Ten years of *Nature Biotechnology* research

Monya Baker

Authors of some of the most highly cited *Nature Biotechnology* papers from the past 10 years discuss their work and the remaining challenges for their fields.

Ten years is not a long time in the life of an industry like biotech, where drug development or commercialization of a crop can take decades. But today's research landscape is almost unrecognizable from that in 1996, when *Nature Biotechnology* was relaunched from its predecessor *Bio/technology*. Publication of the human genome sequence still lay five years off, cDNA arrays had just been described, photolithographically synthesized oligonucleotide chips were under investigation at an obscure company on the West Coast of the United States, the term 'proteomics' had just been coined two years earlier by an Australian researcher at a conference in Siena, Italy, and the first transgenic crops had just been approved for commercial use.

In the following article, *Nature Biotechnology* interviews several of the authors who contributed some of the most highly cited papers from the past ten years. The papers cover a diverse range of research areas—from array technology, proteomics, molecular beacons, fluorescent proteins, quantum dots, embryonic stem cells and gene delivery to agbiotech. The selection of research presented here is not intended to be an exhaustive representation of the most important advances in biotech from the past decade; it is simply a snapshot of the past decade's most highly cited papers in this journal according to the Institute for Scientific Information (Philadelphia, PA, USA). Clearly, research areas not included in this article, such as modeling and systems biology, chemical biology, antibody and protein engineering, biosensors, expression systems, biochemical engineering and environmental biotechnology, will continue to make important contributions to the biotech endeavor in coming years.

Monya Baker is a freelance writer based in San Francisco, California.

The interviews below provide a record of the rather remarkable progress that has been made over the past ten years.

Dawn of the gene expression array

David Lockhart remembers that his hands shook that afternoon in early 1995 as he opened his laboratory notebook and pressed the phone to his ear. After months of work, the Affymetrix scientist had been given a reference sample of ten unidentified genes expressed at four different levels. Now, Gene Brown from Genetics Institute (now part of Wyeth, North Andover, MA, USA) would tell him if the gene expression chip Lockhart's team had designed gave the right answers. After the 40th 'correct,' Lockhart's colleague and officemate, Mark Chee, tried to give him a celebratory 'high five' hand-slap, but furniture got in the way. Over a year into the job, the researchers' desks and file cabinets were still in the same random places where a moving crew had left them.

That blind test led to the first gene expression array and the most highly cited paper in *Nature Biotechnology*[1]. Lockhart and his colleagues had designed over a hundred thousand oligonucleotides as gene detectors and synthesized them directly onto a high-density array using massively parallel combinatorial chemistry. The tool could quantify mRNA molecules' abundance in cells over three orders of magnitude and unambiguously detect RNAs with frequencies as low as one in 300,000 or about one copy per mammalian cell.

But making the chip wasn't the most important part. After all, Stephen Fodor, founder of Affymetrix, and colleagues had already developed the technique to build oligonucleotides on a solid matrix. Lockhart and his colleagues had to figure out how to prepare the sample, pick what oligos to put on the chip, separate signal from noise and analyze results. Key to this was deliberately including oligos with single-nucleotide mismatches to eliminate confounding signals from cross hybridization.

In the 1996 paper, Lockhart had monitored the activity of 118 genes using 130,000 oligos as gene-detecting probes. Today, Affymetrix sells a chip packed with millions of oligos that monitors all known human exons. With the help of laser-mediated transfer, chips can also read a sample extracted from just a few cells[2]. But despite the vastly expanded readout, the chip is fundamentally the same, says Lockhart. What's different is what you can do with the data.

"Even early on, you could see that certain genes would go up and down," he recalls, but making sense of that information was often difficult because little was known about the identified genes. For example, in 1997, Lockhart and his collaborators used a gene expression array, this one containing over 260,000 probes, to detect differences in expression levels of 6000 genes in yeast growing in rich and minimal media and found large differences in both known and uncharacterized genes[3]. "That used to be one of the knocks, you find all these genes that are *associated* with the phenotype. Well what does that

David Lockhart on microarrays: "Ten years ago, the thinking was that that was way too complicated."

mean? Is it a cause? An effect? Something irrelevant coming along for the ride?"

Now, he says, people are better at designing experiments that determine the gene changes that matter. One way to do so is to combine genotyping data with gene expression data and compare differences between phenotypes. Then, researchers can manipulate the expression of identified genes and monitor the effects. This is precisely what Lockhart did as part of a team looking for insights into anxiety. Using gene expression profiling on six inbred mouse strains that exhibit different levels of anxiety, the team identified 17 genes that correlated with anxious behavior. Then, they homed in on two genes involved in the oxidative stress response. Overexpression of these genes increased measures of anxiety, whereas inhibition decreased them[4]. Lockhart says that few people expected gene expression to aid the analysis of as complex a trait as anxiety. "Ten years ago, the thinking was that that was just way too complicated. We were told a lot of times that it would never work."

Despite the exponential growth of genetic information, expression profiling can still turn up uncharacterized genes. Even changes in known genes require further study and validation. Lockhart admits that it's too early for any drug to come from targets identified by combined gene expression and genotyping assays, though of course gene expression is routinely used to identify drugs' potential toxicities and off-target effects, as well as to characterize potential drug targets.

A more serious limitation is that gene expression profiling only gives a snapshot of one level in the hierarchy of cell activity; it looks only at levels of mRNA, not concentrations of proteins, let alone proteins' activation states. But the fact that gene expression profiling raises questions that it can't answer isn't something that can be helped by better technology, says Lockhart.

"The limitations really aren't technical anymore. The limitations these days are kind of old-fashioned, designing the right experiments and getting access to the right samples."

1. Lockhart, D.J. *et al.* Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nat. Biotechnol.* **14**, 1675–1680 (1996).
2. Schutze, K. & Lahr, G. Identification of expressed genes by laser-mediated manipulation of single cells. *Nat. Biotechnol.* **16**, 737–742 (1998).
3. Wodicka, L., Dong, H., Mittmann, M., Ho, M.H. & Lockhart, D.J. Genome-wide expression monitoring in *Saccharomyces cerevisiae*. *Nat. Biotechnol.* **15**, 1359–1367 (1997).
4. Hovatta, I. *et al.* Glyoxalase 1 and glutathione reductase 1 regulate anxiety in mice. *Nature* **438**, 662–666 (2005).

