An expression of interest

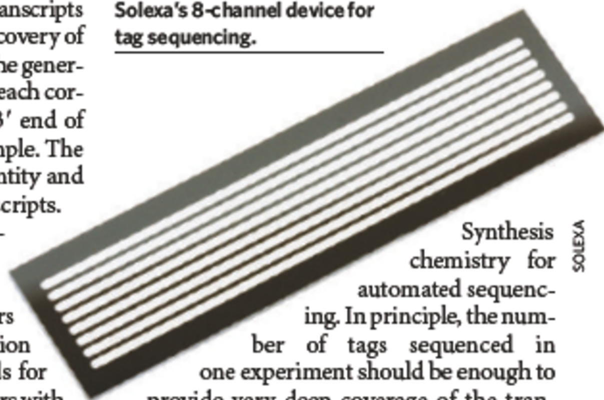
Understanding the regulation of gene expression in stem cells and neurons, and finding ways to manipulate it, are important challenges. Laura Bonetta searches out the tools for the job.

Researchers face a difficult task as they try to realize the therapeutic potential of stem cells and neurons. To better understand how to manipulate these cells, they need to monitor the gene-expression patterns, as well as working out how these genes are controlled. But both stem cells and neurons are not easy to maintain in culture, and it is hard to introduce DNA or RNA molecules into them to target specific genes or pathways.

Nonetheless, researchers have been successful in determining the expression of thousands of genes and comparing expression patterns between different cells or cells grown under different conditions. Such work has allowed them to identify, for example, master regulators of stem-cell differentiation or neuronal survival. In addition, a wide variety of tools has been designed specifically for use in stem cells or neurons to control the expression of a gene of interest and study its function.

One useful technique allows scientists to identify all the genes involved in a particular process, such as the migration or differentiation of stem cells. Serial analysis of gene expression (SAGE) is an open platform for monitoring the expression patterns of thousands of transcripts in one sample and can lead to the discovery of novel genes. The technique relies on the generation of a library of short cDNA 'tags' each corresponding to a sequence near the 3′ end of every transcript in a cell or tissue sample. The tags are sequenced to reveal the identity and quantity of the corresponding transcripts. The National Institutes of Health's Cancer Genome Anatomy Project has generated several SAGE human embryonic stem-cell libraries and offers web-based tools to analyse the expression of genes in these libraries (see 'Tools for expression analysis'). To aid researchers with the construction of SAGE libraries, Invitrogen of Carlsbad, California, sells the I-SAGE and I-SAGE Long kits.

A single SAGE experiment generates about 50,000 tags. But in mid-June Solexa in Hayward, California, will launch a sequence-based expression-analysis platform that can analyse more than one million cDNA tags. The Genome Analysis System uses a tag amplification step on the surface of a glass flow cell and features Solexa's Sequencing-by-Synthesis chemistry for automated sequencing. In principle, the number of tags sequenced in one experiment should be enough to provide very deep coverage of the transcripts expressed in a human cell and so should capture those expressed at very low levels, according to Solexa.



Solexa's 8-channel device for tag sequencing.

SOLEXA

Methods such as SAGE and the ever-popular microarrays look at the complement of transcripts isolated from a population of cells. Using a different method, based on gene-trapping technology, scientists at the Salk Institute for Biological Studies in La Jolla, California, have achieved real-time monitoring of gene expression in individual, living mouse neu-

TOOLS FOR EXPRESSION ANALYSIS

The Mammalian Gene Collection at the National Institutes of Health (NIH) in Bethesda, Maryland, provides full-length, open-reading-frame cDNA clones for human, mouse and rat genes. As part of this effort, NIH-funded scientists have created several cDNA libraries from human embryonic stem-cell lines, which can be studied using tools on the Cancer Genome Anatomy Project website.

- **The Gene Library Summarizer finds all the genes expressed in a single cDNA library or group of cDNA libraries from different tissues, including stem cells. It then classifies the genes as unique or non-unique, and known or unknown.**

- **The cDNA xProfiler compares gene-expression patterns between two pools of libraries.**

- **The Digital Gene Expression Displayer compares the 'degree' of presence of a gene in pool A with its degree of presence in pool B. This comparison is reduced to two numbers: the sequence odds ratio and measure of significance.**

The National Institutes of Health's websites offer tools for expression analysis.

The Cancer Genome Anatomy Project has also generated several long and one short SAGE (serial analysis of gene expression) libraries from human embryonic stem cells. The project's website provides tools to analyse the expression of genes represented in these libraries.

