of translational repression or mRNA degradation to gene silencing seems to differ for each miRNA–mRNA pair. Thus, the final outcome of miRNA regulation is probably affected by other proteins interacting with the targeted mRNA or RISC and counteracting the effects of the miRNA, resulting in differential regulation depending on the proteins present in each tissue[90].

Regulation of silencing pathways

So far, the pictures of RNA silencing pathways that we have built up (shown in Figs 1 and 2) are static. To gain further insight into silencing processes, it is important to incorporate information about how these pathways are regulated. It is already clear that competition between different silencing pathways (for example, competition between endo-siRNAs and miRNAs for LOQS in *D. melanogaster*) is a key step in how each stage of the RNAi mechanism is regulated. Many plant and animal viruses are known to encode suppressor proteins that block

host RNAi, and therefore silencing, at various stages[91]. Cellular proteins can also regulate RNAi. For example, processing to form the human miRNA let-7, which is a tumour suppressor and cell-cycle regulator, is post-transcriptionally inhibited in embryonic cells by the pluripotency factor LIN28, which seems to block the microprocessor-complex-mediated cleavage of pri-let-7 and the Dicer-mediated processing of pre-let-7 in series[92,93]. By contrast, in humans, signalling mediated by the transforming growth factor-β (TGF-β) and bone morphogenetic protein (BMP) family of growth factors rapidly increases the production of mature miR-21 (which is oncogenic), by promoting the processing of pri-miR-21 into pre-miR-21 by DROSHA[94]. More specifically, TGF-β- and BMP-specific signal transducers of the SMAD family are recruited to pri-miR-21 in complex with the RNA helicase p68, a component of the microprocessor complex, facilitating the accumulation of pre-miRNA. In addition, heterogeneous nuclear RNP A1 (hnRNP A1), a well-known regulator of precursor mRNA splicing, also assists DROSHA to crop and release pre-miR-18 efficiently, perhaps by refolding the hairpin or by creating a cleavage site for DROSHA through direct binding to the pri-miRNA[95]. This implies that some hairpins within pri-miRNAs might form and be processed only after the binding of a protein with RNA chaperone activity.

The activity of the RISC can also be regulated. In *A. thaliana*, the non-protein-coding gene *IPS1* (*INDUCED BY PHOSPHATE STARVATION 1*) contains a motif with sequence complementarity to the phosphate-starvation-induced miRNA miR-399, but the pairing is interrupted by a mismatched loop at the expected miRNA cleavage site[96]. *IPS1* mRNA is not cleaved but, instead, sequesters miR-399. Thus, *IPS1* overexpression results in increased accumulation of the target of miR-399, *PHO2* mRNA. The idea of target mimicry introduces unanticipated complexity into the network of RNA-regulatory interactions and raises the possibility that a large number of mRNA-like non-coding RNAs recently identified in humans[97] could be attenuators of the regulation of small-RNA–Argonaute complexes.

Perspective

Recent studies hint that human cells contain a large number of small RNAs similar to miRNAs or siRNAs, with the potential to regulate the expression of almost all human genes. The future challenges in this field are clear. Many questions remain to be answered. How many types of small RNA are there? How are these small RNAs generated? What are their biological functions? How are these pathways regulated? One potential problem is that because many types of small RNA are modified at their 5′ and 3′ ends[98], it is unclear whether the current sequencing technologies are sampling the entire range of small RNAs present in cells. But next-generation sequencing technologies[13] should soon help to uncover the full range of small RNA molecules.

One major challenge will be to identify how specific RNA-binding proteins affect the final outcome of gene regulation by small RNAs, given that RNAs in a cell are usually associated with multiple proteins that regulate many aspects of gene expression. For example, genome-wide *in vivo* approaches using a combination of immunoprecipitation and high-throughput sequencing will be required to establish protein–mRNA interactions or RNP complex occupancy at certain regions of mRNA, where expression is suppressed.

