Two Fantastic Recruiting Opportunities!

POSTDOC CAREERS | August 28, 2015

Be sure to promote your openings to the thousands of scientists who read Science to find out about the latest postdoc opportunities.

Reserve space by August 11, 2015.

FACULTY CAREERS | September 18, 2015

Gear up to recruit for the faculty positions at your university with this much anticipated issue that reaches thousands of Ph.D. scientists looking for positions in academia.

Reserve space by September 1, 2015.
Join the Conversation!

Twitter is a great way to connect with AAAS members and staff about the issues that matter to you most. Be a part of the discussion while staying up-to-date on the latest news and information about your personal member benefits.

Follow us @AAASmember and join the conversation with #AAAS
SCIENCE & DIPLOMACY provides an open access forum for rigorous thought, analysis, and insight to serve stakeholders who develop, implement, or teach all aspects of science and diplomacy.

www.SCIENCEDIPLOMACY.ORG

Register for free to receive regular updates:
www.science diplomacy.org/user/register

@SciDip  www.facebook.com/sciencediplomacy

Senior Advisory Board

Norman P. Neureiter (Chair), AAAS
Peter C. Agre, Johns Hopkins  Sumaya bint El Hasan, Jordan’s RSS
David C. Clary, Oxford  Richard N. Foster, Yale
Steve Clemons, The Atlantic  David A. Hamburg, AAAS
Paula J. Dobriansky, Harvard  Mohamed Hassan, IAP
Esther Dyson, EDventure  Kenneth H. Keller, SAIS

Science & Diplomacy is published by the Center for Science Diplomacy of the American Association for the Advancement of Science (AAAS), the world’s largest general scientific society.
Transcriptomics today: Microarrays, RNA-seq, and more

Gene expression analysis has become routine, thanks to microarrays and now RNA sequencing, which is helping researchers discover novel RNA forms and variants. New technologies promise to reveal even more about RNA and make RNA-based clinical assays common. By Chris Tachibana

When Paul McGettigan needs to characterize a collection of cow uteruses, he begins with transcriptomics—an RNA-level survey of the samples. McGettigan, a research fellow at the School of Agriculture & Food Science, University College Dublin, studies bovine reproductive biology. He starts his projects with an overview of expressed genes to guide subsequent analyses by proteomics, metabolomics, and other methods. “Transcriptomics is the most informative assay to start with,” he says.

Transcriptomics is now easy, cheap, and fast enough to be the kickoff rather than the goal of a project. RNA analysis was once limited to tracking individual transcripts by Northern blots or quantitative PCR. High throughput transcriptomics became possible with microarrays, which detect nucleic acids in a sample by hybridization to probes on microchips. Microarrays are particularly useful for analyzing large mammalian transcriptomes, for example in drug development and clinical research that requires rapidly assessing specific genes in thousands of samples. However, microarrays detect only known sequences, so they can’t be used for discovery. Background hybridization and probe saturation interfere with low-level and high-level detection. Although microarray technology continues to advance, transcriptomics has expanded dramatically in the past few years because of developments in RNA sequencing (RNA-seq).

“For us, RNA-seq has almost completely supplanted microarrays,” says Craig Praul, director of expression analysis, Genomics Core Facility, Penn State University. “At our peak 10 or 15 years ago, we ran 500-plus arrays a year. Last year we did about 30.” Contributing to the decline in microarray use, says Praul, is that grant reviewers consider RNA-seq to be more cutting edge.

Driving RNA-seq: NGS

Use of RNA-seq has exploded because of next generation sequencing (NGS), which can yield readouts of billions of bases a day from a single instrument. Most NGS sequencers use a classic method, reading DNA by detecting bases as DNA polymerase adds them during replication. For RNA-seq, the typical workflow starts with extracting total RNA from a sample and removing the abundant ribosomal RNA. The next steps determine the type of RNA analyzed.

RNA-seq can qualitatively and quantitatively investigate any RNA type including messenger RNAs (mRNAs), microRNAs, small interfering RNAs, and long noncoding RNAs. RNA-seq analysis of RNA isoforms, which are transcribed from the same gene but have different structures, for example because of alternative splicing, are explaining how limited genomes produce complex phenotypes. RNA-seq aids scientists working on unusual model organisms, who use the method to assemble de novo transcriptomes for organisms without sequenced genomes. Most researchers, however, are interested in DGE—differential gene expression—changes in the levels of protein-coding mRNAs in experimental samples versus controls.

