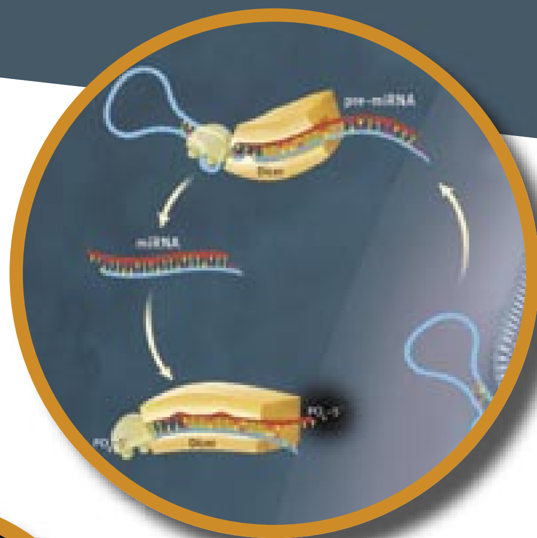
RNAi:
microRNA packs a wallop

MicroRNAs play a hitherto unsuspected, integral role in regulating gene expression. But their short length, sequence similarities, multiplicity of mRNA targets, and unpredictable pairing to those targets make studying them tricky. The good news: Methods for analyzing more familiar RNA species have been adapted for investigation of these extraordinary molecules and improved technologies are constantly becoming available. **By Bruce Goldman**



Wait! Stop. Don't toss out that small-RNA fraction just yet. It took researchers a while to break that habit. Tiny pieces of RNA were long thought to be mere degradation fragments of more important longer molecules.

But it's now obvious that the famous formulation enshrined as the Central Dogma of Molecular Biology – "DNA makes RNA makes proteins" – is only partly true. Small noncoding RNAs called microRNAs (miRNAs), typically about 21 or 22 nucleotides in length, exert a powerful posttranscriptional regulatory influence. By binding to untranslated regions of messenger RNAs (mRNAs), miRNAs stall ribosomal output of the proteins encoded by the target mRNA, fine-tuning translation and, on occasion, marking the target mRNA for degradation. Gregory Hannon at **Cold Spring Harbor Laboratory** and his colleagues recently reported findings implicating miRNA in DNA methylation, too.

From being regarded as an oddity, miRNA has proceeded to the cutting edge of biomedical research. After a sluggish start following isolation of the first miRNA, from *C. elegans* in 1993, the number of human miRNAs logged on the Sanger Institute's miRBase Sequences Database (a central repository of validated miRNAs in the public domain) has exploded in the past few years, and has now reached almost 500.

"Smart scientists that we are, we managed to miss this entire class of incredibly important regulatory molecules until the 21st century," says Eric Lader, associate director of research and development at **Qiagen**, a Germany-based supplier of nucleic acid sample-preparation materials and polymerase chain reaction (PCR) kits.

But researchers are making up for lost time. While only four papers on miRNA were published as recently as 2001, last year the number exceeded 600. "The likely impact of miRNAs in mammalian biology is enormous," says Carl Novina, of the **Dana-Farber Cancer Institute**. "They play roles in differentiation and development, chromosomal segregation and division, cell type-specific functioning, metabolism, and apoptosis." Alterations of miRNA expression, Novina continues, have been correlated with numerous cancers and, at least in mice, with diabetes, spurring the investigation of miRNA as a potential diagnostic and therapeutic tool.

One reason for the field's accelerating advance: The tools were already mostly in place. "Over the last 20 years, an extraordinarily large number of really smart molecular biologists worked out every technique you could ever desire to study small RNAs while somehow scrupulously avoiding discovering too much about them, so that people of my generation would have something to do," marvels Philip Zamore, professor of biomedical sciences and biochemistry at the **University of Massachusetts School of Medicine** in Worcester.

continued ›

IMAGES COURTESY OF KATHARINE SUTLIFF / *SCIENCE*

"We managed to miss this entire class of incredibly important regulatory molecules until the 21st century."