Finally, changes in the activity and specificity of silencing pathways could create quantitative and qualitative genetic variation in gene expression, thereby generating new gene-expression networks. Such changes might have contributed to many processes, including human evolution[16]. Given that all vertebrates have almost exactly the same number of protein-coding genes and therefore cannot readily be distinguished in this way, it might be prophetic that the first small guide RNA to be identified, the *C. elegans* miRNA lin-4, has been found to regulate a gene involved in the timing of development[99,100]. In humans, unlike other mammals, the brain tissue of newborns continues to grow at a similar rate to that of the fetus. This is a good example of a change in developmental timing, and there is much speculation about whether changes in this rate contributed to the evolution of humans as a new species. ■

1. Fire, A. *et al.* Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391,** 806–811 (1998).
2. Stefani, G. & Slack, F. J. Small non-coding RNAs in animal development. *Nature Rev. Mol. Cell Biol.* **9,** 219–230 (2008).
3. Ding, S. W. & Voinnet, O. Antiviral immunity directed by small RNAs. *Cell* **130,** 413–426 (2007).
4. Girard, A. & Hannon, G. J. Conserved themes in small-RNA-mediated transposon control. *Trends Cell Biol.* **18,** 136–148 (2008).
5. Hobert, O. Common logic of transcription factor and microRNA action. *Trends Biochem. Sci.* **29,** 462–468 (2004).
6. Meister, G. & Tuschl, T. Mechanisms of gene silencing by double-stranded RNA. *Nature* **431,** 343–349 (2004).
7. Tomari, Y. & Zamore, P. D. Machines for RNAi. *Genes Dev.* **19,** 517–529 (2005).
8. Hutvagner, G. & Simard, M. J. Argonaute proteins: key players in RNA silencing. *Nature Rev. Mol. Cell Biol.* **9,** 22–32 (2008).
9. Liu, J. *et al.* Argonaute2 is the catalytic engine of mammalian RNAi. *Science* **305,** 1437–1441 (2004).
10. Song, J. J., Smith, S. K., Hannon, G. J. & Joshua-Tor, L. Crystal structure of Argonaute and its implications for RISC slicer activity. *Science* **305,** 1434–1437 (2004).
11. Meister, G. *et al.* Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. *Mol. Cell* **15,** 185–197 (2004).
12. Chapman, E. J. & Carrington, C. Specialization and evolution of endogenous small RNA pathways. *Nature Rev. Genet.* **8,** 884–896 (2007).
13. Mardis, E. R. The impact of next-generation sequencing technology on genetics. *Trends Genet.* **24,** 133–141 (2008).
14. Zhang, H., Kolb, F. A., Jaskiewicz, L., Westhof, E. & Filipowicz, W. Single processing center models for human Dicer and bacterial RNase III. *Cell* **118,** 57–68 (2004).
15. MacRae, I. J., Zhou, K. & Doudna, J. A. Structural determinants of RNA recognition and cleavage by Dicer. *Nature Struct. Mol. Biol.* **14,** 934–940 (2007).
16. Heimberg, A. M. *et al.* MicroRNAs and the advent of vertebrate morphological complexity. *Proc. Natl Acad. Sci. USA* **105,** 2946–2950 (2008).
17. Kim, V. N. MicroRNA biogenesis: coordinated cropping and dicing. *Nature Rev. Mol. Cell Biol.* **6,** 376–385 (2005).
18. Han, J. *et al.* Molecular basis for the recognition of primary microRNAs by the Drosha–DGCR8 complex. *Cell* **125,** 887–901 (2006).
19. Zeng, Y., Yi, R. & Cullen, B. Recognition and cleavage of primary microRNA precursors by the nuclear processing enzyme Drosha. *EMBO J.* **24,** 138–148 (2005).
20. Denli, A. M., Tops, B. B., Plasterk, R. H., Ketting, R. F. & Hannon, G. J. Processing of primary microRNAs by the Microprocessor complex. *Nature* **432,** 231–235 (2004).
