

---

# Two decades of miRNA biology: lessons and challenges

---

ERIC C. LAI

Department of Developmental Biology, Sloan-Kettering Institute, New York, New York 10065, USA

## Introduction

The celebration of 20 years of RNA is opportune to reflect on what has been learned during this time regarding miRNA biology, and challenges that lay ahead. I focus this perspective on genetics, which arguably laid the foundation for the miRNA field. Indeed, genetic studies revealed critical miRNA:target interactions, the fundamental logic of miRNA target recognition, and phenotypically vital usages of miRNA regulation, all prior to formal recognition of the miRNA pathway. Ironically then, deciphering miRNA biology remains as challenging as ever, despite extensive mechanistic insights gained from biochemical and structural approaches, and a wealth of data garnered from genomewide approaches. Nevertheless, accumulating knowledge on miRNA function and ever-expanding genetic resources for manipulating miRNAs and their targets promise an exciting future in understanding in vivo necessities of these tiny RNAs.

The miRNA era exploded in 2001 with the revelation that myriad hairpin-encoding loci generate ~22 nucleotide (nt) RNAs. However, miRNA studies were already well on their way at that point, through developmental analyses in invertebrates. As is evident from their low-numbered gene names, mutants of the founding miRNAs *lin-4* and *let-7* were amongst the earliest *C. elegans* mutants recovered. Notably, the stage for comprehending their cloned sequences was set by genetic inquiries, initially by Chalfie and Ambros, which revealed antagonistic control of heterochronic identity by *lin-4* and its major target *lin-14*.

Ruvkun and colleagues studied gain-of-function *lin-14* 3' UTR deletion mutants that phenocopied *lin-4* loss-of-function, and these were critical to establish the cis-regulatory domain of *lin-4* and its impact on *lin-14* activity. However, there are older precedents for which loss of miRNA-mediated target regulation has phenotypic consequences. In 1923, Bridges isolated the dominant *Drosophila* wing serration mutant *Beadex*. Its eventual cloning showed that recurrent 3' UTR disruptions underlie gain-of-function of *dLMO*. Interestingly, *mir-9a* knockouts exhibit similar wing serra-

tion, and miR-9a directly represses *dLMO*. Strikingly, removal of one *dLMO* allele fully restores *mir-9a* mutant wings, defining it as a critical miR-9a target. Thus, genetics revealed a miRNA-related phenotype nearly a century ago.

## Some personal history of miRNA genetics

I was fortunate to conduct undergraduate thesis research with Gary Ruvkun (MGH/Harvard). During this time, the incredible story unfolded of how Victor Ambros' group cloned *lin-4*, how Victor and Gary recognized the homology of *lin-4* with *lin-14* 3' UTR control sequences, and how this mediated novel RNA-based repression. During graduate studies with Jim Posakony (UCSD), I studied two dominant *Drosophila* sensory organ mutants affecting Notch signaling (*Bearded* and *E(spl)<sup>D</sup>*) that bore 3' lesions. I found their hypermorphic activities were associated with loss of repressive ~7-nt 3' UTR motifs (Brd-, GY- and K-boxes), which were conserved amongst a broader set of Notch pathway target genes. In particular, these motifs mediated transcript destabilization and deadenylation of in vivo sensor transgenes, and genomic transgenes of *Bearded* and *E(spl)m8* mutated for "box" motifs induced gain-of-function phenotypes. Although trans-acting factors for these boxes were unknown, their high degree of evolutionary sequence constraint implied a nucleic acid component. On the basis of the *lin-4/lin-14* precedent, we suggested one possibility was that these boxes might bind complementary RNAs. We even identified instances of fly "box" motifs conserved in mammalian 3' UTRs, and a novel, conserved, repeated 7-nt in worm heterochronic genes (unpublished), suggesting this particular motif length was a shared feature of animal 3' UTRs.

Meanwhile, back in Gary's lab, my friend Brenda Reinhart was struggling to clone a mutant that emerged from genetic screens by Michael Basson (Horvitz lab) and Frank Slack for modifiers of the heterochronic hierarchy. This mutant proved allelic to *let-7*, and following painstaking efforts to rule out potential protein-coding loci, she unexpectedly corresponded to me in early 1998 that *let-7* encoded the second

---

Corresponding author: laie@mskcc.org

Article and publication date are at <http://www.rnajournal.org/cgi/doi/10.1261/rna.051193.115>. Freely available online through the RNA Open Access option.

