With a limited view, you’re simply left guessing.

With GeneChip® Exon and Gene Arrays, you don’t have to guess. Whole-transcript expression arrays allow you to measure the entire transcript, enabling you to detect gene-level expression, exon-level expression, and alternative splicing in a single experiment. Why settle for incomplete results when you can look beyond the 3’ end of a gene? Extrapolation can only take you so far—Affymetrix will take you the rest of the way.

See the real biology at www.affymetrix.com/genechip/wtexpression
Whole Transcriptome Profiling Using the SOLiD™ 3 System

Sequence-based approaches to the study of gene expression have the advantage of querying known as well as previously unknown RNAs in a sample, also termed “hypothesis-neutral” discovery. The only requirement is being able to make cDNA copies of all of the RNA present in the sample, sequencing them, mapping the sequences back to a reference genome, and deducing the structure using bioinformatic tools. Here, we describe the Whole Transcriptome Library Protocol recently released by Applied Biosystems, which allows the rapid construction of strand-specific libraries from a wide range of RNA species. The easy-to-use, sensitive method uses existing commercial products to clone RNA fragments and sequence the resulting cDNAs using the SOLiD™ 3 System. Applied Biosystems open source and freely available analysis tools also facilitate mapping sequences to a reference, counting each short read mapped to a given site, and identifying exon/exon junctions.

Experimental Considerations in Transcriptome Analysis

Successful whole transcriptome analysis depends on RNA quality and efficient, accurate RNA size fractionation, as this will dictate what sequences are generated. There are currently two approaches available for whole transcriptome libraries. One approach is to start with RNA that has been enriched for polyA RNA or RNAs that have polyA tails. Another approach is to start with total RNA. Total RNA contains all the different species of RNA molecules found in the cell (polyA RNA, ncRNA, rRNA, tRNA, etc.). Because of the abundance of the structural RNAs like rRNA and tRNA, these require depletion prior to deep sequencing. These non-target RNAs represent over 90% of total cellular RNA, depleting them from the pool of RNAs enriches the pool for target RNAs of interest. Using rRNA-depleted RNA for whole transcriptome research allows the study of all non-coding RNAs (ncRNA) in addition to coding RNAs (polyA RNA). Recent studies have shown an abundance of ncRNAs in the cell, which suggests the role they may have in the control of gene expression [1].

MATERIALS AND METHODS

RNA Isolation

Five micrograms of total RNA from the Human Brain Reference RNA (Ambion, P/N AM6050) is used as starting material for whole transcriptome library construction. If the library is to be made from polyA RNA, Poly(A)Purist™ Kits (Ambion, P/N AM1916 or AM1922) have been shown to give high-quality polyA RNA. Ribosomal RNA removal can be accomplished using the RiboMinus™ EUkaryote Kit for RNA-Seq (Invitrogen, P/N A10837-08, or A10838-08 for plants). Approximately 0.5 µg of polyA RNA or rRNA-depleted RNA is needed for making libraries.

Whole Transcriptome RNA Library Preparation

The RNA is randomly fragmented using RNase III (Ambion, P/N AM2290), and 100–200 bp fragments are isolated after gel electrophoresis. The RNA fragments are then converted to cDNA libraries in a strand-specific manner using the Whole Transcriptome Library Protocol (solid.appliedbiosystems.com). DNA ‘barcodes’ can be incorporated into the libraries to allow pooling of multiple samples on a single sequencing run if desired.

SOLiD Sequencing

The cDNA libraries are clonally amplified onto beads by emulsion PCR using standard protocols from the SOLiD System User Manual (solid.appliedbiosystems.com). These beads are enriched and deposited onto the surface of a glass slide for sequencing. Current scientific publications estimate that 40–50 million mappable RNA sequences are needed to detect the maximum number of known transcripts from a library constructed from polyA RNA. This number should be considered as the minimum number of sequences needed for a whole transcriptome experiment. The SOLiD 3 System is capable of sequencing more than 400 million individual reads in a single run.

Analysis

The sequences generated by the SOLiD 3 System can be analyzed using a number of analytical tools. Applied Biosystems has developed the Applied Biosystems Whole Transcriptome Analysis Pipeline (solidsoftwaretools.com/qf/project/transcriptome/), which will allow basic analysis such as mapping sequences to a reference, counting the number of sequences mapped to known RNAs, and identifying both known as well as novel exon/exon junctions. The data output from this pipeline is readily imported to the UC Santa Cruz genome browser for visualization of the results. This tool has been designed to allow additional downstream analysis scripts to be developed for further investigation.

RESULTS

Reproducibility and Dynamic Range

The technical reproducibility of the system was measured by comparing the sequencing results from two independent runs of HBR RNA (Figure 1). As can be seen, a Spearman Rank correlation of >0.98 is obtained and more than 22,000 transcripts are detected. Additionally, the levels of 95% of the transcripts do not vary in the two samples by more than 2.3-fold. Figure 1 shows that the entire system, from library generation to sequencing to data analysis, is highly reproducible. This high reproducibility is critical for comparing different samples and will allow fewer technical replicates to be run. A wide dynamic range is very desirable for any gene expression system. Figure 1 also displays the dynamic range of known transcripts detected in these samples, and in this case it is measured to be between 10^5 and 10^6.

Figure 1. PolyA+ Transcript Level Count Reproducibility Scatterplot. The number of known transcripts identified from HBR RNA was obtained from a library that was sequenced completely independently by two different groups. Transcripts were counted as present calls if 1 or more reads mapped in both data sets. The dynamic range of known transcripts detected in these samples is between 10^5 and 10^6. Green lines indicate 2-fold change, and the blue markers darken with higher density.
Reliable Mapping of Sequences to the Genome and to Known RefSeqs
To assess the ability of the SOLiD System to detect known transcripts, a large number of clones were sequenced and mapped back to the RefSeq database. The number of known transcripts detected was calculated as a function of the number of sequence reads required and plotted in Figure 2. The number of sequence tags per kilobase of transcript length (TPKB) required to map to a RefSeq transcript before the RefSeq is called as ‘present’ [2] dictates the fraction of RefSeqs that will be detected at any given number of sequences mapped (Figure 2). The more sequences that are required to map to a transcript per unit length, the higher the confidence of accurately measuring the amount of that transcript present in the sample.