The most common DGE protocols use Illumina NGS systems. From an RNA sample, mRNA is selected by its poly-A tail and fragmented in preparation for Illumina sequencing, which produces “short reads” of hundreds of bases. The mRNA is reverse transcribed into a library of complementary DNA (cDNA) for sequencing. Sequence reads are counted and mapped to a reference transcriptome.

The sequencing part of RNA-seq is straightforward, says Gary Schroth, distinguished scientist at Illumina, which has more than 70% of the global NGS market. The main “pain points” for researchers now, he says, are upstream sample preparation and downstream data analysis. To improve the presequencing steps, Illumina offers the automated NeoPrep system. “Users put purified total RNA into a cartridge,” explains Schroth, “and get quantified and normalized libraries of consistent concentration that are ready to load onto sequencers.” The currently available NeoPrep cartridges prep mRNA for DGE, which is by far the most popular workflow, says Schroth. NeoPrep decreases variability among users, he adds, and reduces hands-on library prep time from hours to about 30 minutes. For downstream analysis, Illumina offers the cloud-based BaseSpace computing platform. BaseSpace includes a selection of apps that are generated and shared by the company or users for tasks such as read annotation or de novo transcriptome assembly.

Illumina RNA-seq for DGE is so common that Schroth
considers DGE and transcriptomics to be separate fields. The majority of RNA-seq studies look at the differential expression of mRNAs, says Schroth. That’s DGE. Studying the 99% of RNA that does not code for protein? Says Schroth, “That’s what I think of as transcriptomics.”

**RNA-seq challenges and solutions**

As RNA-seq data accumulate and researchers compare findings, standardization has become an issue. To assess RNA-seq performance across laboratories, the U.S. Food and Drug Administration coordinated the international Sequencing Quality Control (SEQC) project. The results confirm the versatility of RNA-seq for both DGE and isoform discovery. The findings also offer guidance for interpreting data, such as being cautious about using RNA-seq for absolute quantitation (http://bit.ly/1HTro9U).

Downstream computational analysis of RNA-seq data remains complex because the method generates a staggering amount of data. However, in recent years, says McGettigan, computational developments have reduced data processing from hours to minutes. In addition to commercial programs, researchers can explore free, community-built options at galaxyproject.org, a web-based platform of tools for data-intensive biomedical analyses. McGettigan also recommends the discussion forum SEQanswers (http://seqanswers.com).

Postsequencing analysis of short-read RNA-seq is also complicated by a presequencing step: RNA fragmentation. “The sequence reads don’t necessarily tell you what the full-length RNA looked like,” says Praul. “You can’t always put Humpty Dumpty back together.” Most researchers don’t worry about this, he says. “They just want a big picture—pathways that are activated in their samples or transcript counts from particular genes—before doing a deep dive using other biochemical or molecular techniques.”

But researchers lose crucial data with this approach, says Jonas Korlach, chief scientific officer of Pacific Biosciences. “You miss the full picture by looking at short reads,” he says. “It’s like looking at puzzle pieces without seeing the whole puzzle.” PacBio NGS instruments read unfragmented cDNA from one end to the other, producing long reads of up to 20 kb. PacBio RNA-seq protocols don’t include fragmentation, eliminating the Humpty-Dumpty step of reassembling reads into full-length RNAs. For this reason, PacBio instruments are the system of choice for de novo transcriptome assembly and studying RNA isoforms.

PacBio instruments are less commonly used for DGE than Illumina or Thermo Fisher Scientific Ion Torrent systems, in part because they give fewer reads. But Korlach notes that the PacBio SMRT Cell runs give sufficient reads for targeted studies of genes or gene families. What’s different, says Korlach, is that long reads reveal transcript structures.