Look for these **Upcoming** Articles
Cell Signaling 2 — June 22
Microarray Technologies — August 24
Genomics 2 — October 19

Inclusion of companies in this article does not indicate endorsement by either AAAS or Science*, nor is it meant to imply that their products or services are superior to those of other companies.*

Those tools did require some modification. Early RNA methodologies weren't optimized for ultrashort species such as miRNA, which got tossed out with the buffer solution. Nowadays, companies selling sample-preparation tools, such as Qiagen and **Ambion** (acquired last year by **Applied Biosystems**), offer kits adapted for isolating small RNAs.

Once you've found a given miRNA, figuring out which mRNAs it targets is nontrivial, at least in animals, as only six or seven of the 21 or 22 nucleotides in an miRNA sequence (the seed region) bind to their mRNA targets in a strict antisense fashion. This "fuzzy pairing," with the seed sequence dictating much, but not all, of binding strength, makes it hard to predict mRNA targets computationally—all the more so with the recent discovery that miRNA transcripts get "edited": Enzymatic nucleotide substitutions differentiate them from their DNA sense-strand progenitors.

Finding New miRNAs

The sequencing of the human genome and partial elucidation of miRNAs' structural features made it possible to do computational searches for nucleotide stretches that could fold into characteristic hairpin structures and are conserved across species lines. Unfortunately, the vast majority of computational hits were false positives.

Rosetta Genomics, an Israeli company that presciently began life with an miRNA focus in 2000, created a sophisticated algorithm that located some 11 million hairpin structures, eliminated most of them (those that are thermodynamically unstable, structurally suspect, or located on protein-coding sequences or in repetitive elements), and then filed patents on thousands of candidate miRNAs. Ronit Aharonov, executive vice president for intellectual property and computational biology, says Rosetta has internally validated many miRNAs not yet listed on the Sanger database. In a highly cited 2005 article, Rosetta Genomics scientists claimed that a significant number of these novel miRNAs are conserved among primates but not other mammals, and may therefore have played a major evolutionary role in the higher complexity of primates, including humans.

Massively parallel oligonucleotide sequencing now appears to be outstripping computational approaches for discovering novel miRNAs and is even unveiling entirely new realms of small RNAs. In the past year alone, groups led by David Bartel of **Whitehead Institute at MIT** and by Hannon at Cold Spring Harbor have used this methodology to discover two small RNA classes, each consisting of tens of thousands of members, whose defining features differ considerably from those of miRNA and from one another's.

454 Life Sciences, previously a Curagen subsidiary, was selling its sequencing instrument along with reagents via a marketing agreement with **Roche Diagnostics**, who has now purchased the company outright. Each sequencing run yields between 200,000 and 400,000 readouts of 100 to 250 nucleotides. A newer machine by **Solexa** (now owned by **Illumina**) can deliver up to 20 million short readouts of about 27 nucleotides. Although Solexa's device may truncate slightly longer small-RNA sequences—some species run 30 bases—Hannon, who has used both companies' instruments, says it usually provides enough information to identify a novel miRNA.

While the huge number of sequence readouts per run makes the per sequence cost ultracheap, a single run is still quite expensive. "You don't need a computer to figure out 99.9 percent of all the miRNAs in a human being," comments Zamore. "You just need a big budget."

"You don't need a computer to figure out 99.9 percent of all the miRNAs in a human being. You just need a big budget."

There is a feeling among those in the field that miRNA discovery, per se, is hitting the point of vanishing returns. "The numbers of new miRNAs we're detecting are diminishing," confirms Novina.

What's Different?

Meanwhile, expression profiling of miRNA molecules, once they're found, is a growth area. Microarray techniques for profiling RNA expression, which were quite advanced by the time the study of miRNA began, merely had to be adapted for use with smaller species. Planar arrays are ideal for comparing small numbers of samples for patterns of over- or underexpression of large numbers of miRNAs.