Proteomics goes global

The late eighties and early nineties were depressing times for protein chemists, recalls John Yates, director of the Mass Spectrometry Laboratory at Scripps Research Institute (La Jolla, CA, USA). Colleagues liked to tell him how DNA sequencing and the consequent explosion of genetic information were rapidly transforming his field into a backwater.

But Yates found inspiration rather than despair, exploring how the idea of whole genome sequencing could be applied to the proteome. His team coupled a system of digesting proteins and separating the resulting peptides by cycles of chromatography followed by mass spectrometry in a technique they call MudPIT (multidimensional chromatographic protein identification). With this technique, peptide profiles from heterogeneous mixtures of proteins can be computationally matched with those of known peptides. In 2001, it detected nearly 1,500 distinct proteins in yeast cells, more than anyone had ever seen before in a single analysis[1]. The previous record, of about 300 proteins, had also been set by the Yates team using the more cumbersome, and less sensitive technique of tandem coupling of two-dimensional liquid chromatography[2].

MudPIT could identify many types of proteins that were previously hard to detect. "It was really good at analyzing membrane proteins," says Yates, "Just digesting them off the lipid bilayer, you could show what proteins were present." Subsequently, Yates refined techniques to recover more of the membrane proteins and to identify proteins by phosphorylation state and subcellular location[3]. This ability helped identify 67 previously uncharacterized proteins in the nuclear envelope, several of which mapped to chromosome regions linked with various dystrophies[4].

Identifying proteins is important, but so is quantifying them, says Ruedi Aebersold, a systems biologist at the Swiss Federal Institute of Technology in Zurich. Neither two-dimensional gels nor liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS) could do this, so Aebersold (then at the University of Washington, Seattle) developed a way to incorporate different isotopes into otherwise chemically identical molecules to create isotope-coded affinity tags (ICATs)[5]. The dilution of signal from the labeled molecule relative to a known amount of unlabeled molecules reveals how abundant that molecule is, making mass spectrometry quantitative as well as qualitative. "The general advance of the paper," says Aebersold, was "the widely perceived need to do proteomic measurements quantitatively and a range of robust methods to do that." This ability shifted the overall goals of proteomics to study relative quantities of proteins, as well as how protein content changes temporally.

Aebersold and his colleagues also developed a solid-phase isotope tagging reagent that allowed easier automation and higher sensitivity[6]. Additional refinements allowed the group to determine the relative abundances of almost 500 proteins contained in the microsomal fractions of naive and differentiated human myeloid leukemia cells[7]. Since then, Aebersold and other groups have developed methods to identify and quantify phosphorylated peptides. These provide a way to investigate how proteins are modified (e.g., by phosphorylation) and processed to control their own and other pathways[8,9].

Both Aebersold and Yates think the biggest barriers for proteomics are not technical, but sociological. Yates cites a need to develop standards across laboratory groups to make experiments more repeatable. Aebersold thinks the field needs to shift away from the shotgun approach. "The key idea is to generate once a complete proteomic map and then to navigate for all subsequent experiments in that mapped space, rather than to perpetually rediscover the proteome *de novo* in each experiment." The two goals are intertwined: once experiments are consistently reproducible, investigators will no longer feel the need to rediscover the proteome.

Reudi Aebersold thinks proteomics needs to shift away from the shotgun approach.

1. Washburn, M.P., Wolters, D. & Yates, J.R. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat. Biotechnol.* **9**, 242–247 (2001).
2. Link, A.J. *et al.* Direct analysis of protein complexes using mass spectrometry. *Nat. Biotechnol.* **17**, 676–682 (1999).
3. Wu, C.C., MacCoss, M.J., Howell, K.E. & Yates, J.R. A method for the comprehensive proteomic analysis of membrane proteins. *Nat. Biotechnol.* **21**, 532–538 (2003).
4. Schirmer, E.C., Florens, L., Guan, T., Yates, J.R. III & Gerace, L. Nuclear membrane proteins with potential disease links found by subtractive proteomics. *Science* **301**, 1380–1382 (2003).
5. Gygi, S.P. *et al.* Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.* **17**, 994–999 (1999).
6. Zhou, H., Ranish, J.A., Watts, J.D. & Aebersold, R. Quantitative proteome analysis by solid-phase isotope tagging and mass spectrometry. *Nat. Biotechnol.* **20**, 512–515 (2002).
7. Han, D.K., Eng, J., Zhou, H. & Aebersold, R. Quantitative profiling of differentiation-induced microsomal proteins using isotope-coded affinity tags and mass spectrometry. *Nat. Biotechnol.* **19**, 946–951 (2001).

8. Zhou, H., Watts, J.D. & Aebersold, R. A systematic approach to the analysis of protein phosphorylation. *Nat. Biotechnol.* **19**, 375–378 (2001).
9. Oda, Y., Nagasu, T. & Chait, B.T. Enrichment analysis of phosphorylated proteins as a tool for probing the phosphoproteome. *Nat. Biotechnol.* **19**, 379–382 (2001).

The *tabula rasa* of cells

It isn't often that a scientific publication get its author on the cover of *Time*, but that's what happened to Jamie Thomson when his laboratory at the University of Wisconsin derived the first human embryonic stem (hES) cell lines[1]. Two years later, *Nature Biotechnology* published the second paper describing the production of hES cell lines by Benjamin Reubinoff, Martin Pera and their colleagues at Monash University in Melbourne[2]. This paper confirmed Thomson's results, and went further, showing that hES cells could be differentiated *in vitro*.

The first cells were derived with a medium containing mouse fibroblasts, "which is kind of a standard tissue culture that you use when you don't know what you're doing," explains Thomson, "because they secrete a lot of stuff. Then over the years, we've found that the things that do support mouse ES cells don't support hES cells." Pera says hES cell culture has progressed on many fronts, particularly "refining the culture system to make it easier to expand the cells and defining the medium to eliminate animal products." The remaining challenges, he says, are scaling up the culture and creating reliable techniques to grow up an entire culture from a single ES cell. Already, Pera says, many promising culture methodologies have been described; now they need to be assessed by multiple research groups. In January of this year, Thomson's group described the first fully defined xeno-free medium for culturing human stem cells and isolated two cell lines derived in fully defined medium, though both lines had chromosomal abnormalities[3].

