News from the RNA World

Microinformation

WITOLD FILIPOWICZ
Friedrich Miescher Institute for Biomedical Research and University of Basel, Switzerland
Foreign member, PAS
Witold.Filipowicz@fmi.ch
Earlier this year Prof. Witold Filipowicz received the prestigious Lifetime Achievement in Science Award from the international RNA Society

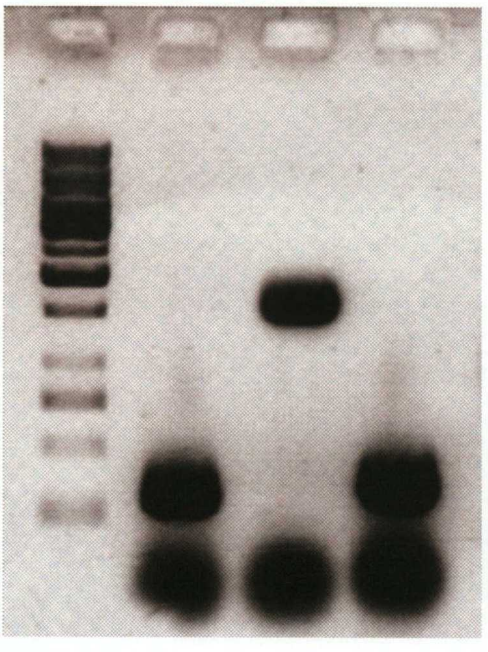
It was long believed that RNA was a molecule specialized in transmitting information from DNA to proteins. However, in recent years scientists have been discovering hitherto unknown complexities of RNA

Prof. Witold Filipowicz, interviewed by *Academia* about his many years of studying RNA, says that there are two types of scientists: those who "grind" (spend their entire lives researching a narrow field), and those who "cherry-pick" (jumping from discipline to discipline, they are often accused of opportunism). He says of himself that he is a bit of both types.

I have devoted my entire career to studying RNA, although our understanding of this family of molecules has become much broader in that time. When I was starting out, we knew about transfer RNA, ribosomal RNA, and messenger RNA (tRNA, rRNA, and mRNA). The real adventure started about 30 years ago with the discovery of different types of regulatory RNAs that did not fit into any of the above categories. It also turned out that RNA can have enzymatic properties, acting as a ribozyme. This blurred the boundaries between RNA, DNA, and proteins. Previously we believed that RNA simply transfers information from DNA to protein, or functions as the genetic material in certain viruses. Instead, it turned out that it can also contain information or act as an enzyme. This popularized earlier hypotheses that our world was originally ruled by RNA rather than DNA. The molecule is capable of self-replication: it can create new molecules from its own matrix, including new ribozymes. These discoveries led huge strides to be made in RNA research, and attracted thousands of new scientists interested in the analysis of RNA and studying its evolution.

RNA research has allowed us to solve the paradox of non-coding regions of the genome. There have been various hypotheses attempting to explain why just approximately 1% of the genomes of humans and many other mammals encodes proteins. It was believed that the rest is junk DNA or selfish DNA: superfluous ballast; an evolutionary relict. Today we know that the remainder, or at least a significant part of it, carries information for RNA that does not code for proteins; that those DNA sequences are transcribed to RNA and play a variety of extremely important regulatory functions. Of course, like all breakthrough discoveries, this hypothesis was treated with a large dose of skepticism to start with. The non-coding DNA was referred to as noise; scientists believed that RNA polymerase skims right over it, and the resulting superfluous products are broken down.

We now know that there are tens of thousands of non-coding RNA molecules.



Patrycja Dołowy

Acrylamide gels were once run down to the size of tRNA. Other RNAs which appeared below that on the gel were believed to be degradation products. Now many researchers are reconsidering their old gels

Mr Checker/Thomas Wydra, Wikipedia Commons

The more complex the eukaryotic organism, the more alternative ways of folding mRNA precursors. The fruit fly (*Drosophila melanogaster*), shown here, is a model organism in molecular biology research

Some are extremely long, running into the thousands of nucleotides, others range between 100-300 nucleotides, and there is also a whole world of small RNAs, including microRNAs, small interfering RNAs, and piwi-interacting RNAs (miRNAs, siRNAs, and piRNAs, respectively). This discovery came as a major surprise, since it had evaded researchers for so many years. It was understandable for small RNAs, since they all simply "run off" from our fractionation gels. Acrylamide gels were once run down to the size of the tRNA. Other RNAs which appeared on the gel below tRNA were believed to be degradation products. Now many people are reconsidering their old gels - "I saw them, they were there!" - but at the time no one realized how important they were.