Ambion was one of the first companies to focus on the miRNA tools market. This company and its diagnostics-and-services spinoff, **Asuragen**, have licensed Rosetta Genomics' candidate miRNAs for their microarrays, which feature capture probes representing 13,349 sequences: several hundred of them from the public domain but the vast majority from Rosetta. Scott Hunicke-Smith, vice president and general manager of Asuragen's service group, says, "You can update a database every month. But you can't update content on an array cost-effectively every month. We think most of the near-term future releases of the Sanger database will probably already be represented by these sequences." Based on its own preliminary initial validation studies (looking for expression in a small number of normal human tissues), Asuragen has already identified three to five times the number of miRNAs that show up in the Sanger database. Hunicke-Smith thinks this multiple could go to as high as 10 when more tissues are examined.

Invitrogen is another big player in the miRNA-microarray space. Invitrogen's microarray methodology makes it easy to detect differences in miRNA expression between two samples—for example tumor versus normal tissue—whose respective miRNA fractions fluoresce simultaneously at different frequencies, thus allowing comparison of samples processed under identical conditions. Invitrogen sells the microarray itself, the labeling kit, and a real-time PCR kit for validating array data, says Invitrogen's research and development lead, Christopher Adams. In May 2006 Invitrogen launched a kit for specifically amplifying small RNAs. "A lot of people in the stem cell field don't have a lot of cellular material available—in some cases as few as 100 cells," Adams notes. "Evaluating miRNA expression in those circumstances can be a real challenge." The ultimate goal, he says, is to profile a single cell without any amplification or other manipulation.

For some academic laboratories, an alternative to buying arrays is to buy the probes and have them spotted onto arrays at their institutions' own facilities. One leading probe supplier is **Exiqon**, a Danish company that manufactures "locked nucleic acids" (LNAs). Unlike RNA, whose ribose backbone is quite flexible, continued ›

Exiqon's synthetic oligonucleotide is chemically constrained, locking the probe's bases into a hybridization-friendly conformation. Whereas miRNAs' short size makes optimizing probe-melting temperatures tricky, individual LNAs' compositions can be modulated to give them very similar melting temperatures.

When a researcher doesn't need to look at the entire universe of known miRNAs, but rather wants to select a set of them for, say, exploring a diagnostic application, **Luminex** offers a new bead-based technology that can multiplex up to a thousand samples—less costly per run than microarrays—testing each against probes representing 320 miRNAs from the Sanger database. "The number of miRNAs that will turn out to be relevant for a particular condition may be 20, or 10, or five," says Keld Sorensen, director of research and development. "We expect scientists profiling miRNA for a particular medical condition to want to run large numbers of samples. If you have a thousand samples, you're going to want a multiplex system."

Unlike sequencing methodologies, hybridization techniques such as Luminex's confront the dual challenge posed by miRNAs' short sequences and the fact that they often cluster in families whose members may differ by as little as a single nucleotide. That doesn't leave much room for discrimination. To tell them apart, Luminex turned to Exiqon's capture probes, Sorensen says.

Time and Place

An miRNA's role in development practically guarantees that it will often exhibit transient or highly tissue-specific expression. Chemistries that increase the stringency of binding, as LNAs do, enhance in situ analysis, performed to locate cells or tissues where a particular miRNA is being expressed at a particular time in the developmental cycle.

If locating miRNAs is getting easier, finding out what they do there is still difficult. Attempts to learn about miRNA function usually boil down to painstaking, labor-intensive experimentation, says Bartel of Whitehead Institute. "You can use antisense reagents that will inhibit the miRNA," he says. "But because each miRNA has hundreds of targets, you'll still be at a bit of a loss as to what's the real mechanism for any effect you see." Also, while miRNA function can be transiently knocked down with standard or modified antisense or RNAi approaches, cell division dilutes such agents. Furthermore, different cell types vary in their susceptibility to transfection with these substances. "There's no good, universal, or conventional way of transfecting multiple, different cell lines," says Robert Setterquist, senior manager of research and development at Ambion. "Each of them has to be optimized. And once you get that worked out, you still may not be able to get good miRNA inhibition."

How Do miRNAs Work?

As the number of relevant miRNAs in humans approaches saturation, and as ever more of them are subjected to expression profiling and characterization within particular tissues or during various developmental stages and disease states, attention is turning to the thorny matter of target identification.