But Thomson thinks that the major barriers for culturing hES cells have already been overcome. "There will be improvements," he says, "but it will be kind of diminishing returns from now on." In particular, the twin specters of genetic instability and tumorigenicity will be mastered, though not exorcised, through better technique, says Thomson, who was part of the team that first described the emergence of chromosoma abnormalities in cultured hES cells[4]. "If you're really careful with the culture conditions, the cells are quite stable. It's a concern that has to be managed," he adds, "but it is ultimately manageable."

"The odd abnormal cell may not be that much of a problem if it doesn't have an advantage" that lets it outcompete healthy cells, agrees Pera. Nonetheless, abnormal cells will arise even in ideal culture conditions; the key will be identifying and removing potentially dangerous cells before they are used in patients. In this issue, Pera and colleagues[5] show that five hES cell sublines with chromosomal abnormalities all express the CD30 receptor and that the protein's expression in normal hES cells prevents apoptosis, but Pera says additional biomarkers will be essential to weed out cells likely to run amok. Understanding, and preventing, the process that sets them down that path is more important.

Of course, deriving and culturing hES cells is really a means to an end. The ultimate goal is understanding and controlling differentiation well enough that the cells can be used to generate tissue for human therapies. "We just basically need more people beavering away at it," says Thomson, who believes that developmental biologists will be able to make most clinically relevant cell types within a decade. "Ultimately, we're going to arrive at a molecular blueprint of the pluripotent stage and then we'll know what switches we need to tweak," agrees Pera, though he is quick to point out that the ability to create differentiated cells is a very far cry from the ability to use the cells for therapy. But he warns that there will be many bumps in the road towards creating differentiated tissues, stemming largely from our ignorance of basic biology. "Stem cell culture isn't an absolute reflection of embryonic development: it's sort of a caricature."

Martin Pera: "Stem cell culture isn't an absolute reflection of embryonic development; it's sort of a caricature."

1. Thomson, J.A. *et al.* Embryonic stem cell lines derived from human blastocysts. *Science* **282**, 1145–1147 (1998).
2. Reubinoff, B.E. *et al.* Embryonic stem cell lines from human blastocysts: somatic differentiation *in vitro*. *Nat. Biotechnol.* **18**, 399–404 (2000).
3. Ludwig, T.E. *et al.* Derivation of human embryonic stem cells in defined conditions. *Nat. Biotechnol.* **24**, 185-187 (2006).
4. Draper, J.S. *et al.* Recurrent gain of chromosomes 17q and 12 in cultured human embryonic stem cells. *Nat. Biotechnol.* **22**, 53–54 (2004).
5. Herszfeld, D. *et al.* CD30 is a survival factor and biomarker for transformed human pluripotent stem cells, *Nat. Biotechnol.* **24**, 345-355 (2006).

Beacons of light

Ten years ago, a pair of scientists described a new kind of probe for quantifying nucleic acids that could do it in real-time and in solution[1]. No purification, no gels, no radioactivity (and no background smudges to contend with). They did it by combining the specificity of nucleic acid hybridization with the ability of nucleic acids to take on different conformations. This was the birth of molecular beacons. In the intervening years, beacons (which are fluorogenic hairpin oligonucleotide probes) have been applied in various ways—for quantitative PCR, for pathogen detection, for assaying single nucleotide polymorphism and mutations, for incorporation into self-reporting arrays, to name a few. The simple beacon itself has evolved by adding gold particles to broaden its specificity and sensitivity, primers so that it can both prime and detect nucleic acid amplification in a unimolecular and hence faster reaction, and even enzymes so that it can potentially prime, amplify and detect a sequence all in one tidy little package.

Sangi Tyagi and Fred Kramer of the Public Health Research Institute (Newark, NJ, USA) were working at the time on Qβ replicase amplification of RNA, which, like the polymerase chain reaction that later supplanted it, had unacceptably high backgrounds when working with low-abundance RNAs. They had the idea, based on Paul Lizardi's (now at Yale University) work on conformational changes of nucleic acids, to create probes that would allow amplification only when they changed their shape.

Thus, Tyagi, Kramer and Lizardi designed a probe with a stem-loop structure that unfolds when it reacts with its target sequence, the complement for which is encoded in the loop. In the first paper, Tyagi, Kramer and Lizardi showed how their fluorogenic probe, which they dubbed molecular beacon, could assay the accumulation of a target molecule in solution (a PCR fragment of the HIV invertase gene), one of the first

real-time PCR assays in the literature[1]. The choice of the two dyes used in the original molecular beacons turned out to be providential, according to Tyagi. By scouring the literature, he found a fluorogenic peptide assay that used the fluorophore/quencher pair adenosine-5′-EDANS (5-[(2-aminoeth yl)amino]naphthalene-1 sulfonic acid) and DABCYL (4-(4-dimethylaminophenylazo) benzoic acid), used in the design of a protease assay by scientists at Abbott Laboratories (Deerfield, IL, USA)[2]. Later, when he started experimenting with different dye combinations, he was able to create a palette of molecular beacons with different colored dyes, the fluorescence from which could all be quenched by a single 'universal' quencher. Thus, a few years after the first paper, Tyagi and Kramer, together with Diana Bratu, had a second *Nature Biotechnology* paper in which they described their palette, and showed how multicolored molecular beacons could be used simultaneously in a single tube[3].