Our understanding of proteins and the genes that encode them is extensive, thanks to genome sequencing. But now we have something like a whole new genome which needs to be analyzed almost from scratch to find out what it does and how. Over ten years ago I decided to abandon my old research on non-coding medium-size RNAs and start working on small RNAs, which are really fascinating. Small RNAs act as regulators for genes encoding proteins, and as regulators of chromatin structure, development, and differentiation. Almost all biological processes are regulated by miRNAs. We still do not fully understand their mechanism of action, although everything suggests that they regulate translation - the process of protein biosynthesis - at two levels: as regulators of translation as such, especially initiation, but also as stimulators of deadenylation and degradation of mRNA. This is regulation at many levels: attenuating translation, and then enhancing the effect through eliminating mRNA that is no longer needed by its degradation.

RNA research is a fascinating and extremely competitive field. One of the breakthroughs came when scientists realized that genes are not continuous, that they are interrupted by introns. In the last 20 years, the process of splicing of mRNA precursors has been one of biology's most fascinating problems. Researchers have been striving to understand why the human genome contains a similar number of genes as those of the fruit fly or the nematode worm Caenorhabditis elegans, yet it "codes" for such a complex organism. It seems that for human genes there are so many different ways of splicing them that each gene can be seen as a collection of ten different forms of mRNA, encoding different proteins. This is likely to have formed the basis for the evolution of multicellular complex organisms. Another important factors are the non-coding regulatory RNAs, which are still not fully understood, especially the largest ones which have only been discovered in the last two or three years. Around eight to ten thousand have been catalogued so far. They operate on the basis of modules, substructures, just like proteins. Proteins with enzymatic properties also contain additional domains, which can attach them to RNA molecules. It seems that the properties of long RNAs also depend on their structure. They frequently fold into modular structures resembling proteins, and they can also have many domains: for example, one domain can be responsible for targeting the non-coding RNA to a specific location along the chromatin, while another can interact with a protein with enzymatic properties. This is another world of molecules - a world of genes waiting to be disentangled...

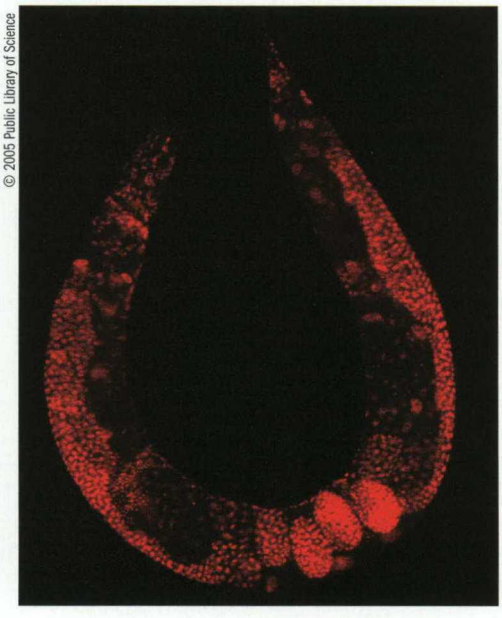
Have proteins been dethroned?

The majority of RNAs regulates the expression of DNA that encodes proteins, either directly or indirectly. These days, nothing important can happen without proteins. Early in evolution, RNA was likely to have been responsible for everything, but now proteins form the basis of the majority of cel-