Typical computational approaches grossly overpredict the number of targets, as a random search of seven-nucleotide sequences will yield more than 500 hits, says Deepak Srivastava, director of the **Gladstone Institute of Cardiovascular Disease** in San Francisco. Srivastava hopes to soon make public a more sophisticated computational algorithm. "Only around 20 or so miRNA targets have been validated at the level of protein regulation. Virtually all of them have something in common: The target site in the mRNA is located in a stretch of RNA that is physically accessible to the miRNA." Srivastava's algorithm, in addition to searching the genome for spots amenable to binding to miRNA seed regions, also screens out DNA stretches whose transcripts would be likely to form secondary RNA structures rendering them inaccessible to a corresponding miRNA. Srivastava says this appears to significantly increase specificity, although perhaps at the expense of sensitivity.

Says Bartel, "Learning the biological function of particular miRNA/ target interactions is certainly possible. But it's a lot of work. I don't see any way around it."

Complicating things, there can be several miRNA-binding sites on a single mRNA transcript. In a study published in early 2007, Bartel's group picked a target mRNA with not one but seven predicted binding sites for miRNA, all of which sat on a section of the molecule whose deletion, via a translocation in the corresponding gene, was known to be associated with oncogenesis in humans. To see if the loss of translational regulation by miRNA was indeed responsible, Bartel and his colleagues created point mutations in each of the mRNA's seven miRNA-target sites – and observed the same tumor phenotype. "You get the same results as if you had that translocation," he says. "That tells us that at least part of what's going on in those human tumors is the loss of suppression of that gene."

Bartel, who has long been out in front of the pack, says he's accustomed to inventing his own techniques, but "companies are making it a lot easier for people who haven't been working on miRNAs for the last six years."

A strong motive in any business, naturally, is to make money. But there are other incentives in play, not least intellectual ones. "When you find out some miRNA is 100-fold downregulated in an early cancer cell," says Lader of Qiagen, "and you know one miRNA could affect the expression of 100 genes, the allure of figuring out what's going on is powerful."

Bruce Goldman, a freelance science writer, lives in San Francisco.

Featured Participants

454 Life Sciences
www.454.com

Ambion
www.ambion.com

Applied Biosystems
www.appliedbiosystems.com

Asuragen
www.asuragen.com

Cold Spring Harbor Laboratory
www.cshl.edu

Dana-Farber Cancer Institute
www.dfci.harvard.edu

Exiqon
www.exiqon.com

Gladstone Institute of Cardiovascular Disease University of California, San Francisco
www.ucsf.edu

Illumina
www.illumina.com

Invitrogen
www.invitrogen.com

Luminex
www.luminexcorp.com

MIT/Whitehead Institute
www.wi.mit.edu

Qiagen
www.qiagen.com

Roche
www.roche.com

Rosetta Genomics
www.rosettagenomics.com

Solexa
Solexa.com

University of Massachusetts, Worcester
www.umassmed.edu

Tissue Homogenizer
The benchtop Precellys 24 is dedicated to the grinding, lysis, and homogenization of biological samples. It can handle difficult samples such as microorganisms and bacteria spores; hard tissues such as teeth and bone; kidney, muscle, and hair; soil samples; plants; and more. It can load up to 24 tubes simultaneously. Protocols are flexible and easy to set. Buffer and samples are added in 2-ml tubes prefilled with specific beads, either glass, ceramic, or metal. The single-use tubes prevent cross-contamination. The high speed and specific motion guarantee homogeneous and efficient grinding for reproducible, high-quality results.
Bertin Technologies
For information +331 39 30 61 69
www.bertin.fr