The 50 base reads generated by the SOLiD 3 System allow the Whole Transcriptome Analysis Pipeline tool to unambiguously identify novel exons.

Genomic DNA Strand Specificity
Recent publications using high-throughput sequencing have shown that as many as 6000 known transcripts are also synthesized from the ‘antisense’ strand of the same DNA region [3]. This work suggests antisense transcription is not the exception, but is common in humans and most likely other higher organisms. Additionally, other types of ncRNAs have been shown to represent a significant fraction of the genome. These transcripts are synthesized from both strands of DNA. Because of the large number of antisense transcripts present, and the need to accurately map ncRNAs, it is important to know which strand of the DNA each RNA transcript maps to [4]. The whole transcriptome system used with the SOLiD System preserves the “strandedness” of the RNA by specific ligation of adapters to either the 5’ or 3’ ends of the RNA molecules before conversion to double-stranded cDNA (Figure 3). This differs from other methods in which the adapters are ligated to the double-stranded cDNA molecules, or methods that utilize random cDNA priming; with such methods it is not possible to know which strand of the DNA the sequences are mapping to, and therefore it is not possible to know which transcript the sequences were derived from [2].

Detection of Known and Novel Exons
Another important feature for whole transcriptome analysis is the ability to analyze all currently annotated exons, as well as novel exons arising from previously unrecognized splice sequences. To achieve this, it is necessary to have uniform coverage across all transcripts, and sufficient coverage to have confidence that the exon has been correctly identified. The large number of unique tags generated by the SOLiD 3 System—greater than 400 million per run—assures high confidence of uniform coverage across the transcriptome.

Figure 2. Fraction of known RefSeqs detected in a library prepared from the MAQC UHR RNA sample. The fraction of known RefSeqs begins to plateau as 30 million mappable tags are detected, when 40–45 million mappable sequences are used for analysis. The length of transcript mapped number of sequence tags per kilobase of transcript length (TPKB) is required to map to a RefSeq transcript. It is observed in the graph that as the number of sequence tags mapped per kilobase of RefSeq increases, a smaller fraction of total RefSeqs are detected.

Conclusion
High-throughput sequencing allows scientists to study the complexity of the RNA synthesized in complex genomes. The massively parallel short-read sequencing technology achieved by the SOLiD 3 System is ideally suited for whole transcriptome analysis. The addition of the Applied Biosystems Whole Transcriptome Analysis Kit enables the detection of known and novel RNA molecules as well as the resolution of strand specificity. The Whole Transcriptome Analysis Pipeline allows the data generated to be mapped and viewed easily. This complete system provides a powerful solution for the study of complex transcriptomes.

References
Highest quality, information-rich interaction data from Biacore™ systems deepen your understanding of molecular mechanisms and interaction pathways and enable you to add function to structure.

Select the perfect solution for your application and draw conclusions with confidence – from the company that continues to set the standard for label-free protein interaction analysis.

For further information or register to have one of our scientific experts contact you, please visit www.gelifesciences.com/biacore-science
Mr. Hedinger, how do you integrate the user in the team?

According to Andreas Hedinger, engineer and Head of the Industry Division at Leica Microsystems, team spirit needs the strengths of the individual, open doors and ambitious joint goals. These goals can only be attained through the planned integration of users, with direct access to their ideas and needs. They therefore form part of the team and take part in developments: for day-to-day cooperation in a spirit of trust and responsibility.

www.leica-microsystems.com

Living up to Life
Panic is unproductive.

Let’s get rational. Tough decisions lie ahead as budgets tighten. But discovery will endure. One name has been here for more than 50 years, providing stability, reliability, and quality to researchers like you: Bio-Rad.

From the beginning we have been a stable presence, offering you products that spark your imagination without burning your budget. Because great results should be within your means.

Proceed with caution. But proceed, nonetheless.

You don’t have to lose your shirt outfitting your lab.
Quite the opposite, in fact.
Visit www.bio-rad.com/economicalm/
for more information.
INFINITE POSSIBILITIES

Cellular Imaging & Analysis

NEB introduces SNAP-tag™ and CLIP-tag™ protein labeling systems. These innovative technologies provide simplicity and extraordinary versatility to the imaging of mammalian proteins in vivo, and to protein capture experiments in vitro. The creation of a single genetic construct generates a fusion protein which, when covalently attached to a variety of fluorophores, biotin, or beads provides a powerful tool for studying the role of proteins in living and fixed cells.

Advantages:

Versatile - Compatible systems enable dual labeling

Flexible - Multiple fluorophores allow for choice & flexibility

Innovative - A range of applications is possible with a single construct
Cutting-edge mTOR Signaling Antibodies

Unparalleled product quality, validation and technical support.

:: Innovative products from Cell Signaling Technology offer unsurpassed sensitivity, specificity and performance.

:: Extensive in-house validation means that optimization is not left up to you, the user.

:: Technical support provided by the same scientists who produce and validate the antibodies — this translates into a thorough, fast and accurate response.

for the most up-to-date product listing visit: www.cellsignal.com
Introducing the GS FLX Titanium Reagents

Length Really Matters

- Obtain sequencing read lengths of 400 to 500 bases.
- Generate more than 1 million sequencing reads per 10-hour instrument run.
- Improve performance by using GS FLX Titanium series reagents — without instrument upgrades.
- Accelerate the pace of discovery with easy-to-use analysis tools for straightforward interpretation of data and biologically meaningful results.

Performance, Results, Impact

Learn more at www.genome-sequencing.com
FROM DISCOVERY TO PUBLISHING
ENDNOTE X2 GETS YOU THERE FASTER.

Don’t waste time in the slow to publish lane. Let EndNote X2 take you there, faster. Revered by millions for its blazing online search speeds, instantaneous Cite While You Write™ technology and organizing references in a snap, EndNote X2 has all the high performance features you could possibly want in an easy-to-use bibliography software, and more.

Now, EndNote X2 can locate full text and connect it with references while you move on down the research road. Within a single window you can group references automatically, search your favorite online databases and transfer to and from your EndNote Web account, all in record time.

