



## Shotgunning the messenger: Single-cell RNA sequencing

New tools are helping scientists sequence and study RNA in unprecedented detail, but each technique has its own strengths and limitations. **By Alan Dove**

**A**t the turn of the 21st century, plummeting costs and rapidly advancing technology sent many scientists rushing to sequence DNA. The field of genomics blossomed, and organisms of all types had their complete genome sequences published. While sequencing genomes has now become routine, the resulting flood of DNA data hasn't proved adequate to fully explain many of the phenomena biologists study.

How does a single genome give rise to all the cells in a complex organism? What genes do cancer cells turn on and off to escape the normal checks on their growth? How do lymphocytes respond to pathogens, and how do pathogens escape that response? Answering those questions requires drilling deeper into cellular information flows by sequencing and tracking the changes of the RNA transcripts produced from genomes.

But RNA is harder to study than DNA: It degrades easily and must be reverse-transcribed to accommodate most sequencing techniques. The transcripts of different genes also vary tremendously in abundance. Fortunately, researchers and laboratory suppliers have been steadily improving RNA sequencing tools, empowering the rapidly growing field of transcriptomics. As the technology advances, scientists are increasingly studying RNA changes within single cells, revealing entirely new levels of cellular behavior.

### In situ veritas

The story of RNA sequencing is one of steadily increasing resolution. A decade ago, most transcriptomics researchers were isolating RNA in bulk from populations of cells, then using reverse transcriptase followed by standard DNA sequencing techniques to reveal the average transcriptome of each population. Today, the field has largely moved toward single-cell RNA sequencing, using various techniques to separate and tag individual cells, then generating individual transcriptome profiles for each one. That approach has uncovered significant cell-to-cell variation in gene expression within populations that were previously considered homogeneous.

In 2014, researchers at Harvard University's **Wyss Institute** in Boston, Massachusetts, took the single-cell approach a step further, sequencing the RNA in cells in situ, to show precisely which cells were transcribing which genes in intact tissue slices and culture plates. The technique entails fixing the cells' RNA in place, then performing the reverse transcription, amplification, and sequencing steps directly on the fixed sample. Because the sequencing process uses fluorescent markers for the four bases, researchers can photograph each step under a fluorescence microscope, process the images with a computer, and map which RNA sequences occur in each cell (1).

"It's doing what a lot of the next-generation sequencers out on the market do; but instead of doing it in a flow cell, it's doing it in the sample itself," says Richie Kohman, lead senior scientist for the Synthetic Biology Platform at the Wyss Institute. **cont. >**

### Upcoming features

Microscopy: Multiphoton Microscopy—March 22 ■ Molecular Biology: Epitranscriptomics—May 17 ■ AI: Neural Networks—September 20

Kohman and his colleagues are now using the technique, called “fluorescent in situ sequencing,” or FISSEQ, in a wide range of projects. In one effort, the team uses genetically engineered viruses to assign unique tags to individual mouse neurons, then performs FISSEQ to detect viral RNAs and map the animal’s neuronal connections. Another project uses gene-editing vectors to generate cellular RNA tags that vary over time as a mouse develops from embryo to adult. Performing FISSEQ on these mice should allow the researchers to trace the lineages of every cell through development.

Though it clearly generates deeper data than sequencing RNA in separated cells, FISSEQ is harder to master. “You have to convert the RNA to DNA, and then you have to get the DNA to [form] a circle, and then you can do the amplification; so all the chemistry required to do that . . . requires a lot of work,” says Kohman. He adds that “the actual sequencing . . . has its own set of challenges, but that also has the advantage of being a somewhat solved problem” due to the growth of high-throughput sequencing technologies.

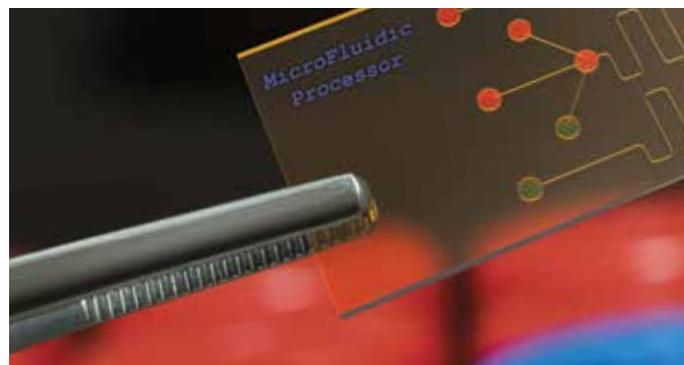
FISSEQ requires standard fluorescence microscopy equipment, plus sophisticated systems for controlling fluid flows across samples on the microscope stage. The fluid-control systems are available commercially, but their prices may be too steep for labs that don’t specialize in microscopy. The enormous datasets generated by FISSEQ also require elaborate bioinformatics tools to analyze them. To answer those needs, the team at the Wyss Institute has founded **ReadCoor**, a company that plans to offer FISSEQ products and services to researchers.