Multi-blocking Oligonucleotides
The Multi-Blocking Morpholino oligonucleotides inhibit the activity of a targeted miRNA by blocking several steps of its maturation. The oligonucleotides inhibit that activity by sterically interfering with: Drosha cleavage of pri-messenger RNA, Dicer cleavage of pre-miRNA, loading of miRNA onto the RNA-induced silencing complex (RISC), and recognition of miRNA targets by the miRNA strand on RISC. The strategy is to synthesize a 31-base Morpholino oligonucleotide that is complementary to the miRNA and extends one base over the flanking sequence past the Drosha cleavage site, with the balance of the bases extending into and complementary to the loop sequence. Because the loop contains many unpaired bases, the Multi-blocking Morpholino has many single-stranded bases available for nucleate pairing and strand invasion. The duplex portion of the miRNA sequence is not perfectly paired, favoring invasion of the miRNA sequence by the Multi-blocking Morpholino.
Gene Tools
For information 541-929-7840
www.gene-tools.com

miRNA Expression Analysis
A complete suite of reagents, instruments, and protocols is dedicated to the investigation of miRNAs by quantitative reverse transcription-polymerase chain reaction (qRT-PCR). The mirVana miRNA Isolation Kit, TaqMan miRNA RT Kit, TaqMan miRNA assays, and 7900HT Fast Real-Time PCR System provide a reliable and ready-to-use approach for quantitation of miRNA expression levels from a variety of sample types. The products offer a lower barrier to getting started and reduced technical variability in experiments backed by solid technical support.
Ambion, an Applied Biosystems Business
For information 512-651-0200
www.ambion.com/catalog/workflows/miRNA

Small Interfering-RNA
FlexiTube siRNA (small interfering-RNA) enhances the flexibility of Qiagen's range of RNA interference (RNAi) solutions by providing a cost-effective option for analysis of small numbers of human or mouse genes. FlexiTube siRNAs are provided in economical 1-nmol amounts. They are designed using neural network technology based on a large set of data from siRNA experiments. The siRNA design is then checked for homology to all other sequences of the genome using an up-to-date, nonredundant sequence database and a proprietary homology analysis tool. Design features include 3' untranslated region/seed region analysis, single nucleotide polymorphism avoidance, and interferon motif avoidance.
Qiagen
For information 800-426-8157
www.qiagen.com/siRNA

Linker Oligonucleotides for miRNA Cloning
Three different adenylated linker oligonucleotides are available for miRNA library construction without the use of ATP. Linker-1 contains a Ban-1 restriction site. Linker-2 contains Ava-I and Sty-I restriction sites. Linker-3 contains EcoR-I and Msp-I restriction sites. All three linkers are modified to prevent self-ligation and can improve the cloning efficiency of miRNAs that have a 5'-phosphate and react unfavorably when attaching linkers using RNA ligase in the presence of ATP. Traditionally, RNA ligase makes use of ATP to adenylate the 5'-end of a single-stranded nucleic acid sequence. The activated adenylated oligomer is then ligated to the 3'-OH of a second single-stranded sequence. In the absence of ATP, these adenylated oligonucleotides containing a pyrophosphate linkage are substrates for T4 RNA ligase.
Integrated DNA Technologies (IDT)
For information 800-328-2661
www.idtdna.com

Eukaryotic mRNA
The capping of in vitro transcribed RNA improves the stability and in vivo translation efficiency of transfected messenger RNA (mRNA) in most eukaryotic cells. Capped mRNA is also more efficiently translated in some in vitro translation system. The ScriptCap m^{7}G Capping System builds the Cap 0 structure found on the 5'-end of most eukaryotic mRNA molecules. Based on the trifunctional vaccinia virus capping enzyme, this system includes all of the components necessary to convert RNA containing a 5'-triphosphate to Cap 0 RNA.
Epicentre Biotechnologies
For information 800-284-8474
www.epibio.com/scriptcapvce.asp

Internet Access to RNA
An array of RNA samples available through an e-commerce site include those from common cancers such as breast and colon cancer. The company plans to add thousands of RNA samples from a wide range of diseases. Each sample of total RNA is from an individual donor and supplied in 5-mg aliquots. Each sample includes clinical data and pathology report details so the clinical context of the RNA can be understood. Each sample has been quality-assured through the use of the Agilent Bioanalyzer and confirmed to have an RNA integrity number of at least 6. This ensures that the materials are suitable for all forms of gene expression analysis, from Northern blots to Affymetrix chips.
Asterand
For information 313-263-0960
solutions.asterand.com