Q&A

MOLECULAR BIOLOGY

The expanding world of small RNAs

Helge Großhans and Witold Filipowicz

Molecular cell biology has long been dominated by a protein-centric view. But the emergence of small, non-coding RNAs challenges this perception. These plentiful RNAs regulate gene expression at different levels, and have essential roles in health and disease.

How many classes of RNA have been identified?

There are three main types of 'classic' RNA: messenger RNA, transfer RNA and ribosomal RNA. mRNAs are translated into proteins, whereas tRNAs and rRNAs have housekeeping roles during mRNA translation. Small RNAs are not translated into proteins. Instead, these 20–30-nucleotide sequences regulate various biological processes, often by interfering with mRNA translation. Small RNAs come in different forms (Box 1), the best-understood classes being small interfering RNAs (siRNAs), microRNAs (miRNAs) and Piwi-associated RNAs (piRNAs).

When were small RNAs discovered?

During the past two to three decades, researchers identified many small regulatory RNAs, ranging in size from roughly 70 to 300 nucleotides, in various organisms. But the 20–30-nucleotide small regulatory RNAs were discovered more recently because, owing to their small size, they are easily missed in biochemical analysis, and they are poor targets for inactivation by classical genetic tools. Moreover, many miRNAs can compensate for each other's function, which makes their identification on the basis of the overt consequences of their absence difficult. The first two miRNAs to be discovered, *lin-4* and *let-7*, were identified in the 1990s through genetic experiments in the worm *Caenorhabditis elegans*. Then siRNAs were identified in animals, plants and fungi as the effector molecules that mediate the process of sequence-specific gene silencing, or RNA interference (RNAi), in response to double-stranded RNA (dsRNA; Fig. 1). The discovery of RNAi earned its two lead researchers, Andrew Fire and Craig Mello, the 2006 Nobel Prize in Physiology or Medicine.

Have we identified all small RNAs?

Hardly. Although hundreds of small RNAs have been identified in plants and animals, high-throughput sequencing and sophisticated bioinformatics tools repeatedly reveal the existence of new miRNAs and siRNAs. The human genome could encode more than a thousand miRNAs, the equivalent of a few per cent of protein-coding genes. Moreover, new classes of small RNAs continue to be discovered.

Do all organisms have small RNAs?

With the notable exception of the budding yeast *Saccharomyces cerevisiae*, almost all eukaryotic organisms (the cells of which contain a nucleus) investigated so far have siRNAs or at least cellular machineries to produce them. Evolutionarily, miRNAs seem to be younger than siRNAs. miRNAs function primarily in multicellular organisms, although they have recently been identified in a unicellular alga, *Chlamydomonas reinhardtii*. Certain DNA viruses also express miRNAs. Although 20–30-nucleotide small RNAs have not been identified in archaea and eubacteria (the cells of which lack a nucleus), Argonaute proteins — the effector molecules of small RNAs — are present in some of these organisms.

How are small RNAs made?

Generally, by fragmentation of longer RNA sequences. The precursors of miRNAs and siRNAs are dsRNAs, which are processed to small RNAs by dedicated sets of enzymes and other proteins (Box 2, page 416).

Where do the dsRNAs come from?

Depending on the class of small RNA, the source of precursor dsRNA differs. For siRNAs, dsRNA can form when complementary DNA strands are transcribed into RNA sequences. Viral infection of a cell can also supply dsRNAs, as many viruses form RNAs of both sense and antisense polarity during replication of their genomes, and this can trigger an RNAi response by the cell, as part of its antiviral defence. By contrast, miRNAs are excised from 'purpose-built', genome-encoded RNA precursors that fold into long hairpins resembling dsRNA. The expression of miRNA-encoding genes and those encoding mRNAs are controlled very similarly and involve the same RNA-synthetic machinery, including the enzyme RNA polymerase II.

Box 1 | Prominent members of the RNA family

Classic RNAs mediating protein synthesis

mRNAs (messenger RNAs) Transcripts of protein-coding genes that act as templates for protein synthesis.

rRNAs (ribosomal RNAs) RNA constituents of the ribonucleoprotein particles known as ribosomes, which mediate the decoding of mRNAs to the amino-acid sequences of proteins.

tRNAs (transfer RNAs) Adapter molecules carrying individual amino acids to the site of protein synthesis that recognize specific codons in mRNA.