21. Gregory, R. I. *et al.* The Microprocessor complex mediates the genesis of microRNAs. *Nature* **432,** 235–240 (2004).
22. Ruby, J. G., Jan, C. H. & Bartel, D. P. Intronic microRNA precursors that bypass Drosha processing. *Nature* **448,** 83–86 (2007).
23. Okamura, K., Hagen, J. W., Duan, H., Tyler, D. M. & Lai, E. C. The mirtron pathway generates microRNA-class regulatory RNAs in *Drosophila*. *Cell* **130,** 89–100 (2007).
24. Lewis, B. P., Burge, C. B. & Bartel, D. P. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* **120,** 15–20 (2005).
25. Brennecke, J., Stark, A., Russell, R. B. & Cohen, S. M. Principles of microRNA–target recognition. *PLoS Biol.* **3,** e85 (2005).
26. Landgraf, P. *et al.* A mammalian microRNA expression atlas based on small RNA library sequencing. *Cell* **129,** 1401–1414 (2007).
27. Morin, R. D. *et al.* Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells. *Genome Res.* **18,** 610–621 (2008).
28. Azuma-Mukai, A. *et al.* Characterization of endogenous human Argonautes and their miRNA partners in RNA silencing. *Proc. Natl Acad. Sci. USA* **105,** 7964–7969 (2008).
29. Faehnle, C. R. & Joshua-Tor, L. Argonautes confront new small RNAs. *Curr. Opin. Chem. Biol.* **11,** 569–577 (2007).
30. Wang, Y., Sheng, G., Juranek, S., Tuschl, T. & Patel, D. J. Structure of the guide-strand-containing argonaute silencing complex. *Nature* **456,** 209–213 (2008).
31. Vagin, V. V. *et al.* A distinct small RNA pathway silences selfish genetic elements in the germline. *Science* **313,** 320–324 (2006).
This paper was the first to describe a Dicer-independent pathway for the biogenesis of the guide strand of small RNAs.
32. Saito, K. *et al.* Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the *Drosophila* genome. *Genes Dev.* **20,** 2214–2222 (2006).
33. Girard, A., Sachidanandam, R., Hannon, G. J. & Carmell, M. A. A germline-specific class of small RNAs binds mammalian Piwi proteins. *Nature* **442,** 199–202 (2006).
34. Aravin, A. *et al.* A novel class of small RNAs bind to MILI protein in mouse testes. *Nature* **442,** 203–207 (2006).
35. Lau, N. C. *et al.* Characterization of the piRNA complex from rat testes. *Science* **313,** 363–367 (2006).
36. Watanabe, T. *et al.* Identification and characterization of two novel classes of small RNAs in the mouse germline: retrotransposon-derived siRNAs in oocytes and germline small RNAs in testes. *Genes Dev.* **20,** 1732–1743 (2006).
37. Yin, H. & Lin, H. An epigenetic activation role of Piwi and a Piwi-associated piRNA in *Drosophila melanogaster*. *Nature* **450,** 304–308 (2007).
38. Saito, K. *et al.* Pimet, the *Drosophila* homolog of HEN1, mediates 2′-O-methylation of Piwi-interacting RNAs at their 3′ ends. *Genes Dev.* **21,** 1603–1608 (2007).
39. Horwich, M. D. *et al.* The *Drosophila* RNA methyltransferase, DmHen1, modifies germline piRNAs and single-stranded siRNAs in RISC. *Curr. Biol.* **17,** 1265–1272 (2007).
40. Kirino, Y. & Mourelatos, Z. Mouse Piwi-interacting RNAs are 2′-O-methylated at their 3′ termini. *Nature Struct. Mol. Biol.* **14,** 347–348 (2007).
41. Ohara, T. *et al.* The 3′ termini of mouse Piwi-interacting RNAs are 2′-O-methylated. *Nature Struct. Mol. Biol.* **14,** 349–350 (2007).
42. Houwing, S. *et al.* A role for Piwi and piRNAs in germ cell maintenance and transposon silencing in zebrafish. *Cell* **129,** 69–82 (2007).
43. Brennecke, J. *et al.* Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. *Cell* **128,** 1089–1103 (2007).