As an example of the power of isoforms, Korlach names a University of California Davis study on the *FMR1* gene responsible for Fragile X syndrome. PacBio RNA-seq was used to discover 16 different *FMR1* RNA isoforms and show that certain isoforms correlate with “premutation carriers” who are at risk for fragile X-associated diseases. Korlach says that PacBio transcriptome sequencing has other clinical applications. Since cDNAs are read intact, the system distinguishes between RNAs that reflect a single mutation vs. multiple mutations in a gene. This can be important in monitoring cancer patients for resistance to targeted therapies.

**Driving developments: clinical applications**

The clinical potential of RNAs as disease and treatment markers is fueling advances in RNA-analysis methods. Jarrett Glasscock, chief executive officer of Cofactor Genomics, a contract research organization and service provider specializing in RNA-seq, confirms a strong industry interest in RNA. Although Cofactor has government and academic clients, Glasscock says, “About 80% of our work is with large pharma companies.” That work focuses on clinical uses of RNA-seq. “Our corporate partners,” Chief Scientific Officer Jon Armstrong says, “are mainly interested in using expression of coding RNAs to identify drug targets, disease biomarkers, and more recently, good or poor responders to a drug.”

But RNA-seq is so versatile that it is also used for exploration and discovery. For example, Cofactor is developing a kit for analyzing circular RNAs, which might function as regulators of microRNAs. “Since microRNAs regulate gene expression,” says Armstrong, “circular RNAs seem to be modulators of modulators for an additional level of control.”

Another area in which transcriptomics can contribute to both basic and applied research is “integromics”—combining genomic, epigenomic, transcriptomic, proteomic, and metabolomic data. Transcriptome researchers study the step in the central dogma between DNA and protein, putting them in an excellent position to be connectors. Cofactor’s work includes integrating transcriptomics and genomics. Glasscock says, “we’re leveraging the fact that RNA expression is downstream of changes in DNA. Integrating RNA data into a study can be very powerful in narrowing down DNA changes of interest to a small handful of candidates.”

continued>
Beyond RNA-seq

Transcriptomic systems are continually updating—which is why most laboratories use their institution’s core services or large service providers such as the Broad Institute rather than investing in their own instruments. In addition, dozens of new methods are appearing to complement, refine, or replace microarrays and RNA-seq.

**Oxford Nanopore Technologies**, for example, just released the portable MinION USB nanopore sequencer, currently available through the MinION Access Programme. Nanopore sequencing does not rely on DNA replication. Instead, as a nucleic acid strand travels through a protein or synthetic pore, base-dependent changes in an ion current across the pore detect the sequence. “Nanopore sequencing has the advantage of reading full-length molecules in realtime,” says Clive Brown, chief technology officer of Oxford Nanopore. “I believe we’re the only technology that can measure molecules directly.” This means the method is suitable for sequencing RNA without conversion to cDNA. Currently, researchers are trying out the MinION for analyzing cDNA, Brown says, and working toward direct sequencing of RNA.

Direct RNA sequencing is complicated because, as Brown says, “RNA contains lots of funky bases.” RNA is heavily modified by methylation and many other epigenetic tags, so the current Oxford Nanopore RNA-sequencing protocol reads an mRNA with an attached cDNA. The cDNA gives the accurate sequence and the mRNA indicates potentially modified bases. Since the types and functions of RNA modifications are largely unknown, nanopore sequencing could spur research in this unexplored area. Brown says Oxford Nanopore is working toward a system that provides RNA counts, sequences, and posttranscriptional modifications. “We don’t know yet which signals indicate which modifications,” he says, “but once we have a model for what the signals mean, we’ll be able to read RNA sequence and modifications directly.”

Other non-RNA-seq methods return to the principle of microarrays: hybridization. Bead-based assays, for example a system from Luminex, tag microspheres with fluorescent barcodes and probes that capture specific transcripts. Incubating the beads with samples, which can be as simple as cell lysates, grabs RNAs for identification and counting by a flow cytometer.

The NanoString system hybridizes two probes to each target transcript: a biotin-labeled capture probe and a fluorescent barcode-labeled reporter probe. Reporter probes hybridize with specific RNAs in a sample and capture probes lock them via avidin onto a static surface. The NanoString nCounter Analysis System counts the immobilized RNAs using their barcodes. Like other hybridization-based systems, NanoString is not for discovery. “NanoString is for targeted transcriptomics,” says Joe Beechem, senior vice president of research and development. “The advantage over NGS is there’s no library to make, no enzymes, no processing. Any time you do processing you introduce bias.”