© 2015 Lai 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/>.

animal small RNA. Our conversations regarding let-7 reinforced the notion the various fly “boxes” might recruit short RNAs. Moreover, Brenda and Amy Pasquinelli later found that let-7 was broadly conserved, exciting first evidence for broader phylogenetic reach of these small RNAs.

Upon reading the landmark miRNA cloning papers from the Ambros, Bartel and Tuschl labs, I was astonished to realize that of the first 21 *Drosophila* miRNAs, 14 exhibited Watson-Crick complementarity at their 5' ends to Brd/GY/K boxes. This suggested that ~7 nt complements, preferentially involving nucleotides 2–8 of miRNAs (i.e., “seed” regions as coined by Bartel and colleagues), are critical for target recognition. This concept was broadly extended by data from many disciplines, including (1) experimental demonstrations that seed complements suffice for miRNA repression, (2) bioinformatic observations that miRNA seed complements are highly over-represented amongst conserved 3' UTR motifs, and (3) structural observations that the seed is pre-organized to enhance target interaction.

### miRNA biology in the post-genomic age

Strategies for studying miRNA biology have changed since 2001. Not only are more miRNAs known, many types of evidence and reagents can guide miRNA analyses.

For each well-conserved miRNA, there is an assortment of target predictions. Strikingly, most conserved miRNAs have accumulated large cohorts of conserved seed-complementary sites. This seems remarkable, and did not have to be this way. For example, it might have turned out that most transcripts evolved to avoid seed matches, leading to a deficit instead of the observed excess of conserved seed matches.

There are also genomewide methods to measure the impact of miRNAs on gene expression. Although the ultimate readout is to alter protein production, miRNAs can reduce steady-state transcript levels. Accordingly, transcriptome measurements provide broad insights into the regulatory networks of miRNAs. Such studies confirm the notion that the seed is a predominant, although not exclusive, determinant of target specificity.

Finally, there exists abundant expression data, which help prioritize miRNAs relevant to particular settings. These include microarray/sequencing analysis of diverse cell-types, tissues, and manipulations/stresses/disease conditions. There is also information from in situ hybridization, promoter-reporter transgenes, and miRNA sensor constructs, which provide cell-by-cell data on miRNA transcription and/or activity.

These methods are part of the common framework that researchers apply when studying miRNAs. So, why is it still so difficult to deconvolve in vivo requirements for miRNAs? Two major challenges include difficulties in recognizing phenotypes incurred by miRNA loss, and associating causal targets to miRNA phenotypes.

### Understanding the ‘modest’ phenotypic effects of most miRNA knockouts

Thanks to persistent efforts, there exist deletions for most miRNAs in *C. elegans* (predominantly from a joint effort of the Ambros, Bartel and Horvitz labs) and *D. melanogaster* (predominantly from the Cohen lab). As well, some dozens of mouse miRNA knockouts exist. A sobering conclusion of these collections is the paucity of dramatic phenotypes. For example, although some of the earliest mutants in worms (*let-7*) and flies (*bantam*) were lethal, there were few lethals amongst the systematic miRNA knockout projects, and few with overt developmental abnormalities. By comparison, 20%–30% of protein-coding genes in these invertebrates are essential. In mice, with only rare exceptions, most miRNA knockouts are viable, fertile, and seemingly largely normal.

Of course, the literature abounds with miRNA mutants with compelling developmental, physiological, and/or behavioral phenotypes. miRNAs are indeed “useful.” But truthfully, most of these phenotypes are subtle when compared to those routinely obtained with protein-coding mutants. Can we rationalize these modest effects? Molecular redundancy of miRNAs is documented, but seems not a satisfactory explanation. After all, some of the most striking miRNA mutant phenotypes occur with individuals within multicopy families (e.g., nematode *lin-4* and *let-7*, *Drosophila mir-9a*), and the Horvitz lab reported that compound deletions for 12/15 *C. elegans* miRNA families still lack phenotypes.