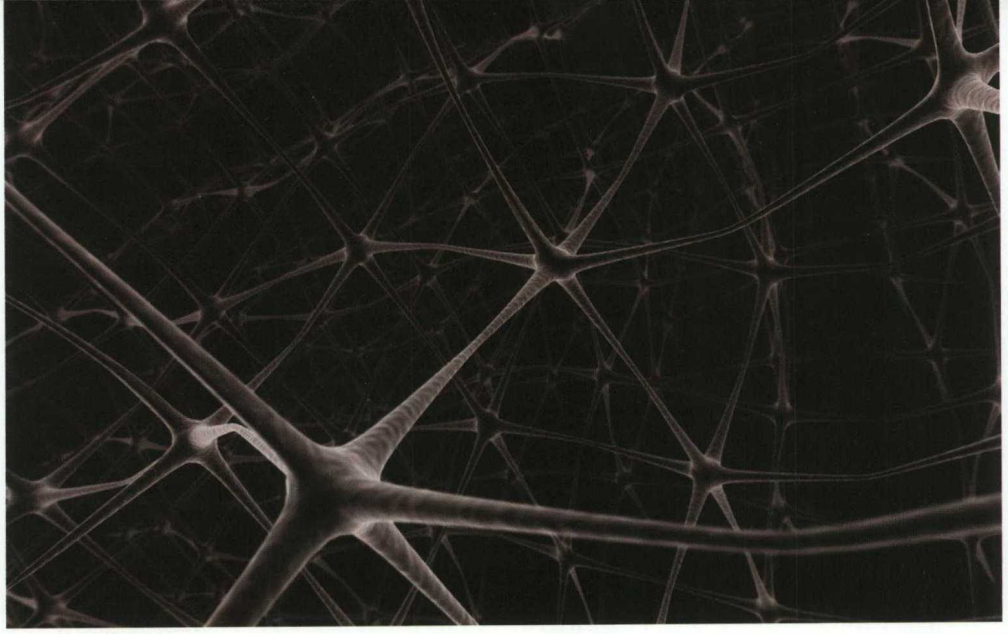
lular functions. However, there are also boundary structures such as ribonucleoproteins, complexes of RNA and proteins. They are reminders of the early days of RNA, and there are hundreds of them. They operate at different levels; for example, mRNA splicing is fully controlled by small ribonucleoprotein particles. Proteins have largely taken over the enzymatic functions of RNA, although RNA remains amazingly useful. It can hybridize to sequences of other RNAs by base pairing. This makes RNA such a universal regulator. Small RNAs frequently contain short nucleotide fragments which give them a local specificity, allowing them to attach to DNA or mRNA sequences, for example at the boundary between introns and exons. The same protein enzyme can be associated with various types of RNA, which then guides it to different locations – it is an extremely economical process. There is no need to build an entire collection of protein enzymes, which must undergo transcription, translation, and folding; it is sufficient to attach the protein to various "guide" RNAs of 20 or 100 nucleotides in length, which will give the protein specificity. Ribosomal RNA, which I used to study, is modified in many locations by methylation or by an exchange of uridine for pseudouridine. All these modifications (there are approximately a hundred methylations and pseudouridylations) are directed by small RNAs which target the same enzyme to different locations in rRNA to carry out its modification. Instead of a hundred different protein enzymes – a huge cost in terms of energy – the cell simply produces 100 short RNAs which determine the specificity of where the very same enzyme can catalyze identical reaction. Nature can be very smart!

Let's talk about epigenetics.

Chromatin regulation is managed not so much by miRNAs, regarded as post-transcriptional regulators (although some papers do indicate that this happens), but by other class of small RNAs, in particular siRNAs; they are formed through a mechanism similar to that forming miRNAs. Chromatin regulation by small RNAs is still poorly understood in mammals; however, the process has been studied in depth in plants and certain fungi. Chromatin structure is very dynamic, and the regulation of its dynamics is quite complex. Until recently it was believed that histones simply organize DNA into structures which can be packed into chromosomes. We now know that histones are the subject of dozens of modifications such as methylation, acetylation or ubiquitination. These modifications determine the availability of chromatin to RNA polymerases – the possibility of transcription and formation of specific proteins. Small RNAs – siRNAs – play an extremely important role, locally converting active euchromatin into (generally) inactive heterochromatin. This mechanism is very well understood in the fission yeast Schizosaccharomyces pombe, *particularly for the heterochromatin located near centromeres. This central part of chromosomes is very tightly packed. During mitosis, when chromosomes separate, it must be heterochromatic. The centromeric regions are transcribed in opposite directions, and a double-stranded RNA is formed from the non-coding transcripts. It is then cleaved by the Dicer nuclease into small siRNAs. These siRNAs then guide into place the entire complex of other proteins, which stimulate heterochromatization and lead to the methylation of histones on residues responsible for gene silencing.* S. pombe, *evolutionarily distinct from baker's yeast, has an interesting cellular cycle. Division occurs along a short axis (as in the majority of bacteria), producing two cells of identical size. In* S. pombe*'s cellular cycle, siRNAs are involved not just in the regulation of centromeres, but also many protein-coding genes. Transcription of many genes occurs in two directions – sense and antisense – to create the double-stranded RNA which is immediately recognized by Dicer. This acts on a feedback basis: gene silencing is brief, occurring only during a specific phase of the cell cycle. These are examples of very exquisite types of regulations, ones hard to even imagine a decade ago.*