So now that you’re up to speed on what EndNote X2 has to offer, it’s easy to see how EndNote makes you the leader of the pack.

Download your free demo or buy online today www.endnote.com

800-722-1227 • 760-438-5526 • rs.info@thomson.com

© Copyright 2008 Thomson Reuters. EndNote is a registered trademark of Thomson Reuters. All trademarks are the property of their respective companies.
Work in tandem for absolute total protein quantitation.

Accurately identify and quantify thousands of proteins in a given sample – in a single experiment. Use the Thermo Scientific Tandem Mass Tag (TMT®) Kits with your mass spectrometry instrument for high-throughput protein biomarker discovery.

- ID and quantitate multiple samples from cells, tissues or biological fluids; efficiently label samples, including membrane and post-translationally modified proteins
- Efficient transition from method development to multiplex quantitation
- Identify disease biomarkers
- Quantitatively analyze proteins for which no antibodies are available

To learn more, visit www.thermo.com/TMT or call 815-968-0747 or 800-874-3723

A total solution for quantitative protein expression analysis. Thermo Scientific TMT Protein Labeling Chemistry complements our industry-leading mass spectrometry instruments and software.

Moving science forward
The SOLiD™ 3 System empowers you to discover new insights into gene regulation. With a hypothesis-generating, highly-sensitive assay that preserves the strand information of the original RNA molecule, the SOLiD 3 System enables you to detect novel transcripts, including non-coding RNA and distinguish strand-specific expression patterns. The combination of 400 million sequence tags and sample multiplexing capabilities allows for cost effective analysis of multiple samples in a single run while maintaining the sensitivity to detect low abundance transcripts. Combined with the SOLiD Small RNA Expression Kits, Whole Transcriptome Analysis Kits and application specific data analysis tools, the SOLiD 3 System provides a comprehensive solution that takes you from RNA to results.

For more SOLID PROOF, visit solid.appliedbiosystems.com
Over 8,000 new monoclonal and polyclonal antibodies with more added every day.

- Validation using a variety of applications -- WB, IHC, IF, ELISA
- Broad coverage of both primary and secondary antibodies
- Convenient target searches with advanced web tools
- World-renowned technical service

Go to sigma.com/antibody for more information.

Our Innovation, Your Research — Shaping the Future of Life Science
Challenging Experiments?
Advanced TC™ cell culture vessels from Greiner Bio-One

- Innovative polymer modification improves cellular adhesion
- Positive effects on cell functionality and performance
- Enhanced propagation of fastidious cells
- Improved cell expansion under limited growth conditions
- Better assay consistency
- Long-term stability and storage at room temperature
Why go with the flow when you can leap ahead?

Make a quantum leap with Millipore's breakthrough portfolio of FlowCellect™ cytometry kits optimized to monitor stem cell differentiation, GPCR expression and signaling pathways. Visit www.millipore.com/flowcytometry to learn how these assays combine with the Guava® system to provide turnkey solutions.


THE EXPERTISE OF CHEMICON® AND UPSTATE® IS NOW PART OF MILLIPORE
Each of your valuable samples deserves the best treatment. See for yourself how the Eppendorf tips will save time and reduce costs.

With respect to material, fit, design and operating forces our tips set new standards. The close environment of each sample should be adapted to its specific quality and purity needs. This can involve a specific purity level or the absence of certain substances, but also stability, reliability, or geometry. The Eppendorf tips are designed to cover all of the specific needs of your samples!

Dispensing at its best! Eppendorf Combitips® plus
- Precise dispensing of liquids with high vapor pressure or higher viscosity
- Contamination-free dispensing due to positive-displacement principle
- Dispensing from 1 µl to 50 ml through 9 different sizes
- Automatic Combitip recognition with Multipette® plus and Multipette® stream/Xstream®

Learn more about Eppendorf Combitips plus: www.eppendorf.com/CBT

Make the best of it!

Top quality for your sample
The new Thermo Scientific NanoDrop 2000 and 2000c Spectrophotometers offer true micro-sample analysis, with sample size capability as low as 0.5 µl and a measurement time of less than five seconds. Either of these is the perfect instrument for all your quantitation needs—DNA, RNA, proteins and more. Providing full spectrum UV-Vis results, both instruments can analyze samples with concentrations greater than 15,000 ng/µl (dsDNA) without dilutions. Innovative software makes it easy to build your own methods, design reports and export data. And with both pedestal and cuvette capability, the NanoDrop™ 2000c is the one spectrophotometer that does it all.

**Test-drive the NanoDrop 2000 or 2000c in your own lab!**

Visit [www.nanodrop.com](http://www.nanodrop.com) to schedule your test-drive. Try out an instrument and run your own samples. It’s completely free.

* Available only in US and Canada
MDS Analytical Technologies introduces the Arcturus® Live Cell Module for its ArcturusXT® Laser Capture Microdissection (LCM) instrument. The ArcturusXT® instrument, the industry leader for tissue-based microdissection, is now capable of imaging and microdissecting living cells from a Petri dish format.

- Live Cell Microdissection from Petri Dish
  - Isolate cells for subsequent reculture
  - Process collected cells for downstream molecular analysis
- Live Cell Imaging from Petri Dish
  - Nikon Eclipse Ti-E research microscope base
  - Optional high resolution camera with MetaMorph® imaging software
- Modular Petri Dish Stage Insert
  - User-installable
  - Easy exchange for standard tissue microdissection
    - Standard slides (25 mm)
    - Large format slides (38 mm and 50 mm)


The Axon GenePix® 4300A and GenePix® 4400A microarray scanning and analysis systems from MDS Analytical Technologies combine increased imaging resolution with a host of enhancements to offer the highest performance of any slide-based microarray scanners. Our open platforms will enable image acquisition of any fluorescent microarray, even the newest ultra-high density formats.