Since 1996, over 350 papers have been published on molecular beacons, five US patents have been issued, 45 companies have licensed the technology, and several diagnostics have been approved, including the first real-time probe for HIV viral load measurements. There were some pitfalls, particularly for *in vivo* use that led some to question their reliability, and sensitivity and penetrability remain challenging. Some scientists have designed around the permeability problem—adding a cell-permeating peptide, for example, has allowed Gang Bao of Emory University to detect cancer-related mRNAs in live cells[4]—Tyagi has persisted with the original design and instead found situations that exploit its inherent properties. In fruitfly oocytes, for example, specific mRNAs are concentrated in different regions of the cell. These clusters of RNAs can be detected with molecular beacons injected into the oocyte. In his most recent study, Tyagi and colleagues engineered an RNA molecule with a kite tail-like addition possessing 96 tandem molecular beacon target sites, which enabled him to track individual mRNA molecules from the site of transcription through the nucleus to the cytoplasm. "Cell biologists are really excited," he says, as this work has provided new insights into the dynamics of RNA metabolism.

Sanjay Tyagi, who with Fred Kramer, developed the original concept of molecular beacons.

1. Tyagi, S. & Kramer, F.R. Molecular beacons: probes that fluoresce upon hybridization. *Nat. Biotechnol.* **14**, 303–308 (1996).
2. Matayoshi, E.D., Wang, G.T., Krafft, G.A. & Erickson, J. Novel fluorogenic substrates for assaying retroviral proteases by resonance energy transfer. *Science* **247**, 954–958 (1999).
3. Tyagi, S., Bratu, D.P. & Kramer, F.R. Multicolor molecular beacons for allele discrimination. *Nat. Biotechnol.* **16**, 49–53 (1998).
4. Nitin, N., Santangelo, P.J., Kim, G., Nie, S. & Bao, G. Peptide-linked molecular beacons for efficient delivery and rapid mRNA detection in living cells. *Nucleic Acids Res.* **32**, e58 (2004).

More bright ideas

Fluorescent proteins have been used as tools to address a range of important problems: as tags for cell sorting to agents for tracking organelles, and as probes for studying the localization, expression and interaction of proteins. Dozens of varieties exist, offering a range of colors and intensities, some that can be activated and deactivated at will. Some fluorescent proteins can be converted from one color to another and back again. Multiple fluorescent proteins can be used in a single cell.

It's easy to forget that the complex applications and permutations began with a small, bioluminescent jellyfish that lives off the coast of Washington State. Over a decade ago, Douglas Prasher and colleagues showed that green fluorescent protein (GFP) from *Aequorea victoria* could be used to track gene expression without any need for exogenous substrates or cofactors[1]. However, the technique had practical limitations.

For one thing, natural GFP emits too little light to be efficient for tracking proteins in eukaryotic cells. By shuffling GFP DNA in *Escherichia coli* and selecting for the brightest colonies, Willem Stemmer and colleagues boosted the protein's signal about 40-fold and created a more soluble protein to boot[2]. That same year, George Phillips and his colleagues solved a high-resolution structure of GFP[3]. The structure showed the fluorophore near the center of a barrel-like β-can fold and identified several adjacent amino acids that affect spectral properties, opening the door to rational engineering and mutagenesis.

But human innovation is no match for nature's, and researchers were eager for natural homologs to illuminate the mechanism of fluorescence and to expand proteins' spectral properties. One breakthrough came from Sergey Lukyanov and colleagues at the Russian Academy of Sciences. Whereas others were combing beaches for bioluminescent organisms, Lukyanov and colleagues probed non-bioluminescent corals obtained from aquarium shops in Moscow and very quickly cloned six new GFP homologs[4]. Though these proteins had only 30% or less sequence homology with GFP, they still had the barrel conformation. Furthermore, one of these proteins, DsRed, fluoresced red, an unprecedented color of emission, expanding the breadth of wavelengths that could be used. A few years later, two groups, headed by Atsushi Miyawaki and Benjamin Glick, used random and directed mutagenesis to isolate variants of yellow and the red fluorescent protein DsRed, leading to the development of functionally improved forms[5,6].

Meanwhile, Lukyanov's team had been optimizing and expanding the uses of fluorescent proteins. Realizing that even very close homologs can have very different spectral wavelengths, the team created a 'photoswitchable' protein[7] that changes color from cyan to green when irradiated by light of 405 nm wavelength. This change in the ratio of cyan to green can be used to track the movement of tagged proteins, if fused with proteins and then activated in select subcellular locations.

Lukyanov thinks fluorescent proteins will still yield many and more-effective tools. For example, available far-red fluorescent proteins are dim, and brighter versions need to be invented or discovered. In addition, natural fluorescent proteins usually function as dimers or tetramers, limiting their uses in tagging proteins; monomeric variants of some of these have still not been made.

Above all, Lukyanov thinks the biological function and evolution of these proteins should be better studied, as the natural world could still yield exploitable ideas, including natural versions of photoswitchable, photoactivatable and phototiming proteins.

And although he expects existing protein variants to improve, he also believes entirely new categories will emerge. In January this year, his team described KillerRed, a genetically encoded photosensitizer[8] derived from the chromoprotein of the hydrozoan *Anemonia sulcata*[8]. When exposed to green light, the protein can

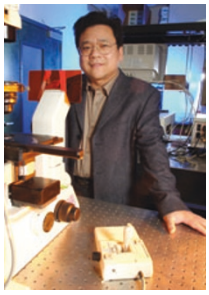
Sergey Lukyanov, who was part of the original team that isolated fluorescent proteins from corals.

kill cells and inactivate enzymes to which it has been fused. Such innovations mark the rise of "a principally novel type of fluorescent protein," says Lukyanov, "that can be used as active light-operated 'manipulators' rather than passive reporters."

1. Chalfie, M., Tu, Y., Euskirchen, G., Ward, W.W. & Prasher, D.C. Green fluorescent protein as a marker for gene expression. *Science* **263**, 802–805 (1994).
2. Crameri, A., Whitehorn, E.A., Tate, E. & Stemmer, W.P. Improved green fluorescent protein by molecular evolution using DNA shuffling. *Nat. Biotechnol.* **14**, 315–319 (1996).
3. Yang, F., Moss, L.G. & Phillips, G.N. Jr. The molecular structure of green fluorescent protein. *Nat. Biotechnol.* **14**, 1246–1251 (1996).
4. Matz, M.V. *et al.* Fluorescent proteins from nonbioluminescent *Anthozoa* species. *Nat. Biotechnol.* **17**, 969–973 (1999).
5. Nagai, T. *et al.* A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. *Nat. Biotechnol.* **20**, 87–90 (1999).
6. Bevis, B.J. & Glick, B.S. Rapidly maturing variants of the *Discosoma* red fluorescent protein (DsRed). *Nat. Biotechnol.* **20**, 83–87 (2002).
7. Chudakov, D.M. *et al.* Photoswitchable cyan fluorescent protein for protein tracking. *Nat. Biotechnol.* **22**, 1435–1439 (2004).
8. Bulina, M.E. *et al.* A genetically encoded photosensitizer. *Nat. Biotechnol.* **24**, 95–99 (2006).

