
MicroRNAs: heralds of the noncoding RNA revolution

AMY E. PASQUINELLI

Division of Biology, University of California San Diego, La Jolla, California 92093-0349, USA

The last decade of the 20th century witnessed the birth of a scientific revolution, as the first indications of a broad layer of gene regulation by noncoding RNAs (ncRNAs) began to take shape. The timely launch of a journal entirely dedicated to RNA research, *RNA*, thus seems prescient for the explosion of discoveries revolving around ncRNAs that would be made over the next 20 years. Perhaps at the base of the current intense interest in ncRNAs are the small RNAs that serve as guides in RNAi and miRNA silencing pathways. During the 1990's, puzzling observations followed by methodical studies in plants, worms, flies, and mold culminated in the revelation that exogenously supplied ncRNAs could silence the expression of just about any gene of interest through a process aptly named RNA interference (RNAi). Additionally, in 1993 a tiny encoded RNA called *lin-4* was reported to use partial base-pairing to repress the expression of a protein coding gene in a pathway essential for development in *Caenorhabditis elegans*. Thus, by the end of the last century it was apparent that plants and animals harbored a mechanism for using ncRNAs to silence the expression of mRNAs and that, at least in worms, there was also endogenous production of a ncRNA to achieve this type of regulation.

While in retrospect it seems obvious that there must be other RNA genes like *lin-4* that engaged the RNA silencing machinery, the next member of what would come to be known as the microRNA (miRNA) family of ncRNAs did not emerge until the turn of the century. This RNA, again discovered through genetic studies of development in worms, bears the ominous name of *let-7* based on the lethal phenotype exhibited by worms with mutations in this ncRNA gene. Serendipitously, in late 1999 the recent public availability of vast data from the human and *Drosophila* sequencing projects enabled Gary Ruvkun to discover potential homologs of the worm *let-7* RNA. A key feature of the *let-7* RNA is that the entire 22 nucleotide (nt) sequence is perfectly conserved between worms and humans. Not only did this trait make it possible to uncover potential homologs with such limited sequence information, but it also allowed the existence of such RNAs to be confirmed by stringent Northern blotting methods. This experimental strategy would have

failed for most other miRNAs, which typically exhibit inexact nucleotide identity across species.

The revelation that a gene like *let-7* existed in many animal species fueled the race to discover more small ncRNA genes of this ilk. Soon hundreds of new small RNAs were identified by biochemical and computational methods in plants and animals. A unifying feature of these ncRNAs was their approximate length of 22 nt, and, thus, they were appropriately christened microRNAs.

The establishment of an entirely new class of RNAs inspired many RNA researchers to turn their attention not only to identifying more miRNAs but also to elucidating how they are expressed and function. In less than a decade after the foundation of the miRNA field of study, the basic molecular mechanisms behind miRNA production and regulation of target gene expression have been detailed. Across species, most miRNA genes are initially transcribed into primary transcripts that are hundreds to thousands of nucleotides long. From these nascent transcripts, the hairpin precursor form of the miRNA is excised and then serves as the substrate for a final processing step to release the ~22 nt mature miRNA, which is incorporated into a protein complex often referred to as the miRNA induced silencing complex (miRISC). Within this complex, miRNAs guide Argonaute and associated proteins to complementary sequences in mRNAs. In plants, pairing between the miRNA and target site is often nearly perfect, whereas in animals it is almost always only partial. The reason for this discrepancy in the pairing behavior between kingdoms is a mystery. Nearly perfect complementarity between the miRNA and its target site can induce Argonaute to cleave the bound mRNA. In cases where target recognition involves looser pairing with the miRNA, miRISC triggers mRNA degradation through deadenylation and decapping events. Translational repression can precede mRNA decay, although this is rarely the predominant mechanism used to block gene expression of most endogenous miRNA targets studied to date.

In many organisms, including humans, a large fraction of the protein-coding genes are subject to regulation by miRNAs. Consequently, miRNAs, as a class, are essential because deficiencies in the factors involved in miRNA

Corresponding author: apasquinelli@ucsd.edu

Article and publication date are at <http://www.rnajournal.org/cgi/doi/10.1261/rna.049981.115>. Freely available online through the RNA Open Access option.

© 2015 Pasquinelli This article, published in *RNA*, is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

biogenesis and function often have fatal consequences. Yet, loss of a single miRNA gene very often results in no discernible phenotypes in laboratory settings. This observation has been attributed to redundancy, where related miRNAs in the genome can compensate for each other, and to the possibility that in many cases miRNAs play fine-tuning roles where their effects on gene expression are too subtle for gross phenotypic outcomes to result from loss of individual miRNAs. Nonetheless, there are striking examples of altered miRNA expression in the right context causing dramatic cellular and organismal abnormalities. In addition to the original miRNAs, which were uncovered because mutations in the *lin-4* and *let-7* genes caused profound developmental defects in worms, the mis-expression of specific miRNA genes has been found to alter mammalian cell biology. For example, deletion of miR-155 in mice causes defective immune cell function, and overexpression of this miRNA has been implicated in driving oncogenesis in a variety of tissues, including breast, lung, and blood.

Understanding the biological functions of specific miRNAs in their proper contexts remains an outstanding challenge in the field. This problem is further complicated by the difficulty in assigning physiologically relevant targets to miRNAs that use limited base-pairing to achieve specificity. Rapid progress has been made in determining motifs involved in miRNA-target interactions, such as the importance of pairing of the 5' end of miRNAs to sites in mRNA 3' untranslated regions (3' UTRs), but deviations from these patterns are common enough to warrant further investigations into the rules that govern functional miRNA interactions in vivo. An important consideration that has recently come to the forefront is the stoichiometry between specific miRNAs and their possible target sites within a cell. The idea that competing endogenous RNAs (ceRNAs) contain sequences that divert miRNAs from other target sites means that not only the absolute quantity of a mature miRNA but

also the amount of miRNA molecules free to bind relevant targets must be calculated to predict the regulatory outcome of a given target. While the ceRNA hypothesis has gained momentum, opposing views have proposed that miRNAs typically far outnumber the binding sites that would be needed to substantially titrate their activity. Careful quantification of miRNA and potential ceRNA synthesis and stability will be needed to resolve this issue in specific cellular contexts.

In addition to founding an entirely new field of study, the discovery of miRNAs bore two broader scientific consequences. First, the recognition that miRNA genes reside in the genomes of diverse species initiated the hunt for other regulatory RNAs, which led to the identification of dozens of new RNA classes. Some of these RNAs, such as piwi RNAs (piRNAs), endogenous small interfering RNAs (endo-siRNAs), and tRNA-derived smRNAs (tsRNAs), function as relatively small (<50 nt) RNAs, while another general class of ncRNAs, called long ncRNAs (lncRNAs), contains members that are typically over 200 nt. Thus, since the establishment of the *RNA* journal 20 years ago there has been an explosion in the number and types of regulatory RNAs, instigating a paradigm shift in the way gene regulation is understood. Second, the pivotal role of *C. elegans* research in revealing the existence of miRNAs highlights the necessity of continued studies in model organisms. Notably, the first miRNA and its target had no apparent homologs outside of worms and, hence, garnered little attention beyond a few *C. elegans* labs. Yet, these worm genes proved to be heralds of a field that now embraces knowledge gained from miRNA studies in any organism and strives to use this information to exploit miRNAs as biomarkers and therapeutic targets in human disease. Perhaps the most important outcome of the discovery of miRNAs and the many other new ncRNAs is that it illustrates that surprises still lurk in the natural world and the possibility of uncovering one is tangible.