In the beginning there was RNA, at least according to some theorists. While life may have originated with self-replicating RNA molecules, molecular biologists have long focused on DNA. There were good reasons for that bias though. Besides being widely perceived as carrying the definitive copy of life’s instructions, DNA was also much easier to work with in the laboratory.

Now, though, both of those rationalizations are crumbling. The discovery of an entire layer of gene regulation based on non-protein–coding pieces of RNA, and the simultaneous development of new tools and techniques for working with those molecules, have spurred intense interest in understanding and exploiting the RNA world. In the past few years, the field’s innovations have ranged from solving simple problems, such as preserving RNA preparations for future analysis, to opening entirely new frontiers with methods to sequence all of the RNA transcripts in a single cell. The latter method has become so straightforward that some researchers are even using it to replace transcript-profiling chips. Meanwhile, established techniques such as silencing genes with short interfering RNA (siRNA) have been overhauled for improved reliability.

**Decoded Messages**

A fundamental advance in studying RNA came in 2008, when researchers described a technique for sequencing all of the transcripts in a population of cells. The method, RNA-Seq, entails reverse-transcribing purified messenger RNA, then using next generation sequencing tools to sequence all of the resulting cDNA. The result is a complete sequence of the cells’ transcriptomes.

Sequencing entire transcriptomes is powerful, but the original versions of RNA-Seq had some limitations. “In 2009, our first RNA-Seq kit required at least a microgram of RNA, and it was preferably very high quality [RNA], otherwise you wouldn’t get good results,” says Gary Schroth, distinguished scientist at Illumina in San Diego, California.

Illumina has since refined the procedure, and now offers kits for performing RNA-Seq on as little as 100 ng of RNA, allowing researchers to sequence transcriptomes of small tissue samples. The company also offers reagents for removing ribosomal RNA, one of the major contaminants that complicated first-generation RNA-Seq.

In an evolution of the technique called Smart-Seq, researchers have even managed to sequence the transcriptomes of individual cells. “The notion of being able to do single-cell [RNA sequencing] is something that a few years ago I would’ve said ‘no way,’ but it turns out that it’s actually not that difficult to go down to those levels … when you have an isolated, free-floating cell,” says Schroth. Smart-Seq doesn’t work on crude tissue preparations, though, so RNA-Seq remains the preferred method for studying those samples.

As transcriptome sequencing continues developing and sequencing costs continue falling, many biologists are now using RNA-Seq for routine transcriptome profiling, a job previously reserved for RNA profiling chips. The chips work by hybridizing RNA to an array of short oligonucleotides to determine which transcripts are present. While that approach has been a workhorse of transcriptome analysis for years, it can only identify the RNA sequences the chipmaker predicted a cell would produce. Because RNA-Seq is unbiased, it often reveals alternatively spliced RNA forms and novel transcripts that don’t show up on chips, and allows investigators to dig more deeply into the data. Schroth, whose company makes equipment for both types of analysis, argues that the chips may still have an advantage for some researchers processing large numbers of samples.

Regardless of the transcript profiling strategies they use, biologists are now keenly aware that transcription does not guarantee that a gene product is expressed. In particular, the RNA interference (RNAi) system produces numerous short RNA pieces that bind expressed transcripts and target them for destruction. Researchers initially sought to study this system by sequencing the cell’s population of short RNAs and quantifying the relative abundance of different sequences. That proved to be tricky.

“What we found doing some in vitro experiments is that some cell populations of small RNAs just aren’t represented as well in sequencing libraries as others,” says Brett Robb, senior scientist at New England Biolabs (NEB) in Ipswich, Massachusetts. In particular, many animal and plant cells modify the 3’ ends of their small RNAs, and standard sequencing...
methods underrepresent these modified molecules. To solve that, Robb and his colleagues optimized a ligation-based technique that sticks a specially modified DNA adapter onto the small RNAs before sequencing.

As scientists continue looking more deeply at posttranscriptional gene regulation, they’re discovering additional species of RNA. “A lot of the things we’ve learned as we’ve been studying small RNAs will be pretty useful for some of the newer things like these long noncoding RNAs (lncRNAs) that have been pretty hot recently,” says Robb.