### Tiny bubbles

Many researchers don’t need the spatial detail of an in-situ technique, especially in fields such as immunology, where the cells are naturally free-floating. For these scientists, multiple vendors offer systems that make single-cell RNA sequencing highly accessible. As these companies compete, they also improve their products continuously.

**10X Genomics** in Pleasanton, California, for example, was one of the pioneers of user-friendly single-cell RNA sequencing. To use the company’s system, researchers simply load a sample of suspended cells onto a proprietary microfluidic chip. The microfluidic device separates the sample into thousands of tiny droplets, each containing a single cell and a single gel bead. The beads carry unique oligonucleotide “barcodes.” These droplets then undergo standard reverse transcription and sequencing, producing a dataset in which each cell’s complete transcriptome is linked to its individual barcode. Software included with the system lets researchers browse and visualize the data, or they can export it to process it with their own algorithms.

The original 10X system focused on single-cell RNA sequencing, but newer versions add simultaneous DNA sequencing and chromatin analysis. “We added this capability to look at protein expression, where we can barcode antibodies or other proteins that may interact with the cell and read those out at the same time we are looking at [gene] expression and maybe their particular [DNA] clonotype,” says Ben Hindson, chief scientific officer and cofounder of 10X.



**The microfluidic device separates the sample into thousands of tiny droplets, each containing a single cell and a single gel bead.**

In the system’s current iteration, a given experiment can look at two features of a particular cell simultaneously, for example, tracking the expression of a protein on a lymphocyte’s surface and the same cell’s RNA transcriptome. Hindson expects future systems to allow more simultaneous analyses, extracting even more data from each cell. “The natural thing for the field as we look into the future is to get as much information as you can from a single run,” he says.

The speed of the microfluidic system allows investigators to study huge numbers of individual cells. Each chip has eight lanes that can accommodate 10,000 cells each, for a potential throughput of 80,000 cells per run. Indeed, the company has demonstrated the ability to separate, barcode, and sequence the complete transcriptomes of over a million cells in a matter of a few days.

Scientists planning to replicate such an ambitious effort, though, should be prepared for some sticker shock. Hindson explains that while the 10X system itself is priced to fit within academic capital budgets and grant proposals, the cost of sequencing every RNA molecule in every cell scales up rapidly with the size of the experiment.

Fortunately, once the cells are separated and barcoded, researchers can decide what level of sequencing their experiment needs. An immunologist who just wants to classify cells into specific categories might want fewer sequencing reads on more cells, while a molecular biologist intent on new discoveries could generate more thorough sequencing reads on fewer cells.

### Playing the odds

Other manufacturers offer single-cell sequencing tools too, and each has its own strategy for sorting and barcoding cells. Experts advise scientists who are just starting to perform RNA sequencing to compare the options carefully, as they come with different strengths and limitations.

Laboratory-supply giant **BD Biosciences** uses a simple but robust cell-barcoding approach based on a statistical phenomenon called the Poisson distribution. In the company’s Rhapsody cell-indexing system, experimenters dilute **cont.**>

## Featured participants

**10X Genomics**

www.10xgenomics.com

**BD Biosciences**

www.bdbiosciences.com/us/home

**ProteinSimple**

www.proteinsimple.com

**ReadCoor**

www.readcoor.com

**Wyss Institute**

wyss.harvard.edu

each sample to contain about 20,000 free-floating cells, then load the sample onto a specially engineered cartridge containing 200,000 microscopic wells. At a dilution of only 0.1 cell per well, the Poisson distribution means that the wells that do contain cells are overwhelmingly likely to contain a single cell. The researcher then loads the cartridge with oligonucleotide-tagged beads, each carrying a unique barcode with a polyadenylation tag to bind messenger RNA.