- **The SAGE Anatomic Viewer displays the relative expression of a given gene in different cells. After selecting a gene, the user can compare the expression in the pool of SAGE embryonic stem-cell libraries to all other major cell lines derived from tissues in the body, both normal and cancerous, or to a mixture of cell lines and tissues. The viewer also provides access to the Ludwig Transcript Viewer, which depicts the transcript that contains the tag and three other theoretical tags downstream of the 3′ end; and to the Digital Northern, which provides the expression of a gene in each library it is found in.**

- **The SAGE Digital Gene Expression Displayer distinguishes significant differences in gene-expression profiles between two pools of SAGE libraries.**

- **The SAGE Absolute Level Lister lists all SAGE libraries and links to the distribution of transcript expression levels in any given library.**

L.B.

©2006 Nature Publishing Group

ronal stem cells. "When you work with a therapeutically relevant cell system you want to keep it as close to its natural state as possible," says Carrolee Barlow, who helped devise the system. Barlow, who is now at Brain Cells in San Diego, California, and her colleagues created a library of stem cells, each with a single retrovirus randomly integrated in its genome. The virus was often integrated within or near a 'trapped' gene, and carried with it a reporter gene that gave off a fluorescent signal when expressed. By detecting fluorescence, the researchers could correlate the expression of a trapped gene with a specific phenotype.

"By PCR or microarray you are looking at genes associated with a specific phenotype," says Barlow. "But by analysing expression in real time you can identify the genes actually driving those phenotypic changes."

GENE EXPRESS

Liquid handling for Gene Express's Standardized Expression Measurement Center technology.

Narrowing the search

Large-scale gene-expression screens usually result in a subset of genes that warrant further analysis, with the polymerase chain reaction (PCR) being the usual choice for the first round of follow-up studies. In particular, real-time PCR allows products to be detected as they are being made, which provides a quantitative measurement of expression levels. With some platforms, it is also possible to look at four or five gene targets in a single reaction tube, increasing throughput and lowering costs. For example, QIAGEN in Venlo, the Netherlands, sells a proprietary PCR buffer solution that increases the specificity of each primer for the respective target sequence, even when several different primer sets are present in a single reaction mix. "It does not matter what primer–probe combinations you use; our product allows for specificity," says Kenneth Dwyer, marketing manager at QIAGEN.

Seegene in Seoul, South Korea, has developed a primer based on dual specific oligonucleotides that allows the length of the primer sequence to be longer than traditional primers in real-time PCR, thereby increasing specificity. "You can use more than five sets of primers in one reaction tube and never have any problems," says Seegene's founder and chief executive, Jong-Yoon Chun. "There is no need for optimization." The product, called GeneXP, is sold in kits for studying the expression of specific gene families.

For research that may lead to drug development, Gene Express in Toledo, Ohio, markets a technology known as *StaRT*-PCR. Although similar in principle to many other quantitative PCR assays, "the key differentiation is the ability to have an internal standard", according to the company's chief executive Gerald Vardzel. The technology relies on a standardized mix of competitive cDNAs included in all the reaction mixtures, which allows numerical values to be assigned to gene-expression levels and for comparisons to be made across the drug-development pipeline. Gene Express provides the technology as a service through its Standardized Expression Measurement Center.

SuperArray Bioscience in Frederick, Maryland, sells 96-well plates that include real-time PCR primer sets for different panels of pathway- or disease-focused genes. "Even people in the field may not know all the genes related to a particular process. We have done the work for them in identifying the genes of interest," says Sean Yu, vice-president of operations.

The company also sells cDNA and oligonucleotide microarrays for specific sets of genes, including one containing cell-type specific markers for human embryonic stem cells and another that carries representative markers for some of the neural phenotypes. The latter can distinguish between dopaminergic neurons, glial cells and pluripotent stem cells by their gene-expression profiles in a concentration-dependent manner. "The goal was to

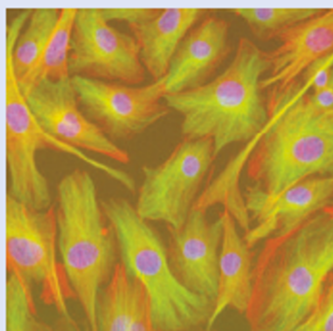
HIGH-CONTENT SCREENING

Analysis of stem-cell growth and differentiation requires experiments that manipulate a large number of variables. Typically these involve looking at cell morphology and proliferation, as well as the expression patterns of specific genes or proteins. But Cellomics in Pittsburgh, Pennsylvania, has developed a method for high-content screening that is based on automated image acquisition, processing and analysis.

"The ability to look at many things at once saves time and takes away the subjectivity," says Judy Masucci, marketing director at Cellomics. "A user sets up the criteria for what is a 'hit' and goes home. The next day the machine has cranked through all the plates and the analysis is done."