44. Gunawardane, L. S. *et al.* A slicer-mediated mechanism for repeat-associated siRNA 5′ end formation in *Drosophila*. *Science* **315,** 1587–1590 (2007).
References 43 and 44 were the first to describe slicer-mediated small RNA production.
45. Aravin, A. A., Sachidanandam, R., Girard, A., Fejes-Toth, K. & Hannon, G. J. Developmentally regulated piRNA clusters implicate MILI in transposon control. *Science* **316,** 744–747 (2007).
46. Ruby, J. G. *et al.* Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in *C. elegans*. *Cell* **127,** 1193–1207 (2006).
47. Batista, P. J. *et al.* PRG-1 and 21U-RNAs interact to form the piRNA complex required for fertility in *C. elegans*. *Mol. Cell* **31,** 67–78 (2008).
48. Das, P. P. *et al.* Piwi and piRNAs act upstream of an endogenous siRNA pathway to suppress Tc3 transposon mobility in the *Caenorhabditis elegans* germline. *Mol. Cell* **31,** 79–90 (2008).
49. Mochizuki, K. & Gorovsky, M. A. A Dicer-like protein in *Tetrahymena* has distinct functions in genome rearrangement, chromosome segregation, and meiotic prophase. *Genes Dev.* **19,** 77–89 (2005).
50. Yigit, E. *et al.* Analysis of the *C. elegans* Argonaute family reveals that distinct Argonautes act sequentially during RNAi. *Cell* **127,** 747–757 (2006).
This paper shows that RNAi occurs by a two-step pathway in *C. elegans*.
51. Pak, J. & Fire, A. Distinct populations of primary and secondary effectors during RNAi in *C. elegans*. *Science* **315,** 241–244 (2007).
52. Sijen, T., Steiner, F. A., Thijssen, K. L. & Plasterk, R. H. Secondary siRNAs result from unprimed RNA synthesis and form a distinct class. *Science* **315,** 244–247 (2007).
53. 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.* **26,** 5007–5019 (2007).
54. Czech, B. *et al.* An endogenous small interfering RNA pathway in *Drosophila*. *Nature* **453,** 798–802 (2008).
55. Kawamura, Y. *et al.* *Drosophila* endogenous small RNAs bind to Argonaute 2 in somatic cells. *Nature* **453,** 793–797 (2008).
56. Okamura, K. *et al.* The *Drosophila* hairpin RNA pathway generates endogenous short interfering RNAs. *Nature* **453,** 803–806 (2008).
57. Ghildiyal, M. *et al.* Endogenous siRNAs derived from transposons and mRNAs in *Drosophila* somatic cells. *Science* **320,** 1077–1081 (2008).
58. Lee, Y. S. *et al.* Distinct roles for *Drosophila* Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways. Cell **117,** 69–81 (2004).
59. Okamura, K., Ishizuka, A., Siomi, H. & Siomi, M. C. Distinct roles for Argonaute proteins in small RNA-directed RNA cleavage pathways. *Genes Dev.* **18,** 1655–1666 (2004).
60. Saito, K., Ishizuka, A., Siomi, H. & Siomi, M. C. Processing of pre-microRNAs by the Dicer-1-Loquacious complex in *Drosophila* cells. *PLoS Biol.* **3,** e235 (2005).
61. Förstemann, K. *et al.* Normal microRNA maturation and germ-line stem cell maintenance requires Loquacious, a double-stranded RNA-binding domain protein. *PLoS Biol.* **3,** e236 (2005).
62. Liu, Q. *et al.* R2D2, a bridge between the initiation and effector steps of the *Drosophila* RNAi pathway. *Science* **301,** 1921–1925 (2003).
63. Tam, O. H. *et al.* Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. *Nature* **453,** 534–538 (2008).
64. Watanabe, T. *et al.* Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. *Nature* **453,** 539–543 (2008).