**Axon GenePix 4300A and GenePix 4400A Scanners**

- Increased resolution: 2.5 µm per-pixel imaging in the GenePix 4400A system for the highest-density arrays; 5 µm resolution in the GenePix 4300A system can be upgraded after purchase
- Controlled uniformity: obtain more consistently reproducible results
- Improved sensitivity: revised optics give unmatched limit of detection
- Superior configurability: choose from four lasers and employ up to 16 different emission filters
- Enhanced automation: minimize your interaction and your time with a fully automated system

**Axon GenePix Pro 7 Software**

- Powerful multiplexing: independent image acquisition and analysis of multiple arrays per slide
- Complete hardware compatibility: control any Axon GenePix scanner, new or old
- Operating system flexibility: use 32- or 64-bit versions of Microsoft Windows Vista or Microsoft Windows XP

For more information, visit [http://www.moleculardevices.com/pages/instruments/microarray_main.html](http://www.moleculardevices.com/pages/instruments/microarray_main.html).
Cancer is a complex family of diseases, characterized by the deregulation or dysregulation of the normal control pathways for cellular growth and/or apoptosis. Traditional research programs have focused on identifying and quantifying environmental and inherited factors associated with cancers found in particular tissues. Despite many advances, these approaches have historically been limited in scope due to technological limitations or excessive cost. With next-generation genomic platforms, scientists are now able to cost-effectively assay individual cancer genomes and characterize them in terms of the global genetic, epigenetic, and transcriptional changes. In-depth characterization of these events—and the relationships between them—will lead to better understanding of the mechanisms of tumorigenesis, metastasis, and therapeutic response. In this timely webinar, a panel of distinguished scientists will share their latest advances in cancer genomics and offer their views on the road ahead for this important area of research.

During the webinar, you will:

• Obtain a general overview of current technologies applied in cancer genomics research
• Learn about new technologies being used to advance the study of cancer
• See recent data provided by experts in the field
• Have your questions answered live!

Participating Experts:

Sean Grimmond, Ph.D.
Institute for Molecular Bioscience
The University of Queensland, Australia

David Wheeler, Ph.D.
Baylor College of Medicine
Houston, Texas

John McPherson, Ph.D.
Ontario Institute for Cancer Research
Toronto, Canada

Join our panel of experts in a live discussion. Register to participate.

Questions can be submitted live to the panel during the webinar or in advance via e-mail provided with registration. To register, visit

www.sciencemag.org/webinar

Webinar sponsored by Applied Biosystems
robust & reliable

Now it’s easier to go green.

GoTaq® Hot Start with Green Buffer:

- High-performance PCR buffer
- Built-in gel loading buffer and tracking dyes
- Master mix and standalone formats

www.gotaqgreen.com


GoTaq® Hot Start Polymerase with Colorless Flexi Buffer (C) and with Green Flexi Buffer (G) outperforms antibody (I) or chemically modified (A) competitor hot-start DNA polymerases for amplification of a 2.4kb fragment of the human APC gene.

Purchase GoTaq Hot Start, and Promega will plant a tree*

*offer valid from March 1 to May 31, 2009. See Web site for details.
This annual international research prize recognizes accomplishments in neurobiology research based on methods of molecular and cell biology. The winner and finalists are selected by a committee of independent scientists, chaired by the Editor-in-Chief of *Science*. Past winners include postdoctoral scholars and assistant professors.

To be eligible, you must be 35 years of age or younger. If you’re selected as this year’s winner, you will receive $25,000, have your work published in the prestigious journal *Science* and be invited to visit Eppendorf in Hamburg, Germany.

Deadline for entries
June 15, 2009

For more information
www.eppendorf.com/prize
NEW JERSEY COMMISSION ON BRAIN INJURY RESEARCH

GRANT AVAILABILITY

1. Individual Research Grants not to exceed $495,000 over 3 years.

2. Multi Investigator Research Grants not to exceed $1,980,000 over 3 years.

3. Postdoctoral/Graduate Student Fellowship Research Grants:
   • Postdoctoral Fellowship grants not to exceed a $162,000 stipend over 3 years based on the postgraduate level of experience.
   • Graduate Student Fellowship Research Grants not to exceed a $72,000 stipend over 3 years.

4. Pilot Research Grants not to exceed $165,000 over 2 years.

Application forms and details contact: www.nj.gov/health/njcbir

New Jersey Commission on Brain Injury Research
PO Box 360
Market and Warren Streets
Trenton, New Jersey 08625-0360
Tel: 609-633-6465
E-mail: njcbir@doh.state.nj.us

Letter of Intent deadline: June 1, 2009
Closing date for applications: September 1, 2009
Can you apply the laws of science to the science of laws?

We’re looking for PhD-level scientists and engineers to interface with California Legislators and policy-makers in the area of science and technology. If you are interested in the public policy decision-making process, we encourage you to apply for one of our fellowships.

Applications will be accepted for the 2009-2010 program until 12 pm PDT on May 29, 2009. For more information, minimum requirements and an application, visit us at fellows.ccst.us

Lambda LS Xenon Arc Lamp Brilliant!

Superior quality and exceptional performance are yours with the Lambda LS stand-alone xenon light source. Especially useful for fluorescence, the system features a uniform, collimated beam of light with a flat spectral output from 340-700nm. Available with a powerful 175W or 300W lamp, IR eliminating cold mirror and regulated DC power supply, it outshines the competition.

A NEW WAY TO LOOK AT CANCER

For a short time, Science is making a free, limited edition poster available to scientists interested in genomics, cancer research, and related fields. The poster identifies and explains some of the latest technologies utilizing comparative genomic hybridization techniques.

You can request delivery of the poster by mail when you visit www.sciencemag.org/products/aCGHposter.

The poster is also available as a downloadable PDF on the Science magazine website. Visit www.sciencemag.org/products today.

This poster is brought to you by the Science/AAAS Business Office and sponsored by Agilent.
**CALL FOR PAPERS**
Submit your research now to be one of the first to be considered for publication in the inaugural issue of *Science Translational Medicine*!

*Science Translational Medicine*, to be published online weekly beginning in the fourth quarter 2009, focuses on the conversion of basic biomedical research into practical applications, thus bridging the research-to-application gap.

The editors of *Science Translational Medicine* are accepting manuscripts for review in the following areas: cancer, cardiovascular disease, metabolism/diabetes/obesity, neuroscience/neurology/psychiatry, immunology/vaccines, infectious diseases, policy, behavior, bioengineering, physics, chemical genomics/drug discovery, imaging, applied physical sciences, medical nanotechnology, drug delivery, biomarkers, gene therapy/regenerative medicine, toxicology and pharmacokinetics, data mining, cell culture, animal, and human studies, medical informatics, and other interdisciplinary approaches to medicine.