Quantum leaps

Quantum dots have properties imagers long for. The nanocrystals absorb a broad spectrum of light, but the light they emit can be tuned to any narrow frequency. Compared to organic dyes, they can be 20 times brighter and orders of magnitude more stable to photobleaching. But at several nanometers across, the dots are closer in size to proteins than small molecules; a quality that could limit their uses in cells. In 2003, researchers at Quantum Dot (Hayward, CA, USA) showed that dots could label cancer markers, actin and nuclear antigens inside cancer cells[1]. The same year, researchers at Rockefeller University (New York) showed that antibodies could be conjugated to quantum dots, and that cells taking in the conjugates grew for over a week[2]. In 2004, in what one commentator hailed as a possible precursor to an 'optical biopsy,' a team led by Shuming Nie, director of the Emory-Georgia Tech Nanotechnology Center (Atlanta, GA, USA) demonstrated for the first time that quantum dots could be used to image cancer in living animals[3]. To do so, Nie and his team encased the quantum dots in an amphiphilic triblock shell. This allows multiple functionalities to be attached to each dot, and provides a general platform to add molecules that target tumors, enhance permeability or bestow other properties. In nude mice with human prostate cancer xenografts, the dots accumulated in tumors, allowing sensitive and multicolor fluorescence imaging.



Shuming Nie: "We want something that is clinically realistic."

Others also started to explore the applications of quantum dots in imaging. In perhaps the first paper to apply quantum dots to a surgical procedure, Harvard oncologist John Frangioni showed that quantum dots could transform a difficult, error-prone surgery used to detect cancer spread[4]—sentinel lymph node biopsy—into a straightforward procedure that no longer requires radioactive tracers. Key to this was the design of quantum dots that emit near-infrared light, which is scattered less by tissues than is visible light. Combined with an imaging system that integrated visible and near-infrared light, the dots clearly identified sentinel lymph nodes as deep as 1 cm below the skin in a living, 35-kg pig, allowing the nodes to be removed with minimal dissection.

But Frangioni doesn't expect quantum dots to guide human surgery anytime soon, if ever. Instead, he hopes to replace the dots with albumin attached to indocyanine green, a near infrared fluorescent dye. These probes won't work as well as the quantum dots, which can be engineered to a size that will lodge in the sentinel nodes, but that's not the point. "Now that we've proved principle," he says, "we want something that's clinically realistic."

Frangioni thinks near-infrared systems will eventually image tissues as deep as 4 cm below the skin, but that penetrance will depend not on fluorophores, but on detection systems. For quantum dots, he says, the most significant barrier is potential toxicity. Dots are typically made from heavy metals. They are too big to be removed by the kidney and too foreign to be metabolized by the liver, so they remain in the body indefinitely. To create a formulation suitable for humans, Frangioni and Massachusetts Institute of Technology collaborator Moungi Bawendi are developing smaller dots made from different elements.

Over the next five to ten years, both Frangioni and Nie believe quantum dots have the greatest potential for *in vitro* biomedical applications, particularly more sensitive detection of biomarkers. In fact, Nie's group has already developed microbeads encrusted with biodetecting quantum dots emitting different colors and intensities[5]. Theoretically, these beads could generate millions of distinct colors, a capacity beyond any imaginable practical applications. So far, about 20 different types can be detected at once, though Nie thinks the ultimate number could be as high as a thousand. The technology has already been commercialized by Crystalplex (Pittsburgh, PA, USA). Nie thinks issues of penetration and toxicity will likely keep *in vivo* applications at bay for several years, but he's not giving up. In fact, he is already putting his coatings onto near-infrared quantum dots like the ones designed by Frangioni.

1. Wu, X. *et al.* Immunofluorescent labeling of cancer marker Her2 and other cellular targets with semiconductor quantum dots. *Nat. Biotechnol.* **22**, 41–46 (2003).
2. Jaiswal, J.K. *et al.* Long-term multiple color imaging of live cells using quantum dot bioconjugates. *Nat Biotechnol.* **21**, 47-51 (2003).
3. Gao X. *et al. In vivo* cancer targeting and imaging with semiconductor quantum dots. *Nat. Biotechnol.* **22**, 969–976 (2004).
4. Kim S. *et al.* Near-infrared fluorescent type II quantum dots for sentinel lymph node mapping. *Nat. Biotechnol.* **22**, 93–97 (2004).
5. Han, M., Gao, X., Su, J.Z. & Nie, S. Quantum-dot-tagged microbeads for multiplexed optical coding of biomolecules. *Nat. Biotechnol.* **19**, 631–635 (2001).

The delivery problem

Delivering drugs, nucleic acids and proteins into cells in a safe and efficient manner has been, and remains, one of the most challenging problems in biotech. Robert Debs still wants to get DNA into cells, but he's largely through fiddling with cationic liposomes. In 1997, Debs, of San Francisco's California Pacific Medical Center Research Institute, reported that big, multilayered vesicles studded with cholesterol could improve the efficiency of gene delivery in living animals by as much as 1,740-fold[1]. A few months later, Nancy Templeton and her colleagues, at Baylor College of Medicine in Austin, Texas, showed how processes like sonication, heating and optimizing DNA/liposome ratios could boost systemic gene delivery[2]. Debs and colleagues were subsequently able to increase efficiency

by a couple of orders of magnitude more, but found that modifications that made the liposomes more efficient at delivering DNA also made them more toxic.

The solution, Debs says, is not figuring out how to put more DNA into a cell but to put more powerful DNA into a cell. "Rather than looking at the DNA carrier, we're looking at the DNA vector." Finding the best vector will eventually mean figuring out, tissue by tissue and condition by condition, what regulates DNA delivery, cellular processing and, most of all, expression.