Because the NanoString method does not require polymerase activity, says Beechem, it works in less-than-ideal conditions like crude lysates, plasma, or formalin-fixed paraffin-embedded (FFPE) samples, which is how clinical tissue specimens are often stored. “We can work with damaged or old samples as long as we can retrieve RNA of 100 nucleotides or so,” says Beechem. “We’ve done transcriptomics on paraffin tissue samples that were 50 years old.” Although NanoString is focused on developing clinical assays, Beechem says that the technology also serves researchers who work on unusual organisms: “carrots, oysters, things like that.” If you know transcript sequences, he says, the company can develop an assay to measure them.

An intriguing feature of the NanoString system is its ability to identify RNAs in a heterogeneous sample, even one that contains cells from different species. Design the right probes, says Beechem, and from a single sample, you can study transcript-level interactions between hosts and pathogens, hosts and microbiomes, or tumor cells and the immune cells that respond to them. NanoString is also in the integromics game, Beechem says. With collaborators at Massachusetts General Hospital, NanoString is adding protein-counting ability to their RNA-counting assays for quantitative transcriptomics and proteomics from a single sample.

Other advances in RNA analysis include developments in microfluidics to separate cells for single-cell transcriptomics. Spatially resolved transcriptomic methods determine both the sequence and the location of RNAs within tissue samples or cells. Transcripts are localized by in situ hybridization to labeled probes, by RNA-seq on RNA extracted from a series of tissue sections, or even in situ RNA-seq on a sample of fixed cells.

Innovations also continue in RNA-seq. New protocols are available to enrich for low-abundance mRNAs, target specific genes, or select ribosomes to ensure analysis of only translated RNAs. Discovery using RNA-seq is being combined with high-resolution microarray quantitation for fully genomewide expression profiling. For now, though, the most common RNA experiments are DGE analysis of target genes using RNA-seq. “It’s so straightforward,” says Gary Schroth. “These days, RNA-seq is as easy as microarrays—but generates data that are much more information-rich.”

*Chris Tachibana is a science writer based in Seattle, USA, and Copenhagen, Denmark.*

DOI: 10.1126/science.opms.p1500095
RNA-Seq Assays
The ultrasensitive Precise Assays have been designed for targeted RNA-seq experiments using single cells. Based on the patented Molecular Indexing technology, the Precise Assays examine large numbers of standard- or low-input mRNA samples from precious samples, or whenever absolute quantitation is required. The assays combine molecular and sample indexing in 96- and 384-sample formats, enabling customers to sequence up to 4,608 samples on one sequencing run without investing in new equipment or extensive training. Cellular Research is collaborating with Seven Bridges Genomics to offer customers a complete analysis pipeline, which provides an integrated cloud-based data platform for management, execution, and collaboration of the RNA-seq experiments in a robust, easy-to-use informatics platform. The robust design and data analysis pipelines offer a simple turnkey solution for targeted RNA-seq experiments.

Cellular Research
For info: 650-752-6144
www.cellular-research.com

RNA Prep Kit
A new RNA library preparation kit for next generation sequencing (NGS) workflows enables users to generate a more complete view of the transcriptome by detecting more low-abundance and unique transcripts. The KAPA Stranded RNA-Seq Kit with RiboErase uses a targeted enzymatic depletion method to remove ribosomal RNA (rRNA) and thereby improve sequencing efficiency. In RNA library preparation, even small amounts of residual rRNA can dramatically reduce sequencing efficiency and data quality. Using enzymatic depletion, the KAPA Stranded RNA-Seq Kit with RiboErase reduces rRNA at industry-leading rates of 99.98%. The kit provides an alternative to mRNA enrichment, since rRNA depletion allows for the analysis of exonic regions, precursor mRNAs, and long noncoding RNAs. The KAPA Stranded RNA-Seq Kit with RiboErase is available for use with human, mouse, and rat samples and is robust and reproducible with flexible input amounts. Kits also contain a low-bias engineered enzyme called KAPA HiFi and include a streamlined, “with-bead” protocol.