"The wells are designed such that [they] can really only hold a single cell and a single bead, so any of the excess beads are washed out of the cartridge with a simple wash," explains Stephen Kulisch, senior director for applied markets at BD in Franklin Lakes, New Jersey. An optional imaging system allows investigators to check the efficiency of the process. "You can image the array and get some really good statistics on multiplets and how many cells you captured and how many bead-cell combos you have," says Kulisch.

Once individual cells are suitably paired with barcoded beads, the system follows the same general path as other RNA-sequencing methods—lysing the cells, reverse-transcribing the RNA, then sequencing the resulting DNA. Because the DNA molecules bind the polyadenylated barcodes with which they shared a well, the system's software can trace each sequence back to the cell that produced it.

Like 10X, BD is also working to extract even more data from each cell. The company now offers a catalog of antibodies tagged with their own molecular barcodes. By incubating cells with these antibodies before loading them into the Rhapsody cartridge, scientists can obtain complete transcriptome sequences and protein-expression data from the same cell.

While fluorescence-activated cell sorters also enable single-cell protein tagging, the BD system drastically expands the number of tags one can use in an experiment. "The advantage of this technology over flow [cytometry] is that you can do really high numbers of targets," says Kulisch, adding that "we've already demonstrated internally 40 simultaneous proteins, but partners are doing up to 100 simultaneous proteins as well." Flow cytometers are limited to the number of colors of fluorescence they can track, typically only a few per run.

For labs that don't already have cell sorters, barcoded antibodies are also much easier to use. Kulisch estimates that the complete Rhapsody system can be installed on a bench and running within a few hours of delivery, while systems without the imaging option are "about as easy as a gel rig" to set up.

Like the 10X system, Rhapsody is priced to fit individual lab budgets, with ongoing costs driven primarily by the expense of sequencing.

**Go Western, young scientist**

Validating single-cell RNA sequencing by checking protein levels is quickly becoming standard practice. "We see a lot of our customers are using 10X instruments . . . or things like that, and now they are basically taking the next step to get protein validation, which is being requested by reviewers and publications," says Kelly Gardner, a director of marketing at **ProteinSimple** in San Jose, California.

The product Gardner sells, called Milo, satisfies even the pickiest third reviewer by taking protein tracking to its logical extreme: single-cell Western blots. The system consists of a glass slide with a thin polyacrylamide gel on it. Over 6,000 microwells dot the gel. Similar to the BD platform and others, Milo uses a Poisson distribution and controlled loading to distribute the sample. "Many of the wells are empty, and then some of the wells have a single cell," says Gardner.

Once seeded with cells, the slide goes into the benchtop Milo instrument, which lyses the cells, performs gel electrophoresis on the proteins, then uses a proprietary, ultraviolet-activated compound in the gel to crosslink the proteins in place. That eliminates the inherent protein losses of the transfer step of traditional Western blots, and achieves the sensitivity necessary to probe proteins in single cells. After that, the procedure continues like a standard protein blot, with a primary antibody followed by a fluorescently tagged secondary antibody. A microarray scanner reads the results.

"You end up with an array of single-cell separations; we then have a software package called Scout that can take all those images and detect peaks and quantify the peak area in each single cell," says Gardner. The current system cannot provide simultaneous RNA sequence information from the same cells, but Gardner says the company is working on that.

In the meantime, processed Milo slides can be stored for up to nine months, allowing researchers to split their samples, perform RNA sequencing and single cell Westerns in parallel, then compare the results. The most common use for the Milo system is for researchers who are already doing single-cell RNA sequencing to discover subsets of transcripts whose expression they want to validate.

Like other equipment makers, ProteinSimple tries to make their products accessible to investigators who need them. User-friendliness is also a major focus, and Gardner says, "We have even nonscientists get started, and they're running assays on their first try."

Regardless of their approach, experts in the field agree that single-cell RNA analysis is driving a huge wave of discoveries. "It seems like every week there are new papers coming out in big journals . . . where they have this novel resolution that gives a fresh perspective," says Hindson.

**Reference**

1. Lee *et al.*, *Science* **343**, 1360–1363 (2014).

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



### Cloning Kit

Gibson Assembly Cloning allows for the insertion of one or more DNA fragments into virtually any location of any vector without the need for compatible restriction sites. In a single round of cloning, join

multiple DNA fragments (1-15) to create seamless constructs without subcloning. With easy reaction setup and less than

---