That understanding would improve not just liposome-mediated gene delivery, but also delivery of naked DNA and, in some cases, viral vectors. The type of delivery mode used, whether injection or electroporation or adsorption, will also depend on the target tissue. "Each target organ requires a very different gene delivery method," says Jun-ichi Miyazaki, professor of stem cell regulation at Osaka University Graduate School of Medicine (Osaka, Japan). In 1998, Miyazaki and colleagues transferred genes into muscle by delivering electric pulses close to the injection site of the plasmid DNA, increasing the rate of gene expression by about 1,000-fold over simple DNA injection[3]. But since electroporation won't work in inaccessible internal organs like the liver where blood flow quickly sweeps away injected DNA, other methods are necessary. "None of the current methods," says Miyazaki "is sufficient in the specificity and duration of expression."

"What limits us now in our nonviral applications is that we have an almost infinite number of combinations and permutations we can test," says Debs, but he believes that this limitation will eventually become a strength. "The nonviral approaches are inherently more powerful because they are more versatile," he contends.

Viral vectors, though, with their higher transduction efficiencies and ability to stably transduce cells, currently have the upper hand in therapeutic applications. But, according to Miyazaki , plasmids have other advantages in that they are faster and cheaper to produce, can carry larger genes and can be designed to ameliorate autoimmune reactions. Deb sums this up: "If you're looking short-term, the majority of applications will be viral; in the long term, none will," he says. "One of the reasons we're alive is that the immune system has learned to recognize and blunt these viruses."

But Didier Trono, at the Swiss Institute of Technology Lausanne, thinks evolution argues on the side of the microbes. "Viruses have done this job for millions of years, and I trust more evolution than [rational] design" no matter how smart the creator of the nonviral vector, he says. Of course, virologists are doing plenty of designing themselves, and important advances have appeared in *Nature Biotechnology*.

Daniel Curiel and colleagues at the University of Alabama, Birmingham demonstrated that adenoviral vectors could be tuned to enter specific cell types and not others. They used a neutralizing antibody fragment to stop the virus from latching onto the widespread cellular receptor for the adenovirus fiber, and redirected the virus to infect cells expressing the folate receptor by conjugating folate to the neutralizing fragment[4].

Compared with adenoviruses, lentiviruses promise to be more broadly applicable because they can infect non-proliferating cells. In 1996, Trono and colleagues published the first lentiviral vector system based on HIV-1; however, many researchers feared using this system because of the possibility that HIV-1 could be reconstituted inside cells[5]. The next year, Trono modified the system by deleting about two-thirds of the HIV genetic information[6]. This multiply attenuated virus could still transduce differentiated cell types efficiently, and lentiviral vectors have since become a widely popular research tool for expressing transgenes, siRNA and even for creating transgenic animals. Indeed, VIRxSYS Corporation (Gaithersburg, MD, USA) currently has a lentiviral gene delivery vector in clinical trials for HIV/AIDS.

Trono says that one issue keeping lentiviruses from wider use is the possibility for them to cause insertional mutagenesis just as has been shown for simple retroviral vectors. But he says lentiviral vectors have already overcome bigger barriers. "In just a few years, a proposal that was initially viewed by many as rather bold, not to say irresponsible, led to a tool that is routinely used."

1. Liu, Y. *et al.* Factors influencing the efficiency of cationic liposome-mediated intravenous gene delivery. *Nat. Biotechnol.* **15**, 167–173 (1997).
2. Douglas, J.T. *et al.* Targeted gene delivery by tropism-modified adenoviral vectors. *Nat. Biotechnol.* **14**, 1574–1578 (1996).
3. Aihara, H. & Miyazaki, J. Gene transfer into muscle by electroporation *in vivo*. *Nat. Biotechnol.* **16**, 867–870 (1998).
4. Douglas, J.T. *et al.* Targeted gene delivery by tropism-modified adenoviral vectors. *Nat. Biotechnol.* **14**, 1574–1578 (1996).
5. Naldini, L. *et al.* *In vivo* gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science* **272**, 263–267 (1996).
6. Zufferey, R. *et al.* Multiply attenuated lentiviral vector achieves efficient gene delivery *in vivo*. *Nat. Biotechnol.* **15**, 871–875 (1997).

Silence is golden

Molecular biologist John Rossi fought HIV with ribozymes—catalytic RNAs that target specific nucleic acid sequences—for years with little success. Then, in 2001 he heard a lecture by Tom Tuschl, then of the Max Plank Institute for Biophysical Chemistry in Göttingen, about a new kind of gene silencing that functions in mammalian cells. RNA interference (RNAi) uses short RNA molecules to waylay mRNA before it can be translated into protein. The technique works exquisitely—assuming the necessary short RNA molecules can get into cells. Rossi rushed back to the lab to map out plans to engineer cells to make the appropriate RNA themselves. "We'd never seen knockdown that was so potent," he recalls. "We immediately shifted to going after the virus [with small interfering RNAs]."

In 2002, Rossi's group at City of Hope Beckman Research Institute (Duarte, CA, USA) described, for the first time, the use of RNAi against HIV in human cells. In transfected cells, a plasmid designed to produce small interfering RNAs (siRNAs) inhibited the expression of HIV DNA by four orders of magnitude[1]. Another paper in the same issue of *Nature Biotechnology* showed that intracellular gene expression could be targeted by siRNAs produced from plasmid DNA[2].

The following year, Mark Kay, a Stanford pediatrician and geneticist, showed that RNAi could inhibit viral replication in an animal model[3]. Kay and colleagues used one plasmid to introduce genes from the hepatitis B virus (which is not infectious in mice) and another to produce small hairpin RNAs that silenced viral genes. About 40% of hepatocytes expressed the viral transgenes in response to the first plasmid. If the second plasmid was included, levels of detectable hepatitis B virus antigens in hepatocytes fell by 99%, and replication of the viral genomes was inhibited. Kay is now looking at using adeno-associated virus as a vector in mouse models of human hepatitis infection.