The lncRNAs are pieces of RNA over 200 nucleotides long that don’t encode proteins but instead appear to regulate transcription and translation in multiple ways. In large-scale sequencing projects, scientists have estimated that the human genome encodes at least tens of thousands of lncRNAs, suggesting that these molecules represent another major layer of gene regulation.

The lncRNAs are double stranded and can be encoded by either strand of the genomic DNA. That posed a major problem for early lncRNA researchers, who often had difficulty figuring out which DNA strand a particular lncRNA came from. In response, NEB has now developed a kit called NEBNext Ultra to produce strand-specific sequencing libraries.

**A Moment of Silence**

In addition to sequencing and studying noncoding RNAs, researchers have also been using them as tools to probe gene functions. Knocking down a target protein’s expression transiently with synthetic siRNA oligonucleotides has become a standard laboratory technique, and drug developers continue to explore ways to use the RNAi machinery to treat diseases.

Indeed, RNAi-based therapies have followed the same cycle as many other technologies adopted by the biopharmaceutical industry. “When RNAi was discovered of course there was a lot of excitement, and then it’s [been] the rise and the fall and then the rise of RNAi,” says Nitin Puri, senior product manager for RNAi technologies at Life Technologies in Carlsbad, California.

The initial promise of being able to transiently shut down expression of any gene soon gave way to reality, as researchers discovered that their synthetic siRNAs often targeted multiple transcripts, and often lacked the potency of more traditional drugs. Since that initial letdown, the field has gradually recovered and investigators have begun to address some of those problems. “We see a lot more interest [in siRNA] now with researchers in the last two to three years, because there’s this realization of how this technology is helping researchers really understand and get more insights into their high throughput screening,” says Puri.

Using improved oligonucleotide design algorithms and chemical modifications, for example, Life Technologies can now synthesize potent and specific siRNAs against individual or multiple genes. By performing high throughput screens with these new siRNAs, scientists can quickly narrow the list of gene products that might be responsible for a particular phenotype.

Even a good high throughput screen still generates a lot of artifacts, though, often highlighting numerous genes that aren’t directly related to the phenotype the investigator wants to study. New strategies for searching through the data may help. “We are working with researchers who are developing new approaches, especially bioinformatic approaches, to weed out some of the nonessential genes,” says Puri.

Researchers have also learned to keep their expectations for siRNA realistic. “Those who’ve been in the field for awhile and using RNAi, they’ve come to understand and be much more accepting of its natural limitations,” says Louise Baskin, senior product manager at Thermo Fisher in Waltham, Massachusetts, adding that “it’s a promiscuous mechanism by nature—we do what we can to make it very specific, but there’s always a chance that something’s happening that’s unanticipated.”

Even highly specific siRNAs against important gene products can fail to produce a phenotype, for reasons biologists know all too well. “One of the frustrations with doing gene-by-gene knockdown is that our cells are very smart, they have secondary pathways and redundant mechanisms to save themselves,” says Baskin. To maximize a siRNA screen’s chances of success, she suggests using panels of siRNAs against multiple gene targets in a pathway instead of just one. Thermo Fisher supports that strategy with premade libraries of siRNAs to target any known gene in humans, rats, or mice, and the company recently added libraries of siRNAs against long noncoding RNAs as well.

For researchers who want to work with primary cells or intact tissues, just getting siRNAs to their targets can also be a problem, as these cells often...
### Featured Participants

<table>
<thead>
<tr>
<th>Institute</th>
<th>Website</th>
</tr>
</thead>
<tbody>
<tr>
<td>Broad Institute</td>
<td><a href="http://www.broadinstitute.com">www.broadinstitute.com</a></td>
</tr>
<tr>
<td>Sigma-Aldrich</td>
<td><a href="http://www.sigmaaldrich.com">www.sigmaaldrich.com</a></td>
</tr>
<tr>
<td>Illumina</td>
<td><a href="http://www.illumina.com">www.illumina.com</a></td>
</tr>
<tr>
<td>Thermo Fisher</td>
<td><a href="http://www.thermoscientificbio.com">www.thermoscientificbio.com</a></td>
</tr>
<tr>
<td>Life Technologies</td>
<td><a href="http://www.lifetechnologies.com">www.lifetechnologies.com</a></td>
</tr>
<tr>
<td>TransOMIC</td>
<td><a href="http://www.transomic.com">www.transomic.com</a></td>
</tr>
<tr>
<td>New England Biolabs</td>
<td><a href="http://www.neb.com">www.neb.com</a></td>
</tr>
</tbody>
</table>

---

As the technologies for studying noncoding RNA continue to develop, experts in the field see a promising future for the RNA world.