The ArrayScan VTI HSC Reader takes images of cells on slides or in microplates using multiple fluorescence channels. For example, human embryonic stem-cell colonies have been cultured on a feeder layer of embryonic mouse fibroblasts in 6-well microplates. The cells were fixed and processed so that their nuclear DNA was labelled. In addition, two other markers — the Oct-4 transcription factor (a marker of pluripotency) and phospho-histone H3 (a mitotic cell marker) — were labelled using standard immunofluorescence procedures. Analysis of images combining these parameters quantified differences in pluripotency among the various cells, the company says.

According to Masucci, what distinguishes the Cellomics idea from similar platforms is the breadth of image-analysis modules and informatics software. Different analysis modules called BioApplications are used depending on the measurements required. "We have about 20 different BioApplications and, depending on which one you're



CELLOMICS

This image of fluorescently labelled cells (left) was captured using Cellomics's ArrayScan VTI HCS Reader (right).

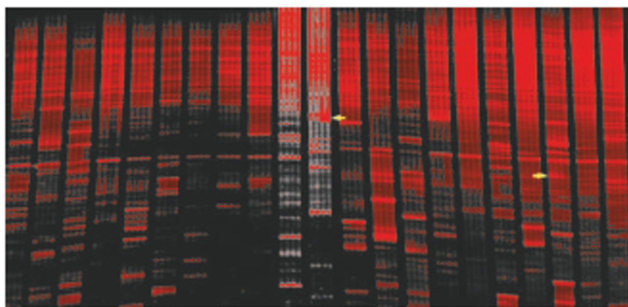
using, you can measure tens to hundreds of different parameters," Masucci says.

Some BioApplications are very general, looking at a wide variety of parameters, whereas others are more streamlined, such as the one for analysing extended neurite outgrowth. The instrument comes in three different platforms. The ArrayScan VTI HCS Reader is an automated fluorescence system primarily for fixed cells. It can analyse up to 26,000 wells a day and offers flexibility in terms of imaging applications. The KineticScan HCS Reader is used for live cell biology. The most recent addition to the family, cellWoRx, is cheaper but less flexible. In addition to instruments and bioinformatics, Cellomics offers a database for storing and managing image data. L.B.

©2006 Nature Publishing Group

provide a practical test for stem-cell differentiation," says Yu.

Along the same lines, OriGene in Rockville, Maryland, offers TissueScan Real-Time PCR panels to study gene expression in human and mouse tissues. Each array contains PCR-ready cDNAs normalized with beta-actin. For example, the Human Brain TissueScan array contains first-strand DNA from 24 tissues in the human brain. The company markets TissueScan Real-Time PCR disease panels for gene-expression analyses across many stages of disease. "It is prohibitively expensive to get hold of high-quality disease tissues," says Rich Hamer, vice-president for business development at OriGene.



GENHUNTER

GenHunter differential display showing changes in gene regulation.

Comparing expression

For comparing genes expressed by two different cells, for example an embryonic stem cell and one that has begun to differentiate, the usual method is differential display. Developed in 1992, the technology works by systematically amplifying portions of mRNAs from two or more samples and resolving the amplified products by denaturing polyacrylamide gel electrophoresis (PAGE). This allows sequence information to be recovered and corresponding cDNAs to be isolated for further molecular and functional characterizations. "It is especially good for samples with limited RNA or species that have not been sequenced," says Jonathan Meade, product manager for GenHunter in Nashville, Tennessee. GenHunter sells two kits for differential display: the RNAimage Kit uses radioactive detection, and the RNAspectra Kit uses fluorescence.

Seegene's GeneFishing DEG (differentially expressed gene) discovery kit also relies on amplification with randomly designed primers. But this method uses annealing control primers, which are longer than primers used in differential display and so provide greater annealing specificity during PCR. The resulting PCR products are sufficiently long to be detected on an agarose gel. "The big advantage is that you can use agarose gels, which are easier to prepare than PAGE," says Chun.

In most cases, microarrays and quantitative PCR technologies cannot assign a gene-expression status to a specific cell type. A better choice for this purpose is fluorescent *in situ* hybridization (FISH). This uses fluorescently labelled DNA oligonucleotide (or RNA) probes to visualize the expression of genes in a single cell with a microscope. Gene expression measured by FISH can be combined with other visual properties of the cell in a high-throughput, automated fashion (see 'High-content screening').