Review the information for authors at [http://sciencemag.org/marketing/stm/papers.dtl](http://sciencemag.org/marketing/stm/papers.dtl)

Submit your research at [www.submit2scitranslmed.org](http://www.submit2scitranslmed.org)

For more information, contact Editor Katrina Kelner, Ph.D. at scitranslmededitors@aaas.org

» **Elias A. Zerhouni, M.D.**
  **Chief Scientific Adviser**
Senior Fellow, Global Health Program, Bill & Melinda Gates Foundation
Former Director, National Institutes of Health

» **Katrina L. Kelner, Ph.D.**
  **Editor**
American Association for the Advancement of Science

**Advisory Board Members**
Kenneth R. Chien, M.D., Ph.D.
Director, Cardiovascular Research Center, Massachusetts General Hospital, Harvard Stem Cell Institute, Harvard Medical School

Harry C. Dietz, M.D.
Professor, Institute of Genetic Medicine, Johns Hopkins Hospital
Investigator, Howard Hughes Medical Institute, Johns Hopkins University School of Medicine

Jeffrey I. Gordon, M.D.
Director, Center for Genome Sciences, Washington University in St. Louis, School of Medicine

Philip Greenland, M.D.
Senior Associate Dean, Clinical and Translational Research, Feinberg School of Medicine
Director, Northwestern University, Clinical and Translational Sciences Institute
Former Editor, Archives of Internal Medicine

Joseph B. Martin, M.D.
Professor, Neurobiology and Co-Chair, Governance, NeuroDiscovery Center, Harvard Medical School
Former Dean, Harvard Medical School

Elizabeth G. Nabel, M.D.
Chief and Principal Investigator, Nabel Lab, Cardiovascular Branch, Vascular Biology Section
Director, National Heart, Lung, and Blood Institute, National Institutes of Health
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.

**SANGER WHO?**

**SEQUENCING THE NEXT GENERATION**

In November 2008 Elaine Mardis of Washington University in St. Louis and colleagues published the complete genome sequence of an individual with acute myeloid leukemia. Coming just a few years after the decade-long, multibillion dollar Human Genome Project, the paper was remarkable on several levels. For one thing, the team sequenced two human genomes, both cancerous and normal, some 140 billion bases in all. More impressive, though, was what the study omitted: the 50 human genomes Mardis sequenced that year (albeit not as deeply) for the 1,000 Genomes Project. “It’s like a whole new world,” she says. Welcome to the sequencing frontier. By Jeffrey M. Perkel

Elaine Mardis’s acute myeloid leukemia work comprised about nine months of collecting 32-base snippets at the rate of about a billion bases per instrument every five days, with five instruments running in parallel, she says. “That seemed like a huge amount [of sequence] at the time,” she recalls.

The instruments in question, Illumina Genome Analyzers, are one of a cadre of so-called next-generation DNA sequencers. Over the past five years they have wrested control of the high-end sequencing market from the once-dominant Sanger dideoxy sequencing chemistry and its workhorse, the 3730xl from Applied Biosystems (now part of Life Technologies, formerly Invitrogen).

Yet today, says Mardis, those heady gigabase-a-week days seem “sort of like ‘ho hum, that took a really long time.’”

Adam Lowe, Illumina’s director of life science product marketing, estimates that his company’s user base “generates about a thousand gigabases per week,” he says, or about 20 times the size of Genbank in 2005.

Harvard University geneticist and next-gen pioneer George Church says the rate of technical improvement in the sequencing arena is unprecedented, about 10-fold per year, and far outpaces Moore’s Law. Illumina reads are now at 75 bases standard, and have been pushed as far as 250 with overlapping paired-end reads (about 40 gigabases per run). Life Technologies has doubled the throughput on its next-gen SOLiD instrument every quarter. At those levels, says Mardis, the study that took some “90-ish” Illumina runs to accomplish in 2008 would require just six or eight today.

Such is life on genomics’ bleeding edge. Rising from relative obscurity in 2005 to Science’s Breakthrough of the Year in 2007, the next-gen sequencing industry now sports five commercial offerings, with several others nearing release. Employing different strategies and addressing different applications, each promises previously unimaginable data output.

Unsurprisingly, the technology is attracting attention. “People can’t seem to get enough of it,” says Church. “When you get a factor of 10,000 [improvement] in four years, people eventually notice.”

Out with the Old…

Prior to 2005, almost all DNA sequencing used a variant of the chemistry first described in 1977 by Fred Sanger.

Sanger’s methodology coopts the normal process of DNA synthesis by blocking the growth of new DNA chains using a sort of molecular brake called a [continued >]

© ISTOCKPHOTO.COM/OSULIO

“it’s like a whole new world.”

Look for these Upcoming Articles

Molecular Diagnostics — May 8
Nucleic Acid Purification and Manipulation — May 15
Technologies for Gene Transfer — June 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.
The applications tend to break apart on read length dependency versus tag density.

“...”

dideoxynucleotide terminator. The resulting pool of molecules, which on average will terminate at every position, can then be sequenced chromatographically, originally on large polyacrylamide gels, and later in hair-thin capillaries.

The ABI 3730 ran 96 capillaries in parallel, each capable of producing between 500 and 1000 bases of high-quality data per run. At that rate, says Mardis, the system could produce about 1.2 megabases per day. Compare that to the gigabase throughputs being generated on new equipment, and it’s clear why sequencing centers are mothballing their old equipment.

But throughput isn’t the only factor; Sanger sequencing is labor-intensive and expensive. DNA to be sequenced first must be cloned, and the resulting libraries maintained. That requires instrumentation and labor, not to mention lab real estate.