It seems most miRNAs are not crucial for broad developmental decisions or core physiologies, at least under laboratory conditions. But, might miRNA mutants exhibit substantial phenotypes, only in particular cells or circumstances? A classic example is the *lsey-6* miRNA, whose mutants the Hobert lab showed fail to specify a single neuron, predominantly via de-repression of *cog-1*. However, *lsey-6* worms do not betray gross abnormalities, and one wonders how long it would take to recognize its defect via reverse genetics. The recent systematic study of *Drosophila* miRNA knockouts by the Cohen lab makes the provocative conclusion that despite their seemingly wildtype appearance, 80% of mutants exhibit at least one confident defect across a panel of phenotypic surveys. Most of these are rather general defects whose cellular basis is unknown, but they reinforce the idea that most miRNA mutants are not actually “normal”.

Alternatively, the purview of miRNAs might not be the core establishment of regulatory networks, but in maintaining their robustness. A formative study by Carthew and colleagues reported that loss of *Drosophila mir-7* has minor consequences on eye development, which were strongly exacerbated when rearing under temperature fluctuation. A corollary is that genetic interventions might be used to sensitize miRNA mutant backgrounds. A striking illustration of this was the finding by the Abbott lab that 25/31 worm miRNA knockouts yielded phenotypes in combination with alteration of other core regulatory factors.

I am certain that closer study of more miRNA knockouts will unveil more phenotypes in the future, information that can be used to leverage other types of genomic and experimental data. In particular, efforts such as sensitizing the genetic background of miRNA mutants, subjecting them to environmental stresses, and performing forward screening using quantitative readouts, especially in metabolic or behavioral assays, may prove enlightening.

### Opposing models: ‘fine-tuning of a target network’ versus ‘key miRNA targets’

A spirited discussion regards whether miRNAs fine-tune large networks of targets, or predominantly repress specific major targets. Both notions are reasonable interpretations of certain data, but problems arise when conflating correlation with causation. The fact that miRNAs subtly alter large target cohorts is the basis of the network model. However, it is an act of faith to assume a phenotype is mediated by modest derepression of a network. It is currently impossible to test this by coordinately manipulating 10s–100s of genes, in a way that mimics their deregulation under miRNA mutant conditions.

The genetic approach is reductionist, and is only possible when miRNAs exhibit epistatic relationships with other genes. Such was the case with the founding miRNA:target pairs *lin-4:lin-14* and *let-7:lin-41*, but it now applies to many additional examples. However, what is the generality of the “critical targets” model? And has this fueled some researchers to over-attribute the consequences of target regulation? It is easy to find literature where a reported target plausibly explains a miRNA-related phenotype, were it not for the fact that quantitative target perturbation by the miRNA is very modest, sometimes <20%.

Advances will come with testing target site mutants. If a specific target truly underlies miRNA function, then mutation of binding site(s) *in cis* should phenocopy miRNA mutation *in trans*. This has traditionally been cumbersome to assess, but CRISPR/Cas9 genome engineering now provides a path. A recent study by the Fulga lab analyzed target site knockouts in zebrafish, *Drosophila*, and human cells. Some phenotypically critical sites were indeed validated, whereas others indeed mediate miRNA regulation but do not underlie the miRNA mutant phenotype. It will be fascinating to reassess more targets from the literature using this stringent approach. Arguably, from a disease perspective, such phenotypically critical miRNA targets/sites are the most interesting ones to know about.

Conceptually, it will be a greater achievement to use genome engineering to test the notion that miRNAs phenotypically fine-tune target networks. Imagine a future when one can routinely engineer animals simultaneously bearing mutations in dozens of miRNA target sites. Only then would we have a chance to prove, or disprove, the concept that a miRNA-mutant phenotype arises only when deleting sufficient numbers of sites. Alternatively, it may end up being that systematic approaches reveal specific critical targets, while 50 other target sites can be mutated concomitantly without phenotypic consequence. We can only dream of results from such experiments.

### Acknowledgments

I regret that referencing of primary literature was not permitted in this brief perspective. I thank Brenda Reinhart and Victor Ambros for critical reading and discussion. E.C.L. is supported by the Burroughs Wellcome Fund and the National Institutes of Health grants R01-GM083300 and R01-NS083833.