Aureon Laboratories in Yonkers, New York, has developed a method called peT-FISH (paraffin embedded tissue FISH) for detecting signals from nascent RNA molecules localized at the transcription site of genes being expressed. "We are looking at the presence of nascent RNA that correlates with early events of gene regulation in response to stimuli," says Paola Capodieci, one of Aureon's scientists. "The transcript is still in the nucleus and we don't know if it will become mRNA or protein. So we can see what happens at the beginning of the expression chain." So far, the method has detected up to five genes in a single cell. One of the advantages of this technology is that tissue is not destroyed. "We preserve the tissue morphology," says Capodieci.

Homing in on individual genes

To study the function of a gene, researchers typically alter its expression, either overexpressing it or knocking it out. A popular way to turn down the expression of specific genes is through RNA interference. This uses synthetic double-stranded RNA oligonucleotides known as small interfering RNAs (siRNAs), short-hairpin RNAs (shRNAs), which are expressed from a polIII promoter on a plasmid

A MICRO PERSPECTIVE ON STEM CELLS

MicroRNAs (miRNAs) are single-stranded RNA molecules roughly 22 nucleotides long that are generated from endogenous hairpin-shaped transcripts encoded in the genomes of humans, animals, plants and viruses. In mammalian cells, the miRNA molecule is incorporated into a protein complex that usually inhibits translation of mRNAs containing sequences partially complementary to it. A number of processes are regulated by miRNAs, including apoptosis, differentiation and the maintenance of stem-cell properties.

To investigate the function of miRNAs in stem cells, scientists at Applied Biosystems in Foster City, California, have developed a real-time PCR-based method to monitor the expression of 216 miRNA molecules in a single mouse embryonic stem cell. "If you study gene expression in stem cells, you don't want to look at a population of cells that may include stem cells and differentiated cells," says Kaiqin Lao, a scientist at the company. The technology Lao developed in collaboration with Azim Surani at the University of Cambridge, UK, involves hand-picking a single cell and, after lysis, performing real-time PCR. "The way we designed the assay is such that you only amplify miRNAs, so you don't need purification," says Lao. Although kits for PCR analysis in single cells are not commercially available, Applied Biosystems has launched a line of TaqMan miRNA assays for quantifying miRNAs by real-time PCR.

Several other companies also offer tools for miRNAs. Ambion in Austin, Texas, has a new line of products "to upregulate and downregulate miRNA expression", explains Ellen Prediger, senior manager of technical publications at the firm. "One can then assess the role specific miRNAs play in biological processes, including disease." Ambion sells sets of Pre-miR miRNA precursors, which are double-stranded, chemically modified nucleic acids designed to mimic endogenous mature miRNA molecules. It also makes anti-miR miRNA inhibitors — nucleic acids designed to specifically target miRNA molecules in cells.

Invitrogen of Carlsbad in California offers a set of cloned miRNAs in plasmid vectors. Unlike short-hairpin RNAs, miRNAs are expressed from polII promoters, and can easily be transferred to lentiviral vectors or vectors with inducible promoters. "We are launching a number of collections of miRNAs for druggable targets," says Peter Welch, the company's research director. L.B.

V. PALOTTA/AMBION

Ambion offers products to aid the investigation of microRNAs.

©2006 Nature Publishing Group

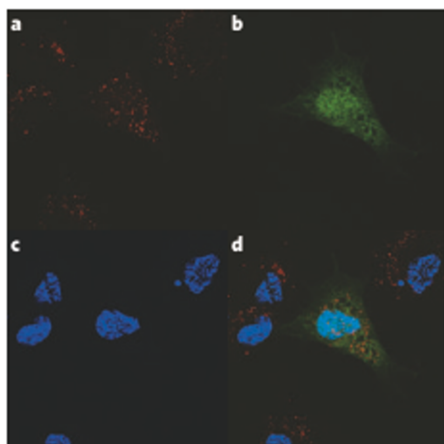
vector, or microRNAs (see 'A micro perspective on stem cells'). But slow-growing stem cells and neuronal cells are difficult to transfect with these molecules.

One way to introduce plasmids or siRNAs into cells is to use lipid-based transfection products. These work by forming a complex with DNA or RNA that interacts with the cell membrane. Invitrogen's Lipofectamine 2000 works with many mammalian cells, and can introduce an siRNA corresponding to the transcription factor Oct-4 in stem cells. "We got about 80% delivery," says Peter Welch, director for research and development.

Several companies have developed transfection reagents to introduce plasmids and siRNAs specifically into cells that are hard to transfect. Mirus Bio in Madison, Wisconsin, has the *Trans*IT LT reagent for neuronal cell lines, which is a combination of an endogenous cellular protein, histone H1 and lipoamine. "The histone H1 is the primary cationic carrier and reduces the amount of lipid needed, reducing toxicity," says James Hagstrom, vice-president for scientific operations at Mirus.