New sequencing technologies employ completely different paradigms. All avoid size-separation in favor of strategies in which fragmented DNA is immobilized in a fixed position and repeatedly interrogated, like an iterative microarray assay. Most, but not all, use polymerase chain reaction to amplify that DNA; Helicos Biosciences, Pacific Biosciences, and ZS Genetics actually read single molecules instead. 454 Life Sciences (part of Roche Applied Science), Illumina, Helicos, and Pacific Biosciences use DNA polymerase to drive their sequencing reactions, but Polonator (Dover Systems), SOLiD (Life Technologies), and Complete Genomics sequence with DNA ligase, and ZS Genetics uses electron microscopy. And whereas most reactions are synchronous—that is, sequential, with a single base interrogated at a time at each position—454 and Pacific Biosciences generate asynchronous reads, such that individual reactions run at their own rates and are not synchronized to one another.

The Longest Read
In 454’s process the DNA to be sequenced is fragmented into 500- to 1,000-base-pair pieces and capped on each end with adaptors. The fragments are then amplified on bead surfaces via emulsion PCR (emPCR), a massively parallel strategy that separately amplifies each fragment inside an aqueous microdroplet in oil emulsion to create a sort of run-time sequencing library. Michael Egholm, chief technology officer and vice president of research and development at 454 Life Sciences calls emPCR one of “several key innovations” at the heart of the company’s success.

Following amplification, the emulsion is broken and the beads placed into the wells of a PicoTiterPlate (PTP), a 6-by-6-cm support of some 3.5 million packed optical fibers etched with wells on one side. “We deposit the beads in each of these [fibers], and cleverly, the holes on the end of the optical fibers are such that there’s only room for one bead,” says Egholm. The result: an immobilized array of 3.5 million beads, each containing millions of identical DNA fragments. Each bead is what Church calls a polony, or polymerase colony.

The sequence itself is read via pyrosequencing, which monitors base-incorporation via the resulting release of pyrophosphate. Pyrosequencing converts that pyrophosphate into ATP, which in turn drives luciferase. As a result, a burst of light is produced whenever a new base is incorporated.

In 454’s Genome Sequencer FLX, this process occurs in a flow cell. Basically, each base is flowed sequentially over the PTP – first A, then C, then G, then T. If the next base in the template is a G, the polymerase must wait until dCTP flows in. At that point, it will incorporate the base and release pyrophosphate, resulting in a flash of light whose intensity is directly proportional to the number of bases added (CCC will yield light three times as intense as C alone). That light is picked up by the optical fibers and transmitted to a camera, which reads the reaction.

Illumina and Helicos don’t use pyrosequencing, yet their processes are largely similar, except the amplification (in the case of Illumina’s technology) occurs directly on the flow cell rather than on beads, and the synthesis uses fluorescently labeled, reversible terminators; the reactions thus pause after each incorporation event (as if using a sort of Sanger sequencing 2.0). Helicos eliminates the amplification step, using what it terms true single molecule sequencing.

On the other hand 454 uses only standard DNA building blocks. As a result, says Egholm, it is both fast and free of background. “It’s almost biblical, there’s light and then there’s no light,” he says. And, producing by far the longest reads of any next-gen instrument, between 400 and 500 bases per bead, Egholm says the FLX can sequence 50 million bases per hour.

The Power of the Short Read
With read lengths approaching those of the 3730, 454 far outpaces the 125 bases Illumina is rolling out, not to mention the SOLiD’s 75, Complete Genomics’ 70, or the Polonator’s 26. As such, it has become the de facto choice for metagenomics, immunogenomics, viral profiling, whole-transcript sequencing, and especially de novo genome sequencing. The technology has been used to sequence and assemble the Arabidopsis and Drosophila genomes from scratch—but not the human; when James Watson’s DNA was decoded in April 2008, that was resequencing, aligning reads to the preexisting reference framework made possible by the Human Genome Project.

On the other hand, the FLX’s 1.25 million reads is a mere fraction of what other instruments can produce. The Polonator yields 200 million to 400 million mappable reads and the SOLiD about 750 million. At the February 2009 Advances in Genome Biology and Technology (AGBT) meeting, Illumina presented data suggesting it could generate 520 million mappable reads per paired-end run. At those levels, a whole different set of applications opens up, including digital RNA profiling, targeted resequencing, and polymorphism discovery.

“The applications tend to break apart on read length dependency versus tag density,” says Kevin McKernan, senior director of scientific operations for SOLiD at Life Technologies. At the Broad Institute of Harvard and MIT, whose 40 3730s, 20 Genome Analyzers, 10 FLXs, 8 SOLiDs, and one Polonator churned out 3 petabases of sequence in 2008, the SOLID tackles “applications that require tons of data,” such as polymorphism discovery and tumor profiling, says continued >
Call for Papers

Science Signaling

*Science Signaling*, from AAAS, the publisher of *Science*, features top-notch, peer-reviewed, original research. The journal publishes leading-edge findings in cellular regulation including:

- Molecular Biology
- Development
- Immunology
- Neuroscience
- Microbiology
- Pharmacology
- Physiology and Medicine
- Biochemistry
- Cell Biology
- Bioinformatics
- Systems biology

Subscribing to *Science Signaling* ensures that you and your lab have the latest cell signal resources. For more information visit [www.ScienceSignaling.org](http://www.ScienceSignaling.org)

**Chief Scientific Editor**

Michael B. Yaffe, M.D., Ph.D.
Associate Professor, Department of Biology
Massachusetts Institute of Technology

Submit your research at:
[www.sciencesignaling.org/about/help/research.dtl](http://www.sciencesignaling.org/about/help/research.dtl)

**HybSelect™**

**CATCH YOUR REGION OF INTEREST FOR YOUR NEXT GENERATION SEQUENCING!**

- Excellent SNP detection
- Deep sequence coverage
- Automation minimizes hands-on time
- Simple and streamlined workflow

**HYBSELECT NOW AVAILABLE**

- as a full-service from febit Analytical Services
- as the newest application for the Geniom RT Analyzer

**febit**

READ, WRITE, UNDERSTAND THE CODE OF LIFE

Europe: phone +49 6221 6510-300  •  info@febit.eu
Americas: phone +1 781 391 4360  •  info@febit.com

[www.febit.com](http://www.febit.com)
Congratulations to the 2009 AAAS Student Poster Competition Winners

AAAS recognizes the winners of the 2009 Student Poster Competition that took place at the AAAS Annual Meeting in Chicago. Their work in a variety of fields displayed originality and understanding that set them apart from their colleagues. First-place winners receive cash prizes thanks to the generous support of Subaru of America, Inc.