Non-coding regulatory RNAs

siRNAs (small interfering RNAs) Small RNAs (20–25 nucleotides in length) formed through cleavage of long double-stranded RNA molecules. siRNAs are particularly important for taming the activity of transposons and combating viral infection, but they can also regulate protein-coding genes. Synthetic siRNAs can also be artificially expressed for experimental purposes.

miRNAs (microRNAs) Small RNAs (20–25 nucleotides in length) that are encoded by specific genes and function in repressing mRNA translation or in mRNA degradation in plants and animals. They are processed from long, single-stranded RNA sequences that fold into hairpin structures.

piRNAs (Piwi-associated RNAs) Small RNAs (25–30 nucleotides in length) that are generated from long single-stranded precursors. They function in association with the Piwi subfamily of Argonaute proteins, and are essential for the development of germ cells.

Longer non-coding RNAs RNAs of 70 to thousands of nucleotides that participate in various cellular processes, including mRNA splicing and ribosome biogenesis. **H.G. & W.F.**

ALL IMAGES R. JORGENSEN, UNIV. ARIZONA, TUCSON

FROM LEFT TO RIGHT: *FUNCT. PLANT BIOL.* **29,** 1501–1506 (2002); *PLANT J.* **13,** 401–409 (1998); *PLANT CELL* **2,** 279–289 (1990)

Figure 1 | First phenotypic description of RNA interference. White sections in petunia flowers represent areas where RNAi has silenced a gene involved in flower coloration.

How do small RNAs function?

They recognize their RNA targets by sequence-specific base-pairing. The outcome of the small-RNA–mRNA association depends on the degree of complementarity between the two sequences. When base-pairing is perfect, or almost perfect, as is the case for siRNAs (and possibly piRNAs), the target mRNA is cleaved in the middle of the small-RNA–mRNA duplex. Most plant miRNAs and some animal miRNAs function similarly. But most animal miRNAs base-pair imprecisely with mRNAs to repress their translation or to induce their breakdown. Irrespective of base-pairing precision, small RNAs rely on proteins of the Argonaute family for their activity. In fact, it is the protein partners of small RNAs that bring about repression of translation or mRNA cleavage; small RNAs act only as guides to tell Argonaute proteins which mRNAs to target.

So can their mRNA targets be predicted through sequence analysis?

Unfortunately, the devil is in the detail — at least in the case of animal miRNAs, which mostly base-pair to their mRNA targets with limited complementarity. Although sequence analyses have revealed some criteria for interaction between miRNA and mRNA, and many bioinformatics tools for target identification are available, sequence-based predictions frequently yield false positives or miss true targets. So identification of bona fide miRNA targets requires extensive experimentation. By contrast, most targets of siRNAs and plant miRNAs can be reliably predicted on the basis of near-perfect sequence complementarity.

To prevent protein synthesis, isn't it simpler to stop mRNA production?

Yes. Intuitively, terminating transcription seems a much more obvious mechanism. But a block in protein production always lags behind a block in transcription — even if transcription is stopped, mRNA sequences that have already been made can still be translated into proteins. So by targeting the existing mRNA pool, small RNAs block or attenuate protein synthesis very rapidly and, occasionally, even reversibly. In addition, because individual small RNAs can simultaneously target tens if not hundreds of mRNAs for different proteins, they are well suited to coordinate the expression of genes that function in the same or related pathways. For example, during zebrafish embryonic development, a specific miRNA, *miR-430*, targets hundreds of mRNAs for rapid degradation, facilitating embryos' transition to a new developmental programme that requires a separate set of proteins. The ability of different miRNAs to concurrently target several sequences of the same mRNA further increases their potential to fine-tune gene expression.

Is all of this small-RNA-mediated regulation post-transcriptional?

No — small RNAs also affect DNA transcription, particularly in plants and fission yeast. They do this by sequence-specific targeting of chromatin (complexes of DNA with histone proteins), converting it to the heterochromatin form that is not easily accessible to the transcriptional machinery. Strikingly, in some lower eukaryotes small RNAs also direct massive genomic DNA rearrangements. In mammals, however, there is currently only limited evidence for small-RNA functions other than post-transcriptional regulation.

Do small RNAs always silence gene expression?

In some conditions, small RNAs may also activate gene expression, although the mechanisms are currently not well understood. Indeed, a liver-specific miRNA, *miR-122*, is even needed for successful replication of the hepatitis C virus.

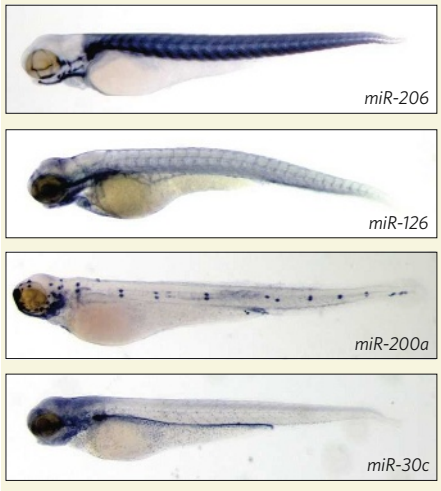
What biological processes do small RNAs regulate?

miRNAs were originally identified in *C. elegans* for their central role in development. Consistent with their function in differentiation and development, expression of many miRNAs is tissue-specific (Fig. 2) or is associated with certain developmental stages. miRNA expression patterns often change in diseases such as cancer. And, as many of the known and predicted miRNA targets have roles in disease, it is widely believed that dysregulation of miRNA expression contributes to disease pathology.