### A Quick Knockout

While synthetic siRNAs provide a convenient way to knock down gene expression transiently, researchers looking for a more lasting effect usually turn to DNA vectors encoding short hairpin RNAs (shRNA). Cells transformed or transfected with one of these vectors transcribe the shRNA in the nucleus, where it is processed through the natural RNAi machinery to form siRNA against a specific transcript. The vector persists in the cell, permanently silencing the target gene product.

By pooling a set of shRNA vectors targeting different transcripts, investigators can produce a population of cells in which each cell has a different gene silenced. Isolating the cells that display a desired phenotype, then sequencing the vectors they’re carrying, provides a quick map of the genes that are likely involved in that phenotype. This pooled approach has drastically reduced the cost and complexity of high throughput shRNA screening. “Instead of taking one shRNA into one well of cells and looking at the results, you put many shRNAs, even thousands, into one plate of cells at a time, and if you plan and carry out your experiment carefully ... you can rapidly screen lots of genes at a much lower cost and much faster,” says Shawn Shafer, functional genomics market segment manager at Sigma Life Sciences in St. Louis, Missouri.

To aid those analyses, scientists in the Broad Institute’s RNAi Consortium have produced libraries of shRNAs targeting 15,000 human and 15,000 mouse genes. The shRNAs are packaged in lentiviral vectors that can transfect a wide range of cells. Both Sigma and Thermo Fisher now offer these libraries to researchers. The consortium is now trying to create at least two highly potent, validated RNAi-based inhibitors for each human and mouse gene, as well as inhibitors for long noncoding RNAs. Those reagents should make high throughput genetic screens even easier.

As the tools improve, researchers are uncovering new complications. Recent data on the mechanism of RNAi suggest that a small “seed sequence” of only seven nucleotides may determine the specificity of small RNAs. That could help explain the promiscuous effects of many siRNA and shRNA inhibitors; a seven-base sequence can potentially bind many more transcripts than a 22-base sequence. “What they’re doing now with pooled screens is requiring two or three shRNAs against the same transcript to pop up as hits, to sort of serve as revalidation of that initial hit,” says Shafer.

Sigma also offers kits to apply the same approach to siRNA experiments. In the company’s EasyRNA protocol, an experimenter can amplify a segment of a transcript and turn it into a pool of siRNAs covering the entire segment. Targeting the transcript with multiple, custom-made siRNAs may help maximize silencing while minimizing off-target effects.

Designing highly potent shRNAs for long-term gene silencing has been a tougher problem. Originally, biologists simply borrowed the algorithms they’d used to design synthetic siRNAs and applied them to shRNA vectors. That didn’t work very well. “If you take a [siRNA] sequence and you put it in and try to express it in a vector, you get a lot of variability in function,” says Andy Crouse, senior product manager at TransOMIC in Huntsville, Alabama.

The difference in performance between a siRNA and a shRNA with the same sequence probably stems from their very different pathways through the cell. Crouse explains that while shRNAs are transcribed and processed by the same cellular machinery that runs the natural RNAi system, synthetic siRNAs bypass much of that process and bind directly to their target transcripts.

For awhile, scientists worked around the problem by simply using multiple shRNA vectors against each transcript they wanted to silence, in the hope that one or a combination of them would work well enough. The ability to do pooled shRNA screens eventually allowed researchers to isolate the most effective shRNAs from a large set. Analyzing the pool of successful shRNAs led to a new algorithm that can accurately predict the best shRNA sequence for silencing a given target. TransOMIC now uses that algorithm to design custom sets of shRNAs, and also prebuilt shRNA libraries that target particular families of genes.

As the technologies for studying noncoding RNA continue to develop, experts in the field see a promising future for the RNA world. “Human complexity cannot be accounted for by our proportionally small number of protein-coding genes ... the complexity that we have is in the percentage of our genome that is transcribed into RNA but never actually makes a protein,” says Thermo’s Baskin. She adds that “it’s really infinite possibilities in terms of how to study and what those might really mean to our understanding of biological systems.”

---

*Alan Dove is a science writer and editor based in Massachusetts.*

DOI: 10.1126/science.opms.p1400081