Meanwhile, Rossi's laboratory has developed a lentiviral vector for delivering siRNAs and hopes to move it into clinical trials. When bone marrow cells are transduced with sequences encoding siRNAs against HIV's *TAT-REV* gene, they differentiate into immune cells that produce siRNAs that target HIV's replication machinery; the inclusion of sequences encoding an RNA decoy

and a ribozyme that interfere with other HIV targets further enhances efficacy. Rossi hopes to file an Investigational New Drug application by mid-year.

RNAi therapies are still so early in development that many issues surrounding safety and efficacy remain unclear, says Rossi. RNAi simply hasn't been studied long enough to quell concerns about long-term toxicities, for one thing. Yet, Rossi cannot hide his enthusiasm for the technology. "This is just a blessing," he says, "it's given my lab a whole new direction."

Beverly Davidson feels the same way. The University of Iowa neurologist works on therapies for genetic neurodegenerative diseases. "I knew I could get genetic material into the regions of the brain that were affected, the hurdle was what do we put there?" Davidson and colleagues delivered genes for siRNA in living mice by injecting recombinant adenoviruses into their brains. "We were able to see silencing of a transgenic allele in the mouse brain," recalls Davidson, "That was a jump-up-and-down kind of day." The team also showed silencing in liver cells and of both endogenous and exogenous genes[4]. Since then, Davidson has been able to use siRNA to ameliorate disease symptoms in mice models for Huntington disease and spinocerebellar ataxia[5,6].

Despite their enthusiasm, all three are quick to rattle off potential barriers to using RNAi as a therapy. One is finding the best sequence of RNA to use. Researchers have largely switched from randomly screening sequences to using algorithms to select the best candidates. For example, researchers at Dharmacon (Lafayette, CO, USA) systematically analyzed 180 siRNAs targeting two genes to find eight characteristics that can help identify potent siRNAs and showed that an algorithm incorporating all the characteristics improved selection[7].

Rational designs have led to big improvements, says Rossi, but there's still a long way to go. "You go through all those gyrations of trying to find a good target site–siRNA combination, and you still don't get something that works at the subnanomolar concentration."

A more worrisome challenge is off-target effects, In fact, Aimee Jackson and her colleagues at Merck (Whitehouse Station, NJ, USA) found that siRNA can silence genes, even if they share as few as 11 complementary nucleotides, a number she has currently revised down to as few as eight[8]. "The dogma was that RNAi was so specific we could use the RNAi to identify and predict off-target effects of small molecules. What we found was that the siRNAs had significantly more off-target effects than small molecules. That was kind of a shock to us." Moreover, she says, the siRNAs are probably also acting through the microRNA pathway, for which expression analysis is inadequate. "There might be off-target effects that we aren't picking up because they are occurring at the protein level," she says.

There is some evidence that chemically modifying RNA will make it more specific, she says, but such modification would work only for siRNAs delivered directly to cells, not for siRNAs cells made endogenously.

Kay and Rossi say off-target issues are less worrisome for viral genes, which share fewer sequences with human genes. Nonetheless, Rossi is currently analyzing genome-wide microarrays in his siRNA-producing immune cells for potential problems. Just as important as minimizing side effects will be determining what off-target effects are acceptable. After all, most small-molecule drugs have some off-target effects, says Kay.

Even those problems presuppose that siRNAs can be safely and reliably delivered to cells. And although the mechanism of RNAi seems to hold constant from worms to mice to humans, the means of using RNAi to fight disease will vary. In dominant genetic diseases, even a partial knockdown of a problematic gene could have a therapeutic effect, but the RNAi therapy will probably need to persist for a lifetime, either by permanently changing cells, chronically administering therapy, or both. Viral therapies, on the other hand, might not work unless viral genes are almost completely suppressed, but RNAi might not need to operate over a long period of time. Also unknown is how many cells RNAi would have to reach. Mouse brains are awfully small, says Davidson. "Can we hit enough of the target cells in a brain that's the size of a human brain?"

Beverly Davidson: "Can we hit enough of the target cells in a brain that's the size of a human brain?"

Kay says delivery is the biggest problem right now, but he's already seen what progress persistence can yield. When he first started in gene therapy about fifteen years ago, transducing liver cells in mice took luck and persistence. High concentrations of the virus had to be infused directly into

Box 1 Digesting the implications of metabolomics

In 2000, *Nature Biotechnology* published one of the first papers demonstrating that metabolic profiling (metabolomics) could be used to compare genotypes[1]. Oliver Fiehn and colleagues at Max Planck Institute of Molecular Plant Physiology (Potsdam, Germany) developed a high-throughput way to analyze hundreds of compounds in *Arabidopsis thaliana* leaves and showed that differences in these small molecules can clearly distinguish different genotypes. Now he's applying these techniques for diverse applications including figuring out how cigarette smoke harms a developing fetus, finding functions for orphan plant genes, and generally mapping metabolic networks.

Key to Fiehn and colleagues' 2000 innovation was a way to volatalize metabolites so that they could be separated by gas chromatography before mass spectrometry. But although this increased the number of compounds that can be analyzed, many still are not identified by molecular structure. "You see things going up and down and you don't know the name for it," says Fiehn, now a molecular biologist at the University of California, Davis. "This is a clear bottleneck."

In most other areas, says Fiehn, metabolomics has surged ahead. He estimates that experiments and analyses that required three weeks in 2000 can now be completed in two days. And the faster experiments include more quality controls, catalog more compounds and permit instant comparisons across data sets. Now, Fiehn and colleagues can look at the metabolites from individual organisms, identify correlations between metabolites, elucidate metabolic networks and potentially link these networks back to genes. Best of all, says Fiehn, today's data can fuel tomorrow's experiments. "The whole idea of having data that is consistent, structured and can be queried wasn't there in 2000, but it's there today."

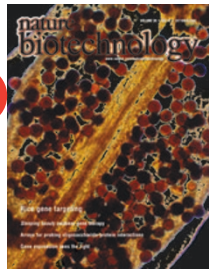
1. Fiehn, O. *et al.* Metabolite profiling for plant functional genomics. *Nat. Biotechnol.* **18**, 1157–1161 (2000).

the liver and even then very few hepatocytes were transduced. Today, the technique has improved so much that nearly 100% of the liver can be transduced simply by injecting the vector into a mouse's tail, a particularly convenient site.