Kapa Biosystems
For info: 855-527-2246
www.kapabiosystems.com

PCR Prep Workstations
MY-PCR Prep provides the molecular biology laboratory technician with a “personal cleanroom” for use in the amplification of DNA and RNA. Between amplifications, My-PCR Prep can be irradiated with shortwave ultraviolet (UV) energy to denature potential contaminants and eliminate their ability to be amplified. The main chamber of MY-PCR Prep is constructed from a continuous piece of polycarbonate to prevent UV light from escaping the chamber during irradiation. Operator access is gained through the folding front polycarbonate sash, overlapped to eliminate gaps in the chamber during UV light irradiation. The work surface is white polypropylene making disinfection and cleaning simple. My-PCR Prep is a Class 100 vertical laminar flow workstation with timed UV light, making it an ideal benchtop personal clean zone for completing polymerase chain reactions (PCR) and RNA amplifications. My-PCR Prep is available in 24”, 32”, and 48” wide models; 110V AC.

Mystaire
For info: 877-328-3912
www.mystaire.com

Transcriptome Assays
Mouse and rat transcriptome assays 1.0 are powerful and flexible tools for measuring a broad range of expression changes across the coding and long noncoding regions of the mouse and rat transcriptomes. The Mouse Transcriptome Assay 1.0 (MTA 1.0) and Rat Transcriptome Assay 1.0 (RTA 1.0) provide a rich data set sufficient to decipher changes at multiple levels in a single experiment from a variety of sample types. With approximately ten detection probes per exon and four probes per exon-exon splice junction, these assays generate the most complete, accurate, and reproducible whole-transcriptome expression data with the flexibility to select the analysis level—gene-level, transcript-level, protein-coding, or long noncoding RNA. To get to biological insights simply, quickly, and seamlessly, the complete mouse and rat transcriptome solutions include intuitive analysis software that transforms the high-quality expression data generated by the assays into biologically meaningful information in minutes.

Affymetrix
For info: 888-362-2447
www.affymetrix.com

Nucleic Acid and Protein Purification System
Low and medium throughput laboratories now have an alternative to manual processes and spin-column methods for nucleic acid purification, thanks to a new automated system. Designed for research labs and small biotech firms looking to simplify DNA, RNA, and protein purification processes, the new KingFisher Duo Prime system builds upon the utility of the existing KingFisher Duo system to help improve reproducibility and deliver high-quality samples with less time and effort than traditional manual methods. With the KingFisher Duo, isolating DNA from 12 blood samples takes as little as 53 minutes and only requires a mere 15 minutes of hands-on time, compared to manual spin-column methods, which take up to 90 minutes of hands-on time. Thermo Scientific KingFisher systems are designed to isolate DNA, RNA, and proteins from a variety of starting materials, including cell-free body fluids, blood, bacteria, cell cultures, tissue, and plant samples.

Thermo Fisher Scientific
For info: 800-556-2323
www.thermoscientific.com/kingfisherduoprime

12 blood samples takes as little as 53 minutes and only requires a mere 15 minutes of hands-on time, compared to manual spin-column methods, which take up to 90 minutes of hands-on time. Thermo Scientific KingFisher systems are designed to isolate DNA, RNA, and proteins from a variety of starting materials, including cell-free body fluids, blood, bacteria, cell cultures, tissue, and plant samples.

Thermo Fisher Scientific
For info: 800-556-2323
www.thermoscientific.com/kingfisherduoprime
WILL YOUR RESEARCH LEAD TO BETTER LIVES FOR PATIENTS?


Find out more about the scope of the journal and submit your research today!

ScienceTranslationalMedicine.org
want new technologies?

antibodies
apoptosis
biomarkers
cancer
cytometry
data
diseases
DNA
epigenetics
genomics
immunotherapies
medicine
microbiomics
microfluidics
microscopy
neuroscience
proteomics
sequencing
toxicology
transcriptomis

watch our webinars

Learn about the latest breakthroughs, new technologies, and ground-breaking research in a variety of fields. Our expert speakers explain their quality research to you and answer questions submitted by live viewers.

VIEW NOW!
webinar.sciencemag.org

Science
AAAS

Brought to you by the Science/AAAS Custom Publishing Office

@SciMagWebinars