In the last 20 years, the process of folding mRNA precursors has been one of biology's most fascinating problems. Researchers have been striving to understand why the human genome contains similar numbers of genes to that of the nematode worm *C. elegans*, shown here, yet codes for a much more complex organism



www.digital-delight.ch

The brain has the highest diversity of miRNA, which is why so many researchers are studying neurons. Local neuron stimulation is essential in long-term memory processes

But let's get back to miRNAs. Their role in post-transcription gene regulation has been studied for around a decade. In humans, there are around a thousand different miRNAs; it is estimated that they regulate over half of all genes encoding proteins, although this has been directly demonstrated for only a few percent of regulated mRNAs. miRNAs function as ribonucleoproteins, and their role is to guide the correct proteins to mRNA. As such, miRNAs operate as guides leading the ribonucleoprotein complex to mRNA by base pairing, association of complementary sequences present in miRNA and mRNA. A very simple and efficient rule. Again, by using miRNA guides the entire machinery leading to silencing of mRNA can be specifically brought to selected targets. RNA has this ability to bind to other RNA or DNA molecules through base pairing – something proteins cannot do. This is why miRNAs evolve so quickly. Based on research that has mainly been conducted in plants, there exist miRNA classes that are still evolutionarily "immature". A small hairpin-shaped RNA molecule (approx. 100 nucleotides long) is a sufficient structure to be processed to miRNA. There are tens of thousands of such hairpin structures resulting from the folding of cellular RNA. Mutations can result in a hairpin to become an efficient substrate for enzymes forming miRNAs and to kick off evolution in a new direction. In Chlamydomonas and C. elegans, there are just around a hundred such "mature" miRNA-producing hairpins. But their new forms evolve in increasingly complex organisms; perhaps this is the reason why so many of the approximately thousand active miRNAs in mammals function in neurons.

Many brain processes are regulated by RNA.

The complexity of brain processes requires extremely sophisticated regulation. For example, axons must find the correct route during development in order to join other neurons. The brain has the highest diversity of miRNAs, which is why so many RNA researchers are studying neurons. Nerve cells have various cellular compartments. Local neuronal stimulation is essential in long-term memory processes. miRNAs frequently act as reversible inhibitors: miRNA attaches to a mRNA molecule, blocking it; when mRNA detaches itself, it regains activity. MiRNAs represent currently the best candidates for regulating translation processes in dendrites. There are data indicating that it is in the neurons, at the synapses, that miRNA molecules block translation, and stimulation of the synapse results in a temporary dissociation of miRNA or its degradation, leading to the activation of the translation process. The most recent research carried out by a post-doctoral fellow in our laboratory, Jacek Król,, shows that miRNAs in neurons have a very fast catabolism. They are constantly broken down and replaced by newly synthesized ones. This is likely to be linked with the activation of translation in neurons. Perhaps miRNAs are broken down locally following activation of the synapse; a new one must then be formed to arrive with a new mRNA. In other tissues miRNAs are exceptionally stable, and neurons appear to be a hive of activity. And yet we still don't know why. There are so many fascinating things to discover in neurobiology. Both small and large RNAs seem to be strongly engaged in gene regulation in neurons. Similarly epigenetics, in particular neuroepigenetics, is a field which will continue to develop in the coming decades. And there is still the question of how, from a single fertilized egg, it is possible to create any tissue from the totipotent cellular stage through chromatin regulation. No doubt, RNA will play a key role in all these processes.

Interview by Patrycja Dołowy
Warsaw, November 2011

Further reading:

Krol J., Loedige I., Filipowicz W. (2010). Regulation of miRNA Biogenesis, Function and Decay. *Nature Reviews Genetics* 11, 597-615.