65. Schwarz, D. S. *et al.* Asymmetry in the assembly of the RNAi enzyme complex. *Cell* **115,** 199–208 (2003).
66. Khvorova, A., Reynolds, A. & Jayasena, S. D. Functional siRNAs and miRNAs exhibit strand bias. *Cell* **115,** 209–216 (2003).
67. Liu, X., Jiang, F., Kalidas, S., Smith, D. & Liu, Q. Dicer-2 and R2D2 coordinately bind siRNA to promote assembly of the siRISC complexes. *RNA* **12,** 1514–1520 (2006).
68. Tomari, Y., Matranga, C., Haley, B., Martinez, N. & Zamore, P. D. A protein sensor for siRNA asymmetry. *Science* **306,** 1377–1380 (2004).
69. Matranga, C., Tomari, Y., Shin, C., Bartel, D. P. & Zamore, P. D. Passenger-strand cleavage facilitates assembly of siRNA into Ago2-containing RNAi enzyme complexes. *Cell* **123,** 607–620 (2005).
70. Rand, T. A., Petersen, S., Du, F. & Wang, X. Argonaute2 cleaves the anti-guide strand of siRNA during RISC activation. *Cell* **123,** 621–629 (2005).
71. Miyoshi, K., Tsukumo, H., Nagami, T., Siomi, H. & Siomi, M. C. Slicer function of *Drosophila* Argonautes and its involvement in RISC formation. *Genes Dev.* **19,** 2837–2848 (2005).
72. Gregory, R. I., Chendrimada, T. P., Cooch, N. & Shiekhattar, R. Human RISC couples microRNA biogenesis and posttranscriptional gene silencing. *Cell* **123,** 631–640 (2005).
73. Maniataki, E. & Mourelatos, Z. A human, ATP-independent, RISC assembly machine fueled by pre-miRNA. *Genes Dev.* **19,** 2979–2990 (2005).
74. Robb, G. B. & Rana, T. M. RNA helicase A interacts with RISC in human cells and functions in RISC loading. *Mol. Cell* **26,** 523–537 (2007).
75. Tomari, Y., Du, T. & Zamore, P. D. Sorting of *Drosophila* small silencing RNAs. *Cell* **130,** 299–308 (2007).
76. Mi, S. *et al.* Sorting of small RNAs into *Arabidopsis* Argonaute complexes is directed by the 5′ terminal nucleotide. *Cell* **133,** 116–127 (2008).
77. Montgomery, T. A. *et al.* Specificity of ARGONAUTE7–miR390 interaction and dual functionality in *TAS3 trans*-acting siRNA formation. *Cell* **133,** 128–141 (2008).
References 76 and 77 show that the sorting of plant miRNAs onto Argonaute proteins depends mainly on the nucleotide at the 5′ end.
78. Pham, J. W. & Sontheimer, E. J. Molecular requirements for RNA-induced silencing complex assembly in the *Drosophila* RNA interference pathway. *J. Biol. Chem.* **280,** 39278–39283 (2005).
79. Weitzer, S. & Martinez, J. The human RNA kinase hClp1 is active on 3′ transfer RNA exons and short interfering RNAs. *Nature* **447,** 222–226 (2007).
80. Paushkin, S. V., Patel, M., Furia, B. S., Peltz, S. W. & Trotta, C. R. Identification of a human endonuclease complex reveals a link between tRNA splicing and pre-mRNA 3′ end formation. *Cell* **117,** 311–321 (2004).
81. Danckwardt, S., Hentze, M. W. & Kulozik, A. E. 3′ end mRNA processing: molecular mechanisms and implications for health and disease. *EMBO J.* **27,** 482–498 (2008).
82. Kennedy, S., Wang, D. & Ruvkun, G. A conserved siRNA-degrading RNase negatively regulates RNA interference in *C. elegans*. *Nature* **427,** 645–649 (2004).
83. Iida, T., Kawaguchi, R. & Nakayama, J. Conserved ribonuclease, Eri1, negatively regulates heterochromatin assembly in fission yeast. *Curr. Biol.* **16,** 1459–1464 (2006).
84. Bühler, M., Haas, W., Gygi, S. P. & Moazed, D. RNAi-dependent and -independent RNA turnover mechanisms contribute to heterochromatic gene silencing. *Cell* **129,** 707–721 (2007).