The solution, say the researchers, is to test many approaches, realizing that many will fail. "RNAi is sort of in a honeymoon period, and there will be some limitations that come out of this," says Kay, "that doesn't mean we shouldn't push the technology."

1. Lee, N.S. *et. al.* Expression of small interfering RNAs targeted against HIV-1 rev transcripts in human cells. *Nat Biotechnol.* **20**, 500–505 (2002).
2. Paul, C.P., Good, P.D., Winer, I. & Engelke, D.R. Effective expression of small interfering RNA in human cells. *Nat. Biotechnol.* **20**, 505–508 (2002).
3. McCaffrey, A.P. *et al.* Inhibition of hepatitis B virus in mice by RNA interference. *Nat. Biotechnol.* **21**, 639–644 (2003).
4. Xia, H., Mao, Q., Paulson, H.L. & Davidson, B.L. siRNA-mediated gene silencing *in vitro* and *in vivo*. *Nat. Biotechnol.* **20**, 1006–1010 (2002).
5. Harper, S.Q. *et al.* RNA interference improves motor and neuropathological abnormalities in a Huntington's disease mouse model. *Proc. Natl. Acad. Sci. USA* **102**, 5820–5825 (2005).
6. Xia, H. *et al.* RNAi suppresses polyglutamine-induced neurodegeneration in a model of spinocerebellar ataxia. *Nat. Med.* **10**, 816–820 (2004).
7. Reynolds, A. *et al.* Rational siRNA design for RNA interference. *Nat Biotechnol.* **22**, 326–330 (2004).
8. Jackson, A.L. *et al.* Expression profiling reveals off-target gene regulation by RNAi. *Nat. Biotechnol.* **21**, 635–637 (2003).

Golden harvests



Technologists are used to trade-offs. Better sensitivity means worse specificity. Higher purity means lower yield. And for plants, greater tolerance to stress means stunted growth. Kazuo Shinozaki and Kazuko Yamaguchi-Shinozaki, a husband-wife team, and colleagues at RIKEN Plant Science Center (Yokohama, Japan), demonstrated that the last trade-off can be overcome[1].

By the late 1990s, many genes that protect plants against harsh environmental conditions had been identified, but researchers balked at the prospect of transferring genes one by one into transgenic plants. Shinozaki and colleagues engineered *Arabidopsis thaliana* to constitutively overexpress the transcription factor DREB1A, which they had previously found to upregulate several genes in response to cold and dehydration. The transgenic plants showed remarkable increases in tolerance to freezing, water stress and salinity, but the plants were dwarf in phenotype. The researchers got around this by expressing the protective gene just when it was needed. When they used the promoter from a dehydration-induced gene to control expression of DREB1A, the resulting plants were even more stress tolerant than constitutively active transgenics and furthermore, grew to a normal size.

"The DREB1A transcription factor is a master switch of stress-inducible gene expression," explains Shinozaki. Though the research was carried out in *A. thaliana*, he says, "our method [for] molecular breeding of abiotic stress tolerance can be applied to real breeding of crops and trees." Shinozaki's colleagues and collaborators are currently using these new insights to probe stress responses in a wide variety of crop plants, including rice, wheat, tobacco, corn, petunia, poplar, canola, tomato, tea, and soybean. It's possible that farmers breeding crops for higher yields have, over the centuries, selected variants less likely to express protective transcription factors[2]. Now, plant scientists have insights on how to tweak coordinated gene expression systems to improve stress tolerance.

Although creating more stress-tolerant crops is likely to benefit farmers everywhere, public mistrust of agricultural biotechnology remains a constraint to acceptance and expansion of the field. This is fuelled by the view that large agribusiness is the primary beneficiary of some of the current crop of transgenics, such as those displaying insect- and pesticide-resistance. Efforts to produce low trans fat soybean oils or to enhance the nutritional content of products are therefore likely to be of increasing importance, but will require more complex genetic engineering than was previously required for single gene traits (e.g., in certain *Bacillus thuringiensis* (Bt) or herbicide-resistant crops) as well as a comprehensive analysis of plant metabolic pathways (see **Box 1**). In this context, the work of Rachel Drake and her colleagues at Syngenta (Basel) in enhancing β-carotene enriched rice[3] (also known as 'Golden Rice') represented a significant step forward. The work continued on from a landmark study in 2000 by Ingo Potrykus and colleagues[4] that aimed to produce a rice strain that could help alleviate vitamin A deficiencies, which can cause blindness among malnourished populations. The problem with the original paper was that the levels of β-carotene (a precursor to vitamin A) produced were too low to be a practical source of the nutrient. Drake and her colleagues examined the metabolic pathway in transgenic rice and found that the original daffodil gene for phytoene synthase was the limiting step in β-carotene accumulation. By systematically substituting the daffodil gene with homologs from other plants, the Syngenta team boosted pro-vitamin A levels in the rice by more than 20-fold, which might be enough to make a practical difference in nutrition. What's more, Syngenta has donated the rights to this rice to the Golden Rice Humanitarian Board.

Kazuo Shinozaki and Kazuko Yamaguchi-Shinozaki of RIKEN Plant Science Center showed that overexpression of transcription factor DREB1A confers remarkable increases in plant tolerance to freezing, water stress and salinity.

1. Kasuga, M., Liu, Q., Miura, S., Yamaguchi-Shinozaki, K. & Shinozaki, K. Improving plant drought, salt, and freezing tolerance by gene transfer of a single stress-inducible transcription factor. *Nat Biotechnol.* **17**, 287–291 (1999).
2. Smirnoff, N. & Bryant, J.A. DREB takes the stress out of growing up. *Nat. Biotechnol.* **17**, 229–230 (1999).
3. Paine, J.A. *et al.* Improving the nutritional value of Golden Rice through increased pro-vitamin A content. *Nat. Biotechnol.* **23**, 482–487 (2005).
4. Ye, X. *et al.* Engineering the provitamin A (beta-carotene) biosynthetic pathway into (carotenoid-free) rice endosperm. *Science* **287**, 303–305 (2000).