It is less clear whether siRNAs have similarly important functions, although in plants they have already been identified as essential players in the regulation of stress resistance. In fission yeast and plants, siRNAs contribute to heterochromatin formation.

And what is the link to viruses?

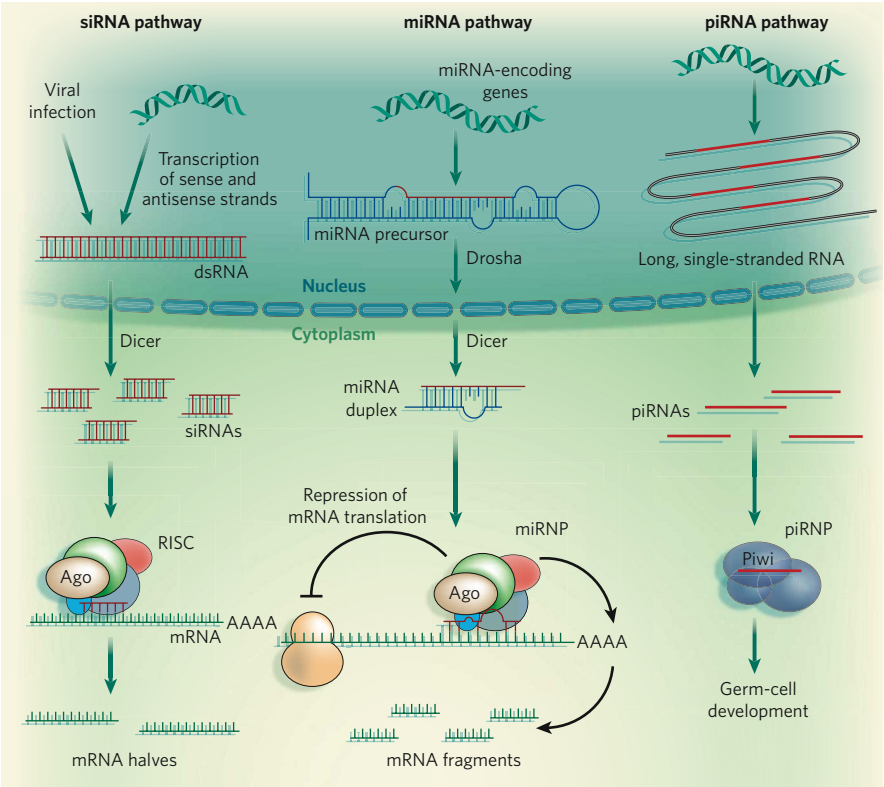
The use of RNAi as a defence mechanism against viruses may have been a driving force in the evolution of the siRNA pathway. In plants, siRNAs are an essential layer of antiviral defence. Also, in plants and invertebrates, siRNAs silence mobile genetic elements called transposons, which would otherwise 'jump' around the genome and disrupt cellular genes. It is not well known whether these small-RNA functions are also crucial in vertebrates, in which the invention of a protein-based adaptive immune response may have reduced reliance on antiviral RNAi activity.



E. WIENHOLDS *ET AL. SCIENCE* **309,** 310–311 (2005)

Figure 2 | miRNAs have tissue-specific functions. In zebrafish embryos, localization patterns of individual miRNAs indicate that their activity could be limited to tissues and organs in which they are expressed. As indicated by blue staining, *miR-206* is mainly expressed in the muscle, *miR-126* in the blood vessels and the heart, *miR-200a* in the lateral-line system (a mechanosensory system detecting water motion) and sensory organs, and *miR-30c* in the kidney precursor.