85. Haley, B. & Zamore, P. D. Kinetic analysis of the RNAi enzyme complex. *Nature Struct. Mol. Biol.* **11,** 599–606 (2004).
86. Ameres, S. L., Martinez, J. & Schroeder, R. Molecular basis for target RNA recognition and cleavage by human RISC. *Cell* **130,** 101–112 (2007).
87. Dreyfuss, G., Kim, V. N. & Kataoka, N. Messenger-RNA-binding proteins and the messages they carry. *Nature Rev. Mol. Cell Biol.* **3,** 195–205 (2002).
88. Filipowicz, W., Bhattacharyya, S. N. & Sonenberg, N. Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nature Rev. Genet.* **9,** 102–114 (2008).
89. Eulalio, A., Behm-Ansmant, I. & Izaurralde, E. P bodies: at the crossroads of post-transcriptional pathways. *Nature Rev. Mol. Cell Biol.* **8,** 9–22 (2007).
90. Kedde, M. *et al.* RNA-binding protein Dnd1 inhibits microRNA access to target mRNA. *Cell* **131,** 1273–1286 (2007).
This paper shows that the final outcome of miRNA regulation is affected by the interaction of proteins other than Argonaute with the target mRNA.
91. Mlotshwa, S., Pruss, G. J. & Vance, V. Small RNAs in viral infection and host defense. *Trends Plant Sci.* **13,** 375–382 (2008).
92. Viswanathan, S. R., Daley, G. Q. & Gregory, R. I. Selective blockade of microRNA processing by Lin28. *Science* **320,** 97–100 (2008).
93. Rybak, A. *et al.* A feedback loop comprising *lin-28* and *let-7* controls pre-*let-7* maturation during neural stem-cell commitment. *Nature Cell Biol.* **10,** 987–993 (2008).
94. Davis, B. N., Hilyard, A. C., Lagna, G. & Hata, A. SMAD proteins control DROSHA-mediated microRNA maturation. *Nature* **454,** 56–61 (2008).
95. Guil, S. & Cáceres, J. F. The multifunctional RNA-binding protein hnRNP A1 is required for processing of miR-18a. *Nature Struct. Mol. Biol.* **14,** 591–596 (2007).
96. Franco-Zorrilla, J. M. *et al.* Target mimicry provides a new mechanism for regulation of microRNA activity. *Nature Genet.* **39,** 1033–1037 (2007).
This paper describes how the activity of miRNAs can be regulated by transcripts that mimic the target transcript.
97. ENCODE Project Consortium. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* **447,** 799–816 (2007).
98. Farazi, T. A., Juranek, S. A. & Tuschl, T. The growing catalog of small RNAs and their association with distinct Argonaute/Piwi family members. *Development* **135,** 1201–1214 (2008).
99. Lee, R. C., Feinbaum, R. L. & Ambros, V. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* **75,** 843–854 (1993).
100. Wightman, B., Ha, I. & Ruvkun, G. Post-transcriptional regulation of the heterochronic gene *lin-14* by *lin-4* mediates temporal pattern formation in *C. elegans*. *Cell* **75,** 855–862 (1993).

Acknowledgements We apologize to colleagues whose relevant primary publications were not cited because of space constraints. We thank Y. Tomari, K. Aoki, Y. Watanabe and all the members of the Siomi laboratory for their comments and critical reading of the manuscript. Work in our laboratory is supported by grants from the Ministry of Education, Culture, Sports, Science and Technology (MEXT, Japan) and the New Energy and Industrial Technology Development Organization (Japan). M.C.S. is associate professor of the Global Centre of Excellence for Human Metabolomics Systems Biology (MEXT).

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence should be addressed to the authors (awa403@sc.itc.keio.ac.jp; siomim@sc.itc.keio.ac.jp).