BRAIN AND BEHAVIOR

Winner (tie): Diane Livio, University of California, Irvine  
*The Strength of Sexual Imprinting Effects in Zebra Finch (Taeniopygia guttata) Populations*

Winner (tie): Maira Soto, University of California, Irvine  
*Characterization of Novel Human Beta-defensins*

ENVIRONMENT AND ECOLOGY

Winner: Rebecca Aicher, University of California, Irvine  
*Soil Nitrogen Affects Convergence in Community Composition in California Grasslands*

Honorable Mention: Jaquan Horton, University of California, Irvine  
*Tough Guts?: The Material Properties of Teleost Intestinal Tissues*

MATH, TECHNOLOGY, AND ENGINEERING

Winner: Ross Barnowski, University Michigan, Ann Arbor  
*Remote Detection of Radioactive Plumes Using Millimeter Wave Technology*

Honorable Mention: Alejandro Campos, University of Rochester  
*Advances in Dust Detection and Removal for Tokamaks*

MEDICINE AND PUBLIC HEALTH

Winner: Eric Howell, Texas Tech University  
*Chemotherapeutic Regulation of the Chernobyl Rodent Apodemus flavicollis*

Honorable Mention: Arun Paul, Rosalind Franklin University of Medicine and Sciences  
*Therapeutic Potential of COX Inhibitors in Treating Kaposi’s Sarcoma Herpes Virus (KSHV) Associated Body Cavity B Cell Lymphoma (BCBL)*

MOLECULAR AND CELLULAR

Winner (tie): Kathleen Mettel, University of Illinois, Urbana-Champaign  
*Plasticity in the Auditory Thalamus Following Exposure to Complex Acoustic Sequences*

Winner (tie): Nagaraj Kerur, Rosalind Franklin University of Medicine and Science  
*KSHV Infection Induces Inflammasome in Human Monocytic THP-1 Cells*

Honorable Mention: Adriana Garcia, University of California, Irvine  
*18F-fallypride MicroPET Imaging To Monitor Pancreatic Beta Cell Loss in Diabetes Mellitus*

PHYSICAL SCIENCES

Winner: Patrick Brown, University of Notre Dame  
*Vertically-Aligned Carbon Nanotube Growth for Energy Storage Applications*

Honorable Mention: Amanda David and Agustin Diaz, University of Puerto Rico, Rio Piedras  
*Encapsulation of Insulin into Inorganic Layered Nanomaterials Envisioned as a Drug Delivery System*

SOCIAL SCIENCES

Winner: Vanashri Nargund, Indiana University  
*The Influence of Secondary Science Teachers’ Beliefs on Classroom Instruction in India*


The Student Poster Competition recognizes the individual efforts of undergraduate and graduate students working toward a degree. Posters are judged at the meeting. Winners in each category receive a cash award, framed certificate, and AAAS membership. Postdoctoral scholars who hold a doctoral degree are not eligible to enter.

Full abstracts can be viewed at www.aaas.org/meetings
Chad Nusbaum, co-director of the institute’s genome sequencing and analysis program.

The SOLiD, along with the Polonator and Complete Genomics’ process, is based on sequencing by ligation, a strategy Church first successfully demonstrated in 2005 on *E. coli*.

The process, Church explains, “is directly mappable to sequencing by polymerase. In both cases you’ve got a template and a primer. In one case polymerase adds a mononucleotide, and in the other case ligase adds an oligonucleotide 6-to-9 bases long, where one of the bases is keyed to the color.”

In general, given a primer-template pair, you add a pool of short oligonucleotides whose sequence is completely random, except that one base corresponds to the fluorescent dye attached to the molecule; you then let ligase make the base call.

Say you are using six-base-long oligos and interrogating base No. 3. Of the 4,096 possible hexamers, 1,024 have an A at position 3 and a corresponding color, 1,024 have a C at that position and a different color, and so on. Only that one oligo whose sequence precisely matches the template will bind strongly enough to be ligated, so that, when the unbound molecules are washed away, the reaction will glow a uniform color. Then, to read the next base, simply denature the primer-template pair, add new primer, and repeat.

One advantage of this approach is that, unlike polymerase-based methods, the bases may be read out of order, thereby eliminating polymerase-induced errors. “In a way, it’s better than the polymerase, where you go sequentially, in the sense that there’s a certain element of random access to this,” Church says. Another advantage: unlike with polymerases, ligase can sequence in both the 5′-to-3′ and 3′-to-5′ directions.

But ligation strategies also produce extremely short fragments, which must then be aligned to a reference. Complete Genomics generates 70 bases by reading a few bases from each of eight start sites; the Polonator interrogates 26 bases by reading two sets each of six and seven bases, respectively, from either end of a longer DNA fragment (a strategy called paired-end sequencing, also supported by Illumina and Life Technologies, which improves the mappability of short sequences by adding phase information). Life Technologies actually garners the longest contiguous reads of any ligation strategy—up to 75 bases, according to data presented at AGBT—by reading two bases at a time at five-base increments, resetting, and repeating the process with a one-base frameshift.

**The Next Next-Generation?**

Other companies are pushing alternate strategies. Like Helicos, Pacific Biosciences is pursuing single molecule sequencing. The company arrays DNA polymerases on the surface of a plate, relying on zero mode waveguides to isolate the individual enzymes and watch as they add base after fluorescent base using a highly multiplexed confocal fluorescence microscope built for the purpose, says founder and chief technical officer Stephen Turner.

“The differentiation with Pacific Biosciences [compared to other polymerase-based strategies] is that we don’t stop the action of the polymerase,” says Turner. “We let it go at its native speed, and we watch in real time and simply record the activity of the polymerase.”

The technology potentially could produce reads far longer than 454’s. The company announced at AGBT the sequencing of the *E. coli* genome with reads averaging 586 bases and as long as 2,805; a commercial launch is planned for 2010.

ZS Genetics literally reads sequences using transmission electron microscopy (EM). A DNA-copying step is used to substitute the normal bases of DNA with variants containing proton-rich atoms (such as 5-bromo-dCTP), then are visualized directly in the EM. Though still in development, says William Glover III, company president and vice president of research and development, “We expect to have, when we launch, in the range of five-to-8,000-base pair reads or better.”