Box 2 | Formation and function of small RNAs



Small RNAs are generally produced by fragmentation of longer precursors. **Small interfering RNAs (siRNAs)** are processed from double-stranded RNAs (dsRNAs) that form by base-pairing of complementary RNAs. An enzyme called Dicer cleaves dsRNA into shorter double-stranded siRNAs that are roughly 20 base pairs long. One siRNA strand then assembles into an effector complex known as an RNA-induced silencing complex (RISC). This complex uses the siRNA guide to identify mRNAs with a sequence perfectly complementary to the siRNA. RISC then cleaves the mRNA in the middle of the mRNA–siRNA duplex, and the resulting mRNA halves are degraded by other cellular enzymes. Under some circumstances, siRNAs might associate into complexes other than RISC (not shown) that function in the nucleus and silence gene transcription. **MicroRNAs (miRNAs)** are processed from specific genome-encoded precursors, which fold into intramolecular hairpins containing imperfectly base-paired segments. The processing generally occurs in two steps, and is catalysed by the enzymes Drosha (in the nucleus) and Dicer (in the cytoplasm). One strand of the resulting miRNA duplex, resembling an siRNA, then incorporates into a RISC-like miRNA–ribonucleoprotein (miRNP) complex. The main components of RISC and miRNPs are proteins of the Argonaute (Ago) family.

Depending on the level of complementarity, miRNAs induce mRNA degradation or repress their translation. Unlike the siRNA pathway, miRNA-mediated degradation is initiated by enzymatic removal of the mRNA poly(A) tail. **Piwi-associated RNAs (piRNAs)** are generated from long, single-stranded precursors in a process independent of Drosha and Dicer. These small RNAs associate with a subfamily of Argonaute proteins called Piwi proteins. Tens of thousands of piRNAs have been identified, although they are far from understood. It is, however, known that, together with their Piwi partners, they are essential for the development of germ cells. **H.G. & W.F.**

Does our collection of small RNAs set us apart from other species?

Some miRNAs are highly conserved, but others vary greatly among organisms; some differ even among primates, for example between apes and humans. With the emerging view that regulation of protein activity could be as vital to evolution as fidelity of the protein sequence itself, it is tempting to speculate that miRNAs influence evolution. Relatively simple requirements for miRNA–mRNA interaction facilitate the development of new regulatory relationships between these sequences, possibly contributing to the evolution of new functions. Perhaps, therefore, it is not surprising that a large fraction of tissue-specific miRNAs operates in the brain.

Are small RNAs restricted to the cells in which they are made?

In *C. elegans* and plants, dsRNAs or siRNAs can move between cells or even longer distances. For example, siRNAs spread through the vascular system of plants, which possibly aids their function in antiviral immunity. And in *C. elegans*, an efficient system of siRNA amplification ensures the maintenance of gene silencing even after the initial 'trigger siRNA' is gone, allowing siRNA to spread to the organism's progeny. A similar amplification system does not occur in mammals. As for miRNAs, their very specific localization patterns (Fig. 2), and the absence of developmental changes in *C. elegans* mutants of the dsRNA transport machinery, suggests that miRNAs are stationary.

Will small RNAs be useful as therapeutic agents or targets?

This is a hot topic of research. siRNAs have the potential to silence disease-relevant genes that cannot be shut down with available drugs. Moreover, the so-called oncomiRs — miRNAs that promote cancer — may themselves be targets for shut-down. But we are still a long way from translating activity observed in a defined experimental system into an effective therapeutic drug. One of the most problematic issues is how to get small RNAs efficiently and specifically to their target site of action in the human body. But regardless of their therapeutic potential, siRNAs have already revolutionized basic biomedical research. The use of synthetic siRNAs, or their short hairpin RNA or dsRNA precursors, allows researchers to repress the function of a gene of interest or even to perform genome-wide RNAi screens to unravel entire biological pathways with unprecedented ease and speed.

So what of the future?

New classes of small RNAs continue to be discovered, and it is unlikely that we have found them all. Even for the known classes, we often have only a very limited understanding of what they do and how they do it. Identification of miRNA targets is another challenge, as is identifying other, currently hypothetical, modes of small-RNA action. Owing to their base-pairing potential, small RNAs could modify local mRNA structures, allowing for alternative mRNA splicing and modulating interactions of mRNAs with proteins. So watch this space. ■

Helge Großhans and Witold Filipowicz are at the Friedrich Miescher Institute for Biomedical Research, CH-4002 Basel, Switzerland.
e-mails: helge.grosshans@fmi.ch; witold.filipowicz@fmi.ch

FURTHER READING
Aravin, A. A, Hannon, G. J. & Brennecke, J. *Science* **318,** 761–764 (2007).
Bushati, N. & Cohen, S. M. *Annu. Rev. Cell. Dev. Biol.* **23,** 175–205 (2007).
Chapman, E. J. & Carrington, J. C. *Nature Rev. Genet.* **8,** 884–896 (2007).
Filipowicz, W., Bhattacharyya, S. N. & Sonenberg, N. *Nature Rev. Genet.* **9,** 102–114 (2008).
Rana, T. M. *Nature Rev. Mol. Cell Biol.* **8,** 23–36 (2007).
Zamore, P. D. & Haley, B. *Science* **309,** 1519–1524 (2005).