This year, Panomics in Fremont, California, will launch a peptide-based reagent called DeliverX, which it claims will have high efficiency and low toxicity. "It has been validated with one primary cell and this summer more primary and suspension cells will be added," says Ian Ley, the company's vice-president for marketing.

When transfection agents fail, electroporation is the brute-force approach for delivering genetic material inside cells. The nucleofector technology marketed by Amaxa Biosystems of Cologne, Germany, is aimed at cell lines that are difficult to transfect. A series of gentle pulses allows transfected DNA directly to enter the nucleus. "Other non-viral transfection methods rely on cell division for the transfer of DNA into the nucleus. Nucleofection provides the ability to transfect even non-dividing cells such as neurons and resting blood cells," says spokeswoman Kimberly Stevenson. Depending on the cell type, a researcher would choose one of many proprietary nucleofector solutions and program the machine with the appropriate parameters. For experiments that require stably transfected cells or *in vivo* work, viral vectors are the obvious choice (see 'Hitching a ride on a virus').



DNA delivery directly into the nucleus visualized with different fluorescent labels.

AMAXA

Regardless of the method of gene expression used, some steps in the process will probably require growing cells in culture, using formulated liquid media supplemented with growth factors and other substances that promote cellular replication and govern differentiation. The conditions and the reagents used may affect the gene-expression programmes of a cell, a concern that is particularly relevant to stem-cell research. "For the stem-cell field to move forward it is necessary to have standardized reagents and assays to allow results to be compared," says Sharon Louis, senior scientist at StemCell Technologies in Vancouver, Canada. "Where we see ourselves is providing standardized tissue-culture reagents." The media reagents provided by the company come with detailed protocols for either maintaining cells in an undifferentiated pluripotent state or for inducing differentiation towards a specific lineage.

Each in their own way, stem cells and neuronal cells present problems for the scientists who want to manipulate them. But many scientists and companies have risen to this challenge by developing increasingly sophisticated tools and technologies. With this tool kit in hand, studying gene expression in stem and neuronal cells is a realistic goal that has been embraced by researchers with an enthusiasm matching the promises of their field. ■

Laura Bonetta is a freelance writer based in the Washington DC area.

HITCHING A RIDE ON A VIRUS

Nucleic acids can be introduced into cells by allowing them to hitch a ride on a virus. Typically this involves engineering a viral vector to carry a gene of interest as well as a selectable marker or reporter gene. The vector is then introduced into a packaging cell line to produce recombinant viral particles, which are used to infect the cells of interest.

Many products on the market provide complete kits that include vector and reporter plasmids, packaging cells, PCR primers, as well as all the necessary buffers and solutions. When working with non-dividing or slow-growing cells, the best choices are systems that are based on lentiviruses or adeno-associated viruses (AAVs).

Lentiviruses, a subclass of retroviruses, can infect cells that are not dividing and are able to integrate DNA into the host cell. The ViraPower lentiviral expression system made by Invitrogen of Carlsbad, California, delivers and actively transports a gene of interest to the nucleus of a broad range of cells, including stem cells and neuronal cells.

Didier Trono has helped to devise a novel lentiviral vector.

AAV vectors can also introduce genes into these cells but there is a downside: the virus typically needs to be co-infected with adenovirus to produce recombinant AAV virions in packaging cells. One possible way around this is to use the AAV helper-free system sold by Stratagene of La Jolla, California. This eliminates the need for the helper virus, greatly simplifying the whole process.

Scientists at the Swiss Federal Institute of Technology in Lausanne have developed a lentiviral vector that they say allows regulated expression of genes or short-hairpin RNAs (shRNAs) in any cell type or tissue either *in vitro* or *in vivo* using ubiquitous or tissue-specific promoters. The technique takes advantage of the repression activity of the Krüppel-associated box (KRAB) domain. This domain is found in many vertebrate transcriptional regulators and, when it is tethered to DNA, it induces a local heterochromatin state that represses transcription.

By coupling the KRAB domain to a tetracycline operator sequence on the viral vector, expression of any gene carried by the virus is blocked. The block is readily removed by adding an antibiotic. This allows gene expression to be switched on and off at will. "The main advantage of this vector is the ability to control polIII promoters, such as those that regulate shRNA expression, and to control polII promoters, which can serve to express transgenes or miRNAs. It provides a means to do drug-inducible shRNA knockdown," says Didier Trono, dean of the School of Life Sciences at the Lausanne institute, who led the team that came up with the technique. "It has a range of applications far greater than existing systems." L.B.

©2006 Nature Publishing Group