Whatever the future, next-generation sequencing has entered the scientific zeitgeist. It has its own X-Prize challenge, and garnered two spots in *The Scientist* magazine’s 2008 top 10 technologies list. Knome is actively selling consumer genomics at $99,500 apiece, while Complete Genomics talks of sequencing one million human genomes in the next five years. Church’s Personal Genome Project has signed up some 10,000 volunteers to have their genomes sequenced and released into the public domain.

Meanwhile, technology development continues on the next next-gen, based on such ideas as nanopores and fluorescence resonance energy transfer between nucleotide and polymerase.

That’s not to say Sanger chemistry is disappearing; some applications simply don’t need next-gen throughput. And with their on-the-fly library generation, next-gen strategies don’t support clone reanalysis. Though the need likely exists for multiple technologies (say, long versus short reads) what remains to be seen is whether the market exists for so many competing technologies. One thing is certain: at the current pace of development, 2009 should be a very interesting year.

---

Jeffrey M. Perkel is a freelance science writer based in Pocatello, Idaho.

DOI: 10.1126/science.opms.p0900033
Digital Array
The Fluidigm 12.765 Digital Array is for performing simple, fast, and reliable high throughput polymerase chain reaction (PCR) applications, including target quantitation, copy number variation, and mutation detection. The 12.765 Digital Array is an integrated fluidic circuit that makes use of a network of integrated channels and valves to divide a mixture of sample and PCR reagents into 765 replicates. The chip is specially designed to quantify target sequences accurately and to detect low-abundance targets that differ by only a base-pair from the wild-type sequence. These targets can be difficult to detect with conventional assays. Because it requires just four simple steps, the array transforms digital PCR into a straightforward, routine approach for applications that demand extreme accuracy.

Sequence Enrichment
The Sequence Enrichment Solution for targeted sequencing of the human genome enables the high-resolution analysis of genetic variation between individuals within populations at a superior level. It improves sample uniformity and reduces the selection bias typically associated with targeted sequences. It includes the RDT 1000 system, which generates picoliter-volume discrete polymerase chain reactions in droplets at the rate of 10 million per hour. It amplifies hundreds to thousands of genomic loci with high specificity and uniformity. The Sequence Enrichment Solution also includes Custom-Order Content for DNA Primer Libraries, which enable the amplification of hundreds to thousands of genomic loci in a single tube.

RainDance Technologies
For information 781-861-6300
www.RainDancetech.com

Sequence Chemistry Kits
New sequencing chemistry kits and complementary software for the Genome Analyzer enable researchers to generate 40 percent more reads per run and extend read length to more than 75 base pairs. The Mate Pair Library Preparation Kit provides support for generating longer insert paired-end libraries and complements Illumina’s existing short-end paired libraries. These new products enable researchers to generate 10 to 15 gigabases of high-quality data per run, more than doubling the output of the Genome Analyzer. The mate pair library kits and long paired-end reads improve the ability of Illumina sequencers to sequence complementary DNA libraries and even open the possibility of de novo sequencing of complex organisms. The flexible mate pair technique also allows researchers to generate paired-end insert libraries measuring two to five kilobases in order to more comprehensively catalog large structural variations.

Illumina
For information 858-332-4055
www.illumina.com

SNP and INDEL Detection Module
A new module for NextGENe software is designed for use with the Roche Applied Science Genome Sequencer FLX System to identify single nucleotide polymorphisms (SNPs) and sequence insertions and deletions (INDELs). The module addresses the homopolymer-based errors of the FLX system by making use of the system’s high coverage to statistically polish and correct the inherent system errors. NextGENe’s alignment tool can accurately align reads with long INDELs and identify them as mutations. Another detection application can identify SNPs by accurately aligning sample reads with a reference. The sequence alignment tool also provides information about amino acid changes, exon-intron boundaries, copy numbers, and methylation sites.

SoftGenetics
For information 814-237-9340
www.softgenetics.com

Genomics Workstation
The Zephyr Genomics Workstation is an easy-to-use, powerful automation tool for molecular biology applications. The workstation includes preinstalled methods for automating many widely used nucleic acid purification chemistries, as well the manufacturer’s methods for routine applications such as polymerase chain reaction setup and sample normalization. This capability removes sample preparation bottlenecks in next generation sequencing, microRNA analysis, genotyping, and gene expression studies. The system also includes several features to minimize common errors and simplify operation, such as a graphical user interface, partial tip loading, and an ultrasonic detector. The interface guides the user through method selection and proper deck setup, eliminating many of the common errors associated with high throughput processing. It can process as few as eight and as many as 96 samples simultaneously.

Caliper Life Sciences
For information 877-522-2447
www.caliperLS.com

Correction: An image that appeared on page 1570 of the Technology Feature in the December 12, 2008, issue of Science (volume 322) was not correctly credited. The image credit should have read: Top and bottom, image courtesy of R&D Systems, Inc.; Middle, image courtesy of Cell Signaling Technology.
When the left brain collaborates with the right brain, science emerges with art to enhance communication and understanding of research results—illustrating concepts, depicting phenomena and drawing conclusions.

The National Science Foundation (NSF) and the journal Science, published by the American Association for the Advancement of Science, invite you to participate in the seventh annual International Science & Engineering Visualization Challenge. The competition recognizes scientists, engineers, visualization specialists and artists for producing or commissioning innovative work in visual communication.

Winners in each category will be published in the February 19, 2010 issue of Science and Science Online, and will be displayed on the NSF Web site.

Award Categories
- Photographs/Pictures
- Illustrations/Drawings
- Informational/Explanatory Graphics
- Interactive Media
- Non-Interactive Media

R&D Systems
High Performance Antibodies.

We deliver.

For research use only. Not for use in diagnostic procedures.

For details about the images above, please visit our website at www.RnDSystems.com/go/AntibodyAd

USA & Canada  R&D Systems, Inc.  Tel: (800) 343-7475  info@RnDSystems.com
Europe  R&D Systems Europe, Ltd.  Tel: +44 (0)1235 529449  info@RnDSystems.co.uk